SU(2)$\times$SU(2) shift operators and representations of SO(5)

J. W. B. Hughes
Department of Applied Mathematics, Queen Mary College, Mile End Road, London E1 4NS, England

(Received 4 November 1981; accepted for publication 2 April 1982)

SU(2)$\times$SU(2) shift operators analogous to the SU(2) shift operators developed and used by the author for the classification and analysis of representations of Lie algebras in an SU(2) or SO(3) basis are obtained for the SU(2)$\times$SU(2) Lie algebra in the case where one has an additional set of operators forming an irreducible four-dimensional tensor representation of SU(2)$\times$SU(2). The shift operators obtained are used to treat the representations of SO(5) in an SU(2)$\times$SU(2) basis.

PACS numbers: 02.20.Qs, 02.20.Sv

1. INTRODUCTION

In a previous paper,[1] Hughes and Yadegar showed how from the generators of an SU(2) [or SO(3)] group and a set of operators $\{T(j,\mu)\}$ transforming as an irreducible tensor representation of dimensions $(2j+1)$ of the SU(2) Lie algebra, one could construct shift operators O^k which, when acting to the right upon eigenstates of the SU(2) representations, shift the value of l $[l(l+1)$ being the value of the SU(2) Casimir invariant] by k, where k can take on any value in the range $-j, -j+1,\ldots,j-1,j$.

These shift operators have been used by the author and collaborators to classify and analyze irreducible representations of numerous Lie algebras, both compact and noncompact, in an SU(2) or SO(3) basis.[2–7] They have also been particularly useful in tackling state labeling problems such as arise for SU(3)$\supset$SO(3),[8–11] and in a series of papers by Vanden Berghe and De Meyer[12–21] have been used to solve the state labeling problems that arise in the classification of multipole phonon states using the SO(5), G(2) and SO(7) Lie algebras. The shift operators have also been used to treat the superalgebras Osp(2,1)[22] and Spl(2,1).[23]

One disadvantage of the SU(2) shift operators is that when the Lie algebra G has dimension greater than about 10 the degeneracy of the SU(2) or SO(3) subalgebra's representations in a given irreducible representation of G becomes rather high and several additional commuting state labeling operators are required to distinguish between the degenerate SU(2) states. For a given Lie algebra G the most convenient subalgebra H with respect to which to analyze the representations of G would be a maximal one, but in general it would be extremely difficult to write down H shift operators analogous to those constructed for SU(2) since, for instance, a knowledge of the Clebsch–Gordan coefficients of H would be required. However, if H is just a direct product $(\times \mathrm{SU}(2))^n$ then the shift operators for H should be obtainable without too much difficulty using the fact that, for each individual SU(2) in the direct product, they must behave like the already familiar SU(2) shift operators of Hughes and Yadegar.[1] The Lie algebras B(n) of the SO($2n+1$) groups and C(n) of the Sp($2n$) groups both possess $(\times \mathrm{SU}(2))^n$ subalgebras, as does the exceptional Lie algebra G(2). This fact can be seen easily by inspection of their root systems. So for these Lie algebras generalized $(\times \mathrm{SU}(2))^n$ shift operators should be a useful tool which can be constructed without too much difficulty. The Lie algebras A(n) of SU($n+1$) and D(n) of SO($2n$) do not contain $(\times \mathrm{SU}(2))^n$ subalgebras but will still contain $(\times \mathrm{SU}(2))^m$ subalgebras with $m<n$—for instance for A(n), $m=[n/2]$.

In this paper we consider the simplest case of SO(5), whose Lie algebra contains an SU(2)$\times$SU(2) subalgebra [the Lie algebras of SO(4), SU(2)$\times$SU(2), and SO(3)$\times$SO(3) are, of course, isomorphic]. Apart from the SU(2)$\times$SU(2) generators, the additional SO(5) operators form an irreducible four-dimensional tensor representation $R^{[\frac{1}{2},\frac{1}{2}]}$ of SU(2)$\times$SU(2). In Sec. 2 we construct the SU(2)$\times$SU(2) shift operators for this case. Denoting by $p(p+1)$ and $q(q+1)$ the eigenvalues of the Casimirs $\mathbb{P}^2$ and $\mathbb{Q}^2$ of the two SU(2) subalgebras, the shift operators obtained shift (p,q) by $(\pm\frac{1}{2},\frac{1}{2})$ or $(\pm\frac{1}{2},-\frac{1}{2})$.

In Sec. 3 the mutual commutation relations of the $R^{[\frac{1}{2},\frac{1}{2}]}$ operators are used to obtain relations between SU(2)$\times$SU(2) scalar double products of the shift operators and the SO(5) invariants. The Hermiticity properties of the shift operators are also written down.

In Sec. 4 these properties of the SU(2)$\times$SU(2) shift operators are used to classify and analyze the representations of SO(5) with respect to the SU(2)$\times$SU(2) subalgebras, and also to write down the matrix elements of the $R^{[\frac{1}{2},\frac{1}{2}]}$ generators between arbitrary SU(2)$\times$SU(2) states. The results obtained are in agreement with those found by Kemmer, Pursey, and Williams[24] and by Sharp and Pieper.[25]

2. SU(2)$\times$SU(2) SHIFT OPERATORS

We consider the group SU(2)$^p\times$SU(2)^q generated by the mutually commuting sets $\{p_0,p_\pm\}$ and $\{q_0,q_\pm\}$, respectively, satisfying the usual SU(2) commutation relations. We denote the two Casimirs by $\mathbb{P}^2=p_+p_-+p_0(p_0-1)$ and $\mathbb{Q}^2=q_+q_-+q_0(q_0-1)$, and the eigenvalues of $\mathbb{P}^2$, $\mathbb{Q}^2$, p_0 and q_0 by $p(p+1)$, $q(q+1)$, m and μ respectively.

We now introduce an irreducible four-dimensional tensor representation $R^{[\frac{1}{2},\frac{1}{2}]}$ of SU(2)$\times$SU(2) whose elements satisfy the commutation relations

$$\begin{aligned}
&[p_0,R_{\frac{1}{2},\pm\frac{1}{2}}]=\tfrac{1}{2}R_{\frac{1}{2},\pm\frac{1}{2}}, \quad [p_0,R_{-\frac{1}{2},\pm\frac{1}{2}}]=-\tfrac{1}{2}R_{-\frac{1}{2},\pm\frac{1}{2}},\\
&[p_+,R_{\frac{1}{2},\pm\frac{1}{2}}]=0, \quad [p_-,R_{-\frac{1}{2},\pm\frac{1}{2}}]=0,\\
&[p_+,R_{-\frac{1}{2},\pm\frac{1}{2}}]=R_{\frac{1}{2},\pm\frac{1}{2}}, \quad [p_-,R_{\frac{1}{2},\pm\frac{1}{2}}]=R_{-\frac{1}{2},\pm\frac{1}{2}},\\
&[q_0,R_{\pm\frac{1}{2},\frac{1}{2}}]=\tfrac{1}{2}R_{\pm\frac{1}{2},\frac{1}{2}}, \quad [q_0,R_{\pm\frac{1}{2},-\frac{1}{2}}]=-R_{\pm\frac{1}{2},-\frac{1}{2}},\\
&[q_+,R_{\pm\frac{1}{2},\frac{1}{2}}]=0, \quad [q_-,R_{\pm\frac{1}{2},-\frac{1}{2}}]=0,\\
&[q_+,R_{\pm\frac{1}{2},-\frac{1}{2}}]=R_{\pm\frac{1}{2},\frac{1}{2}}, \quad [q_-,R_{\pm\frac{1}{2},\frac{1}{2}}]=R_{\pm\frac{1}{2},-\frac{1}{2}}. \quad (2.1)
\end{aligned}$$

 0022-2488/83/051015-06$02.50

For the time being we need not consider the mutual commutation relations of the $R^{[\frac12,\frac12]}$ operators since they do not affect the construction of the shift operators. Neither do we need to specify any Hermiticity relations, even for the p's or q's; the shift operators will be equally valid if $SU(2)\times SU(2)$ is replaced by a noncompact version such as $SO(2,1)\times SO(2,1)$, $SU(1,1)\times SU(1,1)$, or $SO(3,1)$. We obtain four $SU(2)\times SU(2)$ shift operators, the method of construction being by analogy with the construction of SU(2) shift operators out of a two-dimensional irreducible set of tensor operators,[1,7,22] to which the $SU(2)\times SU(2)$ shift operators must reduce for $SU(2)^p$, and $SU(2)^q$ individually. We shall not give any details of their construction, but merely write them down here. Denoting by P and Q the operators whose eigenvalues are p and q [so $\mathbb{P}^2 = P(P+1)$, $\mathbb{Q}^2 = Q(Q+1)$], they are

$$O^{\binom{\frac12,\frac12}{\frac12,\frac12}} = R_{\frac12,\frac12}(P+p_0+1)(Q+q_0+1) + R_{-\frac12,-\frac12}q_+p_+ + R_{\frac12,-\frac12}q_+(P+p_0+1) + R_{-\frac12,\frac12}p_+(Q+q_0+1), \tag{2.2}$$

$$O^{\binom{-\frac12,-\frac12}{-\frac12,-\frac12}} = -R_{-\frac12,-\frac12}(P+p_0)(Q+q_0) - R_{\frac12,\frac12}q_-p_- + R_{-\frac12,\frac12}q_-(P+p_0) + R_{\frac12,-\frac12}p_-(q+q_0), \tag{2.3}$$

$$O^{\binom{\frac12,\ \frac12}{-\frac12,-\frac12}} = -R_{\frac12,-\frac12}(P+p_0+1)(Q+q_0) + R_{-\frac12,\frac12}q_-p_+ + R_{\frac12,\frac12}q_-(P+p_0+1) - R_{-\frac12,-\frac12}p_+(Q+q_0), \tag{2.4}$$

$$O^{\binom{-\frac12,-\frac12}{\frac12,\ \frac12}} = -R_{-\frac12,\frac12}(P+p_0)(Q+q_0+1) + R_{\frac12,-\frac12}q_+p_- - R_{-\frac12,-\frac12}q_+(P+p_0) + R_{\frac12,\frac12}p_-(Q+q_0+1), \tag{2.5}$$

The actions of these shift operators when acting to the right on the eigenstates $\left|{p,m \atop q,\mu}\right\rangle$ of $SU(2)^p\times SU(2)^q$ are

$$\begin{aligned}
O^{\binom{\frac12,\frac12}{\frac12,\frac12}}\left|{p,m \atop q,\mu}\right\rangle &\propto \left|{p+\frac12,m+\frac12 \atop q+\frac12,\mu+\frac12}\right\rangle,\\
O^{\binom{-\frac12,-\frac12}{-\frac12,-\frac12}}\left|{p,m \atop q,\mu}\right\rangle &\propto \left|{p-\frac12,m-\frac12 \atop q-\frac12,\mu-\frac12}\right\rangle,\\
O^{\binom{\frac12,\ \frac12}{-\frac12,-\frac12}}\left|{p,m \atop q,\mu}\right\rangle &\propto \left|{p+\frac12,m+\frac12 \atop q-\frac12,\mu-\frac12}\right\rangle,\\
O^{\binom{-\frac12,-\frac12}{\frac12,\ \frac12}}\left|{p,m \atop q,\mu}\right\rangle &\propto \left|{p-\frac12,m-\frac12 \atop q+\frac12,\mu+\frac12}\right\rangle.
\end{aligned} \tag{2.6}$$

In analogy with results obtained by Hughes and Yadegar,[1] using SU(2) shift operators constructed from a two-dimensional irreducible tensor set of operators, one obtains easily the following results:

$$[P,R_{\frac12,\pm\frac12}](2P+1) = R_{-\frac12,\pm\frac12}p_+ + \tfrac12 R_{\frac12,\pm\frac12}(2p_0+1), \tag{2.7}$$

$$[P,R_{-\frac12,\pm\frac12}](2P+1) = R_{\frac12,\pm\frac12}p_- - \tfrac12 R_{-\frac12,\pm\frac12}(2p_0-1), \tag{2.8}$$

$$[Q,R_{\pm\frac12,\frac12}](2Q+1) = R_{\pm\frac12,-\frac12}q_+ + \tfrac12 R_{\pm\frac12,\frac12}(2q_0+1), \tag{2.9}$$

$$[Q,R_{\pm\frac12,-\frac12}](2Q+1) = R_{\pm\frac12,\frac12}q_- - \tfrac12 R_{\pm\frac12,\frac12}(2q_0-1). \tag{2.10}$$

These will be used in the following section to obtain the Hermiticity properties of the $SU(2)\times SU(2)$ shift operators.

3. THE LIE ALGEBRA OF SO(5)

SO(5) is generated by the p's, q's, and $R^{[\frac12,\frac12]}$ operators given in Sec. 2 where, in addition to the commutation relations (2.1), the $R^{[\frac12,\frac12]}$ operators also satisfy

$$\begin{aligned}
&[R_{\frac12,\frac12},R_{-\frac12,-\frac12}] = -p_0-q_0,\\
&[R_{\frac12,-\frac12},R_{-\frac12,\frac12}] = p_0-q_0,\\
&[R_{\frac12,\frac12},R_{\frac12,-\frac12}] = p_+,\ [R_{\frac12,\frac12},R_{-\frac12,\frac12}] = q_+,\\
&[R_{-\frac12,-\frac12},R_{\frac12,-\frac12}] = q_-,\ [R_{-\frac12,-\frac12},R_{-\frac12,\frac12}] = p_-.
\end{aligned} \tag{3.1}$$

In addition the SO(5) generators also satisfy the Hermiticity conditions

$$\begin{aligned}
&p_0^\dagger = p_0,\ p_+^\dagger = p_-,\quad q_0^\dagger = q_0,\ q_+^\dagger = q_-,\\
&R_{\frac12,\frac12}^\dagger = -R_{-\frac12,-\frac12},\ R_{\frac12,-\frac12}^\dagger = R_{-\frac12,\frac12}.
\end{aligned} \tag{3.2}$$

SO(5) possesses two invariants of second and fourth orders in the generators (but both are of second order in the $R^{[\frac12,\frac12]}$ operators; there is no independent invariant of fourth order[24] in the $R^{[\frac12,\frac12]}$). The invariants, which are both Hermitian, are

$$I_2 = R_{\frac12,-\frac12}R_{-\frac12,\frac12} - R_{\frac12,\frac12}R_{-\frac12,-\frac12} + \mathbb{P}^2 + \mathbb{Q}^2 - p_0, \tag{3.3}$$

$$\begin{aligned}
I_4 = {} & 2R^2_{\frac12,\frac12}q_-p_- + 2R^2_{-\frac12,-\frac12}q_+p_+\\
& - 2R^2_{\frac12,-\frac12}q_+p_- - 2R^2_{-\frac12,\frac12}q_-p_+\\
& - 4R_{\frac12,\frac12}R_{-\frac12,\frac12}q_-p_0 + 4R_{\frac12,-\frac12}R_{-\frac12,-\frac12}q_+p_0\\
& - 4R_{\frac12,\frac12}R_{\frac12,-\frac12}p_-q_0\\
& + 4R_{-\frac12,\frac12}R_{-\frac12,-\frac12}p_+q_0 + 4R_{\frac12,\frac12}R_{-\frac12,-\frac12}p_0q_0\\
& + 4R_{\frac12,-\frac12}R_{-\frac12,\frac12}p_0q_0\\
& + (\mathbb{P}^2+\mathbb{Q}^2)(2I_2+1) - 6\mathbb{P}^2\mathbb{Q}^2 + 4\mathbb{Q}^2p_0\\
& + 4\mathbb{P}^2q_0 - 4q_0p_0^2.
\end{aligned} \tag{3.4}$$

Note that the expressions for I_2 and I_4 are not symmetrical in p_0 and q_0; this is because the products of the $R^{[\frac12,\frac12]}$ operators are written in a form which is not symmetrical with respect to the p and q suffixes. Note also that I_2 and I_4 are precisely what are called A^2 and M^4 by Kemmer *et al.*[24]

In order to simplify calculations, we define the following normalized shift operators, valid when acting to the right on $SU(2)\times SU(2)$ states $\left|{p,m \atop q,\mu}\right\rangle$:

$$A^{\binom{\pm\frac12}{\pm\frac12}}_{\binom{p}{q}} = [(p+m+\tfrac12\pm\tfrac12)(q+\mu+\tfrac12\pm\tfrac12)]^{-1/2}\, O^{\binom{\pm\frac12,\pm\frac12}{\pm\frac12,\pm\frac12}}_{\binom{p,m}{q,\mu}}, \tag{3.5}$$

$$A^{\binom{\pm\frac12}{\mp\frac12}}_{\binom{p}{q}} = [(p+m+\tfrac12\pm\tfrac12)(q+\mu+\tfrac12\mp\tfrac12)]^{-1/2}\, O^{\binom{\pm\frac12,\pm\frac12}{\mp\frac12,\mp\frac12}}_{\binom{p,m}{q,\mu}}. \tag{3.6}$$

The advantage of using these normalized operators is that in doing so one has effectively divided out the internal structure of the $SU(2)\times SU(2)$ representations.

Now, making use of Eqs. (3.1), (3.3), and (3.4), we obtain the following expressions for $SU(2)\times SU(2)$ scalar products of the shift operators.

$$A^{\binom{\frac12}{\frac12}}_{\binom{p-\frac12}{q-\frac12}} A^{\binom{-\frac12}{-\frac12}}_{\binom{p}{q}} = -\tfrac12\big(I_4-(p+q)(p+q+1)$$
$$\times(2I_2-(p+q-1)(p+q+2))\big), \qquad (3.7)$$

$$A^{\binom{-\frac12}{-\frac12}}_{\binom{p+\frac12}{q+\frac12}} A^{\binom{\frac12}{\frac12}}_{\binom{p}{q}} = -\tfrac12\big(I_4-(p+q+1)(p+q+2)$$
$$\times(2I_2-(p+q)(p+q+3))\big), \qquad (3.8)$$

$$A^{\binom{\frac12}{-\frac12}}_{\binom{p-\frac12}{q+\frac12}} A^{\binom{-\frac12}{\frac12}}_{\binom{p}{q}} = \tfrac12\big(I_4-(p-q)(p-q-1)$$
$$\times(2I_2-(p-q+1)(p-q-2))\big), \qquad (3.9)$$

$$A^{\binom{-\frac12}{\frac12}}_{\binom{p+\frac12}{q-\frac12}} A^{\binom{\frac12}{-\frac12}}_{\binom{p}{q}} = \tfrac12\big(I_4-(p-q)(p-q+1)$$
$$\times(2I_2-(p-q-1)(p-q+2))\big). \qquad (3.10)$$

The structure of the SO(5) representation is, apart from Hermiticity requirements, contained entirely in these four equations.

Using the Hermiticity relations (3.2) together with Eqs. (2.7)–(2.10), one obtains after some easy calculations the following Hermiticity properties of the shift operators:

$$\left(A^{\binom{\frac12}{\frac12}}\right)^\dagger(2P+1)(2Q+1) = A^{\binom{-\frac12}{-\frac12}}(2P)(2Q), \qquad (3.11)$$
$$\left(A^{\binom{\frac12}{-\frac12}}\right)^\dagger(2P+1)(2Q+1) = A^{\binom{-\frac12}{\frac12}}(2P)(2Q+2). \qquad (3.12)$$

Note that the original O shift operators satisfy precisely the same Hermiticity relations.

Taking matrix elements of Eqs. (3.11) and (3.12) between the $SU(2)\times SU(2)$ eigenstates $\left|{p \atop q}\right\rangle$ (the m and μ labels are omitted for the sake of brevity) we obtain

$$\left\langle {p-\frac12 \atop q\mp\frac12}\right| \left(A^{\binom{\frac12}{\pm\frac12}}\right)^\dagger \left|{p \atop q}\right\rangle = \frac{(2p)(2q+1\mp1)}{(2p+1)(2q+1)} \left\langle {p-\frac12 \atop q\mp\frac12}\right| A^{\binom{-\frac12}{\mp\frac12}} \left|{p \atop q}\right\rangle, \qquad (3.13)$$

$$\left\langle {p+\frac12 \atop q\mp\frac12}\right| \left(A^{\binom{-\frac12}{\pm\frac12}}\right)^\dagger \left|{p \atop q}\right\rangle = \frac{(2p+2)(2q+1\mp1)}{(2p+1)(2q+1)} \left\langle {p+\frac12 \atop q\mp\frac12}\right| A^{\binom{\frac12}{\mp\frac12}} \left|{p \atop q}\right\rangle. \qquad (3.14)$$

Finally, from these two relations one obtains the following formulas for the matrix elements of the double product operators of Eqs. (3.7)–(3.10):

$$\left\langle {p \atop q}\right| A^{\binom{-\frac12}{\mp\frac12}}_{\binom{p+\frac12}{q\pm\frac12}} A^{\binom{\frac12}{\pm\frac12}}_{\binom{p}{q}} \left|{p \atop q}\right\rangle = \frac{(2p+2)(2q+1\pm1)}{(2p+1)(2q+1)} \left| \left\langle {p+\frac12 \atop q\pm\frac12}\right| A^{\binom{\frac12}{\pm\frac12}}_{\binom{p}{q}} \left|{p \atop q}\right\rangle \right|^2, \qquad (3.15)$$

$$\left\langle {p \atop q}\right| A^{\binom{\frac12}{\mp\frac12}}_{\binom{p-\frac12}{q\pm\frac12}} A^{\binom{-\frac12}{\mp\frac12}}_{\binom{p}{q}} \left|{p \atop q}\right\rangle = \frac{(2p)(2q+1\pm1)}{(2p+1)(2q+1)} \left| \left\langle {p-\frac12 \atop q\mp\frac12}\right| A^{\binom{-\frac12}{\pm\frac12}}_{\binom{p}{q}} \left|{p \atop q}\right\rangle \right|^2. \qquad (3.16)$$

These relations show that the double-product operators are all positive semidefinite. Equations (3.7)–(3.10), (3.15), and (3.16) are all that are needed in order to give a classification and analysis of the irreducible representations (I.R.'s) of SO(5). This we do in the following section.

4. IRREDUCIBLE REPRESENTATIONS OF SO(5)

In this section we use the properties of the $SU(2)\times SU(2)$ shift operators given in Sec. 3 to classify the I.R.'s of SO(5). The internal properties of the I.R.'s of $SU(2)\times SU(2)$ can be summarized by the statement that p and q must be nonnegative integers or half-integers. The first step is to determine the maximum and minimum values within a given I.R. of SO(5) of p and q, so we define $\left|{\underline{p} \atop q_1}\right\rangle$, $\left|{p_1 \atop \underline{q}}\right\rangle$, $\left|{\bar p \atop q'}\right\rangle$, and $\left|{p' \atop \bar q}\right\rangle$ to be states of, respectively, minimum p, minimum q, maximum p, and maximum q. These, and all other states of the I.R., are connected to one another by repeated actions of the shift operators $A^{\binom{\frac12}{\pm\frac12}}$ and $A^{\binom{-\frac12}{\pm\frac12}}$, as depicted in Fig. 1. From the way in which the states are connected, one sees that the various p and q values of the above states must obey the relationships

$$q'-\bar q = p'-\bar p,\quad q'-\underline{q} = \bar p - p_1,$$
$$p'-\underline{p} = \bar q - q_1,\quad q_1-\underline{q} = p_1-\underline{p}. \qquad (4.1)$$

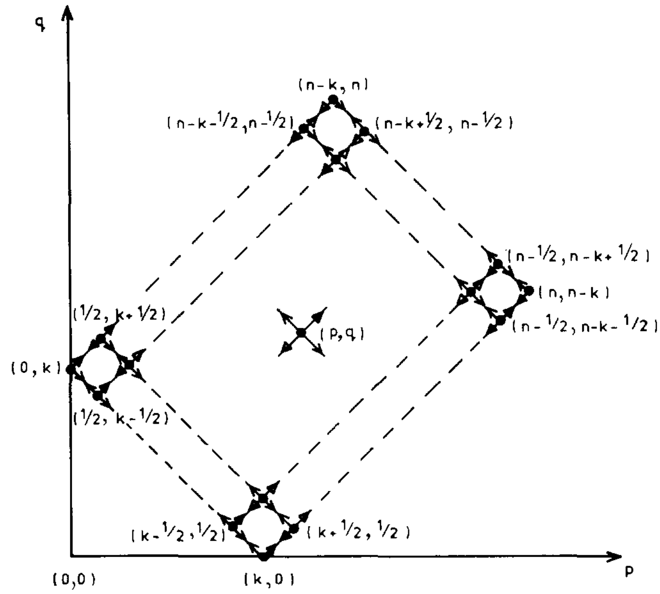

FIG. 1. States (p,q) of the irreducible representations of SO(5) specified by $I_2=(2n^2-2nk+k^2+3n-k)$, $I_4=k(k+1)(2n-k+2)(2n-k+1)$ are represented by a solid circle. The actionsof $A^{\binom{\pm\frac12}{\mp\frac12}}$ are represented by a open arrows, and those of $A^{\binom{\pm\frac12}{\pm\frac12}}$ by closed arrows.

Now we must have $A^{\left(\begin{smallmatrix}\frac12\\ \pm\frac12\end{smallmatrix}\right)}\left|\begin{smallmatrix}\bar p\\ q'\end{smallmatrix}\right\rangle = 0$, and hence also $A^{\left(\begin{smallmatrix}-\frac12\\ \mp\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}\frac12\\ \pm\frac12\end{smallmatrix}\right)}\left|\begin{smallmatrix}\bar p\\ q'\end{smallmatrix}\right\rangle = 0$; so using Eqs. (3.8) and (3.10) we obtain

$$\begin{aligned}0 &= I_4 - (\bar p + q' + 1)(\bar p + q' + 2)\\ &\quad\times(2I_2 - (\bar p + q')(\bar p + q' + 3)),\\ 0 &= I_4 - (\bar p - q')(\bar p - q' + 1)\\ &\quad\times(2I_2 - (\bar p - q^- - 1)(\bar p - q^- + 2)).\end{aligned}$$

Elimination of I_4 yields

$$\begin{aligned}&(\bar p + 1)(2q' + 1)I_2\\ &\quad= (\bar p + 1)(2q' + 1)(\bar p^2 + 2\bar p + q'^2 + q');\end{aligned}$$

so, since we cannot have either $\bar p = -1$ or $q' = -\frac12$,

$$I_2 = (\bar p^2 + 2\bar p + q'^2 + q'). \tag{4.2}$$

Since I_2 is an invariant, this value must apply for all states of the I.R.

In a similar manner, by considering the state $\left|\begin{smallmatrix}\underline{p}'\\ q\end{smallmatrix}\right\rangle$ we obtain

$$I_2 = (\bar q^2 + 2\bar q + p'^2 + p'). \tag{4.3}$$

Next consider the state $\left|\begin{smallmatrix}\underline{p}\\ q_1\end{smallmatrix}\right\rangle$; this must be annihilated by $A^{\left(\begin{smallmatrix}-\frac12\\ \pm\frac12\end{smallmatrix}\right)}$, so also $A^{\left(\begin{smallmatrix}\frac12\\ \mp\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}-\frac12\\ \pm\frac12\end{smallmatrix}\right)}\left|\begin{smallmatrix}\underline{p}\\ q\end{smallmatrix}\right\rangle = 0$. Use of Eqs. (3.7) and (3.9) then yields the relations

$$\begin{aligned}O &= I_4 - (\underline{p} + q_1)(\underline{p} + q_1 + 1)\\ &\quad\times(2I_2 - (\underline{p} + q_1 - 1)(\underline{p} + q_1 + 2))\\ O &= I_4 - (\underline{p} - q_1)(\underline{p} - q_1 - 1)\\ &\quad\times(2I_2 - (\underline{p} - q_1 + 1)(\underline{p} - q_1 - 2)).\end{aligned}$$

Again, elimination of I_4 gives

$$\underline{p}(2q_1 + 1)(I_2 - \underline{p}^2 - q_1^2 - q_1 + 1) = 0$$

or, since $q_1 \neq -\frac12$,

$$\underline{p}(I_2 - \underline{p}^2 - q_1^2 - q_1 + 1) = 0. \tag{4.4}$$

Similarly, from considerations of the state $\left|\begin{smallmatrix}p_1\\ \underline{q}\end{smallmatrix}\right\rangle$, we get

$$\underline{q}(I_2 - \underline{q}^2 - p_1^2 - p_1 + 1) = 0. \tag{4.5}$$

From (4.4) and (4.5) we obtain the following possibilities: (a) $\underline{q} = \underline{p} = 0$; (b) $\underline{p} = 0, I_2 = \underline{q}^2 + p_1^2 + p_1 - 1$, (c) $\underline{q} = 0$, $I_2 = \underline{p}^2 + q_1^2 + q_1 - 1$, (d) $I_2 = \underline{q}^2 + p_1^2 + p_1 - 1 = \underline{p}^2 + q_1^2 + q_1 - 1$. Each possibility must then be used in conjunction with Eqs. (4.1)–(4.3). Possibilities (b), (c), and (d) all lead to conclusions which violate the Hermiticity conditions (for instance, possibility (b) leads to $p_1 = \bar p + 1$ or $\bar p + \frac12$), and we omit the details of the rather cumbersome algebraic manipulations involved. The first possibility, $p = q = 0$, however, works. In this case, Eqs. (4.1) become

$$q_1 = p_1,\ q' = \bar q - \bar p + p',\ q' = \bar p - p_1,\ p' = \bar q - q_1,$$

and these together with Eqs. (4.2) and (4.3) yield (calling $\bar p = n, p_1 = k$)

$$\underline{p} = \underline{q} = 0,\ p_1 = q_1 = k, \bar p = \bar q = n,\ p' = q' = n - k, \tag{4.6}$$

where $n \geqslant k$ and n, k are both non-negative integers or half-integers. Finally, using Eq. (4.2) together with one of the equations for I_4 yields

$$I_2 = 2n^2 - 2nk + k^2 + 3n - k, \tag{4.7}$$

$$I_4 = k(k + 1)(2n - k + 1)(2n - k + 2). \tag{4.8}$$

A given I.R. of SO(5) can then be labeled by the pair (n,k) or, equivalently, by the values given in Eqs. (4.7) and (4.8) for the invariants I_2 and I_4. [Note that in the notation of Kemmer *et al.*,[24] $n = \frac12(k + l)$].

Having determined the ranges of p and q for the I.R. $D^{(n,k)}$ of SO(5), which are as shown in Fig. 1, the next task is to determine the multiplicity of the states. It is already known[24,25] that the states are simple, i.e., (p,q) is nondegenerate, so we shall give here only a sketch of how the simplicity of the states could be demonstrated using our shift operator techniques. The method consists of working out the fourth order product operator

$$G = A^{\left(\begin{smallmatrix}\frac12\\ -\frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p-\frac12\\ q+\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}-\frac12\\ -\frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p\\ q+1\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}-\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p+\frac12\\ q+\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p\\ q\end{smallmatrix}\right)}$$

acting on the state $\left|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle$; G represents a shift around a square of Fig. 1. On calculations of G, one finds that it can be expressed entirely in terms of I_2, I_4, p, and q, which is equivalent to the statement that there are no independent $SU(2)\times SU(2)$ scalar operators of order four in the $R^{[\frac12,\frac12]}$ operators. Thus G is a diagonal operator, so $G\left|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle \propto \left|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle$ for any state $\left|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle$. The simplicity of the (p,q) states can then be proved by induction.

For instance suppose $\left|\begin{smallmatrix}n-k\\ n-1\end{smallmatrix}\right\rangle_1 \propto A^{\left(\begin{smallmatrix}-\frac12\\ -\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}\frac12\\ -\frac12\end{smallmatrix}\right)}\left|\begin{smallmatrix}n-k\\ n\end{smallmatrix}\right\rangle$ and $\left|\begin{smallmatrix}n-k\\ n-1\end{smallmatrix}\right\rangle_2 \propto A^{\left(\begin{smallmatrix}\frac12\\ -\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}-\frac12\\ -\frac12\end{smallmatrix}\right)}\left|\begin{smallmatrix}n-k\\ n\end{smallmatrix}\right\rangle$; then $\left|\begin{smallmatrix}n-k\\ n-1\end{smallmatrix}\right\rangle_2 \propto G\left|\begin{smallmatrix}n-k\\ n-1\end{smallmatrix}\right\rangle_1 \propto \left|\begin{smallmatrix}n-k\\ n-1\end{smallmatrix}\right\rangle_1$, and so the two states are identical, i.e., the eigenvalues $(n - k, n - 1)$ of (P,Q) are simple. By applying the G operator to states corresponding to (p,q) values successively further removed from $(n - k, n)$, i.e., by shifting around squares which are successively further removed from the top square of Fig. 1, one proves by induction that all (p,q) values are simple.

Using Eqs. (3.7)–(3.10) with the values for I_2, I_4 given in Eqs. (4.7) and (4.8), one obtains

$$A^{\left(\begin{smallmatrix}\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p-\frac12\\ q-\frac12\end{smallmatrix}\right)} = -\tfrac12(p + q - k)(p + q + k + 1)(p + q - 2n + k - 1)(p + q + 2n - k + 2), \tag{4.9}$$

$$A^{\left(\begin{smallmatrix}-\frac12\\ -\frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p+\frac12\\ q+\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p\\ q\end{smallmatrix}\right)} = -\tfrac12(p + q + k + 2)(p + q - k + 1)(p + q + 2n - k + 3)(p + q - 2n + k), \tag{4.10}$$

$$A^{\left(\begin{smallmatrix}\frac12\\ -\frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p-\frac12\\ q+\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}-\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p\\ q\end{smallmatrix}\right)} = \tfrac12(p - q + k)(p - q - k - 1)(p - q + 2n - k + 1)(p - q - 2n + k - 2), \tag{4.11}$$

$$A^{\left(\begin{smallmatrix}-\frac12\\ \frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p+\frac12\\ q-\frac12\end{smallmatrix}\right)}A^{\left(\begin{smallmatrix}\frac12\\ -\frac12\end{smallmatrix}\right)}_{\left(\begin{smallmatrix}p\\ q\end{smallmatrix}\right)} = \tfrac12(p - q - k)(p - q + k + 1)(p - q - 2n + k - 1)(p - q + 2n - k + 2). \tag{4.12}$$

Substitution of these results in Eqs.(3.15) and (3.16) then yields

$$\left|\left\langle\begin{matrix}p+\frac12\\ q+\frac12\end{matrix}\right| A^{\binom{\frac12}{\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\right|^2=\frac{(2p+1)(2q+1)(p+q+k+2)(p+q-k+1)(p+q+2n-k+3)(2n-p-q-k)}{2(2p+2)(2q+2)}, \tag{4.13}$$

$$\left|\left\langle\begin{matrix}p-\frac12\\ q-\frac12\end{matrix}\right| A^{\binom{-\frac12}{-\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\right|^2=\frac{(2p+1)(2q+1)(p+q+k+1)(p+q-k)(p+q+2n-k+2)(2n-p-q-k+1)}{2(2p)(2q)}, \tag{4.14}$$

$$\left|\left\langle\begin{matrix}p+\frac12\\ q-\frac12\end{matrix}\right| A^{\binom{\frac12}{-\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\right|^2=\frac{(2p+1)(2q+1)(p-q-k)(p-q+k+1)(p-q-2n+k-1)(p-q+2n-k+2)}{2(2p+2)(2q)}, \tag{4.15}$$

$$\left|\left\langle\begin{matrix}p-\frac12\\ q+\frac12\end{matrix}\right| A^{\binom{-\frac12}{\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\right|^2=\frac{(2p+1)(2q+1)(p-q+k)(p-q-k-1)(p-q+2n-k+1)(p-q-2n+k-2)}{2(2p)(2q+2)}. \tag{4.16}$$

Choosing the relative phases of the state so that $\left\langle\begin{smallmatrix}p\pm\frac12\\ q\pm\frac12\end{smallmatrix}\right|A^{\binom{\frac12}{\pm\frac12}}_{\binom{p}{q}}\left|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle$ are both real and non-negative, we now obtain

$$\left\langle\begin{matrix}p+\frac12\\ q+\frac12\end{matrix}\right| A^{\binom{\frac12}{\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p+1)(2q+1)(p+q+k+2)(p+q-k+1)(p+q+2n-k+3)(2n-p-q-k)}{2(2p+2)(2q+2)}\right]^{1/2}, \tag{4.17}$$

$$\left\langle\begin{matrix}p-\frac12\\ q-\frac12\end{matrix}\right| A^{\binom{-\frac12}{-\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p+1)(2q+1)(p+q-k)(p+q+k+1)(p+q+2n-k+2)(2n-p-q-k+1)}{2(2p)(2q)}\right]^{1/2}, \tag{4.1}$$

$$\left\langle\begin{matrix}p+\frac12\\ q-\frac12\end{matrix}\right| A^{\binom{\frac12}{-\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p+1)(2q+1)(p-q-k)(p-q+k+1)(p-q-2n+k-1)(p-q+2n-k+2)}{2(2p+2)(2q)}\right]^{1/2}, \tag{4.19}$$

$$\left\langle\begin{matrix}p-\frac12\\ q+\frac12\end{matrix}\right| A^{\binom{-\frac12}{\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p+1)(2q+1)(p-q+k)(p-q-k-1)(p-q+2n-k+1)(p-q-2n+k-2)}{2(2p)(2q+2)}\right]^{1/2}. \tag{4.20}$$

By a simple extension of Eq. (3.14) of Hughes and Yadegar,[1] we obtain the following expressions for the reduced matrix of the $R^{[\frac12,\frac12]}$:

$$\left\langle\begin{matrix}p+\frac12\\ q\pm\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p+2)(2q+1\pm1)}{(2p+1)(2q+1)}\right]^{1/2}\left\langle\begin{matrix}p+\frac12\\ q\pm\frac12\end{matrix}\right| A^{\binom{\frac12}{\pm\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle, \tag{4.21}$$

$$\left\langle\begin{matrix}p-\frac12\\ q\pm\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=\left[\frac{(2p)(2q+1\pm1)}{(2p+1)(2q+1)}\right]^{1/2}\left\langle\begin{matrix}p-\frac12\\ q\pm\frac12\end{matrix}\right| A^{\binom{-\frac12}{\pm\frac12}}_{\binom{p}{q}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle, \tag{4.22}$$

Hence, using Eqs. (4.17)–(4.20), we obtain

$$\left\langle\begin{matrix}p+\frac12\\ q+\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=[\tfrac12(p+q+k+2)(p+q-k+1)(p+q+2n-k+3)(2n-p-q-k)]^{1/2}, \tag{4.23}$$

$$\left\langle\begin{matrix}p-\frac12\\ q-\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=[\tfrac12(p+q-k)(p+q+k+1)(p+q+2n-k+2)(2n-p-q-k+1)]^{1/2}, \tag{4.24}$$

$$\left\langle\begin{matrix}p+\frac12\\ q-\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=[\tfrac12(p-q-k)(p-q+k+1)(p-q-2n+k-1)(p-q+2n-k+2)]^{1/2}, \tag{4.25}$$

$$\left\langle\begin{matrix}p-\frac12\\ q+\frac12\end{matrix}\right\|R\left\|\begin{matrix}p\\ q\end{matrix}\right\rangle=[\tfrac12(p-q+k)(p-q-k-1)(p-q+2n-k+1)(p-q-2n+k-2)]^{1/2}. \tag{4.25}$$

Note the sign difference between the expression for $\left\langle\begin{smallmatrix}p-\frac12\\ q-\frac12\end{smallmatrix}\right\|R\left\|\begin{smallmatrix}p\\ q\end{smallmatrix}\right\rangle$ given here and in the paper by Sharp and Pieper.[25] This is due to a slight difference in the definition of reduced matrix elements.

Finally, we obtain the following expression for the actions of the $R^{[\frac12,\frac12]}$ on the states $\left|\begin{smallmatrix}p,m\\ q,\mu\end{smallmatrix}\right\rangle$ of the I.R. $D^{(n,k)}$ of SO(5), in which the matrix elements of $A^{\binom{\frac12}{\pm\frac12}}$ and $A^{\binom{-\frac12}{\pm\frac12}}$ are as given in Eqs. (4.17)–(4.20):

$$\begin{aligned}R_{\frac12,\pm\frac12}\left|\begin{matrix}p,m\\ q,\mu\end{matrix}\right|&=\left[\frac{(p+m+1)(q\pm\mu+1)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix}p+\frac12\\ q+\frac12\end{matrix}\right|A^{\binom{\frac12}{\frac12}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\left|\begin{matrix}p+\frac12,m+\frac12\\ q+\frac12,\mu\pm\frac12\end{matrix}\right\rangle\\ &\mp\left[\frac{(p-m)(q\mp\mu)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix}p-\frac12\\ q-\frac12\end{matrix}\right|A^{\binom{-\frac12}{-\frac12}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\left|\begin{matrix}p-\frac12,m+\frac12\\ q-\frac12,\mu\pm\frac12\end{matrix}\right\rangle\\ &\mp\left[\frac{(p-m)(q\mp\mu)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix}p+\frac12\\ q-\frac12\end{matrix}\right|A^{\binom{\frac12}{-\frac12}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\left|\begin{matrix}p+\frac12,m+\frac12\\ q-\frac12,\mu\pm\frac12\end{matrix}\right\rangle\\ &+\left[\frac{(p-m)(q\mp\mu+1)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix}p-\frac12\\ q+\frac12\end{matrix}\right|A^{\binom{-\frac12}{\frac12}}\left|\begin{matrix}p\\ q\end{matrix}\right\rangle\left|\begin{matrix}p-\frac12,m+\frac12\\ q+\frac12,\mu\pm\frac12\end{matrix}\right\rangle,\end{aligned} \tag{4.27}$$

$$
\begin{aligned}
R_{-\frac{1}{2},\pm\frac{1}{2}}\left|\begin{matrix} p,m \\ q,\mu \end{matrix}\right| = {} & \left[\frac{(p-m+1)(q\pm\mu+1)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix} p+\frac{1}{2} \\ q+\frac{1}{2}\end{matrix}\right|A^{\binom{\frac{1}{2}}{\frac{1}{2}}}\left|\begin{matrix} p \\ q\end{matrix}\right\rangle\left|\begin{matrix} p+\frac{1}{2},m-\frac{1}{2} \\ q+\frac{1}{2},\mu\pm\frac{1}{2}\end{matrix}\right\rangle \\
& \pm\left[\frac{(p+m)(q\mp\mu)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix} p-\frac{1}{2} \\ q-\frac{1}{2}\end{matrix}\right|A^{\binom{-\frac{1}{2}}{-\frac{1}{2}}}\left|\begin{matrix} p \\ q\end{matrix}\right\rangle\left|\begin{matrix} p-\frac{1}{2},m-\frac{1}{2} \\ q+\frac{1}{2},\mu\pm\frac{1}{2}\end{matrix}\right\rangle \\
& \pm\left[\frac{(p-m)(q\mp\mu)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix} p+\frac{1}{2} \\ q-\frac{1}{2}\end{matrix}\right|A^{\binom{\frac{1}{2}}{-\frac{1}{2}}}\left|\begin{matrix} p \\ q\end{matrix}\right\rangle\left|\begin{matrix} p+\frac{1}{2},m-\frac{1}{2} \\ q-\frac{1}{2},\mu\pm\frac{1}{2}\end{matrix}\right\rangle \\
& -\left[\frac{(p+m)(q\pm\mu+1)}{(2p+1)^2(2q+1)^2}\right]^{1/2}\left\langle\begin{matrix} p-\frac{1}{2} \\ q+\frac{1}{2}\end{matrix}\right|A^{\binom{-\frac{1}{2}}{\frac{1}{2}}}\left|\begin{matrix} p \\ q\end{matrix}\right\rangle\left|\begin{matrix} p-\frac{1}{2},m-\frac{1}{2} \\ q+\frac{1}{2},\mu\pm\frac{1}{2}\end{matrix}\right\rangle,
\end{aligned}
\tag{4.28}
$$

5. CONCLUSION

In this paper we have constructed shift operators which play the same role for SO(4) or $SU(2)\times SU(2)$ as the ones constructed by Hughes and Yadegar[1] did for SO(3) or SU(2), and shown by the example of $SO(5)\supset SU(2)\times SU(2)$ how they may be used to classify the irreducible representations of SO(5), obtaining results in agreement with those obtained using different techniques by Kemmer *et al.*[24] and by Sharp and Pieper.[25] It is the author's intention to use similar shift operator techniques to tackle the $G(2)\supset SU(2)\times SU(2)$ problem, and G. Vanden Berghe and H. De Meyer intend to obtain $SU(2)\times SU(2)\times SU(2)$ shift operators for the Lie algebra of SO(7). It is hoped that the results obtained by these techniques will facilitate the calculations involved in the classification of octupole phonon states.

ACKNOWLEDGMENTS

The author would like to thank G. Vanden Berghe and H. De Meyer of the Rijksuniversiteit at Gent for suggesting this problem to him and for many useful discussions.

[1]J. W. B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978).
[2]J. W. B. Hughes, J. Phys. A: Math. Nucl. Gen. **6**, 445 (1973).
[3]J. W. B. Hughes, J. Phys. A: Math. Nucl. Gen. **7**, 779 (1974).
[4]J. W. B. Hughes and J. Yadegar, J. Phys. A: Math. Nucl. Gen. **9**, 1569 (1976).
[5]J. W. B. Hughes and J. Yadegar, J. Phys. A: Math. Nucl. Gen. **9**, 1581 (1976).
[6]N. B. Backhouse and J. W. B. Hughes, J. Math. Phys. **20**, 1545 (1979).
[7]J. W. B. Hughes, J. Math. Phys. **22**, 2775 (1981).
[8]J. W. B. Hughes, J. Phys. A: Math. Nucl. Gen. **6**, 48 (1973).
[9]J. W. B. Hughes, J. Phys. A: Math. Nucl. Gen. **6**, 281 (1973).
[10]H. E. De Meyer, G. Vanden Berghe, and J. W. B. Hughes, J. Math. Phys. **22**, 2360 (1981).
[11]H. E. De Meyer, G. Vanden Berghe, and J. W. B. Hughes, J. Math. Phys. **22**, 2366 (1981).
[12]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 1902 (1980).
[13]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1906 (1980).
[14]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1913 (1980).
[15]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 1967 (1980).
[16]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1973 (1980).
[17]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 2632 (1980).
[18]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 2635 (1980).
[19]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **22**, 2326 (1981).
[20]H. E. De Meyer, G. Vanden Berghe, and J. Van der Jeugt, J. Math. Phys. **23**, 1223 (1982).
[21]G. Vanden Berghe, H. E. De Meyer, and J. Van der Jeugt, J. Math. Phys. **22**, 2332 (1981).
[22]J. W. B. Hughes, J. Math. Phys. **22**, 245 (1981).
[23]J. W. B. Hughes, Ann. Israel Phys. Soc. **3**, 320 (1980).
[24]N. Kemmer, D. L. Pursey, and S. A. Williams, J. Math. Phys. **9**, 1224 (1968).
[25]R. T. Sharp and S. C. Pieper, J. Math. Phys. **9**, 663 (1968).

On the relations between irreducible representations of the hyperoctahedral group and O(4) and SO(4)

M. Baake, B. Gemünden, and R. Oedingen
Physikalisches Institut, Universitat Bonn, Nussallee 12, D-5300 Bonn 1, West Germany

(Received 20 July 1982; accepted for publication 12 November 1982)

In this paper we describe the relations between the irreducible representations of the hyperoctahedral group in four dimensions and irreducible, low-dimensional representations of the orthogonal groups O(4) and SO(4).

PACS numbers: 02.20.Qs, 02.20.Rt

I. INTRODUCTION

In recent times a lot of work has been done on four-dimensional lattices in order to get discrete approximations of the four-dimensional space. The hyperoctahedral lattice can be generated as the set of linear combinations of the four elements of an orthonormal basis with integral coefficients. The first step in the examination of this lattice should be the investigation of its point group. In an earlier paper (see Ref. 1) we described the structure and the complete system of irreducible representations of the point group formed by rotations and reflections, called W_4, and its subgroup of pure rotations, SW_4.

As W_4 is a subgroup of O(4), all representations of O(4) form representations of W_4 when restricted onto that group. Similarly, restrictions of representations of SO(4) onto SW_4 yield representations of SW_4.

It may occur that the restriction of an irreducible representation of the continuous group onto the appropriate finite group is reducible and decomposes into a number of irreducible representations. Our aim is to find all irreducible representations of O(4) and SO(4) which stay irreducible when restricted to W_4 or SW_4, respectively.

The paper is organized as follows. After some preliminaries we describe, in Sec. III, the representations of O(4) and SO(4) in a way which is the most convenient for our purposes. In Sec. IV we determine the irreducible representations of O(4) that stay irreducible after restriction on W_4. For this purpose, we also compare the twofold Kronecker products of the four-dimensional canonical representations of both groups. In Sec. V we do the same with SO(4) and SW_4. Additionally, we show what happens with the low-dimensional representations of SO(4) on the finite subgroup SW_4 (see Table V).

II. PRELIMINARIES

In our notation we follow Miller.[2] If we use the symbol " = " for representations, this only means the equivalence of the representations and not actual identity.

The symmetry group W_4 of the four-dimensional cube, consisting of pure rotations and rotations combined with reflections, is a group of order 384 and can be presented by the wreath product $Z_2 \sim S_4$, which is isomorphic with $(Z_2)^4 \otimes_s S_4$ where $(Z_2)^4$ is the invariant subgroup.

The subgroup of all pure rotations, called SW_4, is of order 192. For details on the structure and the representations of these groups the reader is referred to Ref. 1. Here we present only the characters in the Tables I and II.

For W_4, the characters are denoted by $\chi_k^{(n)}$, where the upper index, n, indicates the dimension of the appropriate representation and the lower one, k, denumerates the characters of different representations of the same dimension.

For SW_4, the characters are denoted by an additional left upper index s: ${}^s\chi_k^{(n)}$. The lower index is omitted if there is only one representation of the given dimension.

The symmetry operations performed on the cube define, in a canonical way, a representation of dimension four which is irreducible and faithful. Therefore, by a theorem of Burnside and Brauer (see Isaacs, Ref. 3), it is possible to obtain all irreducible representations of W_4 and SW_4, respectively, by decomposing multiple Kronecker products of these canonical representations with themselves. The characters of these canonical representations are given by $\chi_1^{(4)}$ (for W_4) and ${}^s\chi_1^{(4)}$ (for SW_4). In the Tables III and IV we list the multiplicities of irreducible representations appearing in m-fold Kronecker products of the canonical representations with themselves. For this purpose, the representations belonging to the characters $\chi_k^{(n)}$ and ${}^s\chi_k^{(n)}$, respectively, are labeled $\tau_k^{(n)}$ and ${}^s\tau_k^{(n)}$, respectively. Furthermore, the canonical representations which appear to be $\tau_1^{(4)}$ and ${}^s\tau_1^{(4)}$ are called T in

TABLE I. Characters of W_4.

S_8-cycles	1^8	$1^6 2$	$1^4 4$	$1^4 2^2$		$1^2 6$	$1^2 2 4$	$1^2 2^3$		$1^2 3^2$	8	4^2		$2^2 4$		$2 6$	$2 3^2$	2^4		
S_4-cycles	1^4	1^4	$1^2 2$	$1^2 2$	1^4	13	$1^2 2$	$1^2 2$	1^4	13	4	4	2^2	2^2	$1^2 2$	13	13	2^2	$1^2 2$	1^4
order	1	4	12	12	6	32	24	24	4	32	48	48	12	24	12	32	32	12	12	1
$\chi_1^{(1)}$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
$\chi_2^{(1)}$	1	-1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	-1	1	-1	1	1	1
$\chi_3^{(1)}$	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	1	1	1	-1	1
$\chi_4^{(1)}$	1	-1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	1	1	-1	1	-1	1
$\chi_1^{(2)}$	2	2	0	0	2	-1	0	0	2	-1	0	0	2	2	0	-1	-1	2	0	2
$\chi_2^{(2)}$	2	-2	0	0	2	1	0	0	-2	-1	0	0	2	-2	0	-1	1	2	0	2
$\chi_1^{(3)}$	3	3	1	1	3	0	1	1	3	0	-1	-1	-1	-1	1	0	0	-1	1	3
$\chi_2^{(3)}$	3	-3	-1	1	3	0	1	-1	-3	0	1	-1	-1	1	-1	0	0	-1	1	3
$\chi_3^{(3)}$	3	3	-1	-1	3	0	-1	-1	3	0	1	1	-1	-1	-1	0	0	-1	-1	3
$\chi_4^{(3)}$	3	-3	1	-1	3	0	-1	1	-3	0	-1	1	-1	1	1	0	0	-1	-1	3
$\chi_1^{(4)}$	4	2	2	2	0	1	0	0	-2	1	0	0	0	0	-2	-1	-1	0	-2	-4
$\chi_2^{(4)}$	4	-2	-2	2	0	-1	0	0	2	1	0	0	0	0	2	-1	1	0	-2	-4
$\chi_3^{(4)}$	4	2	-2	-2	0	1	0	0	-2	1	0	0	0	0	2	-1	-1	0	2	-4
$\chi_4^{(4)}$	4	-2	2	-2	0	-1	0	0	2	1	0	0	0	0	-2	-1	1	0	2	-4
$\chi_1^{(6)}$	6	0	2	0	-2	0	0	-2	0	0	0	0	2	0	2	0	0	-2	0	6
$\chi_2^{(6)}$	6	0	-2	0	-2	0	0	2	0	0	0	0	2	0	-2	0	0	-2	0	6
$\chi_3^{(6)}$	6	0	0	2	-2	0	-2	0	0	0	0	0	-2	0	0	0	0	2	2	6
$\chi_4^{(6)}$	6	0	0	-2	-2	0	2	0	0	0	0	0	-2	0	0	0	0	2	-2	6
$\chi_1^{(8)}$	8	4	0	0	0	-1	0	0	-4	-1	0	0	0	0	0	1	1	0	0	-8
$\chi_2^{(8)}$	8	-4	0	0	0	1	0	0	4	-1	0	0	0	0	0	1	-1	0	0	-8

 0022-2488/83/051021-04$02.50

TABLE II. Characters of SW_4.

S_8-cycles	1^8	$1^4 4$	$1^4 2^2$	$1^2 2^3$	$1^2 3^2$	8		4^2		$2^2 4$	26	2^4	
S_4-cycles	1^4	$1^2 2$	1^4	$1^2 2$	13	4	4	2^2	2^2	$1^2 2$	13	2^2	1^4
order	1	12	6	24	32	24	24	6	6	12	32	12	1
${}^s\chi_1^{(1)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
${}^s\chi_2^{(1)}$	1	−1	1	−1	1	−1	−1	1	1	−1	1	1	1
${}^s\chi^{(2)}$	2	0	2	0	−1	0	0	2	2	0	−1	2	2
${}^s\chi_1^{(3)}$	3	1	3	1	0	−1	−1	−1	−1	1	0	−1	3
${}^s\chi_2^{(3)}$	3	−1	3	−1	0	1	1	−1	−1	−1	0	−1	3
* ${}^s\chi_3^{(3)}$	3	1	−1	−1	0	1	−1	3	−1	1	0	−1	3
* ${}^s\chi_4^{(3)}$	3	1	−1	−1	0	−1	1	−1	3	1	0	−1	3
* ${}^s\chi_5^{(3)}$	3	−1	−1	1	0	−1	1	3	−1	−1	0	−1	3
* ${}^s\chi_6^{(3)}$	3	−1	−1	1	0	1	−1	−1	3	−1	0	−1	3
${}^s\chi_1^{(4)}$	4	2	0	0	1	0	0	0	0	−2	−1	0	−4
${}^s\chi_2^{(4)}$	4	−2	0	0	1	0	0	0	0	2	−1	0	−4
${}^s\chi^{(6)}$	6	0	−2	0	0	0	0	−2	−2	0	0	2	6
${}^s\chi^{(8)}$	8	0	0	0	−1	0	0	0	0	0	1	0	−8

the tables (${}^s\tau_1^{(4)}$ can be obtained as a restriction of $\tau_1^{(4)}$ onto SW_4).

III. REPRESENTATIONS OF O(4) and SO(4)

For the examination of the connection between O(4) and W_4, and SO(4) and SW_4 we need a classification of all irreducible O(4) and SO(4) representations of finite dimension. Starting with the knowledge of the finite-dimensional representations of SU(2) (see Miller, Ref. 2) one can obtain all finite-dimensional representations of SO(4) by means of the isomorphism

$$(\mathrm{SU}(2)\otimes\mathrm{SU}(2))/Z_2 \simeq \mathrm{SO}(4)\,.$$

If $D^{(\mu)}$, $\mu\geqslant 0$ half-integer, denotes the irreducible SU(2) representation of dimension $(2\mu+1)$, we get all irreducible representations of $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$ as

$$D^{(\mu,\nu)}:=D^{(\mu)}\otimes D^{(\nu)},\quad \mu,\nu\geqslant 0 \text{ half-integer}$$

(see Miller, Ref. 2). It is not difficult to prove that $D^{(\mu,\nu)}$ is a faithful representation of $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$ if and only if $2\mu\not\equiv 2\nu \bmod 2$. As we are only interested in single-valued representations of SO(4), we have to select those of $\mathrm{SU}(2)\otimes\mathrm{SU}(2)$ which are not faithful, i.e., $D^{(\mu,\nu)}$ with $2\mu\equiv 2\nu \bmod 2$. Obviously, $D^{(\mu,\nu)}$ has the dimension $(2\mu+1)(2\nu+1)$. Thus, $\{D^{(\mu,\nu)}|\mu,\nu\geqslant 0$ half-integer and $2\mu\equiv 2\nu \bmod 2\}$ is a complete system of finite-dimensional irreducible representations of SO(4).

In order to obtain the O(4) representations from those of SO(4), one possible way is the so-called group extension which is in some sense the inversion of Clifford's theorem (see Weyl, Ref. 4) and makes use of the identity $\mathrm{O}(4)=\mathrm{SO}(4)\cup r\cdot\mathrm{SO}(4)$, where r is a reflection with $r^2=id$.

For $D^{(\mu,\nu)}$ the representation $\widetilde{D}^{(\mu,\nu)}$, defined by

$$\widetilde{D}^{(\mu,\nu)}(g):=D^{(\mu,\nu)}(r\cdot g\cdot r^{-1})\,,$$

is called the "conjugate representation". It is easy to see that the definition of $\widetilde{D}^{(\mu,\nu)}$ is independent of the choice of r, because all reflections s of O(4) can be represented as $s=hr$, r being an arbitrarily chosen reflection and $h\in\mathrm{SO}(4)$. In our case, we find $\widetilde{D}^{(\mu,\nu)}=D^{(\nu,\mu)}$. Since the proof of this identity is not very difficult but rather extensive, it will be omitted here.

For the classification of the O(4) representations we first give the following definitions:

TABLE III. Multiplicity of irreducible representations in n-fold Kronecker products of T with itself, $1\leqslant n\leqslant 10$, for the group W_4.

	$\tau_1^{(1)}$	$\tau_2^{(1)}$	$\tau_3^{(1)}$	$\tau_4^{(1)}$	$\tau_1^{(2)}$	$\tau_2^{(2)}$	$\tau_1^{(3)}$	$\tau_2^{(3)}$	$\tau_3^{(3)}$	$\tau_4^{(3)}$	$\tau_1^{(4)}$	$\tau_2^{(4)}$	$\tau_3^{(4)}$	$\tau_4^{(4)}$	$\tau_1^{(6)}$	$\tau_2^{(6)}$	$\tau_3^{(6)}$	$\tau_4^{(6)}$	$\tau_1^{(8)}$	$\tau_2^{(8)}$
$\tau_1^{(4)}=T$	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
$\otimes^2 T$	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0
$\otimes^3 T$	0	0	0	0	0	0	0	0	0	0	4	1	0	1	0	0	0	0	3	2
$\otimes^4 T$	4	1	0	1	3	2	7	3	3	3	0	0	0	0	10	6	10	6	0	0
$\otimes^5 T$	0	0	0	0	0	0	0	0	0	0	31	20	15	20	0	0	0	0	45	40
$\otimes^6 T$	31	20	15	20	45	40	76	60	60	60	0	0	0	0	136	120	136	120	0	0
$\otimes^7 T$	0	0	0	0	0	0	0	0	0	0	379	336	315	336	0	0	0	0	693	672
$\otimes^8 T$	379	336	315	336	693	672	1072	1008	1008	1008	0	0	0	0	2080	2016	2080	2016	0	0
$\otimes^9 T$	0	0	0	0	0	0	0	0	0	0	5611	5440	5355	5440	0	0	0	0	10965	10880
$\otimes^{10} T$	5611	5440	5355	5440	10965	10880	16576	16320	16320	16320	0	0	0	0	32896	32640	32896	32640	0	0

TABLE IV. Multiplicity of irreducible representations in n-fold Kronecker products of T with itself, $1\leqslant n\leqslant 10$, for the group SW_4.

	${}^s\tau_1^{(1)}$	${}^s\tau_2^{(1)}$	${}^s\tau^{(2)}$	${}^s\tau_1^{(3)}$	${}^s\tau_2^{(3)}$	${}^s\tau_3^{(3)}$	${}^s\tau_4^{(3)}$	${}^s\tau_5^{(3)}$	${}^s\tau_6^{(3)}$	${}^s\tau_1^{(4)}$	${}^s\tau_2^{(4)}$	${}^s\tau^{(6)}$	${}^s\tau^{(8)}$
${}^s\tau_1^{(4)}=T$	0	0	0	0	0	0	0	0	0	1	0	0	0
$\otimes^2 T$	1	0	0	1	0	1	1	0	0	0	0	1	0
$\otimes^3 T$	0	0	0	0	0	0	0	0	0	5	1	0	5
$\otimes^4 T$	5	1	5	10	6	10	10	6	6	0	0	16	0
$\otimes^5 T$	0	0	0	0	0	0	0	0	0	51	35	0	85
$\otimes^6 T$	51	35	85	136	120	136	136	120	120	0	0	256	0
$\otimes^7 T$	0	0	0	0	0	0	0	0	0	715	651	0	1365
$\otimes^8 T$	715	651	1365	2080	2016	2080	2080	2016	2016	0	0	4096	0
$\otimes^9 T$	0	0	0	0	0	0	0	0	0	11051	10795	0	21845
$\otimes^{10} T$	11051	10795	21845	32896	32640	32896	32896	32640	32640	0	0	65536	0

TABLE V. Relations between representations of SO(4) and SW_4.

$$D^{(0,0)}\Big|_{SW_4} = {}^s\tau_1^{(1)}$$

$$D^{(1/2,1/2)}\Big|_{SW_4} = {}^s\tau_1^{(4)}$$

$$D^{(1,1)}\Big|_{SW_4} = {}^s\tau_1^{(3)}\oplus{}^s\tau^{(6)}$$

$$D^{(3/2,3/2)}\Big|_{SW_4} = {}^s\tau_1^{(4)}\oplus{}^s\tau_2^{(4)}\oplus{}^s\tau^{(8)}$$

$$D^{(2,2)}\Big|_{SW_4} = {}^s\tau_1^{(1)}\oplus{}^s\tau_2^{(1)}\oplus{}^s\tau^{(2)}\oplus{}^s\tau_1^{(3)}\oplus{}^s\tau_3^{(3)}\oplus{}^s\tau_4^{(3)}\oplus{}^s\tau_5^{(3)}\oplus{}^s\tau_6^{(3)}\oplus{}^s\tau^{(6)}$$

$$D^{(0,1)}\Big|_{SW_4} = {}^s\tau_4^{(3)}\qquad (\text{fixed})$$

$$D^{(1,0)}\Big|_{SW_4} = {}^s\tau_4^{(3)}$$

$$D^{(0,2)}\Big|_{SW_4} = {}^s\tau^{(2)}\oplus{}^s\tau_5^{(3)}$$

$$D^{(2,0)}\Big|_{SW_4} = {}^s\tau^{(2)}\oplus{}^s\tau_6^{(3)}$$

$$D^{(3/2,1/2)}\Big|_{SW_4} = {}^s\tau^{(8)}$$

$$D^{(1/2,3/2)}\Big|_{SW_4} = {}^s\tau^{(8)}$$

$$D^{(1,2)}\Big|_{SW_4} = {}^s\tau_2^{(3)}\oplus{}^s\tau_4^{(3)}\oplus{}^s\tau_6^{(3)}\oplus{}^s\tau^{(6)}$$

$$D^{(2,1)}\Big|_{SW_4} = {}^s\tau_2^{(3)}\oplus{}^s\tau_3^{(3)}\oplus{}^s\tau_5^{(3)}\oplus{}^s\tau^{(6)}$$

$$U_{+}^{(\mu,\mu)}(g) := \begin{cases} D^{(\mu,\mu)}(g) & \text{if } g \in \mathrm{SO}(4), \\ S \cdot D^{(\mu,\mu)}(rg) & \text{if } g \in \mathrm{O}(4) \setminus \mathrm{SO}(4), \end{cases}$$

with

$$r := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and S being defined by the relations $S^2 = id$ and $T(rgr) = ST(g)S$ for all $g \in \mathrm{SO}(4)$. The existence of an S with these properties can easily be shown.

Furthermore, for $\mu \neq \nu$

$$U_{+}^{(\mu,\nu)}(g) := \begin{cases} D^{(\mu,\nu)}(g) \oplus D^{(\nu,\mu)}(g) & \text{if } g \in \mathrm{SO}(4), \\ R\,(D^{(\mu,\nu)}(rg) \oplus D^{(\nu,\mu)}(rg)) & \text{if } g \in \mathrm{O}(4) \setminus \mathrm{SO}(4), \end{cases}$$

with

$$R := \begin{pmatrix} 0 & id \\ id & 0 \end{pmatrix}.$$

Additionally, we define

$$U_{-}^{(\mu,\nu)}(g) := \det(g) \cdot U_{+}^{(\mu,\nu)}(g)$$

with

$$\det(g) := \begin{cases} +1 & \text{if } g \in \mathrm{SO}(4), \\ -1 & \text{if } g \notin \mathrm{SO}(4). \end{cases}$$

This is also valid for $\mu = \nu$. It can be shown that

$\{ U_{+}^{(\mu,\mu)} \mid \mu \geqslant 0 \text{ half-integer} \} \cup \{ U_{-}^{(\mu,\mu)} \mid \mu \geqslant 0 \text{ half-integer} \}$

$\cup \{ U_{+}^{(\mu,\nu)} \mid 0 \leqslant \mu < \nu \text{ half-integer and } 2\mu \equiv 2\nu \bmod 2 \}$

$\cup \{ U_{-}^{(\mu,\nu)} \mid 0 \leqslant \mu < \nu \text{ half-integer and } 2\mu \equiv 2\nu \bmod 2 \}$

is a complete system of irreducible finite-dimensional representations of O(4). Note that $U_{+}^{(\mu,\nu)}$ and $U_{+}^{(\nu,\mu)}$ are equivalent representations. It is easy to calculate that

$$\dim U_{+}^{(\mu,\mu)} = \dim U_{-}^{(\mu,\mu)} = (2\mu + 1)^2$$

and

$$\dim U_{+}^{(\mu,\nu)} = \dim U_{-}^{(\mu,\nu)} = 2(2\mu + 1)(2\nu + 1).$$

IV. O(4) REPRESENTATIONS RESTRICTED ON W_4

We are now looking for those irreducible representations of O(4) which remain irreducible after restriction on W_4. For a list of the W_4 representations the reader is referred to Ref. 1. W_4 has irreducible representations of dimension one, two, three, four, six, and eight. So we only have to regard the following O(4) representations:

$$U_{+}^{(0,0)};\ U_{-}^{(0,0)};\ U_{+}^{(0,1)};\ U_{-}^{(0,1)};\ U_{+}^{(1/2,1/2)};\ U_{-}^{(1/2,1/2)}.$$

By construction of the O(4) representations the following relations can be seen:

$U_{+}^{(0,0)} | W_4 = \tau_1^{(1)}$ ("identity"),

$U_{-}^{(0,0)} | W_4 = \tau_4^{(1)}$ ("determinant"),

$U_{+}^{(1/2,1/2)} | W_4 = \tau_1^{(4)}$ ("the canonical representation"),

$U_{-}^{(1/2,1/2)} | W_4 = \tau_4^{(4)} = \tau_1^{(4)} \otimes \tau_4^{(1)}$.

Here, $U_{\pm}^{(\mu,\nu)} | W_4$ means the restriction of the O(4) representation $U_{\pm}^{(\mu,\nu)}$ on the finite subgroup W_4.

Furthermore, one finds

$$U_{+}^{(0,1)} | W_4 = U_{-}^{(0,1)} | W_4 = \tau_1^{(6)}.$$

Here $\tau_1^{(6)}$ is the irreducible skew-symmetric part of the twofold Kronecker product of the canonical representation $\tau_1^{(4)}$ with itself. In the same way one obtains $U_{+}^{(0,1)}$ as the skew-symmetric part of the Kronecker product of $U_{+}^{(1/2,1/2)}$ with itself. As, additionally, $\dim U_{+}^{(0,1)} = \dim \tau_1^{(6)}$, we conclude that

$$U_{+}^{(0,1)} | W_4 = \tau_1^{(6)}.$$

Normally, one would expect $U_{-}^{(0,1)}$ to form another six-dimensional representation when restricted to W_4, which should appear to be $\tau_1^{(6)} \otimes \tau_4^{(1)}$. Regarding the character table of W_4, we find that, for all classes of W_4,

$$\det(g) = -1 \text{ implies trace } (\tau_1^{(6)}(g)) = \chi_1^{(6)} = 0,$$

so

$$\tau_1^{(6)} \otimes \tau_4^{(1)} = \tau_1^{(6)},$$

and consequently,

$$U_{+}^{(0,1)} | W_4 = U_{-}^{(0,1)} | W_4.$$

Furthermore, we look at the twofold Kronecker product of $U_{+}^{(1/2,1/2)}$ with itself. We have

$$U_{+}^{(1/2,1/2)} \otimes U_{+}^{(1/2,1/2)} = U_{+}^{(0,0)} \oplus U_{+}^{(0,1)} \oplus U_{+}^{(1,1)}.$$

As shown above, $U_{+}^{(0,1)}$, which appears to be the representation on the space of skew-symmetric tensors of rank 2, stays irreducible after the restriction to W_4.

However, the representation on the space of the traceless symmetric tensors, $U_{+}^{(1,1)}$, decomposes on W_4 in the following manner:

$$U_{+}^{(1,1)} | W_4 = \tau_1^{(3)} \oplus \tau_3^{(6)}$$

(see Table III).

V. SO(4) REPRESENTATIONS RESTRICTED ON SW_4

We shall now proceed in the same way for SO(4) and SW_4, respectively. SW_4 has irreducible representations of the dimensions one, two, three, four, six, and eight. Therefore, we have four possible candidates of SO(4) representation, which stay irreducible when restricted on W_4.

$$D^{(0,0)},\quad D^{(1/2,1/2)},\quad D^{(0,1)},\quad D^{(1,0)}.$$

As in the previous case we directly obtain

$D^{(0,0)} | SW_4 = {}^s\tau^{(1)}$ ("identity"),

$D^{(1/2,1/2)} | SW_4 = {}^s\tau_1^{(4)}$ ("the canonical representation").

Since

$$U_{+}^{(0,1)} | \mathrm{SO}(4) = D^{(0,1)} \oplus D^{(1,0)}$$

and

$$\tau_1^{(6)} | SW_4 = {}^s\tau_3^{(3)} \oplus {}^s\tau_4^{(3)}$$

we conclude that

$$D^{(0,1)} | SW_4 = {}^s\tau_3^{(3)} \text{ and } D^{(1,0)} | SW_4 = {}^s\tau_4^{(3)}.$$

Note that the characters of ${}^s\tau_3^{(3)}$ and ${}^s\tau_4^{(3)}$ (see Table II) differ only on two pairs of conjugacy classes. These classes are of the same order and have the same cycle structure in S_4 and S_8. Thus the last two equations fix these classes and determine the following calculations with ${}^s\tau_3^{(3)}$ and ${}^s\tau_4^{(3)}$.

For more details about the relations between the SW_4 representations and those of SO(4), it is quite useful to examine the decomposition of the multiple Kronecker product of $D^{(1/2,1/2)}$ with itself and to compare the results with the decomposition numbers of Table IV.

For the SU(2) representations the following decomposition rule holds (see Miller, Ref. 2):

$$D^{(\nu)} \otimes D^{(\mu)} = \bigoplus_{l=|\nu-\mu|}^{\nu+\mu} D^{(l)}.$$

Consequently, by the definition of $D^{(\mu,\nu)}$, we obtain the ensuing relation for the SO(4) representations:

$$D^{(\mu,\nu)} \otimes D^{(\mu',\nu')} = \bigoplus_{w=|\mu-\mu'|}^{\mu+\mu'} \bigoplus_{w'=|\nu-\nu'|}^{\nu+\nu'} D^{(w,w')}.$$

Now, it is only a matter of a proof by induction to verify the resultant formula:

$$\otimes^n D^{(1/2,1/2)} := \underbrace{D^{(1/2,1/2)} \otimes \cdots \otimes D^{(1/2,1/2)}}_{n\text{-times}} = \bigoplus_{2\mu=0}^{n} \bigoplus_{2\nu=0}^{n} a_{\mu,\nu}^{(n)} D^{(\mu,\nu)}$$

with $a_{\mu,\nu}^{(n)} := a_\mu^{(n)} \cdot a_\nu^{(n)}$ and

$$a_\nu^{(n)} := \begin{cases} 0 & \text{if } 2\nu \not\equiv n \bmod 2, \\ \binom{n}{(n/2)-\nu} \cdot \dfrac{2\nu+1}{(n/2)+\nu+1} & \text{if } 2\nu \equiv n \bmod 2, \\ 0 & \text{if } \nu < 0 \text{ or } 2\nu > n. \end{cases}$$

Since $D^{(1/2,1/2)}|SW_4 = {}^s\tau_1^{(4)}$ it is clear that

$$(\otimes^n D^{(1/2,1/2)})|SW_4 = \otimes^n\, {}^s\tau_1^{(4)}.$$

In both cases we know the decomposition rules and thus we can compare the irreducible parts and have to attach them to one another. This is possible in a unique way if we additionally make use of the decomposition rules for the twofold Kronecker products of arbitrary representations of SO(4) and SW_4, respectively. Applying this method for $n = 1, 2, 3, 4$, one gets the results listed in Table V. Note that every SW_4 representation appears at least once.

ACKNOWLEDGMENTS

To Professor V. Rittenberg we are very thankful for the support he also gave us for this paper. Furthermore, we wish to thank Professor J. C. Jantzen for giving us a very useful hint on the method of group extension.

[1] M. Baake, B. Gemünden, and R. Oedingen, "Structure and representations of the symmetry group of the four-dimensional cube," J. Math. Phys. **23**, 944 (1982).

[2] W. Miller, *Symmetry Groups and Their Applications* (Academic, New York, 1972).

[3] J. M. Isaacs, *Character Theory of Finite Groups* (Academic, New York, 1976).

[4] H. Weyl, *The Classical Groups* (Princeton U. P., Princeton, NJ, 1946).

Nonscalar extension of shift operator techniques for SU(3) in an O(3) basis. III. Shift operators of second degree in the tensor components

J. Van der Jeugt,[a] H. E. De Meyer,[b] and G. Vanden Berghe
Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit Gent Krijgslaan 281-S9, B-9000 Gent, Belgium

(Received 8 March 1982; accepted for publication 30 April 1982)

Shift operators Q_l^k $(-2\leqslant k\leqslant 2)$ of second degree in the tensor components q_μ $(-2\leqslant\mu\leqslant 2)$ are constructed. Relations connecting quadratic shift operator products of the type $O_{l+k}^j Q_l^k$ or $Q_{l+j}^k O_l^j$, and of the type $Q_{l+k}^j Q_l^k$ are derived. The usefulness of these relations is demonstrated by the example of the O_l^0- and Q_l^0-eigenvalue calculation for various irreducible respresentations (p,q) of SU(3).

PACS numbers: 02.20.Qs, 02.30. + g

1. INTRODUCTION

The construction of orthonormal bases for SU (3) representations in the SU (3)$\supset$O(3) reduction has been the subject of numerous contributions. In most of the approaches one looks for an additional Hermitian operator of which the eigenfunctions form a basis. It has been shown that only two such independent operators exist, and in the present paper we shall denote them by O_l^0 and Q_l^0.

A recursive method for calculating O_l^0 and Q_l^0 eigenvalues has been developed by Hughes.[1,2] The technique essentially relies on a set of relations among products of shift operators which behave as O(3) scalars. These shift operators are constructed out of the three generators $l_0, l_\pm$ of O(3) and the five generators q_μ $(-2\leqslant\mu\leqslant+2)$, which form a five-dimensional tensor representation of O(3). To find a part of the O_l^0-eigenvalue spectrum, Hughes[1,2] needed relations between triple product operators, which were very difficult to construct. In two preceding papers[3,4] (to be referred to as I and II), nonscalar relations between shift operators were introduced. This extension of the shift operator technique provided a lot of advantages and simplifications with respect to the O_l^0-eigenvalue calculation. However, these nonscalar relations did not rule out the complete use of the triple product relations.

In this paper shift operators Q_l^k $(-2\leqslant k\leqslant 2)$, which are quadratic in the generators q_μ $(-2\leqslant\mu\leqslant 2)$, are introduced. They are of fourth degree in the SU (3) generators. The Q_l^s operators can be expressed in terms of products of the type $O_{l+k}^j O_l^k$ $(j+k=s)$. However, the combination of the products $O_{l+k}^j O_l^k$, which are in general operators of sixth degree in the generators, is such that the resulting operator Q_l^s is only of fourth degree. In Sec. 3 relations between shift operators of the type $O_{l+k}^j Q_l^k$ and $Q_{l+j}^k O_l^j$ are derived. Since such operator products are only of seventh degree in the generators, the relations are easier to construct than the triple product relations, where operators $O_{l+k+j}^{-j-k} O_{l+k}^j O_l^k$ of ninth degree are involved.

To complete the shift operator properties, relations between products of the type $O_{l+k}^j Q_l^k$ are constructed in Sec. 4.

In Sec. 5 we show that for the O_l^0-eigenvalue determination, the triple product relations can be completely replaced by the newly derived relations between $O_{l+k}^j Q_l^k$ and $Q_{l+j}^k O_l^j$ operators. We also demonstrate the usefulness of these relations in the calculations of Q_l^0-eigenvlaues, and in deriving a general formula for the Q_l^0-eigenvalues in a case of twofold l-degeneracy.

2. SHIFT OPERATORS FOR SU(3) IN AN O(3) BASIS

A commonly used choice for a generator basis of the group SU(3) is the one consisting of the Cartan subalgebra H_1, H_2, and its root vectors $E_{\pm\alpha}$, $E_{\pm\beta}$, and $E_{\pm\bar\beta}$.[5] SU(3) possess two invariants I_2 and I_3, respectively, of second and third order in the generators, whose eigenvalues serve to specify uniquely its irreducible representations. Every unitary irreducible representation may be labeled by the pair of integers (p,q) satisfying $p\geqslant q\geqslant 0$, and related to I_2 and I_3 by the formulae[5]

$$I_2=\tfrac{1}{9}(p^2+q^2-pq+3p), \tag{2.1}$$

$$I_3=\tfrac{1}{162}(p-2q)(2p+3-q)(p+q+3). \tag{2.2}$$

Here, we shall be concerned with a different generator basis, defined in terms of the above-mentioned generators as follows:

$$\begin{aligned} &l_0=2\sqrt{3}H_1, \quad 1_\pm=2\sqrt{3}(E_{\pm\bar\beta}-E_{\pm\beta}),\\ &q_0=-6H_2, \quad q_{\pm 1}=\mp 3\sqrt{2}(E_{\pm\beta}+E_{\pm\bar\beta}),\\ &q_{\pm 2}=6E_{\pm\alpha}. \end{aligned}\tag{2.3}$$

The relevant commutation relations are

$$\begin{aligned} &[l_0,l_\pm]=\pm l_\pm, \quad [l_+,l_-]=2l_0,\\ &[l_0,q_\mu]=\mu q_\mu,\\ &[l_\pm,q_\mu]=[2\mp\mu)(3\pm\mu)]^{1/2}q_{\mu\pm 1} \quad (\mu=0,\pm 1,\pm 2), \end{aligned}\tag{2.4}$$

and

$$\begin{aligned} &[q_0,q_{\pm 1}]=-3(\sqrt{3}/\sqrt{2})l_\pm, \quad [q_{+1},q_{-1}]=-3l_0,\\ &[q_{\pm 2},q_{\mp 1}]=-3l_\pm, \quad [q_{+2},q_{-2}]=6l_0. \end{aligned}\tag{2.5}$$

The operators l_0 and $l_\pm$ together generate an O(3) subgroup of SU(3), with respect to which the q_μ form a five-dimensional irreducible tensor representation. I_2 and I_3 are given in terms of these generators by the formulae

$$I_2=\tfrac{1}{36}(q_0^2-q_{+1}q_{-1}-q_{-1}q_{+1}+q_{+2}q_{-2}+q_{-2}q_{+2}), \tag{2.6}$$

[a] Research assistant N.F.W.O. (Belgium).
[b] Research associate N.F.W.O. (Belgium).

$$
\begin{aligned}
I_3 = \frac{1}{2\times(36)^2}\{ & 4(q_0^2 - 3q_{+1}q_{-1} - 6q_{+2}q_{-2})q_0 \\
& + 6(6)^{1/2}(q_{+2}q^2_{-1} + q_{-2}q^2_{+1}) \\
& - 9(6)^{1/2}(q_{-2}l^2_+ + q_{+2}l^2_-) \\
& - 18(6)^{1/2}(l_0 - 1)q_{-1}l_+ \\
& + 18(6)^{1/2}(l_0 + 1)q_{+1}l_- \\
& + 18(L^2 - 3l_0^2 + 3l_0 + 10)q_0\},
\end{aligned}
\tag{2.7}
$$

where $L^2 \equiv l_+l_- + l_0^2 - l_0$ is the Casimir of O(3). The irreducible representations of 0(3) will be labeled by l, where $l(l+1)$ is the eigenvalue of L^2.

SU(3) possesses two Hermitian O(3) scalar operators[6] of third and fourth order in the group generators, respectively, O_l^0 and Q_l^0, which however do not commute and so cannot be diagonalized simultaneously, except when acting on states corresponding to nondegenerate l values.

As basis vectors for the representation (p,q) we use the kets $|p,q;l,\lambda_l,m\rangle$ (or $|l,\lambda_l,m\rangle$ if confusion is excluded), where m is the eigenvalue of l_0 and λ_l is the supplementary label needed for their unique specification. We choose λ_l to be the eigenvalue of O_l^0.

An appropriate apparatus for obtaining the O_l^0- and Q_l^0-eigenvalues consists of the l-shift-operator technique, developed by Hughes.[1,2,7] The SU (3) shift operators $O_l^{\pm k}$ $(k = 0,1,2)$ shift the eigenvalues of the O(3) Casimir operator L^2 by $\pm k$, and leave m unchanged, i.e.,

$$
O_l^k|l,\lambda_l,m> \sim |l+k,\lambda_{l+k},m\rangle. \tag{2.8}
$$

Since the eigenvalues of O_l^0 and Q_l^0 are m-independent, we can restrict our attention to SU(3) states which correspond to zero m, and employ the kets $|1,\lambda_l\rangle \equiv |l,\lambda_l,m=0\rangle$. In this case the shift operators read[3]

$$
\begin{aligned}
O_l^0 &= (6)^{1/2}l(l+1)q_0 \\
&\quad - 3(q_{-1}l_+ + q_{+1}l_-) - 3(q_{-2}l^2_+ + q_{+2}l^2_-), \\
O_l^{+1}/(l+1) &= (l+2)(q_{-1}l_+ - q_{+1}l_-) \\
&\quad + (q_{-2}l^2_+ - q_{+2}l^2_-), \\
O_l^{+2}/(l+1)(l+2) &= (6)^{1/2}(l+1)(l+2)q_0 \\
&\quad + 2(l+2)(q_{-1}l_+ + q_{+1}l_-) \\
&\quad + (q_{-2}l^2_+ + q_{+2}l^2_-),
\end{aligned}
$$

and

$$
O_l^{-k} = O^{+k}_{-l-1} \quad (k = 1,2). \tag{2.9}
$$

The O_l^k-operators are linear in the q-generators. In other groups where shift operator techniques have been used,[8,9] only operators linear in the tensor components were considered. In the following we construct shift operators Q_l^k $(k = 0, \pm 1, \pm 2)$ which, on the contrary, are quadratic in the q_μ's.

A first way to find the Q_l^k's, is to prescribe that they must be of the form

$$
\begin{aligned}
Q_l^k = aq_0^2 &+ bq_{+1}q_{-1} + cq_{+2}q_{-2} \\
&+ dq_{-1}q_0l_+ + eq_{-2}q_{+1}l_+ + fq_{+1}q_0l_- \\
&+ gq_{+2}q_{-1}l_- + hq^2_{-1}l^2_+ + iq_{-2}q_0l^2_+ \\
&+ jq^2_{+1}l^2_- + kq_{+2}q_0l^2_- + hl_+l_-,
\end{aligned}
\tag{2.10}
$$

and to require that they satisfy the commutator relation

$$
[L^2, Q_l^k] = k(2l+k+1)Q_l^k. \tag{2.11}
$$

The expression (2.10) is the most general operator up to fourth-order in the generators and quadratic in the q_μ, which commutes with l_0. The linear equations in $a,b,...,h$, generated by (2.11), have a nonzero solution (up to an overall multiplicative constant) for $k = 0, \pm 1, \pm 2$. For the case $k = 0$, one obtains, besides the Q_l^0 operator of Racah, also the Casimir I_2 as a solution. An easier way to construct the Q_l^k is to use the general formula of Hughes,[7] which gives immediately the expression of a shift operator in terms of a $(2j+1)$-dimensional tensor representation $T(j,\mu)$ of O(3). In order to obtain operators which are quadratic in the q_μ, we define a five-dimensional tensor representation as

$$
\begin{aligned}
T(2,\mu) = q_\mu^{(2)} &= [q\times q]^2_\mu \\
&= c\sum_{m_1,m_2}\langle 2m_12m_2|2\mu\rangle q_{m_1}q_{m_2} \quad (-2\leqslant\mu\leqslant 2),
\end{aligned}
\tag{2.12}
$$

where c is an arbitrary constant. If we put $c = -7/\sqrt{2}$ and use the relations (2.5), we obtain

$$
\begin{aligned}
q_0^{(2)} &= q_0^2 - q_{+1}q_{-1} - 2q_{+2}q_{-2} + \tfrac{9}{2}l_0, \\
q^{(2)}_{\pm 1} &= q_{\pm 1}q_0 - 6^{1/2}q_{\pm 2}q_{\mp 1} - \tfrac{9}{2}(3/2)^{1/2}l_\pm, \\
q^{(2)}_{\pm 2} &= -2q_{\pm 2}q_0 + (3/2)^{1/2}q^2_{\pm 1},
\end{aligned}
\tag{2.13}
$$

satisfying the commutation relations $[l\pm, q^\mu_{(2)}] = [(2\mp\mu)\times(3\pm\mu)]^{1/2}q^{(2)}_{\mu\pm 1}$.

In terms of the coupled tensor $q^{(2)}$, the shift operators Q_l^k can be written as follows (again we restrict our attention to operator forms valid when acting upon $m = 0$ states):

$$
\begin{aligned}
Q_l^0 &= 2l(l+1)q_0^{(2)} - 6^{1/2}(q^{(2)}_{-1}l_+ + q^{(2)}_{+1}l_-) \\
&\quad - 6^{1/2}(q^{(2)}_{-2}l^2_+ + q^{(2)}_{+2}l^2_-) - 81l_+l_-, \\
Q_l^{+1}/(l+1) &= (l+2)(q^{(2)}_{-1}l_+ - q^{(2)}_{+1}l_-) \\
&\quad + (q^{(2)}_{-2}l^2_+ - q^{(2)}_{+2}l^2_-), \\
Q_l^{+2}/(l+1)(l+2) &= 6^{1/2}(l+1)(l+2)q_0^{(2)} \\
&\quad + 2(l+2)(q^{(2)}_{-1}l_+ + q^{(2)}_{+1}l_-) \\
&\quad + (q^{(2)}_{-2}l^2_+ + q^{(2)}_{+2}l^2_-),
\end{aligned}
$$

and

$$
Q_l^{-k} = Q^k_{-l-1} \quad (k = 1,2). \tag{2.14}
$$

The shift operators are only determined up to an overall multipicative constant, and the scalar shift operator Q_l^0 is, moreover, only determined up to an additional term in L^2. Here, Q_l^0 corresponds to the choice made by Hughes[1,2] and Racah.[6] Investigation shows that there exist certain relations between the quadratic shift operators Q_l^k and the linear shift operators O_l^k (quadratic and linear refers to the order of q_μ). Those relations, which may be regarded as an extension of Eq. (45) of Ref. 1, are

$$
Q_l^{-2} = [1/2(6)^{1/2}(2l-1)](O_l^{-2}O_l^0 - O_l^0O_l^{-2}), \tag{2.15}
$$

$$
Q_l^{-1} = (1/2(6)^{1/2}l)(O_l^{-1}O_l^0 - O_{l-1}^0O_l^{-1}) \tag{2.16}
$$

$$
\begin{aligned}
&= (1/4(6)^{1/2})(O_{l+1}^{-2}O_l^{+1}/l(l+1)^2 \\
&\quad - O_{l-2}^{+1}O_l^{-2}/l(l-1)^2),
\end{aligned}
\tag{2.17}
$$

$$Q_l^0 = [1/(2l+1)]\,[O_{l-1}^{+1} O_l^{-1}/l^2 - O_{l+1}^{-1} O_l^{+1}/(l+1)^2] + 72l(l+1)I_2 - 6l(l+1)(2l^2+2l+9) \tag{2.18}$$

$$= [1/4(2l+1)]\,[O_{l+2}^{-2} O_l^{+2}/(l+1)^2(l+2)^2 - O_{l-2}^{+2} O_l^{-2}/l^2(l-1)^2] - 72(2l^2+2l+3)I_2 + 6l(l+1)(4l^2+4l+3). \tag{2.19}$$

In this respect the Q_l^k-operators are not independent objects: they can be expressed in terms of product operators of the type $O_{l+j}^k O_l^j$. However, in the next sections we will show that it is sometimes much more practical to work with the Q_l^k instead of the operator products $O_{l+j}^k O_l^j$. The main reason for this lies in the degree of the operators: the Q_l^k operators are in general of fourth degree in the generators of SU(3) [see Eqs. (2.13) and (2.14)] and the products $O_{l+j}^k O_l^j$ are in general of sixth degree in the generators [see Eqs. (2.9)].

3. RELATIONS BETWEEN MIXED PRODUCT OPERATORS

In Refs. 1 and 3 relations between shift operator products were introduced. The relations (35)–(48) of Ref. 1 allow one to calculate in general the O_l^0- and Q_l^0-eigenvalues for cases where there is no l-degeneracy.[2] They contain triple product operators of the type $O_{l+j+k}^{-j-k} O_{l+j}^k O_l^j$, which are in general of ninth degree in the generators, and were extremely tedious to calculate. The relations between scalar and nonscalar quadratic operator products of the type $O_{l+j}^k O_l^j$ [Eqs. (2.2)–(2.8) of Ref. 3] allow one to calculate in general the O_l^0-eigenvalue even in case of twofold and threefold l-degeneracy. However, those relations were not sufficient; triple product relations were still needed. To avoid this difficulty we introduce relations between mixed product operators of the type $O_{l+j}^k Q_l^j$ and $Q_{l+k}^j O_l^k$. To construct such relations all products are reduced to a standard form, which consists in choosing a particular order for the generators of the group. Here we have taken the order $q_{+2} > q_{-2} > q_{+1} > q_{-1} > q_0 > l_+ > l > l_0$. For instance, the standard form of $q_{-1}q_{+2}$ is $q_{+2}q_{-1} + 3l_-$. Once all the terms of the operator products are transformed into the standard form [making use fo the commutators (2.4) and (2.5)], it is rather straightforward to find relations between them. In such relations there can also appear a term in $I_2O_l^s$ and in O_l^s, where $s = j + k$ is the total shift value of the relation. If the relation is of the scalar type (i.e., $s = 0$), a term in I_3, the third order Casimir invariant appears. We give here the expressions of the relations in the case $s \leqslant 0$. Using the transformation rule $O_l^{-k} = O^k_{-(l+1)}$,[7] one can easily obtain the results for $s > 0$. We find one equation with $s = -4$, two with $s = -3$, four with $s = -2$, six with $s = -1$, and eight with $s = 0$/ These are the maximum number of independent equations between the objects $O_{l+j}^k Q_l^j$ and $O_{l+k}^j O_l^k$: any other relation between them must be a linear combination of the independent equations. To restrict the number of formulae, we write down only half of the equations; the other half can be obtained immediately by using the transformation rule "$l \to -l-s-1$." This rule, valid for every relation between shift operators, can be summarized as follows:

(i) all the operator products $O_{l+k}^j Q_l^k$ (resp. $Q_{l+k}^j O_l^k$) are replaced by $Q_{l+j}^k O_l^j$ (resp. $O_{l+j}^k Q_l^j$);

(ii) all the other operators in the equation (e.g., $I_2O_l^s$, O_l^s, I_3) are kept unchanged;

(iii) in every coefficient l is replaced by $-l-s-1$.

A proof of this rule can be found elsewhere.[10] The final results are:

with $s = -4$,

$$O_{l-2}^{-2} Q_l^{-2} - Q_{l-2}^{-2} O_l^{-2} = 0; \tag{3.1}$$

with $s = -3$,

$$(l-3)O_{l-1}^{-2} Q_l^{-1} - (l-1)Q_{l-2}^{-1} O_l^{-2} + 2O_{l-2}^{-1} Q_l^{-2} = 0; \tag{3.2}$$

with $s = -2$,

$$(l+3)O_{l-2}^0 Q_l^{-2} + 12O_{l-1}^{-1} Q_l^{-1} - (l-1)Q_l^{-2} O_l^0 + 216(6)^{1/2}(l-1)I_2 O_1^{-2} + 54(6)^{1/2}(l-1)O_l^{-2} = 0, \tag{3.3}$$

$$\tfrac{1}{2}(6)^{1/2}(l+3)Q_{1-2}^0 O_l^{-2} + 12Q_{l-1}^{-1} O_l^{-1} - \tfrac{1}{2}(6)^{1/2} \times (l-1)O_l^{-2} Q_l^0 + 216(6)^{1/2}(l-1)I_2 O_l^{-2} - 6(6)^{1/2}(l-1)(4l^3+8l^2-20l+39)O_1^{-2} = 0; \tag{3.4}$$

with $s = -1$,

$$(l-5)O_{l-1}^0 Q_l^{-1} - (l-1)Q_l^{-1} O_l^0 + 12O_{1-2}^{+1} Q_l^{-2}/(l-1)^2 - 216(6)^{1/2}(l-1)(l-2)I_2 O_l^{-1} + 18(6)^{1/2} l(l-1)(l-2)(1-4)O_l^{-1} = 0, \tag{3.5}$$

$$-\tfrac{1}{2}(6)^{1/2}(l-5)Q_{l-1}^0 O_l^{-1} + \tfrac{1}{2}(6)^{1/2}(l-1)O_l^{-1} Q_l^0 - 12Q_{l-2}^{+1} O_l^{-2}/(l-1)^2 + 216(6)^{1/2}(l-1)(l-2)I_2 O_l^{-1} - 6(6)^{1/2} l(l-1)(4l^2-32l-3)O_l^{-1} = 0, \tag{3.6}$$

$$(l+4)O_{l-1}^0 Q_l^{-1} - lQ_l^{-1} O_l^0 - 2(6)^{1/2} O_l^{-1} Q_l^0 + 72(6)^{1/2} l(l+4)I_2 O_l^{-1} - 6(6)^{1/2} l(l^3+2l^2+28l+18)O_l^{-1} = 0; \tag{3.7}$$

and with $s = 0$,

$$\tfrac{1}{2}(6)^{1/2}(l-4)O_l^0 Q_l^0 - \tfrac{1}{2}(6)^{1/2} l Q_l^0 O_l^0 - 36O_{l-1}^{+1} Q_l^{-1}/l^2 - 144(6)^{1/2} l(l-2)I_2 O_l^0 + 2\times 6^5 l^2(l+1)(2l-1)I_3 - 12(6)^{1/2} l(l+1)(2l^2+2l+9)O_l^0 = 0, \tag{3.8}$$

$$\tfrac{1}{2}(6)^{1/2}(l-4)Q_l^0O_l^0-\tfrac{1}{2}(6)^{1/2}lO_l^0Q_l^0-36Q_{l-1}^{+1}O_l^{-1}/l^2-144(6)^{1/2}l(l-2)I_2O_l^0$$
$$+2\times 6^5l^2(l+1)(2l-1)I_3-12(6)^{1/2}l(l+1)(2l^2+2l+9)O_l^0=0, \tag{3.9}$$

$$\tfrac{1}{2}(6)^{1/2}(2l-5)O_l^0Q_l^0-\tfrac{1}{2}(6)^{1/2}(2l-1)Q_l^0O_l^0+36O_{l-2}^{+2}Q_l^{-2}/l^2(l-1)^2-144(6)^{1/2}(2l-1)(2l-3)I_2O_l^0$$
$$-2\times 6^5l(2l-1)^2(l-3)I_3+12(6)^{1/2}l(10l^3-52l^2+49l-33)O_l^0=0, \tag{3.10}$$

$$\tfrac{1}{2}(6)^{1/2}(2l-5)Q_l^0O_l^0-\tfrac{1}{2}(6)^{1/2}(2l-1)O_l^0Q_l^0+36Q_{l-2}^{+2}O_l^{-2}/l^2(l-1)^2$$
$$-144(6)^{1/2}(2l-1)(2l-3)I_2O_l^0-2\times 6^5l(2l-1)^2(l-3)I_3+12(6)^{1/2}l(10l^3-52l^2+491-33)O_l^0=0. \tag{3.11}$$

There is a certain connection between the newly derived relations (3.1)–(3.11) and the relations between triple product operators previously given by Hughes.[1] For example, Eq. (3.8) is reproduced by combining Eqs. (40) and (42) of Ref. 1. Indeed, if we eliminate the term $(O_l^0)^3$ from Eqs. (40) and (42) or Ref. 1, and transform the remaining triple product combination as follows:

$$\frac{O_{l-1}^{+1}O_{l-2}^{+1}O_l^{-2}}{l^2(l-1)^2}-\frac{O_{l-1}^{+1}O_{l+1}^{-2}O_l^{+1}}{l^2(l+1)^2}$$
$$=-\frac{O_{l-1}^{+1}}{l}\left(\frac{O_{l+1}^{-2}O_l^{+1}}{l(l+1)^2}\left(-\frac{O_{l-2}^{+1}O_l^{-2}}{l(l-1)^2}\right)\right.$$
$$=-4(6)^{1/2}O_{l-1}^{+1}Q_l^{-1}/l,$$

where we made use of (2.17), we obtain exactly Eq. (3.8). In the same way, Eq. (3.9) can be reproduced from Eq. (4.1) and (4.3) of Ref. 1. An important feature of the relations (3.8)–(3.11) is that I_3 comes in and it is known[3] already that one needs the I_3-eigenvalue to calculate the O_l^0-eigenvalue. Hence, in order to introduce I_3, Hughes needed triple product operators of ninth degree in the generators. We, however, can introduce I_3 in relations that contain only operators of seventh order in the generators, which are much more easy to construct. As a by-product, we obtain from Eqs. (3.8)–(3.11) the following simple relations between shift operator "commutators":

$$O_{l-1}^{+1}Q_l^{-1}-Q_{l-1}^{+1}O_l^{-1}$$
$$=(1/6(6)^{1/2})l^2(l-2)[O_l^0,Q_l^0], \tag{3.12}$$

$$O_{l+1}^{-1}Q_l^{+1}-Q_{l+1}^{-1}O_l^{+1}$$
$$=-(1/6(6)^{1/2})(l+1)^2(l+3)[O_l^0,Q_l^0], \tag{3.13}$$

$$O_{l-2}^{+2}Q_l^{-2}-Q_{l-2}^{+2}O_l^{-2}$$
$$=-(1/6(6)^{1/2})l^2(l-1)^2(2l-3)[O_l^0,Q_l^0], \tag{3.14}$$

$$O_{l+2}^{-2}Q_l^{+2}-Q_{l+2}^{-2}O_l^{+2}$$
$$=(1/6(6)^{1/2})(l+1)^2(l+2)^2(2l+5)[O_l^0,Q_l^0], \tag{3.15}$$

4. RELATIONS BETWEEN PRODUCT OPERATORS OF THE TYPE $Q_{l+j}^kQ_l^j$

It is obvious that there also exist relations between product operators of the type $Q_{l+j}^kQ_l^j(-2\leqslant j,k\leqslant +2)$. The way to construct them is very similar to the one explained in Sec. 3. Due to the fact that a commutator $[q_\mu,q_\nu]$ contains no q-term, it is easy to predict that an $O_{l+m}^nO_l^m$-operator $(m+n=s)$, a Q_l^s-operator, and, if $s=0$, a term in $(I_2)^2$, I_2, and L^2 can appear in the relation.

Finally we obtain one independent relation with $s=-3$, one with $s=-2$, two with $s=-1$, and three with $s=0$. Again we do not give them all explicitly: the missing ones can be easily found from the relations (4.1)–(4.5) by means of the transformation rule "$l\to -l-s-1$."

The results are (valid when acting on $m=0$ states):

with $s=-3$,

$$Q_{l-1}^{-2}Q_l^{-1}-Q_{l-2}^{-1}Q_l^{-2}=0; \tag{4.1}$$

with $s=-2$,

$$-\tfrac{1}{2}(6)^{1/2}(l-1)(l-2)Q_l^{-2}Q_l^0$$
$$+6(2l-1)Q_{l-1}^{-1}Q_l^{-1}+\tfrac{1}{2}(6)^{1/2}l(l+1)Q_{l-2}^0Q_l^{-2}$$
$$+56l(l-1)(2l-1)O_{l-1}^{-1}O_l^{-1}$$
$$-12(6)^{1/2}l(l-1)(l-2)(l+1)(2l-1)Q_l^{-2}=0; \tag{4.2}$$

with $s=-1$,

$$-\tfrac{1}{2}(6)^{1/2}(l-1)(2l-3)Q_l^{-1}Q_l^0$$
$$+\tfrac{1}{2}(6)^{1/2}(2l-1)(l-3)Q_{l-1}^0Q_l^{-1}$$
$$+[6l/(l-1)^2]Q_{l-2}^{+1}Q_l^{-2}$$
$$-54[l(2l-1)/(l-1)]O_{l-2}^{+1}O_l^{-2}$$
$$+6(6)^{1/2}l(l-1)(2l^3-22l^2-3l+27)Q_l^{-1}=0 \tag{4.3}$$

with $s=0$,

$$\tfrac{3}{2}(2l+1)(Q_l^0)^2-[3l(l-2)/(l+1)^2]Q_{l+1}^{-1}Q_l^{+1}$$
$$+[3(l+1)(l+3)/l^2]Q_{l-1}^{+1}Q_l^{-1}$$
$$-6^5l^2(l+1)^2(2l+1)(I_2)^2$$
$$+6l(l+1)(2l+1)(O_l^0)^2$$
$$+27l(l+1)(2l+1)(l^2+l+8)Q_l^0$$
$$-648l^2(l+1)^2(2l+1)(l^2+l+6)I_2$$
$$+54l^2(l+1)^2(2l+1)(2l^4+4l^3+35l^2$$
$$+33l+135)=0,$$

$$\tfrac{3}{2}(l+2)^2(Q_l^0)2 + [3(2l+3)(2l+5)/(l+1)^2]$$
$$\times Q_{l+1}^{-1}Q_l^{+1} + [3(l+3)/(l+1)(l+2)^2]Q_{l+2}^{-2}Q_l^{+2}$$
$$- 6^5(l+1)^2(l+2)^2(2l+3)^2(I_2)^2$$
$$- 6(l+1)(l+2)(2l+3)(O_l^0)^2$$
$$- 9(l+1)(l+2)$$
$$\times(16l^3 + 43l^2 + 39l + 36)Q_l^0$$
$$+ 648(l+1)^2(l+2)(2l+3)^2$$
$$\times(2l^3 + 11l^2 + 9l - 6)I_2$$
$$- 54l(l+1)^2(l+2)(4l^6 + 48l^5$$
$$+ 162l^4 + 304l^3 + 312l^2 + 99l + 108) = 0. \qquad (4.5)$$

When looking back on (2.15)–(2.19), we may conclude that Eqs. (4.1)–(4.5) are in fact equivalent to relations between quartic product operators of the type $O^n_{l+j+k+m}O^m_{l+j+k}O^k_{l+j}O^j_l$. Without the Q^k_l-operators, it would have been unimaginably difficult to find relations between such quartic operator products. Notice, that in Eqs. (4.1)–(4.5), the coefficients of the $Q^j_{l+k}Q^k_l$-operators are the same as the corresponding ones of the $O^j_{l+k}O^k_l$-operators in Eqs. (2.2)–(2.8) of Ref. 3. This is because the Q^k_l-operators can be constructed out of the O^k_l-operators by replacing $q_\mu \to q^{(2)}_\mu$ $(-2 \leqslant \mu \leqslant +2)$. One might think that for this reason the equations in $Q^j_{l+k}Q^k_l$- and $O^j_{l+k}O^k_l$-objects should coincide completely. This is not true, since the commutators $[q^{(2)}_\mu, q^{(2)}_\nu]$ are not the same as the $[q_\mu, q_\nu]$ commutators. Moreover, the operators $q^{(2)}_\mu$ $(\mu = -2,\ldots,+2)$, l_0, $l_\pm$ do not even form a Lie algebra. For instance

$$[q^{(2)}_0, q^{(2)}_{-1}] = 18q^{(2)}_{-1} - 6(6)^{1/2}q^{(2)}_0 l_- - 6q^{(2)}_{-2}l_+ + 6q^{(2)}_{-1}l_0$$
$$- 48(6)^{1/2}q_{-2}q_{+1} - 3(6)^{1/2}q_{+1}q_{-1}l_-$$
$$- 144(6)^{1/2}q_{+2}q_{-2}l_- + 36q_{-2}q_0l_+$$
$$+ 24(6)^{1/2}q_{-2}q_{+1}l_0 - (243(6)^{1/2}/2)l_-l_0$$
$$- (675\sqrt{3}/4\sqrt{2})l_-.$$

These commutators explain the extra terms in (4.1)–(4.5).

5. APPLICATION OF THE SHIFT OPERATOR PRODUCT RELATIONS

In this section we show how the introduced relations of the previous sections simplify the O^0_l- and Q^0_l-eigenvalue determination for general (p,q) representations of SU(3). Before considering the maximum l state, we first remark that since the representations (p,q) and $(p,p-q)$ are mutually contragradient, it suffices to consider representations (p,q) where $p \geqslant 2q$.

For reasons, explained in Sec. 2, we restrict our attention to SU(3) states which correspond to $m = 0$, and employ the kets $|l,\lambda^{(i)}_l\rangle$, where $\lambda^{(i)}_l$ is the O^0_l-eigenvalue:

$$\lambda^{(i)}_l = \langle l,\lambda^{(i)}_l|O^0_l|l,\lambda^{(i)}_l\rangle \quad (i = 1,\ldots,n). \qquad (5.1)$$

If there is no degeneracy $(n = 1)$, we simply write $|l\rangle$. The eigenstates of Q^0_l will be denoted by $|l,\mu^{(i)}_l\rangle$ $(i = 1,\ldots n)$, where

$$\mu^{(i)}_l = \langle l,\mu^{(i)}_l|Q^0_l|l,\mu^{(i)}_l\rangle \qquad (5.2)$$

A. The maximum l-state of (p,q)

Until now, one needed at least on triple product relation in order to calculate λ_p.[2–4] With the aid of Eqs. (3.1)–(3.11) things are going much easier. Indeed, if we let the transformed relations (3.8) and (3.10) act on the $|p\rangle$ state, and if we multiply on the left by $\langle p|$, we obtain respectively,

$$\tfrac{1}{2}(6)^{1/2}(p+5)\lambda_p\mu_p - \tfrac{1}{2}(6)^{1/2}(p+1)\mu_p\lambda_p + 144(6)^{1/2}(p+1)(p+3)\langle I_2\rangle\lambda_p$$
$$- 2\times 6^5(p+1)^2p(2p+3)\langle I_3\rangle + 6(6)^{1/2}p(p+1)(4p^2+4p-9)\lambda_p = 0, \qquad (5.3)$$
$$\tfrac{1}{2}(6)^{1/2}(2p+7)\lambda_p\mu_p - \tfrac{1}{2}(6)^{1/2}(2p+3)\mu_p\lambda_p + 144(6)^{1/2}(2p+3)(2p+5)\langle I_2\rangle\lambda_p$$
$$+ 2\times 6^5(p+1)(2p+3)^2(p+4)\langle I_3\rangle - 6(6)^{1/2}(p+1)(2p+3)(10p^2+67p+96)\lambda_p = 0. \qquad (5.4)$$

$\langle I_2\rangle$ and $\langle I_3\rangle$ are short notations for

$$\langle I_2\rangle = \langle p|I_2|p\rangle = \langle p,q;l,\lambda_l,m|I_2|p,q;l,\lambda_l,m\rangle,$$
$$\langle I_3\rangle = \langle p|I_3|p\rangle = \langle p,q;l,\lambda_l,m|I_3|p,q,l;\lambda_l,m\rangle,$$

whose values are determined by (2.1) and (2.2). Eliminating the product $\lambda_p\mu_p$ from Eqs. (5.3) and (5.4), we obtain

$$[3(6)^{1/2}\langle I_2\rangle - (6)^{1/2}(p+1)(p+3)]\lambda_p$$
$$+ 324(p+1)(2p+3)\langle I_3\rangle = 0, \qquad (5.5)$$

which yields a unique solution

$$\lambda_p = +(6)^{1/2}(p+1)(2p+3)(p-2q). \qquad (5.6)$$

Remark that we did not need the Q^0_l-eigenvalue to determine λ_p. Out of Eq. (5.3) we obtain immediately,

$$\mu_p = -2(p+1)[2p^3 - 2(4q+3)p^2$$
$$+ (8q^2 - 12q + 27)p + 12q^2]. \qquad (5.7)$$

Once λ_p is known, we can calculate the λ_{p-1}, $\lambda^{(i)}_{p-2}$, $\lambda^{(i)}_{p-3}$, and $\lambda^{(i)}_{p-4}$ with the relations in $O^j_{l+k}O^k_l$-products. This is carried out in detail in II.

B. The eigenvalues μ_{p-1} and $\mu^{(i)}_{p-2}$

The eigenvalue μ_{p-1} can still be calculated without the help of a relation between $O^j_{l+k}Q^k_l$-objects. But the proper way to calculate the $\mu^{(i)}_{p-2}$ $(i = 1,2)$, a case where there is a twofold degeneracy, is by using the relations (3.1)–(3.11).

To obtain the μ_{p-1}, one can make use of Eqs. (45) and (46) of Ref. 1 and finally get [see Ref. 2, Eq. (14)]

$$\mu_{p-1} = \langle p-1|Q^0_{p-1}|p-1\rangle$$
$$= -2[2p^4 - 4(2q-1)p^3 + (8q^2 - 28q + 69)p^2$$
$$+ (28q^2 - 12q - 27)p + 12q^2]. \qquad (5.8)$$

For further calculations, we still need the action of Q_p^{-1} on the state $|p\rangle$. This is found immediately from Eq. (2.16), where we make use of (5.6), (II.3.2), and (II.3.3):

$$Q_p^{-1}|p\rangle = -2(6)^{1/2}p(p+1)(2p+1)(p-2q) \times[(p-q)q/(2p-1)]^{1/2}|p-1\rangle. \quad (5.9)$$

In the calculations, the following relationships[8] have been used:

$$\langle l,a_l,m|O_{l+k}^{-k}O_l^{+k}|l,a_l,m\rangle = \frac{1}{\beta_{kl}}\sum_{b_{l+k}}|\langle l+k,b_{l+k},m|O_l^{+k}|l,a_l,m\rangle|^2 = \beta_{kl}\sum_{b_{l+k}}|\langle l,a_l,m|O_{l+k}^{-k}|l+k,b_{l+k},m\rangle|^2, \quad (5.10)$$

with $\beta_{kl}=(2l+1)/(2l+k+1)$.

Note that the choice of the phase of the matrix element $\langle p-1|O_p^{-1}|p\rangle$ determines completely the sign in (5.9). To calculate the $\mu_{p-2}^{(i)}$, we will make use of the relations between product operators of the type $O_{l+k}^{j}Q_l^{k}$ and $Q_{l+j}^{k}O_l^{j}$. First of all, we obtain from Eqs. (2.16), (II.3.2), (II.4.1), and (II.4.9),

$$2(6)^{1/2}(p-1)Q_{p-1}^{-1}|p-1\rangle = 2(6)^{1/2}(2p+1)(p-2q)(b_{p-2}^{(1)}|p-2,\lambda_{p-2}^{(1)}\rangle + b_{p-2}^{(2)}|p-2,\lambda_{p-2}^{(2)}\rangle) - 6(6)^{1/2}\sqrt{\Gamma}(b_{p-2}^{(1)}|p-2,\lambda_{p-2}^{(1)}\rangle - b_{p-2}^{(2)}|p-2,\lambda_{p-2}^{(2)}\rangle), \quad (5.11)$$

where the notation of II is adopted. The subsitution of (II.4.11) in (5.11) gives

$$Q_{p-1}^{-1}|p-1\rangle = \{[p(p+3)(p^2+p-1)-4(p^2-2p-1)(p-q)q]|+\rangle + (p+3)(p-2q)|-\rangle\}/2p(2p+1)[q(p-q)(2p-1)]^{1/2}, \quad (5.12)$$

where $|+\rangle$ and $|-\rangle$ are short notations for

$$|\pm\rangle = a_{p-2}^{(1)}|p-2,\lambda_{p-2}^{(1)}\rangle \pm a_{p-2}^{(2)}|p-2,\lambda_{p-2}^{(2)}\rangle$$

and $a_{p-2}^{(i)}$ is defined in (II.4.1). From Eqs. (2.15) and (5.6) we obtain immediately,

$$2(6)^{1/2}(2p-1)Q_p^{-2} = 2(6)^{1/2}(p+1)\times(p-2q)|+\rangle - 6(6)^{1/2}\sqrt{\Gamma}|-\rangle. \quad (5.13)$$

To solve the eigenvalue problem it suffices to know the actions $Q_{p-2}^{0}|+\rangle$ and $Q_{p-2}^{0}|-\rangle$. A first relation follows from the action of Eq. (3.4) on $|p\rangle$:

$$\tfrac{1}{2}(6)^{1/2}(p+3)Q_{p-2}^{0}|+\rangle + 12Q_{p-1}^{-1}O_p^{-1}|p\rangle - \tfrac{1}{2}(6)^{1/2}(p-1)O_p^{-2}Q_p^{0}|p\rangle + 216(6)^{1/2}(p-1)I_2|+\rangle - 6(6)^{1/2}(p-1)\times(4p^3+8p^2-20p+39)|+\rangle = 0.$$

By using (II.3.3), (5.7), (5.12), and (2.1), this relation ends up in as

$$Q_{p-2}^{0}|+\rangle = -2\{[(4p^5-62p^4+326p^3-487p^2+321p-78) - 4q(p-q)(4p^3+4p^2-19p-1)]|+\rangle + 12(p+1)(p-2q)\Gamma^{1/2}|-\rangle\}/(2p-1). \quad (5.14)$$

A second relation follows from the action of the "$l\to -1-s-1$" transformed relation (3.7) on the $|p-1\rangle$ state:

$$-(p-5)Q_{p-1}^{-1}O_{p-1}^{0}|p-1\rangle + (p-1)O_{p-2}^{0}Q_{p-1}^{-1}|p-1\rangle - 2(6)^{1/2}Q_{p-2}^{0}O_{p-1}^{-1}|p-1\rangle + 72(6)^{1/2}(p-1)(p-5)I_2O_{p-1}^{-1}|p-1\rangle - 6(6)^{1/2}(p-1)(p^3-5p^2+35p-49)O_{p-1}^{-1}|p-1\rangle = 0.$$

To reduce this equation to a useful form we employ Eqs. (II.3.2), (5.12), (II.4.1), (2.1), and (5.14). We finally obtain

$$Q_{p-2}^{0}|-\rangle = -24(p-2q)[(p^2+p-1)(7p^3-10p^2+6p-1) - 4q(p-q)p^2(p-2)]|+\rangle/(2p-1)\Gamma - 2[(4p^5+10p^4+182p^3-415p^2+321p-78) - 4q(p-q)(4p^3+4p^2+17p-1)]|-\rangle/(2p-1). \quad (5.15)$$

The diagonalization of (5.14) and (5.15) is straightforward. There follows:

$$\mu_{p-2}^{(i)} = -2[(2p^4-12p^3+121p^2-165p+78) - 4q(p-q)(p+1)(2p+1)] + (-1)^{i-1}24\gamma^{1/2} \quad (i=1,2), \quad (5.16)$$

where γ is defined by

$$\gamma = p^2(4p^4-4p^3+5p^2-2p+1) - 4q(p-q)\times(2p^4+5p^3-4p^2-2p+1) + 4q^2(p-q)^2p^2.$$

For the (6.2) representation, we obtain

$$\mu_4^{(i)} = -8[133\pm 6(14569)^{1/2}],$$

which is in accordance with the numerical results of Hughes.[2] In the same way, closed formulae can be calculated for $\mu_{p-3}^{(i)}$ $(i=1,2)$. But since our main object was to show the usefulness of the relations between operators of the type $O_{l+k}^{j}Q_l^{k}$, we do not like to report on that here.

C. The low angular momentum states

An intense study on these states was made in I. We want to show here how the triple product relations can again be avoided on account of the newly introduced relations. In the case that p is odd ($q\neq 0$, $q\neq 1$) or p is even and q is odd, we know that the eigenvalues λ_2 and $\lambda_3^{(i)}$ can be expressed in terms of λ_1 (see I.5.15 and I.5.21).

From (I.5.19), we obtain

$$\langle 1|O_2^{-1}O_1^{+1}|1\rangle = 864(4\langle I_2\rangle+1) - 4\lambda_1^2. \quad (5.17)$$

So, there follows from Eq. (2.18),

$$3\langle 1|Q_1^{0}|1\rangle = \lambda_1^2 - 2^2 3^2(12\langle I_2\rangle + 19). \quad (5.18)$$

If we let Eq. (3.8) act on $|1\rangle$, multiply on the left by $\langle 1|$, and substitute the result (5.18), we find the following equation:

$$\lambda_1[\lambda_1^2 - 216(3\langle I_2\rangle + 1)] - 2^5.3^5(6)^{1/2}\langle I_3\rangle = 0, \quad (5.19)$$

which is in agreement with (I.5.24). The solutions for λ_1 are written down in (I.5.25).

ACKNOWLEDGMENTS

We like to express our gratitude to Professor C. C. Grosjean for his continuing interest.

[1]J. W. B. Hughes, J. Phys. A **6**, 48 (1973).
[2]J. W. B. Hughes, J. Phys. A **6**, 281 (1973).
[3]H. E. De Meyer, G. Vanden Berghe, and J. W. B. Hughes, J. Math. Phys. **22**, 2360 (1981).
[4]H. E. De Meyer, G. Vanden Berge, and J. W. B. Hughes, J. Math. Phys. **22**, 2366 (1981).
[5]G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[6]G. Racah, in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, edited by F. Gursey (Gordon and Breach, New York, 1962).
[7]J. W. B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978).
[8]G. Vanden Berghe and H. E. De Meyer, J. Math Phys. **21**, 1902 (1980).
[9]G. Vanden Berghe and H. E. De Meyer, J. Math Phys. **21**, 1967 (1980).
[10]G. Vanden Berghe, H. E. De Meyer, and J. Van der Jeugt, Lett. Nuovo Cimento **33**, 120 (1982).

Indecomposable representations for para-Bose algebra

B. Gruber
Physics Department, Southern Illinois University, Carbondale, Illinois 62901
T. S. Santhanam[a)]
International Centre for Theoretical Physics, Trieste, Italy

(Received 1 March 1982; accepted for publication 10 May 1982)

A general study of the representations of the graded Lie algebra of para-Bose oscillators is given. Besides realizing the standard representations, we also find some interesting indecomposable (not fully reducible) representations.

PACS numbers: 02.20.Sv, 02.20.Qs

I. INTRODUCTION

The algebra of para-Bose oscillators is a prototype version of a graded Lie algebra. The Fock representation has earlier been obtained[1] by realizing that an enveloping algebra in this case is isomorphic to the algebra of Lorentz group SO(2,1) in three dimensions. The representations of the coherent states of para-Bose oscillators have recently been analyzed.[2,3] The present paper deals with an analysis of the representations of this graded Lie algebra of para-Bose oscillators which can be realized on the space of the universal enveloping algebra. Besides obtaining the standard representations we also find some very interesting indecomposable (not fully reducible) representations. We also exhibit some finite-dimensional representations. We follow the method and notations of the general analysis of indecomposable representations carried out by Gruber and Klimyk.[4] Since the method is quite general, it can easily be carried over to the study of other graded Lie algebras as well.

In this paper a general approach is taken in order to find the indecomposable representations for para-Bose oscillators. In Sec. II we summarize the known properties of the algebra of para-Bose oscillators, the Fock representations, and the representations of the coherent states. In Sec. III the most general representation of the algebra of para-Bose oscillators on the space of its enveloping algebra Ω is obtained. Representations which are induced on invariant subspaces as well as quotient spaces are discussed briefly. In Sec. IV, we show how the standard representations are realized in this method. We obtain some interesting indecomposable representations. We also exhibit some finite-dimensional (nonunitary) representations. In Sec. V we discuss the representations of coherent states and some possible generalizations.

II. PARA-BOSE OSCILLATORS

Para-Bose oscillators[5] satisfy (the commutation relations[6]) the equation of motion

$$[a,N]=a, \qquad [a^+,N]=-a^\dagger, \tag{1}$$

where a is the annihilation operator and the "number operator" N is defined by

$$2N=\{a,a^\dagger\}, \tag{2}$$

where the braces { } stand for the anticommutator. The creation operator $a^\dagger$ and a do not satisfy the commutation relation of the normal harmonic oscillator. It can easily be worked out by repeated use of Eqs. (1) and (2) that

$$[a,a^{\dagger 2K}]=2Ka^{\dagger 2K-1} \tag{3}$$

and

$$\{a,a^{\dagger 2K+1}\}=a^{\dagger 2K}(2K+\{a,a^\dagger\}). \tag{4}$$

The normal Fock representation has been obtained earlier[1] by recognizing that

$$\begin{aligned}[]
[H,\tfrac12 a^{\dagger 2}]&=\tfrac12 a^{\dagger 2},\\
[H,\tfrac12 a^2]&=-\tfrac12 a^2,\\
[\tfrac12 a^{\dagger 2},\tfrac12 a^2]&=-2H,
\end{aligned} \tag{5}$$

i.e., $\frac12 a^{\dagger 2}$, $\frac12 a^2$, and $H\equiv\frac12 N$ close, and the algebra is that of the Lorentz group SO(2,1). By using the standard representations of SO(2,1) and the fact that the spectrum of H is positive definite, the representation for a and $a^\dagger$ is obtained by extracting the square root.

If b_0 denotes the lowest value for the spectrum of the Hamiltonian $\mathscr{H}$ related to N by

$$\mathscr{H}=N+b_0, \tag{6}$$

then we obtain

$$\begin{aligned}
a_{2n,2n+1}&=[2(n+b_0)]^{1/2},\\
a_{2n-1,2n}&=(2n)^{1/2}
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
\langle 2n|[a,a^\dagger]|2n\rangle&=2b_0,\\
\langle 2n+1|[a,a^\dagger]|2n+1\rangle&=2(1-b_0),
\end{aligned} \tag{8}$$

so that

$$a=\begin{pmatrix}0 & (2b_0)^{1/2} & 0 & 0 & 0\\ 0 & 0 & 2^{1/2} & 0 & 0\\ 0 & 0 & 0 & [2(b_0+1)]^{1/2} & 0\\ & & & & \ddots\end{pmatrix}. \tag{9}$$

It is clear that for $b_0=\frac12$, the distinction between odd and even matrix elements disappear and we get the standard harmonic oscillator. Properly normalized para-Bose coherent states (eigenstates of a) are obtained as[2,3]

$$|\alpha\rangle=\{D(|\alpha|^2\}^{-1/2}D(\alpha a^\dagger)|0\rangle, \tag{10}$$

where

$$\langle Z|D|0\rangle=(\alpha Z)^{1-b_0}\{I_{b_0}(\alpha Z)+I_{b_0-1}(\alpha Z)\}, \tag{11}$$

where I_ν is the modified Bessel function of the νth order, and

[a)] Permanent address: Matscience, Madras 600020, India.

Z is a complex number. Here $|0\rangle$ is the vacuum (extremal state annihilated by a).

III. REPRESENTATIONS ON Ω

In this section the most general representation of a para-Bose algebra on the space of its universal enveloping algebra Ω will be determined. According to the Poincaré–Birkhoff–Witt theorem a basis for the universal enveloping algebra of the para-Bose algebra can be chosen as[7]

$$\Omega:\{a^{\dagger n}a^m N^r, m,n,r = 0,1,2,\cdots\} \equiv \{X(n,m,r)\}, \tag{12}$$

where $X(n,m,r)$ is an ordered (tensor) product of the elements of $a^\dagger, a$ and N. The values $(n,m,r) = (0,0,0)$ denote the identity operator $\mathbb{1}$ (this corresponds to the vacuum). An element $y \in \Omega$ is called an extremal vector for the representation ρ on Ω if

$$\rho(\beta)y = 0, \quad \beta = a \text{ or } a^\dagger \tag{13}$$

for $\rho(a^\dagger)$ and/or $\rho(a)$. The basis for the universal enveloping algebra can be written as[7]

$$\Omega = \Omega_+ \Omega_- \Omega_N, \tag{14}$$

where $\Omega_+ = a^{\dagger n}, \Omega_- = a^m$, and $\Omega_N = N^r$.

The basic commutation relations Eqs. (1) and (2) can be used to derive

$$a^m N^r = (N+m)^r a^m,$$
$$a^{\dagger n} N^r = (N-n)^r a^{\dagger n}, \tag{15}$$

which in turn can be used to get the following basic relations:

$$\rho(a^\dagger) X(n,m,r) = X(n+1,m,r), \tag{16}$$
$$\rho(a)\, X(2K,m,r) = 2K\, X(2K-1,m,r) + X(2K,m+1,r), \tag{17}$$
$$\rho(a)\, X(2K+1,m,r) = 2(K-m)\, X(2K,m,r) + 2\, X(2K,m,r+1) - X(2K+1,m+1,r), \tag{18}$$
$$\rho(N)\, X(n,m,r) = (n-m)\, X(n,m,r) + X(n,m,r+1). \tag{19}$$

This representation is in general infinite dimensional, with neither a highest nor a lowest weight. Under the action of the operator ρ of this representation the powers m and r remain the same or increase. Thus, each of the subspaces $V(m,r)$ of Ω,

$$V(m,r):\{X(n,m+k,r+k_2)\, m \quad n,k_1,k_2 = \text{non-negative integers}\}, \tag{20}$$

is an invariant subspace with respect to the action of ρ on Ω, and induces subrepresentations on these invariant subspaces. Actually ρ induces representations on the quotient spaces

$$V(m,r)/V(m',r'),$$
$$m \leqslant m', \quad r \leqslant r', \quad V(0,0) = \Omega. \tag{21}$$

The representations which are induced on the invariant subspaces $V(m,r)$ are all algebraically equivalent to the representation ρ. The representations induced by ρ on the quotient spaces

$$\Omega / V(m,r) \tag{22}$$

are algebraically inequivalent to ρ. The representations on these quotient spaces are obtained from Eqs. (16)–(19) by formally setting

$$X(n,m+k_1,r+k_2) \to 0, \quad k_1,k_2 \geqslant 0 \text{ integers}.$$

We will discuss some of the representations induced on the quotient spaces given by Eq. (20) in the following sections.

IV. REPRESENTATION ON $\Omega_+\Omega_-$

The representations on $\Omega_+\Omega_-$ are defined through the relation

$$\rho(N)\mathbb{1} = \lambda \mathbb{1}, \quad \lambda \in \mathbb{C}, \tag{23}$$

i.e., the relation $(N-\lambda)\mathbb{1}$ generates a left ideal I_N of Ω. Then $\Omega|_{I_N} \sim \Omega_+\Omega_-$. A basis for $\Omega_+\Omega_-$ can be chosen as

$$\Omega_+\Omega_-:\{X(n,m), n,m \geqslant 0 \text{ integers}\}. \tag{24}$$

On this space $\Omega_+\Omega_-$, the representation ρ induces the representation ρ':

$$\rho'(a^\dagger)\, X(n,m) = X(n+1,m),$$
$$\rho'(a)\, X(2K,m) = 2K\, X(2K-1,m) + X(2K,m+1),$$
$$\rho'(a)\, X(2K+1,m) = 2(K-m+\lambda)\ X(2K,m) - X(2K+1,m+1), \tag{25}$$
$$\rho'(N)\, X(n,m) = (\lambda + n - m)\, X(n,m).$$

The representation ρ' is infinite dimensional and the operator $\rho'(N)$ is diagonal, and has no extremal vectors. The subspaces

$$V(m):\{X(n,m+K), K \geqslant 0 \text{ integers}\} \tag{26}$$

are invariant with respect to ρ'. The representations which are induced by ρ' on the quotient spaces $\Omega_+\Omega_-/V(m)$ are obtained from the representation Eq. (25) by setting formally $X(n,m+K) \to 0$, for all $K \geqslant 0$ (integers). We shall now consider the special case of representations on $\Omega_+\Omega_-/V(m=0) \sim \Omega_+$. A basis for this is given by

$$\Omega_+:\{a^{\dagger n}, n = 0,1,2,\cdots\}. \tag{27}$$

We obtain the representation $\rho_{\lambda,0}$ from Eq. (25) as (suppressing the indices which should cause no confusion)

$$\rho(N)1 = \lambda \mathbb{1},$$
$$\rho(a^\dagger)\, X(n) = X(n+1),$$
$$\rho(a)\, X(2K) = 2K\, X(2K-1), \tag{28}$$
$$\rho(a)\, X(2K+1) = 2(K+\lambda)\, X(2K),$$
$$\rho(N)\, X(n) = (\lambda + n)\, X(n).$$

It can easily be verified that Eq. (28) satisfies the basic commutation relations (1) and (2). Moreover it holds for the *commutators*

$$[\rho(a),\rho(a^\dagger)]X(2K+1) = 2(1-\lambda)X(2K+1), \tag{29}$$
$$[\rho(a),\rho(a^\dagger)]X(2K) = 2\lambda X(2K), \tag{30}$$

and as before we realize the standard oscillator for $\lambda = \frac{1}{2}$ (actually $\lambda = b_0$). To get a more symmetrical form for a and $a^\dagger$ for the case of *irreducible* representations ($-\lambda \notin \mathbf{N}$) we define

$$
\begin{aligned}
&Y(0) = X(0),\\
&Y(1) = (1/\sqrt{2\lambda})X(1),\\
&Y(2K) = \Big\{\prod_{j=1}^{K}(2j)\Big\}^{-1/2}\Big\{\prod_{i=0}^{K-1}2(i+\lambda)\Big\}^{-1/2}X(2K), \qquad (31)\\
&Y(2K+1) = \Big\{\prod_{j=1}^{K}(2j)\Big\}^{-1/2}\Big\{\prod_{i=0}^{K}2(i+\lambda)\Big\}^{-1/2}X(2K+1),
\end{aligned}
$$

for $K = 1,2,3,\cdots$ to obtain

$$
\begin{aligned}
&\rho(a^\dagger)\, Y(2K) = [2(K+\lambda)]^{1/2}\, Y(2K+1),\\
&\rho(a^\dagger)\, Y(2K+1) = [2(K+1)]^{1/2}\, Y(2K+2),\\
&\rho(a)\, Y(2K) = (2K)^{1/2}\, Y(2K-1), \qquad (32)\\
&\rho(a)\, Y(2K+1) = [2(K+\lambda)]^{1/2}\, Y(2K),\\
&\rho(N)\, Y(n) = (\lambda+n)\, Y(n),
\end{aligned}
$$

which is the standard representation given in Eq. (9). It is interesting to note that we did not extract any square root nor did we explicitly use SO(2,1). These are automatically defined in the enveloping algebra. To appreciate the power of the general method, let us construct some novel representations on the quotient space $\rho_{\lambda 0}/V(m=1)$. The basis is constructed as

$$\{X(n) \equiv a^{\dagger n},\ Y(n) \equiv a^{\dagger n}a,\ n \geqslant 0 \text{ integers}\}. \qquad (33)$$

From the physical point of view, this amounts to taking two vacua, the usual one and the one particle state as the second vacuum. Equation (25) yields the following representations:

$$\rho(a^\dagger)\, X(n) = X(n+1), \qquad (34a)$$

$$\rho(a^\dagger)\, Y(n) = Y(n+1), \qquad (34b)$$

$$\rho(a)\, X(2K) = 2K\, X(2K-1) + Y(2K), \qquad (34c)$$

$$\rho(a)\, Y(2K) = 2K\, Y(2K-1), \qquad (34d)$$

$$\rho(a)\, X(2K+1) = 2(K+\lambda)\, X(2K) - Y(2K+1), \qquad (34e)$$

$$\rho(a)\, Y(2K+1) = 2(K+\lambda-1)\, Y(2K), \qquad (34f)$$

$$\rho(N)\, X(n) = (\lambda+n)\, X(n), \qquad (34g)$$

$$\rho(N)\, Y(n) = (\lambda+n-1)\, Y(n). \qquad (34h)$$

The X and Y are almost decoupled except for the important relations Eqs. (34c) and (34e) which couple these two. The situation can be best explained by Fig. 1.

The representation for a, $a^\dagger$ and N can easily be seen to be the following:

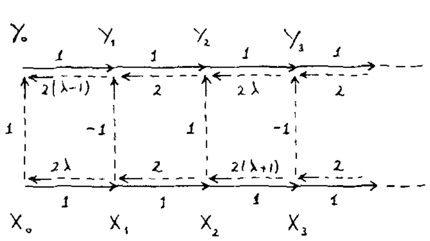

FIG. 1. Representation induced by $\rho_{\lambda 0}$ on the quotient space $\Omega_+/V(m=1)$. The action of $a^\dagger$ is given by solid lines. The action of a is given by dashed lines. The numbers given are the matrix elements of the transition.

$$\rho(a^\dagger) = \begin{bmatrix} 0_2 & 0_2 & & \\ I_2 & 0_2 & & \\ 0_2 & I_2 & 0_2 & \\ 0_2 & 0_2 & I_2 & \ddots \end{bmatrix}, \qquad (35)$$

$$\rho(N) = \begin{bmatrix} \begin{matrix}\lambda & 0\\ 0 & \lambda-1\end{matrix} & & & \\ & \begin{matrix}\lambda+1 & 0\\ 0 & \lambda\end{matrix} & & \\ & & \ddots & \\ & & & \begin{matrix}\lambda+n & 0\\ 0 & \lambda+n-1\end{matrix} \end{bmatrix}, \qquad (36)$$

$$\rho(a) = \begin{bmatrix} \begin{matrix}0 & 0\\ 1 & 0\end{matrix} & \begin{matrix}2\lambda & 0\\ 0 & 2(\lambda-1)\end{matrix} & \\ & \begin{matrix}0 & 0\\ -1 & 0\end{matrix} & \begin{matrix}2 & 0\\ 0 & 2\end{matrix} \\ & & \begin{matrix}0 & 0\\ 1 & 0\end{matrix} \end{bmatrix}, \qquad (37)$$

where

$$0_2 = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}, \qquad I_2 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

and each block acts on the two components

$$\begin{bmatrix} X(n)\\ Y(n)\end{bmatrix}, \quad n = 0,1,2,\cdots.$$

One can easily verify that this representation satisfies the basic commutation relations Eqs. (1) and (2).

From Fig. 1 it is clear that Y_0 is an extremal vector. To see whether there are other extremal vectors (to give rise to an indecomposable representation) we realize that

$$\rho(a)(X_0 + \xi Y_1) = 0, \quad Y_0 + \xi\, 2(\lambda-1)Y_0 = 0, \qquad (38)$$

which gives

$$\xi = -1/2(\lambda-1), \qquad (39)$$

and so

$$
\begin{aligned}
&\rho(a)Y_0 = 0, \qquad (40)\\
&\rho(a)Z = 0, \quad Z \equiv X_0 - (1/2(\lambda-1))\, Y_1.
\end{aligned}
$$

However, the representation is reducible (decomposable) to the sum of

$$
\begin{aligned}
&\{Y_n = a^{\dagger n}Y_0,\ n = 0,1,2,\cdots\},\\
&\text{and} \qquad (41)\\
&\{Z_n = a^{\dagger n}Z_0,\ n = 0,1,2,\cdots\},
\end{aligned}
$$

each of these resulting in the correct commutation relation for the case $\lambda \neq 1$. For $\lambda = 1$, the representation becomes truely indecomposable since ξ blows up making the combination Z_0 not possible, and thus Y_0 *cannot* be reached from Y_1, while Y_1 can be reached from Y_0.

Let us now study the interesting representations which are induced by Eqs. (34) on the quotient space modulo the invariant subspace spanned by the basis elements Y. We find that (setting formally $Y \to 0$)

$$\rho(a^\dagger)\,X(n) = X(n+1),$$
$$\rho(a)\,X(2K) = 2K\,X(2K-1),$$
$$\rho(a)\,X(2K+1) = 2(\lambda + K)\,X(2K),$$
$$\rho(N)\,X(n) = (\lambda + n)\,X(n). \tag{42}$$

From the second and the third of these equations it is clear that for $\lambda = -l$, $l \geqslant 0$ integers, we realize indecomposable representations since $\rho(a)$ does not take $X(2l+1)$ to $X(2l)$. This can be easily visualized from Fig. 2. The point is that while $a^\dagger$ takes the states continuously up, the action of a stops at the point $l+1$ (i.e., $2l+2$ steps from $-l$), i.e., while $a^\dagger$ takes $l \to l+1$, a will not trace back $l+1 \to l$. Equations (42) in turn induce a representation on the quotient space modulo the invariant subspace which is spanned by the basis elements $Y(2l+1)$, $Y(2l+2)$,.... This representation is irreducible and of dimension $2l+1$. On the *quotient space* Eqs. (42) can be cast in a more symmetrical form by defining

$$W(0) = X(0),$$
$$W(1) = (1/\sqrt{-2l}\,)X(1),$$
$$W(2K) = \left\{\prod_{r=1}^{K}(2r)\right\}^{-1/2}\left\{\prod_{r=0}^{K-1}2(-l+r)\right\}^{-1/2} X(2K),$$
$$K = 1,2,\ldots,l,$$
$$W(2K+1) = \left\{\prod_{r=1}^{K}(2r)\right\}^{-1/2} \times \left\{\prod_{r=0}^{K}2(-l+r)\right\}^{-1/2} X(2K+1), \tag{43}$$
$$K = 1,2,\ldots,l-1,$$

and we get

$$\rho(a^\dagger)W(2K) = [2(-l+K)]^{1/2}W(2K+1),$$
$$\rho(a^\dagger)W(2K+1) = [2K+2]^{1/2}W(2K+2),$$
$$\rho(a)W(2K) = [2K]^{1/2}W(2K-1), \tag{44}$$
$$\rho(a)W(2K+1) = [2(-l+K)]^{1/2}W(2K),$$
$$\rho(N)W(n) = (-l+n)W(n).$$

To illustrate the point let us look at the explicit representation for $\lambda = -l = -2, l = 2$ when we get

$$\rho(a^\dagger) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \sqrt{-4} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{-2} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{4} & 0 \end{pmatrix},$$
$$\rho(a) = [\rho(a^\dagger)]^\dagger,$$

and

$$\rho(N) = \begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}. \tag{45}$$

For $l = 1$ we get

$$\rho(a^\dagger) = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{-2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}, \quad \rho(a) = [\rho(a^\dagger)]^\dagger,$$
$$\rho(N) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The situation is reminiscent of para-Fermi algebra (its isomorphism with angular momentum algebra is well known[8]). It can easily be verified that Eq. (45) yields the basic commutation relations. In fact it has been shown earlier[9] that for SU(2), given the spin j, there are two extremal vectors with projections of j equal to $-j$ and $j+1$ and that the usual angular momentum representation is obtained in the $(2j+1)$ finite-dimensional quotient space.

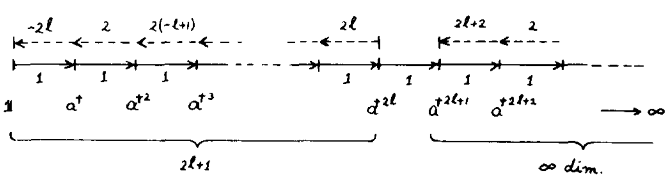

FIG. 2. Indecomposable representation for $\lambda = -l, l \geqslant 0$ integer. The action of $a^\dagger$ is denoted by solid lines and that of a by dotted lines.

V. COHERENT STATES

The representations on $\Omega_+\Omega_N$ are defined through the relation

$$\rho(a)\mathbf{1} = \mu\mathbf{1}, \quad \mu \in \mathbb{C}, \tag{46}$$

i.e., the relation $(a-\mu)\mathbf{1}$ generates a left ideal I_a of Ω. Then $\Omega/I_a \sim \Omega_+\Omega_N$, for which a basis can be chosen as

$$\Omega_+\Omega_N : \{X(n,r), n,r \geqslant 0 \text{ integers}\}. \tag{47}$$

On this space is induced the following representation:

$$\rho(a)\mathbf{1} = \mu\mathbf{1},$$
$$\rho(a^\dagger)\,X(n,r) = X(n+1,r),$$
$$\rho(a)\,X(2K,r) = 2K\,X(2K-1,r) + \mu\sum_{l=0}^{r}\binom{r}{l}X(2K,l),$$
$$\rho(a)\,X(2K+1,r) = 2K\,X(2K,r) + 2\,X(2K,r+1) - \mu\sum_{l=0}^{r}\binom{r}{l}X(2K+1,l), \tag{48}$$
$$\rho(N)\,X(n,r) = X(n,r+1) + n\,X(n,r).$$

This is again an infinite-dimensional representation and has no external vectors. Since Eq. (48) describes basically the action of a, $a^\dagger$, and N on coherent states (eigenstates of a), in order to solve for the coherent states we adopt a different procedure. Guided by Eq. (28) we have to solve the equation

$$\rho(a)\zeta = \sigma\zeta,$$
$$\zeta = \sum_{K=0}^{\infty}\{C_{2K}\,X(2K) + C_{2K+1}\,X(2K+1)\}. \tag{49}$$

Using Eq. (28) we get the recursion relations

$$C_{2K+1} = \left(\frac{\sigma}{2}\right)^2\frac{1}{K(K+\lambda)}C_{2K-1}, \tag{50}$$
$$C_{2K+2} = \left(\frac{\sigma}{2}\right)^2\frac{1}{(K+1)(K+\lambda)}C_{2K},$$

which can be solved to yield

$$C_{2K} = \left(\frac{\sigma}{2}\right)^{2K} \frac{1}{\prod_{r=0}^{K-1} (\lambda + r)K!} C_0, \tag{51}$$

$$C_{2K+1} = \left(\frac{\sigma}{2}\right)^{2K+1} \frac{1}{\prod_{r=0}^{K} (\lambda + r)K!} C_0.$$

Equation (51) can be recast in the form of Eqs. (10) and (11) and gives the coherent states of the para-Bose oscillators. A possible generalization is to look for the solution of the equation

$$\rho(a)E = \sigma E, \tag{52}$$

where

$$E = \sum_{n,m} C_{2n,m} X(2n,m) + \sum_{n,m} C_{2n+1,m} X(2n+1,m). \tag{53}$$

This leads to the recursion relations

$$(2n+2)C_{2n+2,m+1} - C_{2n+1,m} = \sigma C_{2n+1,m+1},$$
$$C_{2n,m-1} + 2(n-m+\lambda)C_{2n+1,m} = \sigma C_{2n,m}. \tag{54}$$

We have not succeeded in solving these except for the case when $\sigma = 0$. Should we succeed in solving them, we will realize some generalized coherent states.

VI. CONCLUSIONS

A general study of the indecomposable representations for the algebra of para-Bose oscillators is made. Besides realizing the standard representations, some novel indecomposable representations are derived.[10] We hope to extend this analysis to other graded algebras elsewhere.

ACKNOWLEDGMENTS

One of the authors (T. S. S.) would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for the hospitality at the International Centre for Theoretical Physics, Trieste.

[1]T. F. Jordan, N. Mukunda, and S. V. Pepper, J. Math. Phys. **4**, 1089 (1963); L. O. Raifeartaigh and C. Ryan, Proc. R. Ir. Acad. Sect. A **62**, 93 (1963).

[2]J. K. Sharma, C. L. Mehta, and E. C. G. Sudarshan, J. Math. Phys. **19**, 2089 (1978).

[3]S. N. Biswas and T. S. Santhanam, J. Aust. Math. Soc. B **22**, 210 (1980).

[4]B. Gruber and A. U. Klimyk, J. Math. Phys. **16**, 1816 (1975).

[5]H. S. Green, Phys. Rev. **90**, 270 (1953).

[6]For a general discussion see S. Kamefuchi, Matscience Report 24 (1963), Madras, India.

[7]N. Jacobson, *Lie Algebras* (Interscience, New York, 1962); J. Dixmier, *Algebras Enveloppantes* (Gauthiers Villars, Paris, 1974). Collection "Cabiers Scientifiques."

[8]C. Ryan and E. C. G. Sudarshan, Nucl. Phys. **47**, 207 (1963). Some finite-dimensional nonunitary representations of para-Bose algebra have recently been obtained by Miss S. N. Uma (unpublished) using Weyl's trick on the corresponding ones of para-Fermi algebra.

[9]B. Gruber, A. U. Klimyk, and Yu. F. Smirnov, Nuovo Cim. (in press).

[10]Professor E. C. G. Sudarshan has also found some indecomposable representations for para-Bose oscillators (unpublished). We thank him for his correspondence.

Harmonic polynomials invariant under a finite subgroup of O(*n*)

A. Ronveaux
Département de Physique, Facultés Universitaires Notre-Dame de la Paix, B-5000 Namur, Belgium
Y. Saint-Aubin
Centre de recherche de mathématiques appliquées, Université de Montréal, Montréal, H3C 3J7, Canada

(Received 28 January 1982; accepted for publication 9 April 1982)

In this paper, an algorithm is described which allows a systematic computation of harmonic polynomials of a given degree invariant under a finite subgroup of the group O(n). An application of the algorithm to the octahedral (cubic) subgroup is given.

PACS numbers: 02.20.Rt, 02.70. + d

1. INTRODUCTION

Computations of physical properties for systems invariant under a (nontrivial) space group or point group are made easier by the use of functions having the same invariance property. Great attention has been paid mainly by physicists working in spectroscopy, crystallography, solid state physics, or quantum chemistry to obtain such invariant functions in various cases (see Refs. 1–19 and references therein). Our task here is to describe an algorithm which allows the calculation of functions invariant under a point group of O(3) [and more generally under any finite subgroup of O(n)] and which are solutions of the Laplace equation. Such functions will be referred to as invariant harmonic polynomials (IHP).

Patera, Sharp, and Winternitz[10] have given a basis for all tensors transforming irreducibly under a given point group [subgroup of O(3)]. They use the standard technique of the so-called Molien function (to be surveyed in the next section). [Many of their results had appeared earlier but their paper is the first systematic and complete study for all the point groups of O(3).] Their Sec. VI describes a prescription to obtain a basis for the IHP out of their integrity basis. (Their technique is based on previous results of Lohe and Hurst.[12]) Our algorithm is in fact equivalent to their prescription; we shall give here the proofs which were skipped in their paper.

To obtain IHP, different approaches have been used. First, the IHP can be expressed as a linear combination of spherical harmonics $Y_l^m(\theta,\varphi)$ or in a Cartesian way as homogeneous polynomials in the variables x, y, z. Secondly, the literature gives analytical[14,15] as well as numerical[15–20] expansions. To our knowledge, the most complete results have been obtained by Dunkl[15] for the analytical aspect and by Fox and Krohn[19,20] for the numerical one. Dunkl has obtained, for the cubic subgroup of O(3), the general expansions of the IHP in terms of the Cartesian coordinates and in terms of Y_l^m. Moreover, he also obtains the expression for the tensors transforming under the alternate representation of the cubic subgroup. The contribution of Fox and Krohn has been to calculate numerically the coefficients of the expansion for the IHP of the cubic subgroup of degree up to 200. (An earlier work[18] dealt with the tetrahedral IHP.)

The next section shows how the information contained in the Molien function combined with the existence of an operator H, the harmonic projector, leads to a very natural algorithm for the computation of the IHP of any finite subgroup of O(n). By applying this algorithm to the case of the cubic IHP, we show in the third section the usefulness of the method. The conclusion points out advantages of the method.

2. THE ALGORITHM FOR THE GENERAL CASE O(*n*)

Roughly speaking, the algorithm splits into three simple steps: (i) compute the Molien function and the integrity basis associated with it. (ii) in order to obtain a basis for the IHP of degree n, compute all the products of an arbitrary number of elements of the integrity basis at the exception of those containing a power of x^2, and (iii) apply the operator H (to be defined below) on each of the products. The result is a basis for the IHP of degree n. The following subsections describe each of these steps.

A. The Molien function and the associated integrity basis

Let P be a vector field over R^n:

$P{:}R^n \to R^r$

the group O(n) acts naturally on R^n. Let G be a finite subgroup of O(n) and Γ_n the representation of G which is the restriction of O(n) to G. The polynomial $P(\mathbf{x})$ is said to be a Γ_r-tensor (Γ_r is an irreducible representation of G) if the following condition is verified:

$$P(\Gamma_n(g)\mathbf{x}) = \Gamma_r(g^{-1})P(\mathbf{x}) \quad \forall g{\in}G. \tag{1}$$

If Γ_r is simply the identity representation, the condition (1) is the usual invariance condition

$$P(\Gamma_n(g)\mathbf{x}) = P(\mathbf{x}) \quad \forall g{\in}G \tag{2}$$

and the Γ_{identity}-tensors are called the invariants.

The Molien function answers the following question: how many independent Γ_s-tensors of degree l are there? This information (for all n) is casted in the Taylor expansion of the Molien function (also called "generating function"):

$$B(\Gamma_n,\Gamma_r;\lambda) = \sum_{i=0}^{\infty} a_i \lambda^i. \tag{3}$$

The coefficient a_l is the number of Γ_r-tensors of degree l. If the representation Γ_n is irreducible [like the natural represenation of the cubic group as a subgroup of O(3)], the function $B(\Gamma_n,\Gamma_r;\lambda)$ is given by

 0022-2488/83/051037-04$02.50

$$B(\Gamma_n,\Gamma_r;\lambda)=\frac{1}{|G|}\sum_{\substack{\text{conjugacy}\\ \text{classes}}}\frac{N_s\chi^*_{sr}}{\det(1-\lambda A_s)}, \tag{4}$$

where $|G|$ is the number of elements of G, N_s the number of elements in the conjugacy classes s, χ^*_{sr} the complex conjugate of the character of the class s in the representation Γ_r and A_s the matrix representing one element of the class s in the representation Γ_n. (The technique of the Molien function is explained in greater details in Burnside.[21]) If the representation is reducible ($\Gamma_n=\Gamma_{n_1}\oplus\Gamma_{n_2}\oplus\cdots\oplus\Gamma_{n_m}$), the generating function is the product of the generating functions for the irreducible parts:

$$B(\Gamma_n,\Gamma_r;\lambda)=B(\Gamma_{n_1},\Gamma_r;\lambda)B(\Gamma_{n_2},\Gamma_r;\lambda)\cdots B(\Gamma_{n_m},\Gamma_r;\lambda). \tag{5}$$

The information contained in the form (3) seems to be the totality of what $B(\Gamma_n,\Gamma_r;\lambda)$ can tell us. However, the result of the explicit computation of (4) can be put in the more informative form

$$B(\Gamma_n,\Gamma_r;\lambda)=\sum_{p\in P}k_p\lambda^p\Big/\prod_{q\in Q}(1-\lambda^q), \tag{6}$$

where k_p, p, and q take positive integer values and P and Q are finite sets. The number of monomials in the denominator is equal to n, the dimension of Γ_n. An integrity basis for the Γ_r-tensors are a finite number of Γ_r-tensors which, joined to a finite number of invariants can span the whole set of Γ_r-tensors by products and linear combination. It is not hard to see that such an integrity basis can be formed by k_p Γ_r-tensors of degree p (for all $p\in P$) and n invariants whose degrees are the elements of Q. Among those invariants there will always be the $x^2=x_1^2+x_2^2+\cdots+x_n^2$ invariant for the subgroups of O(n). (x_i are the coordinates on R^n.)

With the knowledge of what is contained in the integrity basis, one can take the most general homogeneous vector fields and solve the Γ_r-variance condition (1) for a complete set of generators of G. Thus, one has the elements of the integrity basis explicitly.

What should be kept in mind of this subsection is that there exists a finite number of Γ_r-tensors and invariants which span all Γ_r-tensors and that these elements of the integrity basis can be found by the technique of the Molien function. (An example of the results of this technique is given for the cubic group in Sec. 3).

B. A generating function for the IHP

The Molien function gives the number of homogeneous polynomial invariants (or Γ_r-tensors) for a given degree. But this is not exactly what we are looking for. Indeed, we require the invariants to be harmonic, i.e., to verify the equations

$$\Delta I=0, \tag{7}$$

where Δ is the Laplace operator

$$\Delta=\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\cdots+\frac{\partial^2}{\partial x_n^2}. \tag{8}$$

A natural question arises then: Is there an analog of the Molien function for the IHP? The answer to this question is yes; but first, let us recall some basic facts about harmonic polynomials. (For an exhaustive treatment, see Vilenkin.[22])

Let $\mathfrak{R}^{nl}$ be the space of homogeneous polynomials of degree l on R^n. One can define an action of O(n) on the space of functions on R^n by

$$L(g)f(\mathbf{x})=f(g^{-1}\mathbf{x})\quad \forall g\in \mathrm{O}(n). \tag{9}$$

Under this action, the space $\mathfrak{R}^{n,l}$ is invariant: indeed, if $f(\mathbf{x})\in\mathfrak{R}^{n,l}$, $f(g^{-1}\mathbf{x})$ is also in $\mathfrak{R}^{n,l}$. The representation $L^{n,l}$, the restriction of L to $\mathfrak{R}^{n,l}$ is, however, reducible. To convince oneself of that, one can observe that $x^2\mathfrak{R}^{n,l-2}$ is an invariant subspace on $\mathfrak{R}^{n,l}$ [due to the fact that x^2 is invariant under O(n)]. Since O(n) is compact, the representation $L^{n,l}$ is completely reducible.

Another important invariant subspace of $\mathfrak{R}^{n,l}$ is precisely the subspace of harmonic polynomials that will be denoted $\mathfrak{H}^{n,l}$. Its invariance follows from the observations that the operator commutes with the action of $L^{n,l}$. [Note that $(f(g^{-1}\mathbf{x}))$ can be easily computed with the change of variables $\mathbf{x}'=g^{-1}\mathbf{x}$. Under this change of variables, Δ becomes

$$\Delta=\frac{\partial^2}{\partial x_1'^2}+\frac{\partial^2}{\partial x_2'^2}+\cdots+\frac{\partial^2}{\partial x_n'^2} \tag{10}$$

and $\Delta(f(g^{-1}\mathbf{x}))=0$.]

The relationship between $\mathfrak{R}^{n,l}$ and its two invariant subspaces $x^2\mathfrak{R}^{n,l-2}$ and $\mathfrak{H}^{n,l}$ is contained in the following lemma (demonstrated in Vilenkin):

Lemma: The space $\mathfrak{R}^{n,l}$ is the direct sum of the (supplementary) subspaces $r^2\mathfrak{R}^{n,l-2}$ and $\mathfrak{H}^{n,l}$:

$$\mathfrak{R}^{n,l}=r^2\mathfrak{R}^{n,l-2}\oplus\mathfrak{H}^{n,l}. \tag{11}$$

Moreover, $\mathfrak{H}^{n,l}$ is irreducible under the action of $L^{n,l}$.

As we have pointed out in the preceding subsection, all the generating functions $B(\Gamma_n,\Gamma_r;\lambda)$ contain in their denominator the monomial $(1-\lambda^2)$ corresponding to the invariant x^2. Among the invariants (or Γ_r-tensors) of degree l constructed with the elements of the integrity basis, those which contain a power of x^2 lie in the $r^2\mathfrak{R}^{n,l-2}$ subspace of $\mathfrak{R}^{n,l}$. The others are not necessarily in $\mathfrak{H}^{n,l}$ but have at least a component in this subspace.

Lemma: The generating function for the IHP is

$$B_{\mathrm{IHP}}(\Gamma_n,\Gamma_r;\lambda)=(1-\lambda^2)B(\Gamma_n,\Gamma_r;\lambda). \tag{12}$$

(The analogous result for the invariants has also been obtained by Meyer.[13])

Proof: Restricted to the finite subgroup G, the representation $L^{n,l}$ decomposes in irreducible representations:

$$L^{n,l}=a_l^1\Gamma_1\oplus a_l^2\Gamma_2\oplus\cdots\oplus a_l^m\Gamma_m, \tag{13}$$

where m is the number of irreducible representations of G. The a_l^i are the coefficients of the Molien functions:

$$B(\Gamma_n,\Gamma_i;\lambda)=\sum_{l=0}^{\infty}a_l^i\lambda^l. \tag{14}$$

Since $\mathfrak{H}^{n,l}$ is an irreducible subspace of $L^{n,l}$ the restriction of $L^{n,l}$ to this subspace can also be written as a direct sum of the irreducible representations Γ_i:

$$L^{n,l}|_{\mathfrak{H}^{n,l}}=b_l^1\Gamma_1\oplus b_l^2\Gamma_2\oplus\cdots\oplus b_l^m\Gamma_m. \tag{15}$$

The coefficients b_l^i are precisely the numbers we are looking for; indeed, the generating functions $B_{\mathrm{IHP}}(\Gamma_n,\Gamma_r;\lambda)$ for the

IHP will be written

$$B_{\rm IHP}(\Gamma_n,\Gamma_i;\lambda) = \sum_{l=0}^{\infty} b_l^i \lambda^l. \tag{16}$$

By the discussion preceding the lemma, we know that

$$b_l^i \leqslant a_l^i - a_{l-2}^i. \tag{17}$$

Introduce now the generating functions for the dimensions of $L^{n,l}$ and $\mathfrak{H}^{n,l}$:

$$l_n(\lambda) = \sum_{l=0}^{\infty} \lambda^l \dim L^{n,l} \tag{18a}$$

and

$$h_n(\lambda) = \sum_{l=0}^{\infty} \lambda^l \dim \mathfrak{H}^{n,l} = \sum_{l=0}^{\infty} \lambda^l \dim L^{n,l}|_{\mathfrak{H}^{n,l}}. \tag{18b}$$

Using the decomposition (11), $h_n(\lambda)$ can be rewritten as

$$h_n(\lambda) = \sum_{l=0}^{\infty} \lambda^l (\dim L^{n,l} - \dim L^{n,l-2}) \tag{19}$$

which can be expressed in terms of the a_l^i by (13):

$$h_n(\lambda) = \sum_{l=0}^{\infty} \sum_{i=1}^{m} \lambda^l (a^i - a_{l-2}^i) \dim \Gamma_i. \tag{20}$$

(It is to be understood that $a_l^i = 0$ for $l < 0$.) In the same way, (15) gives

$$h_n(\lambda) = \sum_{l=0}^{\infty} \sum_{i=1}^{m} \lambda^l b_l^i \dim \Gamma_i, \tag{21}$$

which finally gives a relation between the a_l^i and the b_l^i:

$$\sum_{i=1}^{m} (a_l^i - a_{l-2}^i) \dim \Gamma_i = \sum_{i=1}^{m} b_l^i \dim \Gamma_i. \tag{22}$$

Subject to the inequalities (17), the b_l^i have to be given by

$$b_l^i = a_l^i - a_{l-2}^i. \tag{23}$$

The generating functions $B_{\rm IHP}$ are then

$$\begin{aligned} B_{\rm IHP}(\Gamma_n,\Gamma_i;\lambda) &= \sum_{l=0}^{\infty} (a_l^i - a_{l-2}^i) \lambda^l \\ &= \sum_{l=0}^{\infty} a_l^i \lambda^l - \lambda^2 \sum_{l=0}^{\infty} a_{l-2}^i \lambda^{l-2} \\ &= (1-\lambda^2) B(\Gamma_n,\Gamma_i;\lambda). \end{aligned} \tag{24}$$

C. The harmonic projector H

Since $\mathfrak{H}^{n,l}$ and $r^2\mathfrak{R}^{n,l-2}$ are supplementary there exists a projector on $\mathfrak{H}^{n,l}$. This harmonic projector $H^{n,l}$ acting on a given vector of $\mathfrak{R}^{n,l}$ gives a vector of the subspace $\mathfrak{H}^{n,l}$. It takes the following differential operator form (see again Vilenkin):

$$H^{n,l} = \sum_{k=0}^{[l/2]} \frac{(-1)^k x^{2k} (n+2l-2k-4)!!}{2^k k! (n+2l-4)!!} \Delta^k. \tag{25}$$

Since the restriction of $L^{n,l}$ of O(n) to the subgroup G commutes with Δ [for example, for a given Γ_r-tensor, we have

$$\Delta P(\Gamma_n(g)x) = \Delta \Gamma_r(g^{-1})(Px) = \Gamma_r(g^{-1}) \Delta P(x)], \tag{26}$$

H preserves the Γ_r-variance of the objects on which it acts. Then if H acts on a basis (constructed with the elements of the integrity basis) for the Γ_r-tensors of degree l, it will project to 0 all the elements containing a factor x^{2k} and give linear independent harmonic Γ_r-tensors out of the others.

We have then proven the validity of our algorithm for obtaining a basis for the invariant (or Γ_r-tensor) harmonic polynomials of degree l:

(i) Obtain the Molien function and the corresponding integrity basis;

(ii) calculate a basis for the subspace of degree l and omit in this list the elements containing a power of x^2; and

(iii) apply H to these elements.

3. AN EXAMPLE: THE IHP FOR THE CUBIC SUBGROUP OF O(3)

This example should make clear the steps described in the preceding section. The first consists in computing the Molien function for the invariants. Putting $\Gamma_i = \Gamma_1$ (identity representation) in (4), one obtains easily (see, for example, Patera, Sharp, and Winternitz[10])

$$B(\Gamma_n,\Gamma_1;\lambda) = (1+\lambda^9)/(1-\lambda^2)(1-\lambda^4)(1-\lambda^6). \tag{27}$$

The corresponding integrity basis is calculated by introducing the most general polynomial of degree 2, 4, 6, and 9 successively in (2). We choose the following elements:

$$\begin{aligned} I_2 &= x^2 + y^2 + z^2, \\ I_4 &= x^4 + y^4 + z^4, \\ I_6 &= x^6 + y^6 + z^6, \\ E_9 &= xyz(x^2-y^2)(y^2-z^2)(z^2-x^2). \end{aligned} \tag{28}$$

The second step produces the basis of the subspace for a given l. Let us take $l = 12$ for this example. The generating function (27) tells us that there are seven invariants of degree 12:

$$\begin{aligned} B(\Gamma_n,\Gamma_1,\lambda) = 1 &+ \lambda^2 + 2\lambda^4 + 3\lambda^6 + 4\lambda^8 + \lambda^9 \\ &+ 5\lambda^{10} + \lambda^{11} + 7\lambda^{12} + \cdots \end{aligned} \tag{29}$$

A basis for this seven dimensional subspace is

$$I_2^6,\quad I_2^4 I_4,\quad I_2^2 I_4^2,\quad I_4^3,\quad I_2^3 I_6,\quad I_2 I_4 I_6,\quad \text{and}\quad I_6^2. \tag{30}$$

We know that the subspace $\mathfrak{H}_{n,l} = \mathfrak{H}_{3,12}$ of dimension 25 is spanned by the polynomials $r^{12} Y_{12}^m$. The restriction of the 25-dimensional irreducible representation of O(3) to the elements of the cubic subgroup G decomposes on the irreducible representations of G as follows (the superscript in parenthesis gives the dimensionality of the representation):

$$\Gamma_{SO(3)}^{(25)}|_{\rm CUBE} = 2\Gamma_1^{(1)} \oplus 1\Gamma_2^{(1)} \oplus 2\Gamma_3^{(2)} \oplus 3\Gamma_4^{(3)} \oplus 3\Gamma_5^{(3)}, \tag{31}$$

the representation Γ_1 being the identity representation. The subspace $\mathfrak{H}_{3,12}$ of harmonic polynomials contains then two invariants. A quick look at (30) tells us that there are precisely two invariants not containing I_2: they are I_4^3 and I_6^2. A last verification can be obtained by calculating $B_{\rm IHP}$:

$$\begin{aligned} B_{\rm IHP}(\Gamma_n,\Gamma_1;\lambda) &= (1-\lambda^2) B(\Gamma_n,\Gamma_1;\lambda) \\ &= 1 + \lambda^4 + \lambda^6 + \lambda^8 + \lambda^9 + \lambda^{10} \\ &\quad + 2\lambda^{12} + \cdots \end{aligned} \tag{32}$$

which also gives 2 as coefficient of λ^{12}.

The third and last step consists in projecting I_4^3 and I_6^2 on $\mathfrak{H}_{3,12}$ by the use of $H^{3,12}$. The operator $H^{3,12}$ is

$$H^{3,12} = \sum_{k=0}^{6} \frac{(-1)^k r^{2k} (23-2k)!!}{2^k k! (23)!!} \Delta^k \tag{33}$$

and the basis of invariant harmonic polynomials of degree 12 is found to be

$$H^{3,12}(I_4^3) = a\{101[x^{12}+y^{12}+z^{12}] - 3333[x^{10}(y^2+z^2)+y^{10}(x^2+z^2)+z^{10}(x^2+y^2)] + 23550[x^8(y^4+z^4)+y^8(x^4+z^4)+z^8(x^4+y^4)] + 8685x^2y^2z^2[x^6+y^6+z^6] - 42609[x^6z^6+y^6z^6+z^6x^6] - 20265[x^6(y^4z^2+y^2z^4)+y^6(x^4z^2+x^2z^4)+z^6(x^4y^2+x^2y^4)] + 101325x^4y^4z^4\},$$

$$H^{3,12}(I_6^2) = b\{24(x^{12}+y^{12}+z^{12}) - 792[x^{10}(y^2+z^2)+y^{10}(x^2+z^2)+z^{10}(x^2+y^2)] + 815[x^8(y^4+z^4)+y^8(x^4+z^4)+z^8(x^4+y^4)] + 30750x^2y^2z^2[x^6+y^6+z^6] + 3262[x^6y^6+y^6z^6+z^6x^6] + 71750[x^6(y^4z^2+y^2z^4)+y^6(x^4z^2+x^2z^4)+z^6(x^4y^2+x^2y^4)] + 358750x^4y^4z^4\}, \tag{34}$$

where a and b are constants.

One can see on this example that the algorithm can be easily programmed with a language manipulating algebraic expressions. This has been done on a DEC 2050 using a REDUCE compilator. Given the expressions H for $n=3$ (l free) and of the Laplacian acting on a general product of the elements of the integrity basis (28),

$$\begin{aligned}\Delta I_2^a I_4^b I_6^c E_9^d &= a(4a-4+16b+24c+36d+6)I_2^{a-1}I_4^bI_6^cE_9^d \\ &+ 16b(b-1)I_2^aI_4^{b-2}I_6^{c+1}E_9^d \\ &+ c(24b+30c+48d)I_2^aI_4^{b+1}I_6^{c-1}E_9^d \\ &+ b(64c+40d+12)I_2^{a+1}I_4^{b-1}I_6^cE_9^d \\ &+ c(-48b+30c-30-12d)I_2^{a+2}I_4^bI_6^{c-1}E_9^d \\ &- 30c(c-1)I_2^{a+3}I_4^{b+1}I_6^{c-2}E_9^d \\ &+ 8bcI_2^{a+4}I_4^{b-1}I_6^{c-1}E_9^d + 6c(c-1)I_2^{a+5}I_4^bI_6^{c-2}E_9^d,\end{aligned} \tag{35}$$

the program generates for each l the invariants subject to the conditions of the step (ii) of the algorithm, applies H on each of them (expressed as a product of elements of the integrity basis), transforms the solutions in Cartesian coordinates and lists the results. We have run the program up to $l=25$ but we did not push further since we did not want to publish any tables. Another version of the program for the pseudoinvariants (transforming under the alternate representation) has also been devised and run for $l \leqslant 25$. The common denominator of the polynomials is factored out; the coefficients are then integer and there is no loss of accuracy.

4. CONCLUSION

To close this paper, we would like to stress some of the advantages of the algorithm.

The algorithm is easy enough to program that anyone who is familiar with the usual programming languages can obtain harmonic polynomials of "reasonable" degrees without any sophisticated numerical methods. The program can be set up in such a way that the coefficients of the IHP are integer and therefore exact. The algorithm clearly works in any dimension n and for any finite subgroups $G \subset O(n)$. Moreover, it is obvious from Sec. 2 how to compute Γ_r-tensor harmonic polynomials for any irreducible representation Γ_r of G. Finally, one important advantage lies in the fact that the invariance (or Γ_r-variance) conditions (2) [or (1)] are solved once and for all (at the step of calculating the integrity basis). This avoids the solution of any high-dimensional system of linear equations.

ACKNOWLEDGMENTS

We would like to thank Jean-Louis Meunier of the Département de mathématiques des Facultés Universitaires Notre-Dame de la Paix de Namur for his help at the programming stage, and J. Patera of the Centre de recherche de mathématiques appliquées, Université de Montréal, for discussions and for kindly reading the manuscript. This work was begun as one of us (Y.S.-A.) was sent to Namur by the Ministère de l'Education du Québec as part of the Echanges belgo-québécois.

[1] T. Molien, Sitzungber. K. Preuss. Akad. Wiss., 1152 (1897).
[2] F. C. von der Lage and H. Bethe, Phys. Rev. **71**, 612 (1947).
[3] D. G. Bell, Rev. Mod. Phys. **26**, 311 (1954).
[4] S. L. Altmann, Proc. Cambridge Philos. Soc. **53**, 343 (1957); S. L. Altman and A. P. Cracknell, Rev. Mod. Phys. **37**, 19 (1965).
[5] W. Döring, Ann. Phys. **1**, 102 (1958); W. Döring and G. Simon, Ann. Phys. **5**, 373 (1960).
[6] J. Muggli, Phys. Kondens. Mater. **12**, 237 (1971).
[7] J. Killinbeck, J. Phys. C **5**, 2497 (1972).
[8] R. P. Bickerstaff and B. G. Wybourne, J. Phys. A **9**, 1051 (1976).
[9] N. J. A. Sloane, Amer. Math. Month. **84**, 82 (1977).
[10] J. Patera and P. Winternitz, J. Chem. Phys. **65**, 2725 (1975); J. Patera, R. T. Sharp, and P. Winternitz, J. Math. Phys. **19**, 2362 (1978). [Corrections to formulas (61)–(64) have appeared in: Y. Saint-Aubin, J. Can. Phys. **58**, 1075 (1980).]
[11] J. Antoine and L. Bodart, Mémoire de Licence, Facultés Universitaires Notre-Dame de la Paix, unpublished (1979).
[12] M. A. Lohe, Ph.D. Thesis, University of Adelaide (1974); M. A. Lohe and C. A. Hurst, J. Math. Phys. **12**, 1882 (1971).
[13] B. Meyer, Can. J. Math. **6**, 135 (1954).
[14] H. Puff, Physica Status Solidi **41**, 11 (1970).
[15] C. F. Dunkl, "Cube group invariant spherical harmonics and Krawtchouk polynomials," preprint of the Department of Mathematics of the University of Virginia (1980).
[16] K. T. Hecht, J. Mol. Spectrosc. **5**, 355 (1960).
[17] J. Moret-Bailly, L. Gautier, and J. Montagutelli, J. Mol. Spectrosc. **15**, 355 (1965).
[18] K. Fox and I. Ozier, J. Chem. Phys. **52**, 5044 (1970).
[19] K. Fox and B. J. Krohn, J. Comput. Phys. **25**, 386 (1977).
[20] B. J. Krohn and K. Fox, Program #347, *Quantum Chemistry Program Exchange* (Department of Chemistry, Indiana University).
[21] W. Burnside, *Theory of Groups of Finite Order* (Dover, New York, 1955).
[22] N. Ja. Vilenkin, *Special functions and the theory of group representations* (American Mathematical Society, Providence, RI, 1968).

A Galerkin method and nonlinear oscillations and waves

Fengsu Chen [a] and D. Y. Hsieh
Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

(Received 25 August 1981; accepted for publication 26 February 1982)

A Galerkin method is developed as a generalization of the variational averaging method to deal with problems with dissipation. Some nonlinear oscillations, nonlinear waves, and nonlinear stability problems are studied to illustrate the application of the new method. It is demonstrated that when the dissipative parameter is small, the solutions agree with those obtained by other established methods.

PACS numbers: 02.30. + g

I. INTRODUCTION

For a wide class of problems on nonlinear oscillations, nonlinear waves, and nonlinear stabilities, variational method has been shown to be a useful tool to obtain approximate asymptotic solutions.[1–4] The method starts with reformulating the problem by an equivalent variational problem; then some appropriately chosen asymptotic trial solutions with adjustable parameters are substituted into the functional to be varied. It is expected that the system of equations governing the adjustable unknown parameters would be simpler than the original problem. An essential step in the variational method is to find the functional to be varied. For many important physical problems, e.g., almost all the problems with dissipation, the functional or the Lagrangian cannot be found, and the variational method is not applicable. This is a serious defect of the method. Some attempts have been made to modify the variational method to accommodate the problems with dissipation.[1,5] In this study we shall develop the method in a more systematic manner.

The method to be developed is a generalization of the Galerkin method. The essence of the Galerkin method may be described as follows.[6] Take the differential equation

$$L[x(t)] = 0. \tag{1}$$

A trial solution is taken in the form

$$x = \sum_{i=1}^{N} c_i x_i(t), \tag{2}$$

where $\{x_i(t)\}$ is a set of given functions. Then choose a set of weighting functions $\{w_i(t)\}$. The parameters $\{c_i\}$ are to be determined by the following set of algebraic equations:

$$\int \left(L\left[\sum_{i=1}^{N} c_i x_i\right]\right)\cdot w_j dt = 0, \quad j = 1,\ldots,N. \tag{3}$$

The weighting functions $\{w_i(t)\}$ were originally chosen by Galerkin[7] to be identical to $\{x_i(t)\}$. Then it may be shown when $\{x_i(t)\}$ form a complete orthogonal set, the solution (2) represents the exact solution, if N is taken to be large enough. However, as an approximate method in practice, N is usually not large and $\{x_i(t)\}$ often does not form a complete set. The weighting functions $\{w_i(t)\}$ are also often chosen to be different from $\{x_i(t)\}$ as dictated mostly by experience or convenience. The flexibility of the choices of $\{w_i(t)\}$ has its merits. But the lack of definiteness is also disturbing when the method is to be applied in an area where one has little experience.

The classical Galerkin method is closely related to the direct variational method. Take Eq. (1); it is often possible to establish an equivalent variational formulation:

$$\Delta J = 0, \tag{4}$$

where the functional $J\{x\}$ is of the form

$$J = \int F(x^{(n)}(t),\ldots,x(t),t)dt. \tag{5}$$

After some manipulation, we obtain from (4)

$$\int L[x(t)]\Delta x dt = 0. \tag{6}$$

Thus Eq. (1) is the Euler–Lagrange equation of the variational problem. When the direct method is employed with the trial solution (2), then

$$\Delta x = \sum_{i=1}^{N} x_i(t)\Delta c_i. \tag{7}$$

Because of the independent variations of Δc_i, we obtain from (6)

$$\int \left(L\left[\sum_{i=1}^{N} c_i x_i\right]\right) x_j dt = 0, \quad j = 1,\ldots,N. \tag{8}$$

Equation (8) is the same as (3) when $\{w_i(t)\}$ are identified with $\{x_i(t)\}$. Thus, when the problem can be formulated in terms of a variational principle, a definite Galerkin method as exemplified by Eq. (8) can always be found, which is equivalent to the variational method. However, there are many problems for which it is impossible or very difficult to find the functional J and to formulate the problem in terms of a variational principle. For these problems, the variational method is not applicable without modification. On the other hand, for the Galerkin method, it is apparent that instead of (6), we may also formulate the problem in terms of the more general relation

$$\int L[x(t)]\Delta f(x) dt = 0, \tag{9}$$

where $f(x)$ is any arbitrary function of x. The proper choice of f would certainly affect the outcome of the analysis. In the

[a] On leave from Wuhan Institute of Hydraulic and Electric Engineering, Wuhan, People's Republic of China.

 0022-2488/83/051041-06$02.50

following, we shall examine these questions and present a scheme of the generalized Galerkin method. Then this Galerkin method will be applied to some nonlinear oscillation and nonlinear stability problems for illustration.

II. GENERAL SCHEME OF THE GALERKIN METHOD

Consider a differential equation schematically represented by

$$L[x(t);\alpha] = 0. \tag{10}$$

where α is a parameter. As examples, take the following differential equations representing oscillations with linear and nonlinear dampings:

$$x'' + 2\alpha x' + x = 0 \tag{11}$$

and

$$x'' + \alpha(x')^3 + x = 0. \tag{12}$$

We shall be interested in those equations, as in the above examples for which an equivalent variational formulation exists if the parameter α is zero. More explicitly, when $\alpha = 0$ a functional J can be found and $\Delta J = 0$ will lead to the following relation:

$$\int L[x(t);0]\Delta x\, dt = 0. \tag{13}$$

Then for α sufficiently small, we expect a similar relation

$$\int L[x(t);\alpha]\Delta x\, dt = 0 \tag{14}$$

is also valid. This is the essence of the general scheme of the proposed Galerkin method. It is an exact formulation when $\alpha = 0$, and it is free from the ambiguity as manifested by (9). It is capable of dealing with dissipative systems, such as (11) and (12), if the dissipation is sufficiently small. Indeed, the method may also be called the variational Galerkin method.

As in the general application of the direct variational method, the trial solution need not be of the form (2). We may take instead

$$x(t) = \phi(t;c_1,\ldots,c_N), \tag{15}$$

where c_i are those adjustable parameters to be varied independently. Then the Galerkin formulation becomes

$$\int L[\phi;\alpha]\frac{\partial\phi}{\partial c_i}\, dt = 0, \quad i = 1,\ldots,N. \tag{16}$$

To treat problems with asymptotic periodic solutions, such as a certain class of nonlinear oscillations, waves, and stability problems, we shall again follow the same averaging scheme as developed in previous studies.[1–3] In essence, we shall make use of whatever prior information there is as much as possible and incorporating it into the form of the trial solution. Then, approximate equations for the adjustable parameters, e.g., amplitudes, phases, etc., are obtained and solved by singling out the secular terms.

Although we have used an ordinary differential equation of a single variable (1) to present the general scheme of the variational Galerkin method, the same procedure can be readily applied to partial differential equations and equations with many variables. The generalization to the problems with more than one dissipative parameter can also be made in the same manner.

III. LINEAR AND NONLINEAR OSCILLATIONS WITH DAMPING

Consider the differential equation of oscillation with linear damping (11). When $\alpha = 0$, the variational formulation will lead to the following relation:

$$\Delta J = 0,$$

where

$$J = \int_0^t \frac{1}{2}[(x')^2 - x^2]dt.$$

Or, equivalently,

$$\int_0^t (x'' + x)\Delta x\, dt = 0. \tag{17}$$

Thus, according to the general scheme presented in Sec. II, the corresponding Galerkin formulation of Eq. (11) is

$$\int_0^t (x'' + 2\alpha x' + x)\Delta x\, dt = 0. \tag{18}$$

Take the trial solution of the form

$$x = A(t)\sin B(t), \tag{19}$$

where $A(t)$ and $B'(t)$ are taken to be slowly varying functions of time. Thus

$$x' = A'\sin B + AB'\cos B, \tag{20}$$

$$x'' = A''\sin B + 2A'B'\cos B + AB''\cos B - AB'^2\sin B, \tag{21}$$

and

$$\Delta x = \sin B\Delta A + A\cos B\Delta B. \tag{22}$$

Substituting these expressions into (18), we obtain

$$\begin{aligned}\int_0^t \{ [(A'' - AB'^2 + 2\alpha A' + A)\sin^2 B \\ + (2A'B' + AB'' + 2\alpha AB')\sin B\cos B]\Delta A \\ + [(2A'B' + AB'' + 2\alpha AB')\cos^2 B \\ + (A'' - AB'^2 + 2\alpha A' + A)\sin B\cos B]A\Delta B\}dt \\ = 0.\end{aligned} \tag{23}$$

Using the averaging scheme as developed in the variational method[1] by retaining only the secular terms, and noting that ΔA and ΔB are independent, we obtain

$$A'' - AB'^2 + 2\alpha A' + A = 0 \tag{24}$$

and

$$2A'B' + AB'' + 2\alpha AB' = 0. \tag{25}$$

Since A and B' are slowly varying functions of time, we shall neglect the terms with A'' and B'' in (24) and (25). Furthermore, α is also a small parameter. Thus we may approximate (24) and (25) by

$$AB'^2 - A = 0 \tag{26}$$

and

$$(A' + \alpha A)B' = 0. \tag{27}$$

Hence we obtain

$$B = t + \beta \tag{28}$$

and

$$A = ae^{-\alpha t}, \tag{29}$$

where a and β are integration constants.

Thus the solution (19) becomes

$$x = ae^{-\alpha t}\sin(t + \beta). \tag{30}$$

The exact solution of the differential equation (11) is readily found to be

$$\begin{aligned} x &= ae^{-\alpha t}\sin[(1-\alpha^2)^{1/2}t + \beta] \\ &= ae^{-\alpha t}\sin(t+\beta) + O(\alpha^2). \end{aligned} \tag{31}$$

Therefore the approximate solution (30) obtained by the Galerkin method agrees with the exact solution at least up to $O(\alpha)$. In fact, for this linear problem, the system (24) and (25) also has the exact solutions

$$A(t) = ae^{-\alpha t}$$

and

$$B(t) = (1-\alpha^2)^{1/2}t + \beta.$$

Let us now apply the Galerkin method to the problem of oscillation with cubic damping (12). Then the corresponding Galerkin formulation is

$$\int_0^t [x'' + \alpha(x')^3 + x]\Delta x \, dt = 0. \tag{32}$$

Take again the trial solution (19), and apply the same averaging scheme as before, we obtain

$$A'' - AB'^2 + A + \frac{3\alpha}{4}(A'^3 + A^2A'B'^2) = 0 \tag{33}$$

and

$$2A'B' + AB'' + \frac{3\alpha}{4}(A^3B'^3 + AA'^2B') = 0. \tag{34}$$

The approximate equations are then

$$(B'^2 - 1)A = 0 \tag{35}$$

and

$$2A'B' + \frac{3\alpha}{4}A^3B'^3 = 0. \tag{36}$$

Thus we obtain from (35)

$$B = t + \beta \tag{37}$$

and

$$A = \frac{1}{[a + (3\alpha/4)t]^{1/2}}, \tag{38}$$

where β and a are integration constants. Thus the approximate solution of (12) is

$$x = \frac{\sin(t+\beta)}{[a + (3\alpha/4)t]^{1/2}}. \tag{39}$$

The approximate solution (39) is the same as that obtained by the method of multiple scale.[8]

IV. DUFFING STABILITY WITH DAMPING

Consider the Duffing equation with damping,

$$u'' + \alpha u' - au + \gamma u^3 = 0, \tag{40}$$

where α, a, and γ are real positive constants. When $\alpha = 0$, the equation represents a system which is linearly unstable and nonlinearly stable, and has been studied by the variational method.[3] Multiply (40) with u' and then integrate; we obtain

$$\left(\frac{du}{dt}\right)^2 = F(u) - \alpha \int_{t_0}^t \left(\frac{du}{dt}\right)^2 dt, \tag{41}$$

where

$$F(u) = C + au^2 - (\gamma/2)u^4 \tag{42}$$

and C is an integration constant to be determined by initial conditions. For instance: $u(t_0) = 0$, $u'(t_0) = \sqrt{C}$. When $\alpha = 0$, Eq. (41) may be integrated to yield solutions in terms of elliptic functions. When α is small, it is expected that $u(t)$ will still exhibit a largely periodic behavior in some definite finite interval of time. From Eq. (41), it is clear that solutions are permitted only if the right-hand side of (41) is positive. It is convenient for our discussion to adjust the constant C in (42) as $C(t_0)$ such that $(t - t_0)$ never exceeds the period of the system at the time. Then C is monotonously decreasing in t_0. The function $F(u)$ for various C's are shown in Fig. 1. When $C > 0$ [Fig. 1(a)], the system is largely oscillating between $\pm u_1$. As t increases, C will decrease and eventually C becomes negative [Fig. 1(c)]. Then the system will oscillate between u_2 and u_1 (or $-u_1$ and $-u_2$). Eventually u will settle down asymptotically to $u_0 = (a/\gamma)^{1/2}$ [or $-(a/\gamma)^{1/2}$], as shown in Fig. 1(d).

Now we apply the Galerkin method to Eq. (40) and obtain

$$\int_0^t (u'' + \alpha u' - au + \gamma u^3)\Delta u \, dt = 0. \tag{43}$$

Let us take the trial function

$$u(t) = A(t)\sin v(t) + B(t), \tag{44}$$

where A, B, and v' are slowly varying functions of time. Thus, for instance,

$$u'(t) = A'\sin v + Av'\cos v + B'$$

and

$$\Delta u = \sin v\Delta A + A\cos v\Delta v + \Delta B.$$

After using the same averaging scheme as before, we obtain

$$A'' + \alpha A' - A[(v')^2 + a - (\tfrac{3}{4}\gamma A^3 + 3\gamma B^2)] = 0, \tag{45}$$

$$2A'v' + Av'' + \alpha Av' = 0, \tag{46}$$

and

$$B'' + \alpha B' - aB + \gamma B^3 + \tfrac{3}{2}\gamma A^2 B = 0. \tag{47}$$

The approximate equations are then

$$(v')^2 = \tfrac{3}{4}\gamma A^2 + 3\gamma B^2 - a, \tag{48}$$

$$A' + (\alpha/2)A = 0, \tag{49}$$

and

$$B[B^2 - (a/\gamma - \tfrac{3}{2}A^2)] = 0. \tag{50}$$

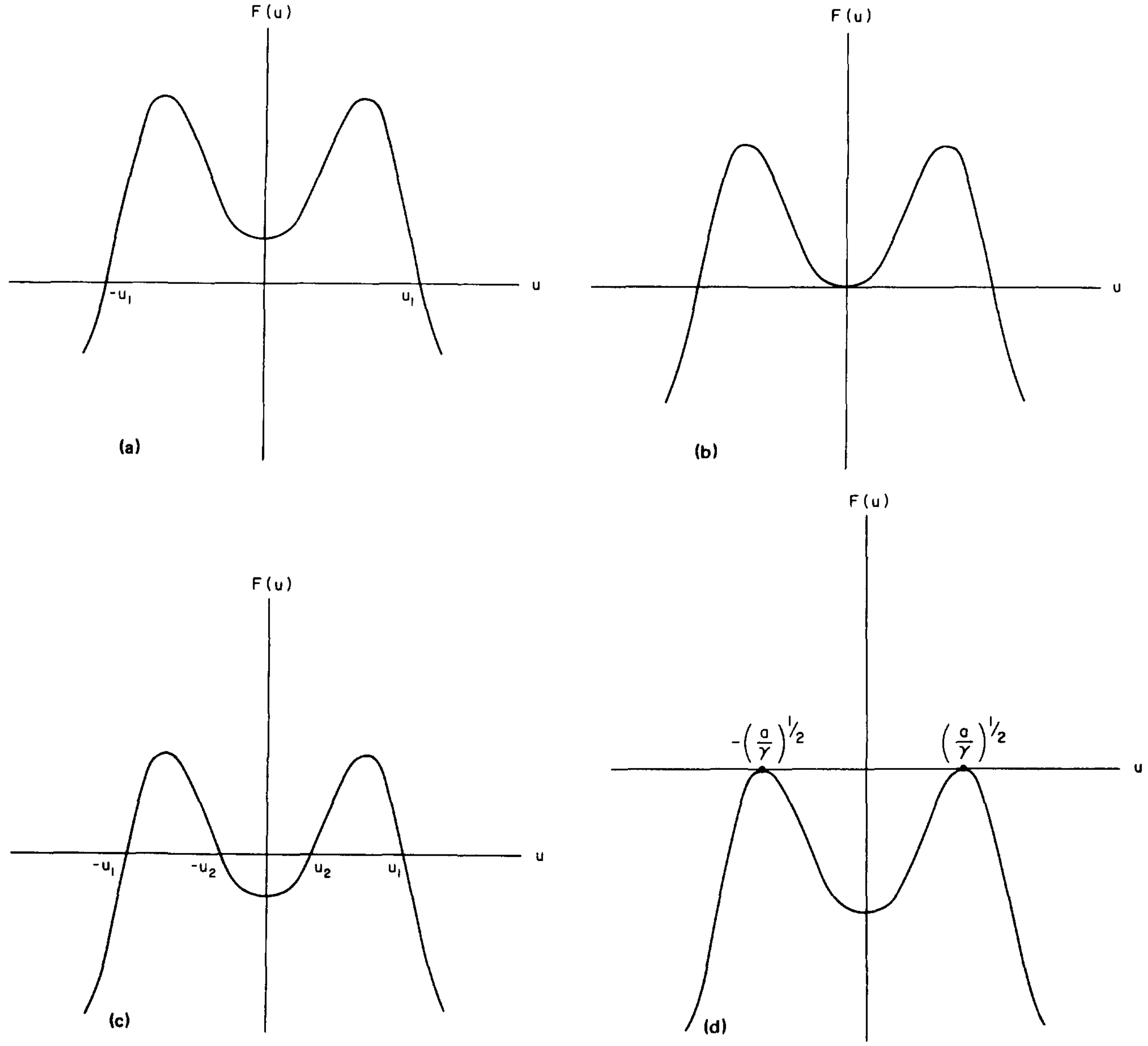

FIG. 1. The schematic representation of $F(u)$. (a): $C > 0$; (b) $C = 0$; (c) $C < 0$; (d) $C = -a^2/2\gamma$.

From (49) we obtain

$$A(t) = A(t_0)e^{-(\alpha/2)(t - t_0)}. \tag{51}$$

Thus the amplitude of the oscillation will diminish slowly but exponentially with time. Since A and B are slowly varying functions of time, we may treat them as constants in any definite small interval of time. Hence (48) can be integrated approximately to yield

$$\nu = \omega t + \phi, \tag{52}$$

where

$$\omega^2 = \tfrac{3}{4}\gamma A^2 + 3\gamma B^2 - a \tag{53}$$

and ω is the frequency of the oscillation.

From (51) we obtain either

$$\text{(i)}\ B = 0 \tag{54}$$

or

$$\text{(ii)}\ B = \pm(a/\gamma - \tfrac{3}{4}A^2)^{1/2}. \tag{55}$$

The case $B = 0$ corresponds to the situation depicted in Fig. 1(a). Since A is decreasing in t, while ω^2 cannot be negative, the solution ceases to be valid when $\omega \rightarrow 0$ or when $A^2 \rightarrow 4a/3\gamma$. Then the mode (i) will switch over to the mode (ii) given by (55), corresponding to the situation depicted in Fig. 1 (c). Since A is again diminishing, eventually for large t, we obtain asymptotically

$$A = 0, \quad B = \pm(a/\gamma)^{1/2},$$

which is the state depicted in Fig. 1(d).

There is no exact analytical solution for Eqs. (40) or (41). The approximate solution attained above by the Galerkin method catches the essential features of the solution except at the transition region, as depicted in Fig. 1(b). We shall briefly sketch how to deal with this transition problem. With C considered as a slowly decreasing function of t_0, we may approximate (41), since α is a small parameter, by

$$\left(\frac{du}{dt}\right)^2 = F(u) \tag{56}$$

and

$$dt = \frac{du}{\sqrt{F(u)}}. \tag{57}$$

Substituting (56) and (57) into the integral in (41), we obtain

$$\left(\frac{du}{dt}\right)^2 = F(u) - \alpha \int^u \sqrt{F(u)}\, du. \tag{58}$$

In the transition region we may take $C = 0$, and thus

$$F(u) = au^2 - \frac{\gamma}{2} u^4. \tag{59}$$

Substitute (59) into (58), then $u(t)$ can be explicitly integrated in terms of quadrature in the transition region.

V. THE KLEIN–GORDON EQUATION WITH DAMPING

Let us now turn to partial differential equations and consider the Klein–Gordon equation with damping:

$$u_{tt} - u_{xx} + \alpha u_t + f(u) = 0. \tag{60}$$

When $\alpha = 0$, this equation has been investigated by the variational method by Whitham[4] and Hsieh.[2] The latter approach can be readily adapted to the Galerkin method, and we shall follow that approach to deal with the case when $\alpha \neq 0$.

The Galerkin formulation of (60) is

$$\int_0^t dt \int_{-\infty}^{\infty} dx [u_{tt} - u_{xx} + \alpha u_t + f(u)] \Delta u = 0. \tag{61}$$

Let us take the trial solution of the following form:

$$u = A(x,t)\phi(S(x,t)), \tag{62}$$

where A, S_t, and S_x are all slowly varying functions of (x,t) and ϕ is a periodic function of S which satisfies the following conditions:

$$\phi(S + 2\pi) = \phi(S), \tag{63}$$

$$\langle \phi \rangle \equiv \frac{1}{2\pi} \int_0^{2\pi} \phi\, dS = 0, \tag{64}$$

$$\langle \phi^2 \rangle = 1. \tag{65}$$

It follows from (63) that

$$\langle \phi \phi' \rangle = \langle \phi' \phi'' \rangle = 0, \tag{66}$$

and we shall denote

$$\langle \phi'^2 \rangle = -\langle \phi \phi'' \rangle = \beta. \tag{67}$$

From (62) we have

$$u_t = A_t \phi + A S_t \phi', \tag{68}$$

$$u_{tt} = A_{tt}\phi + 2A_t S_t \phi' + A S_{tt} \phi' + A S_t^2 \phi'', \tag{69}$$

and a similar set for u_x and u_{xx}. Also we have

$$\Delta u = \phi \Delta A + A \phi' \Delta S. \tag{70}$$

Substituting these expressions in (61) and carrying out the averaging scheme as before, we obtain

$$A_{tt} - A_{xx} + A_t - \beta A (S_t^2 - S_x^2) + g(A) = 0 \tag{71}$$

and

$$\alpha A^2 S_t + (A^2 S_t)_t - (A^2 S_x)_x = 0, \tag{72}$$

where

$$g(A) = \langle \phi f(A\phi) \rangle. \tag{73}$$

Since A is supposed to be a slowly varying function of (x,t) and α is a small parameter, we obtain the approximate equations

$$\beta A (S_t^2 - S_x^2) - g(A) = 0 \tag{74}$$

and

$$\alpha A^2 S_t + S_t (A^2)_t - S_x (A^2)_x = 0. \tag{75}$$

Let us denote

$$S_t = -\omega, \quad S_x = k. \tag{76}$$

When ω and k are slowly varying functions of (x,t), and can be treated as if they are constants, then Eq. (74) represents the nonlinear dispersion relation

$$\omega^2 = k^2 + \frac{g(A)}{\beta A}. \tag{77}$$

Equation (75) becomes

$$A_t + \frac{k}{\omega} A_x + \frac{\alpha}{2} A = 0, \tag{78}$$

which has the general solution

$$A(x,t) = e^{-(\alpha/2)t} a[x - (k/\omega)t], \tag{79}$$

where a is any arbitrary function. Since, from (77) the group velocity C_g is given by

$$C_g = \frac{d\omega}{dk} = \frac{k}{\omega}, \tag{80}$$

(79) can also be written as

$$A(x,t) = e^{-(\alpha/2)t} a(x - C_g t). \tag{81}$$

Thus the amplitude function is propagating with the group velocity and also decays exponentially with damping coefficient $\alpha/2$.

If $f(u)$ is a polynomial of u, i.e.,

$$f(u) = \sum_{n=1}^{N} C_n u^n, \tag{82}$$

then

$$g(A) = \sum_{n=1}^{N} C_n \langle \phi^n \rangle A^n. \tag{83}$$

In order to determine β and $\langle \phi^n \rangle$, or generally $g(A)$, we need to know the periodic function $\phi(s)$. Using the original equation (60) as a guide, we may obtain $\phi(s)$ in the following manner. Let $u = \phi(s)$ and $s(x,t) = kx - \omega t$. Then treat k and w as constants, and neglect the term with α since α is a small parameter. Then Eq. (60) becomes

$$v^2 \phi'' + f(\phi) = 0, \tag{84}$$

where

$$v^2 = \omega^2 - k^2.$$

Equation (84) may be integrated to yield

$$\frac{s}{v} = \int^{\phi} \frac{dy}{\{\int_y^b f(z) dz\}^{1/2}}. \tag{85}$$

The constants b and v can be determined from the conditions (63) and (65). With $\phi(s)$ given, then β and $g(A)$ can be explicitly calculated, and the systems (74) and (75) can be solved. We shall now consider in more detail the linear case, i.e., when $f(u) = u$.

For the linear case, we obtain from (85), (63), and (64) that

$$\phi(s) = \sqrt{2}\sin(s + \psi), \tag{86}$$

where ψ is a constant. From (73) and (67) we have

$$\beta = 1, \quad g(A) = A.$$

Thus (74) becomes

$$S_t^2 - S_x^2 = 1. \tag{87}$$

A complete integral of the above equation is

$$S = kx - (1 + k^2)^{1/2}t + m, \tag{88}$$

where k and m are two arbitrary constants. Substituting (88) into (75), we obtain

$$A_t + \frac{k}{(1 + k^2)^{1/2}}A_x + \frac{\alpha}{2}A = 0, \tag{89}$$

which yields the general solution

$$A(x,t) = e^{-(\alpha/2)t}a\left(x - \frac{k}{(1 + k^2)^{1/2}}t\right). \tag{90}$$

Thus

$$u(x,t) = e^{-(\alpha/2)t}a\left(x - \frac{k}{(1 + k^2)^{1/2}}t\right) \times \sin[kx - (1 + k^2)^{1/2}t + \psi], \tag{91}$$

which represents the travelling solution for the linear Klein–Gordon equation. It may be pointed out that when the linear Klein–Gordon equation is solved by the method of Fourier transform, the general solution can be written as

$$u(x,t) = \int dk\, a(k)e^{(-\alpha/2)t + i[kx - (1 + k^2 - \alpha^2/4)^{1/2}t]}. \tag{92}$$

For travelling waves with wavenumbers in the neighborhood of some definite k, the solutions (91) and (92) agree up to the order $O(\alpha)$.

Equation (87) possesses also a singular solution which is the envelope of the family of solutions represented by (88). This singular solution is

$$S = (t^2 - x^2)^{1/2}. \tag{93}$$

Substituting (93) into (72), we obtain

$$t(A^2)_t + x(A^2)_x + (1 + \alpha t)A^2 = 0. \tag{94}$$

The general solution of (94) is of the form

$$A = e^{-\alpha t}t^{-1/2}b(x/t), \tag{95}$$

where b is an arbitrary function. The solution (95) is again consistent with the asymptotic expression for large t from the solution (92) up to order $O(\alpha)$.

VI. DISCUSSION

The examples of nonlinear oscillations and waves discussed in previous sections demonstrate that the variational Galerkin method can successfully treat problems with dissipation. When comparisons can be made, it is found that the solutions obtained by the Galerkin method agree with those obtained by other established methods so long as the dissipative parameter is small. It may be remembered, in contrast to many other asymptotic methods, that the variational method is not intrinsically a perturbation method and therefore is well adapted to treat a certain class of nonlinear stability problems.[3] However, one major defect of the variational method is its inability to deal with dissipative systems. The proposed Galerkin method, as demonstrated by the examples treated in previous sections, apparently can remedy this defect.

In the scheme of the Galerkin method, we have considered only the cases when the dissipative parameter α is small. An obvious question is how far the small α solutions can be extrapolated to the cases when α is not small. This is a very important question that needs to be investigated. A related problem is a more rigorous establishment of the variational Galerkin method even when α is sufficiently small.

For problems with many adjustable parameters, it is conceivable that there may be more than one possible variational formulation from which to generalize Galerkin schemes. This problem of nonuniqueness is closely related to the question of extrapolation just mentioned, and is different from the nonuniqueness arising from the choices of different $f(x)$ in Eq. (9). For practical applications, however, it may yet be worthwhile to investigate how sensitively the solutions would depend on various choices of $f(x)$ in (9).

It is clear that much work is still needed to answer the many questions raised by the proposed variational Galerkin method and to explore its wide-ranging potentials. In subsequent papers, we shall apply the method to various problems in mechanics and physics, and among others some stability problems in fluid dynamics.

[1] D. Y. Hsieh, J. Math. Phys. **16**, 275–80 (1975).
[2] D. Y. Hsieh, J. Math. Phys. **16**, 1630–36 (1975).
[3] D. Y. Hsieh, J. Math. Phys. **18**, 1093–1104 (1977).
[4] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).
[5] D. Y. Hsieh, J. Acoust. Soc. Am. **58**, 977–82 (1975).
[6] B. A. Finlayson, *The Method of Weighted Residuals and Variational Principles* (Academic, New York, 1972).
[7] B. G. Galerkin, Vestnik Inzhenerov, **1**, 897–908 (1915).
[8] J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics* (Springer, New York, 1981).

The general theory of R-separation for Helmholtz equations

E. G. Kalnins
Mathematics Department, University of Waikato, Hamilton, New Zealand
W. Miller, Jr.[a)]
School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

(Received 18 November 1981; accepted for publication 26 March 1982)

We develop the theory of R-separation for the Helmholtz equation on a pseudo-Riemannian manifold (including the possibility of null coordinates) and show that it, and not ordinary variable separation, is the natural analogy of additive separation for the Hamilton–Jacobi equation. We provide a coordinate-free characterization of variable separation in terms of commuting symmetry operators.

PACS numbers: 02.30. + g, 02.40.Ky

1. INTRODUCTION

Let V_n be a (local) pseudo-Riemannian manifold. The *Helmholtz equation* for V_n is expressed in local coordinates $\{y^j\}$ by

$$\Delta\psi(\mathbf{y}) = E\psi(\mathbf{y}), \tag{1.1}$$

where E is a nonzero constant and Δ is the Hamiltonian or Laplace–Beltrami operator[1]

$$\Delta = \frac{1}{g^{1/2}} \sum_{i,j=1}^{n} \partial_i (g^{1/2} g^{ij} \partial_j). \tag{1.2}$$

Here, $\partial_j = \partial_{y^j}$, the metric on V_n is $ds^2 = \Sigma_{i,j} g_{ij} dy^i dy^j$, $g = \det(g_{ij}) \neq 0$, and $\Sigma_k g^{ik} g_{kj} = \delta^i_j$. The Helmholtz equation is closely associated with the *Hamilton–Jacobi equation*[2]

$$H(\partial_i W) \equiv \sum_{i,j=1}^{n} g^{ij} \partial_i W \partial_j W = E, \tag{1.3}$$

where H is the Hamiltonian function

$$H(p_i) = \sum_{i,j=1}^{n} g^{ij} p_i p_j. \tag{1.4}$$

Both Δ and H are defined independent of local coordinates.

In Ref. 3 the authors presented a theory of orthogonal R-separation for (1.1). [By R-separation we mean separation up to a fixed factor:

$$\psi(\mathbf{y}) = R(\mathbf{y}) \prod_{j=1}^{n} \psi^{(j)}(y^j). \tag{1.5}$$

Ordinary separation corresponds to $R \equiv 1$ and trivial R-separation to $\partial_{ij} \ln R = 0$ for $i \neq j$.] We found necessary and sufficient conditions that an additively separable orthogonal coordinate system for the Hamilton–Jacobi equation will also R-separate the Helmholtz equation. [An R-separable system for (1.1) always separates (1.3).] Further, we found a coordinate-free characterization of orthogonal R-separable coordinate systems in terms of families of commuting symmetry operators for Δ.

In this paper we extend the ideas of Ref. 3 to provide a general theory of R-separation for the Helmholtz equation, encompassing both orthogonal and nonorthogonal coordinate systems. A major new complication is the possibility of type 2 (null) coordinates. Our principal result is Theorem 3, which provides an intrinsic characterization of an R-separable coordinate system in terms of a family of commuting symmetry operators. (In particular, given the operators, expressed in an arbitrary coordinate system, one can compute the R-separable coordinates.)

Although R-separation has long been a useful tool in the study of the Laplace equation [$E = 0$ in (1.1)], its relevance to the Helmholtz equation was, until recently, virtually ignored. Our results show clearly that R-separation, rather than ordinary separation, for the Helmholtz equation is the proper analog to additive separation of the Hamilton–Jacobi equation. In fact, the problem of extending a separable system for (1.3) to an R-separable system for (1.1) reduces to an exercise in quantization theory.

In Sec. 2 we give a precise operational definition of R-separation for the Helmholtz equation. (We expect, though we have not tried to verify, that any coordinate system which R-separates in accordance with some more intuitive definition of separability can be shown to be equivalent to one of our canonical systems.) In Theorem 1 we derive necessary and sufficient conditions that a Hamilton–Jacobi separable system be R-separable for the Helmholtz equation, and we look at the special case of ordinary separation ($R = 1$), obtaining a new generalization of the Robertson condition for orthogonal separability. In Sec. 3 we develop the symmetry operator approach to R-separation and review the corresponding Hamilton–Jacobi theory. Section 4 contains our main result, Theorem 3, which gives the intrinsic symmetry operator characterization of R-separation. Finally, in Sec. 5 we provide some examples of R-separation and briefly discuss the significance of our results.

The theory presented here is local rather than global. All functions are assumed to be locally analytic.

2. TECHNICAL CONSIDERATIONS

Let $\{x^j\}$ be a local coordinate system on the pseudo-Riemannian manifold. We present here an operational definition of R-separation for the Helmholtz equation

$$\Delta\psi \equiv \frac{1}{g^{1/2}} \partial_i (g^{1/2} g^{ij} \partial_j) \psi = E\psi \tag{2.1}$$

in the coordinates $\{x^j\}$ and derive necessary and sufficient conditions for the existence of this phenomenon. Let $(S_{ij}(x^i))$

[a)] Supported in part by NSF Grant MCS 78-26216.

be a *Stäckel matrix*, i.e., an $N \times N$ nonsingular matrix whose ith row depends only on the variable x^i and set $S = \det(S_{ij})$. We divide the coordinates x^j into three disjoint classes: *essential of type 1*, *essential of type 2*, and *ignorable*. We further order the indices so that n_1 coordinates x^a, $1 \leqslant a \leqslant n_1$, are essential of type 1, the n_2 coordinates x^r, $n_1 + 1 \leqslant r \leqslant n_1 + n_2$, are essential of type 2, and the n_3 coordinates x^α, $n_1 + n_2 + 1 \leqslant \alpha \leqslant n_1 + n_2 + n_3 = n$, are ignorable. (In the following, unless otherwise stated, indices a, b, c range from 1 to n_1, indices r, s, t range from $n_1 + 1$ to $n_1 + n_2$, indices α, β, γ range from $n_1 + n_2 + 1$ to n, and indices i, j, k range from 1 to n.) The *ignorable* coordinates are defined to be all x^i such that $\partial_i g^{jk}(\mathbf{x}) = 0$ for all j, k. Finally, set $N = n_1 + n_2$, let $\lambda_1 = -E, \lambda_2, \ldots, \lambda_N$ be complex parameters, and define differential operators K_a, K_r by

$$K_a = \partial_{aa} + l_a(x^a)\partial_a + m_a(x^a) + \sum_{\alpha,\beta} A_a^{\alpha,\beta}(x^a)\partial_{\alpha\beta} + \sum_\alpha n_a^\alpha(x^a)\partial_\alpha + \sum_{i=1}^N \lambda_i S_{ai}(x^a), \qquad (2.2)$$

for $a = 1, \ldots, n_1$ and

$$K_r = 2\sum_\alpha B_r^\alpha(x^r)\partial_{r\alpha} + m_r(x^r) + \sum_{\alpha,\beta} A^{\alpha,\beta}(x^r)\partial_{\alpha\beta} + \sum_\alpha n_r^\alpha(x^r)\partial_\alpha + \sum_{i=1}^N \lambda_i S_{ri}(x^r) \qquad (2.3)$$

for $r = n_1 + 1, \ldots, N$.

We say that the coordinates $\{x^j\}$ are R-separable for the Helmholtz equation (2.1) provided there exist functions $g_k(\mathbf{x})$ and $R(x^a, x^r)$ $(R \neq 0)$ such that

$$R^{-1}\Delta R - E \equiv \sum_{k=1}^N g_k(\mathbf{x}) K_k. \qquad (2.4)$$

Here

$$R^{-1}\Delta R = \Delta + g^{ij}\partial_i \ln R \partial_j + R^{-1}(\Delta R) \qquad (2.5)$$

as an operator, where

$$\Delta = g^{ij}\partial_{ij} + \frac{1}{g^{1/2}}\partial_i(g^{1/2} g^{ij})\partial_j. \qquad (2.6)$$

If the coordinates are R-separable then the function

$$\psi(\mathbf{x}) = R(x^b, x^s) \prod_a \psi^{(a)}(x^a) \prod_r \psi^{(r)}(x^r) \exp\left[\sum \lambda_\alpha x^\alpha\right] \qquad (2.7)$$

is a solution of $\Delta\psi = E\psi$ whenever the $\psi^{(j)}$ satisfy *separation equations*

$$K_a[\psi^{(a)}\exp(\lambda_\alpha x^\alpha)] = 0, \quad a = 1, \ldots, n_1,$$
$$K_r[\psi^{(r)}\exp(\lambda_\alpha x^\alpha)] = 0, \quad r = n_1 + 1, \ldots, N. \qquad (2.8)$$

Here the λ_α are arbitrary complex constants and $\lambda_1, \ldots, \lambda_n$ are the *separation parameters*. Note that the function $\exp(\lambda_\alpha x^\alpha)$ can be factored out of expressions (2.8), thus reducing these expressions to ordinary differential equations. The *type 1* coordinates x^a have the property that the corresponding separation equations are second order ODE's, whereas for *type 2* coordinates x^r the separation equations are first order ODE's. The solutions $\psi(\mathbf{x}, \boldsymbol{\lambda})$ (2.7), depend on the separation parameters λ_i but $R(x^b, x^s)$ is independent of these parameters.

It follows from (2.2)–(2.4) that a necessary condition for R-separation is

$$g_k(\mathbf{x}) = S^{k1}/S, \quad k = 1, \ldots, N \qquad (2.9)$$

where S^{k1} is the $(k,1)$ minor of (S_{ij}).

Thus the metric must take the form

$$g^{ab} = \delta^{ab}\frac{S^{a1}}{S}, \quad g^{ar} = g^{a\alpha} = 0, \quad g^{rs} = 0,$$
$$g^{r\alpha} = B_r^\alpha(x^r)\frac{S^{r1}}{S}, \qquad (2.10)$$
$$g^{\alpha\beta} = \frac{1}{2}\sum_{i=1}^N A_i^{\alpha,\beta}(x^i)\frac{S^{i1}}{S}, \quad \alpha \neq \beta$$
$$g^{\alpha\alpha} = \sum_{i=1}^N A_i^{\alpha,\alpha}(x^i)\frac{S^{i1}}{S}.$$

Note that

$$\begin{array}{cc} & \begin{array}{ccc} n_1 & n_2 & n_3 \end{array} \\ (g^{ij}) = & \begin{pmatrix} \delta^{ab}g^{aa} & 0 & 0 \\ 0 & 0 & g^{r\alpha} \\ 0 & g^{\alpha r} & g^{\alpha\beta} \end{pmatrix} \end{array} \begin{array}{c} n_1 \\ n_2. \\ n_3 \end{array} \qquad (2.11)$$

Conditions (2.10) are necessary but not sufficient for R-separation. Before determining the remaining conditions, however, it is worthwhile to point out the significance of these restrictions on the metric. Consider the *Hamilton–Jacobi equation* associated with the Helmholtz equation (2.1):

$$g^{ij}\partial_i W \partial_j W = E. \qquad (2.12)$$

It has recently been established,[4–7] that conditions (2.10) are necessary and sufficient for (additive) separation of the Hamilton–Jacobi equation in the coordinates $\{x^j\}$

$$W(\mathbf{x}) = \sum_a W^{(a)}(x^a, \boldsymbol{\lambda}) + \sum_r W^{(r)}(x^r, \boldsymbol{\lambda}) + \sum_\alpha \lambda_\alpha x^\alpha \qquad (2.13)$$

Indeed, Benenti[7] has shown that every system which separates (2.12), according to the intuitive definition of Levi-Civita,[8] is equivalent to a system in the canonical form (2.10).

Proposition 1: A coordinate system that is R-separable for the Helmholtz equation is also separable for the Hamilton–Jacobi equation. Let

$$H_i^{-2} = \frac{S^{i1}}{S}, \quad i = 1, \ldots, N. \qquad (2.14)$$

If conditions (2.10) hold then $S^{i1} \neq 0$ since $g \neq 0$. We can associate with our coordinate system $\{x^j\}$ on V_n an orthogonal coordinate system $\{x^1, \ldots, x^N\}$ on a space V_N with metric

$$ds^2 = \sum_{i=1}^N H_i^2 (dx^i)^2. \qquad (2.15)$$

By (2.14), this metric is in Stäckel form.[2] Recall that necessary and sufficient conditions that ds^2 be expressible in the form (2.14) for some Stäckel matrix are (Ref. 1, Appendix 13)

$$\partial_{jk}\ln H_i^{-2} + \partial_j \ln H_i^{-2}\partial_k \ln H_i^{-2} - \partial_j \ln H_i^{-2}\partial_k \ln H_j^{-2} - \partial_k \ln H_i^{-2}\partial_j \ln H_k^{-2} = 0,$$
$$j \neq k; \quad i, j, k = 1, \ldots, N. \qquad (2.16)$$

We further recall some useful results from Ref. 6. Given a metric $ds^2 = \Sigma_i H_i^2 (dx^i)^2$ in Stäckel form, we say that the function $Q(\mathbf{x})$ is a *Stäckel multiplier* for (ds^2) if the metric $d\tilde{s}^2 = Q ds^2$ is also in Stäckel form with respect to the coordinates $\{x^j\}$. It can be shown that Q is a Stäckel multiplier if and only if there exist functions $\psi_j = \psi_j(x^j)$ such that

$$Q(\mathbf{x}) = \sum_{j=1}^{N} \psi_j(x^j) H_j^{-2}. \tag{2.17}$$

Equivalent necessary and sufficient conditions are

$$\partial_{jk} Q - \partial_j Q \partial_k \ln H_j^{-2} - \partial_k Q \partial_j \ln H_k^{-2} = 0, \quad j \neq k. \tag{2.18}$$

We can now reformulate conditions (2.10).

Proposition 2: A necessary requirement for R-separation of (2.1) in the coordinates $\{x^i : i = 1,...,n\}$ is that

$$g^{aa} = H_a^{-2}, \quad g^{r\alpha} = B_r^{\alpha}(x^r) H_r^{-2}, \tag{2.19}$$

and that each $g^{\alpha\beta}$ be a Stäckel multiplier for the Stäckel form metric $ds^2 = \Sigma_{k=1}^{N} H_k^2 (dx^k)^2$. All other matrix elements g^{ij} must vanish.

To obtain sufficient conditions for R-separation we must also demand equality of the coefficients of ∂_j and the zeroth order terms on each side of (2.5):

$$f_a + 2\partial_a \ln R = l_a(x^a), \tag{2.20}$$

$$\sum_r g^{r\alpha}(f_{r\alpha} + 2\partial_r \ln R) = \sum_{k=1}^{N} H_k^{-2} n_k^{\alpha}(x^k), \tag{2.21}$$

$$R^{-1}(\Delta R) = \sum_{k=1}^{N} H_k^{-2} m_k(x^k). \tag{2.22}$$

Here,

$$f_a = \partial_a f, \quad f = \ln(g^{1/2}/S), \tag{2.23}$$

$$f_{r\alpha} = \partial_r \ln(g^{1/2} g^{r\alpha}) = f_r + \partial_r \ln B_r^{\alpha}(x^r).$$

Solving for R from (2.19) we find

$$R = \left(\frac{S}{g}\right)^{1/2} \exp\left[\sum_a A_a(x^a) + Q(x^s)\right], \tag{2.24}$$

and substituting (2.23) into (2.20) and (2.21) we ultimately obtain the following result.

Theorem 1: Necessary and sufficient conditions that the coordinates $\{x^j\}$ be R-separable for the Helmholtz equation

$$\frac{1}{g^{1/2}} \partial_i (g^{1/2} g^{ij} \partial_j \psi) = E\psi$$

are

(1) The requirements of Proposition 2 are satisfied, i.e., the coordinates $\{x^j\}$ are separable for the Hamilton–Jacobi equation $g^{ij}\partial_i W \partial_j W = E$,

(2) $\Sigma_r g^{r\alpha} \partial_r Q$ is a Stäckel multiplier for each α,

(3) $\Sigma_a H_a^{-2}(f_{aa} + \frac{1}{2} f_a^2)$ is a Stäckel multiplier, where $f_a = \partial_a \ln(g^{1/2}/S)$ and S is the determinant of the Stäckel matrix.

If these conditions are satisfied then

$$R(\mathbf{x}) = \left(\frac{S}{g^{1/2}}\right)^{1/2} \exp\left[\sum_a A_a(x^a) + Q(x^s)\right],$$

where the $A_a = A_a(x^a)$ are arbitrary.

We say that the coordinates $\{x^j\}$ are *separable* for the Helmholtz equation provided they are R-separable with $R \equiv 1$. Furthermore, R-separable coordinates are *trivially R-separable* if $R = \Pi_{i=1}^{n} R_i(x^i)$ and (since coordinates are trivially R-separable if and only if they are separable) we regard trivial R-separation as equivalent to ordinary separation.

Especially interesting is the case of ordinary separation. Then $R \equiv 1$ and expression (2.23) becomes

$$\tfrac{1}{2}\ln\left(\frac{g^{1/2}}{S}\right) = \sum_a A_a(x^a) + Q(x^s). \tag{2.25}$$

Corollary 1 (Generalized Robertson Condition): Necessary and sufficient conditions that the coordinates $\{x^j\}$ be separable for the Helmholtz equation are

(1) the coordinates are separable for the Hamilton–Jacobi equation,

(2) $f_{\alpha j} = 0$ for $j = 1,...,N, j \neq \alpha$,

(3) $\Sigma_r g^{r\alpha} f_r$ is a Stäckel multiplier for each α.

Here $f = \ln(g^{1/2}/S)$ and $f_i = \partial_i f$.

The original Robertson condition[9] was concerned with the case of orthogonal separation. (By permitting a type 1 coordinate to be ignorable if necessary, we can identify this case with $n_1 = n$, $n_2 = n_3 = 0$.) Robertson showed that an orthogonal separable system for the Hamilton–Jacobi equation separated the Helmholtz equation if and only if $f_{ab} = 0$ for $a \neq b$. (Since $n_2 = 0$ this agrees with Corollary 1.)

Eisenhart[2] showed that the Robertson condition is equivalent to the requirement

$$R_{ab} = 0, \quad a \neq b \tag{2.26}$$

where R_{ab} is the Ricci tensor expressed in terms of the orthogonal coordinates $\{x^a\}$. (For an explicit definition of the Ricci tensor R_{ij} in terms of the metric g^{ij} together with related computational formulas we refer the reader to Chap. 1 of Eisenhart's text.[1]) Benenti[10] studied nonorthogonal separation for the Helmholtz equation in which no nonignorable null coordinates were allowed ($n_2 = 0$ in our formalism). His requirement for Helmholtz separation agrees with our condition (2). Benenti further showed that his requirement was equivalent to (2.20) again and that $R_{\alpha\alpha} = 0$ automatically for Hamilton–Jacobi separable systems. By a tedious but straightforward computation we have established

Lemma 1: Condition (2) of Corollary 1, namely

$$f_{\alpha j} = 0 \quad \text{for} \quad j = 1,...,N, \quad j \neq \alpha$$

is equivalent to

$$R_{ab} = 0, \quad a \neq b, \quad R_{ar} = 0, \tag{2.27}$$

where R_{ij} is the Ricci tensor for V_n expressed in the coordinates $\{x^j\}$. Furthermore, $R_{\alpha\alpha} = 0$ automatically if $\{x^j\}$ separates the Hamilton–Jacobi equation.

It is perhaps somewhat surprising that requirements (2.25) continue to hold even with the presence of type 2 coordinates. Condition (3) of Corollary 1 appears not to be expressible in terms of the Riemann curvature tensor and its covariant derivatives. However, this condition is vacuous for $n_2 \leqslant 1$. Since $g^{rs} = 0$, type 2 coordinates are null and any two such coordinates are orthogonal. Thus, for separation on a proper Riemannian space V_n we must have $n_2 = 0$ and for a pseudo-Riemannian V_n with signature $(-1, 1^{n-1})$ we must have $n_2 \leqslant 1$.

Corollary 2: In order that Hamilton–Jacobi separable coordinates $\{x^j\}$ separate the Helmholtz equation on a pseudo-Riemannian manifold with signature (1^n) or $(-1,1^{n-1})$ it is necessary and sufficient that

$$R_{ab}=0,\quad a\neq b,\quad R_{ar}=0.$$

3. CONSTANTS OF THE MOTION

Let us suppose that the coordinates $\{x^j\}$ R-separate the Helmholtz equation. Then expanding the corresponding Stäckel matrix in (2.2), (2.3) by the l th, rather than just the 1st, column we obtain operators $\mathscr{A}_l$, $l=1,\ldots,N$, such that $\mathscr{A}_l\psi=-\lambda_l\psi$ for an R-separated solution ψ:

$$\begin{aligned}\mathscr{A}_l=\sum_a\frac{S^{al}}{S}\Big(&\partial_{aa}+f_a\partial_a+\sum_{\alpha,\beta}A_a^{\alpha,\beta}\partial_{\alpha\beta}+\sum_\alpha n_a^\alpha\partial_\alpha\\&+m_a+\tfrac12\partial_a[f_a-l_a]+\tfrac14[f_a^2-l_a^2]\Big)\\+\sum_r\frac{S^{rl}}{S}\Big(&2\sum_\alpha B_r^\alpha\partial_{r\alpha}+\sum_{\alpha,\beta}A_r^{\alpha,\beta}\partial_{\alpha\beta}\\&+\sum_\alpha(n_r^\alpha-2B_r^\alpha\partial_r\ln R)\partial_\alpha+m_r\Big).\end{aligned}\tag{3.1}$$

(Note that $\mathscr{A}_1=\Delta$.) These expressions are not as complicated as they appear. It can be directly verified (and we will show this later) that

$$\begin{aligned}&[\mathscr{A}_l,\mathscr{A}_k]=0,\quad[\mathscr{L}_\alpha,\mathscr{L}_\beta]=0,\\&[\mathscr{A}_l,\mathscr{L}_\alpha]=0,\quad 1\leqslant l,k\leqslant N\end{aligned}\tag{3.2}$$

where

$$\mathscr{L}_\alpha=\partial_\alpha,\quad\alpha=N+1,\ldots,n,\tag{3.3}$$

and $[\mathscr{A},\mathscr{B}]=\mathscr{A}\mathscr{B}-\mathscr{B}\mathscr{A}$. Thus the operators $\mathscr{A}_k$ $(2\leqslant k\leqslant N)$, $\mathscr{L}_\alpha$ form a commuting family of symmetry operators for Δ, i.e., they commute with Δ and with each other. Furthermore, the R-separated solutions of (2.2) are simultaneous eigenfunctions of the symmetry operators:

$$\mathscr{A}_l\psi=-\lambda_l\psi,\quad\mathscr{L}_\alpha\psi=\lambda_\alpha\psi.\tag{3.4}$$

Our construction has started with an R-separable coordinate system $\{x^l\}$ and produced a commuting family of symmetry operators $\{\mathscr{A}_l,\mathscr{L}_\alpha\}$. It is our principal task in this paper to characterize those families of commuting symmetry operators that correspond to R-separation.

In Ref. 6 the authors solved the corresponding problem for the Hamilton–Jacobi equation (2.12). In that case we utilized the natural symplectic structure on the cotangent bundle $\tilde V_n$ of V_n. Corresponding to local coordinates $\{x^i\}$ on V_n we have coordinates $\{x^i,p_i\}$ on the $2n$-dimensional space $\tilde V_n$. The *Poisson bracket* of two functions $F(x^j,p_j)$, $G(x^j,p_j)$ on $\tilde V_n$ is defined by

$$\{F,G\}=\sum_{l=1}^n(\partial_{p_l}F\partial_{x^l}G-\partial_{x^l}F\partial_{p_l}G).\tag{3.5}$$

Let $\{x^i\}$ be a separable coordinate system for the Hamilton–Jacobi equation (2.12) with coordinates of type 1, x^a, of type 2, x^r, and ignorable, x^α, as usual. Then the metric g^{ij} in these coordinates takes the standard form (2.10).

It is convenient at this point to introduce the functions $\rho_j^{(k)}(x^1,\ldots,x^N)$, where

$$\frac{S^{jk}}{S}=\rho_j^{(k)}H_j^{-2},\quad\frac{S^{j1}}{S}=H_j^{-2},\quad 1\leqslant j,k\leqslant N,\tag{3.6}$$

and S_{ij} is the Stäckel matrix corresponding to the separable system $\{x^i\}$. Then $\rho_j^{(1)}=1$ and it can be shown that (Ref. 1, Appendix 13)

$$\partial_i\rho_j^{(k)}=(\rho_i^{(k)}-\rho_j^{(k)})\partial_i\ln H_j^{-2},\quad 1\leqslant i,j,k,\leqslant N.\tag{3.7}$$

Let $H=\Sigma_{i,j}g^{ij}p_ip_j$ be the Hamiltonian corresponding to (2.12). In Ref. 6 we constructed quadratic forms A_l $(A_1=H)$, given by

$$\begin{aligned}A_l=&\sum_a\rho_a^{(l)}H_a^{-2}\Big(p_a^2+\sum_{\alpha,\beta}A_a^{\alpha,\beta}p_\alpha p_\beta\Big)\\&+\sum_r\rho_r^{(l)}H_r^{-2}\Big(\sum_\alpha B_r^\alpha p_rp_\alpha+\sum_{\alpha,\beta}A_r^{\alpha,\beta}p_\alpha p_\beta\Big)\end{aligned}\tag{3.8}$$

for $l=1,\ldots,N$ and n_3 linear forms L_α,

$$L_\alpha=p_\alpha,\quad\alpha=N+1,\ldots,n.\tag{3.9}$$

These polynomials in the p's were shown to satisfy

$$\begin{aligned}&\{A_l,A_k\}=0,\quad\{L_\alpha,L_\beta\}=0,\\&\{A_l,L_\alpha\}=0,\quad l,k=1,\ldots,N,\end{aligned}\tag{3.10}$$

and when evaluated for $p_a=\partial_aW,p_r=\partial_rW,p_\alpha=\partial_\alpha W$ with W a separable solution of (2.12), they satisfy

$$A_l=-\lambda_l,\quad L_\alpha=\lambda_\alpha,\tag{3.11}$$

where $\lambda_1=-E,\ldots,\lambda_n$ are the separation parameters.

Let $a^{ij}(y)$ be a symmetric contravariant 2-tensor on V_n, expressed in terms of local coordinates $\{y^k\}$, and let $g^{ij}(y)$ be the contravariant metric tensor. A *root* $\rho(y)$ of a^{ij} is a solution of the characteristic equation

$$\det(a^{ij}-\rho g^{ij})=0\tag{3.12}$$

and an *eigenform* $\omega=\Sigma\lambda_kdy^k$ corresponding to ρ is a nonzero 1-form such that

$$\sum_{j=1}^n(a^{ij}-\rho g^{ij})\lambda_j=0,\quad i=1,\ldots,n.\tag{3.13}$$

Roots and eigenforms are defined independent of local coordinates.

Note from (3.8) that for a separable system $\{y^j\}$ the $\rho_a^{(l)}$ are simple roots of the A_l with simultaneous eigenforms dx^a, and the $\rho_r^{(l)}$ are roots of multiplicity 2 but with a single eigenform dx^r. Here dx^a,dx^r are also eigenforms for the products $L_\alpha L_\beta$.

Let $\{y^j\}$ be a local coordinate system on a pseudo-Riemannian manifold and let $\omega_{(j)}=\lambda_{i(j)}dy^i$, $1\leqslant j\leqslant n$, be a local basis of 1-forms on V_n. The *dual basis* of vector fields is $\underline{X}^{(h)}=\Lambda^{i(h)}\partial_i$, $1\leqslant h\leqslant n$, where $\Lambda^{i(h)}\lambda_{i(j)}=\delta_{(j)}^{(h)}$. The *inner product* of two 1-forms $\omega_{(j)},\omega_{(k)}$ is $G(j,k)=\lambda_{i(j)}g^{il}\lambda_{l(k)}$. In Ref. 6 we proved

Theorem 2: Let θ be a vector subspace of quadratic forms on V_n such that $H\in\theta$ and

(1) $\{A,B\}=0$ for each $A,B\in\theta$,

(2) there is a basis of 1-forms $\omega_{(j)}=\lambda_{i(j)}dy^i$, $1\leqslant j\leqslant n$, such that

(i) the n_1 forms $\omega_{(a)}$ are simultaneous eigenforms for each $A\in\theta$ with root ρ_a^A,

(ii) the n_2 forms $\omega_{(r)}$ are simultaneous eigenforms for each $A\in\theta$ with root ρ_r^A; the root ρ_r^A has multiplicity 2 but corresponds to only one simultaneous eigenform,

(3) $\{L_\alpha, L_\beta\} = 0$ and $L_\alpha L_\beta \in \theta$, where $L_\alpha = \wedge^{i(\alpha)} p_i$, $\alpha,\beta = n_1 + n_2 + 1,\ldots,n$,
(4) $\{A, L_\alpha\} = 0$ for each $A \in \theta$,
(5) $\underline{\overline{X}}^{(r)}(\lambda_{i(\alpha)} a^{ij} \lambda_{j(\beta)}) = \rho_r{}^A \underline{\overline{X}}^{(r)}(\lambda_{i(\alpha)} g^{ij} \lambda_{j(\beta)})$,
(6) $\dim\theta = \frac{1}{2}(2n + n_3{}^2 - n_3)$, where $n_3 = n - n_1 - n_2$,
(7) $G(a,b) = 0$ if $a \neq b$, and $G(a,r) = G(a,\alpha) = G(r,s) = 0$.

Then there exist local coordinates $\{x^j\}$ for V_n and functions $f^{(j)}(\mathbf{x})$ such that $\omega_{(j)} = f^{(j)} dx^j$ (with a suitable modification of the $\omega_{(\alpha)}$) and the Hamilton–Jacobi equation is separable in these coordinates. Conversely, to every separable coordinate system $\{x^j\}$ for the Hamilton–Jacobi equation there corresponds a subspace θ of quadratic forms on V_n with properties (1)–(7).

In the following section we will show that, with suitable modifications, this result also characterizes R-separable systems for the Helmholtz equation.

4. THE BASIC RESULT

Let Δ be the Hamiltonian operator (1.2), expressed in terms of local coordinates $\{x^j\}$. Suppose $\mathscr{A}$ is a *second order symmetry operator* for Δ, i.e., a differential operator such that $[\mathscr{A},\Delta] = 0$ and which in local coordinates can be written

$$\mathscr{A} = a^{ij}(\mathbf{y})\partial_{ij} + \bar{b}^i(\mathbf{y})\partial_i + c(\mathbf{y}), \quad \partial_i = \partial_{y^i} \tag{4.1}$$

where $a^{ij} = a^{ji}$ and not all a^{ii} vanish. As shown in Ref. 3 we can decompose $\mathscr{A}$ uniquely in the form

$$\mathscr{A} = \mathscr{S} + \mathscr{L}, \tag{4.2}$$

where

$$\mathscr{S} = \frac{1}{g^{1/2}}\partial_i(g^{1/2} a^{ij}\partial_j) + c,$$
$$\mathscr{L} = b^i\partial_i, \tag{4.3}$$
$$[\mathscr{S},\Delta] = [\mathscr{L},\Delta] = 0. \tag{4.4}$$

Furthermore, this decomposition is coordinate independent. Decomposing the operators $\mathscr{A}$, (3.1), in this form we find

$$\mathscr{A}_l = \mathscr{S}_l + \hat{\mathscr{L}}_l,$$
$$\mathscr{S}_l = \frac{1}{g^{1/2}}\partial_i(g^{1/2} a^{ij}_{(l)}\partial_j) + \sum_a \rho_a^{(l)} H_a^{-2}(m_a + \tfrac{1}{2}\partial_a[f_a - l_a] + \tfrac{1}{4}[f_a^2 - l_a^2]) + \sum_r \rho_r^{(l)} H_r^{-2} m_r, \tag{4.5}$$
$$\hat{\mathscr{L}}_l = \Big[\sum_{i=1}^N \rho_i^{(l)} H_i^{-2} n_i^\alpha - \sum_r \rho_r^{(l)} H_r^{-2} B_r^\alpha(\partial_r \ln B_r^\alpha + \partial_r Q)\Big]\partial_\alpha,$$

for $l = 1,\ldots,N$, where

$$A_l = a^{ij}_{(l)} p_i p_j \tag{4.6}$$

is the quadratic form (3.8). Note that $\hat{\mathscr{L}}_l$ is not only a symmetry operator for Δ, but it in addition is *functionally dependent* on the first order symmetries $\mathscr{L}_\alpha$, (3.3). That is, there exist functions $g_l^\alpha(\mathbf{x})$ such that

$$\hat{\mathscr{L}}_l = \sum_\alpha g_l^\alpha(\mathbf{x})\mathscr{L}_\alpha. \tag{4.7}$$

Returning to the general symmetry operator $\mathscr{A}$, (4.1)–(4.4), we can uniquely associate this operator with the quadratic form A on $\tilde{V}_n$, defined in local coordinates by

$$A = \sum_{i,j} a^{ij} p_i p_j. \tag{4.8}$$

We can talk about the *roots* and *eigenforms* of $\mathscr{A}$, meaning by this the roots and eigenforms of A. The following analogy of Theorem 2 holds.

Theorem 3: Let $\{\mathscr{A}_1 = \Delta, \mathscr{A}_2,\ldots,\mathscr{A}_N\}$ be a set of second order symmetry operators for Δ with $\{A_l\}$ linearly independent, and let $\{\mathscr{L}_{N+1},\ldots,\mathscr{L}_n\}$ $(n - N = n_3)$ be a linearly independent set of first order symmetry operators such that
(1) $[\mathscr{A}_l,\mathscr{A}_k] = 0$, $[\mathscr{A}_l,\mathscr{L}_\alpha] = 0$, $[\mathscr{L}_\alpha,\mathscr{L}_\beta] = 0$,
(2) each $\hat{\mathscr{L}}_l$ is functionally dependent on the set $\{\mathscr{L}_\alpha\}$, where $\mathscr{A}_l = \mathscr{S}_l + \hat{\mathscr{L}}_l$ is the canonical decomposition (4.1)–(4.4) of $\mathscr{A}_l$,
(3) no $\mathscr{A}_l$ belongs to the associative algebra generated by $\{\mathscr{L}_\alpha\}$, i.e., $\mathscr{A}_l$ cannot be expressed as $c_l^{\alpha\beta}\mathscr{L}_\alpha\mathscr{L}_\beta$ for constants $c_l^{\alpha\beta}$,
(4) there is a basis of 1-forms $\omega_{(j)} = \lambda_{i(j)} dy^i$, $1 \leqslant j \leqslant n$, such that $(n_1 + n_2 = N)$
(i) the n_1 forms $\omega_{(a)}$ are simultaneous eigenforms for each A_l with root $\rho_a^{(l)}$,
(ii) the n_2 forms $\omega_{(r)}$ are simultaneous eigenforms for each A_l with double root $\rho_r^{(l)}$; the root corresponds to only one eigenform,
(iii) $\mathscr{L}_\alpha = \wedge^{i(\alpha)}\partial_i$,
(5) $\underline{\overline{X}}^{(r)}(\lambda_{i(\alpha)} a^{ij}_{(l)} \lambda_{j(\beta)}) = \rho_r^{(l)} \underline{\overline{X}}^{(r)}(\lambda_{i(\alpha)} g^{ij} \lambda_{j(\beta)})$,
(6) $G(a,b) = 0$ if $a \neq b$, and $G(a,r) = G(a,\alpha) = G(r,s) = 0$.

Then there exist local coordinates $\{x^j\}$ for V_n and functions $f^{(j)}(\mathbf{x})$ such that $\omega_{(j)} = f^{(j)} dx^j$ (with a suitable modification of the $\omega_{(\alpha)}$) and the Helmholtz equation (2.1) is R-separable in these coordinates. Conversely, to every R-separable coordinate system $\{x^j\}$ for the Helmholtz equation there correspond operators $\mathscr{A}_j, \mathscr{L}_\alpha$ on V_n with properties (1)–(6).

Proof: Suppose conditions (1)–(6) are satisfied. Comparing coefficients of the highest order (nonvanishing) derivative terms in condition (1) we find

$$\{A_l, A_k\} = 0, \quad \{A_l, L_\alpha\} = 0, \quad \{L_\alpha, L_\beta\} = 0,$$

where $L_\alpha = \wedge^{i(\alpha)} p_i$. It follows from this and conditions (3)–(6) that the hypotheses of Theorem 2 are satisfied. Indeed the subspace θ is that with basis $\{A_l, L_\alpha L_\beta \alpha \leqslant \beta\}$. Hence, there exists a local coordinate system $\{x^j\}$ such that the functions A_l, L_α can be expressed in the form (3.8). If $A_l = a^{ij}_{(l)} p_i p_j$ then by condition (2) and the fact that $\det(\rho_k^{(l)}) \neq 0$ we can write $\mathscr{A}_l = \mathscr{S}_l + \hat{\mathscr{L}}_l$, where

$$\mathscr{S}_l = \frac{1}{g^{1/2}}\partial_i(g^{1/2} a^{ij}_{(l)}\partial_j) + \sum_{k=1}^N \rho_k^{(l)} H_k^{-2}\xi^k, \tag{4.9}$$
$$\hat{\mathscr{L}}_l = \sum_{k=1}^N \rho_k^{(l)} H_k^{-2}\xi^{k\alpha}\partial_\alpha,$$

and

$$\sum_{k=1}^N H_k^{-2}\xi^k = 0, \quad \sum_{k=1}^N H_k^{-2}\xi^{k\alpha} = 0, \tag{4.10}$$

since $\mathscr{A}_1 = \Delta$ and $\rho_k^{(1)} = 1$.

We have not yet fully utilized condition (1). Since $\mathscr{S}_l$ is self adjoint and $\hat{\mathscr{L}}_l, \mathscr{L}_\alpha$ are skew adjoint,[3] the first two equations in condition (1) yield

$$[\hat{\mathscr{L}}_l, \mathscr{L}_\alpha] = 0, \quad (4.11a)$$

$$[\hat{\mathscr{L}}_l, \mathscr{L}_k] = 0, \quad (4.11b)$$

$$[\mathscr{S}_l, \mathscr{S}_k] = 0, \quad (4.11c)$$

$$[\mathscr{S}_l, \hat{\mathscr{L}}_k] + [\hat{\mathscr{L}}_l, \mathscr{S}_k] = 0. \quad (4.11d)$$

Equation (4.11a) yields $\partial_\alpha \xi^{k\beta} = 0$ and (4.11b) is satisfied identically. Equating coefficients of ∂_{ij} on both sides of (4.11c) we find $\partial_a f_b = \partial_b f_a$, $\partial_r f_a = \partial_r f_{ra}$, a result already known. Equating coefficients of ∂_i on both sides of (4.11c) and using $\det(\rho_k^{(l)}) \neq 0$ we find

$$\partial_b \xi^r = 0, \quad \partial_b(2\xi^a - f_{aa} - \tfrac{1}{2}f_a^2) = 0, \quad a \neq b,$$

$$\partial_s(2\xi^a - f_{aa} - \tfrac{1}{2}f_a^2) = 0,$$

$$B_r^\alpha \partial_r \xi^s = B_s^\alpha \partial_s \xi^r, \quad r \neq s \quad \text{(no sum)}.$$

Since the last equality must hold for all α, we have $\partial_s \xi^r = 0$ for $r \neq s$. Thus

$$\xi^a = \tfrac{1}{2}[f_{aa} + \tfrac{1}{2}f_a^2 + 2P_a(x^a)],$$

$$\xi^r = P_r(x^r)$$

and from (4.10) we see that

$$\sum_a H_a^{-2}(f_{aa} + \tfrac{1}{2}f_a^2) \quad (4.12)$$

is a Stäckel multiplier. Thus condition (3) [and condition (1)] of Theorem 1 are satisfied. [The zeroth order terms in (4.11c) give no new requirements.]

The only constraints remaining to us are (4.11d). Equating coefficients of ∂_{ab} in this expression we find

$$\partial_b \xi^{r\alpha} = 0, \quad \partial_b \xi^{a\alpha} = 0, \quad b \neq a.$$

Equating coefficients of $\partial_{\alpha\beta}$ we find

$$B_r^\beta \partial_r \xi^{a\alpha} + B_r^\alpha \partial_r \xi^{a\beta} = 0,$$

$$B_r^\beta \partial_r \xi^{s\alpha} + B_r^\alpha \partial_r \xi^{s\beta} = B_s^\beta \partial_s \xi^{r\alpha} + B_s^\alpha \partial_s \xi^{r\beta}, \quad r \neq s.$$

Thus

$$\xi^{r\alpha} = T_r^\alpha(x^t), \quad \xi^{a\alpha} = V_a^\alpha(x^a), \quad (4.13)$$

where

$$B_r^\beta \partial_r T_s^\alpha + B_r^\alpha \partial_r T_s^\beta = B_s^\beta \partial_s T_r^\alpha + B_s^\alpha \partial_s T_r^\beta, \quad r \neq s, \quad \text{no sum}. \quad (4.14)$$

To solve relations (4.14) for T_s^α we use the fact that the $n_2 \times n_3$ matrix $(B_r^\beta(x^r))$ has rank n_2. The ignorable coordinates $\{x^\alpha\}$ are not unique. A new set of ignorable coordinates $\{x'^\beta\}$, where $x'^\beta = C_\alpha^\beta x^\alpha$ and (C_α^β) is a nonsingular constant matrix, will do as well. One effect of such a choice of new ignorable coordinates is to provide a new matrix $(B'^\beta_r(x^r))$ constructible from the original matrix by a sequence of elementary column transformations. Conversely, elementary column transformations of (B_r^β) induce transformations of ignorable coordinates. Assuming $n_2 \geqslant 2$ [since otherwise (4.14) is vacuous] we can always choose a new set of ignorable coordinates $\{x'^\beta\}$ such that every matrix element B'^α_r and every 2×2 minor in the new matrix are nonvanishing in a suitably small $\mathbf{x}'$-coordinate neighborhood. Assuming this done and dropping the primes we set $\alpha = \beta$ in (4.14) to obtain

$$\partial_r(T_s^\beta / B_s^\beta) = \partial_s(T_r^\beta / B_r^\beta), \quad r \neq s. \quad (4.15)$$

Substituting this result back into (4.14) and simplifying we obtain

$$\left(\frac{B_s^\alpha B_r^\beta - B_r^\alpha B_s^\beta}{B_s^\alpha B_r^\alpha B_s^\beta B_r^\beta}\right)\left(\partial_r\left(\frac{T_s^\alpha}{B_s^\alpha}\right) - \partial_r\left(\frac{T_s^\beta}{B_s^\beta}\right)\right) = 0. \quad (4.16)$$

It follows from (4.16) that

$$T_s^\alpha = B_s^\alpha(x^s) Z_s + P_s^\alpha(x^s) \quad (4.17)$$

and from (4.15) that $\partial_r Z_s = \partial_s Z_r$, $r \neq s$.

Thus there exists a function $Q(x^t)$ (depending on type 2 variables only) such that $Z_s = -2\partial_s Q$.

We conclude that

$$\xi^{r\alpha} = -2B_r^\alpha \partial_r Q(x^s) + P_r^\alpha(x^r), \quad \xi^{a\alpha} = V_a^\alpha(x^a). \quad (4.18)$$

Substituting this result into (4.10) we see that $\Sigma_r g^{r\alpha}\partial_r Q$ is a Stäckel multiplier. Thus all conditions of Theorem 1 are satisfied and the coordinates $\{\mathbf{x}'\}$ (hence the coordinates $\{\mathbf{x}\}$) R-separate the Helmholtz equation. [We note that the first derivative terms in (4.11d) yield no new restrictions.]

Conversely, if the coordinates $\{x^j\}$ R-separate the Helmholtz equation we can reverse the order of the above argument and verify conditions (1)–(6). Q.E.D.

5. DISCUSSION AND EXAMPLES

Theorem 2 states that a Hamilton–Jacobi separable system $\{x^j\}$ is R-separable for the Helmholtz equation if and only if the involutive family of Killing tensors A_l, L_α corresponds to a commutative family of symmetry operators $\mathscr{A}_l, \mathscr{L}_\alpha$. The technical conditions (2) and (3) of Theorem 1 are necessary and sufficient that such a correspondence exists. In this sense our results have a close relationship with quantization theory.

Note that if the operators $\mathscr{A}_l, \mathscr{L}_\alpha$ satisfy the hypotheses of Theorem 3, except for requirement (2), then the operators $\mathscr{S}_l, \mathscr{L}_\alpha$ define an R-separation of the Helmholtz equation.

Our generalization of variable separation for the Helmholtz equation to R-separation and including null coordinates would be of little value unless nontrivial R-separation exists. In fact, all of the phenomena discussed in this paper do occur. For examples of ordinary separation involving type 2 (null) coordinates see Refs. 4, 5, and 11. For examples (and a theory) of nontrivial orthogonal R-separation see Refs. 3 and 12. Here, we merely recall one example of nonorthogonal R-separation from Ref. 12 to show how it relates to the general theory. The example is a V_4 with local coordinates $(x^1, \ldots, x^4) \equiv (x, y, \alpha, \beta)$ and metric

$$(g^{ij}) = \begin{pmatrix} 0 & 0 & e^x & 1 \\ 0 & 0 & e^y & 1 \\ e^x & e^y & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \quad (5.1)$$

Thus, $n_2 = n_3 = 2$, $n = 4$. The coordinates are easily checked to be Hamilton–Jacobi separable and $f =$

$\ln(g^{1/2}/S) = -\ln(e^y - e^x)$. Since $n_1 = 0$, condition (3) of Theorem 1 is satisfied. We first check ordinary separability. Here $H_x^{-2} = H_y^{-2} = 1$ and $g^{x\alpha}f_x + g^{y\alpha}f_y = -e^x - e^y$, $g^{x\beta}f_x + g^{y\beta}f_y = -1$ so $\Sigma_r g^{r\gamma}f_r$ is always a Stäckel multiplier. It follows that the Helmholtz equation separates in the coordinates $\{x^j\}$. We have shown that $Q = f$ satisfies condition (2) in Theorem 1. However, once we have separation we can achieve further R-separation by choosing Q to be any other function satisfying condition (2). In particular choose $Q = 0$. Then the Helmholtz equation R-separates in the coordinates $\{x^j\}$ with $R = (e^y - e^x)^{1/2}$. (The phenomenon of multiple R-separation for a single coordinate system is possible only if type 2 coordinates are present.) In Ref. 12 we give the operator characterizations of these coordinates in accordance with Theorem 3.

Upon comparison of Theorem 2 and 3 it is clear that R-separation and not just ordinary separation is the appropriate Helmholtz analogy of separation for the Hamilton–Jacobi equation.

ACKNOWLEDGMENT

One of us (W.M.) would like to thank the University of Waikato for a postdoctoral fellowship and generous hospitality that made this collaboration possible.

[1] L. P. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, NJ, 1949).

[2] L. P. Eisenhart, "Separable systems of Stäckel," Ann. Math. **35**, 284–305 (1934).

[3] E. G. Kalnins and W. Miller, Jr., "The Theory of orthogonal R-separation for Helmholtz equations," Adv. Math. (to appear).

[4] C. P. Boyer, E. G. Kalnins, and W. Miller, Jr., "Separable coordinates for four-dimensional Riemannian spaces," Commun. Math. Phys. **59**, 285–302 (1978).

[5] E. G. Kalnins and W. Miller, Jr., "Non-orthogonal separable coordinate systems for the flat 4-space Helmholtz equation," J. Phys. A: Math., **12**, 1129–1147 (1979).

[6] E. G. Kalnins and W. Miller, Jr., "Killing tensors and nonorthogonal variable separation for Hamilton-Jacobi equations," SIAM J. Math. Anal. **12**, 617–638 (1981).

[7] S. Benenti, "Separability structures on Riemannian manifolds," *Proceedings of Conference on Differential Geometrical Methods in Mathematical Physics*, Salamanca 1979, Lecture Notes in Mathematics **836** (Springer-Verlag, Berlin, 1980).

[8] T. Levi-Civita, "Sulla integrazione della equazione di Hamilton-Jacobi per separazione di variabili," Math. Ann. **59**, 383–397 (1904).

[9] H. P. Robertson, "Bemerkung uber separierbare Systeme in der Wellenmechanik," Math. Ann. **98**, 749–752 (1928).

[10] S. Benenti, "Integrabilita per separazione delle variabili delle equazioni alle derivate parziali lineari del secondo ordine interessanti la fisica-matematica," Lincei-Rend. Sci. Fis. Mat. Nat. **62**, 51–60 (1977).

[11] E. G. Kalnins and W. Miller, Jr., "Separable coordinates for three-dimensional complex Riemannian spaces," J. Diff. Geom. **14**, 221–236 (1979).

[12] E. G. Kalnins and W. Miller, Jr., "Some remarkable R-separable coordinate systems for the Helmholtz equation," Lett. Math. Phys. **4**, 469–474 (1980).

Some infinite series of products of Legendre and gamma functions

H. van Haeringen
Department of Mathematics, Delft University of Technology, Delft, The Netherlands

(Received 22 February 1982; accepted for publication 5 March 1982)

We derive closed expressions for some infinite series of products of Legendre functions and gamma functions. A particular series has been used to obtain the partial-wave projected quantum mechanical Coulomb transition matrix in closed analytic form for all partial waves, $l = 0, 1, \cdots$.

PACS numbers: 02.30. + g, 03.65.Nk

1. INTRODUCTION

In this paper we shall derive closed expressions for some infinite series involving products of Legendre functions and products of gamma functions. We have used one of these series to obtain a closed hypergeometric-function expression for the partial-wave projections, for all l, of the Coulomb transition matrix in nonrelativistic quantum mechanics.[1] The need for such a closed expression for the Coulomb T matrices has been the principal motivation for this investigation.

In Sec. 2 we shall state and prove three theorems. Theorem 1 just paves the way for the proof of Theorem 3, which constitutes the main result of this paper. In Sec. 3 we shall briefly consider some interesting particular cases.

2. THREE INFINITE SERIES

Theorem 1: Let $\alpha \in \mathbb{C}\setminus\mathbb{Z}$, $\beta \in \mathbb{R}$, $\varphi \in \mathbb{R}$, $\lambda \in \mathbb{C}$, $\lambda' \in \mathbb{C}$. Let either

(i) $\beta = 0, \quad |\varphi| < \pi,$

or

(ii) $\beta \neq 0, \quad |\beta| + |\varphi/\pi| \leqslant 1, \quad \mathrm{Re}(\lambda + \lambda') > -1,$

or

(iii) $\beta \neq 0, \quad |\beta| + |\varphi/\pi| < 1, \quad \mathrm{Re}(\lambda + \lambda') \geqslant -1.$

Then

$$\sum_{n=-\infty}^{\infty} \frac{(-1)^n}{\alpha - n} \frac{e^{in\varphi}}{\Gamma(\lambda + 1 - n\beta)\Gamma(\lambda' + 1 + n\beta)} = \frac{\pi e^{i\alpha\varphi} \operatorname{csec} \pi\alpha}{\Gamma(\lambda + 1 - \alpha\beta)\Gamma(\lambda' + 1 + \alpha\beta)}. \tag{1}$$

Proof: We define the function $f:(\mathbb{C}\setminus\mathbb{Z})\setminus\{\alpha\} \to \mathbb{C}$ by

$$f(a) := \frac{e^{ia\varphi}}{\alpha - a} \frac{\pi \operatorname{csec} \pi a}{\Gamma(\lambda + 1 - a\beta)\Gamma(\lambda' + 1 + a\beta)}. \tag{2}$$

Then f is a meromorphic function having simple poles in α and in all elements of $\mathbb{Z}$. Let C_n $(n = 1, 2, \cdots)$ be the circles with center 0 and radius $n + \epsilon$ such that the poles of f are avoided. By showing that

$$\lim_{n\to\infty} \int_{C_n} |f(a)| da = 0, \tag{3}$$

we derive

Σ (residues of f) $= 0$,

which proves Eq. (1).

Let us first consider case (i) $\beta = 0$. In this case Eq. (3) reduces to

$$\lim_{n\to\infty} \int_{C_n} \left| \frac{1}{\alpha - a} \frac{e^{ia\varphi}}{\sin a\pi} \right| da = 0 \quad (-\pi < \varphi < \pi),$$

which is easily verified.

In the cases (ii) and (iii) we may assume without real loss of generality that $\beta > 0$. In order to obtain the asymptotic behavior of $f(a)$, we rewrite Eq. (2) by using the equality

$$\Gamma(\lambda + 1 - a\beta)\Gamma(a\beta - \lambda) = \pi \operatorname{csec} \pi(a\beta - \lambda).$$

Then

$$f(a) = \frac{e^{ia\varphi}}{\alpha - a} \frac{\sin \pi(a\beta - \lambda)}{\sin \pi a} \frac{\Gamma(a\beta - \lambda)}{\Gamma(a\beta + \lambda' + 1)} = -\frac{e^{ia\varphi}}{\alpha - a} \frac{\sin \pi(a\beta + \lambda')}{\sin \pi a} \frac{\Gamma(-a\beta - \lambda')}{\Gamma(-a\beta + \lambda + 1)}.$$

If $|\arg a\beta| < \pi - \epsilon (\epsilon > 0)$, $\Gamma(a\beta - \lambda)/\Gamma(a\beta + \lambda' + 1)$ is represented asymptotically by $(a\beta)^{-\lambda - \lambda' - 1}$. If $|\arg(-a\beta)| < \pi - \epsilon$, $\Gamma(-a\beta - \lambda')/\Gamma(-a\beta + \lambda + 1)$ is e-represented asymptotically by $(-a\beta)^{-\lambda - \lambda' - 1}$.

When $\beta + |\varphi/\pi| = 1$, we have $|e^{ia\varphi} \sin \pi(a\beta - \lambda)/\sin \pi a| = O(1)$ on C_n, $n\to\infty$. Hence (3) holds if $\mathrm{Re}(\lambda + \lambda' + 1) > 0$.

When $\beta + |\varphi/\pi| < 1$ one easily verifies that

$$\lim_{n\to\infty} \int_{C_n} \left| \frac{e^{ia\varphi}}{\alpha - a} \frac{\sin \pi(a\beta - \lambda)}{\sin \pi a} \right| da = 0,$$

so that (3) holds if $\mathrm{Re}(\lambda + \lambda' + 1) \geqslant 0$.

Theorem 2: Let $\alpha \in \mathbb{C}\setminus\mathbb{Z}$,

$\beta_i \in \mathbb{R}, \lambda_i, \lambda_i' \in \mathbb{C} \quad (i = 1, 2, \ldots, m).$

Let either

$$\sum_{i=1}^{m} |\beta_i| \leqslant 1, \quad \sum_{i=1}^{m} \mathrm{Re}(\lambda_i + \lambda_i' + 1) > 0,$$

or

$$\sum_{i=1}^{m} |\beta_i| < 1, \quad \sum_{i=1}^{m} \mathrm{Re}(\lambda_i + \lambda_i' + 1) \geqslant 0.$$

Then

$$\sum_{n=-\infty}^{\infty} \frac{(-1)^n}{\alpha - n} \prod_{i=1}^{m} \frac{1}{\Gamma(\lambda_i + 1 - n\beta_i)\Gamma(\lambda_i' + 1 + n\beta_i)} = \pi \operatorname{csec} \pi\alpha \prod_{i=1}^{m} \frac{1}{\Gamma(\lambda_i + 1 - \alpha\beta_i)\Gamma(\lambda_i' + 1 + \alpha\beta_i)}.$$

Proof: The proof is similar to that of Theorem 1. We define

$$f(a) := \frac{\pi \csc \pi\alpha}{\alpha - a} \prod_{i=1}^{m} \frac{1}{\Gamma(\lambda_i + 1 - a\beta_i)\Gamma(\lambda'_i + 1 + a\beta_i)},$$

and we are going to prove Eq. (3). Without real loss of generality we may assume that $\beta_i \neq 0$ $(i = 1, 2, ..., m)$. We rearrange the β's so that

$$0 < \beta_j \leqslant 1, \quad j = 1,2,...,n;$$
$$-1 \leqslant \beta_k < 0, \quad k = n+1,...,m.$$

Then we can rewrite f as follows:

$$\begin{aligned} f(a) &= \pi(\alpha - a)^{-1}(\sin \pi a)^{-1} \\ &\times \prod_{j=1}^{n} \pi^{-1}\sin \pi(a\beta_j - \lambda_j) \\ &\times \Gamma(a\beta_j - \lambda_j)[\Gamma(a\beta_j + \lambda'_j + 1)]^{-1} \\ &\times \prod_{k=n+1}^{m} \pi^{-1}\sin \pi(-a\beta_k - \lambda'_k) \\ &\times \Gamma(-a\beta_k - \lambda'_k)[\Gamma(-a\beta_k + \lambda_k + 1)]^{-1}. \end{aligned}$$

From this expression it can be seen that if $|\arg a| < \pi - \epsilon$ $(\epsilon > 0)$, $f(a)$ is represented asymptotically by

$$\begin{aligned} f(a) &\sim \pi^{1-m}(\alpha - a)^{-1}(\sin \pi a)^{-1} \\ &\times \prod_{j,k} \sin \pi(a\beta_j - \lambda_j) \sin \pi(-a\beta_k - \lambda'_k) \\ &\times (a\beta_j)^{-\lambda_j - \lambda'_j - 1}(-a\beta_k)^{-\lambda_k - \lambda'_k - 1}. \end{aligned}$$

If $\sum_i |\beta_i| \equiv \sum_j \beta_j - \sum_k \beta_k = 1$, we have

$$\Big|\csc \pi a \prod_{j,k} \sin \pi(a\beta_j - \lambda_j) \sin \pi(-a\beta_k - \lambda'_k)\Big| = O(1), \quad a \in C_n, \quad n \to \infty.$$

If $\sum_j \beta_j - \sum_k \beta_k < 1$, it follows that

$$\lim_{n\to\infty} \int_{C_n} \left| \frac{\csc \pi a}{\alpha - a} \prod_{j,k} \sin \pi(a\beta_j - \lambda_j) \times \sin \pi(-a\beta_k - \lambda'_k) \right| = 0.$$

Hence

$$\lim_{n\to\infty} \int_{C_n} |f(a)|\, da = 0$$

if either

$$\sum_i |\beta_i| \leqslant 1, \quad \sum_i \mathrm{Re}(\lambda_i + \lambda'_i + 1) > 0$$

or

$$\sum_i |\beta_i| < 1, \quad \sum_i \mathrm{Re}(\lambda_i + \lambda'_i + 1) \geqslant 0.$$

Definition: Let μ, ν, $z \in \mathbb{C}$. We define

$$\mathfrak{p}_\nu^\mu(z) := \frac{1}{\Gamma(1-\mu)}\, {}_2F_1(-\nu, 1+\nu; 1-\mu; \tfrac{1}{2} - \tfrac{1}{2}z), \tag{4}$$

where the function of the right-hand side is understood to be the analytic continuation with respect to μ in the positive integers. Then $\mathfrak{p}_\nu^\mu(z)$ is, for fixed ν and z, an entire analytic function of μ (cf. Refs. 2–4).

When ν is an integer, $\mathfrak{p}_\nu^\mu(z)$ is a polynomial in z. When $\nu \in \mathbb{C} \setminus \mathbf{Z}$, $\mathfrak{p}_\nu^\mu(z)$ has a branch cut: $z \in (-\infty, -1]$.

Theorem 3: Let $\alpha \in \mathbb{C} \setminus \mathbf{Z}$; $\beta \in \mathbb{R}$; $\lambda, \lambda', \nu, \nu', z, z' \in \mathbb{C}$. If

(i) $\mathrm{Re}\, z > 0$,
(ii) $\mathrm{Re}\, z' > 0$,

and

(iii) either

$$-1 \leqslant \beta \leqslant 1, \quad \mathrm{Re}(\lambda + \lambda') > -1,$$

or

$$-1 < \beta < 1, \quad \mathrm{Re}(\lambda + \lambda') \geqslant -1;$$

then

$$\sum_{n=-\infty}^{\infty} (-1)^n(\alpha - n)^{-1}\mathfrak{p}_\nu^{n\beta - \lambda}(z)\, \mathfrak{p}_{\nu'}^{-n\beta - \lambda'}(z') = \pi \csc(\pi\alpha)\mathfrak{p}_\nu^{\alpha\beta - \lambda}(z)\, \mathfrak{p}_{\nu'}^{-\alpha\beta - \lambda'}(z'). \tag{5}$$

When $\sin \pi\nu = 0$, condition (i) may be omitted. Similarly, when $\sin \pi\nu' = 0$, (ii) may be omitted.

Proof: We shall first prove the theorem under the additional conditions:

$$|1 + z| < 2, \quad |1 + z'| < 2.$$

These conditions can be dropped afterwards by means of analytic continuation. Indeed, in the Appendix we shall prove that the series in Theorem 3 is uniformly convergent when $\mathrm{Re}\, z > 0$, $\mathrm{Re}\, z' > 0$.

When $\sin \pi\nu = 0$ or $\sin \pi\nu' = 0$ the proof is just a simple special case of the general proof of the theorem.

In analogy with the proof of Theorem 1 we define

$$f(a) := \pi(\alpha - a)^{-1}\csc \pi a\, \mathfrak{p}_\nu^{a\beta - \lambda}(z)\mathfrak{p}_{\nu'}^{-a\beta - \lambda'}(z').$$

Then we have to prove that Eq. (3) holds. With this aim we are going to derive a suitable asymptotic estimate of $f(a)$ for $|a| \to \infty$. We shall assume that $\beta > 0$, which means no real loss of generality.

It is well known that

$$\lim_{|c|\to\infty} {}_2F_1(a,b;c;z) = 1,$$

provided that $|z| < 1$ and $|\arg c| < \pi - \epsilon$ $(\epsilon > 0)$. In view of the definition of $\mathfrak{p}_\nu^\mu$ we have

$$\begin{aligned} f(a) &= \frac{\pi \csc(\pi a)(\alpha - a)^{-1}}{\Gamma(\lambda + 1 - a\beta)\Gamma(\lambda' + 1 + a\beta)} \\ &\times {}_2F_1(-\nu, 1+\nu; 1 - a\beta + \lambda; \tfrac{1}{2} - \tfrac{1}{2}z) \\ &\times {}_2F_1(-\nu', 1+\nu'; 1 + a\beta + \lambda'; \tfrac{1}{2} - \tfrac{1}{2}z'). \end{aligned}$$

We rewrite the first hypergeometric function by using formula (17) of Ref. 2, p. 141. Substituting

$$\pi/\Gamma(\lambda + 1 - a\beta) = \Gamma(a\beta - \lambda) \sin \pi(a\beta - \lambda),$$

we obtain

$$\begin{aligned} f(a) &= (\alpha - a)^{-1}\, {}_2F_1(-\nu', 1+\nu'; 1 + a\beta + \lambda'; \tfrac{1}{2} - \tfrac{1}{2}z') \\ &\times \frac{\Gamma(a\beta - \lambda - \nu)}{\Gamma(a\beta + \lambda' + 1)} \frac{\Gamma(a\beta + 1 - \lambda + \nu)}{\Gamma(a\beta + 1 - \lambda)} \\ &\times \frac{\sin \pi(a\beta - \lambda)}{\sin \pi a} \\ &\times [{}_2F_1(-\nu, 1+\nu; a\beta + 1 - \lambda; \tfrac{1}{2} + \tfrac{1}{2}z) \\ &- \frac{\sin \pi\nu}{\sin \pi(a\beta - \lambda)} \left(\frac{1-z}{1+z}\right)^{a\beta - \lambda} \\ &\times {}_2F_1(-\nu, 1+\nu; a\beta + 1 - \lambda; \tfrac{1}{2} - \tfrac{1}{2}z)]. \end{aligned}$$

By observing that

$$\left|\frac{1-z}{1+z}\right|<1, \quad |\tfrac12\pm\tfrac12 z|<1, \quad |\tfrac12\pm\tfrac12 z'|<1,$$

one easily obtains a suitable asymptotic representation of $f(a)$ for $|a|\to\infty$, $|\arg a|<\pi-\epsilon(\epsilon>0)$

$$f(a)\sim(\alpha-a)^{-1}(a\beta)^{-\lambda-\lambda'-1} \times\left[\frac{\sin\pi(a\beta-\lambda)}{\sin\pi a}-\frac{\sin\pi\nu}{\sin\pi a}\left(\frac{1-z}{1+z}\right)^{a\beta-\lambda}\right].$$

The expression in square brackets is bounded for a on the circles C_n, $n\to\infty$, provided $|\arg a|\leqslant\pi/2$. Because of the symmetry ($\lambda\leftrightarrow\lambda'$; $z\leftrightarrow z'$; $a\leftrightarrow -a$; $\alpha\leftrightarrow-\alpha$), a similar fact holds *mutatis mutandis* when $|\arg(-a)|\leqslant\pi/2$. It follows that, for some M,

$$|f(a)|<M|\alpha-a|^{-1}|a|^{-\lambda-\lambda'-1}, \quad a\in C_n.$$

Clearly Eq. (3) holds if $\mathrm{Re}(\lambda+\lambda'+1)>0$.

When the *second* condition of (iii) holds, the proof of the theorem follows by verifying that

$$\lim_{n\to\infty}\int_{C_n}|\alpha-a|^{-1} \times\left|\frac{\sin\pi(a\beta-\lambda)}{\sin\pi a}-\frac{\sin\pi\nu}{\sin\pi a}\left(\frac{1-z}{1+z}\right)^{a\beta-\lambda}\right|da=0$$

when $-1<\beta<1$.

This completes the proof of Theorem 3, except for the restrictions on z and z': $|1+z|<2$, $|1+z'|<2$. The series in Theorem 3 is uniformly convergent with respect to z and z' if $\mathrm{Re}\,z>0$, $\mathrm{Re}\,z'>0$. Since its terms are analytic functions of z and of z', as is its sum, it follows by analytic continuation that the aforementioned restrictions on z and z' can be removed. The proof of the uniform convergence will be given in the Appendix.

3. SOME SPECIAL CASES

In this section we shall consider some special cases of Theorems 1 and 3.

By taking $\lambda=\lambda'=0$ in Theorem 1 we obtain

$$\sum_{n=-\infty}^{\infty}\frac{(-1)^n e^{in\varphi}}{\alpha-n}\frac{\sin\pi n\beta}{\pi n\beta}=\frac{\pi e^{i\alpha\varphi}}{\sin\pi\alpha}\frac{\sin\pi\alpha\beta}{\pi\alpha\beta}, \tag{6}$$

where $\sin x/x$ is understood to be 1 when $x=0$. Clearly, when $\beta=\pm1$ Eq. (6) is valid for $\varphi=0$ only, in conformity with condition (ii) of the theorem.

Taking $\lambda=\lambda'=-\frac12$ we get (using $\Gamma(\frac12+z)\Gamma(\frac12-z)=\pi\sec\pi z$)

$$\sum_{n=-\infty}^{\infty}\frac{(-1)^n}{\alpha-n}e^{in\varphi}\cos\pi n\beta=\frac{\pi e^{i\alpha\varphi}}{\sin\pi\alpha}\cos\pi\alpha\beta, \tag{7}$$

which holds for $|\beta|+|\varphi/\pi|<1$.

By taking, in Theorem 1, $\varphi=0$ and $\alpha\to(\lambda+1)/\beta$, we obtain

$$\sum_{n=-\infty}^{\infty}\frac{(-1)^n\beta}{\Gamma(\lambda+2-n\beta)\Gamma(\lambda'+1+n\beta)}=0, \tag{8}$$

provided that either $-1\leqslant\beta\leqslant1$, $\mathrm{Re}(\lambda+\lambda')>-1$ or $-1<\beta<1$, $\mathrm{Re}(\lambda+\lambda')\geqslant-1$. Note that the factor β in the numerator is necessary for the case $\beta=0$ only.

Now we shall consider some particular cases of Theorem 3. According to the definition of $\mathfrak{p}_\nu^\mu$ [Eq. (4)], we have

$$\mathfrak{p}_\nu^\mu(z)=(z+1)^{-\mu/2}(z-1)^{\mu/2}\mathfrak{P}_\nu^\mu(z),$$

where $\mathfrak{P}_\nu^\mu$ is Legendre's function of the first kind. Hence,

$$\mathfrak{p}_\nu^\mu(z)\mathfrak{p}_\nu^{-\mu}(z)=\mathfrak{P}_\nu^\mu(z)\mathfrak{P}_\nu^{-\mu}(z).$$

By substituting this into Eq. (5) with $\lambda=\lambda'=0$ and $z=z'$, we get

$$\sum_{n=-\infty}^{\infty}\frac{(-1)^n}{\alpha-n}\mathfrak{P}_\nu^{n\beta}(z)\mathfrak{P}_{\nu'}^{-n\beta}(z) =\pi\csc(\pi\alpha)\mathfrak{P}_\nu^{\alpha\beta}(z)\mathfrak{P}_{\nu'}^{-\alpha\beta}(z). \tag{9a}$$

In the particular case $\nu'=\nu$ this can be rewritten as

$$\sum_{n=0}^{\infty}\frac{(-1)^n\alpha\epsilon_n}{\alpha^2-n^2}\mathfrak{P}_\nu^{-n\beta}(z)\mathfrak{P}_\nu^{n\beta}(z) =\pi\csc(\pi\alpha)\mathfrak{P}_\nu^{\alpha\beta}(z)\mathfrak{P}_\nu^{-\alpha\beta}(z), \tag{9b}$$

which is a well-known relation, see, e.g., Refs. 5–8. Here ϵ_n is defined by $\epsilon_0=1$, $\epsilon_n=2$ $(n=1,2,\cdots)$.

Finally we shall consider Eq. (5) (Theorem 3) in the particular case when $\lambda=\lambda'=0$, $\beta=1$, and $\nu=\nu'=l\in\mathbb{N}$. Then we have

$$\sum_{n=-\infty}^{\infty}\frac{(-1)^n}{\alpha-n}\mathfrak{p}_l^n(z)\mathfrak{p}_l^{-n}(z')=\pi\csc(\pi\alpha)\mathfrak{p}_l^\alpha(z)\mathfrak{p}_l^{-\alpha}(z'), \tag{10a}$$

which holds for all complex z and z'. Since l is a nonnegative integer, $\mathfrak{p}_l^\alpha$ can be expressed in terms of the Jacobi polynomial $P_l^{(-\alpha,\alpha)}$:

$$\mathfrak{p}_l^\alpha(z)=P_l^{(-\alpha,\alpha)}(z)\Gamma(l+1)/\Gamma(l-\alpha+1).$$

The infinite sum in Eq. (10a) reduces to a finite sum. We get

$$\sum_{n=-l}^{l}\frac{(-1)^n}{\alpha-n}\frac{1}{(l-n)!(l+n)!}P_l^{(-n,n)}(z)\,P_l^{(n,-n)}(z') =\frac{\pi\csc\pi\alpha}{\Gamma(l-\alpha+1)\Gamma(l+\alpha+1)}P_l^{(-\alpha,\alpha)}(z)P_l^{(\alpha,-\alpha)}(z'), \tag{10b}$$

which holds for $z\in\mathbb{C}$, $z'\in\mathbb{C}$, $\alpha\in\mathbb{C}\setminus\mathbb{Z}$, $l\in\mathbb{N}$.

We have used Eq. (10b) to obtain a closed expression, in terms of hypergeometric functions and elementary functions, of the partial-wave projected Coulomb transition operator in quantum mechanics.[1]

Finally, we note that T. J. Osler[9] has evaluated a number of sums similar to those in this paper.

ACKNOWLEDGMENT

The author is grateful to R. Askey for the reference to T. J. Osler's work.

APPENDIX

In this appendix we shall prove that the series [cf. Eq. (5)]

$$S:=\sum_{n=-\infty}^{\infty}(-1)^n(\alpha-n)^{-1}\mathfrak{p}_\nu^{n\beta-\lambda}(z)\mathfrak{p}_{\nu'}^{-n\beta-\lambda'}(z')$$

is uniformly convergent with respect to z and z' for $\mathrm{Re}\,z>0$,

$\operatorname{Re} z' > 0$, under the conditions of Theorem 3.

By Eq. (4) we have

$$\mathfrak{p}_\nu^\mu(z) = [\Gamma(1-\mu)]^{-1}\,{}_2F_1(-\nu, 1+\nu; 1-\mu; \tfrac{1}{2} - \tfrac{1}{2}z).$$

We rewrite the right member by using,[2] p. 141, Eq. (17):

$$\begin{aligned}\mathfrak{p}_\nu^\mu(z) = {} & \frac{\Gamma(\nu+\mu+1)}{\pi\Gamma(\mu+1)} \\ & \times\Gamma(\mu-\nu)\sin\pi\mu\Big[{}_2F_1(-\nu,1+\nu;1+\mu;\tfrac{1}{2}+\tfrac{1}{2}z) \\ & -\frac{\sin\pi\nu}{\sin\pi\mu}\Big(\frac{1-z}{1+z}\Big)^{\mu} \\ & \times{}_2F_1(-\nu,1+\nu;1+\mu;\tfrac{1}{2}-\tfrac{1}{2}z)\Big].\end{aligned} \tag{A1}$$

We assume first that $0 < \beta \leqslant 1$, and we consider the convergence of S_+, where

$$S_+ = \sum_{n=0}^{\infty}\cdots;\quad S_- = \sum_{n=-\infty}^{-1}\cdots;\quad S = S_+ + S_-.$$

The discussion for the case $-1 \leqslant \beta < 0$ and for S_- is very similar to the present one.

By using Eq. (A1) we have (we assume here $z \notin [1,\infty)$)

$$\begin{aligned} S_+ = {} & \sum_{n=0}^{\infty}(-1)^n\frac{\pi^{-1}}{\alpha-n} \\ & \times{}_2F_1(-\nu',1+\nu';1+n\beta+\lambda';\tfrac{1}{2}-\tfrac{1}{2}z') \\ & \times\frac{\Gamma(\nu+n\beta-\lambda+1)\Gamma(n\beta-\nu-\lambda)}{\Gamma(n\beta+\lambda'+1)\Gamma(n\beta-\lambda+1)} \\ & \times[\sin\pi(n\beta-\lambda) \\ & \times{}_2F_1(-\nu,1+\nu;1+n\beta-\lambda;\tfrac{1}{2}+\tfrac{1}{2}z) \\ & -\sin\pi\nu\Big(\frac{1-z}{1+z}\Big)^{n\beta-\lambda} \\ & \times{}_2F_1(-\nu,1+\nu;1+n\beta-\lambda;\tfrac{1}{2}-\tfrac{1}{2}z)].\end{aligned} \tag{A2}$$

It is well known that for $z\in\mathbb{C}\setminus[1,\infty)$

$${}_2F_1(a,b;c;z) = 1 + abc^{-1}z + \cdots,\quad |c|\to\infty,$$

is an asymptotic expansion provided $\operatorname{Re} c\to\infty$. By using this expansion and

$$\Gamma(z+\alpha)/\Gamma(z+\beta) = z^{\alpha-\beta}[1+O(z^{-1})],\quad |z|\to\infty,\quad |\arg z|<\pi,$$

it follows from (A2) that we have to prove that

$$\sum_{n=1}^{\infty}(-1)^n n^{-2-\lambda-\lambda'} \times\Big[\sin\pi(n\beta-\lambda) - \sin\pi\nu\Big(\frac{1-z}{1+z}\Big)^{n\beta-\lambda}\Big]$$

is uniformly convergent with respect to z.

The first term in square brackets is z-independent. According to Theorem 1, the series of these terms is convergent, and hence uniformly convergent with respect to z.

When $\sin\pi\nu = 0$, the second term vanishes for all z. Suppose now $\sin\pi\nu\neq 0$; we have to prove that

$$\sum_{n=1}^{\infty}(-1)^n n^{-2-\lambda-\lambda'}\Big(\frac{1-z}{1+z}\Big)^{n\beta}$$

is uniformly convergent. From $\operatorname{Re} z>0$ we have $|(1-z)/(1+z)|<1$. It follows that the sum is absolutely convergent and hence uniformly convergent for all δ satisfying

$$\Big|\frac{1-z}{1+z}\Big| \leqslant \delta < 1.$$

This completes the proof of the uniform convergence of S_+ for $\beta > 0$. By interchanging z and z', λ and λ', and ν and ν' it easily follows that S_- is uniformly convergent for $\operatorname{Re} z' > 0$. Finally, the case $\beta < 0$ is obtained by using the new variable of summation $m = -n$.

[1] H. van Haeringen, "The partial-wave projected Coulomb T matrix for all l in closed hypergeometric form," preprint, Delft University of Technology, 1981, in particular Eq. (5.8).

[2] A. Erdélyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.

[3] N. N. Lebedev, *Special Functions and Their Applications* (Dover, New York, 1972).

[4] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products, Corrected and Enlarged Edition* (Academic, New York, 1980).

[5] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966), p. 180.

[6] E. R. Hansen, *A Table of Series and Products* (Prentice-Hall, Englewood Cliffs, N. J., 1975), p. 374.

[7] Unfortunately Refs. 4 and 5, and (cf. Ref. 2) the five volumes of *Bateman's Manuscript Project*, edited by A. Erdélyi: (i) *Higher Transcendental Functions* (McGraw-Hill, New York), Vols. I and II (1953), Vol. III (1955), and (ii) *Tables of Integral Transforms* (McGraw-Hill, New York), Vols. I and II (1954), contain misprints and errors. More than six hundred of these have been collected in Ref. 8.

[8] H. van Haeringen and L. P. Kok, Corrigenda Reports 8201–8204, Delft, Groningen, 1982.

[9] T. J. Osler, SIAM J. Math. Anal. **3**, 1 (1972), especially p. 12. See also T. J. Osler, SIAM J. Appl. Math. **8**, 658 (1970), and J. L. Lavoie, T. J. Osler, and R. Tremblay, SIAM Review **18**, 240 (1976).

A simple proof of a transformation formula for elliptic integrals

B. Bassetti, E. Montaldi, and M. Raciti
Istituto di Scienze Fisiche dell'Università degli Studi di Milano, Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy

(Received 3 August 1982; accepted for publication 3 December 1982)

A very simple proof of a quadratic transformation formula for elliptic integrals found by Carlson in 1977 is given.

PACS numbers: 02.30. – f

Elliptic integrals which occur in many physical problems have been the subject of interesting investigations in the last years. We refer, in particular, to some remarkable papers of Carlson[1-3] concerning both the numerical evaluation of the integrals and their transformation properties. In this context, the formula[2]

$$\int_0^\infty dx[(x+a^2)(x+b^2)(x+c^2)(x+d^2)]^{-1/2}$$
$$=\int_0^\infty dx\{[x+(ab+cd)^2][x+(ac+bd)^2]\times[x+(ad+bc)^2]\}^{-1/2}, \qquad (1)$$

which is a quadratic transformation from the quartic to the cubic case, is of fundamental importance.

The purpose of the present paper is to give a short derivation of Eq. (1) based on a trick which gives a very simple answer to the problem, proposed by Carlson[2] and solved by himself,[3] where Eq. (1) is obtained by a change of variable. At the same time, the role played by the addition theorem emerges in a clear way.

We can put, without loss of generality, $d=1$ in Eq. (1). Let's first consider the right-hand side; by introducing the new variable $t=(a+bc)[x+(a+bc)^2]^{-1/2}$, we readily obtain

$$\int_0^\infty dx\{[x+(c+ab)^2][x+(b+ac)^2][x+(a+bc)^2]\}^{-1/2}$$
$$=\frac{2}{a+bc}\int_0^1 dt\left\{\left[1-\frac{(a^2-b^2)(1-c^2)}{(a+bc)^2}t^2\right]\left[1-\frac{(a^2-c^2)(1-b^2)}{(a+bc)^2}t^2\right]\right\}^{-1/2}$$
$$=\frac{2}{a+bc}F_1\left(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{3}{2};\ \frac{(a^2-b^2)(1-c^2)}{(a+bc)^2},\frac{(a^2-c^2)(1-b^2)}{(a+bc)^2}\right), \qquad (2)$$

where F_1 is the first Appell's hypergeometric series in two variables.[4]

On the other hand, the left-hand side of Eq. (1), setting $t=[(x+a^2)/(x+1)]^{1/2}$, becomes

$$\int_0^\infty dx[(x+1)(x+a^2)(x+b^2)(x+c^2)]^{-1/2}$$
$$=2[(b^2-a^2)(c^2-a^2)]^{-1/2}\int_a^1 dt\left[\left(1-\frac{b^2-1}{b^2-a^2}t^2\right)\times\left(1-\frac{c^2-1}{c^2-a^2}t^2\right)\right]^{-1/2}. \qquad (3)$$

Without loss of generality, we take $0<a<1$ and $b, c>a$. The case $b, c>a>1$ is easily reduced to the previous one, performing the change of variable $x=a^2y$.

Now we consider the problem of finding a function $\tau=\tau(a)$ such that

$$\int_a^1 dt\,[(1-ut^2)(1-vt^2)]^{-1/2}$$
$$=\int_0^{\tau(a)} d\tau[(1-u\tau^2)(1-v\tau^2)]^{-1/2}, \qquad (4)$$

u and v being arbitrary parameters; this just amounts to determining a suitable change of variable.

Equation (4) can be rewritten as

$$\int_0^a dx[(1-ux^2)(1-vx^2)]^{-1/2}$$
$$+\int_0^\tau dx[(1-ux^2)(1-vx^2)]^{-1/2}$$
$$=\int_0^1 dx[(1-ux^2)(1-vx^2)]^{-1/2}.$$

Recalling that[5]

$$z=\int_0^{\mathrm{sn}(z,k)} dx[(1-x^2)(1-k^2x^2)]^{-1/2}$$

and using the addition formula for the Jacobian elliptic function $\mathrm{sn}(z,k)$, we have

$$\tau(a)=\frac{[(1-a^2u)(1-a^2v)]^{1/2}-a[(1-u)(1-v)]^{1/2}}{1-a^2uv}. \qquad (5)$$

Since [cf. Eq. (2)] the right-hand side of Eq. (4) can be expressed in terms of the Appell's function F_1, we obtain

$$\tau(a)F_1(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2};\tau^2(a)u,\tau^2(a)v)$$
$$=\int_a^1 dt\,[(1-ut^2)(1-vt^2)]^{-1/2}. \qquad (6)$$

If we now choose

 0022-2488/83/051058-02$02.50

$$u=\frac{b^2-1}{b^2-a^2}, \quad v=\frac{c^2-1}{c^2-a^2},$$

we immediately recognize that the right-hand sides of Eqs. (2) and (3) are equal and thus formula (1) is proved. We also remark that some results existing in the literature are nothing but particular cases of (4). As a specific example, we quote the formula[6]

$$\int_{-1}^{1} dx\left[\left(1-\frac{2\alpha}{1+\alpha}x^2\right)\left(1-\frac{1+\alpha}{2}x^2\right)\right]^{-1/2} = [2(1+\alpha)]^{1/2}K(\alpha), \tag{7}$$

K being the complete elliptic integral of the first kind. To obtain (7) it suffices to take $a=-1$, $u=2\alpha/(1+\alpha)$, and $v=(1+\alpha)/2$ in (4); this way the left-hand side of (7) becomes equal to $(2/(1+\alpha))^{1/2}K(2\sqrt{\alpha}/(1+\alpha))$, and the desired result follows at once by using a standard quadratic transformation of the hypergeometric function.[7]

We thank D. Zanon for a critical reading of the manuscript.

[1]B. C. Carlson, Duke Math. J. **31**, 405 (1964).
[2]B. C. Carlson, SIAM J. Math. Anal. **8**, 231 (1977).
[3]B. C. Carlson, SIAM J. Math. Anal. **9**, 524 (1978).
[4]A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. I, p. 224, Eq. (6) and p. 231, Eq. (5).
[5]E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University, Cambridge, 1958), pp. 492, 496.
[6]E. Montaldi, Lett. Nuovo Cimento **30**, 403 (1981).
[7]Reference 4, p. 111, Eq. (5).

Expression for $y_{LM}[(\mathbf{r}_1\wedge\mathbf{r}_2)\wedge\mathbf{r}_3]$

T. A. Antone
Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria

(Received 27 May 1981; accepted for publication 3 December 1982)

The solid harmonic $y_{LM}[(\mathbf{r}_1\wedge\mathbf{r}_2)\wedge\mathbf{r}_3]$ is expressed in terms of the spherical harmonics $Y_{L_1M_1}(\hat{\mathbf{r}}_1)$, $Y_{L_2M_2}(\hat{\mathbf{r}}_2)$, and $Y_{L_3M_3}(\hat{\mathbf{r}}_3)$. The calculation of the coefficients in the given expansion in terms of $9-j$ symbols explicitly justifies the form given by Eq. (1).

PACS numbers: 02.30.Gp, 03.65. — w

I. INTRODUCTION

We express the solid harmonic in the form

$$y_{LM}[(\mathbf{r}_1\wedge\mathbf{r}_2)\wedge\mathbf{r}_3]=\sum_{L_1L_2L_{12}L_3LM_1M_2}A(L_1,L_2,L_{12},L_3,L)\langle L_1,M_1;L_2,M_2|L_1L_2L_{12}M_{12}\rangle \times\langle L_{12},M_{12};L_3,M_3|L_{12}L_3LM\rangle Y_{L_1M_1}(\hat{\mathbf{r}}_1)Y_{L_2M_2}(\hat{\mathbf{r}}_2)Y_{L_3M_3}(\hat{\mathbf{r}}_3), \tag{1}$$

where the coefficients $A(L_1,L_2,L_{12},L_3,L)$ are independent of the quantum numbers M_1, M_2, M_3 as well as the angles involved. We find that these coefficients are expressed in terms of the $9-j$ symbols as

$$A(L_1,L_2,L_{12},L_3,L)=(-1)^{L_3}\left(\frac{2L_{12}+1}{2L+1}\right)^{1/2}\sum_{stj}(-1)^jK(s,t,j,L_1,L_2,L_3,L)\begin{Bmatrix}L_1 & L_2 & L_{12}\\ s & t & L_3\\ L-j & j & L\end{Bmatrix}, \tag{2}$$

where j, s, and t take the values

$$j=L,L-1,L-2,\ldots,0,\quad s=j,j-2,j-4,\ldots,1 \text{ or } 0,\quad t=L-j,L-j-2,L-j-4,\ldots,1 \text{ or } 0,$$

and

$$K(s,t,j,L_1,L_2,L_3,L)=4\pi(r_1r_2r_3)^L\left(\frac{(2L_1+1)(2L_2+1)(2L_3+1)(2L)!}{(2j)!(2L-2j)!}\right)^{1/2} \times\frac{2^{(1/2)(s+t-L)}(2L+1)(2s+1)(2t+1)j!(L-j)!}{[\frac12(j-s)]![\frac12(L-j-t)]!(j+s+1)!!(L-j+t+1)!!}\begin{pmatrix}s & L-j & L_1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}t & j & L_2\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}t & s & L_3\\ 0 & 0 & 0\end{pmatrix}. \tag{3}$$

II. CALCULATION OF THE COEFFICIENTS $A(L_1,L_2,L_{12},L_3,L)$

The solid harmonic $y_{kq}(\mathbf{a}+\mathbf{b})$ is expressed by Eq. (12.41) in Talman[1] in terms of the spherical harmonics $Y_{pq_1}(\hat{\mathbf{a}})$ and $Y_{k-p,q-q_1}(\hat{\mathbf{b}})$ as

$$y_{kq}(\mathbf{a}+\mathbf{b})=\sum_{pq_1}(-1)^{k+q}(2k+1)a^pb^{k-p}\times\left(\frac{4\pi(2k)!}{(2p+1)!(2k-2p+1)!}\right)^{1/2}\times\begin{pmatrix}p & k-p & k\\ q_1 & q-q_1 & -q\end{pmatrix}Y_{pq_1}(\hat{\mathbf{a}})Y_{k-p,q-q_1}(\hat{\mathbf{b}}), \tag{4}$$

where in Edmonds[2] we have used Eq. (3.7.3) relating the Clebsch–Gordan coefficients to the $3-j$ symbols in the form

$$\langle j_1,m_1;j_2m_2|j_1j_2j_3m_3\rangle=(-1)^{j_1-j_2+m_3}(2j_3+1)^{1/2}\begin{pmatrix}j_1 & j_2 & j_3\\ m_1 & m_2 & -m_3\end{pmatrix} \tag{5}$$

and Eq. (3.7.17) in the form

$$\begin{pmatrix}p & k-p & k\\ 0 & 0 & 0\end{pmatrix}=(-1)^k\frac{k!}{p!(k-p)!}\left(\frac{(2k-2p)!(2p)!}{(2k+1)!}\right)^{1/2}. \tag{6}$$

We now write our solid harmonic of Eq. (1) as

$$y_{LM}[(\mathbf{r}_1\wedge\mathbf{r}_2)\wedge\mathbf{r}_3]=y_{LM}[(\mathbf{r}_1\cdot\mathbf{r}_3)\mathbf{r}_2-(\mathbf{r}_3\cdot\mathbf{r}_2)\mathbf{r}_1]=Z_{LM}$$

and use Eq. (4) to write it as

$$Z_{LM}=\sum_{jm}(-1)^{j+M}(2L+1)|(\mathbf{r}_1\cdot\mathbf{r}_3)\mathbf{r}_2|^j|(\mathbf{r}_3\cdot\mathbf{r}_2)\mathbf{r}_1|^{L-j}\times\left(\frac{4\pi(2L)!}{(2j+1)!(2L-2j+1)!}\right)^{1/2}\times\begin{pmatrix}j & L-j & L\\ m & M-m & -M\end{pmatrix}\times Y_{jm}(\hat{\mathbf{r}}_2)Y_{L-j,M-m}(\hat{\mathbf{r}}_1). \tag{7}$$

But

 0022-2488/83/051060-02$02.50

$$|(\mathbf{r}_1\cdot\mathbf{r}_3)\mathbf{r}_2|^j = (r_1r_2r_3)^j\cos^j\theta_{13} = 4\pi(r_1r_2r_3)^j\sum_{su}\frac{2^{(1/2)(s-j)}j\,!}{[\frac{1}{2}(j-s)]!(j+s+1)!!}Y^*_{su}(\hat{\mathbf{r}}_1)Y_{su}(\hat{\mathbf{r}}_3),$$

where $s = j, j-2, j-4,...,1$ or 0 and $|u|\leqslant s$.

With a similar expression for $|(\mathbf{r}_3\cdot\mathbf{r}_2)\mathbf{r}_1|^{L-j}$, we write Eq. (7) as

$$\begin{aligned} Z_{LM} = {} & 16\pi^2(2L+1)(r_1r_2r_3)^L\sum_{jmstuv}(-1)^{j+M}\left(\frac{4\pi(2L)!}{(2j+1)!(2L-2j+1)!}\right)^{1/2} \\ & \times\frac{2^{(1/2)(s+t-L)}j\,!(L-j)!}{[\frac{1}{2}(j-s)]![\frac{1}{2}(L-j-t)]!(j+s+1)!!(L-j+t+1)!!}\begin{pmatrix} j & L-j & L \\ m & M-m & -M\end{pmatrix} \\ & \times Y^*_{su}(\hat{\mathbf{r}}_1)Y_{L-j,M-m}(\hat{\mathbf{r}}_1)Y_{tv}(\hat{\mathbf{r}}_2)Y_{jm}(\hat{\mathbf{r}}_2)Y^*_{tv}(\hat{\mathbf{r}}_3)Y_{su}(\hat{\mathbf{r}}_3),\end{aligned} \tag{8}$$

where $t = L-j, L-j-2, L-j-4,...,1$ or 0 and $|v|\leqslant t$.

Using Eq. (4.6.5) in Edmonds[2] in the form

$$\begin{aligned} Y_{j_1m_1}(\hat{\mathbf{r}})Y_{j_2m_2}(\hat{\mathbf{r}}) = {} & \sum_{j_3m_3}(-1)^{m_3}\left(\frac{(2j_1+1)(2j_2+1)(2j_3+1)}{4\pi}\right)^{1/2} \\ & \times\begin{pmatrix} j_1 & j_2 & j_3 \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3\end{pmatrix}Y_{j_3m_3}(\hat{\mathbf{r}}),\end{aligned} \tag{9}$$

we write Eq. (8) as

$$\begin{aligned} Z_{LM} = {} & \sum_{stuvjmL_1L_2L_3M_1M_2}(-1)^{j+u+v}K(s,t,j,L_1,L_2,L_3,L)\begin{pmatrix} s & L-j & L_1 \\ -u & M-m & -M_1\end{pmatrix} \\ & \times\begin{pmatrix} t & j & L_2 \\ v & m & -M_2\end{pmatrix}\begin{pmatrix} t & s & L_3 \\ -v & u & -M_3\end{pmatrix}\begin{pmatrix} j & L-j & L \\ m & M-m & -M\end{pmatrix} \\ & \times Y_{L_1M_1}(\hat{\mathbf{r}}_1)Y_{L_2M_2}(\hat{\mathbf{r}}_2)Y_{L_3M_3}(\hat{\mathbf{r}}_3),\end{aligned} \tag{10}$$

where $K(s,t,j,L_1,L_2,L_3,L)$ is given by Eq. (3). Equation (10) shows that the coefficients of the spherical harmonics vanish unless $u = M-m-M_1$, $v = M_2-m$ and the values $(s+L-j+L_1)$, $(t+j+L_2)$, and $(t+s+L_3)$ are even integers.

Using a relation between the $3-j$ and $9-j$ symbols given by de-Shalit[3] in the form

$$\begin{aligned} & \sum_{uvm}\begin{pmatrix} s & L-j & L_1 \\ -u & M-m & -M_1\end{pmatrix}\begin{pmatrix} t & j & L_2 \\ v & m & -M_2\end{pmatrix}\begin{pmatrix} t & s & L_3 \\ v & -u & M_3\end{pmatrix}\begin{pmatrix} j & L-j & L \\ m & M-m & -M\end{pmatrix} \\ & = \sum_{L_{12}M_{12}}(-1)^{L_1+L_2+L_{12}}(2L_{12}+1)\begin{pmatrix} L_1 & L_2 & L_{12} \\ M_1 & M_2 & -M_{12}\end{pmatrix}\begin{pmatrix} L_{12} & L_3 & L \\ M_{12} & M_3 & -M\end{pmatrix} \\ & \times\begin{Bmatrix} L_1 & L_2 & L_{12} \\ s & t & L_3 \\ L-j & j & L\end{Bmatrix},\end{aligned} \tag{11}$$

we write Eq. (10) as

$$\begin{aligned} Z_{LM} = {} & \sum_{stjL_1L_2L_{12}L_3M_1M_2}(-1)^{j+M_3+L_1+L_2+L_{12}}(2L_{12}+1)K(s,t,j,L_1,L_2,L_3,L) \\ & \times\begin{pmatrix} L_1 & L_2 & L_{12} \\ M_1 & M_2 & -M_{12}\end{pmatrix}\begin{pmatrix} L_{12} & L_3 & L \\ M_{12} & M_3 & -M\end{pmatrix}\begin{Bmatrix} L_1 & L_2 & L_{12} \\ s & t & L_3 \\ L-j & j & L\end{Bmatrix} \\ & \times Y_{L_1M_1}(\hat{\mathbf{r}}_1)Y_{L_2M_2}(\hat{\mathbf{r}}_2)Y_{L_3M_3}(\hat{\mathbf{r}}_3).\end{aligned} \tag{12}$$

Comparing Eqs. (1) and (12), we arrive at the required result given by Eq. (2), noting Eq. (5).

[1]J. D. Talman, *Special Functions* (Benjamin, New York, 1968).
[2]A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University, Princeton, NJ, 1960).
[3]A. de-Shalit, Phys. Rev. **91**, 1479 (1953).

Superposition principles for matrix Riccati equations[a]

J. Harnad and P. Winternitz
Centre de Recherche de Mathématiques Appliquées, Université de Montréal, C.P.6128, Succ. A, Montréal, Québec, Canada H3C 3J7

R. L. Anderson
Department of Physics and Astronomy, University of Georgia, Athens, Georgia 30602

(Received 18 August 1981; accepted for publication 2 April 1982)

A superposition rule is obtained for the matrix Riccati equation (MRE) $\dot W = A + WB + CW + WDW$ [where $W(t)$, $A(t)$, $B(t)$, $C(t)$, and $D(t)$ are real $n \times n$ matrices], expressing the general solution in terms of five known solutions for all $n \geqslant 2$. The symplectic MRE ($W = W^T$, $A = A^T$, $D = D^T$, $C = B^T$) is treated separately, and a superposition rule is derived involving only four known solutions. For the "unitary" and GL(n,$\mathbb{R}$) subcases (with $D = A$ and $C = B^T$, or $D = -A$ and $C = B^T$, respectively), superposition rules are obtained involving only two solutions. The derivation of these results is based upon an interpretation of the MRE in terms of the action of the groups SL($2n$,$\mathbb{R}$), SP($2n$,$\mathbb{R}$), U(n), and GL(n,$\mathbb{R}$) on the Grassman manifold $G_n(\mathbb{R}^{2n})$.

PACS numbers: 02.30.Hq

I. INTRODUCTION

The main purpose of this paper is to derive superposition rules for the matrix Riccati equation (MRE)

$$\dot W(t) = A + WB + CW + WDW, \tag{1.1}$$

where $W(t)$ is a rectangular $n \times k$ matrix function of $t \in \mathbb{R}$ and A, B, C, and D are given, t-dependent coefficient matrices of dimension $n \times k$, $k \times k$, $n \times n$, and $k \times n$, respectively. This work is part of a general program[1–3] involving the study of systems of ordinary nonlinear differential equations for which superposition rules exist; that is, for which it is possible to express the general solution in terms of a finite number of particular solutions.

In a previous work,[3] it was shown how such superposition principles could be derived for certain systems of first-order ODE's satisfying a group theoretical characterization due to Lie.[4] Geometrically, such equations describe the flows of time-dependent vector fields induced by an infinitesimal group action. The superposition rule derives from the fact that these flows lie within the individual group orbits. Two classes of examples for which such superposition principles were obtained[1–3] consist of the coupled Riccati equations characterized by the infinitesimal projective transformations expressed in affine coordinates and by pseudoorthogonal transformations acting conformally on certain projective quadrics. These were referred to as vector Riccati equations of the projective and conformal type, respectively.

In Sec. II we show that Eq. (1.1) also satisfies Lie's characterization, with the underlying infinitesimal group action that of SL($n + k$,$\mathbb{R}$) on the Grassman manifold $G_k(\mathbb{R}^{n+k})$ of k planes in $\mathbb{R}^{n+k}$. For $k = 1$, the space becomes $\mathbb{R}P^n$, and Eq. (1.1) reduces to the projective Riccati equation.

Another case of particular interest is $k = n$, for which the MRE (1.1) involves $n \times n$ square matrices only. This case is treated in detail by two different methods in Sec III, where it is shown how a superposition principle may be derived for arbitrary $n \geqslant 2$, expressing the general solution in terms of only *five* known solutions, chosen arbitrarily up to certain specified independence conditions. The admissible initial conditions for these five solutions form a dense set in the Cartesian product $[G_n(\mathbb{R}^{2n})]^5$. A simple characterization is given for such generic sets, and two algorithms leading to the superposition law are presented, one of them providing an explicit formula.

Section IV is devoted to certain special cases of the square matrix Riccati equation. We first assume that the matrices in (1.1) satisfy

$$W = W^T, \quad A = A^T, \quad D = D^T, \quad B^T = C, \tag{1.2}$$

which define "symplectic Riccati equations," related to the minimal orbit of the subgroup Sp($2n$,$\mathbb{R}$) $\subset$ SL($2n$,$\mathbb{R}$). The general solution, respecting the condition $W = W^T$, can be expressed in terms of only four solutions. Under further restrictions on the coefficients in (1.1), we obtain Riccati equations related to the group U(n) $\subset$ Sp($2n$,$\mathbb{R}$) and GL(n,$\mathbb{R}$) $\subset$ Sp($2n$,$\mathbb{R}$) for which two known solutions suffice.

Such symplectic Riccati equations have an interesting relationship to classical mechanics since they involve infinitesimal symplectic transformations acting upon the Lagrangian subspaces of a symplectic vector space.[5,6] They arise in particular in the Hamiltonian formulation of the optimal control problem with a linear system and quadratic cost functional.[7] For other numerous applications of matrix Riccati equations as well as a survey of the existing theory and references to the original literature, see Ref. 8.

II. MATRIX RICCATI EQUATIONS

The homogeneous coordinates for a point $p \in G_k(\mathbb{R}^{n+k})$ are given by the components of a $[(k + n) \times k]$-dimensional matrix:

$$\begin{pmatrix} X \\ Y \end{pmatrix}, \quad X \in \mathbb{R}^{n \times k}, \quad Y \in \mathbb{R}^{k \times k},$$

[a] Work supported in part by the Natural Sciences and Research Council of Canada and the Fonds FCAC pour l'aide et soutien à la recherche.

whose columns span the k plane defining p. The point p is thus identified with the equivalence class $[\binom{X}{Y}]$ under the relation:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \begin{pmatrix} XG \\ YG \end{pmatrix}, \quad G \in \mathrm{GL}(k,\mathbb{R}), \tag{2.1}$$

identifying different bases for the same space. The action of an element

$$g = \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \in \mathrm{SL}(k+n,\mathbb{R}), \quad \det g = 1 \tag{2.2}$$

upon $G_k(\mathbb{R}^{n+k})$ is obtained by the projection $\pi:\binom{X}{Y} \to [\binom{X}{Y}]$ from the linear action:

$$g:\begin{pmatrix} X \\ Y \end{pmatrix} \to \begin{pmatrix} X' \\ Y' \end{pmatrix} = \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}. \tag{2.3}$$

On the affine subspace defined by $\det Y \neq 0$ we may define coordinates

$$W = XY^{-1} \in \mathbb{R}^{n\times k}, \tag{2.4}$$

in terms of which this action is given by the matrix linear fractional transformations:

$$g: W \to W' = (MW + N)(PW + Q)^{-1}. \tag{2.5}$$

The infinitesimal group action is given by the homomorphism

$$\phi: \mathrm{sl}(n+k,\mathbb{R}) \to \mathfrak{X}(G_n(\mathbb{R}^{n+k}))$$

from the Lie algebra $\mathrm{sl}(n+k,\mathbb{R})$ to the algebra of smooth vector fields on $G_n(\mathbb{R}^{n+k})$ defined by

$$\phi(\xi)f(p) = -\frac{d}{dt} f(e^{t\xi}p)|_{t=0}, \tag{2.6}$$

$$\xi \in \mathrm{sl}(n+k,\mathbb{R}), \quad f \in C^\infty(G_n(\mathbb{R}^{n+k})).$$

Expressed in affine coordinates, the image of

$$\xi \equiv \begin{pmatrix} C & A \\ -D & -B \end{pmatrix}, \tag{2.7}$$

$$A \in \mathbb{R}^{n\times k}, \quad B \in \mathbb{R}^{k\times k},$$

$$C \in \mathbb{R}^{n\times n}, \quad D \in \mathbb{R}^{k\times n}, \quad \mathrm{tr}C - \mathrm{tr}B = 0$$

is the vector field

$$\phi(\xi) = -(A_{\alpha\mu} + W_{\alpha\nu}B_{\nu\mu} + C_{\alpha\beta}W_{\beta\mu} + W_{\alpha\nu}D_{\nu\beta}W_{\beta\mu})\frac{\partial}{\partial W_{\alpha\mu}}. \tag{2.8}$$

For a given curve in $\mathrm{sl}(n+k,\mathbb{R})$

$$\xi(t) = \begin{pmatrix} C(t) & A(t) \\ -D(t) & -B(t) \end{pmatrix}, \tag{2.9}$$

Eq. (1.1) defines the flow of the time-dependent vector field $\phi(\xi(t))$. The general solution is therefore of the form

$$W(t) = [M(t)W_0 + N(t)][P(t)W_0 + Q(t)]^{-1}, \tag{2.10}$$

where the curve in the group $\mathrm{SL}(n+k,\mathbb{R})$

$$g(t) = \begin{pmatrix} M(t) & N(t) \\ P(t) & Q(t) \end{pmatrix}, \tag{2.11}$$

$$M(t) \in \mathbb{R}^{n\times n}, \quad N(t) \in \mathbb{R}^{n\times k}, \quad P(t) \in \mathbb{R}^{k\times n}, \quad Q(t) \in \mathbb{R}^{k\times k},$$

is the unique solution to

$$\begin{pmatrix} \dot M & \dot P \\ \dot N & \dot Q \end{pmatrix} = \begin{pmatrix} C & A \\ -D & -B \end{pmatrix} \begin{pmatrix} M & N \\ P & Q \end{pmatrix} \tag{2.12}$$

for some arbitrarily fixed initial condition

$$g(t_0) = \begin{pmatrix} M_0 & N_0 \\ P_0 & Q_0 \end{pmatrix}, \tag{2.13}$$

and $W_0 \in \mathbb{R}^{n\times k}$ is a constant matrix determined by the initial condition for $W(t)$.

Equivalently, we may consider the linear action of $\mathrm{sl}(n+k,\mathbb{R})$ on $\mathbb{R}^{(n+k)\times k}$ and the corresponding differential equation:

$$\begin{pmatrix} \dot X(t) \\ \dot Y(t) \end{pmatrix} = \begin{pmatrix} C(t) & A(t) \\ -D(t) & -B(t) \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}, \tag{2.14}$$

$$X(t) \in \mathbb{R}^{n\times k}, \quad Y(t) \in \mathbb{R}^{k\times k}.$$

All solutions to Eq. (1.1) may be obtained by regarding $\binom{X(t)}{Y(t)}$ as the homogeneous coordinates for a curve in $G_k(\mathbb{R}^{n+k})$, with affine coordinates

$$W(t) = X(t)Y^{-1}(t), \tag{2.15}$$

defined whenever $Y(t)$ is nonsingular.

The determination of a superposition rule for Eq. (1.1) amounts to finding the group element $g(t)$ entering in Eq. (2.10), up to a coset in the isotropy group of W_0, by "solving" the corresponding relations for $\{M,N,P,Q\}$ in terms of a sufficient number of known solutions $\{W_i(t)\}$.

To summarize, the matrix Riccati equation (1.1) for rectangular matrices $W \in \mathbb{R}^{n\times k}$ is associated with the Lie algebra $\mathrm{sl}(n+k,\mathbb{R})$. All solutions to (1.1) can be obtained by solving the linear equations (2.14) for $X(t)$ and $Y(t)$; the solutions of the MRE (1.1) are then given by (2.15). A superposition principle is obtained from Eq. (2.10), once M, N, P, and Q are determined in terms of a sufficient number of known solutions to the MRE. This can be done for arbitrary k and n, following methods developed earlier.[3]

The most interesting case is that of square matrix Riccati equations ($k = n$), which is treated in detail in the next section.

III. SUPERPOSITION PRINCIPLES FOR SQUARE MATRIX RICCATI EQUATIONS

We now restrict ourselves to the square MRE, i.e., Eqs. (1.1) for $n = k$:

$$W(t), A(t), B(t), C(t), D(t) \in \mathbb{R}^{n\times n}. \tag{3.1}$$

We shall first establish that the number of solutions needed for the superposition formula is five (for any $n \geqslant 2$) and specify the restrictions imposed on these five solutions (or their initial conditions). We shall call any set of five solutions satisfying the established conditions a *fundamental set of solutions* for the MRE. Next, we use three of the solutions belonging to the fundamental set to reduce the MRE to a particularly simple linear matrix equation for the *matrix anharmonic ratio* of four solutions,

$$R = (W_2 - W_3)^{-1}(W_3 - W_1)(W_1 - W)^{-1}(W - W_2). \tag{3.2}$$

Finally we derive two different forms of the superposition formula, expressing the general solution $W(t)$ of the MRE in terms of five generically chosen particular solutions (the fundamental set).

A. The fundamental set of solutions

Following the general procedure outlined elsewhere,[2,3] we write the solution of the MRE (1.1) in the form

$$W(t) = (MU+N)(PU+Q)^{-1} \tag{3.3}$$

where $U\in\mathbb{R}^{n\times n}$ is a constant matrix. The group-valued function

$$g(t) = \begin{pmatrix} M(t) & N(t) \\ P(t) & Q(t) \end{pmatrix} \in \mathrm{SL}(2n,\mathbb{R}), \quad M,N,P,Q\in\mathbb{R}^{n\times n}, \tag{3.4}$$

must be determined in terms of a finite number m of known solutions. The group element $g(t)$ satisfies the system of linear equations (2.12) for arbitrarily chosen initial conditions (2.13) and A,B,C,D now have values in $\mathbb{R}^{n\times n}$. This arbitrariness amounts to the fact that the matrix $g(t)$ can at any stage be replaced by $g(t)g_0^{-1}$, where g_0 is an arbitrary constant matrix and U is redefined accordingly. The matrix U in (3.3) together with the choice of initial condition $g(t_0)$ will determine the initial conditions for the solution $W(t)$ for $t = t_0$. We will not necessarily impose

$$W(t_0) = U, \quad \text{i.e.,} \quad g(t_0) = \begin{pmatrix} M(t_0) & N(t_0) \\ P(t_0) & Q(t_0) \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{3.5}$$

but rather make full use of the liberty in the choice of $g(t_0)$. We shall, however, call U the "initial condition matrix."

The number of solutions m in the fundamental set is determined as the lowest number for which the $\mathrm{SL}(2n,\mathbb{R})$ action (2.5) on the Cartesian product $[G_n(\mathbb{R}^{2n})]^m$ of m Grassmanians is free (except possibly on some singular orbits of lower dimension). The joint stabilizer $G_0^m \subset \mathrm{SL}(2n,\mathbb{R})$ of the m initial condition matrices $U_1,\ldots,U_m$ should thus be just the identity transformation. That is, the equations

$$U_i = (M_0U_i + N_0)(P_0U_i + Q_0)^{-1}, \quad i = 1,2,\ldots,m, \tag{3.6}$$

where U_i, M_0, N_0, P_0, and Q_0 are constant $n\times n$ real matrices, should imply

$$\begin{pmatrix} M_0 & N_0 \\ P_0 & Q_0 \end{pmatrix} = \lambda \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \quad \lambda\in\mathbb{R}. \tag{3.7}$$

We shall now take five initial condition matrices $U_1,\ldots,U_5$ and transform them into a convenient standard form by a constant matrix linear fractional transformation [making use of the arbitrariness in the initial conditions for $g(t)$].

Let us start with three matrices satisfying

$$\det(U_i - U_k)\neq 0, \quad i,k = 1,2,3. \tag{3.8}$$

By a matrix linear fractional transformation with constant coefficients

$$g_0 = \begin{pmatrix} I & , & -U_2 \\ (U_3-U_2)(U_2-U_1)^{-1} & , & -(U_3-U_2)(U_3-U_1)^{-1}U_1 \end{pmatrix} \tag{3.9}$$

we can transform U_1, U_2, and U_3 to the "standard" form

$$U_1^s \to \infty, \quad U_2^s = 0, \quad U_3^s = I \tag{3.10}$$

($U_1^s\to\infty$ should be interpreted as a point on the Grassmannian with homogeneous coordinates $X = I$, $Y = 0$). With no loss of generality, we can assume that the solutions $W_1(t)$, $W_2(t)$, and $W_3(t)$ correspond to the initial condition matrices (3.10), since the transformation matrix (3.9) is absorbed into the definition of $g(t)$.

The stabilizers in $\mathrm{SL}(2n,\mathbb{R})$ of U_1^s, of the pair U_1^s and U_2^s and of the triplet U_1^s, U_2^s, and U_3^s are, respectively,

$$G_0^1 = \begin{pmatrix} M & N \\ 0 & Q \end{pmatrix}, \quad G_0^2 = \begin{pmatrix} M & 0 \\ 0 & Q \end{pmatrix}, \quad G_0^3 = \begin{pmatrix} Q & 0 \\ 0 & Q \end{pmatrix}. \tag{3.11}$$

Notice that, when expressed in terms of affine coordinates, W on the Grassmannian the group $G_0^1 \sim [\mathrm{SL}(n,\mathbb{R}) \otimes \mathrm{SL}(n,\mathbb{R}) \otimes \mathbb{R}^1] \ltimes \mathbb{R}^{n\times n}$ has a linear affine (inhomogeneous) action, $G_0^2 \sim \mathrm{SL}(n,\mathbb{R}) \otimes \mathrm{SL}(n,\mathbb{R}) \otimes \mathbb{R}^1$ a linear homogeneous action, and $G_0^3 \sim \mathrm{SL}(n,\mathbb{R}) \otimes Z^2$ acts linearly and by conjugation. For example if $g\in G_0^3$ (2.5) reduces to

$$W' = QWQ^{-1}. \tag{3.12}$$

According to the general group-theoretical method,[2,3] knowledge of a given set of solutions permits the reduction of the underlying equations to the type associated with the stabilizer of their initial conditions. In the present case, one, two, or three known solutions, therefore, make it possible to reduce the MRE to a linear inhomogeneous equation, a linear homogeneous equation, or a commutator type linear homogeneous equation, respectively (see below).

In the special case of $n = 1$, G_0^3 already represents the identity transformation; hence three solutions suffice to express the general solution of the ordinary Riccati equation in closed form. For $n\geqslant 2$ we must further reduce the stabilizer by adding more solutions. Since G_0^3 acts as in (3.12), the additional solutions must be such that the equations

$$U_a = g_0 U_a g_0^{-1}, \quad g_0\in G_0^3, \quad a = 4,5 \tag{3.13}$$

should imply $G = \lambda I$ ($\lambda\in\mathbb{R}$). It is well known (Schur's lemma[9,10]) that over an algebraically closed field two matrices, U_4 and U_5, would suffice in (3.13) to imply $G = \lambda I$ if and only if they have no common irreducible invariant subspaces. Over the field of real numbers the situation is slightly more complicated; however, two generically chosen matrices will still suffice to reduce g_0 in (3.13) to the identity transformation. Indeed, assume that U_4 has all eigenvalues different, n_1 of them real and n_2 complex conjugate pairs. We can use the stabilizer G_0^3 of U_1^s, U_2^s, and U_3^s to reduce U_4 to its Jordan canonical form, in this case

$$
U_4^s = \begin{pmatrix}
\lambda_1 & & & & & & & & \\
& \ddots & & & & & & & \\
& & \lambda_{n_1} & & & & & & \\
& & & a_{n_1+1} & b_{n_1+1} & & & & \\
& & & -b_{n_1+1} & a_{n_1+1} & & & & \\
& & & & & \ddots & & & \\
& & & & & & a_{n_1+n_2} & b_{n_1+n_2} \\
& & & & & & -b_{n_1+n_2} & a_{n_1+n_2}
\end{pmatrix}, \tag{3.14}
$$

$$\lambda_i, a_i, b_i \in \mathbb{R}, \quad b_i > 0, \quad n_1 + 2n_2 = n.$$

Without loss of generality we assume that $U_1,\dots,U_4$ are in the form (3.10) and (3.14) and again absorb the constant matrix simultaneously transforming U_i into U_i^s $(i = 1,\dots,4)$ into the initial conditions for $g(t)$. The simultaneous stabilizer of $U_1^s,\dots,U_4^s$ is

$$G_0^4 = \left\{ \begin{pmatrix} Q_D & 0 \\ 0 & Q_D \end{pmatrix} \right\} \tag{3.15}$$

where Q_D is any block diagonal matrix of the same structure as U_4^s in (3.14).

Finally, let U_5 have no common irreducible invariant subspaces with U_4^s. This condition can be described simply in terms of graph theory. The irreducible invariant subspaces of U_4^s are clearly all one- or two-dimensional. Let us associate a point P_i $(i = 1,\dots,n_1 + n_2)$ with each invariant subspace of U_4^s and introduce an edge $(P_i P_k)$ whenever the matrix U_5 connects the subspaces i,k. If the obtained graph is a connected one, then U_5 and U_4^s have no common irreducible invariant subspaces, and Eqs. (3.13) will together imply that $G = \lambda I$.

Let us summarize the results and make some comments.

(1) The subgroup of SL($2n$,$\mathbb{R}$), leaving the five initial condition matrices $U_1,\dots,U_5$ specified above invariant, is the identity group. The group element (3.4) can be reconstructed from the knowledge of five such solutions (for $n \geqslant 2$) since formula (2.5) can be interpreted as an SL($2n$,$\mathbb{R}$) transformation, taking the five initial conditions into the five solutions at some t (in a neighborhood of $t = t_0$).

(2) The five solutions contained in a "fundamental" set of solutions can be generically chosen, i.e., their initial data are arbitrary, except for the followng conditions (which define an open, dense set):

(i) $\det(U_k - U_1) \neq 0$, $k = 2,\dots,5$, $\det(U_2 - U_3) \neq 0$.

(ii) For the initial conditions in (3.3) and (3.4) arranged as in Eq. (3.10), all eigenvalues of U_4 are distinct. (3.16a)

(iii) U_5 does not leave any nontrivial irreducible invariant subspaces of U_4 invariant.

Conditions (3.16a) can be reformulated as follows in terms of the solutions $\{W_1(t),\dots,W_5(t)\}$ related to $\{U_1,\dots,U_5\}$ by Eq. (3.3):

(i) $\det(W_k - W_1) \neq 0$, $\det(W_2 - W_3) \neq 0$, $k = 2,\dots,5$.

(ii) All eigenvalues of the matrix anharmonic ratio

$$R_4 = (W_2 - W_3)^{-1}(W_3 - W_1)(W_1 - W_4)^{-1} \times (W_4 - W_2) \tag{3.16b}$$

and distinct.

(iii) The anharmonic ratio

$$R_5 \equiv (W_2 - W_3)^{-1}(W_3 - W_1)(W_1 - W_5)^{-1}(W_5 - W_2)$$

regarded as a linear map leaves none of the irreducible invariant subspaces of R_4 invariant.

In fact, by continuity, it is sufficient for these conditions to hold for some initial value of t, say $t = t_0$.

In order to show that conditions (i) in Eqs. (3.16a) and (3.16b) are equivalent; note that any linear fractional transformation may be obtained by a composition of ones of the type:

$U \rightarrow AUB$ (A,B nonsingular) (left and right linear transformation),

$U \rightarrow U + C$ (translation),

$U \rightarrow U^{-1}$ (inversion).

Each of these transformations applied to a pair of matrices U,V leaves the property

$$\det(U - V) \neq 0$$

invariant (assuming, of course, that inversion is well-defined). Therefore since $\{W_1(t),\dots,W_5(t)\}$ are related to $\{U_1 \cdots U_5\}$ by the linear fractional transformation (3.3), the relations (i) in (3.16a) and (3.16b) are equivalent. The equivalence of (ii) and (iii) in (3.16a) and (3.16b) is shown in the following section.

The conditions (3.16) play a role analogous to the condition of linear independence for solutions of linear differential equations. With no loss of generality we can take $U_1,\dots,U_4$ in the "standard" form (3.10) and (3.14).

(3) The group SL($2n$,$\mathbb{R}$) acts freely on a fundamental set of five solutions. If one more solution is added, it is possible to form n^2 independent SL($2n$,$\mathbb{R}$) invariants out of these six solutions of the MRE. These invariants are constant in t, and they thus implicitly determine the sixth solution in terms of the original five.

B. Reduction of SL($2n$,$\mathbb{R}$) group action to SL(n,$\mathbb{R}$) conjugacy and the matrix anharmonic ratio

We shall now show explicitly how the use of just three known solutions $W_1(t)$, $W_2(t)$, and $W_3(t)$ with initial conditions satisfying (3.8) reduces the problem of reconstructing the SL($2n$,$\mathbb{R}$) group element (3.4) to that of reconstructing the SL(n,$\mathbb{R}$) element

$$g = \begin{pmatrix} Q(t) & 0 \\ 0 & Q(t) \end{pmatrix} \in \mathrm{SL}(n,\mathbb{R}). \tag{3.17}$$

Because of the freedom of choice of initial conditions in (2.12), we can assume the initial data matrices to be in the form (3.10). Substituting U_1^s, U_2^s, and U_3^s successively into the relation (3.3), we obtain

$$W_1 = MP^{-1}, \quad W_2 = NQ^{-1}, \quad W_3 = (M+N)(P+Q)^{-1}. \tag{3.18}$$

Solving for M, N, and P in terms of W_1, W_2, W_3, and Q, and substituting back into (3.3), we obtain

$$W = [W_1(W_3 - W_1)^{-1}(W_2 - W_3)QU + W_2Q] \times [(W_3 - W_1)^{-1}(W_2 - W_3)QU + Q]^{-1}. \tag{3.19}$$

Formula (3.19) expresses the general solution W in terms of three particular solutions and one unknown matrix $Q \in \mathrm{SL}(n,\mathbb{R})$. From (3.19) we can express QUQ^{-1} in terms of W, W_1, W_2, and W_3. We use the "matrix anharmonic ratio" R introduced in (3.2), and from (3.19) we have

$$R = QUQ^{-1}. \tag{3.20}$$

We have thus obtained the known result[8] that the matrix anharmonic ratio R is conjugate to a constant matrix U. This result is a consequence of the fact that the isotropy group G_0^3 [(3.11)] acts by conjugation on the initial data, and we have thus put the result (3.20) into a group-theoretical context.

The conditions (3.8) imposed on the initial condition matrices ensure that, provided the coefficients A, B, C, and D in the MRE (1.1) are sufficiently regular, the inverses $(W_2 - W_3)^{-1}$ and $(W_3 - W_1)^{-1}$ in (3.2) and (3.19) exist at least within some neighborhood of $t = t_0$. The equivalence of (ii) and (iii) in Eqs. (3.16a) to those in (3.16b) follows from Eq. (3.20), since conjugation preserves the eigenvalues and the property that a pair of matrices have no irreducible invariant subspaces in common.

The anharmonic ratio R is defined for solutions W satisfying

$$\det(W - W_1) \neq 0. \tag{3.21}$$

If this condition is not observed, but we have

$$\det(W - W_2) \neq 0, \tag{3.22}$$

then we can use a different anharmonic ratio

$$\tilde{R} = (W - W_2)^{-1}(W_1 - W)(W_3 - W_1)^{-1} \times (W_2 - W_3) = QVQ^{-1}. \tag{3.23}$$

If both R and $\tilde{R}$ exist, then we have

$$\tilde{R} = R^{-1}, \quad V = U^{-1}. \tag{3.24}$$

For $n = 1$ the MRE reduces to the ordinary Riccati equation and the matrix anharmonic ratio becomes the ordinary anharmonic ratio. The quantities W, R, and U are then all scalars, and (3.20) simply states that the anharmonic ratio of four solutions of the Riccati equations is a constant. For $n \geqslant 2$ we must still determine the conjugating matrix Q, i.e., reconstruct the SL(n,$\mathbb{R}$) group element Q acting as in (3.20).

Note that (3.19) and (3.20) were derived by a partial reconstruction of the SL($2n$,$\mathbb{R}$) group element. Let us rederive these formulas by a different procedure, namely linearizing the MRE. In the process we establish further properties of the matrix anharmonic ratio. Let W_1, W_2, and W_3 again be three solutions of the MRE (1.1) satisfying (locally) the conditions $\det(W_i - W_k) \neq 0$, $i = 1,2,3$.

Perform a matrix fractional linear transformation from W to a new variable S putting

$$S = (W - W_2)^{-1}. \tag{3.25}$$

If W and W_2 satisfy the MRE (1.1), we find that S satisfies a linear inhomogeneous equation of the type corresponding to the stabilizer of one initial condition

$$\dot{S} = -\tilde{B}S - S\tilde{C} - D, \tag{3.26}$$

$$\tilde{B} = B + DW_2, \quad \tilde{C} = C + W_2D. \tag{3.27}$$

A second transformation

$$T = S - S_1 = (W_2 - W)^{-1}(W - W_1)(W_1 - W_2)^{-1} \tag{3.28}$$

introduces the variable T satisfying a linear homogeneous equation, corresponding to the stabilizer of two initial conditions:

$$\dot{T} = -\tilde{B}T - T\tilde{C} \tag{3.29}$$

[S_1 is a solution of (3.26), W, W_1, and W_2 solutions of the MRE (1.1)].

Finally, use the third particular solution W_3 of (1.1), or equivalently, a particular solution T_3 of (3.29) to put

$$R = T_3T^{-1} = (W_2 - W_3)^{-1}(W_3 - W_1) \times (W_1 - W)^{-1}(W - W_2). \tag{3.30}$$

We have again arrived at the matrix anharmonic ratio R; from (3.29) and (3.30) we find that R satisfies a linear commutation type equation, corresponding to the stabilizer of three initial values:

$$\dot{R} = [R,\tilde{B}]. \tag{3.31}$$

The solution R of (3.31) can be written in the form (3.20), where $U \in \mathbb{R}^{n\times n}$ is an arbitrary constant matrix [the same as in (3.19) and (3.20)] related to the initial conditions for the general solution W of the MRE, or, equivalently, to the initial conditions for the anharmonic ratio of the four solutions W, W_1, W_2, and W_3. Equation (3.31) implies that the matrix Q in (3.19) and (3.20) must satisfy the simple linear equation

$$\dot{Q} = -\tilde{B}Q, \quad Q(0) = Q_0 \tag{3.32}$$

(for some initial condition Q_0). Equation (3.32) can of course be solved numerically and in some cases analytically.[11] We are, however, after a superposition law and shall hence express Q (or R) in terms of two more solutions of the MRE (1.1).

Let us make a comment on the group theoretical significance of the matrix anharmonic ratio R. For $n = 1$, R is the SL(2,$\mathbb{R}$) group invariant, the existence of which is guaranteed by the fact that the isotropy subgroup of SL(2,$\mathbb{R}$) for three different points is the identity. For $n \geqslant 2$ the group invariant is not R itself but rather $U = Q^{-1}RQ$, where Q involves two more solutions. The existence of this n^2-dimensional invariant, depending on six solutions is guaranteed by the fact that the isotropy subgroup of SL($2n$,$\mathbb{R}$) leaving five points invariant is the identity subgroup. From R itself we may form the elementary trace invariants $\{\mathrm{Tr} R^i\}_{i=1,\ldots,n}$, which in view of (3.31) or (3.20) remain constant in t, but for $n \geqslant 2$ the number is not large enough to be used to determine W in terms of $\{W_1, W_2, W_3\}$ alone.

C. The superposition formula

1. Reconstruction of the $SL(n,\mathbb{R})$ group element Q

For $n \geqslant 2$, assume that two more solutions W_4 and W_5 of the MRE (1.1) are known and such that the matrices R_4 and R_5 satisfy the conditions (3.16b) (ii) and (iii). Since these are of the form

$$R_a = Q U_a Q^{-1}, \quad a = 4,5, \tag{3.33}$$

for some constant "initial vaue" matrices U_4, U_5, with Q arbitrary up to a constant matrix multiplier on the right, we may, without loss of generality, assume U_4 is the Jordan normal form of R_4 and, hence, after choosing some ordering of the eigenspaces, is of the form U_4^s of Eq. (3.14). We may thus write $Q(t)$ in the form

$$Q(t) = Q_0(t) Q_{\mathrm{D}}(t), \tag{3.34}$$

where Q_{D} is in the stabilizer of U_4^s and $Q_0 \in \mathrm{GL}(n,\mathbb{R})$ is now determined from (3.20),

$$R_4(t) Q_0(t) = Q_0(t) U_4^s, \tag{3.35}$$

as being any matrix which reduces $R_4(t)$ to its Jordan canonical form U_4^s. The columns of $Q_0(t)$ are vectors spanning the irreducible invariant subspaces of $R_4(t)$. The first n_1 columns, corresponding to the real eigenvalues λ_i are determined uniquely, up to normalization, as the real eigenvectors $\mathbf{q}_i(t)$ of $R_4(t)$. The remaining $2n_2$ columns correspond to two-dimensional invariant subspaces, one for each complex conjugate pair of eigenvalues $(a_k \pm i b_k)$. An arbitrary basis can be chosen in each invariant subspace, e.g., by taking the real and imaginary part of each complex eigenvector $\mathbf{r}_k \pm i\mathbf{p}_k$. The entire ambiguity in Q_0 is absorbed in the as-yet unknown matrix Q_{D}. Thus Q_0 can be chosen to be

$$Q_0 = (\mathbf{q}_1, \ldots, \mathbf{q}_{n_1}, \mathbf{r}_{n_1+1}, \mathbf{p}_{n_1+1}, \ldots, \mathbf{r}_{n_1+n_2}, \mathbf{p}_{n_1+n_2}). \tag{3.36}$$

Finally we use the last solution W_5 and the corresponding R_5 to calcuate the $n_1 + 2n_2$ independent real entries in Q_{D} from the system of linear algebraic equations

$$(Q_0^{-1} R_5(t) Q_0) Q_{\mathrm{D}} = Q_{\mathrm{D}} U_5, \tag{3.37}$$

where U_5 is defined to be the value of $Q_0^{-1} R_5(t_0) Q_0$ at some arbitrarily chosen initial time $t = t_0$ This determines Q_{D} and hence Q uniquely (up to an immaterial scalar multiplier).

Finally the superposition formula is obtained in one of two forms

$$\begin{aligned} W &= [W_1 (W_3 - W_1)^{-1} (W_2 - W_3) Q U + W_2 Q] \\ &\quad \times [(W_3 - W_1)^{-1} (W_2 - W_3) Q U + Q]^{-1} \qquad (3.38a) \\ &= [W_2 Q V + W_1 (W_3 - W_1)^{-1} (W_2 - W_3) Q] \\ &\quad \times [Q V + (W_3 - W_1)^{-1} (W_2 - W_3) Q]^{-1}, \qquad (3.38b) \end{aligned}$$

where Q is completely specified as $Q = Q_0 Q_{\mathrm{D}}$ by (3.36) and (3.37) The formulas (3.38a) and (3.38b) are equivalent if $\det U \neq 0$, $\det V \neq 0$; then we have $U = V^{-1}$.

We emphasize that we are dealing with the generic situation and that the fundamental set of solutions $W_1, \ldots, W_5$ consists of five arbitrary solutions satisfying (3.16). [The initial conditions $W_i(t_0)$ $(i = 1,\ldots,5)$ do not have to be chosen in the standard forms (3.10) and (3.14) which we have been using purely for convenience.]

2. Superposition formula for matrix anharmonic ratios

The matrix anharmonic ratio R satisfies Eq. (3.31). The solutions of this equation, in addition to forming a linear space, also form an associative algebra under matrix multiplication. Indeed, the commutator character of this equation assures that the product of two (or more) solutions is again a solution. It follows that not only can we write a linear superposition formula

$$R = \sum_{i=1}^{n^2} c_i R_i, \tag{3.39}$$

involving n^2 linearly independent solutions, but that we can generate n^2 linearly independent solutions of (3.31) as polynomials in terms of a small number of generators of the associative algebra. As a matter of fact, two appropriately chosen solutions R_4 and R_5 of (3.31) will suffice. In terms of the initial data U_4 and U_5 we have

$$R_4 = Q U_4 Q^{-1}, \quad R_5 = Q U_5 Q^{-1} \tag{3.40}$$

and

$$R_4 R_5 = Q U_4 U_5 Q^{-1}. \tag{3.41}$$

If U_4 and U_5 generate polynomially the entire algebra of matrices $U \in \mathbb{R}^{n \times n}$, i.e., n^2 linearly independent initial conditions, then $R_4(t)$ and $R_5(t)$ will generate polynomially the entire associative algebra of solutions of (3.31).

The necessary and sufficient conditions for U_4 and U_5 to generate the entire algebra of matrices in $\mathbb{R}^{n \times n}$ are that[9,10]

(1) R_4 and R_5 have no common invariant irreducible subspaces.

(2) The only matrices simultaneously commuting with R_4 and R_5 are multiples of the identity. (This second condition would be a consequence of the first for an algebraically closed field, but we are working over the field of real numbers.)

It is sufficient to verify that these two conditions be satisfied for some specific value of t, say $t = t_0$. To proceed further, consider the generic case when one of the matrix anharmonic ratios, say R_4, has n different nonvanishing eigenvalues, so that U_4 can be chosen as its Jordan canonical form U_4^s of (3.14). The powers $U_4^s, (U_4^s)^2, \ldots, (U_4^s)^n$ are linearly independent, and linear combinations of them provide us with the n matrices

$$E_1, \ldots, E_{n_1}, \quad S_{n_1+1}, \ldots, S_{n_1+n_2}, \quad J_{n_1+1}, \ldots, J_{n_1+n_2}, \tag{3.42}$$

where

$$\begin{aligned} &(E_a)_{ik} = \delta_{ai} \delta_{ak}, \quad (S_\alpha)_{ik} = \delta_{i\alpha} \delta_{k\alpha} + \delta_{i\alpha+1} \delta_{k\alpha+1}, \\ &(J_\alpha)_{ik} = \delta_{i\alpha} \delta_{k\alpha+1} - \delta_{i\alpha+1} \delta_{k\alpha}, \\ &1 \leqslant a \leqslant n_1, \quad n_1 + 1 \leqslant \alpha \leqslant n_1 + n_2, \quad i,k = 1,\ldots,n_1 + 2n_2. \end{aligned} \tag{3.43}$$

The matrix R_5 must be chosen so as to satisfy the above conditions (1) and (2). This can be assured by requiring that, in the basis where $R_4(t_0) = U_4^s$, the matrix $R_5(t_0) \equiv U_5$ has the property that its nonzero entries define the arcs of a strongly connected oriented graph [we introduce an arc from the point P_i to P_k if $(U_5)_{ki} \neq 0$]. An example of such a U_5 and its graph is

$$U_5 = \begin{pmatrix} 0 & * & 0 & * \\ * & 0 & 0 & 0 \\ 0 & * & 0 & * \\ 0 & 0 & * & 0 \end{pmatrix}, \tag{3.44}$$

The graph is (3.44) is strongly connected as an oriented graph, since we can move from any point to any point following the arrows. Note that this is a stronger requirement than that imposed on U_5 in the previous subsection, where the graph had to be connected, without regard to orientation.

Let $R_5(t_0)$ in the appropriate basis correspond to a strongly connected oriented graph and introduce the notation M_i, $i = 1,...,n_1 + 2n_2$, for the matrices in (3.42). Then the matrices

$$X_{ik} = M_i U_5 M_k, \quad 1 \leqslant i,k \leqslant n_1 + 2n_2 \tag{3.45}$$

will be linearly independent and span $\mathbb{R}^{n\times n}$ if

$$X_{ik} \neq 0, \quad 1 \leqslant i,k \leqslant n_1 + 2n_2. \tag{3.46}$$

More generally, let $X_{ab} = 0$ and let $a\, j_1 j_2 \cdots j_l b$ be a permissible path from a to b on the oriented graph corresponding to U_5. Then we replace the missing element X_{ab} in the basis for $\mathbb{R}^{n\times n}$ by

$$Y_{ab} = X_{aj_1} X_{j_1 j_2} \cdots X_{j_l b}. \tag{3.47}$$

To avoid unnecessary complications, let us first assume that all entries in U_5 are nonzero,

$$(U_5)_{ik} = u_{ik} \neq 0, \tag{3.48}$$

in the basis where U_4 is in its Jordan canonical form (such matrices are dense in $\mathbb{R}^{n\times n}$). The superposition formula for the matrix anharmonic ratio R can then be written as

$$R = \sum_{j,k=1}^{n} a_{jk} R_4^{\,j} R_5 R_4^{\,k}. \tag{3.49}$$

The superposition formula for the MRE is obtained by using (3.2) for R_4 and R_5 in (3.49) and then substituting the expansion (3.49) for R into (3.19):

$$\begin{aligned} W = {} & [W_1(W_3 - W_1)^{-1}(W_2 - W_3)R + W_2] \\ & \times [(W_3 - W_1)^{-1}(W_2 - W_3)R + I]^{-1}. \end{aligned} \tag{3.50}$$

In the more general case when U_5 does not satisfy (3.48) in the basis where U_4 is given by (3.14), we write the superposition formula for the matrix cross ratios as

$$R = \sum_{i,k=1}^{n} a_{ik} C_{ik}, \tag{3.51}$$

where the matrices C_{ik} form a basis for the solutions of Eq. (3.31), defined by

$$C_{ik} = R_4^{\,i} R_5 R_4^{\,j_1} R_4^{\,j_1} R_5 R_4^{\,j_2} \cdots R_4^{\,j_l} R_5 R_4^{\,k}, \tag{3.52}$$

where $(i\, j_1 \cdots j_l k)$ is any permissible path on the graph corresponding to U_5.

IV. SYMPLECTIC MATRIX RICCATI EQUATIONS

If, instead of considering the most general equation associated with the action of $SL(2n,\mathbb{R})$ on $G_n(\mathbb{R}^{2n})$, we restrict ourselves to that associated with certain subgroups, particular forms of the matrix Riccati equation (1.1) arise which may be studied in greater detail. Since the subgroup will not generally act transitively on the entire space, it may be necessary to add further constraints characterizing particular orbits. The previously derived superposition rule will not necessarily respect such constraints, and therefore a new analysis may be necessary. Examples of such reductions are given below and a superposition rule is derived for each case.

A. The symplectic group $SP(2n,\mathbb{R}) \subset SL(2n,\mathbb{R})$

A particularly interesting reduction of Eq. (1.1) is obtained by requiring the curve $\xi(t)$ in Eq. (2.9) for $n = k$, to lie within the symplectic subalgebra defined by

$$\mathrm{sp}(2n,\mathbb{R}) = \{\xi \in \mathrm{sl}(2n,\mathbb{R}) \,|\, \xi K + K\xi^T = 0\}, \tag{4.1}$$

where the symplectic form on $\mathbb{R}^{2n}$ is defined by the matrix

$$K = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}. \tag{4.2}$$

The curve $\xi(t)$ thus has the form

$$\xi(t) = \begin{pmatrix} C(t) & A(t) \\ -D(t) & -C^T(t) \end{pmatrix}, \tag{4.3}$$

where

$$A(t) = A^T(t), \quad D(t) = D^T(t). \tag{4.4}$$

The corresponding curve $g(t)$ in $SL(2n,\mathbb{R})$, as defined by Eqs. (2.11)–(2.13), satisfies the condition

$$gKg^T = K \tag{4.5}$$

provided $g(t_0)$ does, implying that it lies within the symplectic subgroup $SP(2n,\mathbb{R})$. This subgroup no longer acts transitively on the full Grassmanian $G_n(\mathbb{R}^{2n})$, since it preserves the symplectic inner product. However, if we restrict ourselves to the submanifold $G_n^0(\mathbb{R}^{2n})$ of totally isotropic n planes in $\mathbb{R}^{2n}$, the so-called Lagrangian subspaces[5,6] with respect to the symplectic form K, it is easily verified that the group action is well defined and transitive. {These spaces comprise the orbit of minimal dimension $[\frac{1}{2}n(n+1)]$ and play an important role in the geometrical formulation of Hamiltonian dynamics and quantization.[12]} In terms of homogeneous coordinates $\binom{X}{Y}$ the condition of isotropy becomes

$$X^TY - Y^TX = 0, \tag{4.6}$$

which, in affine coordinates $W = XY^{-1}$, is equivalent to

$$W = W^T. \tag{4.7}$$

The reduced Riccati equation is

$$\dot W(t) = A(t) + W(t)C^T(t) + C(t)W(t) + W(t)D(t)W(t), \tag{4.8}$$

$$A = A^T, \quad D = D^T,$$

where the symmetry property

$$W(t) = W^T(t) \tag{4.9}$$

may be seen to persist for all t, provided it holds for $t = t_0$, as a consequence of Eq. (4.8) and the conditions (4.4).

In deriving a superposition rule for Eq. (4.8), we begin by following the same procedures as in the previous section. Given two known solutions $W_1(t)$ and $W_2(t)$, and an arbitrary one, $W(t)$, the quantity

$$T(t) = (W - W_1)^{-1}(W - W_2)(W_1 - W_2)^{-1} \tag{4.10}$$

satisfies the linear equation

$$\dot{T}(t) = -\tilde{C}^T T - T\tilde{C}, \tag{4.11}$$

where

$$\tilde{C}(t) = C(t) + W_1(t)D(t). \tag{4.12}$$

It follows that $T(t)$ develops in t according to

$$T(t) = G(t)T_0 G^T(t), \tag{4.13}$$

where $G(t) \in \mathrm{GL}(n,\mathbb{R})$ satisfies

$$\dot{G}(t) = -\tilde{C}^T G(t), \quad G(t_0) = G_0 \tag{4.14}$$

with some conveniently chosen initial condition G_0. Given a third solution $W_3(t)$ of Eq. (4.8) such that $T_3(t)$ defined correspondingly as

$$T_3(t) = (W_3 - W_1)^{-1}(W_3 - W_2)(W_1 - W_2)^{-1} \tag{4.15}$$

is nonsingular, the arbitrariness in the initial value G_0 implies that the symmetric matrix $T_3(t_0)$ may be assumed, without loss of generality, to be of the form

$$T_3(t_0) = G(t_0)I_{pq}G^T(t_0), \tag{4.16}$$

where

$$I_{pq} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}, \tag{4.17}$$

and hence for arbitrary t

$$T_3(t) = G(t)I_{pq}G^T(t). \tag{4.18}$$

Since $T_3(t)$ is symmetric, it may be diagonalized by an orthogonal transformation and therefore expressed as

$$T_3(t) = O_1(t)D(t)I_{pq}D(t)O_1^T(t) \tag{4.19}$$

where $O_1(t) \in O(n)$ is any orthogonal matrix of eigenvectors of $T_3(t)$ and $D(t)$ is the diagonal matrix diag $(|\lambda_1|^{1/2} \cdots |\lambda_n|^{1/2})$, the λ_i's being the corresponding eigenvalues. A certain nonuniqueness associated with the ordering of and rotations within the subspaces of equal eigenvalue is involved in (4.19), but this is of no importance, and we assume some choice of normalized eigenvectors is made. Now, defining

$$O_2(t) = D^{-1}(t)O_1^T(t)G(t), \tag{4.20}$$

it is easily verified that

$$O_2 I_{pq} O_2^T = I_{pq}; \tag{4.21}$$

that is, $O_2(t) \in O(p,q)$. Now, define the matrix

$$\tilde{T}(t) \equiv D^{-1}O_1^T T O_1 D^{-1} I_{pq}, \tag{4.22}$$

which is self-adjoint with respect to the inner product associated with the quadratic form I_{pq}, i.e.,

$$I_{pq}\tilde{T}^T = \tilde{T}I_{pq} \tag{4.23}$$

and $\tilde{T}(t)$ develops according to

$$\tilde{T}(t) = O_2(t)\tilde{T}_0 O_2^{-1}(t), \tag{4.24}$$

where

$$\tilde{T}_0 = T_0 I_{pq}. \tag{4.25}$$

Let $\tilde{T}_4(t)$ be the corresponding quantity associated with a fourth solution $W(t)$ of Eq. (4.8),

$$\tilde{T}_4(t) = D^{-1}(t)O_1^T(W_4 - W_1)^{-1} \times (W_4 - W_2)(W_1 - W_2)^{-1}O_1 D^{-1} I_{pq} \tag{4.26}$$

and assume that all its eigenvalues are distinct. These eigenvalues are independent of t because of Eq. (4.24). It follows from the self-adjointness (4.23) of $\tilde{T}_4(t)$ and the fact that all eigenvalues are assumed different that none of the eigenvectors are isotropic (zero length). Some of the eigenvalues and eigenvectors may be complex, but because $\tilde{T}_4(t)$ is real, these must come in complex conjugate pairs. Let $\{\mathbf{Q}_\alpha, \overline{\mathbf{Q}}_\alpha\}_{\alpha=1,\ldots,m}$ be the complex eigenvectors, with eigenvalues $\{\mu_\alpha, \bar{\mu}_\alpha\}_{\alpha=1,\ldots,m}$. These may be normalized so that

$$\mathbf{Q}_\alpha^T I_{pq}\mathbf{Q}_\alpha = \overline{\mathbf{Q}}{}^T_\alpha I_{pq}\overline{\mathbf{Q}}_\alpha = +1. \tag{4.27}$$

It follows from self-adjointness that, separating into real and imaginary parts,

$$\mathbf{Q}_\alpha = \frac{1}{\sqrt{2}}(\mathbf{p}_\alpha + i\mathbf{q}_\alpha), \tag{4.28}$$

the vectors $\{\mathbf{p}_\alpha, \mathbf{q}_\alpha\}$ are all mutually I_{pq} orthogonal with normalization:

$$\mathbf{p}_\alpha^T I_{pq}\,\mathbf{p}_\alpha = +1, \quad \mathbf{q}_\alpha^T I_{pq}\mathbf{q}_\alpha = -1 \tag{4.29}$$

The remaining $n - 2m$ real eigenvectors may be split into two sets $\{\mathbf{Q}_i^+, \mathbf{Q}_j^-\}$ with eigenvalues $\{\mu_i^+, \mu_j^-\}$ and normalization:

$$\begin{aligned} \mathbf{Q}_i^+ I_{pq}\mathbf{Q}_i^+ &= +1, \quad i = 1,\ldots,p-m, \\ \mathbf{Q}_j^- I_{pq}\mathbf{Q}_j^- &= -1, \quad j = 1,\ldots,q-m. \end{aligned} \tag{4.30}$$

Forming the matrix with these vectors as columns,

$$Q(t) = (\mathbf{Q}_1^+ \cdots \mathbf{Q}_{p-m}^+\ \mathbf{p}_1 \cdots \mathbf{p}_m \mathbf{q}_m \cdots \mathbf{q}_1, \mathbf{Q}_1^- \cdots \mathbf{Q}_{q-m}^-), \tag{4.31}$$

the orthonormality conditions following from the self-adjointness of T_4 imply

$$Q(t)I_{pq}Q^T(t) = I_{pq}, \tag{4.32}$$

that is, $Q(t) \in O(p,q)$. Moreover, since those columns of $Q(t)$ are the real and imaginary parts of the eigenvectors, we have

$$\tilde{T}_4(t) = Q(t)T_D Q^{-1}(t), \tag{4.33}$$

where

$$
T_{\mathrm{D}} = \begin{pmatrix}
\mu_1^+ & & & & & & & & & & & \\
& \ddots & & & & & & & & & & \\
& & \mu_{p-m}^+ & & & & & & & & & \\
& & & a_1 & & & & & b_1 & & & \\
& & & & \ddots & & & \iddots & & & & \\
& & & & & a_m & b_m & & & & & \\
& & & & & -b_m & a_m & & & & & \\
& & & & \iddots & & & \ddots & & & & \\
& & & -b_1 & & & & & a_1 & & & \\
& & & & & & & & & \mu_1^- & & \\
& & & & & & & & & & \ddots & \\
& & & & & & & & & & & \mu_{q-m}^-
\end{pmatrix} \tag{4.34}
$$

and (a_α, b_α) are the real and imaginary parts of μ_α (all nonindicated entries vanish). Using the remaining arbitrariness of the initial value $O_2(t_0)$ in Eq. (4.24), we may assume, without loss of generality, that

$$\widetilde{T}_0 = T_{\mathrm{D}}. \tag{4.35}$$

Comparing Eqs. (4.24) and (4.33), we find that

$$O_2(t) = Q(t)S, \tag{4.36}$$

where S is in the centralizer of T_{D} in $O(p,q)$. Since the eigenvalues are all distinct, S can only be of the form:

$$S = \mathrm{diag}(\pm 1, \pm 1, \ldots, \pm 1). \tag{4.37}$$

It follows that S must be constant provided all quantities vary smoothly in t and may therefore be assumed equal to 1 by absorbing its value in the arbitrary initial condition. Similarly the finite arbitrariness associated with the ordering of the eigenvectors is absorbed in the initial condition. Thus $O_2(t)$ is determined to equal the matrix $Q(t)$ formed from the normalized eigenvectors of $\widetilde{T}_4$. Combining the inverse of Eq. (4.10) with Eqs. (4.13), (4.15), (4.19), (4.20), and (4.24), we arrive at the following superposition formula, expressing a general solution $W(t)$ in terms of four particular solutions $W_1(t), \ldots, W_4(t)$:

$$
\begin{aligned}
W(t) = {} & [W_1 O_1 D Q T_0 + W_2 (W_2 - W_1)^{-1} O_1 D^{-1} I_{pq} Q I_{pq}] \\
& \times [O_1 D Q T_0 + (W_2 - W_1)^{-1} O_1 D^{-1} I_{pq} Q I_{pq}]^{-1}
\end{aligned} \tag{4.38}
$$

or

$$
\begin{aligned}
W(t) = {} & [W_2 (W_2 - W_1)^{-1} O_1 D^{-1} I_{pq} Q S_0 + W_1 O_1 D Q] \\
& \times [(W_2 - W_1)^{-1} O_1 D^{-1} I_{pq} Q I_{pq} S_0 + O_1 D Q]^{-1},
\end{aligned} \tag{4.39}
$$

where $O_1(t) \in O(n)$ is the orthonormalized matrix of eigenvectors of

$$T_3(t) = (W_3 - W_1)^{-1}(W_3 - W_2)(W_1 - W_2)^{-1} \tag{4.40}$$

with eigenvalues

$$
\begin{gathered}
\lambda_1(t), \ldots, \lambda_p(t) > 0 \quad \text{and} \quad \lambda_{p+1}(t), \ldots, \lambda_n(t) < 0, \\
D(t) = \mathrm{diag}(|\lambda_1(t)|^{1/2}, \ldots, |\lambda_n(t)|^{1/2}), \\
Q(t) \in O(p,q), \quad p + q = n
\end{gathered} \tag{4.41}
$$

is the matrix of Eq. (4.31), consisting of the real and imaginary parts of the eigenvectors of

$$
\begin{aligned}
\widetilde{T}_4(t) = {} & D^{-1}(t) O_1^T(t) [(W_4 - W_1)^{-1}(W_4 - W_2) \\
& \times (W_1 - W_2)^{-1}] O_1(t) D^{-1}(t) I_{pq}
\end{aligned} \tag{4.42}
$$

normalized as in Eqs. (4.27),...,(4.30); T_0 and S_0 are any constant, symmetric matrices. It is assumed that all the quantities $(W_2 - W_1)$, $(W_3 - W_1)$, $(W_4 - W_1)$, and $(W_3 - W_2)$ are nonsingular, and the eigenvalues of $\widetilde{T}_4$ are all distinct. The two forms (4.38) and (4.39) are given in order that all solutions be obtainable without involving infinite limits. Both forms are valid if $T_0 = S_0^{-1}$ is regular.

B. Subgroup $GL(n,\mathbb{R}) \subset SP(2n,\mathbb{R})$

Equation (4.8) may be specialized further by requiring the algebra element $\xi(t)$ to satisfy not only (4.1) but also

$$\xi J = J \xi, \tag{4.43}$$

where

$$J = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \tag{4.44}$$

which implies that it has the form

$$\xi(t) = \begin{pmatrix} C(t) & A(t) \\ A(t) & C(t) \end{pmatrix} \tag{4.45}$$

with

$$A = A^T, \quad C = -C^T. \tag{4.46}$$

The map

$$\varphi: \begin{pmatrix} C & A \\ A & C \end{pmatrix} \to A + C \in \mathrm{gl}(n,\mathbb{R})$$

defines a Lie algebra isomorphism with inverse determined by separating an arbitrary element in $\mathrm{gl}(n,\mathbb{R})$ into its symmetric and antisymmetric parts A and C, respectively. The curve $g(t)$ determined by Eq. (2.12) lies correspondingly in the $GL(n,\mathbb{R})$ subgroup of $SP(2n,\mathbb{R})$ defined by

$$gJ = Jg \tag{4.47}$$

provided the initial value $g(t_0)$ does. This implies that $g(t)$ is of the form

$$g(t) = \begin{pmatrix} M(t) & N(t) \\ N(t) & M(t) \end{pmatrix}, \tag{4.48}$$

where

$$M^TM - N^TN = 1, \quad M^TN = N^TM. \tag{4.49}$$

The reduced Riccati equation now becomes

$$\dot W = A + CW - WC - WAW,$$
$$A = A^T, \quad C = -C^T. \tag{4.50}$$

Assuming that $(W-1)$ is nonsingular, we can make a change of coordinates

$$W \to V \equiv (W+1)(W-1)^{-1}, \tag{4.51}$$

which is equivalent to transforming X and g into block diagonal form. This linearizes the equation directly:

$$\dot V = FV + VF^T, \tag{4.52}$$

where

$$F(t) = C(t) + A(t). \tag{4.53}$$

Since this is of the same form as Eq. (4.11), the procedure described there, after reducing the problem to this case, may be applied here, thereby determining the general solution, $W(t)$, to Eq. (4.50) in terms of two known solutions $W_1(t)$ and $W_2(t)$. Explicitly, we have

$$W(t) = (O_1DQV_0Q^TDO_1^T + 1)(O_1DQV_0Q^TDO_1^T - 1)^{-1}, \tag{4.54}$$

where $O_1(t) \in O(n)$ is the orthonormalized matrix of eigenvectors of $V_1(t)$ with eigenvalues $\lambda_1(t) \cdots \lambda_p(t) > 0$ and $\lambda_{p+1}(t),\ldots,\lambda_n(t) < 0$;

$$D(t) = \operatorname{diag}(|\lambda_1(t)|^{1/2},\ldots,|\lambda_n(t)|^{1/2}), \tag{4.55}$$
$$Q(t) \in O(p,q), \quad p+q=n,$$

is defined by Eq. (4.31) in terms of the eigenvectors $\{\mathbf{Q}_i^+, \mathbf{Q}_j^-, \mathbf{Q}_\alpha, \mathbf{Q}_\alpha^-\}$ of

$$\widetilde V(t) = D^{-1}(t) O_1^T(t) V_2(t) O_1(t) D^{-1}(t) I_{pq}, \tag{4.56}$$

normalized as in Eqs. (4.27)–(4.30), and V_0 is any constant symmetric matrix. Again, it is assumed that $V_1(t)$ is nonsingular and $\widetilde V(t)$ has n distinct eigenvalues with all eigenvectors nonisotropic with respect to I_{pq}.

C. Subgroup U(n)

An alternative reduction of the symplectic group is obtained by requiring, in addition to (4.1) and (4.5), the conditions

$$\xi K - K\xi = 0 \tag{4.57}$$

and

$$gK - Kg = 0. \tag{4.58}$$

This restricts ξ and g respectively to the subalgebra and subgroup consisting of elements of the form

$$\xi = \begin{pmatrix} C & A \\ -A & C \end{pmatrix}, \tag{4.59}$$

$$C = -C^T, \quad A = A^T, \tag{4.60}$$

and

$$g = \begin{pmatrix} M & N \\ -N & M \end{pmatrix}, \tag{4.61}$$

$$MM^T + NN^T = 1, \quad M^TN = N^TM. \tag{4.62}$$

These may be identified with the group U(n) of unitary matrices and its Lie algebra u(n) of anti-Hermitian matrices through the correspondance

$$g \leftrightarrow M + iN \in \mathrm{U}(n), \quad \xi \leftrightarrow C + iA \in \mathrm{u}(n). \tag{4.63}$$

It is also easily verified that this subgroup acts transitively on the Lagrangian subspaces, with isotropy group conjugate to O(n), thereby allowing the identification:

$$G_n^0(\mathbb{R}^n) \sim \mathrm{U}(n)/\mathrm{O}(n).$$

The Riccati equation associated with this action is thus:

$$\dot W = A(t) + C(t)W(t) - W(t)C(t) + W(t)A(t)W(t) \tag{4.64}$$

with

$$C(t) = -C^T(t), \quad A(t) = A^T(t). \tag{4.65}$$

Proceeding as in the previous example, a map taking $\xi(t)$ and $g(t)$ into block diagonal form is equivalent to the transformation

$$W \to V \equiv (W+i)(W-i)^{-1} \tag{4.66}$$

(valid for all W, since the eigenvalues are real). The resulting equation for V is

$$\dot V = FV + VF^T, \tag{4.67}$$

where

$$F = C - iA \in \mathrm{u}(n). \tag{4.68}$$

It follows that the general solution to (4.67) is of the form

$$V = UV_0U^T, \tag{4.69}$$

where

$U \in \mathrm{U}(n)$ satisfies

$$\dot U = (C - iA)U, \quad U(t_0) = U_0 \tag{4.70}$$

for some arbitrarily chosen U_0. Note that the definition (4.66) of V in terms of the symmetric real matrix W implies that it may be diagonalized by an orthogonal transformation to the form $\operatorname{diag}(e^{i\varphi_1},\ldots,e^{i\varphi_n})$. Moreover, a further transformation of the type (4.69) may be applied to map these phases to zero. It follows that given any particular solution $W_1(t)$ of Eq. (4.64) and the corresponding solution $V_1(t)$ to Eq. (4.67) defined through (4.66), the initial value $V_1(t_0)$ may, without loss of generality, be assumed to be of the form

$$V_1(t_0) = U(t_0)U^T(t_0), \tag{4.71}$$

and therefore, for arbitrary t,

$$V_1(t) = U(t)U^T(t)$$
$$= O_1(t)D^2(t)O_1^T(t), \tag{4.72}$$

where $O_1(t) \in O(n)$ is a matrix of orthonormalized eigenvectors of $V_1(t)$ [and hence also of $W_1(t)$], and $D^2(t)$ is the diagonal matrix of eigenvalues $\operatorname{diag}\{(\lambda_1 + i)/(\lambda_i - i),\ldots,(\lambda_n + i)/(\lambda_n - i)\}$, where $\lambda_j(t)$ are the eigenvalues of $W_1(t)$. Making an arbitrary choice of signs in defining $D(t)$, Eq. (4.72) implies that U(t) may be written in the form:

$$\mathrm{U}(t) = O_1(t)D(t)O_2(t), \tag{4.73}$$

where $O_2(t)\in O(n)$. A certain nonuniqueness is again involved in such a decomposition related to the ordering of eigenvectors and rotations among those of equal eigenvalues, but this is of no importance and we assume some choice has been made with all quantities varying smoothly in t. Let $W_2(t)$ be a second solution of (4.64) and

$$V_2(t) = [W_2(t) + i][W_2(t) - i]^{-1}. \tag{4.74}$$

Define

$$\begin{aligned}\widetilde{V}_2(t) &= D^{-1}O_1^T V_2 O_1 D^{-1}\\ &= (D^{-1}O_1^T W_2 O_1 D^{-1} + i)(D^{-1}O_1^T W_2 O_1 D^{-1} - i)^{-1}\\ &= O_2(t)V_2(t_0)O_2^T(t),\end{aligned} \tag{4.75}$$

where, without loss of generality, we can assume $V_2(t_0)$ to be diagonal. Assuming that the eigenvalues of $\widetilde{V}(t)$ are all distinct, which is valid for a dense subset of solutions, the orthogonal matrix is essentially uniquely determined to consist of the orthonormalized eigenvectors of $\widetilde{V}_2(t)$ (hence also of $D^{-1}O_1^T W_2 O_1 D^{-1}$). The nonuniqueness again consists of discrete factors of the type diag$(\pm 1,..., \pm 1)$ and the ordering of the eigenvectors. By the assumption of smoothness in t, these are all constants and may be absorbed in the initial conditions. Combining the inverse of Eq. (4.66) with Eqs. (4.69), (4.72–4.75), we obtain the following explicit superposition formula expressing the general solution $W(t)$ of Eq. (4.64) in terms of two known solutions $W_1(t)$ and $W_2(t)$:

$$W(t) = i\,\frac{O_1DO_2V_0O_2^TDO_1^T + 1}{O_1DO_2V_0O_2^TDO_1^T - 1}, \tag{4.76}$$

where $O_1(t)$ is the matrix of orthonormalized eigenvectors of $W_1(t)$, $D(t)$ is defined in terms of the corresponding eigenvalues $\{\lambda_i\}$ by

$$D^2(t) = \mathrm{diag}\left\{\frac{\lambda_1 + i}{\lambda_1 - i},..., \frac{\lambda_n + i}{\lambda_n - i}\right\}$$

and $O_2(t)$ is the orthonormal matrix of eigenvectors $D^{-1}O_1^T W_2 O_1 D^{-1}$, whose eigenvalues are assumed all distinct.

V. CONCLUSIONS

The main results of this paper are the following.

1. We have shown that the rectangular $n\times k$ matrix Riccati equation (1.1) is interpretable in terms of the infinitesimal action of the group $\mathrm{SL}(n + k,\mathbb{R})$ on the Grassman manifold $G_k(\mathbb{R}^{n+k})$. It follows that a superposition rule exists for the solutions of such equations.

2. For the case of square matrices ($n = k$), we have used this characterization to obtain two different versions of this superposition rule, expressing the general solution for arbitrary $n\geqslant 2$ in terms of only five known particular solutions. The first version is provided by Eqs. (3.38) together with (3.36) and (3.37), the second by Eq. (3.50) together with (3.2) and (3.49) or (3.51).

3. A case of particular interest is the symplectic matrix Riccati equation (4.8), for which we have obtained the superposition rule (4.38), (4.39), involving only four known particular solutions. Two subcases corresponding to the subgroups $\mathrm{GL}(n,\mathbb{R})\subset\mathrm{SP}(2n,\mathbb{R})$ and $\mathrm{U}(n)\subset\mathrm{SP}(2n,\mathbb{R})$ are Eqs. (4.50) and (4.64), respectively, for which we have the superposition rules (4.54) and (4.76), involving only two solutions each.

4. The five solutions forming a "fundamental set of solutions" in the $\mathrm{SL}(2n,\mathbb{R})$, as well as the four solutions in the $\mathrm{SP}(2n,\mathbb{R})$ case are generically chosen [their initial conditions form a dense open set in $[G_n(\mathbb{R}^{2n})]^5$ or $[G_n^0(\mathbb{R}^{2n})]^4$, respectively]. In practical calculations we can, of course, make a convenient choice of initial conditions and avoid complications due to complex eigenvalues of R_4 in the $\mathrm{SL}(2n,\mathbb{R})$ case or negative eigenvalues of T_3 in the $\mathrm{SP}(2n,\mathbb{R})$ one.

Such equations arise in particular in the study of Bäcklund transformations for the generalized nonlinear σ models.[13–15] The application of the present results to numerical studies of the matrix Riccati equations arising in optimal control theory is in preparation.

ACKNOWLEDGMENTS

The authors would like to thank H. Zassenhaus for elucidating the algebraic irreducibility requirements cited in Sec. III and A. Manitius, D. Rand, I. Rosenberg, and S. Shnider for numerous helpful discussions.

[1] R. L. Anderson, Lett. Math. Phys. 4, 1 (1980).

[2] R. L. Anderson, J. Harnad, and P. Winternitz, Lett. Math. Phys. 5, 143, (1981).

[3] R. L. Anderson, J. Harnad, and P. Winternitz, Physica D 4, 164 (1982).

[4] S. Lie and G. Sheffers, *Vorlesungenüber Continuierliche Gruppen* (B. G. Teubner, Leipzig, 1893).

[5] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1978).

[6] R. A. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin Cummings, Reading, Mass., 1978).

[7] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory* (Wiley, New York, 1967).

[8] W. T. Reid, *Riccati Differential Equations* (Academic, New York, 1972).

[9] C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Wiley, New York, 1962).

[10] W. Greub, *Linear Algebra* (Springer, New York, 1975).

[11] M. Sorine and P. Winternitz, to be published.

[12] A. Weinstein, *Lectures on Symplectic Manifolds*, CBMS Conference Series No. 29 (Am. Math. Soc., Providence, R.I., 1977).

[13] A. T. Ogielski, M. K. Prasad, A. Sinha, and L. L. Chau Wang, Phys. Lett. B 91, 387 (1980).

[14] J. Harnad, Y. Saint-Aubin, and S. Shnider, "Superposition of solutions to Bäcklund transformations for the U(n) principal sigma model," CRMA-1074 (Centre de Recherche de Mathématiques Appliquées, Université de Montreal, 1982).

[15] J. Harnad, Y. Saint-Aubin, and S. Shnider, "Quadratic pseudopotentials for GL(n,C) principal sigma models," CRMA-1075 (Centre de Recherche de Mathématiques Appliquées, Université de Montreal, 1982).

Restricted multiple three-wave interactions: Integrable cases of this system and other related systems

C. R. Menyuk, H. H. Chen, and Y. C. Lee
Laboratory for Plasma and Fusion Energy Studies, University of Maryland, College Park, Maryland 20742

(Received 5 October 1982; accepted for publication 6 January 1983)

Restricted multiple three-wave interactions, in which a set of wave triads interact through one shared wave, are discussed. It is shown that this system is integrable when all triads have equal coupling coefficients regardless of the frequency mismatches. This system is then used as a starting point from which to determine integrable cases of a more general class of three-wave interactions.

PACS numbers: 03.20. + i

I. INTRODUCTION

This paper is the second in a series of three papers devoted to determining under what circumstances restricted multiple three-wave interactions may be treated statistically. A minimal condition is that the system be nonintegrable. In the first paper,[1] Painlevé analysis was used to help find integrable cases, and numerical evidence was provided, indicating that when the system has the Painlevé property of possessing only simple movable poles in the complex time plane, the system is integrable and otherwise it is not.

From a mathematical point of view, this system appears to be the simplest possible multiply interacting three-wave system with an arbitrary number of triads and, as such, can be analyzed in detail. Thus, this system provides a useful platform from which to attack certain aspects of more general multiply interacting systems, much as the restricted three-body problem provides a useful platform from which to attack the full three-body problem.[2]

Restricted multiple three-wave interactions are those in which an arbitrarily large set of three-wave triads interact with each other through one wave which they all share. It has already been shown in Ref. 1 that these interactions are described by the Hamiltonian

$$H = \sum_{n=1}^{N} \left[\tfrac{1}{2} \Delta_n J_n - \tfrac{1}{2} \Delta_n J'_n - \epsilon_n (J_0 J_n J'_n)^{1/2} \times \cos(\theta_n - \theta'_n - \theta_0) \right], \tag{1.1}$$

where J_0 and θ_0 are the action-angle variables of the shared wave J_n, θ_n, J'_n, and θ'_n are the action-angle variables of the two other members of the nth triad, $\Delta_n \equiv \omega_n - \omega'_n - \omega_0$ is the frequency mismatch, ϵ_n is the coupling coefficient, and N, the number of triads, may be arbitrarily large. We shall assume $\epsilon_n > 0$, for if any of these coupling coefficients are less than zero, we need merely add π to θ_n to reverse that coupling coefficient's sign. This system appears whenever one has a "test wave" interacting conservatively with a spectrum of otherwise noninteracting waves, and was first used by Watson, West, and Cohen[3] to model the growth of a low frequency internal ocean wave due to its interaction with a spectrum of higher frequency surface waves. Meiss[4] has studied this system using a variety of statistical assumptions and proposed it as a model for both ocean wave and plasma turbulence. Paradoxically, he has also suggested that this system may be completely integrable, a conjecture which was disproved in Ref. 1. In the third paper of this series, we discuss the statistical assumptions and show that they too are not generally valid.

In this paper, we are entirely concerned with demonstrating the integrability of the restricted system, as well as a much larger class of systems, in certain special cases. In Sec. II, we demonstrate that the restricted system is integrable when the coupling coefficients are all equal, but the frequency mismatches are arbitrary. We do so by first taking the apparently complicating step of converting our ordinary differential equations (ODE) into partial differential equations (PDE), and then using a formalism developed by Ablowitz and Haberman[5] to, in effect, find a Lax pair for the system. The system was previously demonstrated to be integrable by Meiss[6] in the simpler case where all coupling coefficients and all mismatches are equal using a trial-or-error approach. In Sec. III, we use the method of asymptotic expansion to extract the required number of integrals of the motion, mutually in involution. The approach is similar to that used by Haberman[7] and many other authors. In Sec. IV, we show that the restricted system may be made integrable, no matter what the coupling coefficients, by adding further waves with appropriate coupling coefficients, and, hence, the restricted system is just one of a large class of systems which can be made integrable by an appropriate choice of coupling coefficients. In Sec. V, we show that in the special case where all the coupling coefficients are equal and all the frequency mismatches are equal, the restricted system can be explicitly integrated in terms of elliptic integrals and their quadratures. Finally, Sec. VI contains a summary.

II. DETERMINATION OF LAX PAIRS FOR THE RESTRICTED SYSTEM WHEN ALL COUPLING COEFFICIENTS ARE EQUAL

Letting $b_0 \equiv (J_0)^{1/2} \exp(-i\theta_0)$, $b_n \equiv (J_n)^{1/2} \exp(-i\theta_n)$, and $b'_n \equiv (J'_n)^{1/2} \exp(-i\theta'_n)$, we find that Eq. (1.1) generates the equations of motion

$$\begin{aligned} \dot{b}_0 &= \tfrac{1}{2} i \sum_{n=1}^{N} \epsilon_n b_n b'^{*}_n, \\ \dot{b}_n &= -\tfrac{1}{2} i \Delta_n b_n + \tfrac{1}{2} i \epsilon_n b_0 b'_n, \\ \dot{b}'_n &= \tfrac{1}{2} i \Delta_n b'_n + \tfrac{1}{2} i \epsilon_n b^{*}_0 b_n, \end{aligned} \tag{2.1}$$

 0022-2488/83/051073-07$02.50

where $n = 1,...,N$, which, from now on, will be understood. We may convert Eq. (2.1) into a PDE by writing

$$b_{0,t} = \gamma_0 b_{0,x} + \tfrac{1}{2} i \sum_{n=1}^{N} \epsilon_n b_n b_n'^{*},$$
$$b_{n,t} = \gamma_n b_{n,x} + \tfrac{1}{2} i\epsilon_n b_0 b_n', \tag{2.2}$$
$$b'_{n,t} = \gamma'_n b'_{n,x} + \tfrac{1}{2} i\epsilon_n b_0^{*} b_n,$$

where b_0, b_n, and b'_n are now functions of both x and t, γ_0, γ_n, and γ'_n are constants, and the subscripts x and t indicate differentiation with respect to x and t. From Eq. (2.2), we may return to Eq. (2.1) by imposing

$$b_0(x,t) = b_0(t),$$
$$b_n(x,t) = b_n(t)\exp(-i\Delta_n x/2\gamma_n), \tag{2.3}$$
$$b'_n(x,t) = b'_n(t)\exp(i\Delta_n x/2\gamma'_n),$$

as well as

$$\gamma'_n = -\gamma_n. \tag{2.4}$$

Equation (2.4) is required if the x variation is to remain unaltered for all time.

Following Ablowitz and Haberman,[5] we now consider the differential matrix eigenvalue problem

$$\psi_x = i\zeta \mathsf{D}\psi + \mathsf{N}\psi, \tag{2.5}$$

where ψ is an M-dimensional vector eigenfunction, ζ is the eigenvalue, and D and N are $M \times M$ matrices. We choose the time variation of ψ such that

$$\psi_t = \mathsf{Q}\psi. \tag{2.6}$$

If we now require that

$$\mathsf{Q}_x = \mathsf{N}_t + i\zeta[\mathsf{D},\mathsf{Q}] + [\mathsf{N},\mathsf{Q}], \tag{2.7}$$

where the square brackets indicate commutation, it immediately follows that $\zeta_t = 0$.

Next, we choose Q to have the form

$$\mathsf{Q} = \mathsf{Q}^{(1)} + \zeta \mathsf{Q}^{(0)}, \tag{2.8}$$

where $\mathsf{Q}^{(0)}$ and $\mathsf{Q}^{(1)}$ are independent of ζ. Expanding Eq.(2.7) in powers of ζ, we then find

$$0 = i[\mathsf{D},\mathsf{Q}^{(0)}], \tag{2.9a}$$
$$\mathsf{Q}^{(0)}_x = i[\mathsf{D},\mathsf{Q}^{(1)}] + [\mathsf{N},\mathsf{Q}^{(0)}], \tag{2.9b}$$
$$\mathsf{Q}^{(1)}_x = \mathsf{N}_t + [\mathsf{N},\mathsf{Q}^{(1)}]. \tag{2.9c}$$

We further choose $D_{ij} = \delta_{ij} a_j$ and $Q^{(0)}_{ij} = \delta_{ij} c_j$, where a_j and c_j are assumed real, so that Eq. (2.9a) is automatically satisfied and $Q^{(0)}_x = 0$. From Eq. (2.9b) we find

$$Q^{(1)}_{ik} = [(c_i - c_k)/i(a_i - a_k)]N_{ik}, \quad \text{for } i \neq k, \tag{2.10}$$

where $a_i \neq a_k$. If $a_i = a_k$ and $c_i = c_k$, then any choice of $Q^{(1)}_{ik}$ and N_{ik} is consistent with Eq.(2.9b), and we shall choose $Q^{(1)}_{ik} = \alpha_{ik} N_{ik}$, where α_{ik} is an arbitrary constant. The diagonal entries of Eq. (2.9c) yield

$$Q^{(1)}_{ii,x} = N_{ii,t}, \tag{2.11}$$

from which we conclude that we are free to choose

$$Q^{(1)}_{ii} = N_{ii} = 0. \tag{2.12}$$

Doing so, the off-diagonal elements of Eq. (2.9c) yield

$$\alpha_{ik} N_{ik,x} = N_{ik,t} + \sum_{j \neq i,k} N_{ij} N_{jk} (\alpha_{jk} - \alpha_{ij}), \tag{2.13}$$

where

$$\alpha_{ik} = (c_k - c_i)/i(a_k - a_i) \tag{2.14}$$

if $a_i \neq a_k$, and is otherwise arbitrary. In order for Eq. (2.13) to represent a three-wave interaction, we must have

$$N_{jk} = \sigma_{jk} N^{*}_{kj}, \quad \text{for } j > k, \tag{2.15}$$

where the σ_{jk} are real normalizing coefficients. In order for the ikth equation to be equivalent to the kith equation, we must have

$$\sigma_{jk}\sigma_{ij} = -\sigma_{ik}, \quad \text{for } i > j > k. \tag{2.16}$$

Equation (2.13) may now be written

$$\alpha_{ik} N_{ik,x} = N_{ik,t} + \sum_{j>k>i} \sigma_{jk} N_{ij} N^{*}_{kj} (\alpha_{kj} - \alpha_{ij})$$
$$+ \sum_{k>j>i} N_{ij} N_{jk} (\alpha_{jk} - \alpha_{ij})$$
$$+ \sum_{k>i>j} \sigma_{ij} N^{*}_{ji} N_{jk} (\alpha_{jk} - \alpha_{ji}). \tag{2.17}$$

Equation (2.17) was first obtained by Ablowitz and Haberman.[5]

Equation (2.17) possesses, by construction, an infinite number of constants of the motion, namely the eigenvalues of Eq. (2.5), and hence may be deemed integrable. The task at hand is to find under what conditions Eq. (2.2) may be written in the form of Eq. (2.17). To do so, we make the following identifications:

$$N_{12} = ib_0,$$
$$N_{1,n+2} = ib_n/\sqrt{2},$$
$$N_{2,n+2} = ib'_n/\sqrt{2}, \tag{2.18}$$
$$N_{ln} = 0, \quad l = 3,4,...,N+2, \ n > l.$$

From the equations for N_{12}, N_{1n}, and N_{2n} we then find

$$\sigma_{n+2,2}(\alpha_{2,n+2} - \alpha_{1,n+2}) = \epsilon_n,$$
$$(\alpha_{2,n+2} - \alpha_{12}) = -\epsilon_n/2, \tag{2.19}$$
$$\sigma_{21}(\alpha_{1,n+2} - \alpha_{12}) = \epsilon_n/2,$$

which may be consistently satisfied by the choice

$$(\alpha_{2,n+2} - \alpha_{1,n+2}) = -\epsilon_n, \quad (\alpha_{2,n+2} - \alpha_{12}) = -\epsilon_n/2,$$
$$(\alpha_{1,n+2} - \alpha_{12}) = \epsilon_n/2, \quad \sigma_{n+2,2} = -1, \quad \sigma_{21} = 1. \tag{2.20}$$

Using Eqs. (2.2) and (2.16), one further deduces

$$\alpha_{1,n+2} = \gamma_n = -\gamma'_n = -\alpha_{2,n+2} = \epsilon_n/2,$$
$$\alpha_{12} = \gamma_0 = 0, \quad \sigma_{n+2,1} = 1. \tag{2.21}$$

Referring now to the equations for N_{ln} ($l = 3,...,N+2, n > l$) in Eq. (2.17), we see that in order to have $N_{ln} = 0$ for all time, i.e., in order to have $N_{ln,t} = 0$, we must have

$$\alpha_{jn} - \alpha_{jl} = 0, \tag{2.22}$$

for $j < l < n$. Setting $j = 1$, we find from Eq. (2.21)

$$\epsilon_1 = \epsilon_2 = \cdots = \epsilon_N, \tag{2.23}$$

so that Eq. (2.2) can be put in the form of Eq. (2.17) only if all the coupling coefficients are equal.

Setting for convenience $\epsilon_1 = \epsilon_2 = \cdots = \epsilon_N = 2$, Eq. (2.2) takes on the form

$$b_{0,t} = i \sum_{n=1}^{N} b_n b_n'^{*},$$
$$b_{n,t} = b_{n,x} + i b_0 b_n', \tag{2.24}$$
$$b_{n,t}' = -b_{n,x}' + i b_0^{*} b_n.$$

The quantities $a_1, a_2, \ldots, a_{N+2}$ and $c_1, c_2, \ldots, c_{N+2}$ may now be determined. Since $\alpha_0 = 0$, we find $c_1 = c_2$, and using the relations

$$\alpha_{1n} = -\alpha_{2n} = \frac{c_n - c_1}{i(a_n - a_1)} = -\frac{c_n - c_2}{i(a_n - a_2)},$$
$$n = 3, \ldots, N+2, \tag{2.25}$$

we conclude

$$a_3 = a_4 = \cdots = a_{N+2} = (a_1 + a_2)/2 \tag{2.26}$$

and, letting $c_1 = c_2 \equiv c$,

$$c_3 = c_4 = \cdots = c_{N+2} = ic + i(a_2 - a_1)/2. \tag{2.27}$$

The quantities a_1, a_2, and c are arbitrary, except that $a_1 = a_2$ is not allowed.

We now have all the information needed to construct D, Q, and N, and from them the standard Lax pairs, L and A. Using Eqs. (2.5) and (2.6) we see that

$$\mathsf{L} = (\mathsf{D}^{-1}\partial_x - \mathsf{D}^{-1}\mathsf{N}),$$
$$\mathsf{A} = (-i\mathsf{D}^{-1}\mathsf{Q}^{(0)}\partial_x + i\mathsf{D}^{-1}\mathsf{N}\mathsf{Q}^{(0)} + \mathsf{Q}^{(1)}). \tag{2.28}$$

Given Eq. (2.3), it is easy to see that the eigenvalue equation

$$\mathsf{L}\psi = i\zeta\psi \tag{2.29}$$

has at any given time the form of a generalized Hill equation in which the periodicities need not be commensurable. In the special case where $\Delta_1 = \Delta_2 = \cdots = \Delta_N$, Eq. (2.29) simplifies enormously, becoming an equation with constant coefficients. In this case, it is well known that the solution of Eq. (2.29) is completely characterized by the eigenvalues and eigenvectors of matrix $\tilde{\mathsf{L}}$, where the operator $\mathsf{I}\partial_x$ is removed from L. We have previously shown that, in this case, the matrix $\tilde{\mathsf{L}}$, can be used, after some modification, to determine the constants of the motion and that it can be obtained without resorting to a PDE.[8] However, in the case where the frequency mismatches are not zero, Eq. (2.29) is considerably more difficult to analyze, and the method of asymptotic expansion of the PDE used in Sec. II appears to be the simplest route to determining the constants of the motion.

The reader will have noted that Eq. (2.23) was a consequence of demanding that any waves not in the restricted system be zero for all time. Relaxing this restriction, to allow more waves into the system, will lead us to integrable cases of a more general system as will be discussed in Sec. IV.

III. CONSTANTS OF THE MOTION

In the limit $\zeta \to \infty$, Eqs. (2.5) and (2.6) reduce to

$$\psi_x = i\zeta\mathsf{D}\psi, \quad \psi_t = \zeta\mathsf{Q}^{(0)}\psi, \tag{3.1}$$

which has for its general solution any superposition of the particular solutions

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \exp(i\zeta a_1 x + \zeta c_1 t), \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \exp(i\zeta a_2 x + \zeta c_2 t), \tag{3.2}$$

etc. If we now consider the solution of Eqs. (2.5) and (2.6) at large but not infinite ζ, we find that their solution may be expressed as[9]

$$\begin{bmatrix} 1 \\ A_0(\zeta,x,t) \\ A_1(\zeta,x,t) \\ \vdots \\ A_N(\zeta,x,t) \end{bmatrix} \times \exp\left[i\zeta a_1 x + \zeta c_1 t + \int_0^x T(\zeta,x',t)\,dx'\right], \tag{3.3}$$

$$\begin{bmatrix} B_0(\zeta,x,t) \\ 1 \\ A_1(\zeta,x,t) \\ \vdots \\ A_N(\zeta,x,t) \end{bmatrix} \times \exp\left[i\zeta a_2 x + \zeta c_2 t + \int_0^x T(\zeta,x',t)\,dx'\right],$$

etc., where each of these quantities may be expanded asymptotically in ζ,

$$T(\zeta,x,t) = \sum_{j=0}^{\infty} \zeta^{-j} T_j(x,t),$$
$$\mathbf{A}(\zeta,x,t) = \sum_{j=1}^{\infty} \zeta^{-j} \mathbf{A}_j(x,t), \tag{3.4}$$
$$B_0(\zeta,x,t) = \sum_{j=1}^{\infty} \zeta^{-j} B_{0,j}(x,t).$$

The functional dependence of T, $\mathbf{A}$, and B_0 is, of course, different in each solution. The integrands inside the exponentials of Eq. (3.3) have 0 as their lower limit, rather than $-\infty$ as was the case in Ref. 9, because $T(\zeta,x,t)$ contains a portion which is constant in x.

Substituting Eqs. (3.3) and (3.4) into Eqs. (2.5) and (2.6), we will obtain for each solution a nonlinear recursion relation of the form

$$P_L(\mathbf{b}, T_j(\mathbf{b}), \mathbf{A}_j(\mathbf{b}), B_{0,j}(\mathbf{b}), x, t) = 0, \tag{3.5a}$$
$$P_A(\mathbf{b}, T_j(\mathbf{b}), \mathbf{A}_j(\mathbf{b}), B_{0,j}(\mathbf{b}), x, t) = \frac{d}{dt}\int_0^x T_j(\mathbf{b})\,dx', \tag{3.5b}$$

where $\mathbf{b} = (b_0, b_0^{*}, b_1, b_1', b_1^{*}, b_1'^{*}, \ldots, b_N, b_N', b_N^{*}, b_N'^{*})$ is the set of variables we are considering. P_L and P_A are polynomial functions of $\mathbf{b}$ and its higher derivatives and are therefore quasiperiodic in x. Dividing Eq. (3.5b) by x and taking the limit as $x \to \infty$, we find

$$0 = \frac{d}{dt}\left[\lim_{x\to\infty} \frac{1}{x}\int_0^x T_j(\mathbf{b})\,dx'\right], \tag{3.6}$$

which is a conservation law. The $T_j(\mathbf{b})$ may be determined from Eq. (3.5a). It turns out that they are independent of x, and, hence, just the constants of the motion we are seeking.

We study first the solution

$$\begin{bmatrix} 1 \\ A_0 \\ A_1 \\ \vdots \\ A_N \end{bmatrix} \exp\left(i\zeta a_1 x + i\zeta c_1 t + \int_0^x T\,dx'\right),$$

which, when substituted into Eq. (2.5), yields the recursion relations

$$T_j = ib_0 A_{0,j} + \frac{i}{\sqrt{2}} \sum_{l=1}^{N} A_{l,j} b_l, \tag{3.7a}$$

$$i(a_2 - a_1) A_{0,j+1} - \partial_x A_{0,j} - \sum_{i<j} A_{0,j-i} T_i$$
$$= ib_0^* \delta_{j,0} - \frac{i}{\sqrt{2}} \sum_{l=1}^{N} b_l' A_{l,j}, \tag{3.7b}$$

$$i(a_{n+2} - a_1) A_{n,j+1} - \partial_x A_{n,j} - \sum_{i<j} A_{n,j-i} T_i$$
$$= \frac{ib_n^*}{\sqrt{2}} \delta_{j,0} - \frac{ib_n'^*}{\sqrt{2}} A_{0,j}. \tag{3.7c}$$

We choose for convenience $i(a_2 - a_1) = 2$ and $i(a_{n+2} - a_1) = 1$, a choice which is consistent with the constraint $a_{n+2} = (a_1 + a_2)/2$. From Eq. (3.7a), we have $T_0 = 0$. Equations (3.7b) and (3.7c) start the iteration process, yielding $A_{0,1} = ib_0^*/2$ and $A_{n,1} = ib_n^*/\sqrt{2}$. Substituting these results into Eq. (3.7a) yields T_1, and we may repeat the process as many times as we wish to obtain T_2, T_3, etc. The first two iterates of this process are

$$T_1 = -\frac{1}{2}\left(b_0 b_0^* + \sum_{n=1}^{N} b_n b_n^*\right), \tag{3.8a}$$

$$T_2 = \frac{1}{4} i \left[\sum_{n=1}^{\infty} (-\Delta_n b_n b_n^* + b_0 b_n^* b_n' + b_0^* b_n b_n'^*)\right], \tag{3.8b}$$

where we have used Eqs. (2.3) and (2.21) to set

$$\partial_x b_n^* = \tfrac{1}{2} i\Delta_n b_n^* \tag{3.9}$$

in Eq. (3.8b). Equation (3.8a) is one of the Manley–Rowe relations for this system, and Eq. (3.8b) is related to the Hamiltonian (1.1).

We could use our recursion relations to obtain further constants. However, it is not difficult to show by induction that the constants obtained in this way are polynomials in the six combinations

$$b_0 b_0^*, \quad \sum_{n=1}^{N} b_n b_n^*, \quad \sum_{n=1}^{N} b_n' b_n'^*,$$
$$\sum_{n=1}^{N} b_0 b_n^* b_n', \quad \sum_{n=1}^{N} b_0^* b_n b_n'^*, \quad \left|\sum_{n=1}^{N} b_n^* b_n'\right|^2, \tag{3.10}$$

from which it follows immediately that we cannot obtain the full complement of necessary constants from just the single solution $B_0 = 1$, $\mathbf{A} \to 0$ as $\zeta \to \infty$, which we have been considering.

We now consider the solutions for which $B_0 \to 0$, $A_{l\neq n} \to 0$, $A_n = 1$ as $\zeta \to \infty$, where $n = 1,\ldots,N$. The recursion relations become

$$T_j = (-ib_n^*/\sqrt{2}) B_{0,j} + (ib_n'^*/\sqrt{2}) A_{0,j}, \tag{3.11a}$$

$$-B_{0,j+1} + \partial_x B_{0,j} + \sum_{i<j} B_{0,j-i} T_i$$
$$= ib_0 A_{0,j} + (i/\sqrt{2}) b_n \delta_{j,0} + (i/\sqrt{2}) \sum_{l\neq n} b_l A_{l,j}, \tag{3.11b}$$

$$A_{0,j+1} + \partial_x A_{0,j} + \sum_{i<j} A_{0,j-i} T_i$$
$$= -ib_0^* B_{0,j} + (i/\sqrt{2}) b_n' \delta_{j,0} + (i/\sqrt{2}) \sum_{l\neq n} b_l' A_{l,j}, \tag{3.11c}$$

$$\partial_x A_{l,j} + \sum_{i<j} A_{l,j-i} T_i$$
$$= -(i/\sqrt{2}) b_l^* B_{0,j} + (i/\sqrt{2}) b_l'^* A_{0,j} \quad (l \neq n), \tag{3.11d}$$

where we have set $a_n - a_2 = a_1 - a_n = 1$. The iteration proceeds as follows. Using Eqs. (3.11b) and (3.11c), we find $B_{0,1} = (-i/\sqrt{2}) b_n$ and $A_{0,1} = (i/\sqrt{2}) b_n'$. Using Eq. (3.11a), we then immediately find the N constants

$$T_1 = \tfrac{1}{2}(b_n b_n^* + b_n' b_n'^*). \tag{3.12}$$

Moving to Eq. (3.11d), we obtain

$$A_{l,1} = -\tfrac{1}{2}\partial_x^{-1}(b_l^* b_n + b_l'^* b_n'), \tag{3.13}$$

which becomes, assuming $\Delta_l \neq \Delta_n$ for $l \neq n$,

$$A_{l,1} = [i/(\Delta_l - \Delta_n)](b_l^* b_n + b_l'^* b_n'). \tag{3.14}$$

Using Eqs. (3.11b) and (3.11c) to obtain $A_{0,2}$ and $B_{0,2}$ and substituting the result into Eq. (3.11a), we find the N additional constants

$$T_2 = \tfrac{1}{2} i \Big\{ \tfrac{1}{2}\Delta_n b_n b_n^* - \tfrac{1}{2}\Delta_n b_n' b_n'^* - b_0 b_n^* b_n'$$
$$- b_0 b_n b_n'^* + \sum_{l\neq n} \left[|b_l b_n^* + b_l' b_n'^*|^2/(\Delta_n - \Delta_l) \right] \Big\}. \tag{3.15}$$

These N constants when all added together produce a multiple of the Hamiltonian. Hence, of the $2N + 2$ constants in Eqs. (3.8), (3.12), and (3.15), only $2N + 1$ are independent. All these constants are in involution, and as a result provide the full complement of $2N + 1$ independent constants in involution needed to demonstrate integrability.

In the case where some of the frequency mismatches are equal, the constants of Eq. (3.15) become singular, and one may proceed as follows to obtain N mutually independent, nonsingular constants, corresponding to those of Eq. (3.15): Let us suppose we have a set of triads such that $\Delta_1 = \Delta_2 = \cdots = \Delta_M \equiv \Delta$, where $M \leqslant N$. Let us also designate the T_2 corresponding to n, $T_2^{(n)}$. We obtain one nonsingular constant by adding all the $T_2^{(n)}$ $(1 \leqslant n \leqslant M)$ to obtain

$$I^{(1)} = \tfrac{1}{2} i \sum_{n=1}^{M} \Big\{ \tfrac{1}{2}\Delta b_n b_n^* - \tfrac{1}{2}\Delta b_n' b_n'^* - b_0 b_n^* b_n'$$
$$- b_0^* b_n b_n'^* + \sum_{l\neq 1,2,\ldots,M} \left[|b_l b_n^* + b_l' b_n'^*|^2/(\Delta - \Delta_l) \right] \Big\}. \tag{3.16}$$

To obtain further nonsingular constants, we set $\Delta_n = \Delta + \epsilon u_n$ $(1 \leqslant n \leqslant M)$, where the u_n's are constants, chosen so that $u_l \neq u_n$ for $l \neq n$, but where they are otherwise arbitrary. Substituting these choices of Δ_n into Eq. (3.15) and extracting the most singular term in ϵ, we find the $M-1$ mutually independent, nonsingular constants

$$I^{(n)} = \sum_{l \neq n, l \leqslant M} \frac{|b_l b_n^* + b_l' b_n'^*|^2}{u_n - u_l} \quad (2 \leqslant n \leqslant M). \tag{3.17}$$

We may proceed in this fashion with each set of triads which have the same frequency mismatch to finally obtain N mutually independent, nonsingular constants $I^{(1)}, I^{(2)}, \ldots, I^{(N)}$. Evidently, if $\Delta_n \neq \Delta_l$ for all $l \neq n$, $I^{(n)} = T_2^{(n)}$. This set of constants can be verified to be mutually in involution and in involution with the constants of Eqs. (3.8) and (3.12).

IV. INTEGRABLE CASES OF A MORE GENERAL SYSTEM

In this section, we are interested in systems whose equations of motion may be written in the form

$$\dot{N}_{ik} = -\gamma_{ik} N_{ik} + \sum_j c_{ijk} N_{ij} N_{jk}, \tag{4.1}$$

where i, j, and k are all integers, $i \neq j \neq k$, N_{ik} is one of the complex wave amplitudes, and γ_{ik} and c_{ijk} are arbitrary. We suppose that the wave N_{ik} has a wave vector $\mathbf{k}_{ik}$ and $N_{ik} = \sigma_{ik} N_{ki}^*$, where $\sigma_{ik} = \pm 1$, so that $\mathbf{k}_{ik} = -\mathbf{k}_{ki}$.

Equation (4.1) is similar to Eq. (2.13). We may obtain an equation of the form of Eq. (4.1) by demanding that the spatial variation be exponential in Eq. (2.13). That is to say, we must have

$$N_{ik}(x,t) = N_{ik}(t) \exp(\Delta_{ik} x). \tag{4.2}$$

The quantity Δ_{ik} may be purely imaginary as was the case in Sec. II, where it corresponded to a frequency mismatch. We can also imagine cases where it is real, in which case it corresponds to dissipation, or where it is complex, in which case it corresponds to a combination of a frequency mismatch and dissipation. In all cases, the consistency condition

$$\Delta_{ik} = \Delta_{ij} + \Delta_{jk} \tag{4.3}$$

must be met, and we have $\Delta_{ik} = \Delta_{ki}^*$.

To generate an integrable equation of the same form as Eq. (4.1), we begin with a case of the restricted system with arbitrary coupling coefficients. We recall from Sec. II that we have assumed a change of variables such that no frequency mismatch or dissipation exist for the shared wave. In the former case, that implies going into a moving frame. In the latter case, that implies a scale change. From Eq. (2.21), we have

$$\alpha_{12} = 0, \quad \alpha_{1,n+2} = -\alpha_{2,n+2} = \epsilon_n/2. \tag{4.4}$$

All the other α_{ij} may be chosen arbitrarily. Having chosen $\Delta_{12} = 0$ and $\Delta_{1,n+2}$ arbitrarily, all the other Δ_{ij} are fixed by the consistency condition (4.3). Finally, using Eq. (2.16) as well as Eqs. (2.20) and (2.21), one finds $\sigma_{ij} = -1$, for both i and $j > 2$.

It follows immediately that the restricted system with arbitrary coupling coefficients and mismatches can always be made integrable by introducing additional waves with appropriate coupling coefficients. Moreover, these coupling coefficients are to a large degree arbitrary because of the arbitrariness in the α_{ij} for both i and $j > 2$.

It is perhaps worth emphasizing that the existence of integrable cases for dissipative systems of special parameter values is not inconsistent with the existence of chaotic solutions and attractors at most parameter values. For example, the Lorenz system, which is integrable in certain special cases, also possesses strange attractor solutions.[10]

We now consider as an example the two triad case with frequency mismatches, but no dissipation. In this case, before any additional waves are introduced, the equations of motion, obtained from Eq. (2.17) are

$$\begin{aligned}
\dot{N}_{12} &= -\epsilon_1 N_{13} N_{23}^* - \epsilon_2 N_{14} N_{24}^*, \\
\dot{N}_{13} &= -\tfrac{1}{2} i \Delta_1 N_{13} + \tfrac{1}{2}\epsilon_1 N_{12} N_{23}, \\
\dot{N}_{14} &= -\tfrac{1}{2} i \Delta_2 N_{14} + \tfrac{1}{2}\epsilon_2 N_{12} N_{24}, \\
\dot{N}_{23} &= \tfrac{1}{2} i \Delta_1 N_{23} - \tfrac{1}{2}\epsilon_1 N_{12}^* N_{13}, \\
\dot{N}_{24} &= \tfrac{1}{2} i \Delta_2 N_{24} - \tfrac{1}{2}\epsilon_2 N_{12}^* N_{14}.
\end{aligned} \tag{4.5}$$

This system may be made integrable by the addition of just one more wave, N_{34}, in which case the equations of motion become

$$\begin{aligned}
\dot{N}_{12} &= -\epsilon_1 N_{13} N_{23}^* - \epsilon_2 N_{14} N_{24}^*, \\
\dot{N}_{13} &= -\tfrac{1}{2} i \Delta_1 N_{13} + \tfrac{1}{2}\epsilon_1 N_{12} N_{23} + \tfrac{1}{2}(\tilde{\epsilon} - \epsilon_1) N_{14} N_{34}^*, \\
\dot{N}_{14} &= -\tfrac{1}{2} i \Delta_2 N_{14} + \tfrac{1}{2}\epsilon_2 N_{12} N_{24} - \tfrac{1}{2}(\tilde{\epsilon} - \epsilon_2) N_{13} N_{34}, \\
\dot{N}_{23} &= \tfrac{1}{2} i \Delta_1 N_{23} - \tfrac{1}{2}\epsilon_1 N_{12}^* N_{13} + \tfrac{1}{2}(\tilde{\epsilon} + \epsilon_1) N_{24} N_{34}^*, \\
\dot{N}_{24} &= \tfrac{1}{2} \Delta_2 N_{24} - \tfrac{1}{2}\epsilon_2 N_{12}^* N_{14} - \tfrac{1}{2}(\tilde{\epsilon} + \epsilon_2) N_{23} N_{34}, \\
\dot{N}_{34} &= \tfrac{1}{2} i \tilde{\epsilon} (\Delta_1/\epsilon_1 - \Delta_2/\epsilon_2) N_{34} \\
&\quad + (\epsilon_2 - \epsilon_1) N_{13}^* N_{14} + (\epsilon_2 - \epsilon_1) N_{23}^* N_{24},
\end{aligned} \tag{4.6}$$

where $\tilde{\epsilon}$ is arbitrary.

V. INTEGRATION OF THE EQUATIONS OF MOTION WHEN ALL COUPLING COEFFICIENTS AND FREQUENCY MISMATCHES ARE EQUAL

In this section, we show that in the special case where $\Delta_1 = \Delta_2 = \cdots = \Delta_n \equiv \Delta$ and $\epsilon_1 = \epsilon_2 = \cdots = \epsilon_n \equiv \epsilon$, the integration of the equation of motion can be reduced to quadratures of elliptic functions. It is rather remarkable that this result holds true. For while we know that the solution is meromorphic,[1] meromorphic functions are generally expressible in terms of abelian integrals, of which elliptic integrals are a very special case.

We begin by reviewing the solution to the single-triad case because we shall need this result. The Hamiltonian is

$$H = \tfrac{1}{2}\Delta J_1 - \tfrac{1}{2}\Delta J_1' - \epsilon (J_1 J_1' J_0)^{1/2} \cos(\theta_1 - \theta_1' - \theta_0), \tag{5.1}$$

so that

$$\dot{J}_0 = -\frac{\partial H}{\partial \theta_0} = \epsilon (J_1 J_1' J_0)^{1/2} \sin(\theta_1 - \theta_1' - \theta_0). \tag{5.2}$$

Combining Eqs. (5.1) and (5.2) and recalling that

$$\begin{aligned}
J_1 &= I_0 - J_0, \\
J_1' &= I_1 - J_1 = I_1 - I_0 + J_0,
\end{aligned} \tag{5.3}$$

where I_0 and I_1 are constants of the motion, we find

$$\begin{aligned}
&(H - \Delta I_0 + \Delta I_1/2 + \Delta J_0)^2 + (\dot{J}_0)^2 \\
&\quad = \epsilon^2 J_0 (I_0 - J_0)(I_1 - I_0 + J_0).
\end{aligned} \tag{5.4}$$

This equation may be rewritten in the form

$$(\dot{J}_0)^2=\epsilon^2(Z_1-J_0)(Z_2-J_0)(Z_3-J_0), \tag{5.5}$$

where $Z_1<Z_2<Z_3$, and has the solution

$$J_0(t)=Z_3-(Z_3-Z_2)\mathrm{sn}^2[\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m], \tag{5.6}$$

where $m=(Z_2-Z_1)/(Z_3-Z_1)$ and δ is arbitrary.[11] It immediately follows from Eq. (5.3) that

$$\begin{aligned} J_1(t)&=I_0-Z_3+(Z_3-Z_2)\mathrm{sn}^2[\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m],\\ J_1'(t)&=(J_1-I_0+Z_3)-(Z_3-Z_2)\\ &\quad\times\mathrm{sn}^2[\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m]. \end{aligned} \tag{5.7}$$

From Eq. (5.1), we also have

$$\begin{aligned} \dot{\theta}_0&=-(\epsilon/2J_0)(J_1J_1'J_0)^{1/2}\cos(\theta_1-\theta_1'-\theta_0)\\ &=\tfrac{1}{2}\Delta+(H-\Delta I_0+\Delta I_1/2)/2J_0, \end{aligned} \tag{5.8}$$

which may be integrated to yield

$$\begin{aligned} \theta_0(t)=\ &\tfrac{1}{2}\Delta t+[(H-\Delta I_0+\Delta I_1/2)/\epsilon Z_3(Z_3-Z_1)^{1/2}]\\ &\times\{\Pi[1-Z_2/Z_3;\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m]\\ &-\Pi[1-Z_2/Z_3;-\delta|m]\}+\delta_0, \end{aligned} \tag{5.9}$$

[8]

where δ_0 is arbitrary. The solutions for $\theta_1(t)$ and $\theta_1'(t)$ may be found in a similar fashion.

Proceeding now to the many-triad case, we find

$$\dot{J}_0=\sum_n\epsilon(J_nJ_n'J_0)^{1/2}\sin(\theta_n-\theta_n'-\theta_0), \tag{5.10a}$$

$$\begin{aligned} &H-\Delta I_0+\tfrac{1}{2}\Delta\sum_n I_n+\Delta J_0\\ &\quad=\sum_n\epsilon(J_nJ_n'J_0)^{1/2}\cos(\theta_n-\theta_n'-\theta_0), \end{aligned} \tag{5.10b}$$

where

$$I_0=J_0+\sum_n J_n,\quad I_n=J_n+J_n' \tag{5.11}$$

are constants of the motion. Squaring and then combining Eqs. (5.10a) and (5.10b), we obtain

$$\begin{aligned} (\dot{J}_0)^2=&-\Big(H-\Delta I_0+\tfrac{1}{2}\Delta\sum_n I_n+\Delta J_0\Big)^2\\ &+\epsilon^2\Big[J_0(I_0-J_0)\Big(\sum_n I_n-I_0+J_0\Big)-\tfrac{1}{2}J_0\sum_{i,j}|I^{ij}|^2\Big], \end{aligned} \tag{5.12}$$

where

$$\begin{aligned} |I^{ij}|^2&=(b_ib_j'-b_i'b_j)(b_i^*b_j'^*-b_i'^*b_j^*)\\ &=(J_iJ_j')+(J_i'J_j)-2(J_iJ_i'J_jJ_j')^{1/2}\\ &\quad\times\cos(\theta_j-\theta_j'-\theta_i+\theta_i') \end{aligned} \tag{5.13}$$

is a constant of the motion. Equation (5.12) may be written in the form

$$(\dot{J}_0)^2=\epsilon^2(Z_1-J_0)(Z_2-J_0)(Z_3-J_0), \tag{5.14}$$

which has as its solution

$$J_0=Z_3-(Z_3-Z_2)\mathrm{sn}^2[\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m]. \tag{5.15}$$

We find similarly

$$\begin{aligned} \theta_0=\ &\tfrac{1}{2}\Delta+\Big[(H-\Delta I_0+\tfrac{1}{2}\Delta\sum_n I_n)/\epsilon Z_3(Z_3-Z_1)^{1/2}\Big]\\ &\times\{\Pi[1-Z_2/Z_3;\tfrac{1}{2}\epsilon(Z_3-Z_1)^{1/2}t-\delta|m]\\ &-\Pi[1-Z_2/Z_3;-\delta|m]\}+\delta_0. \end{aligned} \tag{5.16}$$

Hence, the behavior of the shared wave is the same as in the single-triad case. This fact was first noted by Meiss.[4]

Using now the equations of motion

$$\begin{aligned} \dot{b}_n&=\tfrac{1}{2}i\epsilon b_0b_n'-\tfrac{1}{2}i\Delta b_n,\\ \dot{b}_n'&=\tfrac{1}{2}i\epsilon b_0^*b_n+\tfrac{1}{2}i\Delta b_n' \end{aligned} \tag{5.17}$$

to eliminate b_n', we find

$$\begin{aligned} &\ddot{b}_n-(\dot{b}_0/b_0+\tfrac{1}{2}i\Delta)\dot{b}_n\\ &\quad-(\tfrac{1}{2}i\Delta\dot{b}_0/b_0-\tfrac{1}{4}\epsilon^2b_0b_0^*-\tfrac{1}{4}\Delta^2)b_n=0, \end{aligned} \tag{5.18}$$

which is a complex, linear second-order differential equation. If we can obtain the general solution to this equation, we may determine the solution in which we are interested by imposing the appropriate initial conditions. To obtain the general solution, we begin by noting that Eq. (5.18) is valid in the single-triad case, and hence we know the solution in the case where the initial conditions correspond to a single triad. Comparing the coefficient of each power of J_0 in Eqs. (5.4) and (5.12), we find that

$$\overline{H}-\Delta\overline{I}_0+\tfrac{1}{2}\Delta\overline{I}_1=H-\Delta I_0+\tfrac{1}{2}\Delta\sum_n I_n, \tag{5.19a}$$

$$\overline{I}_0(\overline{I}_1-\overline{I}_0)=I_0\Big(\sum_n I_n-I_0\Big)-\tfrac{1}{2}\sum_{i,j}|I^{ij}|^2, \tag{5.19b}$$

$$2\overline{I}_0-\overline{I}_1=2I_0-\sum_n I_n, \tag{5.19c}$$

where the barred quantities indicate the values that the constants of the motion would have to have in a single-triad system in order for the time behavior of the shared wave to be the same as in the many-triad system. Evidently, one must also have $\overline{\delta}=\delta$ and $\overline{\delta}_0=\delta_0$. As a result, the particular solution of Eq. (5.18) corresponding to single-triad motion is determined to within a phase. [We note that Eq. (5.19a) is consistent with the requirement that Eqs. (5.9) and (5.16) have the same time behavior.]

We will label the particular solution of Eq. (5.18) which we have just identified q_n and the general solution v_nq_n. Substituting v_nq_n into Eq. (5.18), we find

$$\ddot{v}_n+(2\dot{q}_n/q_n-\dot{b}_0/b_0+\tfrac{1}{2}i\Delta)\dot{v}_n=0. \tag{5.20}$$

Integrating Eq. (5.21) twice, we find

$$v_n=\alpha\int_0^t\frac{b_0}{q_n^2}\exp(-\tfrac{1}{2}i\Delta t')\,dt'+\beta, \tag{5.21}$$

where α and β are arbitrary complex constants. The general solution may now be written

$$b_n=\alpha q_n\int_0^t\frac{b_0}{q_n^2}\exp(-\tfrac{1}{2}i\Delta t')\,dt'+\beta q_n. \tag{5.22}$$

We may evidently determine b_n' in a precisely analogous manner which completes our demonstration.

VI. SUMMARY

This paper is the second in a series of papers discussing restricted multiple three-wave interactions. This system appears to be the simplest possible multiply interacting three-wave system involving an arbitrarily large number of waves and, as such, provides a useful platform from which to study aspects of more general interactions. In this paper, we show that a special case of the restricted system, namely that in which all coupling coefficients are equal, is integrable for arbitrary frequency mismatches. We then use the restricted system as a springboard from which to generate integrable cases for a more general class of three-wave systems.

ACKNOWLEDGMENTS

This work was supported by N.S.F. Contract PHY-80-14189 and O.N.R. Contract N000-14-77-C-0590.

[1]C. R. Menyuk, H. H. Chen, and Y. C. Lee, Phys. Rev. A **27**, 1597 (1983).

[2]H. Poincaré, *Les méthodes nouvelles de la méchanique céleste* (Gauthier-Villars, Paris, 1982), Vol. I.

[3]K. M. Watson, B. J. West, and B. I. Cohen, J. Fluid Mech. **77**, 185 (1976).

[4]J. D. Meiss, Ph.D. thesis (Univ. of California, Berkeley, 1980); J. D. Meiss, Phys. Rev. A **19**, 1780 (1979).

[5]M. J. Ablowitz and R. Haberman, J. Math. Phys. **16**, 2301 (1975).

[6]J. D. Meiss, in *A. I. P. Conference Proceedings*, No. 88, edited by M. Tabor and Y. Trève (Am. Inst. of Physics, New York, 1981).

[7]R. Haberman, J. Math. Phys. **18**, 1137 (1977).

[8]C. R. Menyuk, H. H. Chen, and Y. C. Lee, Phys. Rev. A (Rapid Comm.) **26**, 3731 (1982).

[9]See, e.g., H. H. Chen, Y. C. Lee, and C. S. Liu, Phys. Scripta **20**, 490 (1979).

[10]See, e.g., M. Tabor and J. Weiss, Phys. Rev. A **24**, 2157 (1981), for references.

[11]L. M. Milne-Thomson, in *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (U. S. Government Printing Office, Washington, D.C., 1964), Chap. 17.

Manifestly covariant, coordinate-free dyadic expression for planar homogeneous Lorentz transformations

Donald E. Fahnline
The Pennsylvania State University, Altoona, Pennsylvania 16603

(Received 29 September 1982; accepted for publication 18 November 1982)

Parametrizing a planar homogeneous Lorentz transformation P by any timelike or spacelike vector b lying in the transformation plane and its transform $a \equiv Pb$ yields a dyadic expression for P with several advantages: It provides an immediate solution to the problem of finding a homogeneous Lorentz transformation converting a given timelike or spacelike vector into a second similar vector. Its manifestly covariant, coordinate-free form is valid in any Lorentz frame and reduces easily to coordinate form. It unifies timelike (including boosts), spacelike (including pure spatial rotations), and null planar transformations and also orthochronous and nonorthochronous planar transformations into a single form; these classifications depend on the vectors a and b. Only if $a = -b$ does the expression fail, but then its limit as $a \rightarrow -b$ still exists and provides a valid expression for P.

PACS numbers: 03.30. + p

I. INTRODUCTION

This paper discusses the properties of a simple expression for planar homogeneous Lorentz transformations, including pure Lorentz transformations (boosts) and pure spatial rotations[1] as special cases, which has several convenient features. Its manifestly covariant and coordinate-free form makes it valid in any Lorentz frame, and its matrix elements are easily found in any Lorentz frame. It unifies timelike, null, and spacelike planar transformations and also orthochronous and nonorthochronous planar transformations into a single form; other approaches require different forms for the different classifications.[1–6] It provides an immediate solution to the problem of finding a planar transformation P converting a given timelike or spacelike vector b into a second similar vector $a \equiv Pb$, because it parametrizes P in terms of just two such vectors. Only if $a = -b$ does the expression fail, but then its limit as $a \rightarrow -b$ still exists and provides a valid expression for P.

In special relativity a planar transformation is a proper homogeneous Lorentz transformation which, under the active interpretation, changes 4-vectors lying in some 2-flat through the origin into new vectors in the same 2-flat and which leaves vectors in the orthogonal 2-flat unchanged. A 2-flat is a two-dimensional plane in flat spacetime and is timelike if it intersects the null cone along two null lines, null if it lies tangent to the null cone along a single null line, and spacelike if it neither touches nor intersects the null cone. Planar transformations are timelike (including boosts), null, or spacelike (including pure spatial rotations, hereafter simply called rotations) according to the classification of their transformation 2-flat.[2–6]

The next section gives the notation and reviews the properties of a manifestly covariant expression for reflections.[7] The third section uses these properties to provide a simple proof that the expression for planar transformations is always a proper homogeneous Lorentz transformation, derives conditions on the vectors a and b for all possible classifications of planar transformations, discusses the properties of the expression obtained in the limit $a \rightarrow -b$, and shows that these expressions are adequate for representing any planar homogeneous Lorentz transformation. The final section gives several other expressions.

II. HOMOGENEOUS LORENTZ TRANSFORMATIONS AND REFLECTIONS

Relative to any Lorentz frame a (4-) vector x has components $x^\mu = (x^0; x^i)$, where Greek indicies run from 0 for the temporal component to 3 and Latin indices run only from 1 to 3 for the spatial components. The scalar product of two vectors is $x \cdot y \equiv x^\mu y_\mu = g_{\mu\nu} x^\mu y^\nu$, where repeated indices are summed and the metric tensor g has $g^{ii} = -g^{00} = 1, g^{\mu\nu} = 0$ for $\mu \neq \nu$ in all Lorentz frames. A vector x is timelike if $x \cdot x < 0$, spacelike if $x \cdot x > 0$, null if $x \cdot x = 0$ and $x \neq 0$, and zero if $x = 0$.

Under the active interpretation, a homogeneous Lorentz transformation, abbreviated HLT and symbolized by the letters H and G, is a linear transformation of vectors x into new vectors $x' = Hx$ conserving the scalar product: $x' \cdot y' = x \cdot y$. The elements $H^\mu{}_\nu$ of H relative to a Lorentz frame convert the components of x into those of x' both relative to the given Lorentz frame via

$$x'^\mu = H^\mu{}_\nu x^\nu. \tag{1}$$

These elements are real and are subject to the orthogonality conditions

$$H^\mu{}_\alpha H_\nu{}^\alpha = g^\mu{}_\nu = \delta^\mu_\nu, \tag{2}$$

which are equivalent to the requirement that the scalar product be invariant. Equation (2) implies

$$|H| \equiv \det \| H^\mu{}_\nu \| = \pm 1 \tag{3}$$

and

$$(H^0{}_0)^2 = 1 + (H^0{}_i)^2 \geqslant 1. \tag{4}$$

An HLT is proper if $|H| = 1$, improper if $|H| = -1$; it is orthochronous if $H^0{}_0 \geqslant 1$, nonorthochronous if $H^0{}_0 \leqslant -1$.

It follows from the definition that the product HG representing the application of two HLT in succession with G first is also an HLT and that every H has an inverse H^{-1}

which is an HLT. According to Eq. (2), its matrix elements are

$$H^{-1\mu}{}_{\nu} = H_{\nu}{}^{\mu}. \tag{5}$$

Similarity transformations $H' = GHG^{-1}$ yield new HLT with $|H'| = |H|$ and $\mathrm{Tr}H' = \mathrm{Tr}H$, where $\mathrm{Tr}H \equiv H^{\mu}{}_{\mu}$. (If G is interpreted passively, then $x^{\mu'} = G^{\mu'}{}_{\mu}x^{\mu}$ and $H^{\mu'}{}_{\nu'} = G^{\mu'}{}_{\mu}H^{\mu}{}_{\nu}G^{-1\nu}{}_{\nu'}$ give the components of x and the elements of H relative to a new Lorentz frame.)

The simplest HLT is the identity transformation E with $Ex = x$ for all x; it satisfies the invariant scalar product condition trivially. In any Lorentz frame the elements of E are $E^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} = \delta^{\mu}_{\nu}$, and hence one has $|E| = 1$, $\mathrm{Tr}E = 4$, and $E^{0}{}_{0} = 1$. This is equivalent to the dyadic expression

$$E = g^{\mu\nu}e_{\mu}e_{\nu}, \tag{6}$$

where e_{μ} is a tetrad of orthonormal basis vectors and the subscript μ on e_{μ} indicates which basis vector, not which component. However, E is independent of any choice of basis or reference frame, and it will be regarded as a well-defined object requiring no dyadic decomposition.

For any vector b such that $b\cdot b \neq 0$, the dyadic[7]

$$I \equiv I\{b\} = E - 2bb/b\cdot b \tag{7}$$

reflects vectors parallel to b through the origin and leaves unchanged vectors orthogonal to b

$$\begin{aligned} Ix &= x - 2b\cdot x\ b/b\cdot b, \\ &= -x \quad \text{for} \quad x = \phi b, \\ &= x \quad \text{for} \quad b\cdot x = 0. \end{aligned} \tag{8}$$

It is an HLT because it is linear and preserves scalar products

$$\begin{aligned} x'\cdot y' &= (x - 2b\cdot x\ b/b\cdot b)\cdot(y - 2b\cdot y\ b/b\cdot b) \\ &= x\cdot y. \end{aligned}$$

Alternatively, one may use the component form

$$I^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} - 2b^{\mu}b_{\nu}/b\cdot b \tag{9}$$

to check Eq. (2). Substituting Eq. (9) into the definition for the determinant of $I^{\mu}{}_{\nu}$ yields

$$|I| \equiv \epsilon_{\alpha\beta\gamma\delta}I^{\alpha}{}_{0}I^{\beta}{}_{1}I^{\gamma}{}_{2}I^{\delta}{}_{3} = -1, \tag{10}$$

where $\epsilon_{\alpha\beta\gamma\delta}$ is the completely antisymmetric Levi-Civita tensor with $\epsilon_{0123} = 1$. Other properties following from Eqs. (7) or (9) are

$$I^{\mu\nu} = I^{\nu\mu}, \tag{11}$$

$$\mathrm{Tr}I \equiv I^{\mu}{}_{\mu} = 2, \tag{12}$$

$$I\{\phi b\} = I\{b\} \quad \text{for} \quad \phi \neq 0, \tag{13}$$

$$I^2 \equiv II = E, \tag{14}$$

and

$$GI\{b\}G^{-1} = I\{Gb\}. \tag{15}$$

If $a\cdot a = b\cdot b \neq a\cdot b$, one has $(a-b)\cdot(a-b) \neq 0$ and

$$\begin{aligned} I\{a-b\}a &= b, \\ I\{a-b\}b &= a. \end{aligned} \tag{16}$$

Relative to any Lorentz frame, Eq. (9) gives $I^{0}{}_{0} = 1 + 2(b^{0})^2/b\cdot b$. If b is timelike, then $b^{i}b_{i} \geqslant 0$ implies $b\cdot b \geqslant -(b^{0})^2$ and $(b^{0})^2/b\cdot b \leqslant -1$ so that $I^{0}{}_{0} \leqslant -1$; one has $I^{0}{}_{0} = -1$ if and only if $b^{i} = 0$. If b is spacelike, it is obvious that $I^{0}{}_{0} \geqslant 1$ and that $I^{0}{}_{0} = 1$ if and only if $b^{0} = 0$. If b is null, $I\{b\}$ does not exist.

III. PLANAR HOMOGENEOUS LORENTZ TRANSFORMATIONS

A 2-flat $\mathscr{F}$ through the origin consists of all vectors $x = \alpha a + \beta b$, where a and b are any two linearly independent vectors in $\mathscr{F}$ and α and β are variable scalars.[2] A timelike $\mathscr{F}$ intersects the null cone along two null lines and has $\lambda \equiv (a\cdot b)^2 - a\cdot a\, b\cdot b > 0$; a null $\mathscr{F}$ touches the null cone along a single null line and has $\lambda = 0$; a spacelike $\mathscr{F}$ does not touch or intersect the null cone and has $\lambda < 0$. These classifications are invariant under any new choice of basis $a' = \alpha_1 a + \beta_1 b$, $b' = \alpha_2 a + \beta_2 b$ in $\mathscr{F}$ as long as $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$, which guarantees that a' and b' are linearly independent.[6] A 2-flat $\mathscr{F}$ through the origin has a unique orthogonal 2-flat $\mathscr{F}^*$ consisting of all spacetime vectors which are orthogonal to every vector in $\mathscr{F}$. If $\mathscr{F}$ is timelike (spacelike), then $\mathscr{F}^*$ is spacelike (timelike); if $\mathscr{F}$ is null, then $\mathscr{F}^*$ is null and contains the same null line as $\mathscr{F}$.

Planar transformations are proper HLT which convert vectors in some 2-flat $\mathscr{F}$ through the origin into new vectors in $\mathscr{F}$ and leave vectors in the orthogonal 2-flat invariant.[2] Common examples are rotations R and pure Lorentz transformations L (boosts). Rewriting the usual formulae for the matrix elements of a boost

$$\begin{aligned} L^{0}{}_{0} &= \gamma, \\ L^{0}{}_{i} = L^{i}{}_{0} &= -\gamma V^{i}, \\ L^{i}{}_{j} &= \delta^{i}_{j} + (\gamma - 1)V^{i}V^{j}/V^2, \end{aligned}$$

where $\mathbf{V}$ is an ordinary 3-velocity with magnitude V less than the speed of light $c = 1$ and $\gamma \equiv (1 - V^2)^{-1/2}$, yields the manifestly covariant expression

$$L^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} - 2n^{\mu}v_{\nu} + (n^{\mu} + v^{\mu})(n_{\nu} + v_{\nu})/(1 - n\cdot v),$$

where $n^{\mu} \equiv (1; 0,0,0)$ and $v^{\mu} \equiv (\gamma; \gamma V^{i})$.[8] The corresponding dyadic form

$$L = E - 2nv + (n+v)(n+v)/(1 - n\cdot v) \tag{17}$$

generalizes to

$$\begin{aligned} P \equiv P\{a,b\} \equiv E + 2ab/a\cdot a - (a+b)(a+b)/ \\ (a\cdot a + a\cdot b), \end{aligned} \tag{18}$$

where a and b are linearly independent vectors such that $a\cdot a = b\cdot b \neq 0$ and $a\cdot a + a\cdot b \neq 0$. The last inequality guarantees that neither $a + b$ nor $(1 + 2\,a\cdot b/a\cdot a)\,a - b$ is null or zero; hence $I\{a\}$, $I\{b\}$, $I\{a+b\}$, and $I\{(1 + 2\,a\cdot b/a\cdot a)a - b\}$ exist. One may check directly by substitution into Eq. (7) that

$$\begin{aligned} P\{a,b\} &= I\{a\}I\{a+b\} \\ &= I\{a+b\}I\{b\} \\ &= I\{(1 + 2a\cdot b/a\cdot a)a - b\}I\{a\}. \end{aligned} \tag{19}$$

Since I is an HLT, it follows from the first of these equalities and from Eq.(10) that P is an HLT and that $|P| = 1$.

By Eq. (18), P produces the transformations

$$Pb = a,$$
$$Pa = (2a\cdot b/a\cdot a)a - b, \tag{20}$$
$$Px = x \quad \text{for} \quad a\cdot x = b\cdot x = 0.$$

Hence P is always a planar transformation in the two-flat $\mathscr{F}$ through the origin determined by b and $a = Pb$. It is a timelike transformation T, a null transformation N, or a spacelike transformation S according to the classifications of $\mathscr{F}$ determined by $\lambda = (a\cdot b)^2 - (a\cdot a)^2$. If $a = \mathrm{n}$ and $b = v$, where $v\cdot v = -1$, Eq. (18) reduces back to Eq. (17) for a boost

$$L\{v\} \equiv P\{n,v\}. \tag{21}$$

This is orthochronous if $v^0 \geqslant 1$ and nonorthochronous if $v^0 \leqslant -1$. Since $v\cdot v = -1$, $L\{v\}$ has three essential parameters. If $a^0 = b^0 = 0$, Eq. (18) reduces to a rotation R.

One of the advantages of Eq.(18) is that it provides a single form for timelike, null, and spacelike planar transformations and for orthochronous and nonorthochronous planar transformations. Other approaches require different forms for the different classifications.[1-6] A second advantage is that, according to Eq. (20), it provides a simple solution to the problem of constructing an HLT converting a given timelike or spacelike vector b into a second, linearly independent, given vector a, where $a\cdot a = b\cdot b$. (If $a\cdot a = b\cdot b \neq a\cdot b$, then by Eq. (16) the improper transformation $I\{a-b\}$ does the same. Multiplying either of these solutions by transformations leaving a invariant yields other solutions.)

Any vector b' in $\mathscr{F}$

$$b' = \alpha_2 a + \beta_2 b \tag{22}$$

such that $b'\cdot b' \neq 0$ has the transform

$$a' \equiv Pb' = \alpha_1 a + \beta_1 b, \tag{23}$$

where $\alpha_1 = 2\alpha_2 a\cdot b/a\cdot a + \beta_2$ and $\beta_1 = -\alpha_2$ by Eq. (20). They obey

$$a'\cdot a' = b'\cdot b' \neq 0,$$
$$a'\cdot a' + a'\cdot b' = b'\cdot b' \ (a\cdot a + a\cdot b)/a\cdot a \neq 0,$$

and

$$\alpha_1\beta_2 - \alpha_2\beta_1 = b'\cdot b'/a\cdot a \neq 0.$$

The last equation shows that a' and b' are linearly independent. Hence $P\{a', b'\}$ exists, and substituting Eqs. (22) and (23) into Eq. (18) shows

$$P\{a',b'\} = P\{a,b\}. \tag{24}$$

From this and Eq. (21) it follows that if n lies in the transformation 2-flat of a planar transformation P, then P is a boost. Also, the condition $a\cdot a = b\cdot b$ and the arbitrariness of α_2 and β_2 in Eq. (22) imply that only five of the total of eight components of a and b are essential parameters.

The following properties of P follow directly from Eq. (18):

$$P\{\phi a,\phi b\} = P\{a,b\} \quad \text{for} \quad \phi \neq 0, \tag{25}$$
$$P^{-1}\{a,b\} = P\{b,a\} = P\{a, 2a\cdot b\ a/a\cdot a - b\}, \tag{26}$$
$$[P^2 - 2(a\cdot b/a\cdot a)P + E][P - E] = 0, \tag{27}$$
$$GP\{a,b\}G^{-1} = P\{Ga,Gb\}, \tag{28}$$
$$\mathrm{Tr}P = P^{\mu}{}_{\mu} = 2(1 + a\cdot b/a\cdot a), \tag{29}$$
$$\mathrm{Tr}(P^2) = [\mathrm{Tr}P - 2]^2 = 4(a\cdot b/a\cdot a). \tag{30}$$

TABLE I. Planar homogeneous Lorentz transformations.

Classification	Symbol	$\psi \equiv a\cdot b/a\cdot a$	$H^{\mu}{}_{\mu}$	$H^0{}_0$
Orthochronous timelike	T	$\psi > 1$	$P^{\mu}{}_{\mu} > 4$	$T^0{}_0 \geqslant \psi$
Boost	L			$L^0{}_0 = \psi$
Null	N	$\psi = 1$	$P^{\mu}{}_{\mu} = 4$	$N^0{}_0 > \psi$
Identity ($a = b$)	E			$E^0{}_0 = 1$
Spacelike	S	$-1 < \psi < 1$	$0 < P^{\mu}{}_{\mu} < 4$	$S^0{}_0 \geqslant 1$
Rotation	R			$R^0{}_0 = 1$
Exceptional ($a = -b$): spacelike	Π	$\psi = -1$	$\Pi^{\mu}{}_{\mu} = 0$	$\Pi^0{}_0 \geqslant 1$
rotation				$\Pi^0{}_0 = 1$
nonorthochronous boost				$\Pi^0{}_0 = -1$
nonorthochronous timelike				$\Pi^0{}_0 \leqslant -1$
nonorthochronous boost	L	$\psi < -1$	$P^{\mu}{}_{\mu} < 0$	$L^0{}_0 = \psi$
Nonorthochronous timelike	T			$T^0{}_0 \leqslant \psi$

For the special case $(a\cdot b)^2 - (a\cdot a)^2 \neq 0$, the reflections $I\{a-b\}$ and $I\{a - a\cdot b\, b/a\cdot a\}$ exist and one has

$$P\{a,b\} = I\{a-b\}I\{a - a\cdot b\ \ b/a\cdot a\} \tag{31}$$

by Eqs. (7) and (18).

Relative to any given Lorentz frame, Eq. (18) yields

$$P^0{}_0 = 1 + \eta/\delta, \tag{32}$$

where

$$\delta \equiv a\cdot a(a\cdot a + a\cdot b) \neq 0 \tag{33}$$

and

$$\begin{aligned}\eta &\equiv a\cdot a(a^0 + b^0)^2 - 2(a\cdot a + a\cdot b)a^0 b^0\\ &= |a^0\mathbf{b} - b^0\mathbf{a}|^2 \geqslant 0,\end{aligned} \tag{34}$$

and $\mathbf{a}$ and $\mathbf{b}$ are the spatial parts of a and b. The condition $\eta = 0$ holds if and only if $a^0\mathbf{b} = b^0\mathbf{a}$, which is equivalent to $a^0 b = b^0 a$. However, a and b must be linearly independent, and therefore $\eta = 0$ holds if and only if $a^0 = b^0 = 0$ and P is a rotation R. Equations (33), (34), and the three-dimensional Schwarz inequality yield

$$\begin{aligned}2\delta + \eta &= (a\cdot a + a\cdot b)^2 + |\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2\\ &\geqslant (a\cdot a + a\cdot b)^2 > 0.\end{aligned} \tag{35}$$

The equality $2\delta + \eta = (a\cdot a + a\cdot b)^2$ holds if and only if $\mathbf{a} = \mathbf{b}$, $\mathbf{a} = 0$, or $\mathbf{b} = 0$, i.e., if and only if $\mathscr{F}$ contains the temporal axis so that P is a boost L.

A timelike planar transformation $T = P\{a,b\}$ has $(a\cdot b)^2 > (a\cdot a)^2 > 0$. Hence either $a\cdot b/a\cdot a > 1$ or $a\cdot b/a\cdot a < -1$ must hold. In the first case Eq. (29) yields $\mathrm{Tr}T > 4$, Eq. (33) yields $\delta > 0$, and Eqs. (32) and (35) yield $T^0{}_0 \geqslant a\cdot b/a\cdot a > 1$ so that T is orthochronous. If in addition $T^0{}_0 = a\cdot b/a\cdot a$, one has $2\delta + \eta = (a\cdot a + a\cdot b)^2$ and T is an orthochronous boost. The vectors a and b may be both timelike or both spacelike; if they are timelike, then they are both future pointing or both past pointing. In the second case Eq. (29) yields $\mathrm{Tr}T < 0$, and Eq. (33) yields $\delta < 0$. It then follows from Eqs. (32) and (35) that $T^0{}_0 \leqslant a\cdot b/a\cdot a < -1$ so that T is nonorthochronous. The additional condition $T^0{}_0 = a\cdot b/a\cdot a$ now holds if and only if T is a nonorthochronous boost. If a and b are timelike, one is future pointing and the other is past pointing.

A spacelike planar transformation $S = P\{a,b\}$ has $(a\cdot b)^2 < (a\cdot a)^2$ and also $a\cdot a > 0$, since a spacelike 2-flat contains only spacelike vectors. It follows that $-1 < a\cdot b/a\cdot a < 1$, that $0 < \mathrm{Tr}S < 4$, that $\delta > 0$, and that $S^0{}_0 \geqslant 1$. Also, $S^0{}_0 = 1$ holds if and only if $\eta = 0$, i.e., if and only if S is a rotation R by the discussion following Eq. (34).

A null planar transformation $N = P\{a,b\}$ has $(a\cdot b)^2 = (a\cdot a)^2$, which reduces to $a\cdot b = a\cdot a > 0$ because $a\cdot a + a\cdot b \neq 0$ and because a null 2-flat contains no timelike vectors. Hence one has $\mathrm{Tr}N = 4$, $\eta > 0$, $\delta = 2(a\cdot a)^2 > 0$, and $N^0{}_0 > a\cdot b/a\cdot a = 1$. The last result distinguishes N from E, which has $E^0{}_0 = 1$. Table I summarizes the classifications of planar transformations.

The classifications of a given $P\{a,b\}$ as timelike, null, or spacelike and as orthochronous or nonorthochronous depend only on scalar products involving a and b. It therefore follows from Eq. (28) and the invariance of scalar products that these classifications are invariant under similarity transformations. However, for a timelike transformation $T \equiv P\{a,b\}$ with a and b given, one may pick any timelike unit vector v_B in the transformation 2-flat $\mathscr{F}$ and define $v_A \equiv Tv_B$ to obtain

$$L_A T L_A{}^{-1} = L_A P\{v_A, v_B\} L_A{}^{-1} = L\{v\} \tag{36}$$

via Eqs. (24), (28), (20), and (21), where $L_A \equiv L\{v_A\}$ and $v \equiv L_A v_B$. It also follows from Eq. (29) and the invariance of the trace under similarity transformations that

$$\gamma \equiv v^0 = n\cdot v/n\cdot n = a\cdot b/a\cdot a. \tag{37}$$

This construction fails if $v_A = \pm n$, but then T is already a boost. Thus a similarity transformation by a boost suffices for changing a timelike transformation into a boost.

Similarly, a spacelike transformation $S \equiv P\{a,b\}$ with a and b given has a timelike pointwise invariant 2-flat $\mathscr{F}^*$ from which one may pick any timelike unit vector v and construct $L \equiv L\{v\}$ by Eq. (21). Then Eq. (28) implies

$$S' \equiv LSL^{-1} = P\{a', b'\}, \tag{38}$$

where $a' \equiv La$ and $b' \equiv Lb$, and Eq. (20) yields

$$S'n = LSL^{-1}Lv = LSv = Lv = n. \tag{39}$$

It follows that S' is a rotation, that a' and b' are pure spatial vectors because Eqs. (39) and (20) imply $a'\cdot n = b'\cdot n = 0$, and that the angle of rotation ϕ of $R \equiv S'$ has

$$\cos\phi = a'\cdot b'/a'\cdot a' = a\cdot b/a\cdot a. \tag{40}$$

The ambiguity in ϕ due to $\cos\phi = \cos(2\pi - \phi)$ corresponds to the ambiguity in the direction in which b' rotates into a'.

If b and a are linearly dependent so that $a = \phi b$, then the condition $a\cdot a = b\cdot b$ implies $a = \pm b$. Although the definition of P in Eq. (18) requires that b and a be linearly independent so as to define a 2-flat, the case $a = b$ causes no difficulty because P simply reduces to E. The second possibility $a = -b$ violates the condition $a\cdot a + a\cdot b \neq 0$ and makes Eq. (18) indeterminate, but it still has a meaningful limit for $a \to -b$ if a and b determine a spacelike or timelike 2-flat. Given $(a\cdot b)^2 - (a\cdot a)^2 \neq 0$ and $a \neq -b$, express the vector a as a linear combination of b and any fixed vector c which is linearly independent of b and which lies in the 2-flat determined by a and b:

$$a = \beta b + \gamma c. \tag{41}$$

The condition $a\cdot a = b\cdot b$ and the desired limit $a \to -b$ as $\gamma \to 0$ require

$$\beta = -1 - \gamma b\cdot c/b\cdot b \tag{42}$$

to first order in γ. Substituting Eqs. (41) and (42) into Eq. (18) gives

$$\begin{aligned}\Pi\{c,b\} &\equiv \lim_{\gamma\to 0}\,[P\{a,b\}]\\ &= E + 2[b\cdot bcc - b\cdot c(bc + cb)\\ &\qquad + c\cdot cbb]/[(b\cdot c)^2 - b\cdot bc\cdot c],\end{aligned} \tag{43}$$

where $(b\cdot c)^2 - b\cdot bc\cdot c \neq 0$ because b and c still determine the original timelike or spacelike 2-flat.

Consider the dyadic expression for Π in Eq. (43) for any two vectors b and c, including null vectors, such that $(b\cdot c)^2 - b\cdot bc\cdot c \neq 0$. This inequality implies that b and c must be linearly independent, nonzero vectors determining a timelike or spacelike 2-flat $\mathscr{F}$. The substitution of

$$d \equiv \gamma_1 c + \beta_1 b, \quad e \equiv \gamma_2 c + \beta_2 b, \tag{44}$$

where $\gamma_1\beta_2 - \gamma_2\beta_1 \neq 0$ so that d and e are linearly independent and $(d\cdot e)^2 - d\cdot d\, e\cdot e \neq 0$, into Eq. (43) yields

$$\Pi \equiv \Pi\{c,b\} = \Pi\{d,e\}. \tag{45}$$

Hence, if c is null, it is always possible to reparametrize Π with vectors d and e, where $d\cdot d \neq 0$. Then the vector $e - e\cdot d\; d/d\cdot d$ exists, is orthogonal to d, and is not null or zero. Hence $I\{d\}$ and $I\{e - e\cdot d\, d/d\cdot d\}$ exist, and one may check that

$$\Pi = \Pi\{d,e\} = I\{d\}I\{e - e\cdot d\;\; d/d\cdot d\} \tag{46}$$

by Eqs. (7) and (43). It follows that Π is a proper HLT. By Eq. (43), it produces the transformations

$$\Pi b = -b, \quad \Pi c = -c$$
$$\Pi x = x \quad \text{for} \quad b\cdot x = c\cdot x = 0. \tag{47}$$

Hence Π is a planar transformation reversing the direction of all vectors in the two-flat $\mathscr{F}$ determined by b and c. Equations (44) and (45) imply that $\Pi\{b,c\}$ has only four essential parameters.

Equation (43) implies the following relations:

$$\Pi\{c,b\} = \Pi\{b,c\} = \Pi^{-1}\{b,c\}, \tag{48}$$

$$\Pi\{\gamma c,\beta b\} = \Pi\{c,b\} \quad \text{for} \quad \gamma \neq 0,\ \beta \neq 0, \tag{49}$$

$$\Pi^2 - E = 0, \tag{50}$$

$$G\Pi\{c,b\}G^{-1} = \Pi\{Gc,Gb\}, \tag{51}$$

$$\Pi^{\mu\nu} = \Pi^{\nu\mu}, \tag{52}$$

$$\mathrm{Tr}\Pi = 0, \tag{53}$$

$$\mathrm{Tr}(\Pi^2) = (\mathrm{Tr}\Pi - 2)^2 = 4, \tag{54}$$

and

$$\Pi\{c,b\}\Pi\{d,e\} = -E, \tag{55}$$

where d and e are orthogonal to c and b. For $a\cdot a \neq 0$ and $c\cdot c \neq 0$, one has

$$I\{a\}I\{c\} = P\{a, 2a\cdot c\;\, c/c\cdot c - a\} \quad \text{for} \quad a\cdot c \neq 0 \tag{56}$$
$$= \Pi\{a,c\} \quad \text{for} \quad a\cdot c = 0 \tag{57}$$

by Eqs. (7), (18), and (43). If $a\cdot a = b\cdot b \neq 0$ and $(a\cdot b)^2 - (a\cdot a)^2 \neq 0$, then Eqs. (18) and (43) give

$$P\{a,b\} = \Pi\{a,b\}P\{-a,b\} = P\{a,-b\}\Pi\{a,b\}. \tag{58}$$

Under the same conditions, Eqs. (58) and (55) imply

$$-P\{a,b\} = \Pi\{d,e\}P\{-a,b\}, \tag{59}$$

where d and e are orthogonal to a and b.

Relative to any Lorentz frame, Eq. (43) gives

$$\Pi^0{}_0 = 1 - 2\eta/\delta, \tag{60}$$

where

$$\delta \equiv (b\cdot c)^2 - b\cdot b\;\, c\cdot c \neq 0 \tag{61}$$

and

$$\eta \equiv b\cdot b\,(c^0)^2 - 2b\cdot c b^0 c^0 + c\cdot c(b^0)^2$$
$$= |c^0\mathbf{b} - b^0\mathbf{c}| \geqslant 0. \tag{62}$$

As before, one has $\eta = 0$ if and only if b and c are pure spatial vectors. The Schwarz inequality provides

$$\eta - \delta = |\mathbf{b}|^2|\mathbf{c}|^2 - (\mathbf{b}\cdot\mathbf{c})^2 \geqslant 0, \tag{63}$$

where the equality holds if and only if the transformation plane contains n. The condition $\delta > 0$ that the transformation plane be timelike thus implies $\Pi^0{}_0 \leqslant -1$, and $\Pi^0{}_0 = -1$ implies that the transformation plane contains n. The condition $\delta < 0$ for a spacelike plane implies $\Pi^0{}_0 \geqslant 1$, and $\Pi^0{}_0 = 1$ implies that the transformation is a rotation by π radians. Table I also includes these classifications.

Suppose that H is an HLT possessing at least a pointwise invariant 2-flat $\mathscr{F}^*$ through the origin; then one of the following constructions expresses H as E, P, or Π if $|H| = 1$ or as I if $|H| = -1$. These constructions use the simple results

$$Hr = r \text{ and } r\cdot b = 0 \text{ imply } r\cdot(Hb) = 0, \tag{64}$$

because $r\cdot(Hb) = (H^{-1}r)\cdot b = r\cdot b = 0$, and

$$Hb = \phi b \text{ and } b\cdot b \neq 0 \quad \text{imply} \quad \phi = \pm 1, \tag{65}$$

because $(Hb)\cdot(Hb) = b\cdot b$.

Case 1: Let $\mathscr{F}^*$ be timelike or spacelike and pick any two linearly independent vectors x and y in $\mathscr{F}^*$. If at least one of them, say x, is not null, define

$$r \equiv x, \quad s \equiv y - x\cdot y x/x\cdot x;$$

if both are null, define

$$r \equiv x + y, \quad s \equiv x - y.$$

Then r and s are linearly independent, orthogonal, timelike or spacelike vectors in $\mathscr{F}^*$ with $Hr = r$ and $Hs = s$. Similarly, construct two linearly independent, orthogonal, timelike or spacelike vectors b and c in $\mathscr{F}$, the timelike or spacelike 2-flat orthogonal to $\mathscr{F}^*$. If Hb is linearly dependent on b, Eq. (65) gives $Hb = \pm b$; otherwise, Hb is linearly independent of b.

Subcase A: Let $Hb = b$. Since b, c, r, and s are orthogonal and not null, they form a basis. Hence Eq. (64) implies $Hc = \phi c$, where $\phi = \pm 1$ by Eq. (65), and one has either $H = E$ or $H = I\{c\}$.

Subcase B: Let $Hb = -b$ and define $H' \equiv I\{b\}H$. Then H' satisfies the conditions for Subcase A by Eq. (8), and it follows that $H' = E$ and $H = I\{b\}$ by Eq. (14) or that $H' = I\{c\}$ and $H = I\{b\}\, I\{c\} = \Pi\{b,c\}$ by Eq. (57).

Subcase C: Let $a \equiv Hb$ be linearly independent of b. Then one has $a\cdot a = b\cdot b \neq 0$ and also $(a\cdot b)^2 - (a\cdot a)^2 \neq 0$, because a is in $\mathscr{F}$ by Eq. (64) and $\mathscr{F}$ is not null. It follows that $a\cdot a + a\cdot b \neq 0$ and that $P\{a,b\}$ exists. For c use $c \equiv a - a\cdot b b/b\cdot b$, which exists in $\mathscr{F}$, is not null because $c\cdot c = [(a\cdot a)^2 - (a\cdot b)^2]/a\cdot a \neq 0$, and is orthogonal to b. Then $H' \equiv P^{-1}\{a,b\}H$ satisfies the conditions for Subcase A by Eq. (20), and one has $H' = E$ or $I\{a - a\cdot b\, b/b\cdot b\}$. It follows that either $H = P\{a,b\}$ or $H = P\{a,b\}\, I\{a - a\cdot b\, b/b\cdot b\} = I\{a - b\}$ by Eq. (31).

Case 2: Let $\mathscr{F}^*$ be a null 2-flat and pick from it any two linearly independent vectors x and y; hence they must satisfy $(x\cdot y)^2 - x\cdot y\; y\cdot y = 0$. This equation and the fact that $x\cdot y \neq 0$ for two linearly independent null vectors imply that x and y cannot both be null. Hence at least one of them, say x, is not null. Define the vectors

$$z \equiv y - x\cdot y\;\; x/x\cdot x, \quad r \equiv x - n\cdot x\;\; z/n\cdot z.$$

Then z exists in $\mathscr{F}^*$, is null, and is orthogonal to x; consequently, r exists in $\mathscr{F}^*$ and is a spacelike vector orthogonal to n and z. It also follows that $Hz = z$ and $Hr = r$. The 2-flat $\mathscr{F}$ orthogonal to $\mathscr{F}^*$ is null and also contains z, and one can similarly construct a spacelike vector b in $\mathscr{F}$ which is orthogonal to n, z, and r. Then the vectors n, z, b, and r are linearly independent and form a basis for spacetime. As in Case 1, one has $Hb = \pm b$ or else Hb is linearly independent of b.

Subcase A: Let $Hb = b$. Equation (64) implies that

$$Hn = \nu n + \zeta z + \beta b + \rho r$$

has $\beta = \rho = 0$. Then $z\cdot n = (Hz)\cdot(Hn) = z\cdot(Hn) = \nu z\cdot n$ yields $\nu = 1$, and $-1 = (Hn)\cdot(Hn) = (n + \zeta z)\cdot(n + \zeta z)$ yields $\zeta = 0$. Hence one has $Hn = n$ and $H = E$.

Subcase B: Let $Hb = -b$ and define $H' \equiv I\{b\}H$. Then H' satisfies the conditions of Subcase A by Eq. (8), and it follows that $H' = E$ and $H = I\{b\}$.

Subcase C: Let $a \equiv Hb$ be linearly independent of b. If $a\cdot a + a\cdot b \neq 0$, then $P\{a,b\}$ exists, $H' \equiv P^{-1}\{a,b\}H$ must equal E by Eq. (20) and subcase A, and $H = P\{a,b\}$. If $a\cdot a + a\cdot b = 0$, one has $(a - b)\cdot(a - b) = 4a\cdot a = 4b\cdot b \neq 0$. Hence $I\{a - b\}$ exists, $H' \equiv I\{a - b\}H$ must equal E by Eq. (16) and Subcase A, and $H = I\{a - b\}$. This completes the constructions.

IV. OTHER EXPRESSIONS FOR PLANAR TRANSFORMATIONS

Although the forms for P and Π given by Eqs. (18) and (43) are together sufficient for expressing any planar HLT, other forms are sometimes useful. For example, if $a\cdot a = b\cdot b \neq 0$ and $(a\cdot b)^2 \neq (a\cdot a)^2$, define

$$\alpha \equiv a\cdot b/a\cdot a \pm [(a\cdot b/a\cdot a)^2 - 1]^{1/2},$$

$$z_1 \equiv \alpha a - b,$$

and

$$z_2 \equiv a/\alpha - b \qquad (66)$$

so that $\alpha \neq 0$ or ± 1 and $z_1\cdot z_1 = z_2\cdot z_2 = 0$. Hence z_1 and z_2 are linearly independent null vectors, and Eq. (66) is invertable:

$$a = (z_1 + z_2)/(\alpha - \alpha^{-1}),$$
$$b = (\alpha^{-1} z_1 + \alpha z_2)/(\alpha - \alpha^{-1}). \qquad (67)$$

Substituting these results into Eq. (18) yields

$$P\{a,b\} = E + [(\alpha - 1)z_1 z_2 + (\alpha^{-1} - 1)z_2 z_1]/z_1\cdot z_2 \qquad (68)$$

with

$$Pz_1 = \alpha z_1, \quad Pz_2 = z_2/\alpha. \qquad (69)$$

If $P\{a,b\}$ is timelike, then α, z_1, and z_2 are real. If $P\{a,b\}$ is spacelike, one has $\alpha\alpha^* = 1$ and $z_2 = z_1^*$, where * here denotes complex conjugation.

Alternatively, one can express $P\{a,b\}$ in terms of z_1 and b using Eq. (66). This is most useful for null $P\{a,b\}$, which have $a\cdot b = a\cdot a$ so that Eq. (66) yields

$$\alpha = 1,$$

$$z \equiv z_1 = z_2 = a - b,$$

and

$$z\cdot z = z\cdot a = z\cdot b = 0. \qquad (70)$$

Eliminating a from Eq. (18) yields

$$N \equiv P\{a,b\} = E + (2zb - 2bz - zz)/2b\cdot b, \qquad (71)$$

with

$$Nz = z, \quad Nb = b + z. \qquad (72)$$

The antisymmetric part of Eq. (18) for $P\{a,b\}$ is the bivector

$$A \equiv (P\{a,b\} - P\{b,a\})/2$$
$$= (ab - ba)/a\cdot a. \qquad (73)$$

Noting that

$$A^2 = [a\cdot b(ab + ba) - a\cdot a(aa + bb)]/(a\cdot a)^2 \qquad (74)$$

and

$$\mathrm{Tr}(A^2) = 2[(a\cdot b/a\cdot a)^2 - 1], \qquad (75)$$

one obtains from Eq.(18)

$$P\{a,b\} = E + A + A^2/\{1 \pm [1 + \mathrm{Tr}(A^2)/2]^{1/2}\}. \qquad (76)$$

Thus the statement[5,6] that the antisymmetric part of P determines P is not quite correct because of the ambiguous sign in the denominator of the third term of this equation. The ambiguity is due to the fact that $P\{a,b\}$ and $P\{-b,a\}$, when it exists, have the same antisymmetric bivector A; for example, a rotation by ϕ radians about the z axis has the same A as a rotation by $\pi - \phi$ radians about the z axis [i.e., $\sin\phi = \sin(\pi - \phi)$]. According to Eq.(75), P is spacelike for $\mathrm{Tr}(A^2) < 0$; then the positive sign in Eq. (76) corresponds to $0 < \phi \leqslant \pi/2$ and the negative sign to $\pi/2 < \phi < \pi$, where ϕ is given by Eq. (40). The transformation P is timelike for $\mathrm{Tr}(A^2) > 0$, and it is orthochronous for the positive sign and nonorthochronous for the negative sign. For $\mathrm{Tr}(A^2) = 0$, the transformation is null and exists only for the positive sign. Equation (76) differs from the various expressions given by Bazanski[5] and by Rao, Saroja,, and Rao[6] in that A is constructed from a pair of unnormalized and nonorthogonal vectors rather than from an orthonormal pair. As a result, it contains no extra scalar paremeter for the angle or pseudo-angle of the transformation and it is able to represent all three classifications of planar transformations.

For $a = \pm b$, one has $A = 0$ and Eq. (74) yields $P = E$ for the positive sign and is indeterminant for the negative sign, exactly as for Eq. (18). The singularity may again be removed by taking the limit $a \to -b$ on a timelike or spacelike 2-flat. This yields

$$\Pi = E - A^2/\mathrm{Tr}(A^2), \qquad (77)$$

where $A \equiv cb - bc$ and c and b are linearly independent vectors in the 2-flat. Rao, Saroja, and Rao[6] give a similar expression for the spacelike case only.

V. CONCLUSIONS

It has been shown that the dyadic $P\{a,b\}$ defined in Eq. (18) and its limit $\Pi\{c,b\}$ as $a \to -b$ given by Eq. (43) are always planar transformations and that they are sufficient for representing any planar transformation regardless of its classification as timelike, null, or spacelike or as orthochronous or nonorthochronous. These classifications depend only on the vectors a and b or c and b as summarized in Table I.

Rao, Saroja, and Rao[6] prove that the equation

$$\mathrm{Tr}H = 1 + \xi/2,$$

where ξ is the sum of all principal minors of second order of H, is a necessary and sufficient condition for a proper orthochronous HLT to be planar. Equations (30) and (54) give a slightly simpler form of this condition

$$\mathrm{Tr}(H^2) = (\mathrm{Tr}H - 2)^2$$

and show that it is a necessary condition for nonorthochronous proper planar transformations as well. A future article will discuss its sufficiency and provide methods for decomposing an arbitrary HLT, including improper and nonorthochronous HLT, into the product of various pairs of planar transformations or reflections.

[1]For a reference discussing both pure spatial rotations and pure Lorentz transformations see H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1980), 2nd ed., pp. 147, 281.

[2]Two-flats, timelike transformations, spacelike rotations, and null transformations, under the name singular transformations, are discussed in J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1972), 2nd. ed., pp. 60–1, 79–80, 434–7.

[3]E. Wigner, Ann. Math. **40**, 149 (1939). Wigner's solution to the eigenvalue problem for proper orthochronous HLT gives classifications equivalent to those described here. His case (b) corresponds to spacelike transformations, case (c) to timelike transformations, and his transformation with an elementary divisor corresponds to a null transformation.

[4]H. M. Schwartz, Am. J. Phys. **31**, 864 (1963).

[5]S. L. Bazanski, J. Math. Phys. **6**, 1201 (1965).

[6]K. N. S. Rao, D. Saroja, and A. V. G. Rao, J. Math. Phys. **22**, 2167 (1981).

[7]A. O. Barut, *Electrodynamics and Classical Theory of Fields and Particles* (MacMillan, New York, 1964), pp. 14, 15.

[8]D. E. Fahnline, Am. J. Phys. **50**, 818 (1982); J. Krause, J. Math. Phys. **18**, 889 (1977).

Can poles change color?[a)]

Jacek Tafel and Andrzej Trautman[b)]
Instytut Fizyki Teoretycznej, Uniwersytetu Warszawskiego, Hoza 69, Warsaw, Poland

(Received 16 June 1982; accepted for publication 11 August 1982)

The definition of the total nonabelian charge ("color") in a classical Yang–Mills theory is shown to require a careful analysis of the boundary conditions at infinity imposed on the potentials and on gauge transformations. The color current of a nonabelian plane wave is found to be different from zero in the transverse gauge, though it vanishes in the null gauge. The color charge of a single pole, described by the Liénard–Wiechert potentials, is constant by virtue of the Yang–Mills equations. An approximate computation indicates that the total color charge of a system of particles may change in time, as a result of radiation. To make this result meaningful, it is necessary to find a method of fixing the allowed gauge transformations to those having a direction-independent limit at infinity.

PACS numbers: 03.50.Kk

I. INTRODUCTION

In a classical gauge theory of the Yang–Mills type, sources of the field have a nonabelian charge density. More precisely, the sources are described by a 3-form j with values in the Lie algebra of the structure (gauge) group. If the group is nonabelian, the current j does not, by itself, satisfy a differential conservation law; it has to be supplemented by another current i constructed out of quantities referring only to the gauge configuration. The latter current may be interpreted as representing the density of the nonabelian charges residing in the gauge field itself. We use the name "color" for this nonabelian charge, but our considerations have little, if anything, to do with chromodynamics. Our nonabelian charges may equally well be associated with "flavors" and, in particular, the charges occurring in gauge theories of weak interactions.

One expects that, upon integration of $i + j$ over a space region Ω, it should be possible to find its total color content. By means of the Gauss law, the total color is represented as the flux, across the boundary of Ω, of the Lie-algebra-valued electric field. The problems considered in this paper are the following: What is the dependence of the total color on the choice of gauge? Can the total nonabelian charge of a system of particles change in time, as a result of radiation? Both these problems have analogs in Einstein's theory of gravitation. The question of color radiation is analogous to that of gravitational radiation. From this point of view, the Yang–Mills theory may be considered as a simplified model of general relativity.

The nature of the difficulties one encounters when trying to define total color, even in the static case, can be seen as follows. Consider a gauge potential A, function of the spherical coordinates r,θ,φ. For an isolated system, one expects that, at large distances,

$$A = O(r^{-1}) \tag{1.1}$$

and the corresponding electric field E is of order $O(r^{-2})$. If A is replaced by its gauge transform $A' = S^{-1}AS + S^{-1}dS$, then E changes into $E' = S^{-1}ES$. A gauge transformation function S satisfying

$$S(r,\theta,\varphi) = a(\theta,\varphi)[I + O(r^{-1})] \tag{1.2}$$

is compatible with (1.1) because $A' = a^{-1}Aa + a^{-1}da + O(r^{-2})$ and $a^{-1}da = O(r^{-1})$. This transformation, however, induces such a change in E,

$$E' = a(\theta,\varphi)^{-1}Ea(\theta,\varphi) + O(r^{-3}) \tag{1.3}$$

that the flux of E' bears no simple relation to that of E. In the theory of general relativity, a similar problem occurs, but is not as acute as in the case of the Yang–Mills theory.[1] In Einstein's theory, the coefficients Γ of a linear connection constitute an analog of A: they transform under changes of the local frames in a manner similar to A. There is also an important difference: the Yang–Mills equations contain second derivatives of A, whereas Γ appears in Einstein's equations differentiated only once. As a result, for a static configuration, Γ falls off faster, $\Gamma = O(r^{-2})$, and an arbitrary function a occurring in (1.2) is not allowed here.

Presumably, the arbitrary function a can be eliminated by a more detailed analysis of the gauge potentials. For example, if it can be shown that the $1/r$ part of a static potential is spherically symmetric, then one can restrict a by requiring that the spherical symmetry be explicit.

In this paper, we leave aside the question of how the direction-dependent gauge transformation can be eliminated and concentrate on the study of the dynamics of Yang–Mills fields in the wave zone. The purpose of the work is to determine the rate of change of a "retarded" total color charge. We use an asymptotic expansion method developed for the study of gravitational radiation by Bondi, *et al.*[2] and Sachs.[3] The method has also been used in the context of the Yang–Mills theory to prove peeling-off theorems for gauge fields.[4]

II. NOTATION[5]

All gauge configurations considered here are defined on the Minkowski space $\mathbb{R}^4$ with its standard metric $g_{\mu\nu}\, dx^\mu dx^\nu = dt^2 - dx^2 - dy^2 - dz^2$ and orientation given

a) Research supported in part by the Polish Ministry of Science, Higher Education and Technology under grant MR-I-7 to Warsaw University.
b) Professor invited to the University of Montreal thanks to funds from la Campagne du Cinquantenaire (March–June 1982).

by the volume 4-form $dt \wedge dx \wedge dy \wedge dz$. A Lie group G is assumed to be the structure (gauge) group of the theory and $\mathfrak{g}$ denotes its Lie algebra. A gauge potential A is a $\mathfrak{g}$-valued 1-form on $\mathbb{R}^4$,

$$A = A^i_\mu e_i \, dx^\mu, \tag{2.1}$$

where (e_i) is a linear basis in $\mathfrak{g}$. The field strengths are

$$F = dA + \tfrac{1}{2}[A,A] = \tfrac{1}{2} F^i_{\mu\nu} e_i \, dx^\mu \wedge dx^\nu. \tag{2.2}$$

The four-dimensional Hodge dual of F is denoted by $\check{F}$. If F is represented in terms of its electric and magnetic components,

$$\begin{aligned} F = {} & dt \wedge (E_x\, dx + E_y\, dy + E_z\, dz) \\ & - B_x\, dy \wedge dz - B_y\, dz \wedge dx - B_z\, dx \wedge dy, \end{aligned} \tag{2.3}$$

then

$$\begin{aligned} \check{F} = {} & E_x\, dy \wedge dz + E_y\, dz \wedge dx \\ & + E_z\, dx \wedge dy + dt \wedge (B_x\, dx + B_y\, dy + B_z\, dz). \end{aligned} \tag{2.4}$$

The Yang–Mills equations are

$$d\check{F} + [A,\check{F}] = 4\pi j, \tag{2.5}$$

where j is the $\mathfrak{g}$-valued 3-form describing the sources.

If $S:\mathbb{R}^4 \to G$ is a function corresponding to a gauge transformation, then

$$A' = S^{-1}AS + S^{-1}dS \quad \text{and} \quad F' = S^{-1}FS \tag{2.6}$$

are the transformed potential and field strengths, respectively.

If G is either abelian or semisimple and compact, then its Lie algebra $\mathfrak{g}$ is compact, i.e., it admits a scalar product, $\mathfrak{g} \times \mathfrak{g} \ni (X,Y) \mapsto (X|Y) \in \mathbb{R}$, which is invariant,

$$([Z,X]|Y) + (X|[Z,Y]) = 0 \quad \text{for any } X,Y,Z \in \mathfrak{g}, \tag{2.7}$$

and positive-definite. If G is semisimple and compact one can indeed take the (negative of the) Killing–Cartan form on $\mathfrak{g}$ as such a scalar product. By applying (2.7) it is straightforward to prove the following:

Lemma: If the elements X and Y of a compact Lie algebra are such that $[X,Y] = X$, then $X = 0$.

If (e_i) is a linear basis in $\mathfrak{g}$, then X may be written as $X^i e_i$ and

$$(X|Y) = h_{ij} X^i X^j, \tag{2.8}$$

where

$$h_{ij} = (e_i|e_j). \tag{2.9}$$

The structure constants of $\mathfrak{g}$ relative to (e_i) are defined by

$$[e_i, e_j] = c^k_{ij} e_k$$

and the condition of invariance (2.7) is equivalent to

$$h_{kl} c^l_{ij} + h_{jl} c^l_{ik} = 0. \tag{2.10}$$

The two-dimensional unit sphere S_2 has a metric $d\theta^2 + \sin^2\theta\, d\varphi^2$ and a surface element $d\theta \wedge \sin\theta\, d\varphi$. The two-dimensional Hodge dual on S_2 will be denoted by a star; thus

$$\begin{aligned} &{}^\star 1 = d\theta \wedge \sin\theta\, d\varphi, \quad {}^\star d\theta = \sin\theta\, d\varphi, \\ &{}^\star \sin\theta\, d\varphi = -d\theta. \end{aligned} \tag{2.11}$$

There exist a few useful relations between four-dimensional and two-dimensional duals. Let $r = (x^2 + y^2 + z^2)^{1/2}$ and $u = t - r$ be the "retarded time." In coordinates (u,r,θ,φ) the Minkowski metric is $du^2 + 2\, du\, dr - r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)$ and $dt \wedge dx \wedge dy \wedge dz = du \wedge dr \wedge r d\theta \wedge r \sin\theta\, d\varphi$. One easily shows that, for any 1-form α on S_2, linear in $d\theta$ and $d\varphi$, the 4-dual of $du \wedge \alpha$ is equal to $du \wedge {}^\star\alpha$, whereas the 4-dual of $dr \wedge \alpha$ is $-(du + dr) \wedge {}^\star\alpha$. Similarly, the 4-dual of $du \wedge dr$ is ${}^\star r^2$.

III. THE CONSERVATION LAW

It is clear from the Yang–Mills field equation (2.5) that the current j of the sources is not conserved by itself: in general $dj \neq 0$. The "total current"

$$J = j - (1/4\pi)[A,\check{F}] \tag{3.1}$$

is conserved,

$$dJ = 0, \tag{3.2}$$

but contains a highly gauge-dependent field contribution,

$$i = -(1/4\pi)[A,\check{F}], \tag{3.3}$$

analogous to the pseudotensor of energy and momentum of the gravitational field in Einstein's theory. Let Σ_R be the surface (boundary) of a ball $\Omega_R \subset \mathbb{R}^3$ of radius R. The total nonabelian charge q contained in Ω_R at time t may be formally defined as

$$q(t,R) = \int_{\Omega_R} J \quad \text{at } t = \text{const.} \tag{3.4}$$

and, by virtue of Eq. (2.5), expressed as a surface integral,

$$q(t,R) = (1/4\pi) \int_{\Sigma_R} \check{F} \quad \text{at } t = \text{const.} \tag{3.5}$$

The Gauss law (3.5) is analogous to the expression of total energy and momentum of a gravitational configuration by means of a surface integral of the Von Freud superpotential. The rate of change of color, $\dot{q} = \partial q/\partial t$ is given by

$$\dot{q}(t,R) = \int_{\Sigma_R} \frac{\partial}{\partial t} \lrcorner\, i \quad \text{at } t = \text{const.} \tag{3.6}$$

provided that the sources are spatially bounded and R is sufficiently large so that $j = 0$ on Σ_R. If both A and F tend to 0 sufficiently fast as $R \to \infty$, then

$$q_\infty(t) = \lim_{R \to \infty} q(t,R) \tag{3.7}$$

is well-defined and conserved by (3.6), $\dot{q}_\infty(t) = 0$. It is known, however, that such a description is not adequate when radiation is present. In this case, one expects a suitably defined total charge to change in the course of time and both A and F to behave as $1/r$ at large r. Making use of the retarded time $u = t - r$, one can define

$$q_{\text{ret}}(u) = \lim_{R \to \infty} q(u + R, R). \tag{3.8}$$

The charge q_{ret}, which is defined on the "future null infinity" in the sense of Penrose,[6] may depend on u even though $\dot{q}_\infty(t) = 0$.

IV. TWO SIMPLE EXAMPLES

The current i, describing the color carried by a classical gluon wave, depends on the choice of gauge to such an extent that it may always be reduced to 0 at a spacetime point. Moreover, in special cases it may be zero throughout spacetime even though the corresponding solution of the Yang–Mills equations is believed to represent a truly colored wave. This difficulty will be illustrated on the following.

Example (*i*): Let $v = t - z$ and let H be a $\mathfrak{g}$-valued function of x, y, and v. The potential[7]

$$A = H\,dv \tag{4.1}$$

is a solution of the sourceless Yang–Mills equations if, and only if, H satisfies the Laplace equation in x and y,

$$\Delta H = 0. \tag{4.2}$$

In particular, the solution $H = a(v)x + b(v)y$ corresponds to Coleman's nonabelian plane waves.[8] It follows from (4.1) that the field and its dual are

$$F = \left(dx\,\frac{\partial H}{\partial x} + dy\,\frac{\partial H}{\partial y}\right) \wedge dv,$$
$$\check{F} = \left(dy\,\frac{\partial H}{\partial x} - dx\,\frac{\partial H}{\partial y}\right) \wedge dv \tag{4.3}$$

so that $[A,\check{F}] = 0$ and the current i vanishes. Consider now the same configuration in a different gauge, defined as follows. Let S be a G-valued function of x, y, and v, defined as a solution of the equation

$$S^{-1}\dot{S} + S^{-1}HS = 0, \quad \text{where } \dot{S} = \frac{\partial S}{\partial v}. \tag{4.4}$$

The potential A' obtained from the potential (4.1) by transforming it with S is

$$A' = M\,dx + N\,dy, \tag{4.5}$$

where

$$M = S^{-1}\frac{\partial S}{\partial x} \quad \text{and} \quad N = S^{-1}\frac{\partial S}{\partial y}. \tag{4.6}$$

The transformed field strengths are

$$F' = dv \wedge (\dot{M}\,dx + \dot{N}\,dy),$$
$$\check{F}' = dv \wedge (\dot{M}\,dy - \dot{N}\,dx) \tag{4.7}$$

and

$$4\pi i' = -[A',\check{F}'] = ([M,\dot{M}] + [N,\dot{N}])dv \wedge dx \wedge dy. \tag{4.8}$$

In the transverse gauge (4.5) the current i' has a structure similar to that of the stress–energy vector-valued 3-form t_μ, given by

$$8\pi t_\mu = -((\dot{M}|\dot{M}) + (\dot{N}|\dot{N}))\frac{\partial v}{\partial x^\mu}\,dv \wedge dx \wedge dy. \tag{4.9}$$

The solution given by Eqs. (4.2)–(4.6) may be thus interpreted as representing a wave, moving with the velocity of light, endowed with energy and color densities proportional to $(\dot{M}|\dot{M}) + (\dot{N}|\dot{N})$ and $[M,\dot{M}] + [N,\dot{N}]$, respectively. From the appearance of the commutators in (4.8) one infers that radiation of color—if it exists—is a nonabelian phenomenon, requiring time-dependent and noncommuting sources.

A single pole particle of color q might radiate its charge away if $[q,\dot{q}]$ could be different from 0, but this is not the case, as follows from:

Example (*ii*)[9]: Let $z^\mu(s)$ be the coordinates of a timelike word line z, parametrized by its proper time s. One defines two functions u and r on the Minkowski space $\mathbb{R}^4$ as follows. For any $\xi \in \mathbb{R}^4$, let $u(\xi)$ be the value of s corresponding to the intersection of the word line z with the past-oriented light cone of vertex at ξ and let

$$r(\xi) = g_{\mu\nu}(\xi^\mu - z^\mu(u))\dot{z}^\nu(u). \tag{4.10}$$

Consider a pole particle of an *a priori* time-dependent color $q(u) \in \mathfrak{g}$ moving along z. The Liénard–Wiechert potential,

$$A = q(u)r^{-1}\dot{z}_\mu(u)\,d\xi^\mu, \tag{4.11}$$

is well-defined outside z, i.e., for $r \neq 0$, and the Yang–Mills equation in that region implies

$$\dot{q} + [q,\dot{q}] = 0. \tag{4.12}$$

Assuming that G compact, one obtains from the lemma

$$\dot{q} = 0. \tag{4.13}$$

This implies $i = 0$ so that the wave corresponding to (4.11) is colorless, although it transports energy.

V. BOUNDARY CONDITIONS AND GAUGE TRANSFORMATIONS

The formal definitions of q, Eqs. (3.4) and (3.5), have little meaning because of their unwieldy gauge dependence.[1,10] The total nonabelian charge should be an element of the Lie algebra $\mathfrak{g}$, defined up to "global gauge transformations," i.e., up to replacements of $q \in \mathfrak{g}$ by $a^{-1}qa$, where $a \in G$. If q is so defined, then one can construct out of it invariants, such as $(q|q)$, which provide gauge-independent, global characteristics of the system. A possible way of obtaining such a definition is suggested by the theory of general relativity where one considers total energy contained in "all of space" and expresses it by a surface integral over "a sphere at infinity." This is equivalent, in our case, to taking limits of q as $R \to \infty$, such as those given by Eqs. (3.7) and (3.8). If it can be shown that gauge transformation functions S may be meaningfully restricted to those having for $R \to \infty$ a limit independent of the direction along which one goes to infinity, then the limit of the surface integral (3.5) provides a total charge transforming, under changes of gauge, in the desired manner. The class of allowed gauge transformations depends on gauge configurations under study. In particular, in the case we consider, the asymptotic behavior $(R \to \infty)$ of the gauge transformation functions should be adapted to that of the potentials.[11,12]

We shall make a specific assumption about the behavior of the potentials at large distances from the sources. The assumption may be justified by reference to what is known in linear field theories and by the successes of a similar hypothesis in the theory of gravitation.

From now on we shall use exclusively coordinates $u = t - r$, r, θ, and φ, with respect to which the Minkowski metric is $du^2 + 2du\,dr - r^2(d\theta^2 + \sin^2\theta\,d\varphi^2)$. We consider gauge configurations which may be described by potentials of the form[13]:

$$A = \sum_{k=1}^{\infty} r^{-k} A_k, \tag{5.1}$$

where each A_k is a g-valued form linear in du, dr, $r\,d\theta$, and $r\sin\theta\,d\varphi$, with coefficients depending only on u, θ, and φ. In particular,

$$A_1 = K\,du + L\,dr + r\omega, \tag{5.2}$$

where

$$\omega = M\,d\theta + N\sin\theta\,d\varphi \tag{5.3}$$

and K, L, M, and N are g-valued functions of u, θ, and φ. The form (5.1) is related to, but not equivalent with, the property of the solution to represent outgoing waves and to satisfy the Sommerfeld radiation condition.[3,14] The field strengths corresponding to (5.1) admit a similar expansion,

$$F = \sum_{k=1}^{\infty} r^{-k} F_k, \tag{5.4}$$

where

$$F_1 = du \wedge (\dot{L}\,dr + r\dot{\omega}) \tag{5.5}$$

and the dot denotes, from now on, a derivative with respect to u. We might have included in (5.1) a term of the form $H(u,\theta,\varphi)\,du$, which would have contributed to F_1. Such a term can, however, be gauge transformed, and absorbed in ω, in a manner similar to what has been done in Example (*i*) of Section IV.

A gauge transformation induced by a function of the form

$$S = a(\theta,\varphi)[I + r^{-1}\alpha(u,\theta,\varphi) + \cdots] \tag{5.6}$$

preserves (5.1). If gauge-transformed quantities are distinguished by primes, then

$$K' = a^{-1}Ka + \dot{\alpha}, \quad L' = a^{-1}La, \tag{5.7}$$

$$\omega' = a^{-1}\omega a + a^{-1}\,da. \tag{5.8}$$

It is important to note that, *a priori*, a may be an arbitrary smooth function on S_2. The occurrence of such a *function* makes it difficult to define the total nonabelian charge: at large distances, the field strengths F transform according to $F' = a^{-1}Fa$. In particular, the radial component θ of the $1/r^2$ part of the electric field transforms in this way. Therefore, its surface integral may be changed, essentially at will, by choosing an appropriate function a.[10,12]

In addition to the "generic" transformation functions (5.6) there may be some special ones, also preserving (5.1). For example:

(i) If c is a central element of g, then $S = \exp(c \log r)$ induces the change $A \mapsto A' = A + cr^{-1}\,dr$. Transformations of this form are used to eliminate the $r^{-1}L\,dr$ term from the electromagnetic potential.

(ii) If G contains $SL(2,\mathbb{R})$ as a subgroup, then its Lie algebra admits two nonzero elements X and Y such that $[X,Y] = X$. The potential $A = r^{-1}(Xu + Y)\,dt$ is a solution of the Yang–Mills equations of the form (5.1).[9] The function $S = \exp(tX)$ transforms it, however, to the Coulomb form $A' = r^{-1}Y\,dt$.

VI. THE ASYMPTOTIC EXPANSION

We now consider in more detail the asymptotic behavior of a gauge configuration produced by localized, time-dependent sources.[15] We assume that the current j falls off, at large distances, as r^{-4} or faster. In analogy with (5.2) we write

$$A_2 = P\,du + Q\,dr + r\pi, \tag{6.1}$$

where π is a g-valued 1-form linear in $d\theta$ and $d\varphi$; P, Q, and the coefficients of π depend on u, θ, and φ.

Under the generic gauge transformations (5.6), the form ω behaves like a gauge potential on S_2. One can associate with it a field strength,

$$d'\omega + \tfrac{1}{2}[\omega,\omega] = -B_r\,d\theta \wedge \sin\theta\,d\varphi, \tag{6.2}$$

where d' denotes the restriction of the exterior derivative to S_2,

$$-B_r = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}N\sin\theta - \frac{1}{\sin\theta}\frac{\partial M}{\partial\varphi} + [M,N]. \tag{6.3}$$

Moreover, for any u-dependent, g-valued differential form $\Phi = \Phi^i e_i$ on S_2, transforming as $\Phi \mapsto a^{-1}\Phi a$ under (5.6), one defines its gauge derivative with respect to ω as

$$D\Phi^i = d'\Phi^i + c^i_{jk}\omega^j \wedge \Phi^k. \tag{6.4}$$

Let Ψ denote the left-hand side of the Yang–Mills equations (2.5). The g-valued 3-form Ψ may be written as

$$\Psi = r^2(R\,du + U\,dr) \wedge d\theta \wedge \sin\theta\,d\varphi + du \wedge dr \wedge r\Xi, \tag{6.5}$$

where R and U are g-valued functions and Ξ is a g-valued form, linear in $d\theta$ and $d\varphi$. If the expansions (5.1) and (5.4) are introduced, then

$$R = \sum_{k=1}^{\infty} r^{-k}R_k,$$

$$U = \sum_{k=1}^{\infty} r^{-k}U_k, \quad \text{and} \quad \Xi = \sum_{k=1}^{\infty} r^{-k}\Xi_k, \tag{6.6}$$

where the quantities with subscripts are constant along the null rays $u,\theta,\varphi = \text{const}$. The Yang–Mills equations without sources are now equivalent to the infinite system $R_k = U_k = \Xi_k = 0$, $k = 1,2,\cdots$. Among these equations we consider all those which involve only A_1 and A_2.

In the lowest order ($k = 1$), U_1 and Ξ_1 are identically zero and

$$R_1 = \ddot{L}. \tag{6.7}$$

For $k = 2$ one obtains the following:

$$U_2 = \dot{L} + [L,\dot{L}]. \tag{6.8}$$

Let us assume from now on that g admits a positive-definite, invariant scalar product. Remembering the lemma, one obtains from $U_2 = 0$

$$\dot{L} = 0. \tag{6.9}$$

Equation (6.9) will be assumed to hold from now on. The equation $R_2 = 0$ is equivalent to

$$^{*}\dot{E}_r = D\,{}^{*}\dot{\omega}, \tag{6.10}$$

where

$$E_r = \dot{Q} + K + [K,L]. \tag{6.11}$$

The function α occurring in the gauge transformation (5.7) may be used to reduce K' to zero so that $E'_r = \dot{Q}'$. Such a choice of gauge may be convenient in computations, but will not be adopted here because it requires time-dependent potentials for a static, Coulomb-like configuration. The equation $\Xi_2 = 0$ is equivalent to

$$[L,\dot{\omega}] = 0 \tag{6.12}$$

and $U_3 = 0$ gives

$$D\,{}^\star DL = [L,{}^\star E_r]. \tag{6.13}$$

Since the invariant scalar product of L with $[L$, anything$]$ vanishes, Eq. (6.13) implies

$$h_{ij} L^i D\,{}^\star DL^j = 0 \tag{6.14}$$

or, equivalently,

$$d(h_{ij} L^i\,{}^\star DL^j) - h_{ij} DL^i \wedge {}^\star DL^j = 0. \tag{6.15}$$

The invariance condition (2.10) has been used, in conjunction with Eq. (6.4), to go over from (6.14) to (6.15). By integrating both sides of Eq. (6.15) over S_2 and taking into account that h is positive definite, one obtains

$$DL = 0 \tag{6.16}$$

so that Eq. (6.13) reduces to

$$[L,E_r] = 0. \tag{6.17}$$

The last equation one has to consider is $\Xi_3 = 0$, or

$$\dot{\pi} + [\dot{\pi},L] = \tfrac{1}{2}(DK + [DK,L] + {}^\star DB_r - \frac{\partial}{\partial u} DQ - [\dot{\omega},Q]). \tag{6.18}$$

It follows from (6.9) and (6.16) that $\|L\|^2 = (L|L)$ is constant; if it is nonzero then one can define the projection q_L of color in the direction of L by

$$4\pi q_L = \int_{S_2} ({}^\star E_r|L)/\|L\|. \tag{6.19}$$

Since L transforms according to (5.7), the integral (6.19) is well-defined (gauge-independent). Moreover, from Eqs. (6.10) and (6.16) is follows that q_L does not depend on u. The same is true of the "magnetic" (dual) charge m_L,

$$4\pi m_L = \int_{S_2} ({}^\star B_r|L)/\|L\|. \tag{6.20}$$

For completeness, we give the explicit form of the $1/r$ and $1/r^2$ terms in the field strengths after the field equations (6.9) and (6.16) have been taken into account:

$$F = du \wedge \dot{\omega} + r^{-2} E_r\, du \wedge dr + r^{-1} du \wedge (\dot{\pi} - DK) - {}^\star B_r + O(r^{-3}). \tag{6.21}$$

It is clear from (6.21) that $r^{-2}E_r$ and $r^{-2}B_r$ are the radial components of, respectively, the electric and magnetic $1/r^2$ parts of the field strengths.

All solutions to our equations can be divided into two classes depending on whether $L \neq 0$ or $L = 0$.

(i) If $L \neq 0$ then one can choose (6.9), (6.10), and (6.16)–(6.18), with E_r defined by (6.11), as independent equations. This is a rather strong system of equations: e.g., for $G = \mathrm{SU}(2)$ it implies that E_r, B_r and $\dot{\omega}$ are all parallel to L.

E. T. Newman raised the following problem: Are there solutions of the Yang–Mills equations, of the form considered in this paper, for which L cannot be reduced to zero by a gauge transformation? We have no complete answer to this question, but wish to make the following comments.

(a) The Lorentz condition $d\check{A} = 0$ implies the following restriction

$$E_r = L + [K,L] + {}^\star d'\,{}^\star\omega \tag{6.22}$$

which is incompatible with $L = 0$ if $q \neq 0$.

(b) If the gauge potential is of the form

$$A = r^{-1} L\, dr + \omega + \bar{A}, \tag{6.23}$$

Eq. (6.16) holds, and if

$$[L,\bar{A}] = 0, \tag{6.24}$$

then the gauge transformation induced by

$$S = \exp(-L \log r) \tag{6.25}$$

eliminates the L term without affecting ω and $\bar{A}$, i.e.,

$$S^{-1}AS + S^{-1}dS = \omega + \bar{A}. \tag{6.26}$$

One cannot however, expect (6.24) to hold in general and $S^{-1}\bar{A}S$ may contain $\log r$ terms prohibited by the assumption inherent in Eq. (5.1).

(c) The analog of L vanishes for the Robinson–Trautman solutions[16] of Einstein's field equations.

(ii) If $L = 0$ then the field equations reduce to only two,

$$\dot{Q} + \dot{K} = {}^\star D^\star \dot{\omega}, \tag{6.27}$$

$$\dot{\pi} = \frac{1}{2}(DK + {}^\star DB_r - \frac{\partial}{\partial u} DQ - [\dot{\omega},Q]). \tag{6.28}$$

Given arbitrary K and ω one can integrate (6.27) and (6.28) to find Q and π, respectively. Formally, the total (retarded) color charge and its rate of change may be computed from

$$4\pi q_{\rm ret}(u) = \int_{S_2} {}^\star E_r = \int_{S_2} (Q + K)\, d\theta \wedge \sin\theta\, d\varphi, \tag{6.29}$$

$$4\pi \dot{q}_{\rm ret}(u) = \int_{S_2} D^\star \dot{\omega} = \int_{S_2} [\omega,{}^\star\dot{\omega}]. \tag{6.30}$$

The significance of these formulas is limited by the occurrence of the arbitrary function a: $S_2 \to G$ in the gauge transformation (5.6). The function a can be restricted to be a constant if (i) the solution is spherically symmetric, or (ii) $A = O(r^{-2})$, i.e., $A_1 = 0$. In the case of the Liénard–Wiechert solution one also does not encounter a difficulty because of the spherical symmetry of the Coulomb, r^{-2} part of the field in each of the instantaneous rest systems of the particle. These simple examples suggest that it may be possible to eliminate the direction-dependent function a by reference to some properties of the gauge configurations, e.g., those at past or future infinity.

ACKNOWLEDGMENTS

Both authors, at different times, had enjoyed the hospitality of the mathematics community at Université de Montréal where this paper was completed. The first author had an opportunity to discuss its contents during his visits to the Université Catholique at Louvain-la-Neuve. The second author enjoyed similar discussions at ICTP in Trieste, at the universities in Vienna, Edmonton, Chicago, Pittsburgh, Stony Brook, and at New York University. We have been influenced by conversations with P. Aichelburg, J. Anandan, R. Beig, J. Harnad, R. Jackiw, J. Winicour, and P. Win-

ternitz. We are grateful to J. Łopuszanski who showed us the preprint of a paper by S. Schlieder where the ambiguity in the definition of nonabelian charges is discussed in detail. We thank J. Kowalski–Glikman who checked some of the computations. We acknowledge with particular gratitude criticism by E. T. Newman which led to the present, revised version of the paper.

[1]B. D. Bramson, Proc. Roy. Soc. London Ser. A **341**, 463 (1975).

[2]H. Bondi, M. G. J. Van der Burg, and A. W. K. Metzner, Proc. Roy. Soc. London Ser. A **269**, 21 (1962).

[3]R. K. Sachs, Phys. Rev. **128**, 2851 (1962).

[4]R. Roskies, Phys. Rev. D **15**, 722 (1977); J. Anandan and K. P. Tod, Phys. Rev. D **18**, 1144 (1978).

[5]A. Trautman, "Fiber Bundles, Gauge Fields, and Gravitation," in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980), Vol. I.

[6]R. Penrose, Proc. Roy. Soc. London Ser. A **284**, 159 (1965); in *Relativity, Groups and Topology* (Gordon and Breach, New York, 1964).

[7]R. Guven, Phys. Rev. D **19**, 471 (1979); W. B. Campbell and T. A. Morgan, Phys. Lett. B **84**, 87 (1979); A. Trautman, J. Phys. A **13**, L1 (1981).

[8]S. Coleman, Phys. Lett. B **70**, 59 (1977).

[9]A. Trautman, Phys. Rev. Lett. **46**, 875 (1981).

[10]S. Schlieder, Nuovo Cimento A **63**, 137 (1981).

[11]M. Lüscher, Nucl. Phys. B **140**, 429 (1978).

[12]R. Jackiw, Rev. Mod. Phys. **52**, 661 (1980).

[13]A. I. Janis and E. T. Newman, J. Math. Phys. **6**, 902 (1965); W. G. Dixon, J. Math. Phys. **11**, 1238 (1970); E. T. Newman, Phys. Rev. D **18**, 2901 (1978); **22**, 3023 (1980).

[14]D. J. Lamb, J. Math. Phys. **7**, 458 (1966).

[15]A. Trautman, Acta Phys. Austriaca Suppl. **23**, 401 (1981).

[16]I. Robinson and A. Trautman, Proc. Roy. Soc. London Ser. A **265**, 463 (1962).

Time evolution kernels: uniform asymptotic expansions

T. A. Osborn
Department of Physics, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada
Y. Fujiwara
School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455

(Received 3 August 1982; accepted for publication 29 October 1982)

For a wide class of self-adjoint Schrödinger Hamiltonians, a detailed description of the time evolution kernel is obtained. In a setting of a d-dimensional Euclidean space without boundaries, the Schrödinger Hamiltonian H is the sum of the negative Laplacian plus a real-valued local potential $v(x)$. The class of potentials studied is the family of bounded and continuous functions that are formed from the Fourier transforms of complex bounded measures. These potentials are suitable for the N-body problem, since they do not necessarily decrease as $|x|\rightarrow\infty$. An asymptotic expansion in the complex parameter z, around $z=0$, is derived for the family of kernels $U_z(x,y)$ corresponding to the analytic semigroup $\{e^{-zH}:\mathrm{Re}\, z>0\}$, which is uniform in the coordinate variables x and y. The asymptotic expansion has a simple semiclassical interpretation. Furthermore, an explicit bound for the remainder term in the asymptotic expansion is found. The expansion and the remainder term bound continue to the time axis boundary $z=it/\hbar\,(t\neq 0)$ of the analytic semigroup domain.

PACS numbers: 03.65. — w

I. INTRODUCTION AND SUMMARY

Suppose H is the generator of time evolution for the N-body Schrödinger problem in nonrelativistic quantum mechanics. This paper investigates kernel representations of the analytic semigroup family

$$\{e^{-zH}:z\in\mathbb{C},\mathrm{Re}\, z>0\} \tag{1.1}$$

as well as that of the time evolution operator family $\{e^{-i(t/\hbar)H}:t\in\mathbb{R},\ t\neq 0\}$. In particular, a uniform asymptotic expansion in powers of z of the semigroup kernels is found. Because the coefficient functions of the asymptotic expansion are polynomials in Planck's constant h, the asymptotic expansion may be interpreted as a semiclassical expansion in the limit $h\rightarrow 0$.[1–4] The asymptotic expansion is accompanied by a remainder term with an explicit bound. This expansion and the error bound are valid on the time axis as well as the analytic semigroup domain $\mathrm{Re}\, z>0$. A novel feature of the asymptotic expansion is that it is uniformly valid for all values of the coordinates x and y.

We choose a mathematical setting of the Schrödinger problem that is sufficiently general to include N-body quantum mechanics. Take x to be a position vector in a d-dimensional Euclidean space $\mathbb{R}^d$. If each individual particle has a mass m and moves in three dimensions, then $d=3N$. The quantum scale factor in this situation is

$$q=\hbar^2/2m. \tag{1.2}$$

Wave functions for spinless particles are elements of the Hilbert space $\mathscr{H}=L^2(\mathbb{R}^d)$. This Hilbert space has inner product (f,g), where f,g are elements of $\mathscr{H}$. The inner product is taken to be antilinear in the left argument and linear in the right. The symbol $\|f\|$ denotes the norm $(f,f)^{1/2}$ on $\mathscr{H}$. If A is a linear operator $A:\mathscr{H}\rightarrow\mathscr{H}$, then $\|A\|$ is taken as the operator norm. Other $L^p(\mathbb{R}^d)$ norms will be introduced as they are needed and denoted by the symbol of $\|\cdot\|_{L^p}$.

Consider the Hamiltonians that define the Schrödinger problem. If Δ_x is the Laplacian in $\mathbb{R}^d$, then take H_0 to be the unique self-adjoint extension on $\mathscr{H}$ of $-q\Delta_x$. The full Hamiltonian H is defined by the operator sum

$$H=H_0+V, \tag{1.3}$$

where the perturbation V is also assumed to be a bounded ($\|V\|<\infty$) and self-adjoint operator; namely V is the operator determined by multiplication with a continuous, bounded and real-valued function $v(x)$,

$$(Vf)(x)=v(x)f(x),\quad f(x)\in L^2(\mathbb{R}^d). \tag{1.4}$$

Since V is bounded, H and H_0 share a common domain $\mathscr{D}\subset\mathscr{H}$. However, it is not assumed that $v(x)$ has any decay as $|x|\rightarrow\infty$. Such a decay assumption would prohibit the treatment of the N-body problem, which is characterized by nondecaying potentials.

Take Λ to be the spectrum of H and $\{E_\lambda:\lambda\in\Lambda\}$ to be the unique family of spectral projectors defined by H. Since V is bounded, H is bounded from below; i.e., $H+\|V\|\cdot I\geqslant 0$ and $\Lambda\subseteq[-c,\infty)$ with $c=\|V\|$. The linear bounded operator e^{-zH} on $\mathscr{H}=L^2(\mathbb{R}^d)$ is defined in terms of its spectral integral

$$e^{-zH}=\int_{-c}^{\infty}e^{-z\lambda}dE_\lambda,\quad \mathrm{Re}\, z\geqslant 0. \tag{1.5}$$

The fact that H is unbounded from above implies that the integral (1.5) defines a bounded operator only if $\mathrm{Re}\, z\geqslant 0$. The parameter z has the standard physical interpretation. If z is positive and $z=\beta=(kT)^{-1}$, where k is the Boltzmann's constant and T is the absolute temperature, then $\{e^{-\beta H}:\beta\in\mathbb{R},\beta>0\}$ defines the family of the semigroup operators connected with the heat transport equation. If z is imaginary and $z=it/\hbar$, where t is the time displacement, then $\{e^{-i(t/\hbar)H}:t\in\mathbb{R}\}$ defines the family of the time evolution operators. From here on, in order to simplify the notation for the time evolution operators we will replace $t/\hbar$ by t. On the other hand, if we take a complex domain $D=\{z\in\mathbb{C}:\mathrm{Re}\, z>0\}$, then $\{e^{-zH}:z\in D\}$ defines a family of analytic semigroup operators. In this paper, we treat the ker-

 0022-2488/83/051093-11$02.50

nel representation of the operator e^{-zH} of Eq. (1.5) in the domain $\bar{D}\setminus\{0\} = \{z\in\mathbb{C}:\mathrm{Re}\, z\geqslant 0 \text{ and } z\neq 0\}$, which includes the heat kernels and the time evolution kernels as special cases. In a similar fashion, the unperturbed version of (1.5) is defined by

$$e^{-zH_0} = \int_0^\infty e^{-z\lambda} dE_\lambda^0, \quad \mathrm{Re}\, z\geqslant 0, \tag{1.6}$$

where $\{E_\lambda^0:\lambda\geqslant 0\}$ is the family of spectral projectors connected with H_0.

The kernel $U_z(x,y)$ of e^{-zH} for $z\in\bar{D}\setminus\{0\}$ is described in detail. In particular, since this kernel has an essential singularity at $z=0$, it is useful to introduce the factorization

$$U_z(x,y) = U_z^{(0)}(x-y)F(x,y;z), \tag{1.7}$$

where $U_z^{(0)}(x-y)$ is the convolution kernel corresponding to e^{-zH_0}. The function $F(x,y;z)$ is analytic in D, continuous in $\bar{D} = \{z\in\mathbb{C}:\mathrm{Re}\, z\geqslant 0\}$ and $F(x,y;0) = 1$ for $x,y\in\mathbb{R}^d$. By employing a constructive Born series expansion of the kernel $U_z(x,y)$, it is possible to find an explicit series representation of $F(x,y;z)$. Restructuring this series representation provides a uniform (in $x,y\in\mathbb{R}^d$) asymptotic series for $F(x,y;z)$ in the form

$$F(x,y;z) = \sum_{n=0}^{M-1} ((-z)^n/n!)P_n(x,y) + E_M(x,y;z). \tag{1.8}$$

The coefficient functions $P_n(x,y)$ are all well known[1-3]; they are polynomials in v of the order n and polynomials in the quantum scale parameter q of order $n-1$. The error term $E_M(x,y;z)$ is of order $O(|z|^M)$. A uniform bound for $E_M(x,y;z)$ in $x,y\in\mathbb{R}^d$ is obtained. The number of terms M in the expansion (1.8) is proportional to the number of bounded derivatives the potential $v(x)$ possesses.

The program this paper embarks on is a special case of "local geometrical asymptotics of continuum eigenfunction expansions" recently outlined by Fulling.[5] See also Simon.[6] This program generalizes the classical investigations of the asymptotic density of eigenvalues of the operator H

$$H\psi_j = \lambda_j\psi_j, \quad \psi_j\in\mathscr{H} \tag{1.9}$$

as $\lambda_j\to\infty$. This investigation is carried out by studying the interrelationships between the kernels of the semigroup e^{-zH}, the resolvent $(H-z)^{-1}$, and the measures defined by the spectral kernels $e(x,y;\lambda)$ of the projectors $\{E_\lambda\}$. In our approach the emphasis has been placed on controlling the behavior of the time evolution kernels of e^{-itH}. It is customary[7] to use the $M=1$ and $z=\beta>0$ version of the asymptotic expansion (1.8), together with the Tauberian theorem in order to derive the large λ asymptotic behavior of $e(x,y;\lambda)$. As a rule this Laplace-transform Tauberian approach gives one only the first term of the large λ asymptotic expansion for $e(x,y;\lambda)$. In a subsequent paper, a method to utilize the full M-term expansion of $F(x,y;z)$ will be developed. As a consequence, it will be possible to find an M-term asymptotic expansion of $e(x,y;\lambda)$ that is analogous to Eq. (1.8).

The investigation in this paper is carried out for the class of potentials $v(x)$ that are represented by Fourier transforms of bounded complex measures supported on $\mathbb{R}^d$. Ito[8] and Albeverio and Høegh-Krohn[9] have used this class of potentials to study the Feynman path integral representation of e^{-itH}. We use this class in order to obtain the constructive series representation of the kernels for e^{-zH}. This class of potentials has a number of attractive mathematical properties. (For details, one should refer to Ref. 9).

Let $\mathscr{M}(\mathbb{R}^d)$ be the set of all bounded complex measures defined on the Borel field $\mathscr{B}$ on $\mathbb{R}^d$. For each measure $\mu\in\mathscr{M}(\mathbb{R}^d)$, we can define a potential function by the Fourier transform of μ;

$$v(x) = \int_{\mathbb{R}^d} e^{i\alpha x} d\mu(\alpha), \tag{1.10}$$

where αx denotes the scalar product of two vectors in $\mathbb{R}^d$. The potential thus defined is bounded and continuous, since

$$|v(x)| \leqslant \int_{\mathbb{R}^d} d\,|\mu|(\alpha) = |\mu|(\mathbb{R}^d) < \infty, \quad x\in\mathbb{R}^d, \tag{1.11}$$

where $|\mu|(e)$ $(e\in\mathscr{B})$ is the total variation of $\mu(e)$. We define the norm of μ in $\mathscr{M}(\mathbb{R}^d)$ by

$$\|\mu\| = \int_{\mathbb{R}^d} d\,|\mu|(e) = |\mu|(\mathbb{R}^d). \tag{1.12}$$

Then the operator norm of V in (1.4) satisfies

$$\|V\| \leqslant \|\mu\|, \tag{1.13}$$

so that we can take $c = \|\mu\|$ in (1.5). The transform in (1.10) defines a natural class $\mathscr{F}$ by

$$\mathscr{F} = \{v(x) = \int_{\mathbb{R}^d} e^{i\alpha x} d\mu(\alpha):\mu\in\mathscr{M}(\mathbb{R}^d)\}, \tag{1.14}$$

which is a subset of the space composed of all bounded and continuous functions.[10] The elements of the spaces $\mathscr{F}$ and $\mathscr{M}(\mathbb{R}^d)$ are in one-to-one correspondence. This is a consequence of the uniqueness of the transform (1.10) that states $v(x)=0$ if and only if $\mu=0$.[11]

The reality condition on $v(x)$ is satisfied if the measure obeys the reflection property $\mu(-e) = \overline{\mu(e)}$ for all $e\in\mathscr{B}$, where $-e = \{x\in\mathbb{R}^d: -x\in e\}$ and the bar denotes the complex conjugate. We will denote by $\mathscr{M}^r(\mathbb{R}^d)$ the space of measures $\mu\in\mathscr{M}(\mathbb{R}^d)$ that satisfy the reflection property. $\mathscr{F}^r$ will indicate the Fourier image of $\mathscr{M}^r(\mathbb{R}^d)$.

It is convenient in our study of asymptotic expansions to characterize potentials that have bounded partial derivatives of order M or less. Thus for a positive integer M, we define a set of functions

$$\mathscr{F}^r_M = \{v(x)\in\mathscr{F}^r: \int_{\mathbb{R}^d} |\alpha|^n d\,|\mu|(\alpha) < \infty \ \text{ for } n = 0,1,2,\ldots,M\}, \tag{1.15}$$

where μ is the measure connected with $v(x)$. In fact, if $v(x)\in\mathscr{F}^r_M$, then there exists a smallest finite positive constant K (depending on μ and M) such that

$$\int_{\mathbb{R}^d} |\alpha|^n d\,|\mu|(\alpha) \leqslant K^n\|\mu\| \quad \text{for } n = 0,1,\ldots,M. \tag{1.16}$$

We call K the *bound constant of the measure μ in the space $\mathscr{F}^r_M$*. Suppose D_x^L denotes an arbitrary partial derivative in $\mathbb{R}^d$ multi-indexed by $L = (l_1,l_2,\ldots,l_d)$ with the length $|L| \equiv l_1 + l_2 + \cdots + l_d$ $(l_i\geqslant 0)$. If $\{x_i; i = 1,\ldots,d\}$ are the Cartesian components of x, then D_x^L is the partial derivative

$$D_x^L = \left(\frac{\partial}{\partial x_1}\right)^{l_1}\left(\frac{\partial}{\partial x_2}\right)^{l_2}\cdots\left(\frac{\partial}{\partial x_d}\right)^{l_d}. \tag{1.17}$$

For $v(x)\in\mathcal{F}_M^r$, then

$$|(D_x^L v)(x)| \leqslant K^{|L|}\|\mu\| \quad \text{for all } |L| \leqslant M. \tag{1.18}$$

Potentials in class $\mathcal{F}^r$ are suitable for discussing the N-body problem, since they do not require any decay as $|x|\to\infty$. Further, the class $\mathcal{F}^r$ has periodic potentials. Thus our treatment of asymptotic expansions for the time evolution operators is also applicable to the problem of particles moving in a periodic medium or an almost periodic medium.

The asymptotic expansions we describe have been discussed heuristically in two recent papers.[2,3] In these papers, the asymptotic expansion was obtained formally for the special case of the heat equation operators $e^{-\beta H}$, but no estimate for the error terms were found. However, this first treatment did succeed in completely determining the coefficient functions $P_n(x,y)$. In particular, it was shown that there is a simple algorithm that constructs $P_n(x,y)$ in terms of connected graphs. Further, the physical interpretation of the semiclassical content of the expansion (1.8) and its relationship to the Wigner–Kirkwood semiclassical expansions[3,4] of the quantum partition function has been analyzed in detail.

The construction of this paper is as follows. In Sec. II the integral kernel $U_\beta(x,y)$ of the semigroup $\{e^{-\beta H}\}_{\beta>0}$ is derived by using the iterative formula with respect to the matrix elements of $e^{-\beta H}$ and $e^{-\beta H_0}$. In Sec. III we investigate the analytic property of $U_z(x,y)$ in z and try to extend the integral kernel representation into the whole domain D of the analytic semigroup and to the imaginary time axis $z = it$ $(t\neq 0)$. In Sec IV the uniform asymptotic expansions of the kernel $U_z(x,y)$ and $F(x,y;z)$ are derived together with an explicit estimate of the remainder terms.

II. INTEGRAL KERNELS OF THE SEMIGROUP

In this section, we will introduce the integral kernels which represent the semigroup family $\{e^{-\beta H}\}_{\beta>0}$. The following facts summarize the well-known properties[12] of the semigroup family:

(i) $e^{-\beta_1 H}e^{-\beta_2 H} = e^{-(\beta_1+\beta_2)H}, \quad \beta_1,\beta_2>0,$ (2.1)

(ii) $\underset{\beta\to 0+}{\text{s-lim}}\, e^{-\beta H}u = u, \quad u\in L^2(\mathbb{R}^d),$ (2.2)

(iii) if $u\in\mathcal{D}$, then $e^{-\beta H}u\in\mathcal{D}$ for $\beta>0$ and the strong derivative $(d/d\beta)e^{-\beta H}u$ exists and

$$\frac{d}{d\beta}e^{-\beta H}u = -He^{-\beta H}u, \quad \beta>0. \tag{2.3}$$

These properties and the similar ones for H_0 give us the next lemma about a type of Born series expansion of the matrix elements of $e^{-\beta H}$.

In order to simplify the expressions entering this Born expansion we will use

$$\int_>^1 \{\cdots\}d^n\xi$$

to represent the n-fold iterated integral

$$\int_{1>\xi_1>\cdots>\xi_n>0}\cdots\int\{\cdots\}d\xi_1\,d\xi_2\cdots d\xi_n.$$

Similarly, we use the abbreviation

$$\int_0^1 d^n\xi = \int_0^1 d\xi_1\int_0^1 d\xi_2\cdots\int_0^1 d\xi_n.$$

Finally, the n-fold integral of measures $d\mu(\alpha_i)$ will be indicated by

$$\int d^n\mu = \int d\mu(\alpha_1)\cdots\int d\mu(\alpha_n).$$

Lemma 1: Suppose the operator V is bounded and self-adjoint on $\mathcal{H}$ and $H = H_0 + V$. Let $u,w\in L^2(\mathbb{R}^d)$ and $\beta>0$; then

$$(1)\ (w,e^{-\beta H}u) = (w,e^{-\beta H_0}u) - \int_0^\beta d\beta'(w,e^{-\beta' H}Ve^{-(\beta-\beta')H_0}u), \tag{2.4}$$

$$\begin{aligned}(2)\ (w,e^{-\beta H}u) &= (w,e^{-\beta H_0}u)\\ &= \sum_{n=1}^{N-1}(-\beta)^n\int_>^1 d^n\xi\\ &\quad\times(w,e^{-\beta\xi_n H_0}Ve^{-\beta(\xi_{n-1}-\xi_n)H_0}V\\ &\quad\times\cdots\times Ve^{-\beta(1-\xi_1)H_0}u) + R_N,\end{aligned} \tag{2.5}$$

where

$$R_N = (-\beta)^N\int_>^1 d^N\xi\,(w,e^{-\beta\xi_N H}\times Ve^{-\beta(\xi_{N-1}-\xi_N)H_0}V\ldots Ve^{-\beta(1-\xi_1)H_0}u) \tag{2.6}$$

and R_N has bound

$$|R_N| \leqslant \frac{(\beta\|V\|)^N}{N!}e^{\beta\|V\|}\|w\|\,\|u\|. \tag{2.7}$$

Proof: (1) Take $u,w\in L^2(\mathbb{R}^d)$ and set $f_{w,u}^\beta(\beta')\equiv(w,e^{-\beta' H}Ve^{-(\beta-\beta')H_0}u)$ for $\beta\geqslant\beta'\geqslant 0$. Then using (1.5) and (1.6), we get

$$f_{w,u}^\beta(\beta') = \int_{-c}^\infty\int_0^\infty e^{-\beta'\lambda-(\beta-\beta')\lambda'}d^2(E_\lambda w, VE_{\lambda'}^0 u).$$

Since the integral is bounded by

$$\int_{-c}^\infty\int_0^\infty e^{+\beta c}d^2|(E_\lambda w,VE_{\lambda'}^0 u)| \leqslant e^{\beta c}\|w\|\,\|V\|\,\|u\| < \infty,$$

$f_{w,u}^\beta(\beta')$ is continuous in $\beta\geqslant\beta'\geqslant 0$. On the other hand, if $u\in\mathcal{D}$, then $H = H_0 + V$ and the above property (iii) leads us to

$$\begin{aligned}f_{w,u}^\beta(\beta') &= (w,He^{-\beta' H}e^{-(\beta-\beta')H_0}u) - (w,e^{-\beta' H}H_0e^{-(\beta-\beta')H_0}u)\\ &= -\frac{\partial}{\partial\beta'}(w,e^{-\beta' H}e^{-(\beta-\beta')H_0}u).\end{aligned}$$

Thus $(w,e^{-\beta' H}e^{-(\beta-\beta')H_0}u)$ is differentiable in β' for $\beta>\beta'>0$. Integrating $f_{w,u}^\beta(\beta')$ from 0 to β and using the property (ii), we get (2.4) for $w\in L^2(\mathbb{R}^d)$ and $u\in\mathcal{D}$. The $w,u\in L^2(\mathbb{R}^d)$ case is the consequence of $\bar{\mathcal{D}} = L^2(\mathbb{R}^d)$ and the boundedness of V.

(2) $N = 1$ case is simply the integral variable change of $\beta'\to\xi_1$ by $\beta' = \beta\xi_1$ in (1). The general case is obtained by iterating (1) and changing the integral variables after that. The bound in (2.7) is obtained from $\|e^{-\beta H}\|\leqslant e^{\beta\|V\|}$. □

The matrix elements of the iterated operators in (2.5) will be calculated by using the next lemma. We state this lemma for all $z\in\bar{D}$, instead of $\beta>0$, because of its later usefulness in analytical continuation.

Lemma 2: Let $\hat{u}(k)\in L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)$. Suppose $u(x)$ is the Fourier transform of $\hat{u}(k)$,

$$u(x)=\frac{1}{(2\pi)^{d/2}}\int e^{ikx}\hat{u}(k)dk \quad \text{a.a. } x; \tag{2.8}$$

then

(1) $u(x)\in L^2(\mathbb{R}^d)$ and $\|u\|=\|\hat{u}\|$,

(2) For all $z\in\bar{D}$, we have

$$(e^{-zH_0}u)(x)=\frac{1}{(2\pi)^{d/2}}\int e^{-zqk^2+ikx}\hat{u}(k)dk \quad \text{a.a. } x. \tag{2.9}$$

Thus

$$|(e^{-zH_0}u)(x)|\leqslant\frac{1}{(2\pi)^{d/2}}\|\hat{u}\|_{L^1} \quad \text{a.a. } x \tag{2.10}$$

and

$$\|e^{-zH_0}u\|\leqslant\|\hat{u}\|=\|u\|. \tag{2.11}$$

(3) For all $z\in\bar{D}$ and $1\geqslant\xi_1\geqslant\cdots\xi_n\geqslant 0$ $(n=1,2\cdots)$, we get

$$(e^{-z\xi_nH_0}Ve^{-z(\xi_{n-1}-\xi_n)H_0}V\ldots Ve^{-z(1-\xi_1)H_0}u)(x)$$
$$=\int g_{\xi_1\cdots\xi_n}(x,k;z)\hat{u}(k)dk \quad \text{a.a. } x, \tag{2.12}$$

where

$$\begin{aligned}g_{\xi_1\cdots\xi_n}(x,k;z)&\equiv\frac{1}{(2\pi)^{d/2}}\int d^n\mu\exp\{-zq\xi_n(k+\alpha_1+\cdots+\alpha_n)^2\\&\quad-zq(\xi_{n-1}-\xi_n)(k+\alpha_1+\cdots+\alpha_{n-1})^2-\cdots\\&\quad-zq(1-\xi_1)k^2+i(k+\alpha_1+\cdots+\alpha_n)x\}\end{aligned} \tag{2.13}$$

$$=\frac{1}{(2\pi)^{d/2}}\int d^n\mu\times\exp\left\{-zq\sum_{l,m=0}^{n}\xi_l\wedge\xi_m\alpha_l\alpha_m+i\left(\sum_{l=0}^{n}\alpha_l\right)x\right\} \tag{2.14}$$

with the convention $\xi_0=1,\alpha_0=k$, and $\xi_l\wedge\xi_m\equiv\mathrm{Min}\{\xi_l,\xi_m\}$. The function $g_{\xi_1\cdots\xi_n}(x,k;z)$ has a pointwise bound

$$|g_{\xi_1\cdots\xi_n}(x,k;z)|\leqslant\frac{\|\mu\|^n}{(2\pi)^{d/2}}, \quad x,k\in\mathbb{R}^d. \tag{2.15}$$

Thus (2.12) has the bound

$$|(e^{-z\xi_nH_0}Ve^{-z(\xi_{n-1}-\xi_n)H_0}V\ldots Ve^{-z(1-\xi_1)H_0}u)(x)|\leqslant\frac{\|\mu\|^n}{(2\pi)^{d/2}}\|\hat{u}\|_{L^1} \tag{2.16}$$

for a.a. x.

Furthermore,

$$\|e^{-z\xi_nH_0}Ve^{-z(\xi_{n-1}-\xi_n)H_0}V\ldots Ve^{-z(1-\xi_1)H_0}u\|\leqslant\|\mu\|^n\|\hat{u}\|=\|\mu\|^n\|u\|. \tag{2.17}$$

Proof: Results (1) and (2) are standard (for details see Ref. 13). (3) is the generalized version of (2), since $n=0$ case reduces to (2). We shall prove (3) by induction. We assume (2.12), (2.13) and (2.16), (2.17) for $n\to n-1$. Let the numerical multiplier, ξ_{n-1}, of the left most exponential operator in (2.12) be replaced by $\xi_{n-1}-\xi_n$. Now multiplying by $v(x)$, we get

$$\begin{aligned}&(Ve^{-z(\xi_{n-1}-\xi_n)H_0}V\ldots Ve^{-z(1-\xi_1)H_0}u)(x)\\&\overset{\text{a.e.}}{=}\frac{1}{(2\pi)^{d/2}}\int d\mu(\alpha_n)e^{i\alpha_nx}\int dk\,\hat{u}(k)\int d^{n-1}\mu\\&\quad\times\exp\{-zq(\xi_{n-1}-\xi_n)(k+\alpha_1+\cdots+\alpha_{n-1})^2\\&\quad-\cdots-zq(1-\xi_1)k^2+i(k+\alpha_1+\cdots+\alpha_{n-1})x\}.\end{aligned} \tag{2.18}$$

Here we can change the integral order arbitrarily, since the integral is absolutely convergent. Change the integral variable k into $k-\alpha_1-\cdots-\alpha_n$ and move k-integral to the left-most position. Then (2.18) is written as

$$(Ve^{-z(\xi_{n-1}-\xi_n)H_0}\cdots u)(x)\overset{\text{a.e.}}{=}\frac{1}{(2\pi)^{d/2}}\int dk\,e^{ikx}\Phi_n(k), \tag{2.19}$$

where

$$\begin{aligned}\Phi_n(k)=\int d^n\mu\exp\{&-zq(\xi_{n-1}-\xi_n)(k-\alpha_n)^2-\cdots\\&-zq(1-\xi_1)(k-\alpha_1-\cdots-\alpha_n)^2\}\\&\times\hat{u}(k-\alpha_1-\cdots-\alpha_n).\end{aligned} \tag{2.20}$$

Note that (2.20) satisfies

$$|\Phi_n(k)|\leqslant\int d\,|\mu|\,(\alpha_1)\cdots\int d\,|\mu|\,(\alpha_n)|\hat{u}(k-\alpha_1-\cdots-\alpha_n)| \tag{2.21}$$

since $1\geqslant\xi_1\geqslant\cdots\geqslant\xi_n\geqslant 0$ and $\mathrm{Re}\,z\geqslant 0$. Therefore we get

$$\|\Phi_n\|_{L^1}\leqslant\|\mu\|^n\|\hat{u}\|_{L^1}<\infty \tag{2.22}$$

and

$$\|\Phi_n\|^2\leqslant(\|\mu\|^n\|\hat{u}\|)^2<\infty. \tag{2.23}$$

Estimate (2.23) is an immediate consequence of Hölder's inequality.

Equations (2.19), (2.22), and (2.23) allow us to apply (2) for $u\to Ve^{-z(\xi_{n-1}-\xi_n)H_0}\cdots u$ and $z\to z\xi_n\in\bar{D}$. Then using (2.20), we get

$$\begin{aligned}&(e^{-z\xi_nH_0}Ve^{-z(\xi_{n-1}-\xi_n)H_0}\cdots u)(x)\\&\overset{\text{a.e.}}{=}\frac{1}{(2\pi)^{d/2}}\int dk\,e^{-zq\xi_nk^2+ikx}\int d^n\mu\\&\quad\times\exp\{-zq(\xi_{n-1}-\xi_n)(k-\alpha_n)^2-\cdots\\&\quad-zq(1-\xi_1)(k-\alpha_1-\cdots-\alpha_n)^2\}\hat{u}(k-\alpha_1-\cdots-\alpha_n).\end{aligned} \tag{2.24}$$

A couple of changes of integral order and the integral variable change $k\to k+\alpha_1+\cdots+\alpha_n$ give us (2.12) and (2.13). Equation (2.16) follows from (2.10) with (2.22), and (2.17) follows from (2.11) with (2.23). Equation (2.14) is easily obtained by rearranging the exponential factor of (2.13). □

The following proposition gives us the kernel representation of the semigroup family. $\mathscr{S}(\mathbb{R}^d)$ will denote the Schwartz space on $\mathbb{R}^d$.

Proposition 1: Assume $v\in\mathscr{F}^r$.

(1) For all $u,w\in\mathscr{S}(\mathbb{R}^d)$ and all $\beta>0$,

$$(w,e^{-\beta H}u)=\int dx\int dy\,U_\beta^{(0)}(x-y)F(x,y;\beta)\,\overline{w(x)}u(y), \tag{2.25}$$

where

$$U_\beta^{(0)}(x) \equiv \left[\frac{1}{4\pi\beta q}\right]^{d/2} \exp\left\{-\frac{|x|^2}{4\beta q}\right\} \tag{2.26}$$

and

$$F(x,y;\beta) \equiv 1 + \sum_{n=1}^{\infty} (-\beta)^n \int_>^1 d^n\xi \int d^n\mu \times \exp\left\{-\beta q \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x + \xi_l y)\alpha_l\right\} \tag{2.27}$$

with

$$\theta(\xi_l,\xi_m) \equiv \mathrm{Min}\{\xi_l(1-\xi_m),\xi_m(1-\xi_l)\}. \tag{2.28}$$

$F(x,y;\beta)$ is a uniformly convergent series in x and y for an arbitrary fixed $\beta \geqslant 0$ and has the bound

$$|F(x,y;\beta)| \leqslant e^{\beta\|\mu\|}, \quad x,y \in \mathbb{R}^d, \quad \beta \geqslant 0. \tag{2.29}$$

(2) For all $\beta > 0$,

$$U_\beta(x,y) \equiv U_\beta^{(0)}(x-y)F(x,y;\beta) \tag{2.30}$$

defines a bounded integral operator U_β on $L^2(\mathbb{R}^d)$ with the operator norm $\|U_\beta\| \leqslant e^{\beta\|\mu\|}$ and

$$(e^{-\beta H}u)(x) = \int dy\, U_\beta(x,y)u(y) \quad \text{a.a. } x \tag{2.31}$$

and for all $u \in L^2(\mathbb{R}^d)$.

Proof: (1) Let's start from Lemma 1, statement (2) with $N \to \infty$. If we use Lemma 2, statement (3) with $z = \beta$, then for all $u,w \in \mathscr{S}(\mathbb{R}^d)$ and all $\beta > 0$, we get

$$(w,e^{-\beta H}u) = (\omega,U_\beta^{(0)}u) + \lim_{N\to\infty}\sum_{n=1}^{N}(-\beta)^n \int_>^1 d^n\xi \times \int dx\, \overline{w(x)} \int dk\, g_{\xi_1\cdots\xi_n}(x,k;\beta)\hat{u}(k). \tag{2.32}$$

Here, if we use (2.14), we get

$$\psi_\beta^{\xi_1\cdots\xi_n}(x) \equiv \int dk\, g_{\xi_1\cdots\xi_n}(x,k;\beta)\hat{u}(k) = \int dy\, u(y) \int d^n\mu \frac{1}{(2\pi)^d} \int dk \exp\left\{-\beta q \times \sum_{l,m=0}^{n} (\xi_l \wedge \xi_m)\alpha_l\alpha_m + i\left(\sum_{l=0}^{n}\alpha_l\right)x - iky\right\}, \tag{2.33}$$

where we have used the interchangeability of integral order since the k^2-term in the exponential factor is $-\beta q(\xi_0 \wedge \xi_0)\alpha_0\alpha_0 = -\beta q k^2 \leqslant 0$. The full exponential factor expression is

$$-\beta q \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l - \frac{(x-y)^2}{4\beta q} - \beta q\left(k + \sum_{l=1}^{n}\xi_l\alpha_l - i\frac{x-y}{2\beta q}\right)^2. \tag{2.34}$$

Performing the Gaussian integral over k, we get for (2.33),

$$\psi_\beta^{\xi_1\cdots\xi_n}(x) = \int dy\, u(y)U_\beta^{(0)}(x-y)\int d^n\mu \times \exp\left\{-\beta q \sum_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l\right\}. \tag{2.35}$$

Thus (2.32) becomes

$$(w,e^{-\beta H}u) = (w,U_\beta^{(0)}u) + \lim_{N\to\infty}\sum_{n=1}^{N}(-\beta)^n\int_>^1 d^n\xi \int dx\, \overline{w(x)} \int dy\, u(y)U_\beta^{(0)}(x-y) \times \int d^n\mu \exp\left\{-\beta q\sum_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l\right\}. \tag{2.36}$$

Note that the integral of each term in (2.36) is absolutely convergent, since

$$\sum_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m = \sum_{l=1}^{n}\left(\frac{1}{\xi_l} - \frac{1}{\xi_{l-1}}\right)\left\{\sum_{m=l}^{n}\xi_m\alpha_m\right\}^2 \geqslant 0, \tag{2.37}$$

for $1 = \xi_0 \geqslant \xi_1 \geqslant \cdots \geqslant \xi_n \geqslant 0$, and

$$\int dx \int dy\, |\overline{w(x)}U_\beta^{(0)}(x-y)u(y)| \leqslant \frac{1}{(4\pi\beta q)^{d/2}}\|w\|_{L^1}\|u\|_{L^1} < \infty. \tag{2.38}$$

Furthermore, since

$$\left|\sum_{n=1}^{N}(-\beta)^n\int_>^1 d^n\xi\int d^n\mu \times \exp\left\{-\beta q\sum_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l\right\}\right| \leqslant \sum_{n=1}^{N}\frac{(\beta\|\mu\|)^n}{n!} \leqslant e^{\beta\|\mu\|} - 1 \quad \text{for } N = 1,2\ldots, \tag{2.39}$$

we can take $N \to \infty$ inside of the integrals over x and y. Thus we get (2.25) with (2.26)–(2.28). Equation (2.29) is obtained by taking $N \to \infty$ in (2.39).

(2) Let's consider

$$\psi_\beta(x) = \int U_\beta(x,y)u(y)dy \quad \beta > 0, \ u \in L^2(\mathbb{R}^d). \tag{2.40}$$

Although $U_\beta(x,y)$ is not a convolution kernel, it has a bound which is a convolution kernel; namely

$$|U_\beta(x,y)| \leqslant e^{\beta\|\mu\|}U_\beta^{(0)}(x-y) \quad \text{for } \beta > 0. \tag{2.41}$$

from (2.29). Therefore, the integral in (2.40) has bound

$$|\psi_\beta(x)| \leqslant e^{\beta\|\mu\|}\int U_\beta^{(0)}(x-y)|u(y)|dy.$$

Since $\int U_\beta^{(0)}(x)dx = 1$ and $|u(x)| \in L^2(\mathbb{R}^d)$, the Hausdorff–Young inequality for convolutions[14] gives us

$$\|\psi_\beta\| \leqslant e^{\beta\|\mu\|}\|u\| \tag{2.42}$$

which means that the integral in (2.40) is absolutely convergent for a.e. in x and the integral operator U_β has the operator norm bound of $\|U_\beta\| \leqslant e^{\beta\|\mu\|}$.

Let's prove (2.31). Since both sides of (2.31) define the bounded operators on $L^2(\mathbb{R}^d)$ with common operator norm bound $e^{\beta\|\mu\|}$, we only need to prove it for $u(x) \in \mathscr{S}(\mathbb{R}^d)$. However, this is obvious, since (2.25) means that

$$(w,f) = 0 \quad \text{for all } w \in \mathscr{S}(\mathbb{R}^d),$$

where $f(x) = (e^{-\beta H}u)(x) - \psi_\beta(x) \in L^2(\mathbb{R}^d)$. □

III. INTEGRAL KERNELS OF ANALYTIC SEMIGROUP AND TIME EVOLUTION OPERATORS

The objective of this section is to extend the kernel description $U_\beta(x,y)$ of the semigroup family $\{(e^{-\beta H}:\beta>0\}$ to include the family $\{e^{-zH}:z\in\bar{D}\setminus\{0\}\}$. Observe that the family of time evolution operators $\{e^{-itH}:t\neq 0\}$ is a subset of this enlarged class of operators. The necessity of the omission of the point $z=0$ is obvious because $U_\beta^{(0)}(x-y)$ has an essential singularity at $\beta=0$. The basic method we employ is a mixture of analytic continuation in the form of the matrix elements of e^{-zH} plus a control of the operator norm bound of the integral kernels. The next lemma gives a detailed norm bound for e^{-zH} and establishes the analytic behavior of the matrix elements of e^{-zH}.

Lemma 3: Let V be a bounded self-adjoint operator on $L^2(\mathbb{R}^d)$.

(1) For $z\in\bar{D}$, e^{-zH} is a bounded operator on $L^2(\mathbb{R}^d)$ with norm bound

$$\|e^{-zH}\| \leqslant e^{\operatorname{Re} z\|V\|}. \tag{3.1}$$

(2) For all $w,u\in L^2(\mathbb{R}^d)$

$$f_{w,u}(z) = (w,e^{-zH}u) \tag{3.2}$$

is a holomorphic function of z in D and is continuous in $\bar{D}$. It has bound

$$|f_{w,u}(z)| \leqslant e^{\operatorname{Re} z\|V\|}\|w\|\|u\| \quad \text{for } z\in\bar{D}. \tag{3.3}$$

Proof: (1) is an immediate consequence of (1.5). For the proof of (2), let's set

$$h_{w,u}(z) \equiv e^{-z\|V\|}f_{w,u}(z) = \int_{-c}^{\infty} e^{-z(\|V\|+\lambda)}d(w,E_\lambda u). \tag{3.4}$$

The integral is bounded by

$$\int_{-c}^{\infty} d|(w,E_\lambda u)| \leqslant \|w\|\,\|u\| < \infty,$$

since $|e^{-z(\|V\|+\lambda)}|\leqslant 1$ for $z\in\bar{D}$ and $\lambda\geqslant -\|V\|$. Thus $h_{w,u}(z)$ is continuous in $\bar{D}$. The holomorphy is a consequence of Fubini's theorem on the interchange of order of iterated integrals and Morera's theorem. Choose C to be an arbitrary simple closed rectifiable contour of finite length lying inside of D. Then

$$\oint_C h_{w,u}(z)dz = \int_{-c}^{\infty}\left\{\oint_C e^{-z(\|V\|+\lambda)}dz\right\}d(w,E_\lambda u) = 0.$$

Multiplying $e^{z\|V\|}$ onto $h_{w,u}(z)$, we get the continuity and holomorphy of $f_{w,u}(z)$ in $\bar{D}$ and D, respectively. Equation (3.3) is implied by (1). □

The next step is to define a holomorphic function in D that is suggested by the series of (2.27) in Proposition 1. We have

Definition 1: (1) The function $F(x,y;z):\mathbb{R}^d\times\mathbb{R}^d\times\bar{D}\to\mathbb{C}$ is defined as the absolutely convergent sum

$$F(x,y;z) = \lim_{N\to\infty}\sum_{n=0}^{N} B_n(x,y;z), \tag{3.5}$$

$$\begin{aligned} B_n(x,y;z) = \frac{(-z)^n}{n!}\int_0^1 d^n\xi\int d^n\mu \times\exp\Big\{ &-zq\sum_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m \\ &+ i\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l\Big\}, \end{aligned} \tag{3.6}$$

where $\theta(\xi_l,\xi_m)$ is (2.28) and $B_0(x,y;z)=1$.

(2) The functions $U_z^{(0)}(x)$ and $U_z(x,y)$ for $z\in\bar{D}\setminus\{0\}$ are defined by

$$U_z^{(0)}(x) = \left[\frac{1}{4\pi zq}\right]^{d/2}\exp\left\{-\frac{|x|^2}{4zq}\right\}, \tag{3.7}$$

$$U_z(x,y) = U_z^{(0)}(x-y)F(x,y;z). \tag{3.8}$$

A summary of the properties of $F(x,y;z)$ that follow from the series (3.5) is

Proposition 2: Let $v\in\mathscr{F}^r$. The function $F(x,y;z)$ satisfies

(a) *Boundedness*. Let D_0 be any arbitrary bounded domain $D_0\subset\bar{D}$. The series (3.5) is absolutely and uniformly convergent for all $x,y,z\in\mathbb{R}^d\times\mathbb{R}^d\times D_0$. The sum of (3.5), $F(x,y;z)$ has the bound

$$|F(x,y;z)| \leqslant e^{|z|\|\mu\|} \quad \text{for } z\in\bar{D}. \tag{3.9}$$

(b) *Continuity and holomorphy*. $F(x,y;z)$ is holomorphic in D and continuous in $\bar{D}$. $F(x,y;z)$ is jointly continuous in x,y everywhere in $\mathbb{R}^d\times\mathbb{R}^d$.

Proof: (a) The absolute and uniform convergences are a consequence of the nonnegative nature of $\Sigma_{l,m=1}^{n}\theta(\xi_l,\xi_m)\alpha_l\alpha_m$ in (2.37). For all $z\in\bar{D}$, the series for $F(x,y;z)$ is majorized term-by-term by $(1/n!)(|z|\,\|\mu\|)^n$. So (a) is proved.

(b) Note that each term of (3.5), $B_n(x,y;z)$, is holomorphic in D and continuous in $\bar{D}$ for each fixed $x,y\in\mathbb{R}^d\times\mathbb{R}^d$. These properties are transmitted to $F(x,y;z)$ by the uniformly convergent nature of (a). The joint-continuity in x,y is also a consequence of the uniform convergence relative to x and y plus the fact that each term is jointly continuous. □

The next proposition shows that the result of Proposition 1, statement (1) for $\beta>0$ can be extended to $z\in\bar{D}\setminus\{0\}$ by the analyticity and the continuity of the matrix elements of e^{-zH}.

Proposition 3: Let $v\in\mathscr{F}^r$. For $z\in\bar{D}\setminus\{0\}$ and $u,w\in\mathscr{S}(\mathbb{R}^d)$,

$$(w,e^{-zH}u) = \int dx\int dy\ \overline{w(x)}U_z(x,y)u(y). \tag{3.10}$$

Proof: Let's set

$$f_{w,u}(z) = (w,e^{-zH}u),$$

$$g_{w,u}(z) = \int dx\int dy\ \overline{w(x)}\ U_z(x,y)u(y) = \int dx\int dy\ \overline{w(x)}\ U_z^{(0)}(x-y)F(x,y;z)u(y).$$

Since $u,w\in\mathscr{S}(\mathbb{R}^d)$, Lemma 3 and Proposition 2 lead us to the result that $f_{w,u}(z)$ and $g_{w,u}(z)$ are both holomorphic in D and continuous in $\bar{D}\setminus\{0\}$. On the other hand, Proposition 1, statement (1) shows

$$f_{w,u}(\beta) = g_{w,u}(\beta) \quad \text{for } \beta > 0.$$

By the identity theorem for holomorphic functions we get

$$f_{w,u}(z) = g_{w,u}(z) \quad \text{for } z \in D.$$

The continuity of $f_{w,u}(z)$ and $g_{w,u}(z)$ in $\bar{D} \setminus \{0\}$ completes the proof. □

We need to know several properties of the integral operator defined by

$$\psi_z(x) = \int U_z(x,y)u(y)dy, \tag{3.11}$$

in order to recover the kernel representation of e^{-zH} from (3.10). In particular, the property $\psi_z(x) \in L^2(\mathbb{R}^d)$ plays the essential role. This can be easily shown for $z \in D$, if we use an argument similar to the discussion in the proof of Proposition 1, statement (2). On the other hand the case $z = it$ $(t \neq 0)$ necessitates a more detailed estimate. The next lemma, which is based on Lemma 2, permits us to establish the boundedness of the integral operator in (3.11) for $z \in \bar{D} \setminus \{0\}$ in a restricted sense.

Lemma 4: Let $v \in \mathcal{F}^r$. For $u(x) \in \mathcal{S}(\mathbb{R}^d)$, $z \in \bar{D} \setminus \{0\}$ and $1 \geqslant \xi_1 \geqslant \cdots \geqslant \xi_n \geqslant 0$, we get

$$\psi_z^{\xi_1 \cdots \xi_n}(x) \equiv \int dk g_{\xi_1 \cdots \xi_n}(x,k;z)\hat{u}(k) = \int dy U_z^{(0)}(x-y) h_{\xi_1 \cdots \xi_n}(x,y;z) u(y), \tag{3.12}$$

where $\hat{u}(k) \in \mathcal{S}(\mathbb{R}^d)$ is the Fourier transform of $u(x)$ and

$$h_{\xi_1 \cdots \xi_n}(x,y;z) \equiv \int d^n\mu \exp\Big\{ -zq \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n} ((1-\xi_l)x + \xi_l y)\alpha_l \Big\}. \tag{3.13}$$

$\psi_z^{\xi_1 \cdots \xi_n}(x)$ is a L^2-function and has norm bound

$$\|\psi_z^{\xi_1 \cdots \xi_n}\| \leqslant \|\mu\|^n \|u\|, \quad z \in \bar{D} \setminus \{0\}. \tag{3.14}$$

Proof: Let us set for $B > 0$

$$I_B \equiv \int_{|k|<B} dk\, g_{\xi_1 \cdots \xi_n}(x,k;z)\hat{u}(k).$$

Since the integral range is finite and $\hat{u}(k), u(x) \in \mathcal{S}(\mathbb{R}^d)$, we can follow the same process as used to get (2.35). Thus, we find

$$I_B = \int dy\, u(y) U_z^{(0)}(x-y) \int d^n\mu \exp\Big\{ -zq \times \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m + i\sum_{l=1}^{n}((1-\xi_l)x + \xi_l y)\alpha_l \Big\} J_B, \tag{3.15}$$

where

$$J_B = \frac{(4\pi zq)^{d/2}}{(2\pi)^d} \int_{|k|<B} dk \times \exp\Big\{ -zq\Big(k + \sum_{l=1}^{n} \xi_l\alpha_l - i\frac{x-y}{2zq}\Big)^2 \Big\}. \tag{3.16}$$

Since $\operatorname{Re} z \geqslant 0$ and $z \neq 0$, a simple calculation of the complex integral gives us

$$\lim_{B\to\infty} J_B = \frac{(4\pi zq)^{d/2}}{(2\pi)^d} \left[\frac{\pi}{zq}\right]^{d/2} = 1. \tag{3.17}$$

Thus if we take B large enough, we can find a B-independent constant, which permits us to take the limit $B \to \infty$ in (3.15) inside of the integrals of $\int dy \int d^n\mu$, by Lebesgue's dominated convergence theorem. Equation (3.14) is the combined consequence of (3.12) and Lemma 2, statement 3. □

Proposition 4: Let $v \in \mathcal{F}^r$.

(1) For all $z \in D$, $U_z(x,y)$ in (3.8) defines an integral operator U_z on $L^2(\mathbb{R}^d)$ with the operator norm bound

$$\|U_z\| \leqslant e^{|z|\|\mu\|}. \tag{3.18}$$

(2) For all $t \neq 0$, $U_{it}(x,y)$ defines an integral operator U_{it} on $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ with the operator norm bound

$$\|U_{it}\| \leqslant e^{|t|\|\mu\|}. \tag{3.19}$$

Proof: We first prove the integral transformation of (3.11), $U_z: u \to \psi_z$, defines a bounded operator for $z \in \bar{D} \setminus \{0\}$ with the domain $\mathcal{S}(\mathbb{R}^d)$ and an operator norm bound $\|U_z\| \leqslant e^{|z|\|\mu\|}$. In fact, if $u(x) \in \mathcal{S}(\mathbb{R}^d)$ and $z \in \bar{D} \setminus \{0\}$, the integral in (3.11) is absolutely convergent for $x \in \mathbb{R}^d$, since

$$|U_z(x,y)| \leqslant |U_z^{(0)}(x-y)| e^{|z|\|\mu\|}, \tag{3.20}$$

$$|U_z^{(0)}(x-y)| = \left[\frac{1}{4\pi|z|q}\right]^{d/2} \exp\left[-\frac{\operatorname{Re} z}{4|z|^2 q}(x-y)^2 \right] \tag{3.21}$$

$$\leqslant \left[\frac{1}{4\pi|z|q}\right]^{d/2}, \quad z \in \bar{D} \setminus \{0\}, \tag{3.22}$$

follows from (3.7)–(3.9). Furthermore, $\psi_z(x)$ in (3.11) is seen to be a bounded and continuous function if $z \in \bar{D} \setminus \{0\}$ and $u(x) \in \mathcal{S}(\mathbb{R}^d)$. Note the relationship

$$B_n(x,y;z) = \frac{(-z)^n}{n!} \int_0^1 d^n\xi\, h_{\xi_1 \cdots \xi_n}(x,y;z)$$

between (3.6) and (3.13). These functions have pointwise bounds

$$|h_{\xi_1 \cdots \xi_n}(x,y;z)| \leqslant \|\mu\|^n \tag{3.23}$$

and

$$|B_n(x,y;z)| \leqslant \frac{(|z|\|\mu\|)^n}{n!} \tag{3.24}$$

for $x,y \in \mathbb{R}^d$ and $z \in \bar{D} \setminus \{0\}$. Then (3.11) becomes

$$\begin{aligned} \psi_z(x) &= \int dy\, u(y) U_z^{(0)}(x-y) \lim_{N\to\infty} \sum_{n=0}^{N} B_n(x,y;z) \\ &= \sum_{n=0}^{\infty} \frac{(-z)^n}{n!} \int_0^1 d^n\xi \times \int dy\, u(y) U_z^{(0)}(x-y) h_{\xi_1 \cdots \xi_n}(x,y;z), \end{aligned} \tag{3.25}$$

where the estimates (3.22)–(3.24) and the fact that $u(y) \in \mathcal{S}(\mathbb{R}^d)$ allow us to use Fubini's theorem to justify changing the order of integration. Further, the dominated convergence theorem permits us to interchange the order of the limiting process. If we use $\psi_z^{\xi_1 \cdots \xi_n}$ in Lemma 4, (3.25) is written as

$$\psi_z(x) = \sum_{n=0}^{\infty} \frac{(-z)^n}{n!} \int_0^1 d^n\xi\, \psi_z^{\xi_1 \cdots \xi_n}(x). \tag{3.26}$$

[Note that $\psi_z(x)$ is actually continuous in x since it is the sum of a uniformly convergent series of functions.] Thus we get, for all $z \in \bar{D} \setminus \{0\}$ and all $u(x) \in \mathcal{S}(\mathbb{R}^d)$,

$$\|\psi_z\| \leqslant \sum_{n=0}^{\infty} \frac{|z|^n}{n!} \Big\| \int_0^1 d^n\xi \, \psi_z^{\xi_1 \cdots \xi_n}(x) \Big\|$$
$$\leqslant \sum_{n=0}^{\infty} \frac{|z|^n}{n!} \int_0^1 d^n\xi \, \|\psi_z^{\xi_1 \cdots \xi_n}\|$$
$$\leqslant \sum_{n=0}^{\infty} \frac{(|z|\,\|\mu\|)^n}{n!} \|u\| = e^{|z|\|\mu\|}\|u\|, \tag{3.27}$$

where we have used (3.14).

If we restrict z to belong to D, then the convolution bound of $U_z(x,y)$ by (3.20) and (3.21) makes it possible to define the integral of (3.11) for all $u(y) \in L^2(\mathbb{R}^d)$ and a.a.$x \in \mathbb{R}^d$. [See the proof of Proposition 1, statement (2)]. Since $\mathscr{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, the extension principle for bounded operators leads us to the conclusion of (1).

The same discussion is also applicable in the case of (2), since, in this case, the integral in (3.11) is absolutely convergent for each $x \in \mathbb{R}^d$ by (3.20)and (3.22) if $u(x) \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. □

We note that the operator norm bounds (3.18) and (3.19) for e^{-zH} are less precise than those given in Lemma 3, Eq. (3.1). This situation seems to be inevitable, since we formed the estimate of $\|\psi_z\|$ using a term-by-term norm bound of the Born series in (3.26).

Finally we get the following theorem about kernel representations of the analytic semigroup and the time evolution operators.

Theorem 1: Let $v \in \mathscr{F}^r$.

(1) If $z \in D$ and $u(x) \in L^2(\mathbb{R}^d)$, then

$$(e^{-zH}u)(x) = \int dy \; U_z(x,y)u(y). \tag{3.28}$$

(2) Suppose $t \neq 0$, if $u(x) \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ then

$$(e^{-itH}u)(x) = \int dy \; U_{it}(x,y)u(y) \tag{3.29}$$

and if $u(x) \in L^2(\mathbb{R}^d)$ then

$$(e^{-itH}u)(x) = \text{s} - \lim_{B\to\infty} \int_{|y|<B} dy \; U_{it}(x,y)u(y). \tag{3.30}$$

Relations (3.28) and (3.29) are valid for a.a. $x \in \mathbb{R}^d$.

Proof: Consider (3.28) and (3.29). We only need to prove them for all $u(x) \in \mathscr{S}(\mathbb{R}^d)$, since both integral operators are known to be bounded operators in $L^2(\mathbb{R}^d)$, and $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, respectively, from Proposition 4. However, this is obvious, since Proposition 3 shows

$$(w,f) = 0 \quad \text{for } w \in \mathscr{S}(\mathbb{R}^d)$$

with $f(x) = (e^{-zH}u)(x) - \psi_z(x) \in L^2(\mathbb{R}^d)$ for $z \in \bar{D} \setminus \{0\}$ and $u(x) \in \mathscr{S}(\mathbb{R}^d)$.

Equation (3.30) is proved by taking

$$u_B(x) \equiv \begin{cases} u(x) & \text{for } |x| \leqslant B, \\ 0 & \text{for } |x| > B, \end{cases}$$

and using (3.29) and the fact that $\|u - u_B\| \to 0$ as $B \to \infty$. □

For C^∞ potentials, $v(x)$, results similar to Propositions 2–4 and Theorem 1 for the time-evolution kernel have been recently obtained by Fujiwara,[15] Kitada and Kumano-go,[16] and Zelditch.[17]

IV. UNIFORM ASYMPTOTIC EXPANSIONS

This section derives the small z asymptotic expansion for $F(x,y;z)$. The asymptotic expansion for $F(x,y;z)$ implies a corresponding asymptotic expansion for the kernel $U_z(x,y)$ of the operator e^{-zH}. Our analysis aims at relating the structure of the asymptotic expansions as well as remainder term bounds to the smoothness of the potential $v(x)$. The expansions are all uniformly valid in the coordinate variables $x,y \in \mathbb{R}^d$ and in any compact subset of the analytic semigroup domain D. The expansions remain valid on time axis boundary except for $t = 0$.

Lemma 5: (1) For all $z \in \mathbb{C}$, the exponential function has an N term estimate

$$e^{-z} = \sum_{n=0}^{N-1} \frac{1}{n!}(-z)^n + H_N(z) \quad (N = 1,2,\cdots), \tag{4.1}$$

where

$$H_N(z) = (-z)^N \int_> d^N\xi \, e^{-z\xi_N}. \tag{4.2}$$

For Re $z \geqslant 0$,

$$|H_N(z)| \leqslant \frac{|z|^N}{N!}. \tag{4.3}$$

(2) For $1 \geqslant \xi_1 \geqslant \cdots \geqslant \xi_n \geqslant 0$ and $\alpha_1,\ldots,\alpha_n \in \mathbb{R}^1$ $(n = 1,2,3,\cdots)$, we have

$$0 \leqslant \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m \leqslant \frac{n}{4}\sum_{l=1}^{n} \alpha_l^2. \tag{4.4}$$

Proof: (1) It is simple to check that (4.1) is true for $N = 1$. The general case can be shown by substituting

$$e^{-z\xi_N} = 1 - z\xi_N \int_0^1 d\xi \, e^{-z\xi_N\xi}$$
$$= 1 - z\int_0^{\xi_N} d\xi_{N+1} e^{-z\xi_{N+1}}$$

into (4.2).

(2) Nonnegativeness is (4.4) is already shown in (2.37). To show the upper bound, we use the Schwartz's inequality:

$$\sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)\alpha_l\alpha_m$$
$$\leqslant \Big\{ \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)^2 \Big\}^{1/2} \Big\{ \sum_{l,m=1}^{n} (\alpha_l\alpha_m)^2 \Big\}^{1/2}.$$

For $\xi_l \leqslant \xi_m$, for example, one has

$$\theta(\xi_l,\xi_m) = \xi_l(1-\xi_m) \leqslant \xi_l(1-\xi_l) \leqslant \frac{1}{4}.$$

Thus

$$\Big\{ \sum_{l,m=1}^{n} \theta(\xi_l,\xi_m)^2 \Big\}^{1/2} \leqslant \frac{1}{4}\Big\{ \sum_{l,m=1}^{n} 1 \Big\}^{1/2} = \frac{n}{4}$$

and

$$\Big\{ \sum_{l,m=1}^{n} (\alpha_l\alpha_m)^2 \Big\}^{1/2} \leqslant \Big\{ \sum_{l=1}^{n}\alpha_l^2 \sum_{m=1}^{n}\alpha_m^2 \Big\}^{1/2} = \sum_{l=1}^{n}\alpha_l^2.$$

Combining these results establishes the upper bound in (4.4). □

The basic idea used in obtaining the uniform asymptotic expansion of $F(x,y;z)$ is to find the asymptotic expansions for $B_n(x,y;z)$ and then sum the expansions over n. Once more it is helpful to introduce a couple of abbreviations. Let a_n and b_n denote the recurring real exponential arguments

$$a_n(\xi_1\cdots\xi_n;\alpha_1\cdots\alpha_n)\equiv\sum_{i,j=1}^{n}\theta(\xi_i,\xi_j)\alpha_i\alpha_j \tag{4.5}$$

and

$$b_n(\xi_1\cdots\xi_n;\alpha_1\cdots\alpha_n;x,y)\equiv\sum_{l=1}^{n}((1-\xi_l)x+\xi_l y)\alpha_l. \tag{4.6}$$

For $B_n(x,y;z)$ we have

Proposition 5: Let $v\in\mathcal{F}^{r}_{2M}$ and take K to be the corresponding bound constant in the family $\mathcal{F}^{r}_{2M}$. For all $z\in\bar{D}$ and $n\geqslant 1$,

$$B_n(x,y;z)=\sum_{m=0}^{M-1}\frac{(-z)^{n+m}}{(n+m)!}q^m D_{m,n+m}(x,y)+R_{n,M}(x,y;z). \tag{4.7}$$

The functions $D_{m,n+m}(x,y)$ are real valued, jointly and uniformly continuous in $\mathbb{R}^d\times\mathbb{R}^d$ and represented by the multiple integrals

$$D_{m,n+m}(x,y)=\binom{n+m}{m}\int_0^1 d^n\xi\int d^n\mu(a_n)^m e^{ib_n} \tag{4.8}$$

for $n\geqslant 1$, $m\geqslant 0$. Furthermore $D_{0,0}(x,y)=1$ and $D_{m,m}(x,y)=0$ if $m\geqslant 1$. For $n\geqslant 1$, $D_{m,n+m}(x,y)$ satisfies the estimate

$$|D_{m,n+m}(x,y)|\leqslant\binom{n+m}{m}\|\mu\|^n\left(\frac{n^2K^2}{4}\right)^m \tag{4.9}$$

for $m\leqslant M$. For all $x,y\in\mathbb{R}^d$ and all $z\in\bar{D}$, the error term $R_{n,M}(x,y;z)$ has the bound

$$|R_{n,M}(x,y;z)|\leqslant\frac{(|z|\,\|\mu\|)^n}{n!M!}\left(\frac{|z|qn^2K^2}{4}\right)^M. \tag{4.10}$$

Proof: Start with integral (3.6) that defines $B_n(x,y;z)$. Use Lemma 5, statement (1) for the exponential argument

$$e^{-zqa_n}=\sum_{m=0}^{M-1}\frac{1}{m!}(-zqa_n)^m+H_M(zqa_n). \tag{4.11}$$

Inserting expression (4.5) for a_n into (3.6) gives (4.7) and (4.8). The error term is the integral

$$R_{n,M}(x,y;z)=\frac{(-z)^n}{n!}\int_0^1 d^n\xi\int d^n\mu\, e^{ib_n}H_M(zqa_n). \tag{4.12}$$

Employing estimate (4.3) for H_M and utilizing (4.4) to bound a_n gives us

$$\int d|\mu|(\alpha_1)\cdots\int d|\mu|(\alpha_n)|a_n|^M\leqslant\|\mu\|^n\left(\frac{n^2K^2}{4}\right)^M. \tag{4.13}$$

Equations (4.13) and (4.12) give the estimate (4.10).

The reality of $D_{m,n+m}(x,y)$ arises because, under reflection $\alpha_i\to-\alpha_i (i=1,\ldots,n)$ the integrand in (4.8) changes into its complex conjugate and $d\mu(\alpha_i)\to d\bar{\mu}(\alpha_i)$. The bound (4.9) is an immediate consequence of (4.13). □

Many asymptotic expansion techniques[18] require that not only Eq. (4.7) but also the z-derivative of Eq. (4.7) is meaningful. In this direction we have a corollary of Proposition 5. Let $B_n^{(i)}(x,y;z)$ and $R_{n,M}^{(i)}(x,y;z)$ denote the ith derivative with respect to z of $B_n(x,y;z)$ and $R_{n,M}(x,y)$, respectively. If $z=it$ $(t\neq 0)$ the ith derivative is understood to be $(\partial/i\partial t)^i$.

Corollary 1: Let $v\in\mathcal{F}^{r}_{2(M+i)}$ and K be the related bound constant. For all $z\in\bar{D}$ and $n\geqslant 1$,

$$B_n^{(i)}(x,y;z)=\sum_{m=\max(0,i-n)}^{M-1}\frac{(-1)^{n+m}z^{n+m-i}}{(n+m-i)!}\times q^m D_{m,n+m}(x,y)+R_{n,M}^{(i)}(x,y;z). \tag{4.14}$$

The error term is $O(|z|^{n+M-i})$ and has a bound

$$|R_{n,M}^{(i)}(x,y;z)|\leqslant\frac{(|z|\,\|\mu\|)^n}{n!M!}\left(\frac{|z|qn^2K^2}{4}\right)^M\times\left\{(n+M)\left[\frac{1}{|z|}+\frac{qn^2K^2}{4}\right]\right\}^i. \tag{4.15}$$

Proof: The error term $R_{n,M}^{(i)}(x,y;z)$ is the integral

$$R_{n,M}^{(i)}(x,y;z)=\frac{(-1)^{n+M}}{n!}\int_0^1 d^n\xi\int d^n\mu\int_{>}^{1} d^M\xi'\times(qa_n)^M e^{ib_n}\left(\frac{\partial}{\partial z}\right)^{(i)}\{z^{n+M}e^{-zqa_n\xi'_M}\}.$$

A little algebra shows that the derivative term in the integrand has a bound ($\mathrm{Re}\, z\geqslant 0$, $\xi'_M\geqslant 0$)

$$\left|\left(\frac{\partial}{\partial z}\right)^i\{z^{n+M}e^{-zqa_n\xi'_M}\}\right|\leqslant|z|^{n+M}(n+M)^i\sum_{r=0}^{i}\binom{i}{r}|z|^{r-i}(qa_n\xi'_M)^r.$$

From here on a repetition of the argument of Proposition 5 gives the bound (4.15). □

The asymptotic expansion for the analytic semigroup kernels $U_z(x,y)$ is obtained from Proposition 5, its corollary and Definition 1 for $F(x,y;z)$. If $v\in\mathcal{F}^{r}_{2M}$, then Eq. (3.5) is written

$$F(x,y;z)=\sum_{n=0}^{M-1}B_n(x,y;z)+\sum_{n=M}^{\infty}B_n(x,y;z). \tag{4.16}$$

We find

Theorem 2: Let $v(x)$ be a real-valued bounded and continuous potential represented by a complex bounded measure μ. Let $U_z(x,y)$ and $U_z^{(0)}(x-y)$ be the kernels of the integral operators e^{-zH} and e^{-zH_0}, respectively. If $v(x)\in\mathcal{F}^{r}_{2M}$, then for all $z\in\bar{D}\setminus\{0\}$,

$$U_z(x,y)=U_z^{(0)}(x-y)\left\{\sum_{n=0}^{M-1}\frac{(-z)^n}{n!}P_n(x,y)+E_M(x,y;z)\right\}, \tag{4.17}$$

where coefficient functions $P_n(x,y)$ are $P_0(x,y)=1$,

$$P_n(x,y)=\sum_{m=0}^{n-1}q^m D_{m,n}(x,y)\quad n=1,2,\ldots,M-1, \tag{4.18}$$

and have the x,y uniform bound

$$|P_n(x,y)|\leqslant\left(\|\mu\|+\frac{qn^2K^2}{4}\right)^n. \tag{4.19}$$

Here K is the bound constant of $v(x)$ in $\mathcal{F}^{r}_{2M}$. The remainder term is of order $O(|z|^M)$ and has the bound

$$|E_M(x,y;z)|\leqslant\frac{(|z|\,\|\mu\|)^M}{M!}\times\left\{\left(1+\frac{q(M-1)^2K^2}{4\|\mu\|}\right)^M+\exp(|z|\,\|\mu\|)\right\}. \tag{4.20}$$

Furthermore, if $v(x)\in\mathcal{F}^{r}_{2(M+i)}$, then one may take i derivatives of expansion (4.17) with respect to z. The ith derivative of the remainder term $E_M(x,y;z)$ is of order $O(|z|^{M-i})$.

Proof: Start with Eq. (4.16). Assume $v(x)\in\mathscr{F}^r_{2(M+i)}$ with the bound constant K. We estimate the sum over $n=(M,...,\infty)$ of $B^{(i)}_n(x,y;z)$. Set

$$S^{(i)}_M(x,y;z)=\sum_{n=M}^{\infty}B^{(i)}_n(x,y;z).$$

From Definition 1 of $B_n(x,y;z)$, we have

$$B^{(i)}_n(x,y;z)$$
$$=\frac{(-1)^n}{n!}\int_0^1 d^n\xi\int d^n\mu\Big(\frac{\partial}{\partial z}\Big)^i\{z^n e^{-zqa_n}\}e^{ib_n}.$$

Using

$$\Big|\Big(\frac{\partial}{\partial z}\Big)^i\{z^n e^{-zqa_n}\}\Big|\leqslant n^i|z|^n\sum_{r=0}^{i}\binom{i}{r}\Big[\frac{1}{|z|}\Big]^{i-r}(qa_n)^r$$

and (4.13) gives

$$|B^{(i)}_n(x,y;z)|\leqslant\frac{(|z|\|\mu\|)^n n^{3i}}{n!}\Big(\frac{1}{|z|}+\frac{qK^2}{4}\Big)^i.$$

In order to estimate the sum over n, note that for $\gamma\geqslant 0$ and $\alpha>0$, one has

$$\sum_{n\geqslant M}^{\infty}\frac{\alpha^n n^\gamma}{n!}\leqslant\frac{\{\alpha\exp(\gamma/e)\}^M}{M!}\exp\{\alpha\exp(\gamma/e)\}.$$

This follows, since $n^\gamma=\exp(\gamma\ln n)\leqslant\exp(\gamma n e^{-1})$. Setting $\gamma=3i$, $\alpha=|z|\,\|\mu\|$ leads us to

$$|S^{(i)}_M(x,y;z)|$$
$$\leqslant\Big(\frac{1}{|z|}+\frac{qK^2}{4}\Big)^i\frac{\{|z|\|\mu\|\exp(3i/e)\}^M}{M!}$$
$$\times\exp\{|z|\|\mu\|\exp(3i/e)\}. \tag{4.21}$$

This sum is of order $O(|z|^{M-i})$. It is valid for all $z\in\bar{D}\setminus\{0\}$. If $i=0$, then it is valid for all $z\in\bar{D}$ and forms the second part of the two expressions bounding $E_M(x,y;z)$.

The $M-1$ leading terms of expression (4.17) come from the sum

$$L_M(x,y;z)=\sum_{n=0}^{M-1}B_n(x,y;z).$$

Rewrite this sum by using (4.7) to find

$$L_M(x,y;z)=\sum_{n=0}^{M-1}\sum_{m=0}^{M-1-n}\frac{(-z)^{n+m}}{(n+m)!}$$
$$\times q^m D_{m,n+m}(x,y)+S'_M(x,y;z),$$

where

$$S'_M(x,y;z)=\sum_{n=0}^{M-1}R_{n,M-n}(x,y;z).$$

Changing the summation index $n+m\to n$ yields

$$L_M(x,y;z)=\sum_{n=0}^{M-1}\frac{(-z)^n}{n!}P_n(x,y)+S'_M(x,y;z),$$

where $P_n(x,y)$ is given by the expression (4.18). The $P_n(x,y)$ bound results from (4.9).

The last step is to estimate $S'_M(x,y;z)$. Assume $v(x)\in\mathscr{F}^r_{2M}$. Use the bound (4.10) for $R_{n,M-n}(x,y;z)$.

$$|S'_M(x,y;z)|\leqslant\sum_{n=0}^{M-1}\frac{(|z|\|\mu\|)^n}{n!(M-n)!}\Big(\frac{|z|qn^2K^2}{4}\Big)^{M-n}$$
$$\leqslant\frac{|z|^M}{M!}\Big(\|\mu\|+\frac{q(M-1)^2K^2}{4}\Big)^M.$$

This estimate for $S'_M(x,y;z)$ is the first factor in the estimate (4.20) for $E_M(x,y;z)$. Thus (4.17) is demonstrated if $v(x)\in\mathscr{F}^r_{2M}$.

To obtain the ith derivative of expansion (4.17), assume $v(x)\in\mathscr{F}^r_{2(M+i)}$. The bound (4.21) controls $|S^{(i)}_M(x,y;z)|$. For $S'^{(i)}_M(x,y;z)$ use the estimates (4.15) of Corollary 1 to demonstrate that $S'^{(i)}_M(x,y;z)$ is uniformly bounded and continuous in x,y. Since,

$$|E^{(i)}_M(x,y;z)|\leqslant|S^{(i)}_M(x,y;z)|+|S'^{(i)}_M(x,y;z)|,$$

we have that the ith derivative of the error term $E_M(x,y;z)$ is of order $O(|z|^{M-i})$. Lastly, note that $U^{(0)}_z(x,y)$ is holomorphic in z for all $z\in D$ and infinitely differentiable in t for $z=it$ and $t\neq 0$. Combining these facts implies that (4.17) may be differentiated term by term i-times if $v(x)\in\mathscr{F}^r_{2(M+i)}$. □

Theorem 2 states the basic results of this paper. The very detailed estimates in the theorem provide a method of calculating the time evolution kernels to high order in t ($z=it$) and have the advantage of possessing a known bound for the total error. The results of Theorem 2 show precisely how a $2M$-times differentiable potential leads us to an M-term asymptotic expansion of the exact time evolution kernel.

The semiclassical content of the asymptotic expansion is best seen from the diagonal form

$$U_z(x,x)=\frac{1}{(4\pi zq)^{d/2}}\Big\{\sum_{n=0}^{M-1}\frac{(-z)^n}{n!}P_n(x,x)+E_M(x,x;z)\Big\}. \tag{4.22}$$

This diagonal form is uniquely defined because all the functions $P_n(x,y)$ and $E_M(x,y;z)$ are jointly continuous in x,y. From the expression (4.18) for $P_n(x,y)$, it is seen that $P_n(x,y)$ is a polynomial of order $n-1$ in the quantum scale parameter q; hence, each term of $P_n(x,y)$ has a semiclassical interpretation. The classical component of $P_n(x,y)$, is just $D_{0,n}(x,y)$, whose diagonal value[2] is $D_{0,n}(x,x)=\{v(x)\}^n$. For a more detailed physical interpretation of this expansion and its applications, one should consult Refs. 2 and 3.

If $z\in D$ and is not on the time axis, then expansions like (4.22) and (4.17) have been derived for much more general partial differential equations and boundary conditions than we have studied here.[19,20] One should note, however, that our aim of this paper is to give the precise uniform asymptotic expansion of the time evolution kernel with a complete remainder term bound, not assuming the decay of the potentials at infinity. This has been accomplished by introducing a special class of smooth potentials.

ACKNOWLEDGMENTS

The authors are indebted to both Steve Fulling and Rod Wong for numerous discussions of the problems described in this paper. This work was supported in part by a grant from the Natural Sciences and Engineering Research Council, Canada. One of the authors (Y. F.) acknowledges that part of this work was done at the University of Minnesota under DOE Contract No. DE-AC02-79-ER10364.

[1]A. M. Perelomov, Ann. Inst. Henri Poincaré **24**, 161 (1976).
[2]S. F. J. Wilk, Y. Fujiwara, and T. A. Osborn, Phys. Rev. A **24**, 2187 (1981).
[3]Y. Fujiwara, T. A. Osborn, and S. F. J. Wilk, Phys. Rev. A **25**, 14 (1982).
[4]E. P. Wigner, Phys. Rev. **40**, 749 (1932); J. G. Kirkwood, Phys. Rev. **44**, 31 (1933).
[5]S. A. Fulling, SIAM J. Math. Anal. **13**, 891 (1982); "The Local Geometrical Asymptotics of Continuum Eigenfunction Expansions. II and III," Texas A&M preprint (1982).
[6]B. Simon, Bull. Am. Math. Soc. **1**, 447 (1982).
[7]C. Clark, SIAM Rev. **9**, 627 (1967).
[8]K. Ito, in Proc. Fourth Berkeley Symp. Math. Stat. and Prob. (Univ. California P., Berkeley, 1961), Vol. II, p. 227; K. Ito, in Proc. Fifth Berkeley Symp. on Math. Stat. and Prob. (Univ. California P., Berkeley, 1967), Vol. II, Part 1, p. 145.
[9]S. A. Albeverio and R. J. Høegh-Krohn, *Mathematical Theory of Feynman Path Integrals* (Springer-Verlag, Berlin, 1976).
[10]Y. Katznelson, *Harmonic Analysis* (Wiley, New York, 1968), p. 139.
[11]W. Rudin, *Fourier Analysis on Groups* (Interscience, New York, 1961), Chap. 1.
[12]E. Hille and R. S. Phillips, *Functional Analysis and Semi Groups* (American Mathematical Society, Providence, RI, 1957).
[13]W. O. Amrein, J. M. Jauch, and K. B. Sinha, *Scattering Theory in Quantum Mechanics* (Benjamin, Reading, MA, 1977), Chap. 3.
[14]S. Mizohata, *The Theory of Partial Differential Equations* (Cambridge U.P., Cambridge, 1973), p. 43.
[15]D. Fujiwara, J. Anal. Math. **35**, 41 (1979).
[16]H. Kitada, J. Fac. Sci. Univ. Tokyo IA, **27**, 193 (1980); H. Kitada and H. Kumano-go, Osaka J. Math. **18**, 291 (1981).
[17]S. Zelditch, "Reconstruction of Singularities for Solutions of Schrödinger Equations," Ph.D. Thesis, Univ. of California, Berkeley, 1981.
[18]R. Wong, SIAM Rev. **22**, 401 (1980).
[19]P. Greiner, Arch. Ration. Mech. Anal. **41**, 163–218 (1971).
[20]P. B. Gilkey, Compositio Math. **38**, 201 (1979).

Time evolution of the Wigner function

Francisco Soto-Eguibar[a] and Pierre Claverie
Laboratoire de Chimie Quantique, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, 75005-Paris, France

(Received 31 March 1981; accepted for publication 13 May 1982)

In this paper we give a partial answer to the problem: When does an initially non-negative Wigner function remain non-negative under the effect of the time evolution? We show that, for pure states, this is the case for linear systems only; to prove this we use the fact that the Wigner function is non-negative if and only if the wavefunction is Gaussian. We also prove that the Green's solution of the evolution equation of the Wigner function, which in the framework of probability theory corresponds to the conditional probability density, takes on negative values. We utilize a theorem, about moments, borrowed from Pawula. We conclude that the Wigner phase-space formulation of quantum mechanics cannot receive a genuine probabilistic interpretation.

PACS numbers: 03.65 – w

I. INTRODUCTION

The existence and properties of phase-space distribution functions in quantum mechanics are closely related to the question of its reformulation in terms of classical concepts and to the existence of an acceptable classical limit.

The choice of a phase-space distribution function has a high degree of arbitrariness, because it is equivalent to the choice of a correspondence rule.[1–4] Several phase-space distribution functions have been proposed (as many as correspondence rules),[2] but none of them gives the correct quantum mechanical expectation values for all observables when calculated through phase-space integration.[2,4]

The most widely known and employed phase-space "distribution" is the Wigner function, which is associated with the Weyl correspondence rule.[1–5] This function gives the correct quantum mechanical marginal distributions for p and q, and thus, predicts the correct expectation values for any observable of the form $F(\hat{P}) + G(\hat{Q})$. However, this is not the case for observables of the form $F(\hat{P}^n\hat{Q}^m)$; for example, if the classical Hamiltonian is used, it gives a nonzero value for the standard deviation of energy of the first excited state of the simple harmonic oscillator.

Another shortcoming of this function is that, in general, it can take negative values and it cannot be considered a true probability distribution. For systems in a pure state it has been shown that a necessary and sufficient condition for the Wigner function to be non-negative is that the corresponding Schrödinger state function is the exponential of a quadratic form: For the one-dimensional case see Hudson,[6] and for the generalization to arbitrary dimension see Soto and Claverie[7] (for the sake of completeness, let us mention that Piquet[8] also considered the one-dimensional case, but his proof was partly erroneous, see discussion in Ref. 7). Such Gaussian functions (modulo a linear canonical transformation) are also called coherent states.[9] At the present time, a similar characterization does not exist for mixed states.

Nevertheless, using the Weyl transformation and the Wigner function it is possible to construct in phase space an alternative form of quantum mechanics[1,3] whose physical meaning has not yet been clarified.

The aim of this paper is to gain some insight about the problems of this phase-space formulation. In fact, Moyal thought[1] that for the phase-space formulation to be consistent, it should be possible to prove that if a state admits initially a non-negative Wigner function, then the function evolved from it will be non-negative at any time; he gave a "proof" of this property for an isolated system with at least one cyclic coordinate, and in a pure state; this proof is wrong, as we shall show, and the question whether an initially non-negative Wigner function remains non-negative was therefore unanswered at this stage. In this paper we give a partial answer by proving that:

(i) For pure states only the linear systems have this property.

(ii) For nonlinear systems the Green's solution of the evolution equation in phase space (which is the Weyl transform of the von Neumann equation) takes negative values for short times.

Both proofs proceed by reduction *ad absurdum*. The first one uses the fact, already mentioned, that only the coherent states have a non-negative Wigner function. The second one is based upon a theorem borrowed from Pawula[10–12] that deals with the properties of the moments of a probability distribution.

The structure of the paper is the following: In Sec. II we introduce briefly the Wigner phase-space formulation of quantum mechanics. We refute, in Sec. III, Moyal's "proof" that an isolated system with at least one cyclic coordinate and in a pure state preserves the initially non-negative character of the Wigner function. The fact that, for pure states, only the linear systems have this property, is proved in Sec. IV. In Sec. V the Pawula theorem is enunciated, and using it, we prove (Sec. VI) that for arbitrary nonlinear systems the Green's solution of the phase-space evolution equation takes on negative values already for short times. The last section (Sec. VII) is devoted to the conclusions.

[a] Permanent address: Instituto de Fisica, UNAM (Universidad Nacional Autonoma de Mexico), Apdo. Postal 20-364, Mexico 20 D.F., Mexico.

 0022-2488/83/051104-06$02.50

II. THE WIGNER PHASE-SPACE FORMULATION

The Weyl correspondence rule associates with every operator $\hat{A}(t)$ a function $a(p,q,t)$ in phase-space which is given by

$$a(p,q,t)=\frac{1}{h^3}\operatorname{Tr}\left\{A(t)\int\int du\,dv\,e^{(i/\hbar)[(q-\hat{Q})\cdot u+(p-\hat{P})\cdot v]}\right\},\tag{1}$$

where integration is over all the corresponding space. If we have another operator $\hat{B}(t)$, it is easy to show[3] that

$$\operatorname{Tr}\hat{A}\hat{B}=\frac{1}{h^3}\int\int a(p,q,t)b(p,q,t)\,dp\,dq,\tag{2}$$

where $b(p,q,t)$ is the function which corresponds to $\hat{B}(t)$.

The Wigner function is defined as the Weyl transform of the density matrix $\hat{\rho}(t)$ divided by h^3 (Refs. 1–5); from (1) we thus see that it has the following expression:

$$F(p,q,t)=\frac{1}{h^6}\operatorname{Tr}\left\{\hat{\rho}(t)\int\int du\,dv\,e^{(i/\hbar)[(q-\hat{Q})\cdot u+(p-\hat{P})\cdot v]}\right\}.\tag{3}$$

In the case of a system in a pure state, with the wavefunction $\Psi(q)$ in configuration space, it has the well-known form given by Wigner[5]:

$$F(p,q,t)=\frac{1}{h^3}\int dv\,\Psi^*(q+\tfrac{1}{2}v,t)e^{(i/\hbar)p\cdot v}\Psi(q-\tfrac{1}{2}v,t).\tag{3a}$$

The expression for the expectation value of an operator $\hat{A}(t)$ is easily deduced from the quantum expression $\langle\hat{A}(t)\rangle=\operatorname{Tr}(\hat{A}\hat{\rho})$ and from Eq. (2). We get,

$$\langle\hat{A}(t)\rangle=\frac{1}{h^3}\int\int a(p,q,t)F(p,q,t)\,dp\,dq\equiv\langle a(p,q,t)\rangle.\tag{4}$$

We thus see that the Wigner function has the properties of a probability distribution in phase-space with the exception that in general it is not non-negative. We already said in the Introduction, that only for pure states does there exist an answer to the question: When is the Wigner function non-negative? This is the case if and only if the state wavefunction in configuration space is of the Gaussian form[6,7]

$$\Psi(q_1,\dots,q_n)=\exp\{-\tfrac{1}{2}[\mathbf{q}^\dagger\mathbb{A}\mathbf{q}+2\mathbf{b}\cdot\mathbf{q}+c]\},\tag{5}$$

where $\mathbb{A}$ is a complex matrix with $|\mathrm{Re}\,\mathbb{A}|>0$, $\mathbf{b}$ is an arbitrary complex vector, and c a constant that ensures the normalization.

The evolution equation of the Wigner function is found by performing the Weyl transformation of the von Neumann equation:

$$\frac{\partial\hat{\rho}}{\partial t}=-\frac{i}{\hbar}[H,\hat{\rho}],\tag{6}$$

where H is the time-independent Hamiltonian. We thus get

$$\frac{\partial F(p,q,t)}{\partial t}=\frac{2}{\hbar}\left[\sin\frac{\hbar}{2}\left(\frac{\partial}{\partial p},\frac{\partial}{\partial q}\right)\right]H_w(p,q)F(p,q,t)\tag{7a}$$

$$=-\frac{2}{\hbar}\left[\sin\frac{\hbar}{2}\left(\frac{\partial}{\partial p},\frac{\partial}{\partial q}\right)\right]H_w(p,q)F(p,q,t),\tag{7b}$$

where $H_w(p,q)$ is the Weyl-tranformed Hamiltonian and the notation $(\partial/\partial q,\partial/\partial p)H_wF$ means the Poisson bracket of the functions H_w and F:

$$\left(\frac{\partial}{\partial p},\frac{\partial}{\partial q}\right)H_w(p,q)F(p,q,t)=\{H_w(p,q),F(p,q,t)\}_{\mathrm{PB}}=\frac{\partial H_w(p,q)}{\partial q}\frac{\partial F(p,q,t)}{\partial p}-\frac{\partial H_w(p,q)}{\partial p}\frac{\partial F(p,q,t)}{\partial q}.\tag{8a}$$

We use now the following more explicit notation of Baker[13]:

$$\left(\frac{\partial}{\partial p},\frac{\partial}{\partial q}\right)H_wF=\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2}-\frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\times H_w(p_1,q_1)F(p_2,q_2)\Big|_{\substack{p_1=p_2=p,\\ q_1=q_2=q}}\tag{8b}$$

and then the complete operator $(2/\hbar)\sin(\hbar/2)(\partial/\partial p,\partial/\partial q)$ is defined through the formal series expansion of the sine:

$$\frac{2}{\hbar}\left[\sin\frac{\hbar}{2}\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2}-\frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\right]\times H_w(p_1,q_1)F(p_2,q_2)\Big|_{\substack{p_1=p_2=p,\\ q_1=q_2=q}}=\frac{2}{\hbar}\left\{\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\times\left[\frac{\hbar}{2}\left(\frac{\partial}{\partial q_1}\frac{\partial}{\partial p_2}-\frac{\partial}{\partial p_1}\frac{\partial}{\partial q_2}\right)\right]^{2n+1}\right\}\times H_w(p_1,q_1)F(p_2,q_2)\Big|_{\substack{p_1=p_2=p,\\ q_1=q_2=q}}\tag{8c}$$

The evolution equation (7) can be derived in another way. We shall recall only the main steps of this derivation; for details see Secs. 6 and 7 of Ref. 1. The fundamental relation connecting the probability distribution $F(p,q,t)$ and $F(p,q,0)$ is

$$F(p,q,t)=\int\int K(p,q|p_0,q_0;t)F(p_0,q_0,0)\,dp_0\,dq_0,\tag{9}$$

where $K(p,q|p_0,q_0;t)$ is the conditional distribution function. From (9), Moyal shows[1] that $F(p,q,t)$ satisfies the following evolution equation (which in fact is a particular case of the Pawula equations as we shall see later):

$$\frac{\partial F(p,q,t)}{\partial t}=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^{n+m}}{n!m!}\left(\frac{\partial}{\partial p}\right)^n\left(\frac{\partial}{\partial q}\right)^m[\alpha_{n,m}(p,q)F(p,q,t)],\tag{10}$$

where the coefficients $\alpha_{n,m}(p,q)$ are the "derivative moments" given by

$$\alpha_{n,m}(p,q)=\lim_{t\to 0}\int\int(\eta-p)^n(\xi-q)^m\frac{K(\eta,\xi|p,q;t)}{t}\,d\eta\,d\xi,\tag{11}$$

and where

$$K(\eta,\xi|p,q;t) = \delta(\eta - p)\delta(\xi - q) + \sum_{k=0}^{\infty} \frac{t^{k+1}}{(k+1)!} \int \cdots_{(k)} \int S(\eta,\xi|\eta_1,\xi_1) S(\eta_1,\xi_1|\eta_2,\xi_2) \cdots S(\eta_k,\xi_k|p,q)\, d\eta_1\, d\xi_1 \cdots d\eta_k\, d\xi_k, \tag{12}$$

with

$$S(\eta,\xi|p,q) = \frac{i}{h^3} \int\int [h(p + \tfrac{1}{2}u, q - \tfrac{1}{2}v) - h(p - \tfrac{1}{2}u; q + \tfrac{1}{2}v)] e^{-(i/h)[u(p-\eta) + v(q-\xi)]}\, du\, dv. \tag{13}$$

In principle Eqs. (11)–(13) could be used to find the coefficients $\alpha_{n,m}(p,q)$, but in practice this is really hard to do; here we already know them from the other deduction, comparing (10) with (7) we have

$$\alpha_{n,m}(p,q) = 0 \quad \text{if } n + m \text{ is even}, \tag{14a}$$

$$\alpha_{n,m}(p,q) = (-1)^{(n+3m-1)/2} \times \left(\frac{\hbar}{2}\right)^{n+m-1} \left(\frac{\partial}{\partial p}\right)^m \left(\frac{\partial}{\partial q}\right)^n H_w(p,q), \quad \text{if } n + m \text{ is odd}. \tag{14b}$$

We remark that in this deduction the non-negativity of the distribution function has not been assumed.

III. DISPROOF OF MOYAL'S ARGUMENT

In his important paper (Ref. 1, Sec. 15) Moyal made some claims concerning the conservation of the non-negative character of the Wigner function under the effect of time evolution. We shall now see that these claims are unfounded. Moyal's statement was the following: An isolated system with at least one cyclic coordinate θ and in a pure state preserves the initially non-negative character of the Wigner function. The argument (Ref. 1, Sec. 15, p. 117) goes as follows:

(a) A canonical transformation is made from the original coordinate system p_i, q_i to the system (g, θ, P_i, Q_i), where g is the conjugate of θ and P_i, and Q_i are the other (transformed) moments and coordinates. The transformed Hamiltonian $H(g, \theta, P_i, Q_i)$ is then such that

$$\frac{\partial H}{\partial \theta} = 0, \quad \frac{\partial H}{\partial g} = \text{const} = \omega. \tag{15}$$

(b) The evolution equation (7) is written in the (g, θ, P_i, Q_i) system.

(c) ω being a constant, the evolution equation [taken under form (7b)] can be written

$$\frac{\partial F(g,\theta,P_i,Q_i,t)}{\partial t} + \omega \frac{\partial F(g,\theta,P_i,Q_i,t)}{\partial \theta} + \frac{2}{\hbar}\left[\sin \frac{\hbar}{2}\left(\frac{\partial}{\partial P}, \frac{\partial}{\partial Q}\right)\right] H(g,\theta,P_i,Q_i) F(g,\theta,P_i,Q_i) = 0. \tag{16}$$

(d) In Eq. (16) the variables t and θ can be separated from the others through the substitution

$$F(g,\theta,P_i,Q_i,t) = F_1(\theta,t) F_r(g,P_i,Q_i) \tag{17}$$

which gives the following two equations:

$$\frac{1}{F_1(\theta,t)}\left[\frac{\partial F_1(\theta,t)}{\partial t} + \frac{\partial F_1(\theta,t)}{\partial \theta}\right] = 2i\mu, \tag{18a}$$

$$\frac{1}{F_r(g,P_i,Q_i)}\left[\sin \frac{\hbar}{2}\left(\frac{\partial}{\partial P}, \frac{\partial}{\partial Q}\right)\right] \times H(g,\theta,P_i,Q_i) F_r(g,P_i,Q_i) = \frac{\hbar}{i}\mu, \tag{18b}$$

where μ is a separation constant

(e) Solving (18a) we find that the time dependence is of the form

$$\exp[i\mu(t + \theta/\omega)], \tag{19a}$$

Comparing this time dependence with the expansion of the Wigner function in energy eigenfunctions (see formula 8.1 of Ref. 1) we find

$$F(g,\theta,P_i,Q_i) = \sum_{j,k} a_j^* a_k (g,P_i,Q_i) \times \exp\left(i \frac{E_j - E_k}{h}\left(t + \frac{\theta}{\omega}\right)\right) \tag{19b}$$

where a_j are the coefficients in the expansion of the wavefunction in energy eigenfunctions ψ_j and the $F_{jk}(p,q)$ are defined as follows (see formula 4.11 of Ref. 1):

$$F_{jk}(\mathbf{p},\mathbf{q}) = \frac{1}{h^n} \int_{R^n} \psi_j^*\left(\mathbf{q} + \frac{1}{2}\mathbf{v}\right) \times \exp\left(\frac{i}{h}\mathbf{p}\cdot\mathbf{v}\right) \psi_k\left(\mathbf{q} - \frac{1}{2}\mathbf{v}\right) d\mathbf{v}.$$

From (19b) Moyal concludes that if $F > 0$ whatever θ at $t = 0$, it must be non-negative for all $t > 0$.

This "proof" is wrong due to the following two mistakes:

(1) The first and most serious mistake occurs in step (b), where Moyal writes the evolution equation (7) in the canonical transformed system (q,θ,P_i,Q_i), which amounts to assuming that Weyl's correspondence rule is covariant with respect to canonical transformation over the classical phase-space. But this covariance just does *not* hold (whatever quantization rule is assumed), as shown first by Van Hove (Ref. 14, Chap. VI, Sec. 23), so that the evolution equation (7) is valid *only* in terms of the usual Cartesian coordinates P_i, Q_i (see Sec. 6, p. 105 of Ref. 1). The reader may easily build for himself an illustration of this situation by treating a two-dimensional harmonic oscillator, first in Cartesian and then in polar coordinates.

(2) Even if step (b) were correct, the proof would still be wrong because step (c) is also erroneous. In this step, ω is considered a constant, independent of the coordinates and

moments, and this is not generally true. Of course, it is a constant of motion, but, except for the case of the harmonic oscillator, it is not constant with respect to the variables of the system. Therefore, as a general rule, nonzero derivatives of $\partial H/\partial g = \omega$ would be generated by the operator $(2/h)$ sin $[(h/2)(\partial/\partial p, \partial/\partial q)]$ of Eq. (7b), and these derivatives are lacking in Eq. (16). Only for a linear system (harmonic oscillator) would ω be a genuine constant. In actual fact, for such a system, the evolution equation for the Wigner distribution function just reduces to the usual Liouville equation, as may be deduced from Eqs. (14), and this equation actually preserves the non-negative character of the distribution function. From the analysis presented in this section, we may therefore conclude that Moyal's statement holds true for the trivial case of the harmonic oscillator, but, owing to the incorrect character of his proof, the question remains open, at this stage, as concerns the general case (nonlinear systems). We shall see in the next section that, for nonlinear systems, the answer is actually negative.

IV. THE PURE STATE CASE

In this section we consider only systems in a pure state. We prove that only the linear systems have the property that if at the initial time (taken here as 0) they are in a state with a non-negative Wigner function, then this function will be non-negative for all positive times; or equivalently, that if we have a nonlinear system with initial condition:

$$\Psi(\mathbf{q},0) = \exp\{-\tfrac{1}{2}[\mathbf{q}^\dagger \mathbb{A}\mathbf{q} + 2\mathbf{b}\cdot\mathbf{q} + c]\}, \tag{20}$$

where $\mathbb{A}$ is a complex matrix with $|\mathrm{Re}\mathbb{A}| > 0$, $\mathbf{b}$ an arbitrary complex vector, and c a normalization constant. Then for any time $t > 0$ (with the possible exception of a discrete set of times) the Wigner function will take negative values. This result may be considered as a generalization of a similar property proved by Guichardet for coherent states, through very different methods (see Ref. 15, Chap. 2, Sec. 2.2, Lemmas 2.1). The connection between Guichardet's result and ours deserves some developments, which are presented in the Appendix.

We know[6,7] that for the Wigner function to be non-negative at any $t > 0$, the wavefunction must be of the form

$$\Psi(\mathbf{q},t) = \exp\{-\tfrac{1}{2}[\mathbf{q}^\dagger \mathbb{A}(t)\mathbf{q} + 2\mathbf{b}(t)\cdot\mathbf{q} + c(t)]\}, \tag{21}$$

with $\mathbb{A}(t)$, $\mathbf{b}(t)$, and $c(t)$ complex functions of time and $|\mathrm{Re}\mathbb{A}(t)| > 0$ for all $t > 0$.

Thus, we must find which are the systems that have (21) as solution with initial condition (20). For that purpose, we substitute (21) in the Schrödinger equation with an arbitrary potential $V(\mathbf{q})$ and we find the following equation (the Einstein summation convention is used):

$$\left[\frac{\hbar^2}{2m}\mathbb{A}^2(t) - i\frac{\hbar}{2}\dot{\mathbb{A}}(t)\right]_{ij} q_i\, q_j + \left[\frac{\hbar^2}{m}\mathbb{A}(t)\mathbf{b}(t) - i\hbar\,\dot{\mathbf{b}}(t)\right]_i q_i + \left[\frac{\hbar^2}{2m}\mathbf{b}^2(t) - \frac{\hbar^2}{2m}\mathrm{Tr}\,\mathbb{A}(t) - i\frac{\hbar}{2}\dot{c}(t)\right] = V(\mathbf{q}), \tag{22a}$$

with initial conditions

$$\mathbb{A}(0) = \mathbb{A},\ \mathbf{b}(0) = \mathbf{b},\ \text{and}\ c(0) = c. \tag{22b}$$

Since the powers of the q_i's are linearly independent, this equation can be satisfied at any time if and only if $V(\mathbf{q}) = \alpha_{ij}q_i q_j + \beta_i q_i + \gamma$, i.e., if the system is a linear one, which is the announced result.

V. THE PAWULA THEOREM

Pawula[10–12] has derived generalized Fokker–Planck equations for the conditional probability density functions of arbitrary random processes and found the conditions under which these equations are of finite order. In this section we expose his results.

Let $\mathbf{y}(t)$ denote an M-dimensional vector whose components are the M random variables $y_i(t)$ ($i = 1,\dots, M$) belonging to different random processes, and let $p(\mathbf{y}, t\,|\mathbf{Y},T)$ be its conditional probability density function conditioned by an arbitrary set $(\mathbf{Y},T)$ of k values $\mathbf{Y}$ with k times of occurrence denoted by T. Thus, $\mathbf{Y}$ stands for $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \dots, \mathbf{y}^{(k)}\}$ and T stands for $\{t^{(1)}, t^{(2)}, \dots, t^{(k)}\}$.

For $t \notin T$, the transition probability density function satisfies the following generalized Fokker–Planck equations:

$$\frac{\partial}{\partial t}P(\mathbf{y},t\,|\mathbf{Y},T) = \sum_{\substack{n_1,\dots,n_M=0\\ \left(\sum_{i=1}^M n_i \neq 0\right)}}^{\infty}\left[\prod_{i=1}^{M}\frac{(-1)^{n_i}}{n_i!}\left(\frac{\partial}{\partial y_i}\right)^{n_i}\right][A^{\pm}_{n_1,\dots,n_M}P(\mathbf{y},t\,|\mathbf{Y},T)], \tag{23}$$

where the derivate conditional moments are

$$A^{\pm}_{n_1,\dots,n_M} = \lim_{\Delta t\to 0^\pm}\frac{1}{\Delta t}E\left[\prod_{i=1}^{M}\{y_i(t+\Delta t) - y_i(t)\}^{n_i}|\mathbf{y},t,\mathbf{Y},T)\right], \tag{24}$$

where the superscript $+$ (resp. $-$) in $A^{\pm}$ corresponds to the choice of the limit $\Delta t\to 0^+$ (resp. $\Delta t \to 0^-$).

Here some words of explanation are needed in order to prevent confusion. The A^+ coefficients correspond to the familiar Fokker–Planck equation of the theory of Markov processes ("forward" equation), but the A^- coefficients do *not* correspond to the so-called *backward* equation (see, e.g., Arnold, Ref. 16, Chap. 2). Indeed, the backward equation for Markov processes has the same coefficients A^+, but in front of the derivation operators instead of lying inside. The A^- coefficients correspond to the Fokker–Planck equation for the *time-reversed* process. (See Nelson, Ref. 17, Chap. 13. Our A^+ coefficients correspond to Nelson's b and D, while our A^- correspond to Nelson's b^* and D^*). For the sake of definiteness, we shall consider from now on the usual Fokker–Planck equation, i.e., the A^+ coefficients (corresponding to $\Delta t \to 0^+$), and we therefore shall omit the $+$ superscript.

We now consider conditions under which Eq. (23) is of finite order in the variables y_i; these conditions are given by the so-called Pawula theorem.

First of all, we must define the one-dimensional derivate moments $U^{(i)}_{n_i}$ (i = 1, 2,..., M):

$$U^{(i)}_{n_i} = \lim_{\Delta t\to 0}\frac{1}{\Delta t}E \times [\{y_i(t+\Delta t) - y_i(t)\}^{n_i}|y_i,t;Y_{[i]},T_{[i]}], \tag{25}$$

where $t \notin T_{[i]} \subset T$ and $T_{[i]} \subset Y$. It is easy to see that

$$U^{(i)}_{n_i} = E\left[A_{0,0\ldots,n_i,\ldots,0} | y_i, t; Y_{[i]}, T_{[i]}\right], \tag{26}$$

where $E[A | y_i, t; Y_{[i]} T_{[i]}]$ means that averaging of A has been performed with respect to $y_j (j \neq i)$ and all variables $y_i^{(\alpha)}$ of the set Y except those belonging to the subset $Y_{[i]}$ (hence, the result can depend only on the subset of times $T_{[i]}$ corresponding to the subset $Y_{[i]}$). We can now state the:

Pawula theorem: If each of the one-dimensional derivate moments $U^{(i)}_{n_i}$ is finite and vanishes for some even n_i, then

$$A_{n_1,\ldots,n_M} = 0 \text{ (with probability 1)} \tag{27}$$

for every set $\{n_i\}$ such that $\Sigma_{i=1}^{M} n_i \geqslant 3$.

We finish this section with two remarks about this theorem:

(i) it is independent of any derivation of Eq. (23); (ii) it is valid for moments in general and not only for derivate moments as has been enunciated here (see Lemma 1, Sec. III of Ref. 11). It is therefore a theorem about moments in general.

VI. THE GREEN'S SOLUTION OF THE EVOLUTION EQUATION

Let us analyze the relations of the Moyal evolution equation (10) and the Pawula equation (29). First we remark that the function $K(p, q | p_0', q_0'; t)$ [Eqs. (9)–(13)] is the distribution function $F(p, q, t)$ which corresponds to the initial condition $F(p_0, q_0, 0) = \delta(p_0 - p_0')\delta(q_0 - q_0')$, i.e., it is the Green's solution of Eq. (10).

Comparing (10) and (11) with (29) and (30) we see that if $K(p, q | p_0, q_0; t)$ is a true probability density, then Moyal's equation (10) is a particular case of Pawula's equation (29) and we can utilize the theorem of the preceding section; thus we suppose that $K(p, q | p_0, q_0; t)$ is a true probability density, i.e., that $K(p, q | p_0, q_0; t)$ is non-negative for all t.

From Eqs. (14) and (26) we have

$$U^{(p)}_{2n} = U^{(q)}_{2n} = 0, \quad n = 1, 2, 3, \ldots, \tag{28}$$

i.e., all the even one-dimensional moments vanish. Also from (14) and (26) we see that the odd one-dimensional moments do not necessarily vanish and we can suppose that for real physical systems they are all finite. Thus, all the conditions required for the application of the Pawula theorem are fulfilled, and we must have, from Eq. (27), $\alpha_{n,m}(p, q) = 0$ for all n, m such that $n + m \geqslant 3$. Going back again to Eqs. (14) we see that only the linear systems satisfy this condition; thus for nonlinear systems we have a contradiction, whose solution is that the hypothesis made about $K(p, q | p_0, q_0; t)$ is false; in other words, $K(p, q | p_0, q_0; t)$ is *not* a true conditional probability density. Then we have two alternatives:

(i) $K(p, q | p_0, q_0; t)$ is not normalizable;

(ii) $K(p, q | p_0, q_0; t)$ takes on negative values.

The first one must be rejected because it implies that the moments $\alpha_{n,m}(p, q)$ would be infinite and this is not the case; thus we have only the second alternative, i.e., $K(p, q | p_0, q_0; t)$ is not non-negative.

We remark that, strictly speaking, this conclusion has been established for short times only, because it has been deduced from a property of the *derivate* moments, which are defined in the limit $t \to 0$.

Another remark is that this result is valid for an arbitrary number of dimensions and for both pure and mixed states.

Finally, the fact that $K(p, q | p_0, q_0; t)$ takes negative values for short times does not automatically imply that the Wigner function has the same property. In actual fact, from (9), we see that an integration is involved, and we do not know yet when this integration gives a negative value.

VII. CONCLUSION

We conclude that the Wigner phase-space formulation of quantum mechanics is not probabilistically consistent. Although for the general case we have not shown that the Wigner function will lose its initially non-negative character, the fact that its evolution kernel [Eq. (9)] can take negative values is enough for drawing the conclusion above.

We may distinguish a "static" and a "dynamical" aspect in the "non-positive" character of the Wigner function: the static aspect refers to the fact that the Wigner function corresponding to some given quantum state may be non-positive, while the dynamical aspect refers to the fact that, even if we take some non-negative Wigner function, we may lose this non-negative character under the effect of the time evolution governed by Moyal's Eq. (7) (itself deduced from the von Neumann evolution equation through the Weyl correspondence rule). The dynamical aspect may be considered as still more important than the static aspect because, even if we find some non-negative Wigner distribution, this non-negative character is not kept under time evolution. The importance of this issue was clearly appreciated by Moyal (Ref. 1, Sec. 15), but the fact that the quantum law of evolution implied a negative answer (i.e., nonconservation of the positive character) was not at all evident, as shown by the fact that Moyal attempted (without success, as we showed in the present work) to prove the opposite property.

Finally, we want to emphasize that, due to the general character of the Pawula theorem, the proof given in Sec. VI may, at least in principle, be extended to any of the phase-space formulations given by Cohen.[2]

ACKNOWLEDGMENTS

We express our thanks to L. Pesquera for many helpful discussions about the Wigner function and the Fokker–Planck type evolution equations, and to the referee for very valuable comments and additional references (3b and 13–15).

APPENDIX: CONNECTION BETWEEN THE CONSERVATION OF THE POSITIVE CHARACTER OF THE WIGNER FUNCTION UNDER TIME EVOLUTION AND THE PROPERTIES OF COHERENT STATES

This appendix discusses the relationship between the results of our Sec. IV and some properties of coherent states as formulated by Guichardet.[15] The mathematical developments provided by this author are based upon the so-called "symmetric Hilbert spaces" SH (more familiar to the quantum physicist under the name of "Fock space") associated with a given Hilbert space H, and these developments may look at first sight rather unrelated to the present work. We therefore feel it appropriate to describe in a sufficiently detailed way the connection between his work and ours. First of all, it must be recalled that the usual Hilbert space L^2 of the functions of one variable may be considered as a "symmetric" (Fock) Hilbert space built from the one-dimensional space R^1 (the set of the real numbers) (see Ref. 15, Example 2.1), and similarly the space of the functions of n variables (possibly with some prescribed symmetry) will be considered as a "symmetric" (Fock) space built from the n-dimensional space R^n. This property enables us to apply Guichardet's general results to the usual n-particle Hilbert space of quantum-mechanical wavefunctions. The so-called coherent states, denoted EXP a are introduced (Ref. 15, Definition 2.2); in the one-dimensional case, after multiplication by the "basic measure" $\exp(-y^2/2)$, they just become Gaussians with arbitrary centers (but standard deviation fixed to unity),

$$\exp(-y^2/2)\mathrm{EXP}(a) = \exp(-y^2/2)\exp(ay - a^2/2) = \exp[-(y-a)^2/2].$$

The problem is now to find the group $\mathscr{G}_H$ of the unitary operators which leave globally invariant the set of all coherent states (possibly multiplied by some arbitrary constant). The answer is precisely provided by Lemma 2.1 of Ref. 15: $\mathscr{G}_H$ is made from the operators

$$U_{A,b,c} = cU_{I,b,1}\,U_{A,0,1}.$$

(1) c is a complex constant with modulus 1.

(2) $U_{A,0,1}$ denotes the natural extension to the symmetric Hilbert space SH of the unitary operator A acting on the basic Hilbert space H, namely

$$U_{A,0,1}(\mathrm{EXP}\ a) = \mathrm{EXP}(Aa)$$

(since, in our case, $H = R^n$, A is just a familiar finite-dimensional unitary transformation). Note that A does not act upon the variable y of the Gaussian function EXP a, but on the parameter a which defines the center of the Gaussian.

(3) Finally, $U_{I,b,1}$ (where b denotes an element of the basic Hilbert space H) is defined through

$$U_{I,b,1}(\mathrm{EXP}\ a) = \exp[-\tfrac{1}{2}\|b\|^2 - (a|b)]\mathrm{EXP}(a+b),$$

i.e., $U_{I,b,1}$ realizes an arbitrary translation (with translation vector b) of the center a of the Gaussian function [the scalar factor in front of EXP $(a+b)$ merely preserves the normalization].

To sum up, apart from the multiplication by some constant c (with modulus 1), the unitary transformations mapping coherent states onto coherent states just correspond to an arbitrary displacement of the center of the Gaussian associated with the coherent state ($U_{A,0,1}$ corresponds to rotations, $U_{I,b,1}$ to translations). Now, all such mappings may precisely be generated by the evolution operator $\exp(iHt)$ where H is the Hamiltonian of the harmonic oscillator whose ground state is the basic Gaussian measure $\exp(-y^2/2)$, and no other Hamiltonian may exhibit the same property. Indeed, the set of coherent states spans the whole symmetric Hilbert space ($SH = L^2$). Thus, an evolution operator $\exp(iHt)$, and hence, its infinitesimal generator iHt, is entirely defined by its action upon the coherent states, and therefore, if some mapping may be expressed as $\exp(iH_{\mathrm{harm}}t)$ where H_{harm} is a harmonic oscillator Hamiltonian, no other Hamiltonian may generate just the same mapping. We are thus led to the following conclusion as a corollary of Guichardet's Lemma 2.1: Any evolution operator which changes coherent states into coherent states (corresponding to some given basic harmonic Hamiltonian H_{harm}) just corresponds to the evolution generated by this basic harmonic Hamiltonian, i.e., it is of the form $\exp(iH_{\mathrm{harm}}t)$. In our Sec. IV, we derive just the same conclusion, but from less restrictive assumptions, since we only assume that Gaussian functions are changed into Gaussian functions *without requiring that initial and final Gaussians have the same dispersion*, i.e., without assuming that they are coherent states of *one and the same* basic harmonic oscillator.

[1]J. E. Moyal, Proc. Cambridge Philos. Soc. **45**, 99–124 (1949).
[2]L. Cohen, J. Math. Phys. **7**, 781–786 (1966).
[3](a) H. Groenewold, Physica **12**, 405–460 (1946); (b) J. C. T. Pool, J. Math. Phys. **7**, 66–76 (1966); (c) S. de Groot, *La transformation de Weyl et la fonction de Wigner: une forme alternative de la mécanique quantique* (les Presses de l'Université de Montréal, Montréal, 1974).
[4]T. A. Brody, "Problems and Promises of the Ensemble Interpretation of Quantum Mechanics," Preprint IFUNAM 80-08, Universidad Nacional Autonoma de Mexico, 1980.
[5]E. Wigner, Phys. Rev. **40**, 749–759 (1932).
[6]R. L. Hudson, Rep. Math. Phys. **6**, 249 (1974).
[7]F. Soto and P. Claverie, "When is the Wigner function of multidimensional systems nonnegative?," J. Math. Phys. **24**, 97 (1983).
[8]C. Piquet, C. R. Acad. Sci. Paris Ser. A **279**, 107 (1974).
[9](a) R. J. Glauber, Phys. Rev. **131**, 2766–2788 (1963); (b) J. R. Klauder and E. C. G. Sundarshan, *Fundamentals of Quantum Optics*, (Benjamin, New York, 1968), Chap. 7.
[10]R. F. Pawula, "Generalizations and extensions of the Fokker–Planck–Kolmogorov equations," doctoral thesis, California Institute of Technology, 1965.
[11]R. F. Pawula, IEEE Trans. Inform. Theory **13**, 33–41 (1967).
[12]R. F. Pawula, Phys. Rev. **162**, 186–188 (1967).
[13]G. A. Baker, Jr., Phys. Rev. **109**, 2198–2206 (1958).
[14]L. Van Hove, Bull. Cl. Sci. Acad. R. Belg. Mém. in 8° (2nd series) **26** (6), 1–102 (1951).
[15]A. Guichardet, *Symmetric Hilbert Spaces and Related Topics*, Lectures Notes in Mathematics, Vol. 261 (Springer, New York, 1972).
[16]L. Arnold, *Stochastic Differential Equations* (Wiley-Interscience, New York, 1974).
[17]E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U. P., Princeton, NJ, 1967).

Transfer matrices for one-dimensional potentials

Asher Peres
Department of Physics, Technion—Israel Institute of Technology, 32 000 Haifa, Israel

(Received 11 December 1981; accepted for publication 28 February 1982)

The one-dimensional Schrödinger equation can be written as a first-order multicomponent equation by considering ψ and $d\psi/dx$, or combinations thereof, as independent variables. A potential barrier is then represented by a matrix belonging to one of the homomorphic groups SU(1,1), SO(2,1), Sp(2,R), or SL(2,R). The relationship between these groups is clarified. In various applications, one of them may turn out more convenient than others. In particular, SO(2,1), which is obtained by using as a basis some bilinear combinations of ψ and $d\psi/dx$, leads to remarkable results: The Schrödinger wavefunction is represented by a trajectory on a unit hyperboloid; a periodic potential corresponds to a pseudorotation around a fixed axis; a random potential gives a random walk on the hyperboloid. This method can also be used to calculate bound states (in potential wells) and may have many other interesting applications.

PACS numbers: 03.65. – w, 02.20. + b

I. INTRODUCTION

One-dimensional physics[1] is a convenient theoretical laboratory to test analytical and numerical methods before applying them to the real world.

In this paper, group theory is used to discuss some properties of the one-dimensional time-independent Schrödinger equation

$$(-\hbar^2/2m)\psi'' + V(x)\psi = E\psi. \tag{1}$$

A potential barrier (or potential well) is represented by a transfer matrix belonging to one of the four (homomorphic) noncompact groups[2] SU(1,1), SO(2,1) Sp(2,R), and SL(2,R). The relationship between these groups is clarified, and it is shown that in various applications one of them may turn out more convenient than the others.

The outline of this article is as follows. In Sec. II, we write ψ as the sum of forward and backward amplitudes. These amplitudes can then be used as a basis to define transfer matrices, belonging to the two-dimensional representation of SU(1,1). Section III discusses the infinitesimal generators of the SU(1,1) group and the resulting finite transformations. Higher-dimensional representations are introduced, leading to the homomorphic group SO(2,1). The latter has a special status in this work, as shown in Sec. IV: The Schrödinger wavefunction can be represented by a trajectory on a unit hyperboloid, involving only two real first-order equations (the third variable, a phase, has been eliminated). Section V is devoted to transmission through disordered chains, a topic of high current interest. In Sec. VI, we examine the effect of arbitrarily shifting the zero of the energy scale. The total energy must however remain positive (as in all the preceding sections). Negative energies, in particular bound states, are discussed in Sec. VII. They involve representations of Sp(2,R). Finally, Sec. VIII presents a formalism based on SL(2,R) which is valid for both positive and negative energies, at the cost of using expressions which are not dimensionally homogeneous.

Throughout this paper, scalar products are defined, which involve various *indefinite* metrics. The resulting transfer matrices are not normal (they do not commute with their adjoints) but "pseudo-normal" with respect to the appropriate indefinite metric. Some properties of pseudo-normal matrices are briefly reviewed in Appendix A, and some properties of "average" transfer matrices are discussed in Appendix B.

II. FORWARD AND BACKWARD AMPLITUDES

For a free particle, $V = 0$ and $E > 0$, the solution of Eq. (1) is

$$\psi = Fe^{ikx} + Ge^{-ikx}, \tag{2}$$

where F and G are constants and

$$k = (2mE)^{1/2}/\hbar. \tag{3}$$

The first term in (2) is the forward amplitude (positive x direction); the second one is the backward amplitude. This decomposition of ψ can be generalized for arbitrary $V(x)$. Define

$$f = (\psi + \psi'/ik)/2 \tag{4a}$$

and

$$g = (\psi - \psi'/ik)/2, \tag{4b}$$

with k still given by Eq. (3). (This decomposition is similar to the one used for the Klein–Gordon equation by Feshbach and Villars,[3] and it will likewise lead to the introduction of an indefinite metric.) It is easily seen that, for $V = 0$, we have $f = Fe^{ikx}$ and $g = Ge^{-ikx}$. For $V \neq 0$, the physical meaning of f and g is illustrated by Fig. 1. However, it must be pointed out that the decomposition $\psi = f + g$ is *not* invariant under a shift of the energy scale

$$V \to V + \Delta, \quad E \to E + \Delta, \tag{5}$$

although the Schrödinger equation (1) is, of course, invariant under this rescaling. This problem will be discussed in Sec. VI and the case $E < 0$ in Sec. VII.

Differentiating (4) with respect to x and using (1), we obtain

$$f' = ik\,[f - (f+g)V/2E\,], \tag{6a}$$

$$g' = -ik\,[g - (f+g)V/2E\,]. \tag{6b}$$

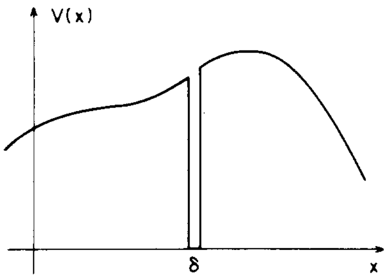

FIG. 1. The physical meaning of f and g can be understood by making a narrow "cut" in the potential $V(x)$, namely assuming $V(x) = 0$ in a small segment of length δ. (This is somewhat analogous to the old-fashioned way of defining **E** and **D** in a dielectric by cutting small cavities of various shapes.) This procedure leaves ψ'' finite, and therefore does not affect ψ nor ψ' in the limit $\delta \to 0$. Inside the "cut," Eq. (2) holds *exactly*, and Fe^{ikx} and Ge^{-ikx} are equal to the local values of f and g.

These equations can be simplified by introducing a two-component object

$$\Psi = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{7}$$

a dimensionless potential

$$u(x) = V(x)/E, \tag{8}$$

and a dimensionless *length* parameter

$$t = -kx \tag{9}$$

(the minus sign for later convenience). Differentiation with respect to t will be denoted by a dot, as if t were the time (this cannot cause any confusion, since our problem is time independent). With these notations, we obtain

$$i\dot{\Psi} = \begin{pmatrix} 1 - u/2 & -u/2 \\ u/2 & -1 + u/2 \end{pmatrix} \Psi \tag{10}$$

or

$$i\dot{\Psi} = [(1 - u/2)\sigma_z - (u/2) i\sigma_y] \Psi, \tag{11}$$

where σ_z and σ_y are the standard Pauli matrices. (Of course, no spin is involved in this problem.)

If we think of t as the time, the operator on the rhs of (11) is the "Hamiltonian," and at first sight it is surprising that the latter is not Hermitian. However, we must remember that we are not discussing a time evolution (where we would expect $\int |f + g|^2 dx$ to be constant) but the spatial dependence of forward and backward amplitudes. The conserved quantity is the total *current* density[4]

$$j = (\hbar/2im)(\bar{\psi}\psi' - \bar{\psi}'\psi), \tag{12}$$

which can be written, by virtue of (4) and (6), as

$$j = (\hbar k/m)(|f|^2 - |g|^2) = (\hbar k/m)\Psi^\dagger \sigma_z \Psi. \tag{13}$$

It is convenient to normalize j as $\hbar k/m$ so that

$$\Psi^\dagger \sigma_z \Psi = 1. \tag{14}$$

We see that σ_z plays the role of an *indefinite metric*[3] for the normalization of Ψ.

With this indefinite metric, the "Hamiltonian" (11) is pseudo-Hermitian, because

$$(i\sigma_y)^\dagger = \sigma_z (i\sigma_y) \sigma_z. \tag{15}$$

(In this paper, the prefix "pseudo" will be used to mean: "with respect to the indefinite metric being used.") For finite t, Eq. (11) generates a pseudo-unitary transformation

$$\Psi(t) = A(t)\Psi(0), \tag{16}$$

where the transfer matrix $A(t)$ can be written as[5–8]

$$A(t) = \begin{pmatrix} M & \bar{N} \\ N & \bar{M} \end{pmatrix}, \tag{17}$$

with

$$|M|^2 - |N|^2 = 1. \tag{18}$$

Therefore, $A(t)$ belongs to the SU(1,1) group.[2]

The physical significance of the coefficients M and N is illustrated by Fig. 2. We have a barrier penetration problem, where the transmitted amplitude $\psi = e^{ik(x-a)}$ has been normalized in compliance with Eq. (14). We write the incoming plus reflected amplitude as $\psi = Me^{ik(x-b)} + Ne^{-ik(x-b)}$ and current conservation ensures Eq. (18). The explicit values of M and N for a given barrier can be obtained by integrating the Schrödinger equation (1) or (10), starting from the *known* value of Ψ at $x = a$ and proceeding toward $x = b$. [Going in the negative x direction is like going in the *positive* t direction. This is the reason it was convenient to put a minus sign in Eq. (9).] We thereby determine the first column of Eq. (17). The second column is then obtained by replacing ψ by $\bar{\psi}$ (this corresponds to a time reversal, i.e., to a pure incoming wave at $x = a$).

If there are several potential barriers with transfer matrices C, B, A, say (in that order, from left to right) the overall transfer matrix is CBA. Note that the transfer matrix of an "empty" region ($V = 0$) of length L is

$$A_0 = \begin{pmatrix} e^{-ikL} & 0 \\ 0 & e^{ikL} \end{pmatrix}. \tag{19}$$

An important case is when the *same* barrier is repeated n times (for infinite n, we would have a periodic potential). The behavior of A^n for large n is controlled by the eigenvalues of A. The characteristic equation for the matrix (17) is

$$\lambda^2 - \lambda(M + \bar{M}) + 1 = 0. \tag{20}$$

Three cases must be distinguished, depending on the value of $M + \bar{M} = \mathrm{Tr}A$ (which is in general a function of E).

If $|\mathrm{Re}M| > 1$, we may write $\mathrm{Re}M = \pm\cosh(\theta/2)$ and the eigenvalues of A are $\lambda_1 = \pm e^{\theta/2}$ and $\lambda_2 = 1/\lambda_1$ (in the

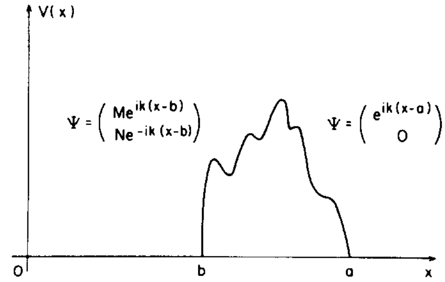

FIG. 2. A plane wave impinges at b on a potential barrier with compact support $b<x<a$. The transmitted part is normalized to unity at $x = a$. The incoming and reflected amplitudes at $x = b$ are M and N, respectively. The transmission probability is $|M|^{-2}$ and the reflection probability $|N/M|^2$.

next section, it will become clear why we write $\theta/2$ rather than simply θ). The largest eigenvalue of A^n then is $\pm e^{n\theta/2}$, implying that the matrix elements M and N grow exponentially with n. The transmission probability of the barrier, $|M|^{-2}$, decreases as $e^{-n\theta}$.

On the other hand, if $|\mathrm{Re}M| < 1$, we may write $\mathrm{Re}M = \cos(\theta/2)$ and then $\lambda_1 = 1/\lambda_2 = e^{i\theta/2}$. The eigenvalues of A^n thus are $e^{\pm in\theta/2}$. If θ/π is rational, there will be some finite n for which $A^n = I$. If not, there will still be finite n for which A^n will be arbitrarily close to the unit matrix (see Appendix A). Therefore, the transmission probability oscillates and (quasi-) periodically returns to 1 as n increases. The energy E is said to be in a *conduction band* (the preceding case corresponds to a forbidden band, but it is really forbidden only for infinite n).

Finally, in the exceptional case $|\mathrm{Re}M| = 1$, the matrix elements of A^n grow *linearly* rather than exponentially with n, see Eq. (A7).

From Eqs. (17) and (18) we obtain

$$A^{-1} = \begin{pmatrix} \bar{M} & -\bar{N} \\ -N & M \end{pmatrix} = \sigma_z A^\dagger \sigma_z. \tag{21}$$

Therefore, A is pseudo-unitary (with respect to the metric σ_z) and in particular is pseudonormal. Some properties of pseudonormal matrices are discussed in Appendix A.

III. INFINITESIMAL GENERATORS AND FINITE TRANSFORMATIONS

An SU(1,1) matrix very close to the unit matrix I can be written as

$$A \simeq I - s_X\xi - s_Y\eta + s_T\tau, \tag{22}$$

where the *real* numbers ξ, η, τ are very small, and

$$s_X = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{23a}$$

$$s_Y = \tfrac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{23b}$$

$$s_T = \tfrac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{23c}$$

It is convenient to define a "vector" $\boldsymbol{\theta} = (\xi,\eta,\tau)$ and to write

$$\mathbf{s}\cdot\boldsymbol{\theta} \equiv -s_X\xi - s_Y\eta + s_T\tau. \tag{24}$$

The signs in (24) have been arbitrarily chosen (for later convenience), but no sign convention can alter the fact that the three commutation relations

$$[s_X, s_Y] = s_T, \tag{25a}$$

$$[s_Y, s_T] = -s_X, \tag{25b}$$

$$[s_T, s_X] = -s_Y, \tag{25c}$$

must have different signs on their right-hand sides.

Note that the s are pseudo-anti-Hermitian with respect to the metric σ_z,

$$\mathbf{s}^\dagger = -\sigma_z \mathbf{s} \sigma_z, \tag{26}$$

and that their anticommutators are

$$s_m s_n + s_n s_m = -G_{mn}/4, \tag{27}$$

where

$$G_{XX} = G_{YY} = -1, \tag{28a}$$

$$G_{TT} = 1, \tag{28b}$$

and the other $G_{mn} = 0$. It follows that

$$(\mathbf{s}\cdot\boldsymbol{\theta})^2 = -\theta^2/4, \tag{29}$$

where

$$\theta^2 = G_{mn}\theta_m\theta_n = \tau^2 - \xi^2 - \eta^2. \tag{30}$$

A *finite* transfer matrix can now be written as

$$\exp(\mathbf{s}\cdot\boldsymbol{\theta}) \equiv 1 + (\mathbf{s}\cdot\boldsymbol{\theta}) + (\mathbf{s}\cdot\boldsymbol{\theta})^2/2! + (\mathbf{s}\cdot\boldsymbol{\theta})^3/3! + \cdots \tag{31}$$

$$= \cos(\theta/2) + \sin(\theta/2)(\mathbf{s}\cdot\boldsymbol{\theta}/\theta), \tag{32}$$

by virtue of (29). Comparing with (17) and (23), we obtain

$$M = \cos(\theta/2) + \sin(\theta/2)(i\tau/\theta), \tag{33a}$$

$$N = -\sin(\theta/2)(\xi + i\eta)/\theta. \tag{33b}$$

These formulas are convenient when $\theta^2 > 0$, i.e., $|\mathrm{Re}M| < 1$. This is the conduction band, as defined in the preceding section. If $\theta^2 < 0$, one should replace $\cos(\theta/2)$ by $\cosh(|\theta|/2)$ and $\sin(\theta/2)/\theta$ by $\sinh(|\theta|/2)/|\theta|$. In the exceptional case $\theta^2 = 0$, one simply has $\exp(\mathbf{s}\cdot\boldsymbol{\theta}) = 1 + \mathbf{s}\cdot\boldsymbol{\theta}$ [see Eq. (A6)].

We now arrive at the main point of this article. The commutation relations (25) can be realized not only by 2×2 matrices, but also by matrices of higher dimensionality, for example

$$S_X = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{34a}$$

$$S_Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \tag{34b}$$

$$S_T = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{34c}$$

The basis for this three-dimensional representation of SU(1,1)—or, as we shall see, of the homomorphic group SO(2,1)—must be bilinear in the one used hitherto (namely f and g), just as the vectors acted upon by the ordinary rotation group SO(3) are bilinear in the two component spinors which are the basis for SU(2).

This can be seen as follows. Define a "vector' $\mathbf{R} = (X,Y,T)$ by[9]

$$X = \bar{f}g + f\bar{g} = (\bar{\psi}\psi - \dot{\bar{\psi}}\dot{\psi})/2, \tag{35a}$$

$$Y = i(\bar{f}g - f\bar{g}) = (\bar{\psi}\dot{\psi} + \dot{\bar{\psi}}\psi)/2, \tag{35b}$$

$$T = \bar{f}f + g\bar{g} = (\bar{\psi}\psi + \dot{\bar{\psi}}\dot{\psi})/2. \tag{35c}$$

It is straightforward to verify that an infinitesimal transformation

$$\Psi \to \Psi' = \Psi + (\mathbf{s}\cdot\boldsymbol{\theta})\Psi \tag{36}$$

indeed yields

$$\mathbf{R} \to \mathbf{R}' = \mathbf{R} + (\mathbf{S}\cdot\boldsymbol{\theta})\mathbf{R}, \tag{37}$$

where $\mathbf{S}\cdot\boldsymbol{\theta}$ is defined as in Eq. (24).

Here, one may be tempted to define a fourth component of $\mathbf{R}$ as $Z = \bar{f}f - g\bar{g}$; however, $Z = 1$ by virtue of (14). For the same reason

$$T^2 - X^2 - Y^2 = 1, \tag{38}$$

i.e., the vector $\mathbf{R}$ is constrained to lie on the upper sheet of a unit hyperboloid (Fig. 3).

We now see that the tensor G_{mn} of Eq. (28) is the indefinite metric of the $\mathbf{R}$ space. The relevant group is SO(2,1). Its representations have been discussed extensively.[10] This group has many applications in physical problems, such as collective motions in a nucleus,[11] superfluidity,[12] coherent states,[13] large-N expansions in quantum mechanics,[14] and, of course, as a subgroup of the Lorentz group. I shall henceforth freely use the "relativistic" terminology, although the problem discussed here has nothing to do with the theory of relativity.

Finite pseudorotations can be written as $\exp(\mathbf{S}\cdot\boldsymbol{\theta})$, just as in Eq. (31), but now there is no simple formula like (29) to go over to Eq. (32). Rather, we shall use a method similar to the one giving the matrix for a finite SO(3) rotation.[15] We *define a vector product* in (2 + 1)-Minkowski space as follows:

$$\boldsymbol{\theta}\times\mathbf{R} \equiv (\mathbf{S}\cdot\boldsymbol{\theta})\mathbf{R}, \tag{39}$$

or, explicitly,

$$\boldsymbol{\theta}\times\mathbf{R} \equiv (-\xi T - \tau Y, \eta T + \tau X, -\xi X + \eta Y). \tag{40}$$

Note that $\boldsymbol{\theta}\times\boldsymbol{\theta}$ does *not* vanish, i.e., $\boldsymbol{\theta}$ is *not* invariant under a pseudorotation through an "angle" $\boldsymbol{\theta}$. Rather, the invariant vector is

$$\boldsymbol{\theta}' = (-\eta, -\xi, \tau). \tag{41}$$

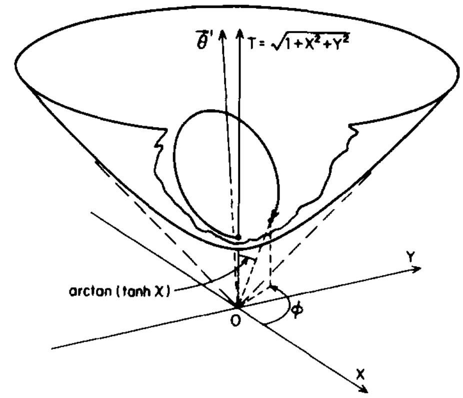

FIG. 3. The unit hyperboloid $T^2 - X^2 - Y^2 = 1$ with a trajectory resulting from a uniform pseudo-rotation around the invariant vector $\boldsymbol{\theta}'$. In this drawing, $\boldsymbol{\theta}'$ is timelike and the trajectory is an ellipse. For spacelike $\boldsymbol{\theta}'$, the trajectory would be a hyperbola.

It satisfies

$$(\boldsymbol{\theta}')' = \boldsymbol{\theta}, \tag{42}$$

$$\boldsymbol{\theta}\times\boldsymbol{\theta}' = 0, \tag{43}$$

and

$$\boldsymbol{\theta}'^2 = \boldsymbol{\theta}^2. \tag{44}$$

A direct caluclation then shows that

$$\boldsymbol{\theta}\times(\boldsymbol{\theta}\times\mathbf{R}) = -\theta^2\mathbf{R} + \boldsymbol{\theta}'(\boldsymbol{\theta}'\cdot\mathbf{R}) \tag{45}$$

and

$$\boldsymbol{\theta}\times(\boldsymbol{\theta}\times(\boldsymbol{\theta}\times\mathbf{R})) = -\theta^2(\boldsymbol{\theta}\times\mathbf{R}). \tag{46}$$

The last equation is formally identical to that for the Euclidean vector product, and we can therefore write[15]

$$\exp(\mathbf{S}\cdot\boldsymbol{\theta})\mathbf{R} = \mathbf{R} + (\boldsymbol{\theta}\times\mathbf{R}) + (\boldsymbol{\theta}\times(\boldsymbol{\theta}\times\mathbf{R}))/2 + - \theta^2(\boldsymbol{\theta}\times\mathbf{R})/3! - \theta^2(\boldsymbol{\theta}\times(\boldsymbol{\theta}\times\mathbf{R}))/4! + \cdots \tag{47}$$

$$= \mathbf{R} + (\sin\theta/\theta)(\boldsymbol{\theta}\times\mathbf{R}) + [(1-\cos\theta)/\theta^2]\boldsymbol{\theta}\times(\boldsymbol{\theta}\times\mathbf{R}), \tag{48}$$

$$= \cos\theta\mathbf{R} + (\sin\theta/\theta)(\boldsymbol{\theta}\times\mathbf{R}) + [(1-\cos\theta)/\theta^2]\boldsymbol{\theta}'(\boldsymbol{\theta}'\cdot\mathbf{R}). \tag{49}$$

[As usual, if $\theta^2 < 0$, we replace $\cos\theta$ by $\cosh|\theta|$ and $\sin\theta/\theta$ by $\sinh|\theta|/|\theta|$. If $\theta^2 = 0$, only the first three terms of (47) appear.]

We thus finally obtain the explicit form of the SO(2,1) transfer matrix:

(50a)

$$A \equiv \exp(\mathbf{S}\cdot\boldsymbol{\theta}) = \begin{pmatrix} -\eta^2 C + \cos\theta & -\xi\eta C - \tau S & -\eta\tau C - \xi S \\ -\xi\eta C + \tau S & -\xi^2 C + \cos\theta & -\xi\tau C + \eta S \\ \eta\tau C - \xi S & \xi\tau C + \eta S & \tau^2 C + \cos\theta \end{pmatrix}, \tag{50b}$$

where

$$S \equiv \sin\theta/\theta, \tag{51}$$

and

$$C \equiv (1-\cos\theta)/\theta^2 \equiv 2\sin^2(\theta/2)/\theta^2. \tag{52}$$

Now, by virtue of (33), it is also possible to write

$$\xi S = -2\mathrm{Re}M\,\mathrm{Re}N, \tag{53a}$$

$$\eta S = -2\mathrm{Re}M\,\mathrm{Im}N, \tag{53b}$$

$$\tau S = +2\mathrm{Re}M\,\mathrm{Im}M, \tag{53c}$$

with similar expressions for $\xi^2 C$, etc. Also

$$\cos\theta = 2\cos^2(\theta/2) - 1 = 2(\mathrm{Re}M)^2 - 1. \tag{54}$$

We thus obtain the equivalent expression

$$A = \begin{pmatrix} \mathrm{Re}(M^2+N^2) & -\mathrm{Im}(M^2+N^2) & 2\mathrm{Re}(M\overline{N}) \\ \mathrm{Im}(M^2-N^2) & \mathrm{Re}(M^2-N^2) & 2\mathrm{Im}(M\overline{N}) \\ 2\mathrm{Re}(MN) & -2\mathrm{Im}(MN) & |M|^2+|N|^2 \end{pmatrix}. \tag{55}$$

This result could also have been obtained directly from (35) and the transformation $f \to Mf + \overline{N}g$ and $g \to Nf + \overline{M}g$ [see Eq. (17)].

Note that the TT component in (55) is always larger than unity, because of (18). This is a general property of Lorentz transformations. Another general property,

$$\det A = 1, \tag{56}$$

is due to the fact that we have a *continuous* transformation generated by the *traceless* matrices (34). Moreover, it follows from (50) and (51) that

$$\mathrm{Tr}A = 1 + 2\cos\theta \tag{57}$$

(or $1 + 2\cosh|\theta|$ if $\theta^2 < 0$.) We can thereby find *all* the eigenvalues of $A = \exp(\mathbf{S}\cdot\boldsymbol{\theta})$: One of them must be unity, corresponding to the invariant eigenvector $\boldsymbol{\theta}'$ [see Eqs. (39) and (43)]. It then follows from the characteristic equation $\det(A - \lambda I) = 0$ that the two other ones are $\lambda_1 = e^{i\theta}$ or $\pm e^{|\theta|}$, and $\lambda_2 = 1/\lambda_1$. The discussion then proceeds just as after Eq. (20). If

$$-1 < \mathrm{Tr}A < 3, \tag{58}$$

the energy is in a conduction band; otherwise, it is in a forbidden band.

IV. SCHRÖDINGER EQUATION ON THE UNIT HYPERBOLOID

In this section, we shall write the one-dimensional time independent Schrödinger equation in terms of the real "vector" $\mathbf{R} = (X,Y,T)$, defined by Eq. (35). The "Hamiltonian" will be a *real* 3×3 matrix, similar to the pseudo-Hermitian "Hamiltonian" (11) used for the "spinor" representation (7). Indeed, Eq. (11) can be rewritten, by virtue of (23), as

$$\dot{\Psi} = [-us_Y - (2-u)s_T]\Psi. \tag{59}$$

Likewise

$$\dot{\mathbf{R}} = [-uS_Y - (2-u)S_T]\mathbf{R}, \tag{60}$$

as can be directly checked from (1) and (35). The "angular velocity"

$$\Omega = -uS_Y - (2-u)S_T = \begin{pmatrix} 0 & 2-u & 0 \\ u-2 & 0 & u \\ 0 & u & 0 \end{pmatrix} \tag{61}$$

is a pseudo-antisymmetric matrix, namely it is antisymmetric with respect to the indefinite metric G_{mn} defined by Eq. (28). The conserved "length" is $T^2 - X^2 - Y^2 = 1$ [see Eq. (38)].

The eigenvalues of Ω are 0 and $2(u-1)^{1/2}$. The invariant eigenvector satisfying $\Omega\boldsymbol{\theta}' = 0$ is

$$\boldsymbol{\theta}' = (u, 0, 2-u), \tag{62}$$

up to a normalization factor. From (62) we have $\boldsymbol{\theta}'^2 = 4(1-u)$. Thus, if $u < 1$ (i.e., $V < E$, the classically allowed region), $\boldsymbol{\theta}'$ is "timelike" and the nonnull eigenvalues of Ω are imaginary. On the other hand, for $u > 1$ (i.e., $V > E$), $\boldsymbol{\theta}'$ is "spacelike" and the nonnull eigenvalues of Ω are real. Note that these properties cannot be affected by a shift of the origin of the energy scale.

These results can be visualized as follows. The antisymmetric part of the matrix Ω is a Euclidean rotation in the XY plane (around the timelike T axis) with angular velocity $(2-u)$. The symmetric part is a Lorentz boost in the YT plane (around the space like X axis) with acceleration u. The Schrödinger equation can therefore be interpreted as a combined rotation and boost of the representative point $\mathbf{R} = (X,Y,T)$ on the unit hyperboloid

$$T = (1 + X^2 + Y^2)^{1/2}. \tag{63}$$

As we see from Eq. (61), Ω consists of a part proportional to $u = V/E$, and a part $\Omega_0 = -2S_T$, causing a rotation of $\mathbf{R}$ even in the absence of a potential. This free rotation can be eliminated by using a rotating coordinate system, in a way similar to Dirac's "interaction picture." Let

$$\Omega = \Omega_0 + u\Omega_1, \tag{64}$$

where $\Omega_1 = S_T - S_Y$. Then

$$\mathbf{R}' = \exp(-\Omega_0 t)\mathbf{R} \tag{65}$$

satisfies

$$\dot{\mathbf{R}}' = u\exp(-\Omega_0 t)\Omega_1\exp(\Omega_0 t)\mathbf{R}' \tag{66}$$

$$= u\begin{pmatrix} 0 & -1 & -\sin 2t \\ 1 & 0 & \cos 2t \\ -\sin 2t & \cos 2t & 0 \end{pmatrix}\mathbf{R}'. \tag{67}$$

This equation, which is exact, may conveniently be used as a starting point for a perturbation expansion, if u is small.

It is also possible to give the Schrödinger equation a remarkably simple form[9,16] by introducing polar coordinates χ and ϕ (see Fig. 3), thereby eliminating the constraint (63). Let

$$T = \cosh\chi \tag{68}$$

and

$$X + iY = \sinh\chi\, e^{i\phi}. \tag{69}$$

We obtain from (60), or directly from (1) and (35),

$$\dot{x} = u\sin\phi, \tag{70}$$

$$\dot{\phi} = u - 2 + u\coth\chi\cos\phi. \tag{71}$$

The surprising property of these first order equations is that *both are real*! It is very easy to convert the Schrödinger equation (1) into a pair of *complex* first-order equations, e.g., (6a), and (6b), but the fact that two *real* equations were obtained implies that some information has been lost. This can indeed be seen if we attempt to solve (35) for ψ, subject to the current normalization

$$i(\dot{\bar{\psi}}\psi - \bar{\psi}\dot{\psi})/2 = 1. \tag{72}$$

We obtain

$$\dot{\psi}/\psi = (Y - i)/(T + X), \tag{73}$$

but there is no algebraic solution for ψ itself. Note that the value of $\dot{\psi}/\psi$ is indifferent to multiplying ψ by a *constant* phase.

Likewise, if we try to solve (35) for f and g, subject to the constraint (14), we obtain

$$f = \cosh(\chi/2)\exp[i(\alpha + \phi)/2], \tag{74a}$$

$$g = \sinh(\chi/2)\exp[i(\alpha - \phi)/2], \tag{74b}$$

and the phase $e^{i\alpha/2}$ remains algebraically undetermined. However, it is not arbitrary! Substitution of (74) in (10) yields (70), (71), and also

$$\dot{\alpha} = -u\cos\phi/\sinh\chi, \tag{75}$$

so that only a constant phase is left undetermined.

The remarkable result here is that we do *not* need α to solve the "reduced" Schrödinger equation (70) and (71). Moreover, in many physical applications[9,16] χ is large, and we can replace $\coth\chi$ by unity in Eq. (71). The latter can then be solved, analytically or numerically, for ϕ and substitution in (70) readily yields χ.

To conclude this section, let us examine the case of a delta-function potential $V = V_0\delta(x)$, an approximation often used in solid state physics. As $\delta(x) = k\delta(t)$, we have

$$u = (2mV_0/\hbar^2 k)\delta(t) \equiv v\delta(t). \qquad (76)$$

This can be considered as the limit of a rectangular potential $u = v/\epsilon$ for $0 < t < \epsilon$, as $\epsilon \to 0$. Now, the transfer matrix for a constant potential u over a dimensionless distance $\Delta t = a$ is given by Eqs. (50)–(52), with

$$\boldsymbol{\theta} = (0, ua, -(2-u)a), \qquad (77)$$

as can be seen from (60). In the present case, $ua = -v$ and we get $\boldsymbol{\theta} = (0,v,v)$, so that $\theta^2 = 0$. We thus have, instead of (51) and (52), $S = 1$ and $C = \frac{1}{2}$, and (50) becomes

$$A(v) = \begin{pmatrix} 1 - v^2/2 & -v & -v^2/2 \\ v & 1 & v \\ v^2/2 & v & 1 + v^2/2 \end{pmatrix}. \qquad (78)$$

This can also be written as

$$A(v) = \exp\begin{pmatrix} 0 & -v & 0 \\ v & 0 & v \\ 0 & v & 0 \end{pmatrix}, \qquad (79)$$

from which it follows that

$$A(v)A(w) = A(v+w), \qquad (80)$$

just as in Eq. (A8). Since $\mathrm{Tr}A(v) = 3$, there is only a single degenerate eigenvalue, $\lambda = 1$, corresponding to a unique null invariant eigenvector

$$\boldsymbol{\theta}' = (-1,0,1). \qquad (81)$$

V. TRANSMISSION THROUGH DISORDERED CHAINS

There has recently been considerable interest in the conduction properties of disordered media. For a one-dimensional chain, it has been argued that the electric resistance is[17–20]

$$(\pi\hbar/e^2)\rho = (\pi\hbar/e^2)|N^2|. \qquad (82)$$

The universal constant $\pi\hbar/e^2$ is 12 906 Ω in engineering units, but the "resistance" (82) has properties very different from those familiar to electrical engineers! E.g., consider a rectangular barrier of length a and height $u = 1 - (\pi/2a)^2$. From Eq. (77) we have $\theta = \pi$ and, therefore, by Eq. (33b), $|N|^2 = \eta^2/\theta^2 = (a/\pi)^2 - \frac{1}{4}$. The "resistance" of this barrier can therefore be made arbitrarily large by increasing a. However, two consecutive barriers, as shown by Fig. 4, have zero "resistance," because $\theta = 2\pi$ and therefore $N = 0$.

Nevertheless, ρ, which is the ratio of reflection to transmission probabilities, has a direct physical meaning and an important physical problem is to find the ensemble average $\langle |N|^2 \rangle$ and higher moments $\langle |N|^{2p} \rangle$ for a given set of random chains. The mathematical techniques developed in the preceding section are ideally suited for this purpose.

Figure 2 shows that at the exit end of the barrier ($x = a$) we have $\mathbf{R}_0 = (0,0,1)$ and at its entrance $T = |M|^2 + |N|^2 = 2|N|^2 + 1$, by Eqs. (18) and (35c). Therefore,

$$\rho = (T-1)/2 = \sinh^2(\chi/2). \qquad (83)$$

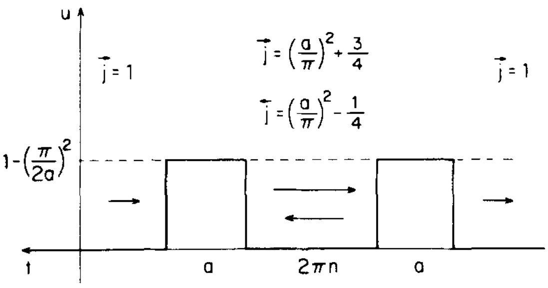

FIG. 4. Two identical barriers of thickness a and height $1 - (\pi/2a)^2$ are separated by an integral number of wavelengths. The transmission probability of *each* barrier can be made arbitrarily small by increasing a, but the *double* barrier has unit transmission probability.

As a simple example, consider an ensemble of one-dimensional chains, consisting of two types of sites represented by transfer matrices A and B, randomly distributed with probabilities a and $b = 1 - a$, respectively. Each site corresponds to an SO(2,1) rotation through a finite angle around some invariant vector, as in Fig. 3. However, even if both rotation axes are "timelike" (i.e., even if the energy is in the conduction band of each one of the sites), a product of these transfer matrices, such as AB, or $ABAAB\cdots$, may have a spacelike invariant vector (and vice versa). That this is in fact the *generic case* for a random product of transfer matrices may be seen as follows. Each one of our random chains is represented by a random walk on the hyperboloid of Fig. 3. As all the points of the hyperboloid are equivalent (and thus equiprobable) under the SO(2,1) group, the locus of $\mathbf{R}$ is almost certain to run away to infinity: Most random chains have very low transmission probabilities.

A quantitative estimate of this property can be obtained by noting that the average $\mathbf{R}$ after n sites simply is

$$\langle \mathbf{R} \rangle = (aA + bB)^n \mathbf{R}_0, \qquad (84)$$

because expanding the parenthesis yields all the configurations such as $ABAAB\cdots$, each one multiplied by the corresponding probability $abaab\cdots$ (we assume here that consecutive sites are uncorrelated). This result is readily generalized to more than two types of sites.

Now, the crucial fact is that the expression $aA + bB$ is *not* an SO(2,1) matrix (unless $A = B$) just as the average of two orthogonal matrices is not in general an orthogonal matrix. Moreover, we shall presently show that if $[A,B] \neq 0$, the largest eigenvalue of $(aA + bB)$, which dominates (84) for large n, is real and *must* exceed unity. The proof is given below in the only physically intersting case, when both A and B have "timelike" eigenvectors (all their eigenvalues are 1 and $e^{\pm i\theta}$).

First, assume that $b \ll 1$ and choose the T axis along the invariant eigenvector of A, so that

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (85)$$

as shown by Eq. (50) with $\xi = \eta = 0$ and $\tau = \theta$. The characteristic polynomial

$$D = \det[(1-b)A + bB - \lambda I] \qquad (86)$$

can be written as

$$D=(1-2\lambda\cos\theta+\lambda^2)[1-\lambda+b(B_{TT}-1)] \\ +(1-\lambda)O(b)+O(b^2), \tag{87}$$

and a solution of $D=0$ is

$$\lambda_1=1+b(B_{TT}-1)+O(b^2). \tag{88}$$

Since $B_{TT}>1$, as noticed after Eq. (55), we see that a perturbation $b(B-A)$ *always* shifts the eigenvalue $\lambda=1$ of A toward $\lambda_1>1$. (The only exception is if $B_{TT}=1$. This, however, implies not only $[A,B]=0$, but also that A and B must be powers of the *same* matrix. In other words, the two "different" sites simply consist of different numbers of *identical* sites.) The situation with b (or a) not infinitesimal is described by Fig. 5, which shows that there is always (at least) one real eigenvalue $\lambda_1>1$.

The next question is what happens to the complex eigenvalue $e^{\pm i\theta}$. It is shown in Appendix B that they may move inside or outside the unit circle, but that the eigenvalue having the largest magnitude is always real. (If it were not so, the *average* resistance of an ensemble of *random* chains would be an oscillatory function of their length, a rather unreasonable proposition.)

Returning to Eqs. (83) and (84), we see that

$$\langle\rho\rangle\sim\lambda_1{}^n, \tag{89}$$

i.e., the average resistance increases exponentially with the length of the random chains. This result, which had been known for a long time,[17] has little physical significance,[21] because this average $\langle\rho\rangle$ is not at all representative of the "typical" ρ: The value of the average is mostly due to exceedingly rare configurations of extremely high resistance which, if found by experimentalists, would be rejected as "broken" chains! (In some computer simulations involving thousands of random chains, it happened that not a single one exceeded the expected average.)

This amazing discrepancy between the average $\langle\rho\rangle$ and the "typical" ρ is best seen by calculating the standard deviation, or more simply

$$\langle\rho^2\rangle=\langle(T-1)^2\rangle/4. \tag{90}$$

We can obtain $\langle T^2\rangle$, or more generally $\langle T^p\rangle$[9,22] by investigating the $(2p+1)$-dimensional representations[10] of $\langle aA+bB\rangle$. E.g., we may take as basis XY, XT, YT, $(X^2-Y^2)/2$, and $(3T^2-1)/2$, which transform linearly under SO(2,1). The same argument as before then leads to

$$\langle\rho^2\rangle\sim\lambda_2{}^n, \tag{91}$$

where λ_2 is the largest eigenvalue of the five dimensional representation of $\langle aA+bB\rangle$. Now, since $\langle\rho^2\rangle\geqslant\langle\rho\rangle^2$, we must have $\lambda_2\geqslant\lambda_1{}^2$. In other words, the standard deviation increases faster than the mean (the distribution of ρ has a very long tail with a disproportionate influence on $\langle\rho\rangle$).

The problem of defining a "typical" ρ (more precisely, of finding a normally distributed function of ρ) has been studied by a number of authors[9,18,20,21] and will not be discussed in the present paper.

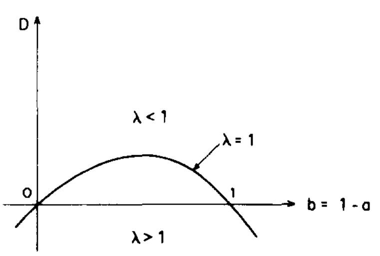

FIG. 5. The characteristic polynomial $D(b,\lambda)=\det(aA+bB-\lambda I)$ for real λ. As shown in the text, the line $\lambda=1$ must have a positive slope at $b=0$ and, for the same reason, a negative slope at $b=1$. As the polynomial D is of degree 3, that line cannot cut the b axis between $b=0$ and $b=1$. All the physical values $0<b<1$ therefore yield $\lambda>1$ for $D=0$.

VI. ENERGY RESCALING

In this section, we examine the consequences of shifting the origin of the energy scale, as in Eq. (5). Indeed, as long as we considered barriers with compact support, such as in Fig. 2, it was natural to set $V=0$ outside the barrier. However, there is no natural energy zero for an infinite chain, e.g., for a periodic potential. We shall first assume that the new energy $E'=\hbar^2k'^2/2m$ is positive, like E. The case $E'<0$ will be discussed in Sec. VII.

It is convenient to define the energy rescaling by

$$k'=e^{\Phi}k, \tag{92}$$

where Φ is a real *constant*. This is accompanied by a (dimensionless) *length* rescaling

$$t'=-k'x=e^{\Phi}t. \tag{93}$$

We then have

$$u'=[(2mV/\hbar^2)+k'^2-k^2]/k'^2 \tag{94a}$$

$$=e^{-2\Phi}(u-1)+1. \tag{94b}$$

Moreover, to retain the current normalization $j'=\hbar k'/m=e^{\Phi}j$, the new Schrödinger wavefunction must be $\psi'=e^{\Phi/2}\psi$. (Throughout this section, a prime means "new," not a derivative with respect to x.)

If we define f' and g' as in Eq. (4), we thus obtain

$$f'+g'=e^{\Phi/2}(f+g), \tag{95a}$$

$$f'-g'=e^{-\Phi/2}(f-g), \tag{95b}$$

whence

$$\Psi'=\begin{pmatrix}\cosh(\Phi/2) & \sinh(\Phi/2)\\ \sinh(\Phi/2) & \cosh(\Phi/2)\end{pmatrix}\Psi \tag{96a}$$

$$=\exp(\Phi s_X)\Psi. \tag{96b}$$

The transformation matrix is both real and pseudo-unitary. The new transfer matrices are given by

$$A'=\exp(\Phi s_X)A\exp(-\Phi s_X), \tag{97}$$

so that their trace (which determines whether the energy is in a conduction band or in a gap) is not affected. These transfer matrices have no other independent invariant because $\mathrm{Tr}(A^2)=(\mathrm{Tr}A)^2-2$.

Higher-dimensional representations transform in a similar way under energy rescaling. E.g., we have

$$\mathbf{R}'=\exp(\Phi S_X)\mathbf{R} \tag{98a}$$

$$=\begin{pmatrix}\cosh\Phi & 0 & \sinh\Phi\\ 0 & 1 & 0\\ \sinh\Phi & 0 & \cosh\Phi\end{pmatrix}\mathbf{R}. \tag{98b}$$

Note that Y is invariant, while $T \pm X$ scale as $e^{\pm \Phi}$, i.e., as $k^{\pm 2}$ or $E^{\pm 1}$.

Looking at Eq. (96), one is naturally tempted to ask why did the Schrödinger equation (11) involve only σ_y and σ_z, but no σ_x. Let us try to generalize (11) as

$$i\dot{\Psi} = [(1 - u/2)\sigma_z - (u/2)i\sigma_y + vi\sigma_x]\Psi. \tag{99}$$

If we impose $f + g = \psi$, we obtain

$$i\dot{\psi} = f - g + iv\psi, \tag{100}$$

so that we *cannot* have $f - g = i\dot{\psi}$ as in (4). Differentiating (100) once more with respect to t and using (99), we finally obtain

$$\ddot{\psi} = (u + \dot{v} + v^2 - 1)\psi, \tag{101}$$

which is a Schrödinger equation with potential

$$V = (u + \dot{v} + v^2)E. \tag{102}$$

Therefore, if V is such as to be conveniently split as in Eq. (102), the Schrödinger equation can be written as in Eq. (99).

VII. NEGATIVE ENERGY AND BOUND STATES

Rescaling the energy as in Eq. (92) can never give $E' < 0$, unless Φ is imaginary. This, however, is unacceptable because it would make the transformation (98) complex, while $\mathbf{R}'$ has to be real. We must therefore start afresh and write the Schrödinger equation as

$$-\psi'' + u\psi = -\kappa^2\psi, \tag{103}$$

where $u = V/E$ as before, and

$$\kappa^2 = -2mE/\hbar^2 > 0. \tag{104}$$

We now define $t = -\kappa x$ so that (103) becomes $\ddot{\psi} = (1 - u)\psi$, instead of $\ddot{\psi} = (u - 1)\psi$, which was valid for $E > 0$.

Instead of (4), we now write

$$f = (\psi - \dot{\psi})/2, \tag{105a}$$

$$g = (\psi + \dot{\psi})/2, \tag{105b}$$

and the Schrödinger equation becomes

$$\dot{\Psi} = [-(1 - u/2)\sigma_z + (u/2)i\sigma_y]\Psi. \tag{106}$$

The matrix

$$\omega = -(1 - u/2)\sigma_z + (u/2)i\sigma_y \tag{107}$$

has two interesting properties: All its elements are *real* (so that all transfer matrices A will now be real, even though Ψ may be complex), and ω is anti-Hermitian with respect to the metric σ_y:

$$\omega^\dagger = -\sigma_y\omega\sigma_y. \tag{108}$$

This means that $\Psi^\dagger\sigma_y\Psi$ is invariant under $\Psi \to A\Psi$ (indeed, $\Psi^\dagger\sigma_y\Psi = -2mj/\kappa\hbar$ must be conserved), i.e., A must belong to the real symplectic group Sp(2,R):

$$A^T\sigma_y A = \sigma_y. \tag{109}$$

In fact, *any* 2×2 real matrix with unit determinant satisfies (109)—Sp(2,R) is isomorphic to SL(2,R). This point will be further discussed in the next section.

Of special interest are normalizable negative energy wavefunctions, which correspond to *bound states*. They behave as $e^{-\kappa x} = e^t$ on the right-hand side of the potential well (the "barrier" will now be called a well) and as $e^{\kappa x} = e^{-t}$ on its left-hand side. We therefore have $f = 0$ and $g = 0$, respectively, on both sides of the well, so that the transfer matrix corresponding to a bound state has the form

$$A_{\rm bs} = \begin{pmatrix} a & b \\ -b^{-1} & 0 \end{pmatrix}. \tag{110}$$

For example, if we consider a square well ($u > 1$ constant over a distance $d = \kappa t$) we have from (106) and (107)

$$\begin{aligned} A &= e^{\omega t} \\ &= \cos[(u-1)^{1/2}t] + \omega(u-1)^{-1/2}\sin[(u-1)^{1/2}t]. \end{aligned} \tag{111}$$

The null matrix element in (110) is

$$\begin{aligned} &\cos[(u-1)^{1/2}t] + (1 - u/2)(u-1)^{-1/2}\sin[(u-1)^{1/2}t] \\ &\quad = 0, \end{aligned} \tag{112}$$

whence

$$\tan[(u-1)^{1/2}t] = -(1 - u/2)(u-1)^{-1/2}. \tag{113}$$

This is identical with the familiar eigenvalue equation

$$\tan kd = 2\kappa k/(k^2 - \kappa^2), \tag{114}$$

where $k^2 = (u-1)\kappa^2 = 2m(E - V)/\hbar^2$. We have thereby found the energy eigenvalues *algebraically*, without having to solve any differential equation.

The same results can also be expressed in the SO(2,1) formalism. By analogy with (35) we define

$$X = \bar{f}g + f\bar{g} = (\bar{\psi}\psi - \bar{\psi}\dot{\psi})/2, \tag{115a}$$

$$Z = \bar{f}f - g\bar{g} = -(\dot{\bar{\psi}}\psi + \bar{\psi}\dot{\psi})/2, \tag{115b}$$

$$T = \bar{f}f + g\bar{g} = (\bar{\psi}\psi + \dot{\bar{\psi}}\dot{\psi})/2. \tag{115c}$$

The fourth combination,

$$i(\bar{f}g - f\bar{g}) = i(\bar{\psi}\dot{\psi} - \dot{\bar{\psi}}\psi)/2, \tag{116}$$

is proportional to the current and can be normalized to 1, unless it is zero (e.g., for a bound state). We thus have

$$T^2 - X^2 - Z^2 = 1 \quad (\text{or } 0). \tag{117}$$

In particular a bound state can be defined by the following boundary conditions (at both ends of the well): $X = 0$, $Z = \pm 1$, and $T = 1$.

VIII. ENERGY INDEPENDENT REPRESENTATION

We have seen that positive energy wavefunctions are conveniently related by transfer matrices belonging to SU(1,1), and negative energy ones by matrices belonging to Sp(2,R). In both cases, it is possible to construct transfer matrices belonging to SO(2,1), but we need different bilinear combinations, (35) and (115), for $E \gtrless 0$.

In this section, we show how to construct transfer matrices in a way which does not depend on the energy. These matrices belong to the group SL(2,R), which is homomorphic to the three other ones. It is in fact *isomorphic* to Sp(2,R), as noted previously.

The SL(2,R) basis $\{f,g\}$ which was used in the preceding section is given by (105), where $\dot{\psi} = (-1/\kappa)(\partial\psi/\partial x)$. It follows that *any* real linear combinations of f and g are also a basis for SL(2,R), because $\det(SAS^{-1}) = \det(A) = 1$, for any nonsingular S. In particular, it is possible to choose as our basis[23] ψ and $\psi' = \partial\psi/\partial x$, even though they are not dimensionally homogeneous.

As an example, let us write down the SL(2,R) matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with basis $\{\psi,\psi'\}$, corresponding to the SU(1,1) matrix (17). We readily obtain from (4)

$$M = [a + ikb + (c/ik) + d]/2, \tag{118a}$$

$$N = [a + ikb - (c/ik) - d]/2, \tag{118b}$$

whence

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \mathrm{Re}(M+N) & Im(M+N)/k \\ -k\,\mathrm{Im}(M-N) & \mathrm{Re}(M-N) \end{pmatrix}. \tag{119}$$

Note that $|M|^2 - |N|^2 = ad - bc = 1$, as expected.

The equations of "motion"

$$\frac{d\psi}{dx} = \psi', \tag{120a}$$

$$\frac{d\psi'}{dx} = \frac{2m}{\hbar^2}(V-E)\psi \tag{120b}$$

are now identical for both signs of E. On the other hand, it is impossible to construct a bilinear "vector" $\mathbf{R}$ such as (35) or (115) which is dimensionally homogeneous and real for *both* signs of E (because some components must involve $E^{1/2}$ for dimensional homogeneity). However, if the latter requirement is *not* imposed, one can define a real basis such as

$$X = (\bar\psi\psi - \bar\psi'\psi')/2, \tag{121a}$$

$$Y = (\bar\psi\psi' + \bar\psi'\psi)/2, \tag{121b}$$

$$T = (\bar\psi\psi + \bar\psi'\psi')/2, \tag{121c}$$

satisfying

$$T^2 - X^2 - Y^2 \equiv [(\bar\psi\psi' - \bar\psi'\psi)/2i]^2 = \text{const.} \tag{122}$$

We again encounter the SO(2,1) group.

The rescaling transformation discussed in Sec. VI is simply given by

$$\begin{pmatrix} e^{\Phi/2} & 0 \\ 0 & e^{-\Phi/2} \end{pmatrix}, \tag{123}$$

for the $\{\psi,\psi'\}$ basis and by Eq. (98b) for the $\{X,Y,T\}$ basis.

IX. CONCLUDING REMARKS

The formalism developed in this paper may also have interesting applications for some problems intermediate between exactly periodic potentials and completely random ones. For example, a potential with two incommensurable periods[24] may be described as generating two simultaneous SO(2,1) rotations around different axes. Another possible application could be the addition of a homogeneous electric field to a periodic potential, a notoriously controversial problem.[25]

In retrospective, one might be surprised that group theory was found useful in a problem with no apparent symmetry. In fact, there is (of course) a "hidden" symmetry: If boundary conditions allow ψ to be genuinely complex for some E (namely, ψ cannot be made real by adjusting a constant phase), then ψ and $\bar\psi$ are linearly independent solutions of the Schrödinger equation (1) and so is any linear combination thereof. These linear combinations can then be used as a manifold for the representation of a continuous group of transformations. As is well known, the underlying symmetry is time reversal—which is implicit in the time-independent Schrödinger equation.

ACKNOWLEDGMENTS

It is a pleasure to thank Micha Revzen and Amiram Ron for suggesting the problem which led to this work. The question discussed in Appendix B was raised by Joshua Zak. This research was partly supported by the Gerard Swope Fund and by the Fund for Encouragement of Research at Technion. Some results included in this paper were obtained during a visit at Bell Laboratories (Murray Hill, N.J.). I am very grateful to John R. Klauder for his warm hospitality.

APPENDIX A: PSEUDONORMAL MATRICES

The transfer matrices which we considered here are not normal (they do not commute with their adjoints) and some of their properties may be unfamiliar to physicists accustomed to Hermitian or unitary matrices. This appendix briefly states some theorems needed in the text. (For the special case of 3×3 matrices, see Ref.26.)

A square matrix A has right eigenvectors $A|v_\lambda\rangle = \lambda|v_\lambda\rangle$ and left eigenvectors $\langle u_\lambda|A = \lambda\langle u_\lambda|$. The eigenvalues λ are the same, since they arise from the same characteristic equation. From

$$\langle u_\lambda|A|v_\mu\rangle = \lambda\langle u_\lambda|v_\mu\rangle = \mu\langle u_\lambda|v_\mu\rangle, \tag{A1}$$

it follows that

$$(\lambda-\mu)\langle u_\lambda|v_\mu\rangle = 0. \tag{A2}$$

Thus, left eigenvectors and right eigenvectors belonging to different eigenvalues are mutually orthogonal.

Let us now assume that all the eigenvalues are different, so that the eigenvectors are linearly independent. We can normalize them by

$$\langle u_\lambda|v_\mu\rangle = \delta_{\lambda\mu}. \tag{A3}$$

(This determines the $\langle u|$ once the $|v\rangle$ are given, and vice versa.) Then the $P_\lambda = |v_\lambda\rangle\langle u_\lambda|$ are a complete set of mutually exclusive projection operators, namely $P_\lambda P_\mu = \delta_{\lambda\mu}P_\mu$ and $\Sigma_\lambda P_\lambda = I$. We can then write

$$A = \Sigma_\lambda \lambda P_\lambda, \tag{A4}$$

and therefore

$$A^n = \Sigma_\lambda \lambda^n P_\lambda. \tag{A5}$$

It follows that A^n is controlled by the largest $|\lambda|$ and increases exponentially with n if the latter is larger than 1.

However, if some eigenvalues are degenerate, the situation becomes radically different. E.g., consider the SU(1,1) matrix

$$A(a) = \begin{pmatrix} 1+ia & ia \\ -ia & 1-ia \end{pmatrix}. \tag{A6}$$

Its only eigenvalue is $\lambda = 1$ and the corresponding eigenvectors are, up to a factor, $|v\rangle = (1,-1)$ and $\langle u| = (1,1)$. Instead of (A5) we have

$$A(a)^n = A(na), \tag{A7}$$

or more generally

$$A(a)A(b) = A(a+b). \quad \text{(A8)}$$

Although the transfer matrices used in this paper are not normal, we can call them "pseudonormal" because they satisfy

$$A^{-1} = \eta A^{\dagger}\eta, \quad \text{(A9)}$$

and therefore

$$[A^{\dagger}, \eta A \eta] = 0. \quad \text{(A10)}$$

Here, η is the indefinite metric corresponding to the representation being used, e.g., $\eta = \sigma_z$ for SU(1,1) matrices or $\eta = G$ [cf. Eq. (28)] for SU(2,1) matrices. As $\eta^{\dagger} = \eta$ and $\eta^2 = 1$ for all representations, we have

$$A^{\dagger}\eta|v_\lambda\rangle = \eta A^{-1}|v_\lambda\rangle = \lambda^{-1}\eta|v_\lambda\rangle, \quad \text{(A11)}$$

whence, taking Hermitian conjugates,

$$\langle v_\lambda|\eta A = \bar{\lambda}^{-1}\langle v_\lambda|\eta. \quad \text{(A12)}$$

It follows that

$$\langle v_\lambda|\eta = \langle u_{1/\bar{\lambda}}|, \quad \text{(A13)}$$

so that if λ is an eigenvalue, $1/\bar{\lambda}$ must also be one and (A13) gives the relationship between the corresponding eigenvectors. In particular, if $\lambda = 1/\bar{\lambda} = e^{i\theta}$, we simply have $\langle v_\lambda|\eta = \langle u_\lambda|$. On the other hand, if λ is real, then $\langle v_\lambda|\eta = \langle u_{1/\lambda}|$.

In summary, we have the following pseudo-orthogonality relations

$$\langle v_\lambda|\eta|v_\mu\rangle = 0, \quad \text{(A14)}$$

if λ and μ are real and $\lambda\mu \neq 1$; or if λ and/or μ is $e^{i\theta}$ and $\lambda \neq \mu$.

APPENDIX B: THE COMPLEX EIGENVALUES OF $(aA + bB)$

It was shown in Sec. V. that the real eigenvalue of $aA + bB$, continuously connected to the eigenvalue $\lambda = 1$ of A and B, satisfies $\lambda_1 > 1$ (see Fig. 5). For small b it is given by Eq. (88), where B_{TT} is the TT matrix element of B in the coordinate system where A has the form given by Eq. (85). With the notations of Appendix A,

$$B_{TT} = \mathrm{Tr}(BP_1) = \langle u_1|B|v_1\rangle. \quad \text{(B1)}$$

In this appendix, we examine the behavior of the two other eigenvalues of $(aA + bB)$, which are continuously connected to the complex conjugate eigenvalues of A and B. It will be shown that if they are also complex conjugate, their absolute value cannot exceed λ_1, so that the matrix elements of $(aA + bB)^n$ indeed behave as $(\lambda_1)^n$ for large n.

Let λ and $\bar{\lambda}$ be these two complex eigenvalues. From the characteristic equation, we have

$$\lambda_1|\lambda|^2 = \det(aA + bB). \quad \text{(B2)}$$

The right-hand side of (B2) is a third degree polynomial in $b = 1 - a$, having the value 1 at $b = 0$ and $b = 1$ (see Fig. 6). As we did in Fig. 5, we first examine the case $b \ll 1$. We have

$$\det(aA + bB) = 1 + b\,[B_{TT} + (B_{XX} + B_{YY})\cos\theta + (B_{YX} - B_{XY})\sin\theta - 3] + O(b^2). \quad \text{(B3)}$$

Now

$$|(B_{XX} + B_{YY})\cos\theta + (B_{YX} - B_{XY})\sin\theta| \leqslant [(B_{XX} + B_{YY})^2 + (B_{YX} - B_{XY})^2]^{1/2}. \quad \text{(B4)}$$

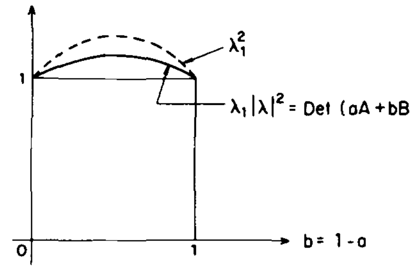

FIG. 6. The solid line is a cubic with a local maximum for $0 < b < 1$. Equation (B6) shows that its slope at $b = 0$ (and likewise at $b = 1$) is less than that of λ_1^2, drawn as a dashed line. The latter also can have only a single maximum for $0 < b < 1$ (see Fig. 5). It is therefore very plausible that these two lines do not intersect between 0 and 1, but a formal proof seems difficult. [As explained after Eq. (88), the opposite result would be physically unreasonable.]

By virtue of Eq. (55), the right-hand side of (B4) simply is $|M|^2 = |N|^2 + 1$, where M and N refer to the B matrix in the coordinate system where A is given by (85). Also

$$\lambda_1 = 1 + b(B_{TT} - 1) = 1 + 2b|N|^2, \quad \text{(B5)}$$

and we obtain from (B3) and (B4)

$$1 \leqslant \det(aA + bB) \leqslant 1 + 4b|N|^2 \leqslant \lambda_1^2. \quad \text{(B6)}$$

It then follows from (B2) that we have

$$\lambda_1^{-2} < \lambda_1^{-1} \leqslant |\lambda|^2 \leqslant \lambda_1 < \lambda_1^2. \quad \text{(B7)}$$

This was proved for small b (and likewise for small a). The situation for finite b is illustrated by Fig. 6, which clearly indicates that the relation (B7) should be valid for all $0 < b < 1$. A formal proof, however, seems very difficult.

[1] E. H. Lieb and D. C. Mattis, Eds., *Mathematical Physics in One Dimension; Exactly Soluble Models of Interacting Particles* (Academic, New York, 1966).

[2] R. Gilmore, *Lie Groups, Lie Algebras and Some of Their Applications* (Wiley, New York, 1974).

[3] H. Feshbach and F. Villars, Rev. Mod. Phys. **30**, 24 (1958), Sec. 2B.

[4] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), p. 26.

[5] R. E. Borland, Proc. Phys. Soc. **78**, 926 (1961), Appendix.

[6] P. Erdös and R. C. Herndon, in *Computational Methods for Large Molecules and Localized States in Solids*, edited by F. Herman *et al.* (Plenum, New York, 1973), p. 275.

[7] E. Abrahams and M. J. Stephen, J. Phys. C **13**, L377 (1980).

[8] B. S. Andereck and E. Abrahams, J. Phys. C **13**, L383 (1980).

[9] A. Peres, M. Revzen, and A. Ron, Phys. Rev. B **24**, 7463 (1981).

[10] A. O. Barut and C. Fronsdal, Proc. Roy. Soc. (Lond.) A **287**, 532 (1965).

[11] S. Goshen and H. J. Lipkin, Ann. Phys. (N.Y.) **6**, 301 (1959).

[12] A. I. Solomon, J. Math. Phys. **12**, 390 (1971).

[13] A. M. Perelomov, Usp. Fiz. Nauk **20**, 23 (1977) [Sov. Phys. Usp. **20**, 703 (1977)].

[14] L. D. Mlodinow and N. Papanicolaou, Ann. Phys. (N.Y.) **128**, 314 (1980).

[15] A. Peres, Am. J. Phys. **48**, 70 (1980).

[16] R. Kree and A. Schmid, preprint, Univ. Karlsruhe, 1981.

[17] R. Landauer, Phil. Mag. **21**, 863 (1970).

[18] P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, Phys. Rev. B **22**, 3519 (1980).

[19] M. Ya. Azbel, Phys. Lett. A **78**, 410 (1980).

[20] E. N. Economou and C. M. Soukoulis, Phys. Rev. Lett. **46**, 618 (1981).

[21] J. Sak and B. Kramer, Phys. Rev. B. **24**, 1761 (1981).

[22] A. D. Stone and J. D. Joannopoulos, Phys. Rev. B **24**, 3592 (1981).

[23] B. I. Halperin, Adv. Chem. Phys. **13**, 123 (1967).

[24] M. Ya. Azbel, Phys. Rev. Lett. **43**, 1954 (1979).

[25] J. Zak, Phys. Rev. B **19**, 1946 (1979).

[26] V. K. Stokes, J. Math. Phys. **21**, 2842 (1980).

A probabilistic formulation of quantum theory. III

Thomas Wenning and Alexander Bach
Institut für Theoretische Physik I, Universität Münster, D-4400 Münster, West Germany

(Received 26 July 1982; accepted for publication 3 December 1982)

A representation of the expectation value of quantum mechanics, which recently has been set up for normal states, is generalized to singular states. It is shown that states can be represented by finitely additive measures on the state space, which are σ-additive if and only if the states are normal.

PACS numbers: 03.65.Ca, 02.50.Cw

I. INTRODUCTION

Recently we have set up a formulation of elementary quantum mechanics by means of classical probability theory.[1,2] In this formulation, the well-known trace formula for the expectation value is replaced by an integral expression quite in analogy to the expectation value in classical statistical mechanics.

For an observable represented by a self-adjoint operator A on a complex separable Hilbert space H, the expectation value reads

$$E(A;W) = \mathrm{tr}(WA), \tag{1}$$

where W is a statistical operator describing the state of the system. As has been shown in Ref. 1, the expectation value can equally well be expressed by the formula

$$E(A;W) = \int_H d\mu_W(\phi) f_A(\phi), \tag{2}$$

where μ_W is a probability measure on H determined by the statistical operator W, and $f_A : H \to \mathbb{R}$ is defined by

$$f_A(\phi) = \langle \phi | A | \phi \rangle. \tag{3}$$

Whereas in Refs. 2 and 3 this formalism has been generalized to unbounded operators, here we are concerned with the representation of singular states within the framework of the probabilistic formulation. To this end, let us remember some results concerning singular states where we confine ourselves, as far as the observables are concerned, to the C^*-algebra $L(H)$ of bounded operators on H.

Let us first fix some notation. By $L_\infty(H)$ we denote the space of compact operators and by $L_{\mathrm{sa}}(H)$ the real Banach space of bounded self-adjoint operators on H. $S(H)$ denotes the set of statistical operators, i.e., the positive, normalized trace-class operators on H. The space of linear, continuous, positive, normalized functionals on $L(H)$ is the state space $E_{L(H)}$ of $L(H)$. For more information on the mathematical background cf., e.g., Ref. 4.

A state ω on $L(H)$ is called normal if

$$\sup\{\omega(A_n)\} = \omega(\sup\{A_n\}) \tag{4}$$

for all increasing nets $\{A_n\}$ of positive elements of $L(H)$ with an upper bound in $L(H)$. It can be shown, cf. Ref. 4, that this definition is equivalent to the following more familiar characterization. A normal state $\omega \in E_{L(H)}$ is determined by the fact that it can be represented unambiguously by a statistical operator $W \in S(H)$ such that $\omega = \omega_W$ and

$$\omega_W(A) = \mathrm{tr}(WA). \tag{5}$$

The subset of normal states is denoted by $N_{L(H)}$ and identified with $S(H)$.

It can be shown, cf. Ref. 4, that every $\omega \in E_{L(H)}$ has a unique decomposition

$$\omega = \lambda\omega_n + (1-\lambda)\omega_s, \tag{6}$$

$\lambda \in [0,1]$, $\omega_n, \omega_s \in E_{L(H)}$ such that $\omega_n \in N_{L(H)}$ and $\omega_s \notin N_{L(H)}$ and the decomposition of ω_s in analogy to Eq. (6) has no normal part. This state ω_s is called the singular part of ω. It is characterized by $\omega_s(A) = 0$ for all $A \in L_\infty(H)$.

The normal states can be characterized from a probabilistic point of view as follows. Consider Lat(H), the lattice of projections on closed subspaces of H. In analogy to a probability measure on a measurable space, a generalized probability measure π is a [0,1]-valued function on Lat(H), normalized to $\pi(\mathbb{1}) = 1$ and satisfying the condition of σ-additivity,

$$\pi\Big(\bigvee_i P_i\Big) = \sum_i \pi(P_i), \tag{7}$$

for any countable family of mutually orthogonal elements P_i of Lat(H). The theorem of Gleason, cf. Ref. 5, assures that for $\dim(H) \geqslant 3$ a generalized probability measure π uniquely is characterized by a statistical operator $W \in S(H)$, $\pi = \pi_W$ such that

$$\pi_W(P) = \mathrm{tr}(WP) \tag{8}$$

holds for all $P \in$ Lat(H). Identifying elements of Lat(H) with random variables in this generalized theory of probability formula (1) is easily reconstructed.

These considerations show that generalized σ-additive probability measures on Lat(H) and normal states can be identified and that singular states lie outside the framework of proposition calculus which is based on Lat(H). Nonetheless, singular states, as well as mixtures of singular and normal states, are necessary and useful for the description of scattering states, ergodic states, or idealized states associated with a point in a continuous spectrum.

The proof of Eq. (2) shows that this formula is confined to normal states. It is the purpose of the present paper to generalize this formalism to include singular states. This is achieved by representing operators by functions and states by functionals. An appropriate extension of these allows the identification of states with finitely additive measures on the

 0022-2488/83/051120-03$02.50

Hilbert space. These results are discussed with respect to their structural status in the framework of the probabilistic formulation of quantum theory.

II. REPRESENTATION OF SINGULAR STATES

To include singular states in the scheme setup in Refs. 1 and 2, we proceed in several steps.

(i) Instead of representation (2), we prefer to express the expectation value as

$$E(A;W) = \int_{H \setminus \{0\}} d\nu_W(\phi) f^A(\phi), \tag{9}$$

where ν_W is a probability measure on $(H \setminus \{0\}, \mathscr{B}(H \setminus \{0\}))$ defined by

$$\nu_W(B) = \int_B d\mu_W(\phi) \|\phi\|^2. \tag{10}$$

$B \in \mathscr{B}(H \setminus \{0\})$ and $f^A : H \setminus \{0\} \to \mathbb{R}$ is given by

$$f^A(\phi) = \|\phi\|^{-2} f_A(\phi) = \langle \phi | A | \phi \rangle / \|\phi\|^2. \tag{11}$$

(ii) Consider $C_b(H \setminus \{0\})$, the Banach space of continuous, real-valued, bounded functions on $H \setminus \{0\}$ with the sup-norm

$$\|f\| = \sup\{|f(\phi)|; \phi \in H \setminus \{0\}\} \tag{12}$$

and define $i: L_{\rm sa}(H) \to C_b(H \setminus \{0\})$ by

$$i(A) = f^A. \tag{13}$$

Obviously, $f^A \in C_b^q(H \setminus \{0\})$, which is defined as the set

$$C_b^q(H \setminus \{0\}) = \{f \in C_b(H \setminus \{0\}); \text{ there is an } A \in L_{\rm sa}(H) \text{ such that } f = f^A\}. \tag{14}$$

The mapping i is injective and $i(A) \in C_b(H \setminus \{0\})$. By virtue of

$$\|i(A)\| = \|A\|, \tag{15}$$

i is bounded and an isometry such that we obtain the following result:

Lemma 1: The mapping i is an isometric isomorphism from $L_{\rm sa}(H)$ onto $C_b^q(H \setminus \{0\})$ and $C_b^q(H \setminus \{0\})$ is a norm-closed subspace of $C_b(H \setminus \{0\})$.

(iii) Let ω be a state on $L(H)$, i.e., $\omega \in E_{L(H)}$. The mapping $\tilde{\omega}$ on $C_b^q(H \setminus \{0\})$,

$$\tilde{\omega}(f^A) = \omega(i^{-1}(f^A)) = \omega(A), \tag{16}$$

defines a linear, positive, continuous functional. It is our aim to extend $\tilde{\omega}$ to all of $C_b(H \setminus \{0\})$, which is a vector lattice, and to apply abstract integration theory there.

Proposition 1: Let ω be a state on $L(H)$. Then $\tilde{\omega}$, defined by (16), has a positive, linear, norm-continuous extension to $C_b(H \setminus \{0\})$.

For the proof, we use a theorem of Krein-Rutmann.[6] It is sufficient to show that $C_b^q(H \setminus \{0\}) \cap C_b^+(H \setminus \{0\})$; $C_b^+(H \setminus \{0\})$, denoting the positive cone of $C_b(H \setminus \{0\})$, contains an interior point of $C_b^+(H \setminus \{0\})$. Take $f^1 \in C_b^q(H \setminus \{0\}), f^1(\phi) = 1$, and verify that the open ball of radius $\frac{1}{2}$ with center f^1, $U_{1/2}(f^1) \subset C_b^q(H \setminus \{0\}) \cap C_b^+(H \setminus \{0\})$, is contained in $C_b^+(H \setminus \{0\})$.

The extension of $\tilde{\omega}$, the existence of which is assured by this theorem, will be denoted by $\hat{\omega}$.

(iv) $C_b(H \setminus \{0\})$ is a Stonian vector lattice (cf., e.g., Ref. 7). An abstract integral, also called a Daniell integral, on a Stonian vector lattice F of functions is a linear, positive, σ-continuous functional j on this lattice, where σ-continuity refers to the property that $f_n \uparrow f, f = \sup\{f_n\}$ implies that

$$\sup\{j(f_n)\} = j(\sup\{f_n\}) \tag{17}$$

for any nondecreasing sequence of functions f_n of F.

Let j now denote an abstract integral on $C_b(H \setminus \{0\})$. According to the general theory (cf., e.g., Ref. 7, Theorem 40.5) there exists a unique finite Borel measure μ on $(H \setminus \{0\}, \mathscr{B}(H \setminus \{0\}))$ such that $C_b(H \setminus \{0\}) \subset L^1(H \setminus \{0\}, \mathscr{B}(H \setminus \{0\}), \mu)$ and

$$j(f) = \int_{H \setminus \{0\}} d\mu(\phi) f(\phi) \tag{18}$$

for all $f \in C_b(H \setminus \{0\})$.

As $\hat{\omega}$ is a linear, positive, norm-continuous functional on the Stonian vector lattice $C_b(H \setminus \{0\})$, we want to analyze its properties more precisely. To this end, let j be a positive, linear functional on a vector lattice F. Then there exists, cf. Ref. 8, a unique decomposition of j into a σ-continuous part j_c and a purely discontinuous (non-σ-continuous) part j_d which cannot be decomposed further into parts containing σ-continuous constituents such that

$$j = j_c + j_d. \tag{19}$$

Equation (18) allows for a representation of j_c in (19).

(v) We now are able to apply these results to the state $\hat{\omega}$ defined on $C_b(H \setminus \{0\})$ in terms of $\omega \in E_{L(H)}$.

Proposition 2: Let $\omega \in E_{L(H)}$ and $\hat{\omega}$ the associated representation as functional on the Stonian vector lattice $C_b(H \setminus \{0\})$. Then the decomposition of ω into a normal and a singular part, $\omega = \lambda \omega_n + (1 - \lambda)\omega_s$, is equivalent to the decomposition of $\hat{\omega}$, i.e., $(\hat{\omega})_c = (\lambda \omega_n)\hat{}$ and $(\hat{\omega})_d = ((1-\lambda)\omega_s)\hat{}$.

Proof: (a) Assume that ω is normal. Then we can use ν_ω, the measure on $H \setminus \{0\}$ set up in step (i), to define $\hat{\omega}$ on $C_b(H \setminus \{0\})$ by

$$\hat{\omega}(f) = \int_{H \setminus \{0\}} d\nu_\omega(\phi) f(\phi). \tag{20}$$

As $f \in C_b(H \setminus \{0\})$, the integral exists. Moreover, due to the fact that abstract integrals and finite Borel measures coincide on $H \setminus \{0\}$, this is already the desired representation.

(b) Let ω be a singular state and $\hat{\omega}$ the associated linear, positive, norm-continuous functional on $C_b(H \setminus \{0\})$. As there exists an increasing net $\{A_n\}$ of positive elements of $L_{\rm sa}(H)$ such that $\sup\{\omega(A_n)\} \neq \omega(\sup\{A_n\})$, we infer that $\sup\{\hat{\omega}(f^{A_n})\} \neq \hat{\omega}(\sup\{f^{A_n}\})$. Here we use the fact that positivity, on a complex Hilbert space, implies self-adjointness such that the f^{A_n} are well defined. The violation of equality shows that $\hat{\omega}$ is not σ-continuous and, by contradiction, we conclude that $\hat{\omega}$ is purely discontinuous.

(c) Let $\omega \in E_{L(H)}$, $\hat{\omega}$ the associated functional on $C_b(H \setminus \{0\})$, and $\hat{\omega} = (\hat{\omega})_c + (\hat{\omega})_d$ the decomposition of $\hat{\omega}$ with respect to σ-continuity. As $(\lambda \omega_n)\hat{}$ is σ-continuous and $((1-\lambda)\omega_s)\hat{}$ is purely discontinuous, the uniqueness of the decompositions gives the assertion.

(vi) According to the procedure set up in Ref. 9, $\hat{\omega}$ can be used to associate with every state ω a positive, finitely additive set function μ_ω on a Boolean lattice $\mathscr{F}$ (depending on $\hat{\omega}$) of subsets of $H \setminus \{0\}$. This family of sets contains $H \setminus \{0\}$ as $1_{H \setminus \{0\}} \in C_b(H \setminus \{0\})$ such that

$$\mu_\omega(H \setminus \{0\}) = \hat{\omega}(1_{H \setminus \{0\}}) = \tilde{\omega}(f^1) = \omega(\mathbb{1}) = 1. \tag{21}$$

Quite in analogy to the decomposition of ω and $\hat{\omega}$, this set function μ_ω can be uniquely decomposed into a σ-additive part μ_ω^c and a purely finitely additive part μ_ω^d such that

$$\mu_\omega = \mu_\omega^c + \mu_\omega^d. \tag{22}$$

Considering the parts $\hat{\omega}_c$ and $\hat{\omega}_d$ separately on $C_b(H \setminus \{0\})$, we obtain Boolean lattices $\mathscr{F}_c$ and $\mathscr{F}_d$ of subsets of $(H \setminus \{0\})$ and finitely additive, positive set functions $\hat{\mu}_\omega^c$ and $\hat{\mu}_\omega^d$ which can be associated with these functionals. Here $\hat{\mu}_\omega^c$ defined on $\mathscr{F}_c$ is σ-additive and $\hat{\mu}_\omega^d$ defined on $\mathscr{F}_d$ is purely finitely additive. As $\mathscr{F} = \mathscr{F}_c \cap \mathscr{F}_d$, the set functions $\hat{\mu}_\omega^c$ and $\hat{\mu}_\omega^d$ can be restricted to $\mathscr{F}$ with the result (cf. Ref. 9) that on $\mathscr{F}$ we have

$$\mu_\omega^c = \hat{\mu}_\omega^c, \quad \mu_\omega^d = \hat{\mu}_\omega^d. \tag{23}$$

For the continuous part of $\hat{\omega}$, namely $\hat{\omega}_c$, we already know that it can be represented by means of a measure, which can be identified with $\hat{\mu}_\omega^c$, on $\mathscr{B}(H \setminus \{0\})$. As the σ-continuity properties of $\hat{\omega}$ transfer to μ_ω, the decomposition of μ_ω allows us to state the central result.

III. CONCLUSION

We have shown that every state $\omega \in E_{L(H)}$ can be represented by a linear, positive, norm-continuous functional $\hat{\omega}$ on $C_b(H \setminus \{0\})$ and therefore represented by a normalized finitely additive set function on $H \setminus \{0\}$. This set function is σ-additive, hence a measure, if and only if ω is normal.

Here it becomes obvious that the analogy between quantum mechanics and classical statistical mechanics as suggested by Eq. (2) is not sufficient. In quantum mechanics there exist, contrary to the classical case, states which cannot be represented by σ-additive measures on the state space.

Let us conclude with some remarks. First of all, we want to state that the representation given by Eq. (2) is not appropriate for the incorporation of singular states as $f_A \in C(H)$, but $C(H) \not\subset L^1(H, \mathscr{B}(H), \mu_W)$ for a given W.

As we have seen, singular states can be represented by finitely additive measures. This may allow for an application of the theory of weak distribution (cf., e.g., Ref. 10) to express the expectation value in this case. It remains an open question whether the formulation of singular states for unbounded operators may be incorporated into the present formalism.

Equations (4) and (17) already show the structural similarity of σ-continuous functionals and normal states. The property of σ-additivity in generalized (quantum) probability can now be identified with classical σ-additivity. Whereas the theorem of Gleason asserts that a generalized σ-additive probability measure can be represented by a statistical operator, we have shown that it can be represented by a σ-additive classical probability measure. On the other hand, singular states may be identified with non σ-additive generalized probability measures on Lat(H) or finitely additive measures on $H \setminus \{0\}$.

ACKNOWLEDGMENTS

We would like to thank Professor J. Kamphusmann and Professor D. Plachky for stimulating discussions and valuable suggestions.

[1] A. Bach, J. Math. Phys. **21**, 789 (1980).
[2] A. Bach and T. Wenning, J. Math. Phys. **23**, 1078 (1982).
[3] J. F. Cyranski, J. Math. Phys. **23**, 1074 (1982).
[4] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Springer, New York, 1979).
[5] V. S. Varadarajan, *The Geometry of Quantum Theory* (Van Nostrand, Toronto, 1968).
[6] H. Schäfer, *Topological Vectorspaces* (Springer, New York, 1971).
[7] H. Bauer, *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie* (de Gruyter, Berlin, 1978).
[8] H. Bauer, "Eine Rieszsche Bandzerlegung im Raum der Bewertungen," Sitzungsber. Math. Naturwiss. Kl. Bayer. Akad. Wiss. Muenchen, 89 (1954).
[9] H. Bauer, "Zur Theorie des Riemann Integrals in lokal kompakten Räumen," Sitzungsber. Math. Naturwiss. Kl. Bayer. Akad. Wiss. Muenchen, 187 (1955).
[10] L. Gross, Trans. Amer. Math. Soc. **105**, 371 (1962).

Nonlinear time-dependent anharmonic oscillator: Asymptotic behavior and connected invariants

D. Besnard, J. R. Burgan, and A. Munier
Commissariat á l'Energie Atomique Limeil, B. P. 27, 94100 Villeneuve St Georges, France
M. R. Feix and E. Fijalkow
Physique Théorique et Modélisation, CNRS, 45045 Orléans, Cédex, France

(Received 10 August 1982; accepted for publication 29 October 1982)

The motion of a particle in a potential decreasing with time as $|X|^n$ is considered. Different time and space rescaling are considered in order to obtain the asymptotic solutions. The validity of adiabatic invariants is discussed. The classical critical case corresponds to the obtainment of self-similar solutions for the quantum problem.

PACS numbers: 03.65.Ge

I. INTRODUCTION

In a preceding paper[1] we have shown how the concept of quasi-invariance allows the study of the asymptotic solutions of equations of the type

$$\frac{d^2x}{dt^2} + A(1+\Omega t)^{-\mu}x + B(1+\Omega t)^{-\nu}x^3 = 0.$$

We show here how this technique may be applied to equations describing the motion of a particle in a potential of the form

$$\phi = K|x|^n(1+\Omega t)^{-p}, \tag{1}$$

where K, Ω, n, and p are four real positive numbers. Moreover, in this paper we study the case $\Omega \to 0$ which corresponds to an adiabatic motion with an assotiated adiabatic invariant. We show that if $p < p_c = n/2 + 1$ this adiabatic invariant is valid for all times. On the other hand for a finite value of Ω we will be able to define an asymptotic invariant, i.e., a quantity independent of time for t large enough. Finally we will comment on the quantum problem and show how this critical case $p_c = n/2 + 1$ can be interpreted through the self-similar solutions of the corresponding Schrödinger equation.

II. GENERALIZED CANONICAL TRANSFORM

Consider the equation of motion

$$\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} = \Gamma \tag{2}$$

and the following transformation (see also Ref. 2):

$$x = \xi C(t), \quad d\theta = dt/A^2, \tag{3}$$

where $C(t)$ and $A(t)$ are two arbitrary functions of time, always positive. We call ξ the new coordinate, θ the new time, and $\eta = d\xi/d\theta$ the new velocity. We have

$$v = \frac{dx}{dt} = \eta\frac{C}{A^2} + \xi\frac{dC}{dt}. \tag{4}$$

Straightforward algebra transforms Eq. (2) into the new equation

$$\frac{d^2\xi}{d\theta^2} + \left(\beta A^2 + 2\frac{A}{C}(A\dot{C} - C\dot{A})\right) \times\frac{d\xi}{d\theta} + \frac{A^4}{C}\left(\frac{d^2C}{dt^2} + \beta\frac{dC}{dt}\right)\xi = \frac{A^4}{C}\Gamma. \tag{5}$$

It is easy to show that transformation (3) forms a continuous Lie group. Two successive applications of transformations of the type (3) defined, respectively, by $C_1(t)$, $A_1(t)$ and $C_2(t)$, $A_2(t)$ lead to a third transformation τ_3 such that

$$C_3 = C_2C_1 \tag{6}$$

$$A_3 = A_2A_1. \tag{7}$$

The proof is straightforward. Equations (6) and (7) are direct consequences of Eq. (3). We must also check that the velocity satisfies relation (4) and that the substitution leads to the same terms as the two successive applications for the drag term [second term of the lhs of Eq. (5)], the transformation field [third term of the lhs of Eq. (5)] and the rescaled physical field [rhs of Eq. (5)].

Equation (5) can now be used for different purposes. A and C may be selected in such a way that a problem with friction ($B \neq 0$) is transformed into a problem without friction. On the other hand, we can prefer a problem exhibiting a friction term and a potential energy independent of time to a frictionless problem where the Hamiltonian is time dependent. This last point of view is adopted here to solve

$$\frac{d^2x}{dt^2} + Kn(\operatorname{sgn} x)|x|^{n-1}(1+\Omega t)^{-p} = 0. \tag{8}$$

Taking Eq. (3) into account Eq. (8) becomes

$$\frac{d^2\xi}{d\theta^2} + 2\frac{A}{C}\left(A\frac{DC}{dt} - C\frac{dA}{dt}\right)\frac{d\xi}{d\theta} + \frac{A^4}{C}\frac{d^2C}{dt^2}\xi + A^4Kn(\operatorname{sgn}\xi)C^{n-2}|\xi|^{n-1}(1+\Omega t)^{-p} = 0. \tag{9}$$

In Eq. (9) we will call "new friction" the coefficient of $d\xi/d\theta$, "transformation field" the third term of the lhs, and "rescaled physical field" the last term with corresponding transformation and rescaled potential. The initial equation (8) has no damping but unless $A = C$ the new equation exhibits a friction term.[3] Select now A and C such that

$$A = (1+\Omega t)^\alpha, \quad C = (1+\Omega t)^\gamma, \tag{10}$$

and impose that the coefficients of the second, third, and fourth term in Eq. (9) are time independent. We must take $\alpha = 1/2$ and $\gamma = (p-2)/(n-2)$ and Eq. (9) is now written

$$\frac{d^2\xi}{d\theta^2} + \frac{2p-n-2}{n-2}\Omega\frac{d\xi}{d\theta} + \frac{(p-2)(p-n)}{(n-2)^2} \times\Omega^2\xi + Kn(\operatorname{sgn}\xi)|\xi|^{n-1} = 0. \tag{11}$$

 0022-2488/83/051123-06$02.50

Equation (11) allows the computation of the asymptotic form of the solution. We consider (Fig. 1) eight cases corresponding, respectively, to eight regions of the n-p space delineated by the straight lines $n=2, p=2, n=p, p=n/2+1$ (see Fig. 2). Moreover on Fig. 1 we give the sign of the coefficient of $d\xi/d\theta$ in (11). $+$ indicates a positive damping, $-$ a negative damping, i.e., a spontaneous increase of the velocity.

III. DISCUSSION OF THE DIFFERENT CASES

(a) In region I the negative damping drags the particle towards high-energy regions where the potential is dominated by the transformation potential. Consequently the asymptotic motion corresponds to that of a free particle.

(b) In region II, although the transformation potential is now repulsive, the negative damping imposes again an asymptotic free particle motion.

(c) In region III the damping becomes positive and two different classes of particles must be considered.

Particles of the first class whose initial energy is large enough to pass the total potential barrier will experience asymptotically the transformation field alone, i.e., their asymptotic motion corresponds again to that of a free particle. Particles of the second class with small initial energy will be trapped, damped, and will fall towards the origin. To obtain more useful information on their behavior, go back to Eq. (9) and select now A and C such that $A=C=(1+\Omega t)^{(p/(n+2)}$. We have

$$\frac{d^2\xi}{d\theta^2}+\frac{p}{n+2}\left(\frac{p}{n+2}-1\right)\times\Omega^2\frac{1}{1+[(n+2-2p)/(n+2)]\Omega\theta^2}\xi+Kn(\operatorname{sgn}\xi)|\xi|^{n-1}=0, \tag{12}$$

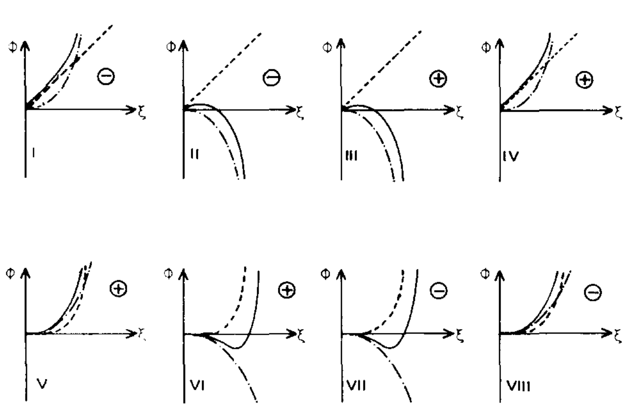

FIG. 1. Potential profiles for the eight regions. ------ rescaled physical, ------ transformation, ——— total. Above the region's number the sign of the friction term is indicated.

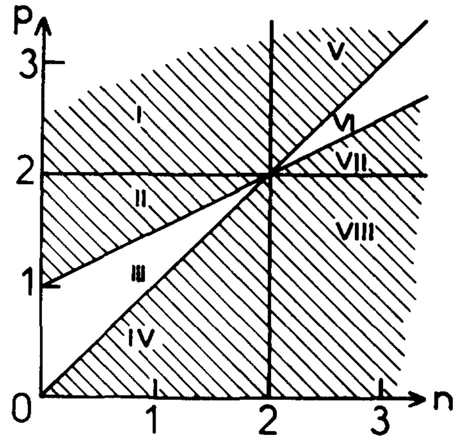

FIG. 2. The eight regions and the four zones of the n-p parameter space.

the new time θ varying now from 0 to ∞. The transformation field in Eq. (12) tends asymptotically to zero and the final motion is periodic. These particles are consequently trapped in the $K|\xi|^n$ potential and the amplitude of their oscillation in x space increases ultimately as $(1+\Omega t)^{p/(n+2)}$.

(d) In region IV the transformation potential is attractive; no particle can escape and the asymptotic motion is the same as the motion of the trapped particle of region III.

(e) In region V the damping is positive, bringing the particles towards the $\xi\to 0$ region where the dominating field is the transformation field. Consequently the asymptotic motion corresponds to that of a free particle.

(f) In region VI the damping being still positive, the particle is driven to the bottom of the potential which is now at a finite distance $\pm\xi_l$. The asymptotic motion is consequently

$$x=\pm\xi_l(1+\Omega t)^{(p-2)/(n-2)}. \tag{13}$$

(g) In region VII the friction is negative again and the dominant field is the rescaled physical field. To gain more information we select A and C as in (c) to obtain Eq. (12). Since $n+2>2p$, the new time θ goes from 0 to ∞ and in this $(\xi,0)$ space the asymptotic motion is a nonlinear periodic motion with constant amplitude. In x space the amplitude increases as $(1+\Omega t)^{p/(n+2)}$.

(h) The same asymptotic motion is obtained in region VIII. Finally the eight regions can be grouped into four zones.

—Zone I, union of regions I, II, and V where the asymptotic motion corresponds to that of a free particle (uniform motion with uniform velocity).

—Zone II (region III) where the asymptotic motion corresponds either to that of a free particle or to a nonlinear oscillation of amplitudes growing as $(1+\Omega t)^{p/(n+2)}$.

—Zone III, union of regions IV, VII, and VIII where the asymptotic motion is a nonlinear oscillation growing again as $(1+\Omega t)^{p/(n+2)}$.

—Zone IV, region VI, where the asymptotic motion is given by Eq. (13).

It is worth noticing that the linear oscillator case ($n=2$) is a degenerate case for which zones II and IV disappear. It should also be pointed out that the limiting case $p = n/2 + 1$ corresponds to a damping equal to zero with Eq (11) written as

$$\frac{d^2\xi}{d\theta^2} - \frac{\Omega^2}{4}\xi + Kn(\operatorname{sgn}\xi)|\xi|^{n-1} = 0. \tag{14}$$

Equation (14) can be integrated taking into account the fact that the new energy is a constant

$$\frac{1}{2}\left(\frac{d\xi}{d\theta}\right)^2 - \frac{\Omega^2}{8}\xi^2 + K|\xi|^n = \epsilon. \tag{15}$$

This limiting case is interesting since it shows clearly how the method can be linked to the stretching (or dilation) group technique.[2] Rewriting the initial equation (8) as

$$\frac{d^2x}{d\tau^2} + Kn(\operatorname{sgn}x)|x|^{n-1}\left(\frac{\tau}{T}\right)^{-p} = 0 \tag{16}$$

with a shifted time τ such that $\tau = t + T = t + \Omega^{-1}$, we consider the stretching transformations

$$\tau = a^\alpha\bar{\tau}, \quad x = a^\beta\bar{x}$$

which leaves Eq. (16) invariant if we take $\beta/\alpha = (2-p)/(2-n)$. Now the existence of such a transformation allows the reduction of the second-order differential equation (16) to an equation of the first order (see Ref. 4 for details) taking as new variable ξ and new function $D(\xi)$:

$$\xi = x\left(\frac{\tau}{T}\right)^{-(2-p)/(2-n)},$$

$$D(\xi) = \frac{dx}{dt}\left(\frac{\tau}{T}\right)^{-(n-p)/(2-n)}. \tag{17}$$

Notice that ξ has *a priori* nothing to do with the one introduced in Eqs. (3) and (10). The new equation is obtained by computing $d\xi/d\tau$ and $dD/d\tau$ [where we replace $d^2x/d\tau^2$ by its value given by (16)]. Forming $dD/d\xi$, we check that τ does not appear any more in the final first-order equation

$$D\frac{dD}{d\xi} + \frac{1}{T}D\frac{n-p}{2-n} - \frac{2-p}{2-n}\xi\frac{dD}{d\xi} + Kn(\operatorname{sgn}\xi)|\xi|^{n-1} = 0. \tag{18}$$

A further integration can be obtained if the second term of the lhs of Eq. (18) is a total differential, i.e., if $p = n/2 + 1$. For this particular value of p we have a time invariant:

$$\tfrac{1}{2}D^2 + K|\xi|^n - (1/2T)D\xi = \text{const}. \tag{19}$$

Introducing in Eq. (19)

$$\eta = D - \xi/2T, \tag{20}$$

we rewrite Eq. (19) as

$$\tfrac{1}{2}\eta^2 - (\Omega^2/8)\xi^2 + K|\xi|^n = \text{const}. \tag{21}$$

It is easy to check that for $p = n/2 + 1$ and ξ and η as given by relations (17) and (20), we can also write

$$\xi = x\left(\frac{\tau}{T}\right)^{-1/2} = x(1+\Omega t)^{-1/2},$$

$$\eta = (1+\Omega t)^{1/2}\frac{dx}{dt} - \frac{\Omega x}{2}(1+\Omega t)^{-1/2}. \tag{22}$$

For this value of p, parameters α and γ of Eq. (10) take the value 1/2. Introducing all this in Eq. (3) we check that ξ and $\eta = d\xi/d\theta$ as given by the quasi-invariant theory are identical with those defined in Eq. (22) and that the constant [in Eq. (21)] can be identified with the new energy of Eq. (15).

IV. ADIABATIC ASYMPTOTIC AND EXACT INVARIANTS

A. Case $p < p_c$ adiabatic invariant

Consider Eq. (12) obtained by taking $A = C$ (this problem is consequently without damping) and $C = (1+\Omega t)^{p/(n+2)}$ leading to time-invariant rescaled physical field. The transformation field goes to zero provided $p < p_c = n/2 + 1$. If $\Omega \to 0$ the transformation field not only goes to zero as $\theta \to \infty$ but is of order Ω^2. The unperturbed motion ξ_0 obtained from setting $\Omega = 0$ is a periodic motion corresponding to the potential $K|\xi|^n$. In this unperturbed motion the (new) energy is conserved. The energy variation connected to the presence of the time-dependent transformation field is

$$\Delta\epsilon = \frac{p(n+2-p)}{(n+2)^2}\Omega^2 \times \int_o^\infty \frac{\xi_0}{(1+[(n+2-2p)/(n+2)]\Omega\theta)^2}\frac{d\xi_0}{d\theta}d\theta. \tag{23}$$

It can be shown that in the limit $\Omega \to 0$ the integral in Eq. (23) goes to a finite limit and $\Delta\epsilon$ is finally a quantity of second order in Ω. The new energy is consequently an adiabatic invariant to zero and first order in Ω.

Notice that

$$K|\xi|^n = k|x|^n(1+\Omega t)^{-p}(1+\Omega t)^{2p/(n+2)}. \tag{24}$$

By Eq. (4) we deduce

$$V^2 = \frac{n^2}{C^2} + 2\xi\eta\frac{\frac{dC}{dt}}{C} + \xi^2\left(\frac{dC}{dt}\right)^2. \tag{25}$$

The first term in the rhs of Eq. (23) is of order Ω^0, the second of order Ω, and the third of order Ω^2. We can write, neglecting terms in Ω and Ω^2.

$$\tfrac{1}{2}V^2 + K|x|^n(1+\Omega t)^{-p} = \tfrac{1}{2}\eta^2 + K|\xi|^n(1+\Omega t)^{-2p/(n+2)}. \tag{26}$$

Finally we define the length of the "expanding box" for a time varying potential through the relation

$$\phi(x,t) = |x|^n(1+\Omega t)^{-p} = B^{-2}\hat{\phi}(x/B), \tag{27}$$

with $B = (1+\Omega t)^\gamma = C$. Identifying we find $\gamma = p/(n+2)$. Later we will justify relation (27) but the result is quite reasonable since we know that the rescaled quantity $\xi = x/C$ is periodic and that the particle "bounces" on the time-independent potential in ξ space. Now Eq. (26) can be written

$$(\tfrac{1}{2}V^2 + K|x|^n(1+\Omega t)^{-p})l^2 = A, \tag{28}$$

where $l = C(t)\,l_0$ is the "length of the box." Equation (28) is the well-known "adiabatic relation" between the energy and the volume (i.e., the length for a one-dimensional system) with $\gamma = 3$. Two final remarks are in order.

Remarks: 1. The "new energy ϵ" is a better invariant than El^2 since ϵ is invariant to zero and the first-order term in Ω, while in the computation of El^2 we have dropped first-order term [the second term in the rhs of (26)].

2. This invariant never breaks down.

B. Case $p<p_c$ asymptotic invariant

If Ω is no longer small we can call ϵ an "asymptotic invariant," i.e., a quantity which remains constant for a long time. The only difference with the case $\Omega \rightarrow 0$ is that the work done by the transformation field vanishing as time goes on is not a small quantity.

Moreover, Eq. (28) remains asymptotically true. It was obtained by transforming ϵ into the energy in the initial space through relations (24) and (25) and by neglecting in Eq. (25) the second and third term of the rhs. This remains true for large times for all values of Ω, a result easily checked since

$$C^{-2} \quad \text{varies as } (1+\Omega t)^{-2p/(n+2)},$$

$$\frac{1}{C}\frac{dC}{dt} \quad \text{varies as } (1+\Omega t)^{-1},$$

$$\left(\frac{dC}{dt}\right)^2 \quad \text{varies as } (1+\Omega t)^{2p/(n+2)}.$$

To sum up these results: for $p<p_c$ we can introduce either an adiabatic invariant (if $\Omega \rightarrow 0$) or at least an asymptotic invariant which in both cases is the new energy.

C. $p=p_c$

For $p=n/2+1$, Eq. (12) and (11) are identical. They have no friction term, a time-independent potential, and the new energy ϵ is an exact invariant. The length of the box varies as $(1+\Omega t)^{1/2}$, the energy decreases as t^{-1}. The possibility of finding invariants for similar equations has been also considered in Ref. 5.

D. $n>p>p_c$

This case correponds to region VI where $|\xi|$ goes to a limiting value $|\xi_l|$. From Eq. (13) we can write asymptotically

$$\frac{1}{2}V^2 + K|x|^n(1+t)^{-p}$$
$$\approx\left(\frac{1}{2}\xi_l^{\,2}\left(\frac{p-2}{n-2}\right)^2\Omega^2 + K|\xi_l|^n\right)(1+\Omega t)^{2(p-n)/(n-2)}. \tag{29}$$

Equation (29) shows three interesting points.

—In region VI the asymptotic motion corresponds to a fixed ratio between potential and kinetic energy.

—for $p=p_c$ this relation agrees with relation (28) and leads to a total energy E varying as $(1+\Omega t)^{-1}$.

—for $p=n$ the energy is a constant. Indeed for $p>n$ the asymptotic motion corresponds to that of a free particle.

V. CASE OF AN EXPANDING BOX

An interesting limiting case is provided by a box made of a potential well (zero potential inside, infinite potential outside the box). The abscissa of the two walls are, respectively, $\pm l_0(1+\Omega t)^\alpha$. (For a solution of the case $\alpha=1$ see Ref. 6). The particle experiences a perfect elastic collision on the moving walls and a free motion otherwise. It can be easily shown that such a situation corresponds to a potential as given by relation (1) with p and $n\rightarrow\infty$ and $p=n\alpha$. In this limit Eq. (11) becomes

$$\frac{d^2\xi}{d\theta^2} + (2\alpha-1)\Omega\frac{d\xi}{d\theta} + \alpha(\alpha-1)\Omega^2\xi = 0 \tag{30}$$

and is valid inside the "new motionless box" extending from $-l_0$ to l_0. The scaling factors are $C(t)=(1+\Omega t)^\alpha$ and $A(t)=(1+\Omega t)^{1/2}$. Four cases must be investigated.

A. $\alpha<1/2$

This corresponds to region VII. In that case it is simpler to take $A(t)=C(t)=(1+\Omega t)^\alpha$ and write the limit form of Eq. (12),

$$\frac{d^2\xi}{d\theta^2} + \alpha(\alpha-1)\Omega^2\frac{\xi}{(1+(1-2\alpha)\Omega\theta)^2} = 0. \tag{31}$$

In the "new box" the asymptotic motion is an endless bouncing on the two walls. If $\Omega\rightarrow 0$ there exists an adiabatic invariant which never breaks down and if Ω is finite an asymptotic invariant.

B. $\alpha=1/2$

Both Equations (30) and (31) show the existence of an exact invariant, the new energy is

$$\epsilon = \frac{1}{2}\left(\frac{d\xi}{d\theta}\right)^2 - \frac{\Omega^2}{8}\xi^2. \tag{32}$$

C. $1>\alpha>1/2$

The description is given by Eq. (30). In the "new box" the particle sticks asymptotically to one of the wall. In the original space this means that eventually after some oscillations, the particle bounces on one of the walls without changing the sign of its velocity. Since the wall velocities decrease with time a new bouncing takes place, slowing down the particle but never changing the sign of the velocity, etc.

D. $\alpha>1$

In the "new box" as given again by Eq. (30) the particle falls down to $\xi=0$ corresponding to a free motion in the original space. Indeed after a transient state the wall velocities increase without limit and the particle does not collide any more.

VI. QUANTUM MECHANIC INTERPRETATION

We have seen that the case $p=n/2+1$ corresponds to a critical regime where we have an exact invariant. In fact it is worth noticing that this case corresponds to the possibility of obtaining a solution of the Schrödinger equation via stretching groups. This last equation reads

$$i\hbar\frac{\partial\psi}{\partial\tau} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + \phi(x,\tau)\psi. \tag{33}$$

Notice that when we use stretching group methods we come back to the shifted time $\tau=t+T=t+\Omega^{-1}$ as in Sec. III(h).

A nontrivial stretching group leaving Eq. (33) invariant,

$$\tau = a^\alpha \bar{\tau}, \quad x = a^\beta \bar{x}, \quad \phi = a^\gamma \bar{\phi}, \quad \text{and} \quad \psi = a^\delta \bar{\psi}, \tag{34}$$

is given by

$$\beta = \alpha/2, \quad \gamma = -\alpha, \quad \delta = -\beta/2 = -\alpha/4, \tag{35}$$

where the last relation comes from the normalization condition

$$\int \psi\psi^* \, dx = 1.$$

Equation (35) indicates that if we introduce the new variable and functions

$$\hat{x} = \frac{x}{(\tau/T)^{1/2}}, \hat{\phi} = \left(\frac{\tau}{T}\right)\Phi, \hat{\psi} = \psi\left(\frac{\tau}{T}\right)^{1/4}, \tag{36}$$

the Schrödinger equation becomes

$$\frac{i\hbar}{T}\left(-\frac{1}{4}\hat{\psi} - \frac{1}{2}\hat{x}\frac{d\hat{\psi}}{d\hat{x}}\right) = -\frac{\hbar^2}{2m}\frac{d^2\hat{\psi}}{d\hat{x}^2} + \hat{\phi}\hat{\psi} \tag{37}$$

($\hat{\psi}$ and $\hat{\phi}$ being functions of $\hat{x}$ only). The first two equations (36) are equivalant to $\xi = \hat{x} = x/C(\tau)$ and to Eq. (27) where we take for $C(\tau)$ the limiting case $C(\tau) = (\tau/T)^{1/2}$. In this case, ψ keeps its initial shape which implies variation of both the kinetic potential energy inversely proportional to the square of the length of the system. It is quite remarkable that the invariance of the quantum equation leads to $\phi = (\tau/T)^{-1}\phi x/(\tau/T)^{1/2}$, i.e., the critical case leads to a definition in agreement with Eq. (27). This last equation was needed to write down the well-known relation (28). This deep and subtle relationship between quantum mechanics and the theory of invariants was noticed at the very beginning of quantum mechanics and still remains somewhat mysterious.

Finally for a class of potentials of the form

$$\phi(x,t) = C^{-2}\hat{\phi}(x/C(t)), \tag{38}$$

we have shown[6] that by taking $x = \xi C(t)$ and $d\theta = dt/C^2$ [i.e., $A = C$ in Eq. (3)], the new Schrödinger equation can be written in the new space time as

$$i\hbar\frac{\partial\mu}{\partial\theta} = -\frac{\hbar^2}{2m}\frac{\partial^2\mu}{\partial\xi^2} + \left(\hat{\phi}(\xi) + \frac{1}{2}C^3\frac{d^2C}{dt^2}\xi^2\right)\mu, \tag{39}$$

where

$$\psi = C^{-1/2}\exp\left(i\frac{m}{\hbar}\frac{1}{2C}\frac{dC}{dt}x^2\right)\mu(\xi,\theta) \tag{40}$$

and with $\int\psi\psi^*\,dx = \int\mu\mu^*\,d\xi = 1$. Some remarks on how these equations are obtained are given in the Appendix.

Now for a potential given by Eq. (1) we substitute $C = (1+\Omega t)^{p/(n+2)}$ in Eq. (38). For $p < p_c$ the potential as in the classical case can be written as the sum of the rescaled physical potential $K|\xi|^n$ and the transformation potential

$$\frac{1}{2}\frac{p}{n+2}\left(\frac{p}{n+2} - 1\right)\Omega^2\xi^2 \times \frac{1}{(1+[(n+2-2p)/(n+2)]\Omega\theta)^2}. \tag{41}$$

Equations (31) and (41) are the quantum mechanics equivalent of relation (12), for $p = p_c = n/2 + 1$, $C = (1+\Omega t)^{1/2}$, and $\psi = (1+\Omega t)^{-1/4}\exp i(m/\hbar)(\Omega/4)\xi^2\mu$. A little algebra shows that the time-independent solution of Eq. (39) ($\partial/\partial\theta = 0$) is identical to the self-similar solution obtained from Eq. (37), since

$$\psi = \hat{\psi}(1+\Omega t)^{-1/4} = (1+\Omega t)^{-1/4}\exp(i(m/\hbar)\Omega/4)\xi^2)\mu(\xi).$$

VII. CONCLUSION

In this paper we have used the quasi-invariant technique to obtain the asymptotic motions of a particle in a potential of the form $K|x|^n(1+\Omega t)^{-p}$. Four zones are delineated according to the values of p and n, a critical value being $p_c = n/2 + 1$. From these solutions we studied the possible "asymptotic" (valid only for large time), "adiabatic" (connected to a smallness of Ω), or "exact" invariants.

For $p < p_c$ we have an asymptotic (for finite Ω) and an adiabatic (for $\Omega \to 0$) invariant. This last invariant never breaks down and correponds to the relation $El^2 = \text{const}$, where E is the total energy and l is a natural length of the "box."

For $p < p_c < n$ we have an "asymptotic invariant" but the corresponding relation is $E(1+\Omega t)^{2(n-p)/(n-2)} = \text{const}$. The adiabatic invariant initially present if Ω is small enough breaks down for large times.

For $p > n$ the particle, asymptotically, is not submitted on the field.

Finally for $p = p_c$ we have an exact invariant. It is interesting to notice that the critical case corresponds to a length of the box increasing as $(1+\Omega t)^{1/2}$ and to the existence of a self-similar solution for the quantum problem. This length increasing as $(1+\Omega t)^{1/2}$ corresponds to the quantum invariant $l^2/t = \text{const}$, a consequence of the Schrödinger equation.

APPENDIX: QUASI-INVARIANT TRANSFORMATIONS IN QUANTUM MECHANICS[7]

For classical motion of a particle we have considered the following transformation

$$\left.\begin{matrix} x \\ v \\ t \end{matrix}\right\} \rightrightarrows \left.\begin{matrix} \xi \\ \eta \\ \theta \end{matrix}\right\} \text{ with } \begin{matrix} x = C(t)\xi + D(t), \\ v = (C/A^2)\eta + \dot{C}\xi + \dot{D}, \\ dt = A^2 d\theta. \end{matrix}$$

Notice that we have introduced the shift $D(t)$ where D is an arbitrary time function.

We want to determine if there exists a corresponding group transformation in quantum mechanics—compatible with the correspondence principle. More precisely we consider a wave function $\psi(x,t)$. Is it possible to transform it into a function $\bar{\psi}(\xi,\theta)$ with

$$\psi(x,t) = \phi(x,t)\bar{\psi}(\xi,\theta), \tag{A1}$$

$$dt = A^2(t)d\theta, \tag{A2}$$

$$x = C_1(t)\xi + D_1(t), \tag{A3}$$

such that

$$\int \psi(x,t)\psi^*(x,t)dx = \int \bar{\psi}(\xi)\bar{\psi}^*(\xi)d\xi = 1, \tag{A4}$$

$$\langle x\rangle = \int x\psi\psi^*\,dx = C(t)\langle\xi\rangle + D(t), \tag{A5}$$

$$\langle \xi \rangle = \int \xi \bar{\psi} \bar{\psi}^* \, d\xi, \tag{A6}$$

$$\langle v \rangle = -\frac{i\hbar}{m} \int \psi^* \frac{\partial \psi}{\partial x} dx = \frac{C}{A^2} \langle \eta \rangle + C \langle \xi \rangle + \dot{D}, \tag{A7}$$

$$\langle \eta \rangle = -\frac{i\hbar}{m} \int \bar{\psi}^* \frac{\partial \bar{\psi}}{\partial \xi} d\xi . \tag{A8}$$

we can suspect that $C = C_1$ and $D = D_1$ if such a transformation satisfying the correspondence principle exists. But it is more rigorous to verify that it is indeed needed. Equations (A1) and (A3) are introduced in Eq. (A5) and we write that the equality must be verified $\forall \psi$. We get

$$C(t) = \phi\phi^* C_1^2, \quad D = D_1. \tag{A9}$$

Introducing now (A1) and (A3) in (A4) we obtain

$$\phi\phi^* C_1 = 1. \tag{A10}$$

From Eqs. (A9) and (A10) we deduce $C_1 = C$, $D_1 = D$. Now from Eq. (A10) we deduce

$$\phi = C^{-1/2} \exp iB(\xi,\theta), \tag{A11}$$

where B is real. Now we introduce (A11) and (A1) into (A7) and obtain

$$-\frac{i\hbar}{m} C^{-1} \int \bar{\psi}^* \frac{\partial \bar{\psi}}{\partial \xi} d\xi + \frac{\hbar}{m} C^{-1} \int \frac{\partial B}{\partial \xi} \bar{\psi}^* \bar{\psi} \, d\xi$$
$$= \frac{C}{A^2} \langle \eta \rangle + \dot{C} \int \xi \bar{\psi}^* \bar{\psi} \, d\xi + \dot{D}.$$

This is an identity for all possible ψ. Consequently

$$C^{-1} = \frac{C}{A^2} \text{ and } \frac{\hbar}{m} \frac{1}{C} \frac{\partial B}{\partial \xi} = \dot{C}\xi + \dot{D}. \tag{A12}$$

From (A12) we deduce $C = A$ and

$$B = (m/2\hbar) C\dot{C}\xi^2 + mC\dot{D}\xi + F(t). \tag{A13}$$

We can state the following proposition.

The most general quasi-invariant transformation in quantum mechanics for which the correspondence principle leads to the transformation indicated at the beginning of this appendix is given by

$$x = \xi C(t) + D(t),$$
$$dt = C^2(t) d\theta, \tag{A14}$$
$$\psi = C^{-1/2} \exp i \left(\frac{m}{2\hbar} C\dot{C}\xi^2 + mC\dot{D}\xi + F \right) \bar{\psi}(\xi,\theta).$$

We see that in the quantum case we must restrict ourselves to the case $C = A$ which corresponds in classical mechanics to the conservation of the phase-space volume element. This condition ensures that the x-p commutator is invariant with

$$[x,mv] = [\xi,m\eta] = i\hbar.$$

We must stick to the Hamiltonian formalism and cannot introduce a friction term as in the classical case. We have demonstrated elsewhere[5] that indeed $\bar{\psi}$ does satisfy a Schrödinger equation with a new potential given by

$$v = C^2 V + (m/2) C^3 \ddot{C} \xi^2 + mC^3 \ddot{D} \xi. \tag{A15}$$

Other terms can be introduced in (A15) but being space independent they do not play any role and introduce simply a space-independent (and consequently unobservable) phase shift in $\bar{\psi}$.

[1]M. P. Moraux, E. Fijalkow, and M. R. Feix, J. Phys. A **14**, 1611 (1981).
[2]H. R. Lewis, Phys Rev. Lett. **18**, 510 (1967).
[3]Strictly speaking $A = kC$ produces the same effect but there is no loss of generality in taking $k = 1$.
[4]G. W. Bluman and J. D. Cole, *Similarity Methods for Differential Equations* (Springer-Verlag, Berlin, 1974).
[5]P. G. L. Leach, J. Math. Phys. **22**, 465 (1981).
[6]J. R. Burgan, M. R. Feix, E. Fijalkow, J. Gutierrez, and A. Munier "Utilisation des groupes de transformation pour la résolution des équations aux dérivées partielles, in *Applied Inverse Problem*, edited by P. C. Sabatier, *Lecture Notes in Physics*, Vol. 85 (Springer-Verlag, Berlin, 1978).
[7]A. Munier, J. R. Burgan, M. Feix, and E. Fijalkow, J. Math. Phys. **2**, 1219–1223 (1981).

Complex-potential description of the damped harmonic oscillator

Pavel Exner[a)]
Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, 141980 Dubna, USSR

(Received 29 October 1981; accepted for publication 23 April 1982)

The multidimensional damped harmonic oscillator is treated by means of a non-self-adjoint Hamiltonian with complex potential. The propagator referring to the evolution semigroup is evaluated from the Lie–Trotter formula. The one-dimensional case is discussed in detail with the following results: (a) the nondamped limit gives the correct propagator including the Maslov phase factor, (b) for some initial conditions, the classical limit of the solution can differ from the behavior of the classical damped oscillator, the difference being negligible in the case of weak damping, and (c) the point spectrum of the considered pseudo-Hamiltonian is found.

PACS numbers: 03.65.Ge

1. INTRODUCTION

There is a large number of problems ranging from elementary particles to statistical physics, in which the considered systems are dissipative (cf., e.g., Refs. 1–5). The dynamics in such cases can be rarely described fully, including interaction with the heat reservoir (decay products, compound nucleus channel, etc.); usually one is forced to express influence of these degrees of freedom by means of phenomenological Lagrangians or Hamiltonians. They can be constructed in different ways: as time-dependent, nonlinear (e.g., Refs. 3 and 6) or non-self-adjoint; in particular, Hamiltonians with complex potentials are popular in practical calculations in nuclear physics.

Recently we have shown how to incorporate description of a dissipative system S via a phenomenological non-self-adjoint Hamiltonian H into the standard quantum theoretical framework.[7] If H is closed and iH generates a continuous contractive semigroup (such operators we called *pseudo-Hamiltonians*), then by minimal unitary dilation of this semigroup we obtain objects which are naturally interpretable as the state Hilbert space of a larger isolated system Σ containing S and the unitary evolution group of Σ. The well-known difficulty with the spectrum of the corresponding total Hamiltonian (see Refs. 4,8 and references therein) means that the semigroup evolution of S is necessarily approximative[7]; however, this approximation is good enough for almost all applications.[9,10]

In the present paper, we apply the pseudo-Hamiltonian approach to the case of multidimensional damped harmonic oscillator. There are, of course, many possibilities how to choose H; some complex structures have been already studied.[11] We shall use the most natural choice $H = -\frac{1}{2}\Delta + x\cdot(A - iW)x$, where A, W are strictly positive matrices (strict positivity of A is assumed for convenience, in fact, the proofs can be carried out for positive A as well). We assume neither a time-dependent frequency,[6] nor any driving force, stochastic or not.[6,12] On the other hand, we consider oscillators of an arbitrary dimension d; the generalization to the $d > 1$ case is nontrivial, because A, W need not be simultaneously diagonalizable. This multidimensionality together with the special choice of H could be of some interest for the old problem of constructing a field theory with basic quanta metastable.

One has to check first that our H is a pseudo-Hamiltonian in the sense of the above definition. If the damping part could be regarded as a perturbation to the undamped oscillator, the Kato–Rellich-type lemma would be applicable. In general, however, it is not so. Thus we use a trick based on a successive application of the lemma; this trick might appear to be useful for some self-adjointness proofs too.

The main result of the paper is an *explicit integral-operator expression of the evolution semigroup* corresponding to H. After some preliminaries, we prove it in Secs. 5 and 6. The method is based on Feynman-type path integrals in the sense of Nelson, i.e., defined by the Lie–Trotter formula.[13,14] The same result, however, is obtained with some other definitions of the path integral, for instance, the one of Truman[15,16] or that using the "uniform" Trotter formula.[17]

Discussion of the obtained results is limited in the present paper essentially to the one-dimensional case. The first problem concerns the *nondamped limit*: we show that it gives the correct Feynman propagator including the phase factor[18,19]; thus we find in the present case an alternative and very natural way of deriving the Maslov correction. Further, we shall discuss the *classical limit*. Let us notice that comparing to common practice[3,11,12,20,21] we did not obtain our pseudo-Hamiltonian by some kind of quantization of the classical damped oscillator (CDO). According to our opinion, such an approach makes sense only if there is a reasonable similarity between the classical and quantum mechanisms of damping. In general, this is not the case; thus there is no *a priori* reason why the classical limit should reproduce the exact behavior of CDO. We shall illustrate it on an example: for our damped oscillator and special Gaussian wavepackets, the classical limit gives trajectories of CDO but corresponding to changed initial conditions; the difference vanishes in the weak-damping limit. Finally, we shall find the *point spectrum* of H, which is of the form of the undamped-oscillator spectrum rotated around the origin to the lower complex half-plane. The eigenvectors, however, are no longer orthogonal because H is not normal.

2. SOME NOTATION AND CONVENTIONS

$$Q^2 = \sum_{j=1}^{d} Q_j^2,$$

where $(Q_j\psi)(x) = x_j\psi(x)$,

a) On leave of absence from the Nuclear Centre, Charles University, Areál Troja, 18000 Prague—Pelc-Tyrolka, Czechoslovakia.

 0022-2488/83/051129-07$02.50

$$P^2=\sum_{j=1}^d P_j^2=-\Delta,$$

where $P_j=F_d^{-1}Q_jF_d$ and F_d is the d-dimensional Fourier-Plancherel operator,

$$v_1(x)=x\cdot Ax,\quad v_2(x)=x\cdot Wx,$$

where A,W are real positive symmetric $d\times d$ matrices (more exactly, positive symmetric operators on $\mathbb{R}^d$) and

$$v(x)=v_1(x)-iv_2(x)=x\cdot Bx,$$

$$V_i:(V_i\psi)(x)=v_i(x)\psi(x),\quad V=V_1-iV_2,$$

$$H_1=\tfrac12P^2+V_1,$$

$$H_2=H_1\restriction\mathscr{S}(\mathbb{R}^d),\quad H_3=H_2-iV_2=H\restriction\mathscr{S}(\mathbb{R}^d),$$

$$H=H_1-iV_2=\tfrac12P^2+V,$$

$\mathscr{M}(\mathscr{H})$ is the set of all (finite) complex Borel measures on a real Hilbert space $\mathscr{H}$,

$\mathscr{F}(\mathscr{H})$ is the set of functions $f:f(\gamma)=\int_{\mathscr{H}}\exp(i(\gamma,\gamma'))\,d\mu(\gamma')$,

where $\mu\in\mathscr{M}(\mathscr{H})$ and $(.,.)$ is the inner product in $\mathscr{H}$.

In what follows, square roots of complex numbers and matrices will appear frequently. It is useful to make an overal choice of the branch: we prefer to work with $(e^{i\varphi})^{1/2}=\exp(\frac12 i\varphi)$, $0\leqslant\varphi<2\pi$. There is a particular case which should be mentioned: when complex frequencies are considered, it is more natural to have their real parts positive, at least from the viewpoint of the nondamped limit. We shall use therefore $\Omega=-(2B)^{1/2}$ with the square root understood in the above sense.

3. THE PSEUDO-HAMILTONIAN PROPERTY OF H

As mentioned above, through this section we assume the matrices A,W to be strictly positive (as operators on $\mathbb{R}^d$). The eigenvalues of A are α_j, $j=1,...,d$, so $\alpha=\min_{1<j<d}\alpha_j>0$. The inequalities

$$\alpha^2\|Q^2\psi\|^2\leqslant\|V_1\psi\|^2=\sum_{j,k=1}^d\alpha_j\alpha_k\|Q_jQ_k\psi\|^2\leqslant\|A\|^2\|Q^2\psi\|^2$$

show that $D(V_1)=D(Q^2)$, analogously $D(V_2)=D(Q^2)$, i.e.,

$$D(H)=D(H_1)=D(P^2)\cap D(Q^2). \tag{1}$$

Proposition 1: H_1 is self-adjoint.

Proof: We notice first that H_2 is essentially self-adjoint (e.s.a.) due to existence of a complete set of eigenvectors $\subset\mathscr{S}(\mathbb{R}^d)$. Both P^2 and V_1 are self-adjoint and therefore closed so that $H_1\subset\overline{H}_2$. In order to prove the opposite inclusion we shall verify that there is $b>0$ such that

$$\tfrac14\|P^2\psi\|^2+\|V_1\psi\|^2\leqslant\|H_2\psi\|^2+b\|\psi\|^2,\quad\psi\in\mathscr{S}(\mathbb{R}^d). \tag{2}$$

We have $(P_i\psi)(x)=-i\partial\psi(x)/\partial x_i$ for these ψ, i.e.,

$$((P_jQ_k-Q_kP_j)\psi)(x)=-i\delta_{jk}\psi(x)\psi(x),\quad\psi\in\mathscr{S}(\mathbb{R}^d). \tag{3}$$

We choose a basis in $\mathbb{R}^d$ so that A is diagonal. Then

$$(\psi,(P^2V_1+V_1P^2)\psi)\geqslant\sum_{j=1}^d\alpha_j(\psi,(P_j^2Q_j^2+Q_j^2P_j^2)\psi),$$

because $(\psi,P_j^2Q_k^2\psi)\geqslant0$ for $j\neq k$ due to the relations (3), which further imply

$$(\psi,(P^2V_1+V_1P^2)\psi)\geqslant\tfrac12\sum_{j=1}^d\alpha_j\|(P_j^2Q_j^2+Q_j^2P_j^2)\psi\|^2-\tfrac32\|\psi\|^2\mathrm{Tr}\,A.$$

Thus (2) holds if $b\geqslant\frac34\,\mathrm{Tr}\,A$. Assume now $\psi\in D(\overline{H}_2)$. If $\{\psi_n\}$ is a sequence $\subset\mathscr{S}(\mathbb{R}^d)$, $\psi_n\to\psi$, then $\{H_2\psi_n\}$ converges too, i.e., $\|H_2\psi_n-H_2\psi_m\|\to0$ with $n,m\to\infty$. The inequality (2) shows that also $\{P^2\psi_n\}$ and $\{V_1\psi_n\}$ converges, however, both P^2, V_1 are closed and $\mathscr{S}(\mathbb{R}^d)\subset D(P^2)\cap D(V_1)$ so that $\psi\in D(P^2)\cap D(V_1)=D(H_1)$. ■

The pseudo-Hamiltonian property of H will be proved below by successive applications of the following perturbative lemma[22]:

Proposition 2: Let G be a densely defined closable operator on a Hilbert space $\mathscr{H}$ such that $\overline{G}$ is a pseudo-Hamiltonian. Let further C be closed and accretive, $D(C)\supset D(G)$, and assume that there exist non-negative $a<1$, b such that

$$\|C\psi\|^2\leqslant a^2\|G\psi\|^2+b^2\|\psi\|^2,\quad\psi\in D(G). \tag{4}$$

Then $D(\overline{G})\subset D(C)$ and the operator $\overline{G}-iC$ defined on $D(\overline{G})$ is closed and belongs to the class of pseudo-Hamiltonians.

One must exhibit conditions under which (4) is fulfilled in the case under consideration:

Proposition 3: (a) Let $b^2\leqslant\frac12\alpha\|W\|^{-1}$, then there is a positive c such that

$$\|bV_2\psi\|^2\leqslant\tfrac12\|H_1\psi\|^2+c\|\psi\|^2,\quad\psi\in\mathscr{S}(\mathbb{R}^d). \tag{5a}$$

(b) Let $a>0$ and $b^2\leqslant\frac12a^2$, then there is a positive c such that

$$\|bV_2\psi\|^2\leqslant\tfrac12\|(H_1-iaV_2)\psi\|^2+c\|\psi\|^2,\quad\psi\in\mathscr{S}(\mathbb{R}^d). \tag{5b}$$

Proof: We have to find c for which

$$I\equiv(\psi,(\tfrac12(H_1+iaV_2)(H_1-iaV_2)-bV_2^2+c)\psi)$$

is non-negative independently of $\psi\in\mathscr{S}(\mathbb{R}^d)$. We choose again a basis in $\mathbb{R}^d$ so that A is diagonal and denote by W_{jk} the corresponding matrix elements of W. Expressing $(V_2P^2-P^2V_2)\psi$ and $(V_1P^2+P^2V_1)\psi$ from (3) and omitting the positive term $\frac18(\psi,P^4\psi)$, we obtain

$$I\geqslant\Big(\psi,\Big[\tfrac18\sum_{j=1}^d\alpha_j(P_kQ_j+Q_jP_k)^2-\tfrac38\,\mathrm{Tr}\,A+\tfrac12\Big(\sum_{j=1}^d\alpha_jQ_j\Big)^2+(\tfrac12a^2-b^2)\Big(\sum_{j,k=1}^d W_{jk}Q_jQ_k\Big)^2-\frac a4\sum_{j,k=1}^dW_{jk}(P_kQ_j+Q_jP_k)+c\Big]\psi\Big).$$

Assume first $a=0$ and $b^2\leqslant\frac12\alpha\|W\|^{-1}$, then the last inequality yields

$$\geqslant(\psi,[c-\tfrac38\,\mathrm{Tr}\,A+(\tfrac12\alpha-b^2\|W\|)Q^2]\psi)\geqslant(c-\tfrac38\,\mathrm{Tr}\,A)\|\psi\|^2$$

so that (5a) holds if $c\geqslant\frac38\,\mathrm{Tr}\,A$. On the other hand, if $a^2\geqslant2b^2$, then

$$I\geqslant\Big(\psi,\Big[\sum_{j,k=1}^d(8\alpha_j)^{-1}(\alpha_j(P_kQ_j+Q_jP_k)-aW_{jk})^2-\tfrac18a^2\sum_{j,k=1}^d\alpha_j^{-2}W_{jk}^2+c-\tfrac38\,\mathrm{Tr}\,A\Big]\psi\Big),$$

and therefore (5b) holds if $c \geqslant \frac{3}{8}\,\mathrm{Tr}\,A + \frac{1}{8}a^2 \sum_{j,k=1}^{d} \alpha_j^{-2} W_{jk}^2$. ∎

Combining now the above three auxiliary statements, we can prove the main result of this section:

Theorem 1: Let A, W be strictly positive so that (1) holds, then H is closed and belongs to the class of pseudo-Hamiltonians. Moreover, $\mathscr{S}(\mathbb{R}^d)$ is a core for H, i.e., $H = \overline{H}_3$.

Proof: (a) If $\alpha \geqslant 2\|W\|$, then there is $c > 0$ such that (5a) with $b = 1$ holds. The operator V_2 is positive, and therefore accretive, $D(V_2) \supset D(H_1)$, and $H_1 = \overline{H}_2$ is a pseudo-Hamiltonian due to Proposition 1. Applying then Proposition 2 to $G = H_2$, $C = V_2$ we see that for $H = H_1 - iV_2$ the assertion is valid. (b) If $\alpha < 2\|W\|$ we choose k positive, $2\|W\|k^2 \leqslant \alpha$, and n natural so that

$$k(1 + 2^{-1/2})^{n-1} = 1. \qquad (\star)$$

The same argument as above shows that the operator $H_1 - ikV_2$ with the domain $D(H_1)$ is closed and belongs to the pseudo-Hamiltonian class. Moreover, this operator equals $\overline{H_2 - ikV_2}$: obviously $\overline{H_2 - ikV_2} \subset H_1 - ikV_2$; on the other hand, for an arbitrary $\varphi \in D(H_1)$ and a sequence $\{\varphi_n\} \subset \mathscr{S}(\mathbb{R}^d)$, $\varphi_n \to \varphi$, we have $(H_2 - ikV_2)\varphi_n = H_1\varphi_n - ikV_2\varphi_n$ so that $\varphi \in D\,\overline{(H_2 - ikV_2)}$.

(c) The proof is completed by induction: assume that the assertion holds for $H_{1j} = \overline{H}_{2j}$, where $\times H_{sj} = H_s - ik \times (1 + 2^{-1/2})^{j-1} V_2$. The assumptions of Proposition 3(b) are fulfilled for $a = 2^{1/2} b = k(1 + 2^{-1/2})^{j-1}$; thus (5b) together with Proposition 2 imply that the assertion holds for

$$H_{1j} - ik\,2^{-1/2}(1 + 2^{-1/2})^{j-1} = H_{1,j+1}$$

as well. In the same way as above one proves $H_{1,j+1} = \overline{H}_{2,j+1}$. Since the assertion is valid for $H_{11} = \overline{H}_{21}$ due to (b), the same is true for H_{1j} corresponding to any natural j, in particular for $H_{1n} = \overline{H}_{2n}$ which equals $H = \overline{H}_3$ in view of ($\star$). ∎

4. AN AUXILIARY INTEGRAL FORMULA

In the next sections, the following integral will be useful

$$I_N(M,\eta) = \int_{\mathbb{R}^N} \exp\left\{\frac{i}{2}\xi\cdot M\xi + i\xi\cdot\eta\right\} d\xi, \qquad (6)$$

where M is a symmetric $N \times N$ matrix the imaginary part of which is assumed strictly positive, $\mathrm{Im}\,\xi\cdot M\xi > 0$ for each nonzero $\xi \in \mathbb{R}^N$, and $\eta \in \mathbb{C}^N$.

Proposition 4: Under the stated assumptions, the integral (6) equals

$$I_N(M,\eta) = (2\pi i)^{N/2}(\det M)^{-1/2}\exp\left\{-\frac{i}{2}\eta\cdot M^{-1}\eta\right\}. \qquad (7)$$

Proof based on analytic continuation[23]: we denote $M_\lambda = \lambda M_1 + iM_2$, $\eta_\lambda = \lambda\eta_1 + i\eta_2$, where $M_1 = \mathrm{Re}\,M = \frac{1}{2}(M + M^T)$, etc. so that $M = M_1$, $\eta = \eta_1$. Due to the assumption, $\epsilon M_1 + M_2$ is strictly positive for all real ϵ with small enough modulus. Thus there is $\epsilon_0 > 0$ such that $I_N(M_\lambda, \eta_\lambda)$ exists in the strip $S = \{\lambda : |\mathrm{Im}\,\lambda| < \epsilon_0\}$. Moreover the function $\lambda \mapsto I_N(M_\lambda, \eta_\lambda)$ is easily seen to be analytic in S.[24] For each $\lambda = i\epsilon$, $|\epsilon| < \epsilon_0$, one can choose a basis in $\mathbb{R}^N$ in which M_λ is diagonal so that by Fubini's theorem we obtain[25]

$$I_N(M_{i\epsilon}, \eta_{i\epsilon}) = \prod_{j=1}^{N}\int_{\mathbb{R}} \exp\left\{\frac{i}{2}m_j\xi_j^2 + i\xi_j z_j\right\} d\xi_j = \prod_{j=1}^{N}(2\pi i m_j^{-1})^{1/2}\exp\left\{-\frac{i}{2}m_j^{-1}z_j^2\right\},$$

where m_j, z_j are eigenvalues of $M_{i\epsilon}$ and components of $\eta_{i\epsilon}$, respectively. Then due to the proved analycity, (7) holds for all M_λ, η_λ with $\lambda \in S$, in particular for $\lambda = 1$. ∎

5. THE PROPAGATOR

The continuous contractive semigroup corresponding to our pseudo-Hamiltonian H can be expressed explicitly. This is the content of the following theorem which will be proved in the next section:

Theorem 2: Let A, W be strictly positive and denote $\Omega = -(2B)^{1/2}$, $B = A - iW$. Then for each $t \geqslant 0$, $\exp(-iHt) = V_t$, where $\{V_t : t \geqslant 0\}$ is a contractive semigroup which acts on an arbitrary $\varphi \in L^2(\mathbb{R}^d)$ according to the relation

$$(V_t\varphi)(x) = \int_{\mathbb{R}^d} G_t(x,y)\varphi(y)\,dy, \quad t > 0, \qquad (8a)$$

$$G_t(x,y) = (2\pi i)^{-d/2}(\det(\Omega^{-1}\sin\Omega t))^{-1/2} \times \exp\left\{\frac{i}{2}[x\cdot(\Omega\cot\Omega t)x + y\cdot(\Omega\cot\Omega t)y\right\} - iy\cdot(\Omega\csc\Omega t)x]. \qquad (8b)$$

One has to verify first that (8) makes sense:

Lemma 5.1: Let A be positive, W strictly positive, $t > 0$; then Ω is regular and the real quadratic forms $x \mapsto \mathrm{Im}\,x\cdot Mx$ with $M = -\Omega^{-1}\tan\Omega t$, $-\Omega\tan\Omega t$, and $\Omega\cot\Omega t$ are strictly positive.

Proof: Suppose first $d = 1$. We have $3\pi/2 \leqslant \arg B < 2\pi$ due to the assumption so $0 < \nu \leqslant \omega$ hold for $\Omega = \omega - i\nu$. The $-\mathrm{Im}\,\Omega^{-1}\tan\Omega t = C(t)\,(\omega\tanh\nu t\cos^{-2}\omega t - \nu\tan\omega t\cosh^{-2}\nu t)$ with $C(t) > 0$, and the inequalities $\alpha^{-1}\sin\alpha < 1 < \beta^{-1}\sinh\beta$ for nonzero α, β imply

$$-\mathrm{Im}\,\Omega^{-1}\tan\Omega t = C(t)[2t\cos^2\omega t\cosh^2\nu t]^{-1}[\omega t\sinh(2\nu t) - \nu t\sin(2\omega t)] > 0. \qquad (\star)$$

Analogously we obtain positivity of $\mathrm{Im}\,\Omega\cot\Omega t$ and $-\mathrm{Im}\,\Omega\tan\Omega t$. As for the case $d > 1$, regularity of Ω follows from symmetry of Ω, which gives $|\Omega x|^2 = x\cdot Bx$, and from strict positivity of W. A real quadratic form is strictly positive if all eigenvalues of its matrix are positive.[26] They equal $-\mathrm{Im}\,\omega_j^{-1}\tan\omega_j t$ in the first case,[27] where ω_j are eigenvalues of Ω. Further, each eigenvalue $\beta_j = \frac{1}{2}\omega_j^2$ of B satisfies $\mathrm{Im}\,\beta_j < 0$; otherwise $y\cdot Wy = -\mathrm{Im}\,\beta_{j_0}|y|^2 \leqslant 0$ for some nonzero y in contradiction with the assumption. Thus ($\star$) gives $-\mathrm{Im}\,\omega_j^{-1}\tan\omega_j t > 0$ for all j; in the same way the assertion is obtained for the other two forms. ∎

Lemma 5.2: Let A, W be as in Lemma 5.1, then $\det(\Omega^{-1}\sin\Omega t)$ and $\det(\cos\Omega t)$ are nonzero for each $t > 0$.

Proof: It is sufficient to check that all eigenvalues of both the matrices are nonzero: they equal $\omega_j^{-1}\tan\omega_j t$ and $\cos\omega_j t$, $j = 1,\dots,d$, respectively. Further, $\mathrm{Im}\,\beta_j < 0$ implies $\mathrm{Im}\,\omega_j \neq 0$, but sin and cos have no zeros outside the real axis. ∎

Proposition 5: Let A,W be as in Lemma 5.1, let further V_t be given by (8) and $V_0 = I$. Then $\{V_t : t \geqslant 0\}$ is a semigroup of bounded operators on $L^2(\mathbb{R}^d)$.

Proof: According to Lemma 5.1 there exist positive c_1, c_2 (depending on t) such that

$$|G_t(x,y)| \leqslant c_1 \exp(-c_2(x^2+y^2)). \tag{9}$$

This inequality together with the Fubini theorem imply

$$\|V_t\varphi\|^2 \leqslant c_1^2 \int_{\mathbb{R}^{3d}} |\varphi(y)||\varphi(z)| \times \exp(-c_2(2x^2+y^2+z^2))\,dx\,dy\,dz,$$

so integration over x and the Schwarz inequality gives

$$\|V_t\varphi\| \leqslant c_1\left(\frac{\pi}{2c_2}\right)^{d/2}\|\varphi\| \tag{10}$$

for each $\varphi \in L^2(\mathbb{R}^d)$. As for the semigroup property, in view of $V_0 = I$ and of (9) it is sufficient to verify

$$G_{t+s}(x,z) = \int_{\mathbb{R}^d} G_s(x,y)G_t(y,z)\,dy$$

for all $t,s > 0$; it follows from Proposition 4 with $M = \Omega(\cot \Omega t + \cot \Omega s)$, $\eta = -\Omega((\operatorname{cosec} \Omega s)x + (\operatorname{cosec} \Omega t)z)$ and from the matrix functional-calculus rules.[27] ■

The following equivalent expression for V_t will be useful:

Proposition 6: Let A,W be as in Lemma 5.1; then for all $t>0$ and $\varphi \in L^2(\mathbb{R}^d)$

$$(V_t\varphi)(x) = \int_{\mathbb{R}^d} F_t(x,y)(F_d\varphi)(y)\,dy, \tag{11a}$$

$$F_t(x,y) = (2\pi)^{-d/2}(\det(\cos \Omega t))^{-1/2} \times \exp\{-(i/2)[x\cdot(\Omega \tan \Omega t)x + y\cdot(\Omega^{-1}\tan \Omega t)y\} + iy\cdot(\sec \Omega t)x], \tag{11b}$$

where F_d is the Fourier-Plancherel operator.

Proof: Let first $\mathscr{F}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, $\varphi(x) = \int_{\mathbb{R}^d} e^{ix\cdot y}d\nu(y)$ with $\nu \in \mathscr{M}(\mathbb{R}^d)$; then (11a) can be rewritten as

$$(V_t\varphi)(x) = (2\pi)^{d/2}\int_{\mathbb{R}^d} F_t(x,y)\,d\nu(y). \tag{12}$$

In order to prove this, we use (9) together with boundedness of φ, $|\varphi(x)| \leqslant |\nu|(\mathbb{R}^d)$. Then the Fubini theorem applied to (8) gives (12) with

$$F_t(x,y) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} G_t(x,z)e^{iy\cdot z}\,dz$$
$$= (4\pi^2 i)^{-d/2}(\det(\Omega^{-1}\sin \Omega t))^{-1/2} \times \exp\{(i/2)x\cdot(\Omega \cot \Omega t)x\} \cdot I_d(\Omega \cot \Omega t, y - (\Omega \csc \Omega t)x).$$

Using now Proposition 4 and the matrix functional-calculus rules, we get (11b). Let us assume further an arbitrary $\varphi \in L^2(\mathbb{R}^d)$, and construct the following sequence:

$$\varphi_n : \varphi_n(x) = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{ix\cdot y}\hat{\varphi}_n(y)\,dy,$$

$$\hat{\varphi}_n(y) = \begin{cases} (F_d\varphi)(y) & \ldots\ |y| \leqslant n \text{ and } |(F_d\varphi)(y)| \leqslant n, \\ n & \ldots\ |y| \leqslant n \text{ and } |(F_d\varphi)(y)| > n, \\ 0 & \ldots\ |y| > n. \end{cases}$$

Clearly $F_d\varphi_n = \hat{\varphi}_n$ and $\hat{\varphi}_n \in L(\mathbb{R}^d)$ so the assertion is valid for φ_n. The sequence $\{\hat{\varphi}_n\}$ converges pointwise to $F_d\varphi$; further, $|\hat{\varphi}_n(y)| \leqslant |(F_d\varphi)(y)|$ and $F_t(x,\cdot) \in L^2(\mathbb{R}^d)$ so that

$$\lim_{n\to\infty}(V_t\varphi_n)(x) = \int_{\mathbb{R}^d} F_t(x,y)(F_d\varphi)(y)\,dy. \tag{**}$$

One verifies easily that $\hat{\varphi}_n \to F_d\varphi$ in the L^2-norm too. Since F_d is unitary and V_t is bounded due to (10), we obtain $V_t\varphi_n \to V_t\varphi$; then there exists a subsequence $\{V_t\varphi_{n_k}\}$ which converges to $V_t\varphi$ pointwise and the assertion follows from (**). ■

In order to prove Theorem 2 in a straightforward way, one has to check first strong continuity of $\{V_t\}$ or equivalently[28]

$$\lim_{t\to 0+}(\psi, V_t\varphi) = (\psi,\varphi) \tag{13}$$

for all $\psi, \varphi \in L^2(\mathbb{R}^d)$. Further, the generator of $\{V_t\}$ must be calculated and shown to coincide with H. Proposition 6 shows that (13) is valid for $\psi, \varphi \in L^2(\mathbb{R}^d) \cap L(\mathbb{R}^d)$. Using further the matrix functional-calculus rules, one can verify that for $\varphi \in \mathscr{F}(\mathbb{R}^d)$, $\psi : \psi(x,t) = (V_t\psi)(x)$ solves in $\mathbb{R}^d \times (0,\infty)$ the Schrödinger equation with potential $v(x) = \frac{1}{2}x\cdot\Omega^2 x$ and initial data φ. However, the remaining part of such a proof seems to be complicated, and therefore we choose another way: to express $\exp(-iHt)$ and identify it with V_t.

6. exp($-iHt$) BY LIE–TROTTER FORMULA

We shall assume again both A,W to be strictly positive, $t>0$, and abbreviate $S_n^t = \exp(-iH_0 t)\exp(-iVt)$, where $H_0 = \frac{1}{2}P^2$ is the free Hamiltonian. Since $iH = iH_0 + iV$ generates a continuous contractive semigroup due to Theorem 1, the Lie–Trotter formula for semigroups asserts[29]

$$s\text{-}\lim_{n\to\infty} S_n^t = \exp(-iHt). \tag{14}$$

Our goal is to prove that the lhs of (30) coincides with V_t. For $\varphi \in L^2(\mathbb{R}^d)$ we have[30]

$$(S_n^t\varphi)(x) = (2\pi i\delta)^{-nd/2}\int_{\mathbb{R}^{nd}} \exp\left\{\frac{i}{2\delta}\sum_{k=0}^{n-1}(\gamma_{k+1}-\gamma_k)^2 - i\delta\sum_{k=0}^{n-1}\gamma_k\cdot B\gamma_k\right\}\varphi(\gamma_0)d\gamma_0\ldots d\gamma_{n-1},$$

where $\gamma_n = x$ and $\delta = t/n$. Modulus of the integrand is majorized by $|\varphi(\gamma_0)|\exp\{-\delta\,\Sigma_{k=0}^n \gamma_k\cdot W\gamma_k\}$; thus, the rhs makes sense and the integrations may be interchanged arbitrarily. Assume now $\varphi \in \mathscr{F}(\mathbb{R}^d)$, $\varphi(x) = \int_{\mathbb{R}^d} e^{ix\cdot y}d\nu(y)$, then substituting $\gamma_k = \alpha_k\delta^{1/2}$, $k = 0,1,\ldots,n-1$, and rearranging the integral, we obtain

$$(S_n^t\varphi)(x) = (2\pi i)^{-nd/2}\int_{\mathbb{R}^d} d\nu(y)\exp\left(\frac{i}{2}\delta x^2\right)I_{nd}(M_n,\eta), \tag{15a}$$

where $\eta = (y\delta^{1/2},0,\ldots,0,-x\delta^{-1/2})$ and $M_n = M_n(\delta)$ is the $nd \times nd$ matrix

$$M_n = \begin{pmatrix} I-2\delta^2 B & -I & 0 & 0 & \cdots & 0 \\ -I & 2I-2\delta^2 B & -I & 0 & \cdots & 0 \\ 0 & -I & 2I-2\delta^2 B & -I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 2I-2\delta^2 B \end{pmatrix} \tag{15b}$$

which obviously fulfills the assumption of Proposition 4. Thus one has to calculate $\det M_n$ and the corner blocks of M_n^{-1}. Let us denote by $B_\lambda = A - i\lambda W$ and M_n^λ the corresponding matrix (15b). For small enough δ, there is $\epsilon_0 > 0$ such that M_n^λ is regular in $S = \{\lambda \neq i\epsilon: |\epsilon| \geqslant \epsilon_0\}$. Consequently, $\lambda \mapsto \det N_n^\lambda$ and the matrix function $\lambda \mapsto (M_n)^{-1}$ are analytic in S. If $\lambda = i\epsilon$, $|\epsilon| < \epsilon_0$, one can diagonalize M_n^λ and calculate the needed expressions. Continuing the results analytically[23] to the point $\lambda = 1$, we obtain

$$\begin{aligned}(S_n^t \varphi)(x) = {} & (\det[d(M_n)])^{-1/2} \int_{\mathbb{R}^d} d\nu(y) \\ & \times \exp\Big\{ \frac{i}{2\delta} x \cdot d(M_n)^{-1} [d(M_{n-1}) \\ & - d(M_n)] x - \frac{i\delta}{2} y \cdot d(M_{n-1})^{-1} \\ & \times d(K_{n-1}) y + iy \cdot d(M_n)^{-1} x \Big\},\end{aligned} \tag{16}$$

where $d(M_n)$, $d(K_{n-1})$ are the "block determinants" of M_n and its lower-right $(n-1)d \times (n-1)d$ submatrix, respectively. They satisfy the relations

$$d(K_{n-1}) = (2I - \delta^2 \Omega^2) d(K_{n-2}) - d(K_{n-3}),$$
$$d(M_n) = (I - \delta^2 \Omega^2) d(K_{n-1}) - d(K_{n-2}),$$

which can be seen easily to have the following solutions

$$d(K_{n-1}) = \sum_{j=0}^{\infty} (-1)^j \binom{n+j}{2j+1} (\delta\Omega)^{2j},$$
$$d(M_n) = \sum_{j=0}^{\infty} (-1)^j \binom{n+j}{2j} (\delta\Omega)^{2j}.$$

Let us turn now to the limits. Assume first $\delta d(K_{n-1}(\delta))$ with $\delta = t/n$. This sum converges (because it is finite), further,

$$\frac{t}{n} d\left(K_{n-1}\left(\frac{t}{n}\right)\right) = \Omega^{-1} \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} c_{nj} (\Omega t)^{2j+1},$$

where $c_{nj} = \Pi_{k=1}^j (1 - k^2 n^{-2})$ so $0 \leqslant c_{nj} \leqslant 1$ for all n, j, and therefore the convergence is uniform with respect to n. Thus, we have

$$\lim_{n\to\infty} \frac{t}{n} d\left(K_{n-1}\left(\frac{t}{n}\right)\right) = \Omega^{-1} \sin \Omega t;$$

similarly one obtains $\lim_{n\to\infty} d(M_n(t/n)) = \cos \Omega t$ and $\lim_{n\to\infty} (n/t)[d(M_{n-1}(t/n)) - d(M_n(t/n))] = \Omega \sin \Omega t$. These relations together with (16), (11b), and (12) imply

$$\lim_{n\to\infty} (S_n^t \varphi)(x) = (V_t \varphi)(x)$$

for $\varphi \in \mathscr{F}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. On the other hand, $\lim_{n\to\infty} S_n^t \varphi = \exp(-iHt)\varphi$ for these φ due to (14) so there exists a subsequence $\{S_{n_k}^t \varphi\}$ which converges to $\exp(-iHt)\varphi$ pointwise a.e. in $\mathbb{R}^d$. Consequently, we have

$$V_t \varphi = \exp(-iHt)\varphi \tag{17}$$

for all $\varphi \in \mathscr{F}(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. This set is, however, dense in $L^2(\mathbb{R}^d)$ (containing, e.g., $\mathscr{S}(\mathbb{R}^d)$) and the operators V_t, $\exp(-iHt)$ are bounded due to Proposition 5 and Theorem 1, respectively. Thus, (17) holds for each $\varphi \in L^2(\mathbb{R}^d)$ too, and the proof of Theorem 2 is finished.

7. THE NONDAMPED LIMIT AND MASLOV CORRECTION

For the sake of simplicity we shall limit ourselves further to the one-dimensional case. It is known that the Feynman's propagator formula for the nondamped harmonic oscillator must be corrected by jumps in phase at every half-period:

$$K_t(x,y) = K_t^F(x,y) M(t), \tag{18}$$

where

$$K_t^F(x,y) = (2\pi i)^{-1/2} \left(\frac{\omega}{|\sin \omega t|} \right)^{1/2} \times \exp\left\{ \frac{i\omega}{2 \sin \omega t} [(x^2 + y^2)\cos \omega t - 2xy] \right\}, \tag{18a}$$

$$M(t) = \exp\left\{ -\frac{\pi i}{2} \operatorname{Ent} \frac{\omega t}{\pi} \right\} \tag{18b}$$

if $t \neq \frac{1}{2}k\tau$ (we assume $m = \hbar = 1$) and

$$K_t(x,y) = \exp\{ -\frac{\pi i}{2} k \} \delta(x - (-1)^k y) \tag{19}$$

if $t = \frac{1}{2}k\tau$ (see Ref. 19 for further references).

We shall show that the *Maslov correction* (18b) emerges naturally in the nondamped limit of the above results:

Proposition 7: Let $d = 1$ and $\Omega = \omega - i\nu$ with ω, ν positive. Then, if $\omega t \neq k\pi$, $k = 0,1,2,\ldots$, and $\varphi \in L^2(\mathbb{R})$ has a compact support, we have

$$\lim_{\nu\to 0+} (V_t \varphi)(x) = \int_{\mathbb{R}} K_t(x,y) \varphi(y)\, dy. \tag{20}$$

On the other hand, it holds

$$\lim_{\nu\to 0+} (V_t \psi)(x) = \exp\left\{ -\frac{\pi i}{2} k \right\} \psi((-1)^k x) \tag{21}$$

for $t = k\pi/\omega$ and $\psi \in \mathscr{S}(\mathbb{R})$.

Proof: Let $\omega t \neq k\pi$ and consider (8) with $d = 1$ and $\Omega = \omega - i\nu$. We denote

$$h_x(y) = \exp\{ (i\Omega/2 \sin \Omega t)(y^2 \cos \Omega t - 2xy) \};$$

then

$$|h_x(y)| = \exp\left\{ \frac{\omega \nu t}{2|\sin \Omega t|^2} \left[(y^2 \cos \omega t - 2xy \cosh \nu t) \frac{\sin \omega t}{\omega t} - (y^2 \cosh \nu t - 2xy \cos \omega t) \frac{\sin \nu t}{\nu t} \right] \right\}$$

so that

$$|h_x(y)| \leqslant \exp\{\omega|y|(|y| + 2|x|)\sinh \nu t \sin^{-2}\omega t\},$$

and therefore the dominated convergence theorem can be applied if φ has a compact support. It implies

$$\lim_{\nu\to 0+}(V_t\varphi)(x) = \lim_{\nu\to 0+}\exp\left\{\frac{i}{2}g_\nu(t)\right\}\int_{\mathbb{R}} K_t^F(x,y)\varphi(y)\,dy, \tag{22}$$

where

$$g_\nu(t) = \arg(\Omega/\sin\Omega t), \text{i.e.},$$

$$g_\nu(t) = \arctan(\tanh\nu t\cot\omega t) - \arctan(\nu/\omega) - k\pi \tag{23a}$$

for $k\pi < \omega t < (k+1)\pi$. The term $-k\pi$ is chosen so that the rhs is continuous in the points $t = k\pi/\omega$ and tends to zero with $t\to 0+$ which certainly must be true for g_ν. It is easy to see that g_ν is decreasing; its shape for three values of ν/ω is sketched on Fig. 1. For fixed t, (23a) gives

$$\lim_{\nu\to 0+} g_\nu(t) = -k\pi \quad \text{for } k\pi < \omega t < (k+1)\pi; \tag{23b}$$

this relation together with (18b) and (22) gives (20).

Let now in turn $\omega t = k\pi$. We take $\psi\in\mathscr{S}(\mathbb{R})$ and express $(V_t\psi)(x)$ from Proposition 6. Since

$$\left|\exp\left\{-\frac{y^2 t\tanh\nu t}{2(k\pi - i\nu t)} + \frac{i(-1)^k xy}{\cosh\nu t}\right\}\right| \leqslant 1,$$

the dominated convergence theorem can be again applied which gives

$$\lim_{\nu\to 0+}(V_t\psi)(x) = (2\pi)^{-1/2}\exp\left\{-\frac{\pi i}{2}k\right\} \times\int_{\mathbb{R}}\exp\{i(-1)^k xy\}(F\psi)(y)\,dy,$$

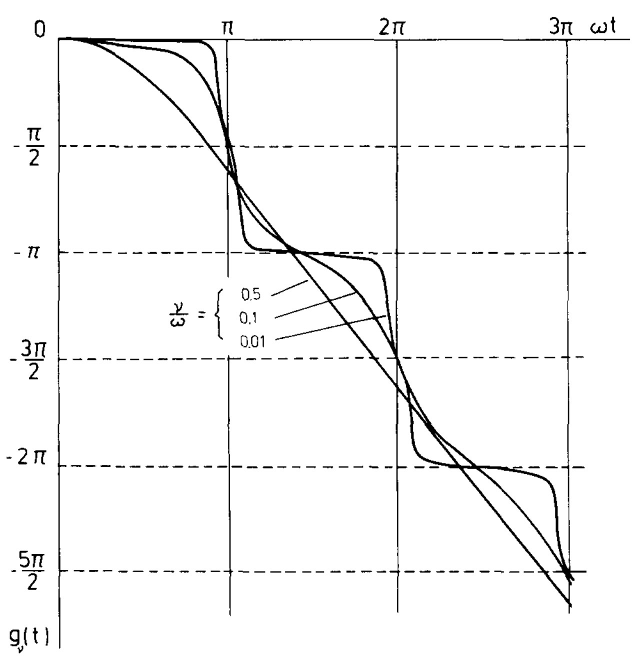

FIG. 1. The function g_ν.

where $F = F_1$ is the Fourier–Plancherel operator [so $F\psi\in\mathscr{S}(\mathbb{R})$]. Using further $(F^2\psi)(x) = \psi(-x)$ for k odd, we arrive at (21). ■

8. THE CLASSICAL LIMIT

As mentioned above, we limit ourselves to the case when the initial wavepackets are Gaussian, especially such obtained by shifting the "ground state." We take $\varphi = \varphi_{L,\alpha,\kappa}$:

$$\varphi(x) = (\pi l^2)^{-1/4}\exp\{-(2L^2)^{-1}(x-\alpha)^2 + (i/\hbar)\kappa x\} \tag{24a}$$

with L complex, $\operatorname{Re} L^2 \geqslant 0$, $l^{-2} = |L|^{-4}\operatorname{Re} L^2$, and α, κ real. Expectations and dispersions of position and momentum are the following

$$\langle Q\rangle_\varphi = \alpha, \quad \langle P\rangle_\varphi = \kappa,$$

$$(\Delta Q)_\varphi = 2^{-1/2}l, \quad (\Delta P)_\varphi = 2^{-1/2}\hbar l|L|^{-2}. \tag{25}$$

The propagator referring to arbitrary m and $\hbar$ is obtained from (8b) by substitutions $t\to\hbar t/m$, $\Omega\to m\Omega/\hbar$. Applying now Theorem 2 with this modification and Proposition 4, we obtain

$$\begin{aligned}(V_t\varphi)(x) = {}& (\pi l^2)^{-1/4}(\cos\Omega t + i\Lambda^2 L^{-2}\sin\Omega t)^{-1/2}\\ &\times\exp\{-i(2\Lambda^2)^{-1}\\ &\times\frac{\sin\Omega t - i\Lambda^2 L^{-2}\cos\Omega t}{\cos\Omega t + i\Lambda^2 L^{-2}\sin\Omega t}\\ &\times[x^2 - 2xz\Lambda^2(\sin\Omega t\\ &- i\Lambda^2 L^{-2}\cos\Omega t)^{-1}\\ &+ \Lambda^4 z^2(\sin\Omega t - i\Lambda^2 L^{-2}\cos\Omega t)^{-1}\\ &\times\sin\Omega t\} - \tfrac{1}{2}\alpha^2 L^{-2}],\end{aligned} \tag{26a}$$

where $\Lambda^2 = \hbar/m\Omega$ and $z = \kappa\hbar^{-1} - i\alpha L^{-2}$. Further, we choose L as follows

$$L^2 = \Lambda^2 = \hbar/m\Omega, \tag{24b}$$

and denote as above $\Omega = \omega - i\nu$; then (26a) can be simplified into the form

$$\begin{aligned}(V_t\varphi)(x) = {}& (\pi\lambda^2)^{-1/4}\exp\{-i\tfrac{1}{2}\Omega t\\ &- \tfrac{1}{2}\Lambda^{-2}[x - (\alpha + (i/\hbar)\kappa\Lambda^2)e^{-i\Omega t}]^2\\ &+ \tfrac{1}{2}\Lambda^{-2}(\alpha + (i/\hbar)\kappa\Lambda^2)^2\\ &\times e^{-i\Omega t}\cos\Omega t - \tfrac{1}{2}\alpha^2\Lambda^{-2}\},\end{aligned} \tag{26b}$$

where $\lambda^2 = \hbar/m\omega$. The probability density is given by

$$|(V_t\varphi)(x)|^2 = (\pi\lambda^2)^{-1/2} \times\exp\{-\nu t - \lambda^{-2}(x - x_0(t))^2 + y(t)\}, \tag{27}$$

where

$$x_0(t) = [\alpha\cos\omega t + (m\omega)^{-1}(\kappa - m\alpha\nu)\sin\omega t]e^{-\nu t}, \tag{28}$$

and

$$y(t) = \tfrac{1}{2}\gamma^2\lambda^{-2} - \tfrac{1}{2}\lambda^{-2} - \tfrac{1}{2}\lambda^{-2}[(\beta^2 - \gamma^2)\cos 2\omega t + (\nu/\omega)(\alpha^2 - \gamma^2)\sin 2\omega t - \alpha^2 - \beta^2]e^{-2\nu t}$$

with

$$\beta = (m\omega)^{-1}(\kappa - m\alpha\nu), \quad \gamma^{-1} = m|\Omega|\kappa^{-1}.$$

Thus we have obtained the Gaussian-shaped function with the following properties:

(i) height of the peak decreases with time, for large t

approximately as $e^{-\nu t}$;

(ii) its width λ does not change, it is negligible in the classical limit when $\alpha^2 + \beta^2 \gg \lambda^2$;

(iii) the peak travels along $x = x_0(t)$ which is the trajectory of the classical damped oscillator with the initial position $x_0(0) = \alpha$, however, the corresponding initial momentum is $m\dot{x}_0(0) = \kappa - 2m\alpha\nu$ instead of κ. Denoting by $x_c(\cdot)$ the trajectory of CDO with initial conditions (α,κ), we have $x_c(t) - x_0(t) = 2\alpha\nu\omega^{-1}e^{-\nu t}\sin\omega t$ so that the difference is negligible in the case of weak damping, $\nu \ll \omega$.

9. THE POINT SPECTRUM OF H

Proposition 8: Let $d = 1, \Omega = \omega - i\nu$, the $\sigma_p(H)$ consists of eigenvectors

$$\psi_n(x) = N_{nn}^{-1/2} H_n(\Omega^{1/2}x)\exp(-\tfrac{1}{2}\Omega x^2), \quad n = 0,1,2,..., \tag{29a}$$

where H_n are the Hermite polynomials and

$$H\psi_n = \lambda_n \psi_n, \quad \lambda_n = \Omega(n + \tfrac{1}{2}). \tag{29b}$$

In general, the eigenvectors (29a) are not orthonormal:

$$(\psi_n, \psi_m) = N_{nn}^{-1/2} N_{mm}^{-1/2} N_{nm}, \text{ where}$$

$$N_{n,n+2s+1} = 0,$$

$$N_{n,n+2s} = \left(\frac{\pi}{\omega}\right)^{1/2} \frac{n!(n+2s)!}{(n+s)!}\omega^{-(n+2s)}|\Omega|^n \Omega^s \sum_{k=0}^{[n/2]} \sum_{l=0}^{[n/2]+s} (-1)^{k+l} \cdot \binom{2(n+s-k-l)}{n-2k}\binom{k+l}{k}\binom{n+s}{k+l}\omega^{k+l}(\bar{\Omega})^{-k}\Omega^{-l}, \tag{30}$$

$s = 0,1,2,...$, and $[\cdot]$ denotes the entire part.

Proof: The relations (29) and (30) can be checked by straightforward computation. Let us show that H has no other eigenvalues. For an arbitrary complex λ,[31] the equation $\psi'' + (2\lambda - \Omega^2 x^2)\psi = 0$ has the following fundamental solution:

$$\psi_\lambda(x) = [\alpha\Phi(\tfrac{1}{4} - \gamma, \tfrac{1}{2}; \Omega x^2) + \beta x \Phi(\tfrac{3}{4} - \gamma, \tfrac{3}{2}; \Omega x^2)]\exp(-\tfrac{1}{2}\Omega x^2), \tag{31}$$

where $2\Omega\gamma = \lambda$ and Φ is the degenerate hypergeometric function. We have $\mathrm{Re}\,\Omega x^2 = \omega x \geqslant 0$; thus, the asymptotic behavior of (31) [except for the cases, when one of the functions in (31) reduces to (29a) and the other is absent] is given by[32]

$$\psi_\lambda(x) = C(\alpha,\beta,\lambda,\Omega)x^{1/4-\gamma}\exp(\tfrac{1}{2}\Omega x^2)[1 + O(|x|^{-1})]$$

for large $|x|$, where $C(\alpha,\beta,\lambda,\Omega)$ is nonzero unless $\alpha = \beta = 0$. Consequently, (29a) are the only solutions to the above equation contained in $L^2(\mathbb{R})$. ■

In conclusion, let us make some remarks. It is easy to see that $P = \{\psi_n\}_{n=0}^{\infty}$ is complete in $L^2(\mathbb{R})$ so that for $\lambda \neq \lambda_n$, $n = 0,1,2,...$, $(H - \lambda)P_{\mathrm{lin}} = P_{\mathrm{lin}}$ is dense and *H has no residual spectrum*. The problem of absence of continuous spectrum will be considered separately. Proposition 8 determines, of course, also $\sigma_p(H)$ for the multidimensional oscillator in the case when $\Omega^2 = 2(A - iW)$ can be diagonalized. Moreover, some results remain true even if A, W are not simultaneously diagonalizable. For instance, one can check easily that the "ground state" vector

$$\psi_0(x) = \pi^{-d/4}(\det(\mathrm{Re}\,\Omega))^{1/4}\exp(-\tfrac{1}{2}x\cdot\Omega x)$$

corresponds to the eigenvalue $\frac{1}{2}\,\mathrm{Tr}\,\Omega$ for any A, W which satisfy the assumptions of Theorems 1 and 2; notice that it is not a minimum-uncertainty state [cf. (25)].

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. G. I. Kolerov who convinced me of the effectivity of path-integral methods. I am further grateful to Dr. J. Blank, Dr. L. A. Dadashev, Dr. E.-M. Ilgenfritz, and Professor O. G. Smolianov for useful discussions and valuable comments, and to the referee for his constructive criticism.

[1]R. P. Feynman and F. L. Vernon, Jr., Ann. Phys. **24**, 118–173 (1963).

[2]D. N. Williams, Commun. Math. Phys. **21**, 314–333 (1971).

[3]R. W. Haase, J. Math. Phys. **16**, 2005–2011 (1975).

[4]E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).

[5]K. Yasue, Ann. Phys. **114**, 479–496 (1978).

[6]D. C. Khandekar and S. V. Lawande, J. Math. Phys. **20**, 1870–1877 (1979).

[7]J. Blank, P. Exner, and M. Havlíček, Czechoslovak J. Phys. B **29**, 1325–1341 (1979).

[8]P. Exner, Commun. Math. Phys. **50**, 1–10 (1976).

[9]J. Dolejší and P. Exner, Czechoslovak J. Phys. B **27**, 855–862 (1977).

[10]P. Exner, Rep. Math. Phys. **17**, 275–285 (1980).

[11]H. Dekker, Phys. Lett. A **80**, 369–371 (1980), H. Dekker, Phys. Rep. **80**, 1–112 (1981).

[12]J. R. Brinati and S. S. Mizrahi, J. Math. Phys. **21**, 2154–2158 (1980).

[13]E. Nelson, J. Math. Phys. **5**, 332–343 (1964).

[14]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, II. Fourier Analysis, Self-adjointness* (Academic, New York, 1975), Secs. X.8 and X.11.

[15]A. Truman, J. Math. Phys. **18**, 1499–1509 (1977).

[16]P. Exner and G. I. Kolerov, Phys. Lett. A **83**, 203–206 (1981).

[17]P. Exner and G. I. Kolerov, Lett. Math. Phys. **6**, 153–159 (1982).

[18]P. Combe, G. Rideau, R. Rodriguez, and M. Sirugue-Collin, Rep. Math. Phys. **13**, 279–294 (1978).

[19]P. A. Horváthy, Int. J. Theor. Phys. **18**, 245–250 (1979).

[20]B.-G. Englert, Ann. Phys. **129**, 1–21 (1980).

[21]S. Okubo, "Non-canonical quantization of some dissipative systems," Preprint UR-763 (Rochester, 1980).

[22]Cf. Ref. 7 or Ref. 14, Sec. X.8.

[23]The analyticity argument was suggested by the referee; it is much simpler than the direct evaluation presented in the preliminary version of this paper (Preprints JINR E2-81-608, 609).

[24]For each $\lambda \in S$, the integrand in $I_N(M_\lambda, \eta_\lambda)$ is majorized by an exponentially decreasing function so $dI_N(M_\lambda, \eta_\lambda)/d\lambda$ exits.

[25]I. S. Gradshtein and I. M. Ryzhik, *Tables of integrals, sums, series and products* (Nauka, Moscow, 1971), 3.323.

[26]F. R. Gantmakher, *Theory of matrices* in Russian (Nauka, Moscow, 1966); Sec. X.5.

[27]Reference 26, Chap. V.

[28]K. Yosida, *Functional analysis* (Springer, Berlin, 1965), Theorem IX. 1.1.

[29]H. Trotter, Proc. Am. Math. Soc. **10**, 545–551 (1959) or Ref. 14, Theorem. X.51; in fact we need only the special case of the LT-formula considered by Nelson in Ref. 13.

[30]Reference 14, Secs. IX.7 and X.11.

[31]In fact, we have to consider Im $\lambda \leqslant 0$ only because the remaining λ belong to the resolvent set $\sigma(H)$ due to Theorem 1.

[32]*Handbook of mathematical functions*, edited by M. A. Abramowitz and I. Stegun (National Bureau of Standards, 1964), 13.1.4.

Extended continued fractions and energies of the anharmonic oscillators

M. Znojil
Nuclear Physics Institute, Czechoslovak Academy of Sciences, 250 68 Řež near Prague, Czechoslovakia

(Received 5 February 1982; accepted for publication 21 May 1982)

We describe the analytic solution to the Schrödinger eigenvalue problem for the class of the central potentials $V(r) = \Sigma_{\delta\in Z} a_\delta r^\delta$, where $a_{-2} > -1/4$, $a_{\max \delta} > 0$, Z is an arbitrary finite set of the integer or rational exponents, $-2 \leqslant \delta_1 < \delta_2 < \cdots < \delta_I$, and the couplings a_δ satisfy only one auxiliary (formal, "superconfinement") restriction of the type $a_{\delta_{I-1}} > 0$. The formalism is based on an analysis of the asymptotic behavior of the explicit regular solution $\psi(r)$ and issues in the formulation of the "secular" equation $1/L_1(E) = 0$ which determines the binding energies. The main result is the rigorous construction of $L_1(E)$ as a generalized ("extended") and convergent continued fraction. The method cannot be applied to the $a_{\delta_{I-1}} < 0$ cases—this disproves the closely related Hill-determinant approach as conjectured recently by Singh *et al.* for the simplest potentials with $Z = \{-2,2,4,6\}$ and $Z = \{-2,-1,1,2\}$.

PACS numbers: 03.65.Ge

I. INTRODUCTION

In the quantum field theory, the radial Schrödinger equation

$$-\frac{d^2}{dr^2}\psi(r) + \frac{l(l+1)}{r^2}\psi(r) + V(r)\psi(r) = E\psi(r)$$

$$l(l+1) = g_{-1} > -\tfrac{1}{4} \tag{1.1}$$

with the potential

$$V(r) = \sum_{j=1}^{2q+1} g_j r^{2j}, \quad g_{2g+1} = a^2 > 0, \tag{1.2}$$

appears in models with the simplest class of interactions.[1] It possesses the elementary solutions in some particular cases.[2] The more standard applications of Eq. (1.1) with the radial coordinate $r\in(0,\infty)$ and centrifugal interpretation of $g_{-1} = l(l+1), l = 0,1,\ldots$, range from the perturbed harmonic oscillations of various systems[3] up to the structure of quarkonium.[4]

The knowledge of the particular solutions to Eq. $(1.1)^2$ is insufficient in most cases, and the nonnumerical construction of the complete solution represents a challenge to the mathematical physicists: The simple-minded perturbation expansions fail to give the convergent results.[5]

To overcome the methodical difficulties connected with the anharmonic equation (1.1), a number of alternative approaches have been developed recently—let us mention here just the moment recursions,[6] $p - x$ symmetrization,[7] matrix continued fractions[8] and the continued-fraction method suggested for $q = 1$ by Singh *et al.*[9], generalized to any $q \geqslant 1$ in Ref. 10, and to an arbitrary fractionally anharmonic oscillator (FAO):

$$V(r) = \sum_{\delta\in Z} a_\delta r^\delta = \sum_{i=1}^{I} \gamma_i r^{-2+m_i/n_i},$$

$$0 \leqslant m_1/n_1 < \cdots < m_I/n_I \tag{1.3}$$

in Ref. 11.

Our present paper has been inspired by the misleading Hill-determinant interpretation of the results in Refs. 9, 10, which was criticized by a few authors.[12] In brief, our intention is to clarify the situation and to prove in a rigorous way the following:

(A) For the potential (1.2), the secular equation of the Hill-determinant type may be written in the form

$$1/L_1(E) = 0, \tag{1.4}$$

where $L_1(E)$ is an extended continued fraction[13] (ECF) to be defined below.

(B) For the potential (1.2), the roots of Eq. (1.4) coincide with the Schrödinger eigenvalues provided that the couplings satisfy the "superconfinement" restriction $g_{2q} > 0$ or

$$g_{2q-m_0} > 0, \quad g_{2q-m_0+i} = 0, \quad i = 1, 2, \ldots, m_0, \tag{1.5}$$

for some nonnegative integer $m_0 \leqslant q - 1$.

(C) The validity of the statements (A) and (B) may simply be extended to the general FAO potential (1.3).

The material is organized as follows. In Sec. II, we summarize or reformulate some of the results of Refs. 10 and 11 and describe the analytic form of the general solution to the differential equation (1.1) with the FAO potential (1.3). In Sec. III, we present an analysis of the asymptotic structure of $\psi(r)$ and complement it in Sec. IV by a thorough investigation of the other related formal questions. Finally, it is relatively straightforward to complete our discussion in Sec. V and to identify the convergent ECF quantity $L_1(E)$ with the physical "Green's" function in all the SFAO ("superconfining" FAO) cases. In this way, Eq. (1.4) may be interpreted as an analytic sum of the (divergent, Brillouin–Wigner) perturbation expansion of the binding energy, or as a direct, though implicit, analog of the well-known harmonic-oscillator quantization requirement

$$E = \hbar\omega(2n + l + \tfrac{3}{2}).$$

II. GENERAL SOLUTION TO THE DIFFERENTIAL EQUATION

By the direct insertion, we may easily verify that the explicit analytic solution of the radial Schrödinger equation

 0022-2488/83/051136-06$02.50

(1.1) with the generalized harmonic potential Eq. (1.2) has the closed form

$$\psi(r) = r^{\nu_\pm} \exp[-f(r)]\varphi(r), \quad \nu_\pm = \tfrac{1}{2} \pm (\tfrac{1}{4} + g_{-1})^{1/2}, \tag{2.1}$$

$$f(r) = \sum_{j=1}^{q+1} \frac{\beta_j r^{2j}}{2j}, \quad \varphi(r) = \sum_{n=0}^{\infty} h_{n+1} r^{2n},$$

where the Taylor coefficients have the form

$$h_{n+1} = h_1 \frac{\Gamma(\nu + \frac{1}{2}) \det Q(n)}{4^n n! \Gamma(n + \nu + \frac{1}{2})} \tag{2.2}$$

and the $(q+2)$-diagonal, $(n \times n)$-dimensional matrix $Q(n)$ is defined by the prescription $(E = -g_0)$

$$\begin{aligned} &Q(n)_{jj+1} = B_j = -4j(j + \nu - \tfrac{1}{2}), \quad j = 1,2,\dots,n-1, \\ &Q(n)_{j+kj} = C^{(k)}_{j+k} = 4\beta_{k+1}(j - 1 + \nu/2) + G_k, \\ &\quad j = 1, 2, \dots, n-k, \\ &G_k = g_k - \sum_{i=1}^{k} \beta_i \beta_{k+1-i} + \beta_{k+1}(2k+1), \quad k = 0,1,\dots,q. \end{aligned} \tag{2.3}$$

The function $f(r)$ is closely related to the potential, $V(r) = [\partial_r f(r)]^2 + O(r^{2q})$, $r \gg 1$, i.e.,

$$\begin{aligned} &\beta_{q+1} = a = g^{1/2}_{2q+1}, \\ &\beta_k = \frac{1}{2\alpha}\Big(g_{q+k} - \sum_{j=k+1}^{q} \beta_j \beta_{k+q+1-j}\Big), \\ &k = q, q-1, \dots, 1. \end{aligned} \tag{2.4}$$

When we introduce the new (barred) variables[11]

$$\bar{r} = r^p, \quad \bar{\psi}(\bar{r}) = r^{(p-1)/2}\psi(r), \quad 1 \leqslant p \leqslant 2q+2, \tag{2.5}$$

the radial Schrödinger equation (1.1) transforms into itself, with the barred centrifugal and energy parameters

$$\bar{g}_{-1} = (\tfrac{1}{4} + g_{-1})/p^2 - \tfrac{1}{4}, \quad \bar{E} = -g_{p-1}/p^2 \tag{2.6}$$

and with the new, broader class of potentials

$$\begin{aligned} &\bar{V}(\bar{r}) = \bar{g}_{-1}\bar{r}^{-2} + V_{(-)}(\bar{r}) + V_{(+)}(\bar{r}), \\ &V_{(\pm)}(x) = p^{-2} \sum_{m=1}^{M_\pm} g_{p-1\pm m} x^{\pm 2m/p} \\ &M_+ = 2q + 2 - p, \quad M_- = p - 1 \end{aligned} \tag{2.7}$$

equivalent to the FAO forces (1.3) and containing Eq. (1.2) as a $p = 1$ special case, of course.

Theorem 1: For any FAO potential written in the form of Eq. (2.7), the regular solution $\bar{\psi}(\bar{r})$ to the radial Schrödinger equation is given by Eq. (2.5), where $\psi(r)$ is defined by Eqs. (2.1)–(2.4) and $\nu = \nu_+$. Similarly the irregular solution is obtained when $\nu = \nu_-$.

Proof: By insertion, we find that $\bar{\psi}(\bar{r})$ is a solution. The regularity of the new equation in the origin is guaranteed since $\bar{g}_{-1} > -\frac{1}{4}$ if and only if $g_{-1} > -\frac{1}{4}$, and an estimate $\psi(r) \sim r^{\nu_+}$, $r \ll 1$, is equivalent to $\bar{\psi}(\bar{r}) \sim \bar{r}^{\bar{\nu}_+}$, $\bar{r} \ll 1$, since $\bar{\nu}_\pm = \frac{1}{2} + (\nu_\pm - \frac{1}{2})/p$. QED

Let us add the following three remarks.

(i) For the given set $Z_{\rm FAO}$ of exponents $\delta_i = n_i/m_i$, i is useful to determine the minimal value of q in practice. This reconstruction of Eq. (2.7) from $V_{\rm FAO}$ may be done in two steps. First, we divide all δ's in $Z_{\rm FAO}$ by 2, find their minimal common denominator M ($\delta_i = 2N_i/M$), and put $p = M \times T$, where T should be a minimal positive integer. Second, we must guarantee that the energy term is not omitted from the Schrödinger equation and that the leading exponent $2N_I/M = \max(\delta,0)$ has the particular form $2(2q + 2 - p)/p$, i.e., $(N_I + M)T = 2q + 2$. Thus, for even $N_I + M = 2k$ we choose $T = 1$ and put $q = k - 1$ while for odd $N_I + M = 2k + 1$ we must take $T = 2$ and $q = 2k$. This procedure defines the minimal parameter q uniquely—a few examples are shown in Table I.

TABLE I. Sample of the "complexities" of the simplest FAO one-term potentials $V(r) = gr^{N/M}$.

M \ N	1	2	3	4	5	6
−4	3	1	3	0	3	1
−3	2	2	0	2	2	0
−2	1	0	1	0	—	—
−1	0	0	—	—	—	—
1	2	0	4	2	6	1
2	4	2	6	0	8	4
3	6	1	2	4	10	0
4	8	4	10	2	12	6

(ii) The choice of $p = 2q + 2$ in Eq. (2.5) is a little bit exceptional since $V(\infty) = V_{(-)}(\infty) = 0$ and our "spectral" restriction $g_{2q+1} > 0$ eliminates in effect the continuous part of the spectrum (cf., e.g., the well-known $q = 0$ harmonic oscillator→Coulomb-potential transformation). At the same time, when $g_{2q-m} = 0$, $m = 0, 1, 2, \dots, m_0 - 1$, and $g_{2q-m_0} \neq 0$, the "subdominant anharmonicity" m_0 is to be added to the classification of $V = V_{\rm FAO}$ and ψ by the "complexity" q and the "fractionality" p.

(iii) The choice of $p = 1$ with $g_{-1} = 0$ seems to be also exceptional—it admits the ($r \leftrightarrow -r$ symmetric) one-dimensional interpretation of Eq. (1.1) and $\psi(r)$, with $r \in (-\infty, \infty)$ and negative and positive parity for $\nu = \nu_+$ and $\nu = \nu_-$, respectively. For $q = 0$, the series $\varphi(r)$ in Eq. (2.1) coincides with the well-known confluent hypergeometric function, while for $q = 1$, it reproduces precisely the sextic-oscillator functions of Ref. 9.

III. ASYMPTOTIC BEHAVIOR OF THE REGULAR SOLUTIONS

In analogy with the Singh's papers,[9] we may achieve in principle even the termination of the infinite series $\varphi(r)$ in Eq. (2.1) for some potentials.[2] For this purpose, it is sufficient to require the validity of the $q + 1$ independent algebraic conditions (determinantal constraints)

$$h_{N+i} = 0, \quad i = 1, 2, \dots, q+1. \tag{3.1}$$

They fix the energy plus q couplings—the thorough discussion of this peculiarity may be found in Refs. 2 and shows that we get at most two polynomial bound states when $q > 3$. The simplest harmonic-oscillator and Coulombic $q = 0$ forces appear to be the only exceptions giving the complete set of the terminating solutions (Laguerre polynomials) accompanied by the explicit formula (i.e., $h_{N+1} = 0$) for the energy.

Reverting the preceding argument, we may conclude that the normalizable $q \geqslant 1$ bound-state solutions should be represented by an infinite series in general. In accord with the standard textbooks,[14] this series is convergent and its absolute value grows in the asymptotic region, provided that the energy E does not coincide with the physical eigenvalues E_n. At $E = E_n$ the value of $\psi(r_0)$ changes sign in the limit $r_0 \to \infty$—this may be used in the numerical determination of the discrete spectrum.

In the present paper, the structure of $\psi(r)$ and its asymptotic behavior in the vicinity of E_n will be clarified by the decomposition of the regular solution Eq. (2.1) into the $q+1$ partial infinite summations

$$\psi(r) = \sum_{j=0}^{q} \chi_j(r),$$
$$\chi_j(r) = r^{\nu_+} e^{-f(r)} \sum_{k=0}^{\infty} h_{n_j(k)+1} r^{2n_j(k)}, \tag{3.2}$$
$$n_j(k) = k\cdot(q+1) + j.$$

Let us emphasize that the value of fractionality p is irrelevant in this context, of course—it concerns just the interpretation of one of the constants $(-g_{p-1}/p^2)$ as energy.

Proposition 1: The necessary and sufficient condition for the FAO regular wavefunction $\psi(r) = \chi_0(r) + \cdots + \chi_q(r)$ to be normalizable is either its termination or an asymptotic cancellation of its $q+1$ growing exponential components $\chi_j(r) \sim \exp[f(r)]$, $r \gg 1$.

Proof: We introduce an auxiliary sequence $L_n^{(N)}$ defined by the recurrences

$$L_k^{(N)} = -B_{k-1} \Big/ \left[C_k^{(0)} + \sum_{j=1}^{q} \left(\prod_{m=0}^{j-1} L_{k+m+1}^{(N)} \right) C_{k+j}^{(j)} \right], \tag{3.3}$$

$L_{N+1} = L_{N+2} = \cdots = 0, \quad k = N, N-1, \ldots, 1, \quad B_0 \equiv 1,$

abbreviate the combinations

$$U_{m+k}^{(m)} = C_{m+k}^{(m)} + L_{m+k+1}^{(N)} U_{m+k+1}^{(m+1)}, \quad 1 \leqslant m \leqslant q-1, \tag{3.4}$$

$$U_{q+k}^{(q)} = C_{q+k}^{(q)}, \quad U_{m+k}^{(m)} = 0, \quad m > q, \quad k = 1,2,\ldots,N-m,$$

and define formally the product decomposition of the matrix Eq. (2.3)

$$Q(N) = -X(N)Y(N)Z(N) \tag{3.5}$$

with the nonzero elements

$$X(N)_{kk} = 1, \quad Z(N)_{kk} = B_{k-1},$$
$$Y(N)_{kk} = 1/L_k^{(N)}, \quad 1 \leqslant k \leqslant N,$$
$$X(N)_{k\,k+1} = -L_{k+1}^{(N)}, \quad Z(N)_{k+m\,k} = -L_{k+m}^{(N)} U_{m+k}^{(m)}, \tag{3.6}$$
$$1 \leqslant m \leqslant \min(q, N-k), \quad 1 \leqslant k \leqslant N-1.$$

For the Taylor coefficients (2.2) we then get

$$h_{N+1} = h_1 / L_1^{(N)} L_2^{(N)} \cdots L_N^{(N)}. \tag{3.7}$$

As a rule, the singularities of the type $L_k^{(N)} = \infty$ are to be understood as limits of the regular cases.

In the $n \gg 1$ asymptotic region we have $C_n^{(j)} = 4n\beta_{j+1}$, $j = 0,1,,\ldots,q$, $B_n = -4n^2$ and

$$L_n^{(N)} = n \Big/ \left(\beta_1 + \sum_{j=1}^{q} \beta_{j+1} \prod_{m=1}^{j} L_{n+m}^{(N)} \right) \tag{3.8}$$

within the $1 + O(1/n)$ error bounds. As a consequence we obtain the first few elements of the sequence $L_n^{(N)}$ in the form

$$L_{N-(q+1)k}^{(N)} = N/\beta_1^{(k)}, \quad L_{N-(q+1)k-j}^{(N)} = \beta_j^{(k)}/\beta_{j+1}^{(k)},$$
$$j = 1, 2, \ldots, q, \tag{3.9}$$

where $\beta_j^{(0)} = \beta_j$ and

$$\beta_j^{(k+)} = \beta_j + \frac{1}{\alpha} \sum_{m=1}^{q+1-j} \beta_{q+1-m}^{(k)} \beta_{j+m} \tag{3.10}$$
$$j = 1,2,\ldots,q, \quad \beta_{q+1}^{(k+1)} = \alpha, \quad k = 0,1,\ldots.$$

The products $L_{N-(q+1)k}^{(N)} \times L_{N-(q+1)k-1}^{(N)} \times \cdots \times L_{N-(q+1)k-q}^{(N)} = [N + O(1)]/\alpha$ may therefore be inserted into the definition (3.7) and the coefficients h_{N+1} and their N dependence may be given the form

$$h_{N+1} = [(\alpha\Delta)^{N\Delta}/\Gamma(N\Delta+1)] b_{N+1}, \quad \Delta = 1/(q+1). \tag{3.11}$$

In the asymptotic region $N \gg 1$, the new coefficients $b_{n+1} \sim b_{N-q}$ exhibit at most the $(q+1)$-periodic oscillations so that the $q+1$ infinite partial summations in Eq. (3.2) exhibit the same exponential behavior $\exp[2f(r)]$ for $r \gg 1$. QED

Let us note that for $q = 0$ the cancellation is impossible $(\psi \equiv \chi_0)$ while, for $q \geqslant 1$, the termination of some particular states (with $b_N = 0$, $N \gg 1$, and $b_{N+1}^{(N)} = 0$; cf. Ref. 2) belongs in fact to the same exceptional phenomena as a nonexistence of our fundamental decomposition (3.5) and will not be discussed here in any detail. Hence the only way how to normalize the harmonic oscillator is the textbook termination of the infinite series $\varphi(r)$ while, for the anharmonic oscillators, the universal mechanism of satisfying the physical boundary conditions is represented by the asymptotic cancellation of χ's. For $q = 1$, this cancellation was shown in Refs. 9 and 12 to follow from the convergence of the analytic continued fractions $L_n = L_n^{(\infty)} = \lim_{N\to\infty} L_n^{(N)}$. Similar limits (ECF quantities[13]) may be defined for $q > 1$ as well—their relevant properties will be described in the next section.

IV. EXTENDED CONTINUED FRACTIONS AND THEIR CONVERGENCE

For large $n = O(N) \gg 1$, the finite approximants $L_n^{(N)}$ depend only on the difference $N - n$ and oscillate—this is a consequence of Eqs. (3.8)–(3.10). Provided that the ECF limit $L_n^{(\infty)}$ exists, these $(q+1)$-periodic oscillations must be suppressed, i.e., $L_n \doteq L_{n+1} = \cdots \doteq L_{n+q} \doteq P_n$ for $O(1) \ll n \ll O(N)$. The P_n's [stationary points of the q-to-one mapping Eq. (3.3)] have to satisfy the algebraic equation

$$P_n = n/(\beta_1 + \beta_2 P_n + \cdots + \beta_{q+1} P_n^q) \tag{4.1}$$

and may be given explicitly by the asymptotic expansion of the type

$$P_n = P_n^{(\epsilon)} = \epsilon\left(\frac{n}{\alpha}\right)^{\Delta} - \frac{\beta_q \Delta}{\alpha} + O\left(\frac{1}{n}\Delta\right), \tag{4.2}$$

$$\Delta = \frac{1}{(q+1)}, \quad \epsilon = \epsilon_m = \exp(2\pi i m \Delta), \quad m = 0,1,\ldots,q.$$

As a consequence, we arrive at an asymptotic estimate

$$L_n^{(N)} = P_n^{(\epsilon)}\left(1 + O\left(\frac{1}{n}\right)\right)_j \quad N \gg n, \tag{4.3}$$

which enables us to accelerate the convergence: We put

$$L_n = P_n^{(\epsilon)} + R_n \tag{4.4}$$

and redefine the initialization, $R_{N+1}^{(N)} = 0$ in the limit $N \to \infty$.

Proposition 2: Provided that $0 \leqslant m_0 \leqslant q-1$ and

$$\beta_{q-m} = 0, \quad m = 0, 1, \ldots, m_0 - 1, \quad \beta_{q-m_0}\epsilon^{q-m_0} > 0, \tag{4.5}$$

the finite ECF approximants $R_n^{(N)} = L_n - P_n^{(\epsilon)}$ defined by the recurrences (3.3) and initialization $R_{N+1}^{(N)} = 0$ are convergent and determine the ECF function $R_n^{(\infty)}$.

Proof: The new nonlinear ECF recurrences for $R_n^{(N)}$ follow directly from an insertion of the definition (4.4) into the old recurrences (3.3). They have the asymptotic form

$$R_n^{(N)} = -\frac{P_n^2 S_n(R_{n+1}^{(N)},\ldots,R_{n+q}^{(N)})[1 + O(1/n)]}{n + P_n S_n(R_{n+1}^{(N)},\ldots,R_{n+q}^{(N)})}, \tag{4.6}$$

$$\begin{aligned} &S_n(R_{n+1},\ldots,R_{n+q}) \\ &= \beta_2(L_{n+1} - P_n) + \cdots + \beta_{q+1}(L_{n+1}\cdots L_{n+q} - P_n^q) \\ &= \sum_{S=1}^{q} \beta_{S+1} P_n^{S-1} \sum_{k=1}^{S} R_{n+k}\Bigg[1 + \sum_{l=1}^{k-1}\sum_{t_1=1}^{l} \\ &\quad \cdots \sum_{t_l=1}^{t_l-1}\prod_{m=1}^{l}\frac{R_{n+l+t_m-m}}{P_n} + O\left(\frac{1}{n}\right)\Bigg], \end{aligned}$$

with the zero (i.e., $O(1/n^{1-\Delta})$) stationary point.

Having in mind the error bounds, we may specify the vicinity of "zero" in the form

$$R_n^{(N)} = O(n^{(1-\Delta)/2}/n^{1-\Delta}) \gg O(1/n^{1-\Delta}). \tag{4.7}$$

In the asymptotic region, this permits the linearization of the ECF recurrences (4.6),

$$\begin{aligned} R_n &= -(t_1 R_{n+1} + \cdots + t_q R_{n+q}) + O(1/n^{1-\Delta}) \\ t_k &= 1 - (\beta_k P_n^k + \cdots + \beta_1 P_n)/n + O(1/n), \\ k &= 1, 2, \ldots, q. \end{aligned} \tag{4.8}$$

Let us consider any initialization

$$R_{N+i}^{(N)} = \rho_i = O(1/N^{(1-\Delta)/2})$$

and rewrite the linearized recurrences (4.8) in the matrix form

$$\begin{pmatrix} 1 & 0 & \cdots & & & \\ t_1 & 1 & 0 & & \cdots & \\ & \cdots & & & & \\ t_{q-1} & \cdots & 1 & 0 & \cdots & \\ t_q & t_{q-1} & \cdots & & & \\ 0 & t_q & \cdots & & & \\ & \cdots & & & & \end{pmatrix}\begin{pmatrix} R_N \\ R_{N-1} \\ \cdots \\ R_{N-q+1} \\ R_{N-q} \\ R_{N-q-1} \\ \cdots \end{pmatrix} = \begin{pmatrix} \tilde\rho_1 \\ \tilde\rho_2 \\ \cdots \\ \tilde\rho_q \\ 0 \\ 0 \\ \cdots \end{pmatrix}, \tag{4.9}$$

where $\tilde\rho_i$, $i = 1,2,\ldots,q$, are some simple linear $O(1/N^{(1-\Delta)/2})$ combinations of ρ_i's. Using the decomposition of the type

$$\begin{aligned} &\begin{pmatrix} 1 & 0 & \cdots & & \\ t_1 & 1 & 0 & \cdots & \\ t_2 & t_1 & 1 & 0 & \cdots \\ & & \cdots & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & & \\ -a_1 & 1 & 0 & \cdots & \\ 0 & -a_1 & 1 & 0 & \cdots \\ & & \cdots & & \end{pmatrix} \\ &\quad \times\cdots\times\begin{pmatrix} 1 & 0 & \cdots & & \\ -a_q & 1 & 0 & \cdots & \\ 0 & -a_q & 1 & 0 & \cdots \\ & & \cdots & & \end{pmatrix} \end{aligned} \tag{4.10}$$

where $a_1, \ldots, a_q$ are complex numbers in general, we may invert easily the left-hand-side matrix in Eq. (4.9) and get

$$\begin{aligned} &\begin{pmatrix} R_N \\ R_{N-1} \\ R_{N-2} \\ \cdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & & \\ a_q & 1 & 0 & \cdots & \\ a_q^2 & a_q & 1 & 0 & \cdots \\ & & \cdots & & \end{pmatrix}\times\cdots \\ &\quad \times\begin{pmatrix} 1 & 0 & \cdots & & \\ a_1 & 1 & 0 & \cdots & \\ a_1^2 & a_1 & 1 & 0 & \cdots \\ & & \cdots & & \end{pmatrix}\times\begin{pmatrix} \tilde\rho_1 \\ \tilde\rho_2 \\ \tilde\rho_3 \\ \cdots \end{pmatrix}. \end{aligned} \tag{4.11}$$

In the next step, we may prove by induction that the right-hand-side product $M_{m,n} = M_{m+k,n+k}$, $k = 1,2,\ldots$, of matrices in Eq. (4.11) is zero for $n > m$ and its nonzero elements may be written in the compact form

$$\begin{aligned} M_{m+k\,m} &= \frac{a_1^{q+k}}{(a_1 - a_2)(a_1 - a_3)\cdots(a_1 - a_q)} \\ &\quad + \frac{a_2^{q+k}}{(a_2 - a_1)(a_2 - a_3)\cdots(a_2 - a_q)} + \cdots \\ &\quad + \frac{a_q^{q+k}}{(a_q - a_1)(a_q - a_2)\cdots(a_q - a_{q-1})}, \\ m &= 1,2,\ldots,\quad k = 0,1,\ldots. \end{aligned} \tag{4.12}$$

Hence, the k dependence of R_{N-k} is given by the powers of a_i^k. In the final step of the proof, it is sufficient to show that $\|a_i\| < 1$, $i = 1,2,\ldots,q$.

By straightforward algebra, we get $t_1 = -a_1 - a_2 - \cdots - a_q, \ldots, t_q = (-1)^q a_1 a_2 \cdots a_q$ in Eq. (4.10) so that a_i's are equal to the roots of the algebraic equation $x^q + t_1 x^{q-1} + \cdots + t_q = 0$ which may be written in the form

$$\begin{aligned} &\frac{1 - x^{q+1}}{1 - x} - \frac{\beta_1 P_n}{n}\frac{1 - x^q}{1 - x} - \cdots - \frac{\beta_q P_n^q}{n}\frac{1 - x}{1 - x} \\ &\quad = O\left(\frac{1}{n}\right), \quad x \neq 1. \end{aligned} \tag{4.13}$$

Now, assuming that n is sufficiently large and Eq. (4.5) is satisfied, we get

$$1 - x^{q+1} = \frac{\beta_{q-m_0} P_n^{q-m_0}}{n}(1 - x^{m_0+1})[1 + O(1/n^{\Delta})]. \tag{4.14}$$

Since the right-hand-side of Eq. (4.14) may be made arbitrarily small by the choice of $n \gg 1$, we may write

$$a_l = \epsilon_l(1+\eta), \quad \epsilon_l = e^{2\pi i \Delta l},$$
$$1 \leqslant l \leqslant q, \quad \|\eta\| \ll 1. \tag{4.15}$$

Inserting this expression into Eq. (4.14), we obtain

$$\eta = \beta_{q-m_0} P_n^{q-m_0}(\epsilon_l^{m_0+1} - 1)\Delta / n + O(\eta^2). \tag{4.16}$$

In accord with our assumption (4.5), the real part of η is always negative so that $\mathrm{Re}(1+\eta) < 1$ and $\|a_l\| < 1$ for $n < \infty$. This completes the proof since, in accord with Eq. (4.15), the denominators in Eq. (4.12) are all nonzero, and $R_{N-k} = O(\max\|a_i^k\|) = O(\exp(-\mathrm{const}k/n^\Delta))$ may be made arbitrarily small within the $N^\Delta \ll k \ll N$ asymptotic region. QED

(i) On the particular $q = 1$ example, the content of Proposition 2 may be illustrated very simply—the ECF degenerates to an ordinary continued fraction,[15] the choice between the two roots of the quadratic Eq. (4.1) is given by the rule $\epsilon = \mathrm{sgn}\,\beta_q$ and both the initializations $L_{N+1}^{(N)} = 0$ and $R_{N+1}^{(N)} = 0$ are equivalent so that the transition $L_n \to R_n$ [Eq. (4.4)] is merely an acceleration of convergence.

(ii) For $q > 1$, the values of $P_n^{(\epsilon)}$ may be complex, causing a divergence of $L_n^{(\infty)}$, which need not necessarily be accompanied by the divergence of $R_n^{(\infty)}$ *a priori*. The exact proofs of equivalence would be extremely complicated even for the real ϵ's ($\epsilon = \pm 1$)—an example of a $q = 2$—analog of Proposition 2 for $L_n^{(\infty)}$'s may be found in Ref. 10. Fortunately, the present weaker convergence of $R_n^{(\infty)}$ is fully sufficient in the context of the next section.

V. ENERGIES OF THE BOUND STATES

The determinantal equation (2.2) for h_{n+1} is less suitable in the asymptotic region—let us switch back to the original recurrences. They may be given the form of the infinite-dimensional matrix equation

$$Q(\infty)\begin{pmatrix} h_1 \\ h_2 \\ \cdots \end{pmatrix} = 0. \tag{5.1}$$

Assuming that the product decomposition (3.5) of $Q(N)$ exists also in the limit $N \to \infty$, we obtain the fully equivalent form of Eq. (5.1):

$$Y(\infty)Z(\infty)\begin{pmatrix} h_1 \\ h_2 \\ \cdots \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \cdots \end{pmatrix}, \tag{5.2}$$

where the right-hand-side vector is defined in terms of the convergent ECF's by the formulas $X(\infty)\omega = 0$, i.e.,

$$\omega_k = h_1 \prod_{j=1}^{k} L_j^{(\infty)}, \quad k = 1, 2, \ldots, \tag{5.3}$$

and it can, of course, be nonzero in the $N = \infty$ case. Hence we do not need to require *a priori* the zero Hill determinant $\det[Q(\infty)] = 0$, contrary to the intuitive expectations of Ref. 9, where Eq. (5.1) was misinterpreted as an eigenvalue problem.

The validity of the $g_2 > 0$ part of the Singh's results is almost surprising—the physical energies coincide indeed with the zeros of the Hill determinant.[12] Let us show that this is a peculiarity of all the SFAO potentials [= FAO's restricted by the "superconfinement" or "subdominant positivity" requirement (1.5)].

Theorem 2: When we put $\epsilon = 1$ in the ECF definition (4.3), the roots E_n of the "secular" equation

$$1/R_1^{(\infty)}(E) = 0 \tag{5.4}$$

determine the physical SFAO bound-state energies.

Proof: First, we have to prove the convergence of the ECF quantities $R_n^{(\infty)}$. This is an easy task because Eq. (1.5) implies Eq. (4.5), and we may apply Proposition 2. Next, we may factor out all the dominant n-dependence in the recurrences Eq. (5.1). In detail, estimates of Eq. (4.3) and

$$L_n U_n^{(q-k)} = 4n\alpha P_n^{k+1} + 4n\,\beta_q P_n^k(1 - k\Delta) + \cdots,$$
$$k = 0, 1, \ldots, q-1 \tag{5.5}$$

make it possible for us to rewrite Eq. (5.2), i.e., the Taylor recurrences

$$h_k B_{k-1} / L_k^{(\infty)} = \omega_k + \sum_{i=1}^{\min(q,k-1)} U_k^{(i)} h_{k-i}, \quad k = 1, 2, \ldots, \tag{5.6}$$

in the $k \gg 1$ asymptotic form. Indeed, using Eq. (3.11) and the Stirling formula in the form

$$[(k/\Delta)!]^\Delta = k!\Delta^{-k-\Delta/2}(2\pi k)^{(\Delta-1)/2}[1 + O(1/k)], \tag{5.7}$$

we obtain the relation

$$\sum_{j=0}^{q} \epsilon^j b_{N+j} = \frac{\mathrm{const}}{L_1^{(\infty)}}\left[1 + O\left(\frac{1}{N^\Delta}\right)\right], \tag{5.8}$$

reflecting the contribution of ω_N's, the size of which is comparable with the b_N's.

Finally, the $q+1$ b's may be considered almost constant in the $N \gg 1$ asymptotic region, $b_{N+j} \sim b_j$. Hence, in accord with Eq. (5.8) and for the particular value $\epsilon = 1$, the sign of the superposition of χ_j's for $r \gg 1$ is determined by the sign of $L_1(E)$. This sign changes exactly at the root E_n of Eq. (5.4) since the poles of $L_1(E)$ and $R_1(E)$ coincide. With respect to the standard oscillation theorems,[14] a new node appears in ψ whenever E crosses E_n—the value E_n coincides precisely with the SFAO bound-state energy. QED

(i) The proof given above is the main result of our paper. The Theorem 2 may be interpreted as a rigorous foundation of the Hill-determinant interpretation[9] of the eigenvalue problem with the following strong warning: The physical energies E_n in Ref. 9 coincide with the Hill-determinant zeros by mere chance—for the non-SFAO potentials with $g_{2q-m_0} < 0$, the Singh-type interpretation of $\det[E - Q(\infty)]$ is completely misleading and gives the unphysical energies.[12]

(ii) Indeed, for $\epsilon \neq 1$, there is no relation between the root of Eqs. (5.4) and (1.4) and the asymptotic behavior of ψ. The superposition of b_j's or χ_j's is in no way related to Eq. (5.8): For $\epsilon = -1$ and $q = 1$, the asymptotic cancellation of χ_0 and χ_1 is even minimal[12] at $E = E_n$—the asymptotically *decreasing* exponential component of $\psi(r)$ seems to be suppressed. Thus, the choice of $\epsilon = +1$ in Theorem 2 is unique, and we cannot remove the "superconfinement" restriction within the present ECF framework.

(iii) The main merit of our "secular" Eqs. (5.4) or (1.4) is their analytic and compact structure. Nevertheless, in contrast to the original expectations,[9] the purely numerical exploitation of Eq. (5.4) is also possible in principle.

From this numerical point of view, the transition $L_n^{(N)} \to R_n^{(N)}$ is extremely useful. When we consider the illustrative example $\beta_i = 1$ with $L_{N+1}^{(N)} = 0$ and

$$\begin{aligned} L_{N-(q+1)k}^{(N)} &= \frac{q!k!}{(q+k)!} N, \\ L_{N-(q+1)k-j}^{(N)} &= 1 + \frac{k}{q+1-j}, \\ j &= 1, 2, ..., q, \quad k = 0, 1, ... \end{aligned} \tag{5.9}$$

[cf. Eq. (3.9)], we see that the use of $R_n^{(N)}$'s eliminates the large $(q+1)$-periodic oscillations $L_n^{(N)} - L_{n+1}^{(N)} = O(n)$ which survive in the asymptotic region for at least as many as $O(N^{\Delta})$ iterations of the old recurrences, Eq. (3.3).

(iv) Of course, the smaller oscillations are present in $R_n^{(N)}$'s as well. Fortunately, they decrease for the higher cutoffs N. Moreover, they may systematically be suppressed either by the next (k th) substraction of the form

$$\begin{aligned} &R_n[k+1] = R_n[k] - P_n[k+1], \\ &R_n[0] = L_n^{(N)}, \quad P_n[1] = P_n^{(1)}, \\ &R_n[1] = R_n^{(N)}, \cdots, \end{aligned} \tag{5.10}$$

or by an averaging over the $q+1$ neighboring cutoffs $N = N_0 + i$, $i = 1, 2, ..., q+1$. The efficiency of the latter type of "smoothing" was found empirically in a somewhat related context in Ref. 6. When combined with the subtractions (5.10), it was tested in Ref. 10 for $q = 2$ and proved to be comparable even with the specialized methods.

VI. CONCLUDING REMARKS

We have proved that Eq. (5.4) defines the energies of the anharmonic oscillators which belong to the SFAO class, i.e., which are restricted by Eq. (1.5). We have also proved that the ECF representation of the underlying "Green's" function $R_1^{(\infty)}(E)$ is convergent. In the conclusion we would like to add the following remarks:

(i) Rather surprisingly, the acceleration of convergence (quasiequivalence $L_n \to R_n \to R_n[2] \to \cdots$) makes the analytic ECF formalism well suited even for the numerical computations.

(ii) The immanent "superconfinement" restriction resembles the similar property of the JWKB method[16] and cannot be removed without the deep modifications of the ECF method. The possibility to revert the ECF recurrences is under current investigation at present.

(iii) The partial (SFAO) coincidence of the present secular equation with its Hill-determinant precursors[9,10] may be characterized as a lucky chance and attributed to the very special type of the Schrödinger boundary conditions imposed on $\psi(r)$ at $r \to \infty$.

(iv) Contrary to the *a priori* expectations based on the analogy with the $q = 0$ harmonic oscillator, the essential technicality leading to the successful completion of the standard power-series method is not only the factorization of the wavefunction $\psi = \exp(-\text{polynomial}) \times \text{power series}$, but also the ECF shortening of the Taylor recurrences to their final $(q+1)$-term form. The SFAO restriction (or the ECF convergence from the more formal point of view) is in fact just the condition of stability of the underlying decomposition $Q(\infty) = X(\infty)Y(\infty)Z(\infty)$.

(v) The exceptional (terminating) solutions[2] which correspond to the singular cases in our formulas ($L_k \to \infty$ etc.) are to be interpreted as limits of the neighboring fully regular cases. This is compatible with the spirit of the classical theory of the analytic continued fractions.[15]

(vi) The classification of the FAO and SFAO forces by means of the "complexity" q, "fractionality" p, and "subdominant anharmonicity" m_0 is an interesting byproduct of our considerations. It leads to an unusual partial ordering $q = 0, q = 1, \cdots$ of some standard anharmonic models.

[1] C. A. Aragao de Carvalho, Nucl. Phys. B **119**, 401 (1977); J. R. Fulco and L. Masperi, Phys. Rev. D **19**, 690 (1979).
[2] G. P. Flessas and K. P. Das, Phys. Lett. A **78**, 19 (1980); E. Magyari, Phys. Lett. A **81**, 116 (1981); M. Znojil, J. Phys. A: Math. Gen. **15**, 2111 (1982).
[3] F. T. Hioe, D. MacMillan, and E. W. Montroll, Phys. Rep. **43**, 305–35 (1978).
[4] C. Quigg and J. L. Rosner, Phys. Rep. **56**, 167 (1979).
[5] B. Simon, Ann. Phys. (N.Y.) **58**, 76 (1969).
[6] J. L. Richardson and R. Blankenbecler, Phys. Rev. D **19**, 496 (1979).
[7] I. G. Halliday and P. Suranyi, Phys. Rev. D **21**, 1529 (1980); M. Znojil, Phys. Rev. D **24**, 903 (1981).
[8] S. Graffi and V. Grecchi, Lett. Nuovo Cimento **12**, 425 (1975).
[9] V. Singh, S. N. Biswas, and K. Datta, Phys. Rev. D **18**, 1901 (1978); Lett. Math. Phys. **3**, 73 (1979); **4**, 131 (1980)—with A. Rampal; D. P. Datta and S. Mukherjee, J. Phys. A: Math. Gen. **13**, 3161 (1980).
[10] M. Znojil, Lett. Math. Phys. **5**, 405 (1981).
[11] M. Znojil, submitted to Lett. Math. Phys.
[12] Private communications; also M. Znojil and G. P. Flessas, submitted to J. Phys. A.
[13] M. Znojil, J. Phys. A: Math. Gen. **9**, 1 (1976); **13**, 2375 (1980).
[14] E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[15] H. S. Wall, *Analytic Theory of Continued Fractions* (Van Nostrand, New York, 1948).
[16] T. Arai, Progr. Theor. Phys. **59**, 311 (1978).

On the number of unnatural parity bound states of the H^- ion[a)]

H. Grosse
Institut für Theoretische Physik, Universität Wien, A-1090 Wien, Austria
L. Pittner
Institut für Theoretische Physik, Universität Graz, Graz, Austria

(Received 22 December 1981; accepted for publication 5 March 1982)

We show rigorously that the H^- ion possesses exactly one (three times degenerate) bound state in the unnatural parity sector.

PACS numbers: 03.65.Ge, 03.65.Db

1. INTRODUCTION

Although the existence of the discrete ground state $(1s)^2\ {}^1S$ of the nonrelativistic H^- ion at an energy of -0.52775 a.u. (atomic units) was proved fifty years ago (Bethe[1], Hylleraas[2]) and also relativistic corrections (except the Lamb shift) have been calculated with high accuracy (Pekeris[3]), the rigorous proof that the discrete bound state is the only one has been performed only a few years ago (Hill[4]), including also corrections due to the finite proton mass.

The point spectrum of the H^- ion is of considerable interest both from the physical and from the mathematical viewpoint. The bound state of this ion accounts for the long-wavelength continuous absorption in the solar atmosphere. The H^- ion is the only negative ion for which rigorous lower bounds on energy eigenvalues and upper bounds on the number of eigenvalues can be derived with present techniques, except for results concerning the finiteness of the discrete spectrum (Antonets, Zhislin, and Shereshevskii[5]), the absence of the discrete spectrum in the special case of no symmetry and large ionization (Ruskai[6]) and some results for a certain class of potentials (Grosse[7] and Klaus and Simon[8]).

In the fixed (infinite mass) proton approximation, the Hamilton operator of the H^- ion reads

$$H = H_0 + V, \quad H_0 = \frac{p_1^2}{2} - \frac{1}{r_1} + \frac{p_2^2}{2} - \frac{1}{r_2},$$
$$V = \frac{1}{r_{12}}, \quad r_{12} = |\mathbf{x}_1 - \mathbf{x}_2|, \tag{1.1}$$

defined on the usual dense domain of self-adjointness in $\mathscr{H} = L^2(\mathbb{R}^6)$. The point spectrum of the unperturbed Hamilton operator H_0 consists of the eigenvalues $-(1/n_1^2 + 1/n_2^2)/2$ with integers n_1 and n_2 $(n_i \geqslant 1)$, which are $4n_1^2 n_2^2$ or $2n_1^2(n_1^2 - 1)$ times degenerate for $n_1 \neq n_2$ or $n_1 = n_2$, respectively. Since the total wave function must be antisymmetric with respect to the exchange of spin and spatial coordinates, we may restrict the Hamilton operator H to the symmetric sector $\mathscr{H}_s$ and the antisymmetric sector $\mathscr{H}_a$ of $\mathscr{H}$, respectively.

The spectrum of the unperturbed Hamilton operator H_0 is highly degenerate because of its large commutant. Each combination of electron angular momenta $\mathbf{L}_1$ and $\mathbf{L}_2$ to $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2$ with $|l_1 - l_2| \leqslant l \leqslant l_1 + l_2$ determines the parity $P = P_1 P_2$, $P_i = (-1)^{l_i}$, $i = 1,2$, which need not be equal to $(-1)^l$. States of the H^- ion with parity $P = (-1)^l$ or $(-1)^{l+1}$ will be called states of natural or unnatural parity. Each of the two sectors $\mathscr{H}_s$ and $\mathscr{H}_a$ is then decomposed into the two subspaces of natural and unnatural parity, so that the Hamiltonian H is reduced by each of these four subspaces of $\mathscr{H}$, because P commutes with H.

Whereas natural parity states with energy above $-\frac{1}{2}$ can decay into one free electron and the hydrogen ground state (Auger effect), this decay is impossible for states of unnatural parity below $-\frac{1}{8}$ because of the natural parity of the final state. Other decays, e.g., radiative transitions of unnatural parity states may occur, although implying larger life times than Auger transitions.

Let us denote the unnatural parity subspaces of the symmetric and antisymmetric sector by $\mathscr{H}_s^u$ and $\mathscr{H}_a^u$, respectively. The ground state of H_0 restricted to $\mathscr{H}_a^u$ with binding energy $-\frac{1}{4}$ is three times degenerate, because the lowest angular momentum eigenvalues yielding unnatural parity are $l_1 = l_2 = l = 1$, i.e., both electrons are in $(2p)$ states; the lowest states within $\mathscr{H}_s^u$ have energy $-\frac{13}{72}$. The lowest threshold of H restricted to the unnatural parity sector lies at the value $-\frac{1}{8}$ and corresponds to one free electron and a $(2p)$ hydrogen atom.

In the following we shall transpose some method of estimating the number of discrete energy eigenvalues from the natural parity subspace to unnatural parity. Variational computations (Drake[9]) have shown the existence of at least one discrete energy eigenstate with unnatural parity, lying below -0.125350 a.u. with an estimated root-mean square radius of 20.3 a.u.

At first it becomes obvious that Hill's method can be applied only if the ground state onto which one would like to project in order to bind from below the Hamilton operator is nondegenerate, which is not true in the unnatural parity subspaces $\mathscr{H}_s^u$ and $\mathscr{H}_a^u$. But next we observe that to each choice of total angular momentum, i.e., l and m fixed, there exists exactly one possible choice for l_2 if $l_1 = 1$, namely $l_2 = l$, in order to obtain an unnatural parity state, and analogously for l_1 exchanged with l_2; or, in other words, coupling $Y_{1,m_1}(\Omega_1)$ with $Y_{l,m-m_1}(\Omega_2)$ with $l \geqslant 1$ yields states with total angular momentum $l-1$, l, and $l+1$, but only the states with total angular momentum l have unnatural parity; therefore, if one of the two electrons has angular momentum $l_i = 1$, there exists exactly one linear combination of products

a)Work supported in part by the "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich", Project. Nr. 3569.

of spherical harmonics, which we shall denote by $\mathscr{Y}^{(i)}_{l,m}(\Omega_1,\Omega_2)$, with unnatural parity and total angular momentum eigenvalues l, m,

$$\mathscr{Y}^{(1)}_{l,m}(\Omega_1,\Omega_2)=\sum_{m_1=0,\pm 1} c^{1,m_1}_{l,m}Y_{1,m_1}(\Omega_1)Y_{l,m-m_1}(\Omega_2), \quad (1.2)$$
$$\mathscr{Y}^{(2)}_{l,m}(\Omega_1,\Omega_2)=\mathscr{Y}^{(1)}_{l,m}(\Omega_2,\Omega_1),$$

with appropriate Clebsch–Gordan coefficients.

The lowest possible choice of total angular momentum in the unnatural parity subspace, i.e., $l=1$, which leads to an antisymmetric combination of spherical harmonics with respect to the exchange $\Omega_1\leftrightarrow\Omega_2$, namely

$$\begin{aligned}
&\mathscr{Y}^{(1)}_{1,1}(\Omega_1,\Omega_2)\\
&\quad=\frac{1}{\sqrt{2}}\,[Y_{1,1}(\Omega_1)Y_{1,0}(\Omega_2)-Y_{1,0}(\Omega_1)Y_{1,1}(\Omega_2)]\\
&\mathscr{Y}^{(1)}_{1,0}(\Omega_1,\Omega_2)\\
&\quad=\frac{1}{\sqrt{2}}\,[Y_{1,1}(\Omega_1)Y_{1,-1}(\Omega_2)-Y_{1,-1}(\Omega_1)Y_{1,1}(\Omega_2)] \quad (1.3)\\
&\mathscr{Y}^{(1)}_{1,-1}(\Omega_1,\Omega_2)\\
&\quad=\frac{1}{\sqrt{2}}\,[Y_{1,0}(\Omega_1)Y_{1,-1}(\Omega_2)-Y_{1,-1}(\Omega_1)Y_{1,0}(\Omega_2)],
\end{aligned}$$

is much more difficult to handle both in the symmetric and antisymmetric sector of $\mathscr{H}$. Therefore we shall start with the simpler cases with total angular momentum $l\geqslant 2$.

2. TOTAL ANGULAR MOMENTUM $l\geqslant 2$

We start with the decomposition

$$\mathscr{H}=L^2(\mathbb{R}^6)=L^2(\mathbb{R}_+;r_1^2dr_1)\otimes L^2(\mathbb{R}_+;r_2^2dr_2)$$
$$\otimes L^2(d\Omega_1)\otimes L^2(d\Omega_2) \quad (2.1)$$

and restrict our investigations to one choice of total angular momentum, i.e., we keep $l\geqslant 2$ and m fixed. Using the projectors

$$P^{(i)}_{l,m}=|\mathscr{Y}^{(i)}_{l,m}\rangle\langle\mathscr{Y}^{(i)}_{l,m}|,\quad i=1,2,$$
$$\text{on}\quad L^2(d\Omega_1)\otimes L^2(d\Omega_2), \quad (2.2)$$

the main idea consists in trying to minorize the potential V by using the well-known projection method (Hill,[4] and Thirring[10]) and projecting onto the subspace $R^{(i)}_{l,m}\cdot\mathscr{H}$ where

$$R^{(1)}_{l,m}=(P_\Phi\otimes 1)\otimes P^{(1)}_{l,m}$$

and

$$R^{(2)}_{l,m}=(1\otimes P_\Phi)\otimes P^{(2)}_{l,m},$$

with

$$P_\Phi=|\Phi\rangle\langle\Phi|\ \text{ on } L^2(\mathbb{R}_+,r^2dr),$$

$$\Phi(r)=\frac{r}{\sqrt{24}}\,e^{-r/2}. \quad (2.3)$$

The operators $R^{(i)}_{l,m}$ project $i=1,2$ onto a p state, leaving the radial part of the wave function of the other particle free in such a way as to obtain an unnatural parity two-particle state with total angular momentum l,m.

With the aid of the new projectors

$$Q^{(i)}_{l,m}=V^{-1/2}R^{(i)}_{l,m}(R^{(i)}_{l,m}V^{-1}R^{(i)}_{l,m})^{-1}R^{(i)}_{l,m}V^{-1/2},$$
$$i=1,2, \quad (2.4)$$

(note V is positive so $V^{1/2}$ exists) one can use the obvious operator inequalities

$$V\geqslant V^{1/2}Q^{(i)}_{l,m}V^{1/2},\quad i=1,1, \quad (2.5)$$

to obtain lower bounds (Bazley,[11] Thirring,[10] and Bazley and Fox[12]),

$$V\geqslant R^{(i)}_{l,m}(R^{(i)}_{l,m}V^{-1}R^{(i)}_{l,m})^{-1}R^{(i)}_{l,m},\quad i=1,2, \quad (2.6)$$

which imply minorization of the ordered discrete energy eigenvalues via the min-max principle. Then one obtains the operator estimates

$$V\geqslant(P_\Phi\otimes 1)V_{l,m}(r_2)\otimes P^{(1)}_{l,m}$$

and $\quad (2.7)$

$$V\geqslant(1\otimes P_\Phi)V_{l,m}(r_1)\otimes P^{(2)}_{l,m}$$

with the effective potentials

$$V^{-1}_{l,m}(r_2)=\int_0^\infty dr_1r_1^2\Phi^2(r_1)$$
$$\times\int d\Omega_1d\Omega_2|\mathscr{Y}^{(1)}_{l,m}(\Omega_1,\Omega_2)|^2|\mathbf{x}_1-\mathbf{x}_2|. \quad (2.8)$$

Using the expansion of $|\mathbf{x}_1-\mathbf{x}_2|$ into products of spherical harmonics and definition (1.2) one easily evaluates the potentials $V_{l,m}(r)$, but for total angular momentum $l\geqslant 3$ one can exclude bound states with the aid of rather crude estimates.

A. Case $l\geqslant 3$

Since in this case the centrifugal force causes an additional repulsive potential, we may try to replace first the potential V by the lower bound

$$\frac{1}{|\mathbf{x}_1-\mathbf{x}_2|}\geqslant\frac{1}{r_1+r_2}=V_L. \quad (2.9)$$

Then the same procedure as above can be applied to the potential V_L, now yielding

$$V\geqslant(P_\Phi\otimes 1)U(r_2)\otimes P^{(1)}_{l,m}+(1\otimes P_\Phi)U(r_1)\otimes P^{(2)}_{l,m}, \quad (2.10)$$

with the potential U given by

$$U^{-1}(r_2)=\int_0^\infty dr_1r_1^2\Phi^2(r_1)(r_1+r_2)=r_2+5, \quad (2.11)$$

where we have used the fact that $P^{(1)}_{l,m}\cdot P^{(2)}_{l,m}=0$ for $l\geqslant 2$ and especially the angular independence of the lower bound (2.9).

Insertion of the lower bound (2.10) into the Hamilton operator then leads to the estimate

$$H\,|_{\mathscr{H}^u}\geqslant h(r_1)R^{(2)}_{l,m}+R^{(1)}_{l,m}h(r_2)$$
$$-\tfrac{1}{8}(R^{(1)}_{l,m}+R^{(2)}_{l,m})-\tfrac{1}{9}(R^{(1)}_{l,m}+R^{(2)}_{l,m})^{\perp} \quad (2.12)$$

$$h(r_i)=\frac{p_i^2}{2}-\frac{1}{r_i}+U(r_i),\quad i=1,2,$$

where $\mathscr{H}^u$ denotes the unnatural parity sector of $\mathscr{H}$. In order to exclude the existence of eigenstates of H with unnatural parity and total angular momentum $l\geqslant 3$ below the energy $-\frac{1}{8}$, it therefore suffices to prove that the one-particle

Schrödinger operator acting in the Hilbert space $L^2(\mathbb{R}_+,dr)$

$$-\frac{1}{2}\frac{d^2}{dr^2}+\frac{l(l+1)}{2r^2}-\frac{1}{r}+\frac{1}{r+5}\geqslant 0$$

$$\text{for}\quad l\geqslant 3 \tag{2.13}$$

should be positive, an estimate of a rather trivial kind.

B. Case $l = 2$

Here we do not use (2.9) but project the full interaction onto the subspaces spanned by

$$\mathscr{Y}_{\pm}(\Omega_1,\Omega_2)=\frac{1}{\sqrt{2}}\,(\mathscr{Y}^{(1)}_{2,m}(\Omega_1,\Omega_2))\,\mathscr{Y}^{(2)}_{2,m}(\Omega_1,\Omega_2),$$

$$P_{\pm}=|\mathscr{Y}_{\pm}\rangle\langle\mathscr{Y}_{\pm}|. \tag{2.14}$$

Here we suppress the index m since all is clearly independent of it. Using the fact that

$$P_{\pm}V^{-1}P_{\mp}=0 \tag{2.15}$$

we obtain a lower bound as follows:

$$V\geqslant V_+P_++V_-P_-$$

with

$$V_{\pm}^{-1}(r_1,r_2)=\int d\Omega_1 d\Omega_2|\mathscr{Y}_{\pm}(\Omega_1,\Omega_2)|^2|\mathbf{x}_1-\mathbf{x}_2|. \tag{2.1.6}$$

Explicit calculation of $V_{\pm}$ is trivial by noting the expansion of V^{-1} into products of spherical harmonics:

$$|\mathbf{x}_1-\mathbf{x}_2|=\sum_{l=0}^{\infty}k_l(r_1,r_2)\sum_{m=-l}^{l}Y_{l,m}(\Omega_1)Y^*_{l,m}(\Omega_2)\cdot 4\pi, \tag{2.17}$$

where the kernels k_l are explicitly given by

$$(2l+1)k_l(r_1,r_2)=\frac{1}{2l+3}\frac{r_<^{l+2}}{r_>^{l+1}}-\frac{1}{2l-1}\frac{r_<^{l}}{r_>^{l-1}},$$

$$l=0,1,2,\ldots, \tag{2.18}$$

$$r_<=\min(r_1,r_2),\quad r_>=\max(r_1,r_2).$$

Using, for instance, the expression for $\mathscr{Y}^{(1)}_{2,0}$:

$$\mathscr{Y}^{(1)}_{2,0}(\Omega_1,\Omega_2)=\frac{1}{\sqrt{2}}\,[Y_{1,1}(\Omega_1)Y_{2,-1}(\Omega_2)-Y_{1,-1}(\Omega_1)Y_{2,1}(\Omega_2)] \tag{2.19}$$

in Eqs. (2.14) and (2.16) gives

$$V_{\pm}^{-1}(r_1,r_2)=(k_0-k_2\mp\tfrac{3}{5}k_1\pm\tfrac{3}{5}k_3)(r_1,r_2). \tag{2.20}$$

Next we use $V_-\geqslant V_+$ in Eq. (2.16) and get

$$V\geqslant V_+\cdot(P_++P_-). \tag{2.21}$$

Adding H_0 to both sides of (2.21) and taking expectation values in states of the form

$$\psi_{\pm}(\mathbf{x}_1,\mathbf{x}_2)=\frac{1}{\sqrt{2}}\,[\mathscr{Y}^{(1)}_{2,m}(\Omega_1,\Omega_2)\chi(r_1,r_2)\pm\mathscr{Y}^{(2)}_{2,m}(\Omega_1,\Omega_2)\chi(r_2,r_1)], \tag{2.22}$$

$$\chi(r_1,r_2)=\Phi(r_1)f(r_2),\quad f\in L^2(\mathbb{R}_+;r_2^2dr^2),$$

shows that for both the symmetric and antisymmeric subspace one obtains the same lower bound

$$\langle\psi_{\pm},V\psi_{\pm}\rangle\geqslant\int_0^\infty dr_1r_1^2\int_0^\infty dr_2r_2^2|\chi(r_1,r_2)|^2V_+(r_1,r_2). \tag{2.23}$$

Projecting once more particle one or two onto the radial part of the p wave function reduces the problem of excluding bound states below energy $-\frac{1}{8}$ to the question of positivity of the one-particle operator on $L^2(\mathbb{R}_+,dr)$

$$-\frac{1}{2}\frac{d^2}{dr^2}+\frac{3}{r^2}-\frac{1}{r}+U_+(r)\geqslant 0, \tag{2.24}$$

where the repulsive potential $U_+(r)$ is defined through

$$U_{+}^{-1}(r_2)=\int_0^\infty dr_1r_1^2\Phi^2(r_1)V_{+}^{-1}(r_1,r_2). \tag{2.25}$$

A somewhat tedious but straightforward calculation leads to the explicit expression for $U_+(r)$

$$U_{+}^{-1}(r)=r+1+\frac{12}{r}-\frac{12}{r^2}-\frac{48}{r^3}+\frac{144}{r^4}-e^{-r}\left(\frac{12}{r^2}+\frac{96}{r^3}+\frac{144}{r^4}\right). \tag{2.26}$$

To prove (2.24) it is enough to show positivity for a potential which is a lower bound to $U_+(r)$. We take

$$U_+(r)\geqslant\theta(r-3)\cdot(r+1+12/r)^{-1}, \tag{2.27}$$

then the total potential in (2.24) turns out to be positive.

3. TOTAL ANGULAR MOMENTUM $l = 1$

In order to study this case we use the fact that there is one, and only one, angular momentum configuration for fixed total magnetic quantum number m for which at least one of the two electrons is in a p state, namely with $l_1=l_2=l=1$ [see (1.3)], and start replacing the potential V by the lower bound

$$V\geqslant V_L(r_1,r_2)P_{1,m}, \tag{3.1}$$

where the operator $P_{1,m}:=P^{(1)}_{1,m}$ projects onto $\mathscr{Y}_{1,m}:=\mathscr{Y}^{(1)}_{1,m}$ and clearly $P^{(1)}_{1,m}=P^{(2)}_{1,m}$ since $\mathscr{Y}^{(1)}_{1,m}=-\mathscr{Y}^{(2)}_{1,m}$. The lower bound potential V_L is given by

$$V_L^{-1}(r_1,r_2)=\int d\Omega_1\int d\Omega_2|\mathscr{Y}_{1,m}(\Omega_1,\Omega_2)|^2|\mathbf{x}_1-\mathbf{x}_2|$$

$$=(k_0-k_2)(r_1,r_2), \tag{3.2}$$

where the kernels k_l have been defined in (2.18). From now on we consider the Hamiltonian

$$H_L:=H_0+V_LP_{1,m}\leqslant H \tag{3.3}$$

restricted to the subspace $P_{1,m}\mathscr{H}$ with $l=1$ and fixed $m=0,\pm 1$. Without always mentioning $P_{1,m}$ we shall work in $L^2(\mathbb{R}_+,r_1^2dr_1)\times L^2(\mathbb{R}_+,r_2^2dr_2)$. Projecting next one of the two electrons onto the state Φ, we obtain

$$V_L\geqslant V_L^{1/2}(Q_1\vee Q_2)V_L^{1/2}, \tag{3.4}$$

where the projection operators Q_1 and Q_2 are given by

$$Q_1=V_L^{-1/2}(P_\Phi\otimes 1)W(r_2)V_L^{-1/2},\quad P_\Phi=|\Phi\rangle\langle\Phi|, \tag{3.5}$$

$$Q_2=V_L^{-1/2}W(r_1)(1\otimes P_\Phi)V_L^{-1/2},$$

and the new potential W is given by the integral

$$W^{-1}(r_2) = \int_0^\infty dr_1 r_1^2 \Phi^2(r_1) V_L^{-1}(r_1,r_2)$$
$$= r_2 + \frac{12}{r_2} - \frac{48}{r_2^3}$$
$$+ e^{-r_2}\left(1 + \frac{12}{r_2} + \frac{48}{r_2^2} + \frac{48}{r_2^3}\right). \quad (3.6)$$

$Q_1 \vee Q_2$ in Eq. (3.4) projects onto the subspace spanned by $Q_1\mathscr{H}$ and $Q_2\mathscr{H}$. In order to simplify we use the expansion

$$Q_1 \vee Q_2 = \frac{1}{2}(Q_1 + Q_2) + \frac{1}{2} \times \sum_{n=0}^{\infty} (Q_1^{\perp}(Q_2Q_1)^n Q_2 Q_1^{\perp} + Q_2^{\perp}(Q_1Q_2)^n Q_1 Q_2^{\perp}) \quad (3.7)$$

and, since all terms on the right hand side are positive, we may take into account only the first term which gives

$$Q_1 \vee Q_2 \geqslant Q_1 + Q_2 - Q_1Q_2 - Q_2Q_1 + \tfrac{1}{2}(Q_1Q_2Q_1 + Q_2Q_1Q_2). \quad (3.8)$$

To handle the contributions coming from

$$R = V_L^{1/2}(Q_1Q_2 + Q_2Q_1)V_L^{1/2} \quad (3.9)$$

we shall use Hill's procedure (Hill[4]) of rewriting the inverse Coulomb potential into two rank one operators and an integral over separable terms, i.e.,

$$2|\mathbf{x}_1 - \mathbf{x}_2| = (1 + r_1)(1 + r_2) - (1 - r_1)(1 - r_2) - \frac{1}{\pi}\int d^3x' \left(\frac{1}{|\mathbf{x}_1 - \mathbf{x}'|} - \frac{1}{r'}\right) \times \left(\frac{1}{|\mathbf{x}_2 - \mathbf{x}'|} - \frac{1}{r'}\right), \quad (3.10)$$

a representation which is easily checked by noting that the linear potential is the Green's function of $(-\Delta)^2$. Projecting (3.10) onto a particular angular momentum state yields for k_0 and k_2 entering into Eq. (3.2)

$$k_0(r_1,r_2) = \frac{1}{2}(1 + r_1)(1 + r_2) - \frac{1}{2}(1 - r_1)(1 - r_2) - 2\int_0^\infty dr' \left(\frac{r'}{r_{1>}} - 1\right)\left(\frac{r'}{r_{2>}} - 1\right), \quad (3.11)$$
$$k_2(r_1,r_2) = -\frac{2}{25}\int_0^\infty dr' r'^2 \frac{r_{1<}^2}{r_{1>}^3}\frac{r_{2<}^2}{r_{2>}^3};$$

$r_{i>} = \max(r_i,r')$, $r_{i<} = \min(r_i,r')$, $i = 1,2$. In order to use these separable kernels we write down the matrix element

$$\langle\psi|R\psi\rangle = \int_0^\infty dr_1 r_1^2 \int_0^\infty dr_2 r_2^2 \psi^*(r_1,r_2) \times \int_0^\infty dr_1' r_1'^2 \int_0^\infty dr_2' r_2'^2 R(r_1,r_2,r_1',r_2')\psi(r_1',r_2'), \quad (3.12)$$
$$\psi \in L^2(\mathbb{R}_+, r_1^2 dr_1) \otimes L^2(\mathbb{R}_+, r_2^2 dr_2),$$

where the operator R defined in (3.9), which reads explicitly

$$R = (P_\Phi \otimes 1)W(r_2)k(r_1,r_2)W(r_1)(1 \otimes P_\Phi) + (1 \otimes P_\Phi)W(r_1)k(r_1,r_2)W(r_2)(P_\Phi \otimes 1),$$
$$k(r_1,r_2) = (k_0 - k_2)(r_1,r_2), \quad (3.13)$$

has the kernel representation

$$R(r_1,r_2,r_1',r_2') = \Phi(r_1)\Phi(r_2)(W(r_2)k(r_2,r_1')W(r_1') + W(r_1)k(r_1,r_2')W(r_2'))\cdot\Phi(r_1')\Phi(r_2'). \quad (3.14)$$

From Eqs. (3.11) and (3.13) one finds easily that the contribution to $(-R)$ and thus also to the lower bound (3.4) coming from $k_2(r_1,r_2)$ consists of an infinite-rank part which is positive (negative) for antisymmetric (symmetric) radial wave functions, whereas the contribution from $k_0(r_1,r_2)$ consists of a positive (negative) rank-one part and a negative (positive) infinite-rank part for the antisymmetric (symmetric) case.

Our next efforts are devoted to the reduction of the resulting lower bound to suitable one-particle problems; we must treat separately the spin-singlet and -triplet case. Since for $l = 1$ the wave function is always antisymmetric in Ω_1 and Ω_2, the singlet wave functions are antisymmetric with respect to r_1 and r_2, whereas the radial parts of the triplet states are symmetric functions.

A. Singlet sector

Here we may omit the contributions to $(-R)$ coming from $k_2(r_1,r_2)$ and that from the rank one part of $k_0(r_1,r_2)$. One may try to eliminate the rank-one part with the help of the common eigenfunction of Q_1 and Q_2 to eigenvalue one, which is given by

$$Q_1\chi = Q_2\chi = \chi, \quad \int_0^\infty dr_1 r_1^2 \int_0^\infty dr_2 r_2^2 \chi^2(r_1,r_2) = 1, \quad (3.15)$$
$$\chi(r_1,r_2) = cV_L^{-1/2}(r_1,r_2)\Phi(r_1)\Phi(r_2), \quad c^2 = \frac{32}{231}.$$

But unfortunately $k_0 - k_2$ has an infinite-rank positive part. Therefore we use (3.11) and obtain

$$V_L \geqslant (P_\Phi \otimes 1)W(r_2) + W(r_1)(1 \otimes P_\Phi) - R_0, \quad (3.16)$$
$$R_0 = (P_\Phi \otimes 1)W(r_2)\overline{k_0}(r_1,r_2)W(r_1)(1 \otimes P_\Phi) + (1 \leftrightarrow 2)$$
$$\overline{k_0}(r_1,r_2) = k_0(r_1,r_2) - \tfrac{1}{2}(1 + r_1)(1 + r_2).$$

The Coulomb-like operator H_0 can be estimated conveniently by means of the projector

$$(P_\Phi \otimes 1) \vee (1 \otimes P_\Phi) = P_\Phi \otimes 1 + 1 \otimes P_\Phi - P_\Phi \otimes P_\Phi, \quad (3.17)$$

where the last term is zero in the singlet sector; so one obtains

$$H_0 + \tfrac{1}{8} \geqslant (P_\Phi \otimes 1)(p_2^2/2 - 1/r_2) \otimes P_{1,m} + (1 \leftrightarrow 2) \quad (3.18)$$

and inserting the estimate (3.16) the inequality

$$H + \tfrac{1}{8} \geqslant \left[(P_\Phi \otimes 1)(p_2^2/2\,|_{l_2=1} - 1/r_2 + W(r_2) - (P_\Phi \otimes 1)W(r_2)\overline{k_0}(r_1,r_2)W(r_1)(1 \otimes P_\Phi) + (1 \leftrightarrow 2)\right]P_{1,m} \quad (3.19)$$

is obtained. The kinetic energy operator is here restricted to angular momentum one.

Sandwiching (3.19) between singlet states of the form

$2^{-1/2}(\Phi(r_1)f(r_2)-f(r_1)\Phi(r_2)), f\in L^2(\mathbb{R}_+,r^2dr)$ for the radial parts, one finds out that it suffices to count the negative eigenvalues of the one-particle operator

$$h_s=(1-P_\Phi)[p^2/2|_{l=1}-1/r+W(r)-S](1-P_\Phi)$$
$$\text{on } L^2(\mathbb{R}_+,r^2dr), \tag{3.20}$$

where the integral operator S has a kernel

$$s(r_1,r_2)=-\Phi(r_1)W(r_1)\,\overline{k_0}(r_1,r_2)W(r_2)\Phi(r_2). \tag{3.21}$$

In Appendix A we shall prove that this operator h_s on the half-line is nonnegative.

B. Triplet sector

The Coulomb-like part can again be estimated using (3.17); one obtains

$$H_0+\tfrac{1}{8}\geqslant((P_\Phi\otimes 1)(p_2^2/2|_{l_2=1}-1/r_2)+(1\leftrightarrow 2)$$
$$+\tfrac{1}{8}P_\Phi\otimes P_\Phi)P_{1,m}. \tag{3.22}$$

With respect to the potential V_L, this time we omit the infinite-rank part of $k_0(r_1,r_2)$ and thus obtain the estimate

$$V_L\geqslant(P_\Phi\otimes 1)W(r_2)+(P_\Phi\otimes 1)W(r_2)$$
$$\times\,\overline{k_2}(r_1,r_2)W(r_1)(1\otimes P_\Phi)+(1\leftrightarrow 2), \tag{3.23}$$

$$\overline{k_2}(r_1,r_2)=k_2(r_1,r_2)-\tfrac{1}{2}(1+r_1)(1+r_2),$$

where the last term stems from the rank-one part of $k_0(r_1,r_2)$ and we omitted the last two terms of Eq. (3.8).

Again sandwiching the resulting estimate of H, now between triplet states, i.e., symmetric radial wave functions, one ends up with counting the negative eigenvalues of the one-particle operator

$$h_t=\frac{p^2}{2}\bigg|_{l=1}-\frac{1}{r}+W(r)-T+\frac{1}{16}P_\Phi$$
$$\text{on } L^2(\mathbb{R}_+,r^2dr), \tag{3.24}$$

where T denotes the integral operator with kernel

$$t(r_1,r_2)=-\Phi(r_1)W(r_1)\,\overline{k_2}(r_1,r_2)W(r_2)\Phi(r_2). \tag{3.25}$$

In (3.24) we have disregarded the factors $(1+P_\Phi)$ on both sides. In Appendix B we shall prove that this operator h_t is non-negative, except for *one* negative eigenfunction.

Together with the absence of negative eigenvalues of $H+\frac{1}{8}$ in the unnatural parity sector with total angular momentum $l\geqslant 2$, we thus have obtained the following.

Theorem: In the subspace of unnatural parity there exists *exactly one* (three times degenerate) bound state of the Hamiltonian (1.1) below the threshold $-\frac{1}{8}$, which belongs to the triplet sector and carries total angular momentum $l=1$ and $m=0,\pm 1$.

APPENDIX A

Here we are going to study the pure point spectrum of the integro-differential operator h_s defined in Eq. (3.20) using standard Birman–Schwinger techniques, i.e., the discrete eigenvalues $-\epsilon$ of the integro-differential equation

$$(1-P_\Phi)[p^2/2|_{l=1}-\lambda(1/r-W(r)+S)](1-P_\Phi)\psi(r)$$
$$=-\tfrac{1}{2}\epsilon(1-P_\Phi)\psi(r),\quad \epsilon\geqslant 0,\quad \psi\in L^2(\mathbb{R}_+,r^2dr), \tag{A1}$$

with the potential $W(r)$ defined in (3.6) and the integral operator S defined by its kernel (3.21); instead of counting the eigenvalues below energy zero for $\lambda=1$ we may also try to determine the number of characteristic values of the coupling constant λ in the interval $0\leqslant\lambda\leqslant 1$ and then let $\epsilon\searrow 0$.

In order to write explicitly the resolvent $G_\Phi(\epsilon)$ of the restriction of the kinetic energy to the subspace orthogonal to Φ

$$(1-P_\Phi)(p^2|_{l=1}+\epsilon)(1-P_\Phi)G_\Phi(\epsilon)(1-P_\Phi)$$
$$=1-P_\Phi,\quad \epsilon\geqslant 0 \tag{A2}$$

in terms of the resolvent $G(\epsilon)$ defined by

$$(p^2|_{l=1}+\epsilon)G(\epsilon)=1 \text{ on } L^2(\mathbb{R}_+,r^2dr) \tag{A3}$$

we use the well-known operator identity

$$(1-P_\Phi)G_\Phi(\epsilon)(1-P_\Phi)$$
$$=G(\epsilon)-G(\epsilon)P_\Phi(P_\Phi G(\epsilon)P_\Phi)^{-1}P_\Phi G(\epsilon). \tag{A4}$$

With the aid of the square root of this positive resolvent we can now transform our eigenvalue problem (A1) to the following integral equation with symmetric kernel:

$$\frac{1}{\lambda}\chi=[k_1(\epsilon)+k_2(\epsilon)]\chi,\quad \chi=R_\Phi^{-1}(\epsilon)\psi,$$
$$R_\Phi^2(\epsilon)=(1-P_\Phi)G_\Phi(\epsilon)(1-P_\Phi),$$
$$k_1(\epsilon)=2R_\Phi(\epsilon)\Big(\frac{1}{r}-W(r)\Big)R_\Phi(\epsilon),$$
$$k_2(\epsilon)=2R_\Phi(\epsilon)SR_\Phi(\epsilon),\quad 0\leqslant\lambda\leqslant 1,\quad \epsilon\geqslant 0. \tag{A5}$$

Then the number of bound states N_s of Eq. (A1) below the energy zero is limited by the trace

$$N_s\leqslant\lim_{\epsilon\searrow 0}\operatorname{tr}[k_1(\epsilon)+k_2(\epsilon)]^2. \tag{A6}$$

Next we use the property of both integral operators $k_1(\epsilon)$ and $k_2(\epsilon)$ of being positive (here is the reason why we have thrown away repulsive contributions in part 3); we conclude that

$$N_s\leqslant\lim_{\epsilon\searrow 0}[(\operatorname{tr}k_1^2(\epsilon))^{1/2}+\operatorname{tr}k_2(\epsilon)]^2. \tag{A7}$$

The terms on the right-hand side of (A7) can be written more explicitly as

$$A_1=\lim_{\epsilon\searrow 0}\operatorname{tr}k_1^2(\epsilon)$$
$$=4\int_0^\infty dr_1r_1^2\int_0^\infty dr_2r_2^2\Big[\frac{1}{r_1}-W(r_1)\Big]$$
$$\times\Big[\frac{1}{r_2}-W(r_2)\Big]g_\Phi^2(r_1,r_2)$$
$$A_2=\lim_{\epsilon\searrow 0}\operatorname{tr}k_2(\epsilon)$$
$$=2\int_0^\infty dr_1r_1^2\int_0^\infty dr_2r_2^2 s(r_1,r_2)g_\Phi(r_1,r_2) \tag{A8}$$
$$g_\Phi(r_1,r_2)=g(r_1,r_2)-\frac{1}{c_1}\psi(r_1)\psi(r_2).$$

The kernel $g(r_1,r_2)$ is obtained as the $\epsilon\searrow 0$ limit of the Green's function for $l=1$

$$g(r_1,r_2)=\lim_{\epsilon\searrow 0}\epsilon^{1/2}j_1(\epsilon^{1/2}r_<)h_1^{(+)}(\epsilon^{1/2}r_>)=\frac{r_<}{3r_>^2}$$
$$r_<=\min(r_1,r_2),\quad r_>=\max(r_1,r_2), \tag{A9}$$

where j_1 and $h_1^{(+)}$ denote the spherical Bessel functions. $\psi(r)$ and c_1 enter from Eq.(A4) and are explicitly given as

$$\psi(r) = \int_0^\infty dr_2 r_2^2 g(r_1,r_2)\Phi(r_2)$$

$$= c_2\left\{\frac{1}{r_1^2} - e^{-r_1/2}\left(\frac{r_1}{64} + \frac{1}{8} + \frac{1}{2r_1} + \frac{1}{r_1^2}\right)\right\}$$

$$c_1 = \langle\Phi, G(0)\Phi\rangle = \tfrac{28}{3}, \quad c_2 = 64\sqrt{2/3}. \tag{A10}$$

Numerical integration leads to

$$A_1 = 0.595, \quad A_2 = 0.161 \rightarrow N_s \leqslant 0.869, \tag{A11}$$

and to the conclusion written in the text.

APPENDIX B

Here we shall count the number of eigenvalues N_t of the operator (3.24). Since a straightforward procedure leads to $N_t \leqslant 2$ and since one has the feeling that the positive one-dimensional contribution to (3.24) should compensate the negative one-dimensional part of T we proceed as follows: Projecting h_t onto the Hilbert space orthogonal to $P_\Phi \mathscr{H}$ one obtains

$$h_t' = (1 - P_\Phi)h_t(1 - P_\Phi)$$

$$= (1 - P_\Phi)\left[p^2/2|_{l=1} - \frac{1}{r} + W(r) - T\right](1 - P_\Phi)$$

$$\text{on } L^2(\mathbb{R}_+, r^2 dr) \tag{B1}$$

where W and T are as in Eq. (3.24). Next we use the min-max principle to conclude that the number of bound states of h_t is greater by at most one than the number of bound states of h_t'. But counting the bound states of h_t' is a problem analogous to that treated in Appendix A. Therefore we get

$$N_t \leqslant 1 + \lim_{\epsilon \searrow 0} \operatorname{tr}[k_1(\epsilon) + k_3(\epsilon)]^2, \tag{B2}$$

where $k_1(\epsilon)$ has been defined in (A5) and $k_3(\epsilon)$ is similar to $k_2(\epsilon)$ and given by

$$k_3(\epsilon) = 2R_\Phi(\epsilon)TR_\Phi(\epsilon), \quad \epsilon > 0. \tag{B3}$$

As before we use trace inequalities to obtain

$$N_t \leqslant 1 + \lim_{\epsilon \searrow 0}\left[(\operatorname{tr}k_1^2(\epsilon))^{1/2} + \operatorname{tr}k_3(\epsilon)\right]^2, \tag{B4}$$

$\operatorname{tr}k_3(\epsilon)$ entering (B4) reads more explicitly for $\epsilon \searrow 0$:

$$A_3 = \lim_{\epsilon \searrow 0} \operatorname{tr}k_3(\epsilon)$$

$$= 2\int_0^\infty dr_1 r_1^2 \int_0^\infty dr_2 r_2^2 g_\Phi(r_1,r_2) t(r_1,r_2), \tag{B5}$$

where g_Φ has been given in (A8) and t in (3.25). Numerical integration leads to

$$A_3 = 0.0718 \rightarrow N_t \leqslant 1.711, \tag{B6}$$

and $\lim_{\epsilon \searrow 0} \operatorname{tr}k_1^2(\epsilon)$ has been given in (A11).

ACKNOWLEDGMENT

We thank Dr. F. Schöberl for help with the handling of the computer.

[1]H. Bethe, Z. Phys. **57**, 815 (1929)
[2]E. H. Hylleraas, Z. Phys. **63**, 291 (1930).
[3]C. L. Pekeris, Phys. Rev. **126**, 1470 (1962).
[4]R. N. Hill, Phys. Rev. Lett. **38**, 643 (1977); J. Math. Phys. **18**, 2316 (1977).
[5]M. A. Antonets, G. M. Zhislin, and I. A. Shereshevskii, Teor. Mat. Fiz. **16**, 235 (1973) [Theor. Math. Phys. **16**, 800 (1974)].
[6]M. B. Ruskai, Comm. Math. Phys. **82**, 457 (1982).
[7]H. Grosse, J. Phys. A Math. Gen. **10**, 711 (1977).
[8]M. Klaus and B. Simon, Comm. Math. Phys. **78**, 153 (1981).
[9]G. W. F. Drake, Phys. Rev. Lett. **24**, 126 (1970).
[10]W. Thirring, *Quantum Mechanics of Atoms and Molecules* (Springer, New York, 1981).
[11]N. W. Bazley, Proc. Nat. Acad. Sci. USA **45**, 850 (1959).
[12]N. W. Bazley and D. W. Fox, Phys. Rev. **124**, 483 (1961).

Inverse scattering for the reflectivity function

Samuel H. Gray
Amoco Production Company, P. O. Box 591, Tulsa, Oklahoma 74102

(Received 24 September 1982; accepted for publication 30 December 1982)

An inverse method for elastic, electromagnetic, or acoustic waves in a stratified half-space is presented. Rather than transforming the wave equation to one for which quantum inverse scattering methods can be applied in solving for a potential $q(\tau)$, we transform to one where it is suitable to solve for a "reflectivity function," or local reflection coefficient, $\gamma(\tau)$. We show that $\gamma(\tau)$ can be discontinuous, thus improving a result of Balanis, and that discontinuities of $\gamma(\tau)$ match those of the impulse response $R(t)$. We also show the relationship between the scattering kernel of this method and the scattering kernel of the quantum inverse scattering theory.

PACS numbers: 03.65.Ge, 03.65.Nk

I. INTRODUCTION

The Gel'fand–Levitan (Marchenko) integral equation,[1,2] first applied in solving the inverse problem of quantum scattering theory, has also been used to solve analogous problems in elastic and electromagnetic wave propagation theories.[3,4] The approach usually taken in solving these latter inverse problems is to transform the wave equation into either the equation for an elastically braced string (time-dependent)

$$V_{\tau\tau} - V_{tt} - q(\tau)V = 0, \tag{1}$$

or the time-independent Schrödinger equation

$$v_{\tau\tau} + [\omega^2 - q(\tau)]v = 0. \tag{2}$$

Obtaining Eq. (1) or Eq. (2) from the wave equation requires transforming both the dependent and the independent variables. If $q(\tau)\equiv 0$ for $\tau\leqslant 0$, then q can be determined from the Gel'fand–Levitan (Marchenko) equation as follows:

$$q(\tau) = 2\frac{d}{d\tau}K(\tau,\tau), \tag{3}$$

where $K(\tau,t)$ satisfies

$$K(\tau,t) + R(\tau+t) + \int_{-t}^{\tau} R(t+s)K(\tau,s)\,ds = 0 \quad \text{for } |t|<\tau, \tag{4}$$

with $R(t)$ the impulse response of the medium measured at $\tau = 0$. The potential $q(\tau)$ is related to the profile in question—impedance profile on dielectric profile—by a differential equation obtained in the transformations leading to Eq. (1) or (2).

Recently, however, Balanis[5] has devised an inverse scattering theory, which uses an equation similar to Eq. (4), for the equation

$$U_{\tau\tau} - U_{tt} - \gamma(\tau)U_\tau = 0 \tag{5}$$

[where $\gamma(\tau)\equiv 0$ for $\tau\leqslant 0$]. Equation (5), like Eqs. (1) and (2), comes from transforming the wave equation. Specifically, the elastic wave equation

$$(\rho c^2 U_x)_x - \rho U_{tt} = 0 \tag{6}$$

is transformed into Eq. (5) by letting

$$\frac{dx}{d\tau} = c(x); \tag{7}$$

then

$$\gamma(\tau) = -\frac{d}{d\tau}\ln \rho c. \tag{8}$$

Also, the electromagnetic wave equation

$$U_{xx} - [\epsilon(x)\mu_0/c^2]\,U_{tt} = 0 \tag{9}$$

becomes Eq. (5) if

$$\frac{dx}{d\tau} = c[\epsilon(x)\mu_0]^{-1/2}. \tag{10}$$

Here,

$$\gamma(\tau) = -\frac{1}{2}\frac{d}{d\tau}\ln\epsilon. \tag{11}$$

Balanis has also applied his result to the acoustic wave equation

$$U_{xx} - [1/c^2(x)]\,U_{tt} = 0. \tag{12}$$

Here,

$$\frac{dx}{d\tau} = c(x), \tag{13}$$

and

$$\gamma(\tau) = \frac{d}{d\tau}\ln c. \tag{14}$$

For impedance or dielectric profile inversion, inverse scattering on Eq. (5)—i.e., determining $\gamma(\tau)$ in Eq. (5) and the profile in question by Eqs. (8), (11), or (14)—is more appealing than inverse scattering on Eq. (1) or (2). This is because Eq. (5) is more closely related to the wave equation than Eq. (1) or (2) is: Only one change of variable is required. Also, the quantity being sought in Eq. (5), namely $\gamma(\tau)$, has a direct physical interpretation while $q(\tau)$ in Eq. (1) or Eq. (2) does not. For example, in Eq. (14), the reflectivity function[6]

$$\gamma(\tau) = \frac{c'(\tau)}{c(\tau)} = 2\lim_{\Delta\tau\to 0}\left(\frac{1}{\Delta\tau}\frac{\Delta c}{2c+\Delta c}\right) \tag{15}$$

is related to the reflection coefficient at a point in the varying medium, and similarly for Eqs. (8) and (11). Finally, as will be shown, $\gamma(\tau)$ is as continuous or discontinuous as the impulse response $R(t)$, up to and including jump discontinuities.

In obtaining his result, Balanis works in the time domain. Drawing on his previous results,[7] he concludes the validity of the approach for continuous functions $\gamma(\tau)$. The major result of this note is to extend the validity of the result

 0022-2488/83/051148-04$02.50

to include functions with jump discontinuities. We do this by using a combination frequency/time domain approach.

We emphasize that our results do not constitute a theoretical improvement over the approach which transforms the wave equation into Eq. (1) or (2). There, the quantity ultimately being sought—e.g., impedance—possesses two continuous derivatives more than $q(\tau)$, which can sometimes be as singular as a delta function.[3] In our extension of Balanis's theory, the quantity ultimately being sought possesses one continuous derivative more than $\gamma(\tau)$, which can have at most a jump discontinuity. But, as we shall discuss, Balanis's integral equation does appear to offer computational advantages over Eq. (4).

II. A GEL'FAND–LEVITAN (MARCHENKO) INTEGRAL EQUATION AND AN EQUATION FOR $\gamma(\tau)$

Here, we shall derive the equations central to Balanis's theory, namely,

$$\gamma(\tau) = 2\frac{d}{d\tau}\mathscr{K}(\tau,\tau) - \gamma(\tau)\mathscr{K}(\tau,\tau), \tag{16}$$

where

$$\int_0^{t+\tau} R(s)\,ds - \mathscr{K}(\tau,t) + \int_{-t}^{\tau} \mathscr{K}(\tau,s)R(t+s)\,ds = 0 \quad \text{for } |t|<\tau. \tag{17}$$

In addition, we shall show the relationship between $\mathscr{K}(\tau,t)$ and the function $K(\tau,t)$ of Eq. (4). These equations will be derived partially in the frequency domain, partially in the time domain. Our presentation will closely resemble that of Balanis[7] and Scott *et al.* (Ref. 8, Appendix D).

Since Eq. (17) will be derived partially in the frequency domain, we need the time-independent form of Eq. (5); this is

$$u_{\tau\tau} + \omega^2 u - \gamma(\tau)u_\tau = 0. \tag{18}$$

By comparison with the Schrödinger equation (2), it is seen that fundamental solutions of Eq. (18) are

$$f_1(\tau,\omega) = \exp(i\omega\tau) - \int_\tau^\infty \frac{\sin[\omega(\tau-s)]}{\omega}\gamma(s)f_{1s}(s,\omega)\,ds, \tag{19a}$$

$$f_2(\tau,\omega) = \exp(-i\omega\tau) + \int_0^\tau \frac{\sin[\omega(\tau-s)]}{\omega}\gamma(s)f_{2s}(s,\omega)\,ds, \tag{19b}$$

where the subscript s denotes differentiation with respect to s. From the form of Eq. (19),

$$f_1(\tau,\omega) \to \exp(i\omega\tau) \quad \text{as } \tau\to\infty, \tag{20a}$$

$$f_2(\tau,\omega) = \exp(-i\omega\tau) \quad \text{for } \tau\leqslant 0. \tag{20b}$$

Next, we write

$$g(\tau,\omega) = f_2(\tau,\omega) - \exp(-i\omega\tau), \tag{21}$$

and we express the physical wave at a given frequency as a combination of the linearly independent solutions $f_2(\tau,\omega)$ and $f_2(\tau,-\omega)$ of Eq. (18):

$$u(\tau,\omega) = r(\omega)f_2(\tau,\omega) + f_2(\tau,-\omega). \tag{22}$$

Then the response at a point in the medium to the incident wave $\delta(\tau-t)$ is

$$U(\tau,t) = (1/2\pi)\int_{-\infty}^{\infty} u(\tau,\omega)\exp(-i\omega t)\,d\omega. \tag{23}$$

Using Eq. (21) and computing the inverse Fourier transform of Eq. (22) leads to

$$U(\tau,t) = R(\tau+t) + \int_{-\infty}^{\infty} R(t+s)G(\tau,s)\,ds + \delta(t-\tau) + G(\tau,t), \tag{24}$$

where

$$G(\tau,t) = (1/2\pi)\int_{-\infty}^{\infty} g(\tau,\omega)\exp(i\omega t)\,d\omega, \tag{25}$$

$$R(t) = (1/2\pi)\int_{-\infty}^{\infty} r(\omega)\exp(-i\omega t)\,d\omega. \tag{26}$$

Integrating Eq. (24) with respect to t gives

$$\int_{-\infty}^{t} U(\tau,s)\,ds = \int_{-\infty}^{\tau+t} R(s)\,ds - \int_{-\infty}^{\infty} R(t+s)\left[\int_{-\infty}^{s} G(\tau,\sigma)\,d\sigma\right]ds + H(t-\tau) + \int_{-\infty}^{t} G(\tau,s)\,ds, \tag{27}$$

where $H(t)$ is the Heaviside function. [As a check, we note that the t-derivative of Eq. (27) is Eq. (24) as long as $\int_{-\infty}^{t} G(\tau,s)\,ds = 0$ for $t>\tau$.] For this problem with $\gamma(\tau)=0$ for $\tau\leqslant 0$, it turns out that $G(\tau,t)=0$ for $t<-\tau$.[9] Also, the causal impulse response R satisfies $R(s)=0$ for $s<0$ and, as noted above, $\int_{-\tau}^{s} G(\tau,\sigma)\,d\sigma = 0$ for $s>\tau$. Therefore, Eq. (27) can be rewritten as

$$\int_{-\infty}^{t} U(\tau,s)\,ds = \int_{-\infty}^{\tau+t} R(s)\,ds - \int_{-t}^{\tau} R(t+s) \times\left[\int_{-\tau}^{s} G(\tau,\sigma)\,d\sigma\right]ds + H(t-\tau) + \int_{-\tau}^{t} G(\tau,s)\,ds. \tag{28}$$

For $t<\tau$, $\int_{-\infty}^{t} U(\tau,s)\,ds = 0$ and $H(t-\tau)=0$, so that Eq. (28) is the same as Eq. (17) for $|t|<\tau$ if we make the identification

$$\mathscr{K}(\tau,t) = \begin{cases} -\int_{-\tau}^{t} G(\tau,s)\,ds & \text{for } t\leqslant\tau, \\ 0 & \text{for } t>\tau. \end{cases} \tag{29}$$

[We note that since $G(\tau,t)=0$ for $t<-\tau$, $\mathscr{K}(\tau,t)$ is continuous at $t=-\tau$, and since $\mathscr{K}(\tau,t)=0$ for $t>\tau$, $\mathscr{K}(\tau,t)$ is discontinuous at $t=\tau$.] This completes the derivation of Eq. (17).

Next, we show how $\mathscr{K}(\tau,t)$ and the solution $K(\tau,t)$ of Eq. (4) are related. From Eq. (29),

$$G(\tau,t) = -\frac{\partial}{\partial t}\mathscr{K}(\tau,t), \tag{30}$$

where $\mathscr{K}(\tau,t)$ has a jump discontinuity at $t=\tau$. Thus,

$$G(\tau,t) = \delta(\tau-t)\,\mathscr{K}(\tau,\tau) - H(\tau-t)\frac{\partial}{\partial t}\mathscr{K}_1(\tau,t), \tag{31}$$

where $\mathscr{K}_1(\tau,t)\equiv\mathscr{K}(\tau,t)$ for $t<\tau$. Inserting this expression into Eq. (24), we see that

$$U(\tau,t) = (1 + \mathscr{K}(\tau,\tau))\Big[R(\tau+t) + \int_{-t}^{\tau} R(t+s)K(\tau,s)\,ds + \delta(t-\tau) + K(\tau,t)\Big], \tag{32}$$

where, for $t<\tau$,

$$K(\tau,t) = \frac{G(\tau,t)}{1+\mathscr{K}(\tau,\tau)} = -\frac{\partial}{\partial t}\frac{\mathscr{K}(\tau,t)}{1+\mathscr{K}(\tau,\tau)}, \tag{33}$$

or

$$\int_{-\tau}^{t} K(\tau,s)\,ds = -\frac{\mathscr{K}(\tau,t)}{1+\mathscr{K}(\tau,\tau)}. \tag{34}$$

Since Eq. (32) reduces to Eq. (4) when $|t|<\tau$, Eq. (33) or (34) provides the relationship between $\mathscr{K}$ and K for $|t|<\tau$. It will be shown below that Eq. (34) also holds on the line $t=\tau$.

To obtain Eq. (16), we first find an equation satisfied by $g(\tau,\omega)$. This is

$$\Big(\frac{d^2}{d\tau^2}+\omega^2-\gamma\frac{d}{d\tau}\Big)g = \Big(\frac{d^2}{d\tau^2}+\omega^2-\gamma\frac{d}{d\tau}\Big)[f_2-\exp(-i\omega\tau)] = -i\omega\gamma(\tau)\exp(-i\omega\tau). \tag{35}$$

If we divide this equation by $-i\omega$ and Fourier transform, the result is an equation for $\mathscr{K}(\tau,t) = -\int_{-\tau}^{t} G(\tau,s)\,ds$:

$$\Big(\frac{\partial^2}{\partial\tau^2}-\frac{\partial^2}{\partial t^2}-\gamma\frac{\partial}{\partial\tau}\Big)\mathscr{K} = \gamma(\tau)\delta(t-\tau). \tag{36}$$

Writing this equation in the variables $\zeta = t+\tau$ and $\eta = t-\tau$ leads to

$$-4\frac{\partial^2\mathscr{K}}{\partial\zeta\partial\eta} - \gamma\Big(\frac{\zeta-\eta}{2}\Big)\Big(\frac{\partial}{\partial\zeta}-\frac{\partial}{\partial\eta}\Big)\mathscr{K} = \gamma\Big(\frac{\zeta-\eta}{2}\Big)\delta(\eta). \tag{37}$$

Holding ζ fixed, performing the η integration from $-\epsilon$ to ϵ and letting $\epsilon\to 0^+$ gives

$$4\frac{\partial}{\partial\zeta}\mathscr{K}(\zeta,0) - \gamma\Big(\frac{\zeta}{2}\Big)\mathscr{K}(\zeta,0) = \gamma\Big(\frac{\zeta}{2}\Big). \tag{38}$$

In performing the integrals to obtain $\gamma(\zeta/2)$, γ need not be continuous at $\eta=0$. For example, in the integral

$$\int_{-\epsilon}^{\epsilon}\gamma\Big(\frac{\zeta-\eta}{2}\Big)\frac{\partial}{\partial\eta}\mathscr{K}(\zeta,\eta)\,d\eta,$$

$\mathscr{K}(\zeta,\eta)$ has a jump discontinuity at $\eta=0$ $(\tau=t)$ so that $(\partial/\partial\eta)\mathscr{K}$ has a jump discontinuity plus a delta-function discontinuity at $\eta=0$. Integrating the less singular of these contributes zero in the limit $\epsilon\to 0^+$ [even if $\gamma((\zeta-\eta)/2)$ has a jump discontinuity at $\eta=0$]. Integrating the more singular contributes the size of the delta function [namely, $\lim_{\epsilon\to 0^+}\mathscr{K}(\zeta,-\epsilon)$] times the average of the values of γ on either side of $\eta=0$. Rewriting Eq. (38) in the variables τ and t gives

$$2\Big(\frac{\partial}{\partial\tau}+\frac{\partial}{\partial t}\Big)\mathscr{K}(\tau,t) - \gamma(\tau)\mathscr{K}(\tau,t) = \gamma(\tau) \quad \text{for } t=\tau, \tag{39}$$

which is the same as Eq. (16). In Eq. (39), however, $\gamma(\tau)$ means $\lim_{\epsilon\to 0}[\gamma(\tau+\epsilon)+\gamma(\tau-\epsilon)]/2$.

An alternate derivation of Eq. (16) can be obtained by comparing Eq. (32) with a different expression for $U(\tau,t)$. This expression is

$$U(\tau,t) = \exp\Big[\frac{1}{2}\int_0^{\tau}\gamma(s)\,ds\Big]V(\tau,t), \tag{40}$$

where $V(\tau,t)$ satisfies Eq. (1). Since $V(\tau,t)$ is precisely the bracketed term in Eq. (32), it follows that

$$\exp\Big[\frac{1}{2}\int_0^{\tau}\gamma(s)\,ds\Big] = 1+\mathscr{K}(\tau,\tau), \tag{41}$$

or

$$\gamma(\tau) = 2\frac{d}{d\tau}\ln[1+\mathscr{K}(\tau,\tau)], \tag{42}$$

which is equivalent to Eq. (16).

Equation (16), rewritten in the form

$$\gamma(\tau) = 2\frac{(d/d\tau)\mathscr{K}(\tau,\tau)}{1+\mathscr{K}(\tau,\tau)}, \tag{43}$$

provides a means for finding the quantity which is the ultimate goal of the inversion. For example, acoustic impedance $\rho(\tau)c(\tau)$ in Eq. (6) is obtained by integrating Eqs. (8) and (43):

$$\rho(\tau)c(\tau) = \rho(0)c(0)[1+\mathscr{K}(\tau,\tau)]^{-2}. \tag{44}$$

Analogous results have been reported in terms of the more commonly used scattering kernel $K(\tau,t)$.[10,11] The derivations of those results are based on quantum inverse scattering techniques and are independent of the one presented here. For inverse scattering on the elastic wave equation (6), the formula is[10]

$$\rho(\tau)c(\tau) = \rho(0)c(0)\Big[1+\int_{-\tau}^{\tau}K(\tau,t)\,dt\Big]^2. \tag{45}$$

Comparison of Eqs. (44) and (45) shows that

$$1+\mathscr{K}(\tau,\tau) = \Big[1+\int_{-\tau}^{\tau}K(\tau,t)\,dt\Big]^{-1}, \tag{46}$$

which is true if

$$\frac{\mathscr{K}(\tau,\tau)}{[1+\mathscr{K}(\tau,\tau)]} = -\int_{-\tau}^{\tau}K(\tau,t)\,dt. \tag{47}$$

This, in turn, implies that Eq. (34) holds on the line $t=\tau$, thus verifying the assertion made earlier.

Finally, it is easy to show that $\gamma(\tau)$ is as continuous or discontinuous as the impulse response $R(t)$. From the transformation (40) which relates Eq. (5) to Eq. (1), the function $q(\tau)$ of Eq. (1) is equal to $\frac{1}{4}\gamma^2(\tau)-\frac{1}{2}\gamma'(\tau)$; this is as singular as $\gamma'(\tau)$. By comparison with Eq. (3), then, $\gamma(\tau)$ is as singular as $K(\tau,\tau)$, which is in turn as singular as $R(2\tau)$ by Eq. (4).

III. CONCLUSIONS

We have discussed a solution of the inverse-scattering problem for the reflectivity function $\gamma(\tau)$. We have shown that $\gamma(\tau)$ can be discontinuous. However, the quantity ultimately being sought must be continuous.

It has been noted that our results do not in theory improve upon existing methods which apply quantum inverse scattering considerations to a transformed wave equation. But from a numerical standpoint, it would seem that a profile inversion using Eq. (44) and a discretization of Eq. (17)

will be preferable to one using Eqs. (47) and (4). This is because $\mathscr{K}(\tau,t)$ is a smoother function than $K(\tau,t)$, a fact which should enhance the numerical stability of schemes for solving the discretized equation (17).

ACKNOWLEDGMENT

I wish to thank the referee for correcting an important error in an earlier version of this paper.

[1] I. M. Gel'fand and B. M. Levitan, Am. Math. Soc. Transl. **1**, 253 (1955).

[2] Z. S. Agranovich and V. A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963).

[3] J. A. Ware and K. Aki, J. Acoust. Soc. Am. **45**, 911 (1969).

[4] H. E. Moses and C. M. de Ridder, Lincoln Lab. Tech. Rep. No. 322, MIT, Cambridge, 1963.

[5] G. N. Balanis, J. Math. Phys. **23**, 2562 (1982).

[6] M. Foster, Geophys. J. Roy. Astron. Soc. **42**, 519 (1975).

[7] G. N. Balanis, J. Math. Phys. **13**, 1001 (1972).

[8] A. C. Scott, F. Y. F. Chu, and D. W. McLaughlin, Proc. IEEE **61**, 1443 (1973).

[9] R. Burridge, Wave Motion **2**, 305 (1980).

[10] J. G. Berryman and R. R. Greene, Geophysics **45**, 213 (1980).

[11] S. Coen, IEEE Trans. Antennas Propagation **AP-29**, 726 (1981).

The Coulomb Jost states in the momentum representation for all partial waves in closed form

H. van Haeringen
Department of Mathematics, Delft University of Technology, Delft, The Netherlands

(Received 1 October 1981; accepted for publication 5 March 1982)

We obtain the Coulomb Jost states in the momentum representation for all l in exact closed form. These closed expressions consist of combinations of the function ${}_2F_1(1, i\gamma; 1 + i\gamma; \cdot)$, Jacobi polynomials, and other polynomials. We also discuss the relation of the Coulomb Jost states with other quantities that are of interest in the theory of charged-particle scattering.

PACS numbers: 03.65.Nk, 02.30. + g

1. INTRODUCTION

The Coulomb interaction plays an important and, from a mathematical point of view, interesting role in the theory of scattering by charged particles. Two-particle pure Coulomb scattering wave functions in the coordinate representation have been known in closed form for a long time. Often it is advantageous to work in the momentum representation. In this case one needs expressions for all relevant scattering quantities in the momentum representation. The so-called regular solutions of Schrödinger's equation with a pure Coulomb potential are known in momentum space for all l (see below). On the other hand, no closed expressions are known for the irregular solutions, or Jost solutions, in momentum space (except for $l = 0$; see below).

In this paper we shall derive an integral representation [Eq. (5)], a series representation [Eq. (14)], and two hypergeometric-function expressions [Eqs. (8) and (17)–(21)] for the Coulomb Jost states in the momentum representation, which we denote by $\langle p|kl\uparrow\rangle_c$, for all $l = 0,1,\ldots$. Expressions are also obtained for the closely related quantities $\langle p|V_{cl}|kl\uparrow\rangle_c$, where V_{cl} is the Coulomb potential. Preliminary results have been reported in Ref. 1.

There exists an interesting relationship with the partial-wave projected off-shell Coulomb T matrix in the momentum representation, $\langle p|T_{cl}|p'\rangle$. Indeed, one and the same hypergeometric function, $F_{i\gamma}(\cdot) \equiv {}_2F_1(1,i\gamma;1+i\gamma;\cdot)$, plays an important part in the expressions for these quantities.[2] Moreover, $\langle p|kl\uparrow\rangle_c$ can be obtained from $\langle p|T_{cl}|p'\rangle$ by letting p' tend to infinity [Eq. (15)], and also from the Coulomb-modified form factors $\langle p|g^c_{\beta l}\rangle$, where $g_{\beta l}$ are the form factors of the so-called simple separable potentials, (cf. Ref. 2).

We shall perform a check on the hypergeometric-function expressions for $\langle p|kl\uparrow\rangle_c$ by deriving the Coulomb scattering state in momentum space, $\langle p|kl+\rangle_c$, in closed form. Such a closed form is known in the literature.[3,4]

We shall use the conventions and notations of Refs. 1 and 4. In particular, we put $\hbar = 2m = 1$, where m is the reduced mass, and $E \equiv k^2$ denotes the energy. The Coulomb potential is given by $V_c(r) = 2k\gamma/r$, where γ is Sommerfeld's parameter. The momentum variables p and p' are real positive. In this paper we shall assume for convenience that k and γ are real positive, too, which often facilitates the derivations. However, many formulas are also valid for complex k and γ. In some expressions it is essential that k has a (small) *positive* imaginary part. Whenever necessary, we will assume that the limit Im $k\downarrow 0$ is carried out, i.e., we replace k by $k + i\epsilon$ and let $\epsilon\downarrow 0$ [cf. Eq. (25)].

It is important to note that $\langle p|kl\uparrow\rangle_c$ is *not* a solution of the Schrödinger equation in momentum representation. This is related to the (at $r = 0$) singular behavior of the irregular solution, $\langle r|kl\uparrow\rangle_c$, of the Schrödinger equation.

2. THE JOST STATES FOR THE COULOMB POTENTIAL

The Coulomb Jost state in the momentum representation, $\langle p|kl\uparrow\rangle_c$, is defined as the Hankel transform of $\langle r|kl\uparrow\rangle_c$,

$$\langle p|kl\uparrow\rangle_c = \int_0^\infty \langle pl|r\rangle \langle r|kl\uparrow\rangle_c\, r^2\, dr, \tag{1}$$

where

$$\langle pl|r\rangle = (2/\pi)^{1/2}\, i^{-l} j_l(pr). \tag{2}$$

In the coordinate representation we have the well-known expression,

$$\langle r|kl\uparrow\rangle_c = (2/\pi)^{1/2} e^{\pi\gamma/2} (kr)^{-1}\, W_{-i\gamma,l+1/2}(-2ikr). \tag{3}$$

By using the integral representation (Ref. 5, p. 313)

$$W_{-i\gamma,l+1/2}(z) = (1/\Gamma(l+1+i\gamma))\, e^{(-1/2)z} z^{l+1} \times \int_0^\infty e^{-zt}\, t^{l+i\gamma}(1+t)^{l-i\gamma}\, dt$$

and the equality[6]

$$\int_0^\infty r^{l+2} e^{-\alpha r} j_l(pr)\,dr = 2\alpha(2p)^l (l+1)!(\alpha^2+p^2)^{-l-2}, \tag{4}$$

we obtain, after some manipulations,

$$\langle p|kl\uparrow\rangle_c = \frac{2(l+1)f_{cl}}{\pi p(p^2-k^2)}\left(\frac{-2}{v}\right)^{l+1} \times \int_0^1 \frac{(1-t^2)t^{l+i\gamma}}{[(1-ta)(1-t/a)]^{l+2}}\,dt. \tag{5}$$

Here $f_{cl} = f_{cl}(k)$ is the Coulomb Jost function,

$$f_{cl} = e^{\pi\gamma/2}\, l!/\Gamma(l+1+i\gamma),$$

and

$$a = (p-k)/(p+k),$$
$$v = (p^2-k^2)/(2pk).$$

0022-2488/83/051152-05$02.50

From Ref. 7, p. 238 we have, for any potential V_l,

$$\langle p|V_l|kl\uparrow\rangle = (k^2-p^2)\langle p|kl\uparrow\rangle + 2(\pi k)^{-1}(p/k)^l f_l. \quad (6)$$

From (5) and (6) we derive in the Appendix

$$\langle p|V_{cl}|kl\uparrow\rangle_c = \frac{2i\gamma}{\pi p} f_{cl}\left(\frac{-2}{v}\right)^{l+1} \times \int_0^1 \frac{t^{l+i\gamma}}{[(1-ta)(1-t/a)]^{l+1}}\,dt. \quad (7)$$

After carrying out some more manipulations we obtain from Eqs. (5) and (7), respectively, (see the Appendix),

$$\langle p|kl\uparrow\rangle_c = \frac{2f_{cl}}{\pi p(k^2-p^2)}[X_l(x) - x^{l+1} + \{F_{i\gamma}(a)-\tfrac{1}{2}\}P_l^{(-i\gamma,i\gamma)}(u) - \{F_{i\gamma}(1/a)-\tfrac{1}{2}\}P_l^{(i\gamma,-i\gamma)}(u)], \quad (8)$$

$$\langle p|V_{cl}|kl\uparrow\rangle_c = \frac{2f_{cl}}{\pi p}[X_l(x) + \{F_{i\gamma}(a)-\tfrac{1}{2}\}P_l^{(-i\gamma,i\gamma)}(u) - \{F_{i\gamma}(1/a)-\tfrac{1}{2}\}P_l^{(i\gamma,-i\gamma)}(u)]. \quad (9)$$

Here

$$F_{i\gamma}(\cdot) = {}_2F_1(1,i\gamma;1+i\gamma;\cdot),$$
$$x = p/k,$$
$$u = (p^2+k^2)/(2pk) = \tfrac{1}{2}(x+x^{-1}),$$

$P_l^{(\alpha,\beta)}$ is Jacobi's polynomial, and X_l is a simple rational function of x and of γ: $X_l = X_l(x) = X_l(x;\gamma)$. It is defined by the following recursion relation ($X_0 = 0$),

$$(l+1)X_{l+1} = \tfrac{1}{2}(l+1)(x+x^{-1})X_l + \tfrac{1}{2}(x^2-1)\frac{d}{dx}X_l + \gamma x \,\mathrm{Im}\, P_l^{(i\gamma,-i\gamma)}((x+x^{-1})/2), \quad (10)$$

where x and γ are supposed to be real. Denoting $P_l^{(i\gamma,-i\gamma)}(u)$ for the moment by P_l, we have (cf. Ref. 5)

$$P_0 = 1,$$
$$P_1 = u + i\gamma,$$
$$P_2 = \tfrac{1}{2}(3u^2 + 3i\gamma u - 1 - \gamma^2),$$
$$P_3 = \tfrac{1}{6}[15u^3 + 15i\gamma u^2 - u(9+6\gamma^2) - i\gamma(\gamma^2+4)].$$

We have found from (10)

$$X_0 = X_1 = 0,$$
$$X_2 = \tfrac{1}{2}\gamma^2 x,$$
$$X_3 = \tfrac{1}{12}\gamma^2(7x^2+5),$$
$$X_4 = (1/96x)\gamma^2(-4\gamma^2x^2 + 57x^4 + 48x^2 + 35). \quad (11a)$$

We list some interesting properties of X_l:

X_l is real when x and γ are real;

it is a polynomial in γ^2;

its parity is $(-)^{l+1}$: $x^{l+1}X_l(x)$ is even in x;

$$X_l(1) = 1 - \mathrm{Re}\binom{l+i\gamma}{l} \text{ [see Eq. (24)].}$$

We contend that the function Y_l, defined by

$$Y_l(x;\gamma) = 2^{l-2}\,l!\,\gamma^{-2}x^{l-3}X_l(x;\gamma) \quad (11b)$$

is a polynomial in x^2 and in γ^2, with real integer coefficients, that its degree in x^2 is $l-2$, and that its degree in γ^2 is Entier $(\frac{1}{2}l-1)$, for $l = 2,3,4,\cdots$. From Eq. (11a) we have

$$Y_0 = Y_1 = 0, \quad Y_2 = 1,$$
$$Y_3 = 7x^2 + 5, \quad (11c)$$
$$Y_4 = -4\gamma^2x^2 + 57x^4 + 48x^2 + 35.$$

We have obtained two representations for $\langle p|kl\uparrow\rangle_c$ [see Eqs. (5) and (8)]. Now we are going to deduce a third and a fourth one [see Eqs. (14) and (21)]. We shall utilize the equality

$$\lim_{\beta\to\infty}\beta^{2l+2}\langle p|g_{\beta l}^c\rangle = (\pi/2)^{1/2} f_{cl}^{-1}k^{l+1}(p^2-k^2)\langle p|kl\uparrow\rangle_c, \quad (12)$$

which will be proved in the Appendix. Here $g_{\beta l}^c$ is the Coulomb-modified form factor

$$\langle p|g_{\beta l}^c\rangle = \langle p|G_{0l}^{-1}G_{cl}|g_{\beta l}\rangle,$$

where

$$\langle p|g_{\beta l}\rangle = (2/\pi)^{1/2}p^l(p^2+\beta^2)^{-l-1}.$$

By using an expression for $\langle p|g_{\beta l}^c\rangle$ obtained in Ref. 2,

$$\langle p|g_{\beta l}^c\rangle = \frac{(2/\pi)^{1/2}}{p}\left(\frac{-2k/v}{\beta^2+k^2}\right)^{l+1} \times \sum_{n=l+1}^{\infty}\frac{n}{n+i\gamma}\left(\frac{\beta+ik}{\beta-ik}\right)^n C_{n-l-1}^{l+1}(u/v), \quad (13)$$

we get from Eq. (12)

$$\langle p|kl\uparrow\rangle_c = \frac{2f_{cl}}{\pi p(p^2-k^2)}\left(\frac{-2}{v}\right)^{l+1} \times \sum_{n=l+1}^{\infty}\frac{n}{n+i\gamma}C_{n-l-1}^{l+1}(u/v). \quad (14)$$

The infinite series in (13) and (14) are convergent when the energy k^2 is negative and $0 < p \neq p' > 0$. [We have derived Eq. (14) also in a different way: By comparing Eq. (5) with the integral representation for the Coulomb T matrix given by Eq. (24), p. 25 of Ref. 7 we get

$$\lim_{p'\to\infty}p'^{l+2}\langle p'|T_{cl}|p\rangle = \tfrac{1}{2}k\gamma(4k)^{l+1}(p^2-k^2)\langle p|kl\uparrow\rangle_c \times e^{-\pi\gamma/2}\Gamma(l+1+i\gamma)l!/(2l+1). \quad (15)$$

By substituting for $\langle p'|T_{cl}|p\rangle$ the infinite sum containing products of Gegenbauer polynomials given by Eq. (12) of Ref. 2, we easily obtain the verification of Eq. (14).]

By using

$$(-2/v)^{l+1}C_{n-l-1}^{l+1}(u/v) = a^n P_l^{(n,-n)}(u) - (-)^l a^{-n}P_l^{(n,-n)}(-u) \quad (16)$$

we obtain in the same way as in Ref. 2,

$$\langle p|V_{cl}|kl\uparrow\rangle_c = \frac{2f_{cl}}{\pi p}Z_l(a;1), \quad (17)$$

$$Z_l(a;1) = [\,F_{i\gamma}(a) - 1\,]\,P_l^{(-i\gamma,i\gamma)}(u)$$
$$- i\gamma\binom{2l}{l}(1-a^2)^{-l}\sum_{\nu=1}^{l-1} a^\nu/(\nu+i\gamma)$$
$$-(-)^l i\gamma\Gamma(l+1+i\gamma)\sum_{m=0}^{l-1}\binom{l+m}{l}\frac{(a^2-1)^{-m}}{(l-m)!}$$
$$\times\Bigg[\frac{1}{\Gamma(m+1+i\gamma)}\sum_{\nu=1}^{m}\frac{a^\nu}{\nu+i\gamma}$$
$$- a^m\sum_{\mu=0}^{l-m-1}\frac{\mu!}{\Gamma(\mu+m+2+i\gamma)}$$
$$\times\left(\frac{a}{a-1}\right)^{\mu+1}\Bigg] - (-l)^l\{\text{Idem},\ a\to 1/a\}. \qquad (18)$$

Here, and henceforth, $\{\text{Idem}, a\to 1/a\}$ means that all foregoing expressions on the right-hand side should be repeated after a has been replaced by $1/a$ everywhere. We have verified explicitly for $l=0,1,2$ that Eqs. (17) and (18) are in agreement with Eq. (9). Different expressions for $Z_l(a;1)$ are

$$Z_l(a;1) = \frac{i\gamma a}{l+1+i\gamma}\left(\frac{a}{1-a^2}\right)^l\sum_{n=0}^{l}\binom{2l-n}{l}(1-a^2)^n$$
$$\times {}_2F_1(n+1,l+1+i\gamma;l+2+i\gamma;a)$$
$$-(-)^l\{\text{Idem},\ a\to 1/a\}$$
$$= \frac{i\gamma a}{l+1+i\gamma}\left(\frac{a}{1-a}\right)^l\sum_{m=0}^{l}\binom{l+m}{l}(1+a)^{-m}$$
$$\times {}_2F_1(1,m+1+i\gamma;l+2+i\gamma;a)$$
$$-(-)^l\{\text{Idem},\ a\to 1/a\}. \qquad (19)$$

In particular the last expression is convenient for numerical calculation, because the sum of the first two parameters, minus the third one, of the hypergeometric function is equal to $m-l$, which is a nonpositive integer. Consequently the corresponding hypergeometric series converges for $|a|\leqslant 1$. (Except when $m=l$; in that case the value $a=1$ must be excluded.)

From Eqs. (9) and (17) we get

$$Z_l(a;1) = X_l(x) + i\,\mathrm{Im}\,P_l^{(i\gamma,-i\gamma)}(u)$$
$$+ F_{i\gamma}(a)\,P_l^{(-i\gamma,i\gamma)}(u)$$
$$- F_{i\gamma}(1/a)\,P_l^{(i\gamma,-i\gamma)}(u). \qquad (20)$$

By comparing this with Eq. (18) we have

$$X_l(x) - i\,\mathrm{Im}\,P_l^{(i\gamma,-i\gamma)}(u)$$
$$= -i\gamma\binom{2l}{l}(1-a^2)^{-l}\sum_{\nu=1}^{l-1}a^\nu/(\nu+i\gamma)$$
$$-(-)^l i\gamma\Gamma(l+1+i\gamma)\sum_{m=0}^{l-1}\binom{l+m}{l}\frac{(a^2-1)^{-m}}{(l-m)!}$$
$$\times\Bigg[\frac{1}{\Gamma(m+1+i\gamma)}\sum_{\nu=1}^{m}\frac{a^\nu}{\nu+i\gamma} - a^m$$
$$\times\sum_{\mu=0}^{l-m-1}\frac{\mu!}{\Gamma(\mu+m+2+i\gamma)}\left(\frac{a}{a-1}\right)^{\mu+1}\Bigg]$$
$$-(-)^l\{\text{Idem},\ a\to 1/a\}. \qquad (21)$$

We want to derive a simple expression for $X_l(1)$. When $x=1, p=k, a=0$, and $u=1$. From Eq. (21) we obtain

$$X_l(1) - i\,\mathrm{Im}\,P_l^{(i\gamma,-i\gamma)}(1)$$
$$= -\frac{\Gamma(l+1+i\gamma)}{l!}\,i\gamma\sum_{\mu=0}^{l-1}\frac{\mu!}{\Gamma(\mu+2+i\gamma)}.$$

By using

$$P_l^{(i\gamma,-i\gamma)}(1) = \binom{l+i\gamma}{l} \qquad (22)$$

and (cf. Ref. 8)

$$i\gamma\sum_{\mu=0}^{l-1}\frac{\Gamma(\mu+1)}{\Gamma(\mu+2+i\gamma)} = \frac{1}{\Gamma(1+i\gamma)} - \frac{\Gamma(l+1)}{\Gamma(l+1+i\gamma)}, \qquad (23)$$

we get

$$X_l(1) = 1 - \binom{l+i\gamma}{l} + i\,\mathrm{Im}\binom{l+i\gamma}{l},$$

i.e.,

$$X_l(1) = 1 - \mathrm{Re}\binom{l+i\gamma}{l} = 1 - \mathrm{Re}\,P_l^{(i\gamma,-i\gamma)}(1). \qquad (24)$$

Finally we shall derive a simple expression for $\langle p|V_{cl}|kl+\rangle_c$ from the expression for $\langle p|V_{cl}|kl\uparrow\rangle_c$ given by Eq. (9). We have

$$2ie^{-i\sigma_l}\langle p|V_{cl}|kl+\rangle_c$$
$$= e^{i\sigma_l}\langle p|V_{cl}|kl\uparrow\rangle_c - e^{-i\sigma_l}\langle p|V_{cl}|kl\downarrow\rangle_c$$
$$= e^{i\sigma_l}\langle p|V_{cl}|kl\uparrow\rangle_c - \text{c.c.}$$

Since X_l is real and $e^{i\sigma_l}f_{cl}$ is real for real k and γ, we obtain from (9)

$$\langle p|V_{cl}|kl+\rangle_c = i(\pi p)^{-1}f_{cl}^*[\{-1+F_{i\gamma}(1/a)$$
$$+F_{-i\gamma}(a^*)\}\,P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}].$$

We use the equality

$$F_{-i\gamma}(a) + F_{i\gamma}(1/a) = 1 + \Gamma(1+i\gamma)\Gamma(1-i\gamma)(-a)^{i\gamma},$$

and

$$F_{-i\gamma}(a) - F_{-i\gamma}(a^*) = 0,$$
$$(-a)^{i\gamma} = e^{-\pi\gamma}a^{i\gamma}.$$

[Note that the sign of Im k is important here: $a=(p-k-i\epsilon)/(p+k+i\epsilon)$, $\epsilon\downarrow 0$.] In this way we obtain

$$\langle p|V_{cl}|kl+\rangle_c = \frac{i}{\pi p}c_{l\gamma}f_{cl}^{-1}\,[a^{i\gamma}P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}], \qquad (25)$$

where

$$c_{l\gamma}^{-1} = \binom{l+i\gamma}{l}\binom{l-i\gamma}{l} = \prod_{n=1}^{l}(1+\gamma^2/n^2),$$
$$f_{cl}^{-1} = e^{-\pi\gamma/2}\Gamma(l+1+i\gamma)/l!.$$

Equation (25) is in full agreement with a previously obtained result [see Ref. 4, Eq. (7.7)].

Summarizing, we have obtained an integral representation for $\langle p|kl\uparrow\rangle_c$ and for $\langle p|V_{cl}|kl\uparrow\rangle_c$ [Eqs. (5) and (7)], and expressions containing the hypergeometric function $F_{i\gamma}$, the Jacobi polynomial $P_l^{(i\gamma,-i\gamma)}$, and a simple rational function X_l [Eqs. (8) and (9)]. For the function X_l we have given:

(i) explicit expressions for $l=0,1,2,3,4$ [Eq. (11)],
(ii) a recursion relation [Eq. (10)],
(iii) a finite-series expression [Eq. (21)], and
(iv) a simple explicit expression in the special case $p=k$ [Eq. (24)].

ACKNOWLEDGMENT

This work is part of the project "The Coulomb Potential in Quantum Mechanics and Related Topics."

APPENDIX

In this Appendix we shall prove (i) Eq. (7), (ii) Eqs. (9) and (10), and (iii) Eq. (12).

(i) In order to derive Eq. (7) from Eqs. (5) and (6), we introduce the function h_l,

$$h_l(z) = i\gamma \int_0^1 (1 + t^2 - tz)^{-l-1} t^{i\gamma + l} dt, \qquad \text{(A1)}$$

where $z = a + a^{-1}$ [cf. Eq. (7)]. We rewrite h_l and perform integration by parts

$$\begin{aligned} h_l(z) &= i\gamma \int_0^1 (t + t^{-1} - z)^{-l-1} t^{i\gamma - 1} dt \\ &= [(t + t^{-1} - z)^{-l-1} t^{i\gamma}]_0^1 \\ &\quad - \int_0^1 (t + t^{-1} - z)^{-l-2}(l+1)(t^{-2} - 1) t^{i\gamma} dt \\ &= (2 - z)^{-l-1} - (l+1) \\ &\quad \times \int_0^1 (1 + t^2 - tz)^{-l-2}(1 - t^2) t^{i\gamma + l} dt. \qquad \text{(A2)} \end{aligned}$$

By substituting

$$2 - z = 2 - a - a^{-1} = 4k^2/(k^2 - p^2)$$

the proof of Eq. (7) follows easily from Eqs. (A1) and (A2).

(ii) Now we shall derive Eqs. (9) and (10) from Eq. (7). By using Eq. (A1) and inserting

$$a - a^{-1} = 4pk/(k^2 - p^2) = -2/v,$$

Eq. (9) can be rewritten as

$$\begin{aligned} (a - a^{-1})^{l+1} h_l(z) &= X_l(x) + i \operatorname{Im} P_l^{(i\gamma, -i\gamma)}(u) \\ &\quad + F_{i\gamma}(a) P_l^{(-i\gamma, i\gamma)}(u) \\ &\quad - F_{i\gamma}(1/a) P_l^{(i\gamma, -i\gamma)}(u). \qquad \text{(A3)} \end{aligned}$$

We shall prove this equation by induction on l. First we shall verify explicitly that Eq. (A3) is valid for $l = 0$. Inserting $X_0 = 0$, $P_0^{(\cdot,\cdot)} = 1$, and

$$F_{i\gamma}(a) = i\gamma \int_0^1 (1 - ta)^{-1} t^{i\gamma - 1} dt, \qquad \text{(A4)}$$

the right member of Eq. (A3) becomes

$$\begin{aligned} & i\gamma \int_0^1 \left(\frac{1}{1 - ta} - \frac{1}{1 - t/a} \right) t^{i\gamma - 1} dt \\ &= (a - a^{-1}) i\gamma \int_0^1 [1 + t^2 - t(a + a^{-1})]^{-1} t^{i\gamma} dt \\ &= (a - a^{-1}) h_0(z), \end{aligned}$$

which completes the first step of the induction proof.

Now we differentiate both members of Eq. (A3) with respect to $x \equiv p/k$. We recall

$$a - a^{-1} = \frac{4x}{1 - x^2},$$

$$a + a^{-1} = z = 2\,\frac{x^2 + 1}{x^2 - 1},$$

$$a = \frac{x - 1}{x + 1}, \quad u = \frac{x^2 + 1}{2x},$$

and use

$$\frac{d}{dz} F_{i\gamma}(z) = \frac{i\gamma}{z(1 - z)} - \frac{i\gamma}{z} F_{i\gamma}(z), \qquad \text{(A5)}$$

$$\begin{aligned} &(1 - z^2) \frac{d}{dz} P_l^{(i\gamma, -i\gamma)}(z) \\ &\quad = (i\gamma + lz + z) P_l^{(i\gamma, -i\gamma)}(z) - (l+1) P_{l+1}^{(i\gamma, -i\gamma)}(z), \qquad \text{(A6)} \end{aligned}$$

$$\frac{d}{dz} h_l(z) = (l+1)\, h_{l+1}(z). \qquad \text{(A7)}$$

Introducing the (for real γ and u) real functions R_l and I_l by

$$P_l^{(i\gamma, -i\gamma)}(u) \equiv R_l + iI_l$$

we obtain in this way from Eq. (A3),

$$\begin{aligned} & X_l'(x) + i\,\frac{x^2 - 1}{2x^2} I_l' + \frac{2i\gamma}{(x^2 - 1)(1 - a)} (R_l - iI_l) \\ &\quad + \frac{2i\gamma}{(x^2 - 1)(1 - a^{-1})} (R_l + iI_l) \\ &\quad + F_{i\gamma}(a) \left[\frac{-2i\gamma}{x^2 - 1} P_l^{(-i\gamma, i\gamma)}(u) \right. \\ &\quad \left. + \frac{x^2 - 1}{2x^2} \frac{d}{du} P_l^{(-i\gamma, i\gamma)}(u) \right] \\ &\quad - F_{i\gamma}(a^{-1}) \left[\frac{2i\gamma}{x^2 - 1} P_l^{(i\gamma, -i\gamma)}(u) \right. \\ &\quad \left. + \frac{x^2 - 1}{2x^2} \frac{d}{du} P_l^{(i\gamma, -i\gamma)}(u) \right] \\ &\quad = \frac{l+1}{x} \frac{1 + x^2}{1 - x^2} [X_l + iI_l + F_{i\gamma}(a) P_l^{(-i\gamma, i\gamma)}(u) \\ &\quad - F_{i\gamma}(a^{-1}) P_l^{(i\gamma, -i\gamma)}(u)] \\ &\quad - \frac{2(l+1)}{1 - x^2} [X_{l+1} + iI_{l+1} + F_{i\gamma}(a) P_{l+1}^{(-i\gamma, i\gamma)}(u) \\ &\quad - F_{i\gamma}(a^{-1}) P_{l+1}^{(i\gamma, -i\gamma)}(u)]. \qquad \text{(A8)} \end{aligned}$$

The coefficients of the corresponding hypergeometric functions in both members turn out to be equal. By equating real and imaginary parts we obtain

$$(l+1) X_{l+1} = (l+1)\,\frac{x^2 + 1}{2x} X_l + \tfrac{1}{2}(x^2 - 1) X_l' + \gamma x I_l, \qquad \text{(A9)}$$

and

$$\frac{(x^2 - 1)^2}{4x^2} I_l' + \gamma R_l = (l+1) I_{l+1} - \frac{x^2 + 1}{2x} (l+1) I_l. \qquad \text{(A10)}$$

Here Eq. (A9) is just Eq. (10) rewritten and Eq. (A10) is an equality which can be verified independently. We note that I_l' here means $(d/du) I_l$.

(iii) Finally we shall prove Eq. (12). In fact we shall prove a more general relation, which holds for any local potential V:

$$\begin{aligned} & \lim_{\beta \to \infty} \beta^{2l+2} \langle p | g_{\beta l}^V \rangle \\ &\quad = (\pi/2)^{1/2} f_l^{-1} k^{l+1} (p^2 - k^2) \langle p | kl \uparrow \rangle_V, \qquad \text{(A11)} \end{aligned}$$

where f_l is the Jost function corresponding to V. Let G_l and T_l be the resolvent and the transition operator associated with V. Then

$$G_l = G_{0l} + G_{0l} T_l G_{0l},$$

and

$$|g_{\beta l}^V\rangle = G_{0l}^{-1} G_l\, |g_{\beta l}\rangle. \qquad \text{(A12)}$$

Since

$$\langle p | g_{\beta l} \rangle = (2/\pi)^{1/2} p^l (\beta^2 + p^2)^{-l-1}, \qquad \text{(A13)}$$

we also have

$$\lim_{\beta\to\infty} \beta^{2l+2} \langle p|g^V_{\beta l}\rangle$$
$$= (2/\pi)^{1/2}(k^2-p^2)\int_0^\infty \langle p|G_l|p'\rangle\, p'^{l+2}\,dp' \qquad \text{(A14)}$$
$$= (2/\pi)^{1/2}p^l - (2/\pi)^{1/2} \times \int_0^\infty \langle p|T_l|p'\rangle \frac{p'^{l+2}}{p'^2-k^2}\,dp'. \qquad \text{(A15)}$$

In order to prove Eq. (A11), we use

$$\langle r'|G_l|r\rangle = (-1)^{l+1}\tfrac{1}{2}\pi k \langle r_<|kl+\rangle \langle r_>|kl\uparrow\rangle, \qquad \text{(A16)}$$

which holds for any local potential, and

$$\lim_{r\to 0} r^{-l}\langle r|kl+\rangle = (2/\pi)^{1/2} f_l^{-1}(2ik)^l l!/(2l+1)!. \qquad \text{(A17)}$$

The coordinate representation of $g_{\beta l}$ is given by [cf. (A13)]

$$\langle r|g_{\beta l}\rangle = r^{l-1}e^{-\beta r}(i/2)^l/l!. \qquad \text{(A18)}$$

From (A16) and (A18) we derive

$$\lim_{\beta\to\infty}\beta^{2l+2}\langle r|G_l|g_{\beta l}\rangle$$
$$= (-)^{l+1}\tfrac{1}{2}\pi k\,(l!)^{-1}(i/2)^l\langle r|kl\uparrow\rangle \times \lim_{\beta\to\infty}\beta^{2l+2}\int_0^\infty t^{l+1}e^{-t}\beta^{-l-2}\left\langle \frac{t}{\beta}\middle| kl+\right\rangle dt.$$

By using Eq. (A17) and $\int_0^\infty t^{2l+1}e^{-t}\,dt = (2l+1)!$ we easily obtain

$$\lim_{\beta\to\infty}\beta^{2l+2}\langle r|G_l|g_{\beta l}\rangle = -(\pi/2)^{1/2}f_l^{-1}k^{l+1}\langle r|kl\uparrow\rangle, \qquad \text{(A19)}$$

hence

$$\lim_{\beta\to\infty}\beta^{2l+2}\langle p|G_l|g_{\beta l}\rangle = -(\pi/2)^{1/2}f_l^{-1}k^{l+1}\langle p|kl\uparrow\rangle. \qquad \text{(A20)}$$

Substitution of (A12) completes the proof of Eq. (A11).

[1] H. van Haeringen, Phys. Lett. A **86**, 359 (1981).

[2] H. van Haeringen, "Off-shell T matrix for Coulomb plus simple separable potentials for all l in closed form," preprint, Delft University of Technology, 1981.

[3] E. I. Dolinskii and A. M. Mukhamedzhanov, Sov. J. Nucl. Phys. **3**, 180 (1966).

[4] H. van Haeringen, "Long-Range Potentials in Quantum Mechanics," Report 150, University of Groningen, 1979.

[5] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966).

[6] A. Erdélyi, *Tables of Integral Transforms* (McGraw-Hill, New York, 1954), Vol. 2, p. 29, Eq. (4).

[7] H. van Haeringen, Ph. D. thesis (Free University, Amsterdam, 1978).

[8] A. Erdélyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1, p. 191.

The unitarity relations for the off-shell Coulomb T matrix for all partial waves

H. van Haeringen
Department of Mathematics, Delft University of Technology, Delft, The Netherlands

(Received 1 December 1981; accepted for publication 5 March 1982)

By employing a recently obtained expression for the partial-wave projection of the off-shell Coulomb T matrix for all l, we prove that the unitarity relations for the Coulomb T matrix hold provided that they are properly modified with the help of Coulombian asymptotic states.

PACS numbers: 03.65.Nk

The well-known optical theorem establishes a simple relation between on-shell matrix elements of the transition (T) matrix associated with any short-range interaction. For the partial-wave projected T operator T_l this relation is given by

$$\mathrm{Im}\langle k|T_l|k\rangle = -\tfrac{1}{2}\pi k\,|\langle k|T_l|k\rangle|^2, \quad k>0. \tag{1}$$

Here the units are chosen such that $\hbar = 2m = 1$ and the energy-dependence of $T_l = T_l(E)$, where $E=(k+i\epsilon)^2$, $\epsilon\downarrow 0$, is suppressed.

Equation (1) is a direct consequence of the unitarity of the scattering (S) matrix, $S^\dagger S = SS^\dagger = 1$. The so-called unitarity relation for the *off-shell* T matrix (cf. Refs. 1 and 2) is a generalization of Eq. (1). It can be expressed by

$$\mathrm{Im}\langle p|T_l|p'\rangle = -\tfrac{1}{2}\pi k\,\langle p|T_l|k\rangle\langle k|T_l^\dagger|p'\rangle, \quad k>0, \tag{2}$$

where the momenta p and p' are real positive. We note that the off-shell T matrix $\langle p|T_l|p'\rangle$ is symmetric in p and p'.

The unitarity relations (1) and (2) are not valid for the Coulomb T matrix, T_{cl}. As is well known, $\langle p|T_{cl}|p'\rangle$ has no half-shell or on-shell limit (i.e., for $p\to k$, $p'\to k$). There exists a simple prescription for dealing with the half-shell and on-shell singularity of $\langle p|T_{cl}|p'\rangle$. A convenient and consistent notation and prescription is provided by the so-called Coulombian asymptotic state $|k\infty\rangle$. If applied to T_{cl}, this state can be expressed by [see Ref. 3, Eq. (16)]

$$\begin{aligned}\langle p|kl\infty\rangle &\equiv \langle p|k\infty\rangle\\ &= k^{-2}\delta(p-k)[2k/(p-k-i\epsilon)]^{i\gamma}\\ &\quad\times e^{\pi\gamma/2}/\Gamma(1-i\gamma), \quad \epsilon\downarrow 0,\end{aligned} \tag{3}$$

where γ is Sommerfeld's parameter. The connection with the time-dependent Coulomb scattering theory has been given in Ref. 3. By replacing in Eq. (2) $|k\rangle$ by $|k\infty\rangle$ and T_l by T_{cl} we get the unitarity relation for the off-shell Coulomb T matrix:

$$\mathrm{Im}\langle p|T_{cl}|p'\rangle = -\tfrac{1}{2}\pi k\,\langle p|T_{cl}|k\infty\rangle\,\langle k\infty|T_{cl}^\dagger|p'\rangle, \quad k>0. \tag{4}$$

The main purpose of this paper is to prove Eq. (4) by using the explicit expression for $\langle p|T_{cl}|p'\rangle$ that we have recently obtained.[4]

We shall also prove the equality

$$\langle p|V_{cl}|kl+\rangle_c = \langle p|T_{cl}|k\infty\rangle, \tag{5}$$

where $|kl+\rangle_c$ is the Coulomb scattering state with energy k^2.

Furthermore, we shall prove [cf. Eq. (9.73) of Ref. 5]

$$\langle k\infty-|T_{cl}|k\infty\rangle \sim i(\pi k)^{-1}[e^{2i\sigma_l} - e^{2i\sigma_0}(2k/\epsilon)^{2i\gamma}], \quad \epsilon\downarrow 0, \tag{6}$$

where σ_l is the Coulomb phase shift. From Eq. (6) we shall derive that summation of the partial-wave series, *if* summed in the proper way, i.e., with exclusion of the forward direction (see below), just gives the Coulomb scattering amplitude.

In Ref. 4 we have proved

$$\langle p|T_{cl}|p'\rangle = \frac{-k\gamma}{\pi pp'}c_{l\gamma}\left[(i\gamma)^{-1}\mathscr{F}_l + \mathscr{G}_l + \mathscr{L}_l\ln\left(\frac{p+p'}{p-p'}\right)^2\right]. \tag{7}$$

Here

$$c_{l\gamma}^{-1} = \binom{l+i\gamma}{l}\binom{l-i\gamma}{l} = \prod_{n=1}^{l}(1+\gamma^2/n^2),$$

$$\begin{aligned}\mathscr{F}_l = {}& F_{i\gamma}(aa')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u')\\ &+ F_{i\gamma}((aa')^{-1})P_l^{(i\gamma,-i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u')\\ &- F_{i\gamma}(a/a')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u')\\ &- F_{i\gamma}(a'/a)P_l^{(i\gamma,-i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u'),\end{aligned} \tag{8}$$

where ($k = k+i\epsilon, \epsilon\downarrow 0$)

$$a=(p-k)/(p+k), \quad a'=(p'-k)/(p'+k),$$
$$u=(p^2+k^2)/(2pk), \quad u'=(p'^2+k^2)/(2p'k),$$

and

$$F_{i\gamma}(\cdot) = {}_2F_1(1,i\gamma;1+i\gamma;\cdot).$$

Closed expressions for $\mathscr{G}_l$ and $\mathscr{L}_l$ have been given in Ref. 4. Here we need these expressions in the particular case $p'=k$, and moreover we need $\mathrm{Im}\,\mathscr{G}_l$ and $\mathrm{Im}\,\mathscr{L}_l$. Assuming that p, p', k, and γ are real, we have

$$\mathrm{Im}\mathscr{L}_l = 0,$$
$$\mathrm{Im}\mathscr{G}_l = -2\gamma^{-1}\mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u)\mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u'). \tag{9}$$

Furthermore,

$$\mathscr{L}_l(p'=k) = 0,$$
$$\mathscr{G}_l(p'=k) = 2\gamma^{-1}\binom{l-i\gamma}{l}\mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u). \tag{10}$$

These equations are sufficient to prove the Coulomb unitarity relation given by Eq. (4), and to prove Eqs. (5) and (6).

We begin by evaluating $\langle p|T_{cl}|k\infty\rangle$. When $p'\to k$, $a'\to 0$ and $u'\to 1$, hence $P_l^{(i\gamma,-i\gamma)}(u')\to\binom{l+i\gamma}{l}$. Considering Eq. (8), we observe that $F_{i\gamma}(aa')$ and $F_{i\gamma}(a'/a)$ tend to 1. The two remaining hypergeometric functions, $F_{i\gamma}(a/a')$ and $F_{i\gamma}((aa')^{-1})$, have to be transformed. By using the equality

$$F_{-i\gamma}(z) + F_{i\gamma}(1/z) = 1 + \Gamma(1+i\gamma)\Gamma(1-i\gamma)(-z)^{i\gamma}, \tag{11}$$

and Eqs. (8)–(10), we obtain

$$\langle p|T_{cl}|p'\rangle \underset{p'\to k}{\rightarrow} \frac{ik}{\pi pp'} c_{l\gamma} \binom{l+i\gamma}{l} \Gamma(1+i\gamma)\Gamma(1-i\gamma) \times [(-aa')^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) - (-a'/a)^{i\gamma} P_l^{(-i\gamma,i\gamma)}(u)].$$

By applying Eq. (3), we obtain, after some manipulations,

$$\langle p|T_{cl}|k\infty\rangle = \frac{i}{\pi p} e^{-\pi\gamma/2} \frac{|\Gamma(1+i\gamma)|^2 l!}{\Gamma(l+1-i\gamma)} \times \lim_{\epsilon\downarrow 0} [a^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}], \tag{12}$$

which is just equal to $\langle p|V_{cl}|kl+\rangle_c$, according to Eq. (7.7) of Ref. 5. This completes the proof of Eq. (5).

In order to prove Eq. (4), we have to evaluate the left-hand side. From Eqs. (7) and (9) we have

$$\mathrm{Im}\langle p|T_{cl}|p'\rangle = (\pi pp')^{-1} k c_{l\gamma} (\mathrm{Re}\mathscr{F}_l - \gamma\, \mathrm{Im}\mathscr{G}_l). \tag{13}$$

Assuming for definiteness $0<k<p<p'$, we obtain from Eq. (8), by using Eq. (11),

$$\begin{aligned}\mathscr{F}_l + \mathscr{F}_l^* = {} & |\Gamma(1+i\gamma)|^2 [(-aa')^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) P_l^{(i\gamma,-i\gamma)}(u') \\ & - (-a/a')^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) P_l^{(-i\gamma,i\gamma)}(u')] + \text{c.c.} \\ & + [P_l^{(i\gamma,-i\gamma)}(u) - P_l^{(-i\gamma,i\gamma)}(u)] \\ & \times [P_l^{(i\gamma,-i\gamma)}(u') - P_l^{(-i\gamma,i\gamma)}(u')].\end{aligned} \tag{14}$$

A careful analysis of the branch cuts leads to

$$(-aa')^{i\gamma} = e^{-\pi\gamma}|aa'|^{i\gamma}, \quad (-a/a')^{i\gamma} = e^{-\pi\gamma}|a/a'|^{i\gamma}. \tag{15}$$

Details can be found in Appendix F of Ref. 5.

By inserting the expression for $\mathrm{Im}\mathscr{G}_l$ given by (9) into Eq. (13), and using Eqs. (14) and (15), we obtain

$$\begin{aligned}\mathrm{Im}\langle p|T_{cl}|p'\rangle = {} & (2\pi pp')^{-1} k e^{-\pi\gamma} |\Gamma(1+i\gamma)|^2 c_{l\gamma} \\ & \times (|a|^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}) \\ & \times (|a'|^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u') - \text{c.c.}).\end{aligned} \tag{16}$$

The proof of Eq. (4) is easily obtained from Eqs. (12) and (16).

Finally we want to derive Eq. (6). We consider the quantity $\langle p|T_{cl}|k\infty\rangle$ for $p\to k$. From Eq. (12) we get

$$\langle p|T_{cl}|k\infty\rangle \underset{p\to k}{\rightarrow} i(\pi k)^{-1} e^{-\pi\gamma/2} \Gamma(1-i\gamma) \times [a^{i\gamma} e^{2i\sigma_l} - a^{*-i\gamma} e^{2i\sigma_l}], \tag{17}$$

where

$$e^{2i\sigma_l} = \Gamma(l+1+i\gamma)/\Gamma(l+1-i\gamma).$$

We apply $\langle k\infty -|$ to both members of Eq. (17). According to Eq. (3), we have to multiply the right member by

$$a^{-i\gamma} e^{\pi\gamma/2}/\Gamma(1-i\gamma),$$

and take $p\to k$, $\epsilon\downarrow 0$. We point out that

$$a^{-i\gamma} a^{*-i\gamma} \to (2k/\epsilon)^{2i\gamma},$$

whereas $a^{i\gamma}a^{-i\gamma} = 1$. In this way the proof of Eq. (6) is completed (see also Sec. 9D of Ref. 5).

Equation (17) shows that the on-shell limit ($p\to k$) of the physical half-shell partial-wave Coulomb T matrix $\langle p|T_{cl}|k\infty\rangle$ does not exist. Also, the on-shell limits ($p\to k$, or $p'\to k$) of the off-shell partial-wave Coulomb T matrix $\langle p|T_{cl}|p'\rangle$ do not exist. This is analogous to the situation for the three-dimensional Coulomb T matrix $\langle\vec{p}|T_c|\vec{p}'\rangle$. The on-shell limits ($p\to k$, or $p'\to k$) of this quantity do not exist. Also, the on-shell limit of the physical half-shell Coulomb T matrix $\langle\vec{p}|T_c|\vec{k}\infty\rangle$ does not exist (cf. Ref. 6).

Application of the three-dimensional Coulombian asymptotic states $|\vec{k}\infty\rangle$ to the three-dimensional Coulomb T matrix *does* give the right Coulomb scattering amplitude, i.e.,

$$\langle\vec{k}'\infty -|T_c|\vec{k}\infty\rangle = -(2\pi^2)^{-1} f^c(\hat{k}\cdot\hat{k}'), \quad \hat{k}'\neq\hat{k}, \quad k'=k\in\mathbb{R}^+,$$

where

$$f^c(x) = -(\gamma/2k) e^{2i\sigma_0} (\tfrac{1}{2} - \tfrac{1}{2}x)^{-1-i\gamma}. \tag{18}$$

In contrast, application of the partial-wave Coulombian asymptotic states $|k\infty\rangle$ [cf. Eq. (3)] to the partial-wave Coulomb T matrix *does not* lead to well-defined on-shell limits. This is clear from Eq. (6). However, it is interesting to note that the singular part of $\langle k\infty -|T_{cl}|k\infty\rangle$, containing the singular factor $\epsilon^{-2i\gamma}$, is independent of l. Consequently, if we sum the partial-wave series before taking the limit $\epsilon\downarrow 0$, this singular term gives a contribution that is proportional to $\epsilon^{-2i\gamma}\delta(1-\cos\theta)$, as is clear from the equality

$$\sum_{l=0}^{\infty} (l+\tfrac{1}{2}) P_l(x) = \delta(1-x).$$

As discussed by Taylor,[7] the Coulomb partial-wave series $\Sigma(l+\frac{1}{2})P_l(\cos\theta)\exp(2i\sigma_l)$ can be summed in the sense of distributions if only test functions are allowed that vanish in the forward direction, i.e., for $\theta=0$. Formally, this comes down to putting $\delta(1-\cos\theta)=0$. Therefore, the singular part of $\langle k\infty -|T_{cl}|k\infty\rangle$ containing the factor $\epsilon^{-2i\gamma}$ vanishes by this procedure. By summing the remaining nonsingular part in the partial-wave series we get

$$\sum_{l=0}^{\infty} (4\pi)^{-1}(2l+1) P_l(\hat{k}\cdot\hat{k}') \langle k\infty -|T_{cl}|k\infty\rangle = \langle\vec{k}\infty -|T_c|\vec{k}\infty\rangle, \quad k=k'>0, \tag{19}$$

which follows from Refs. 3 and 7, according to

$$\sum_{l=0}^{\infty} (l+\tfrac{1}{2}) P_l(x) \exp(2i\sigma_l) = ik f^c(x), \tag{20}$$

where f^c is given by Eq. (18). The partial-wave series for on-shell T matrices associated with *short*-range potentials can be written as

$$\sum_{l=0}^{\infty} (4\pi)^{-1}(2l+1) P_l(\hat{k}\cdot\hat{k}') \langle k|T_l|k\rangle = \langle\vec{k}|T|\vec{k}'\rangle, \quad k=k'>0. \tag{21}$$

Clearly Eq. (19) is just the Coulomb analog of this equation. In this way we have obtained the satisfactory result that the partial-wave series of the "on-shell" Coulomb T matrix with the Coulombian asymptotic states, if summed in the proper way (i.e., excluding the forward direction), gives the Coulomb scattering amplitude.

In this paper we have derived, by using *exact analytic* expressions, the Coulomb generalization of the following relations familiar in short-range potential scattering theory: (i) The off-shell unitarity relation for the partial-wave T matrix for all l [Eq. (4)]; (ii) The relation $\langle p|V_l|kl+\rangle = \langle p|T_l|k\rangle$ for all l [Eq. (15)]; (iii) The summation of the partial-wave T matrices [Eqs. (6) and (19)]. The restriction to the pure Coulomb potential is not serious. For Coulomb plus short-range potentials one can construct corresponding quantities from the pure Coulomb ones in a well-known way.[3,8]

Nuttall and Stagat[9] have evaluated a modified unitarity relation for the *three-dimensional* Coulomb T matrix for restricted values of the momenta. In the simple and convenient notation introduced in Ref. 3 it is expressed by[10]

$$\mathrm{Im}\langle\vec{p}|T_c|\vec{p}'\rangle = -\frac{1}{2}\pi k \int \langle\vec{p}|T_c|\vec{k}\infty\rangle\,\langle\vec{k}\infty|T_c^\dagger|\vec{p}'\rangle\, d\hat{k}\,, \tag{22}$$

where the integration is over the unit sphere, $|\hat{k}| = 1$. The similarity of Eqs. (4) and (22) is consistent and satisfactory. The restriction on the momenta in Ref. 9 can be easily removed,[10] as has been noted by Chen and Chen.[11] These authors have shown that the off-shell unitarity relation for the three-dimensional Coulomb T-matrix does not have a well-defined on-shell limit. See also their review article on off-shell Coulomb amplitudes.[12]

Our proof of Eq. (4) presents a complement to this work on the three-dimensional Coulomb unitarity relation. This explicit proof could be given since a closed formula for $\langle p|T_{cl}|p'\rangle$ became recently available.[4]

In conclusion it appears that there is no "violence of unitarity" in Coulomb scattering. One only has to deal with singularities that are more complicated than poles and Dirac delta functions (distributions). Instead one encounters in Coulomb-scattering quantities branch-point singularities and more complicated (than δ) distributions, respectively. The modification we have introduced takes care of these singularities in a well-defined and elegant manner.

ACKNOWLEDGMENTS

It is a pleasure to thank L. P. Kok for discussions and comments. This work is part of the project "The Coulomb Potential in Quantum Mechanics and Related Topics."

[1] R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).

[2] K. M. Watson and J. Nuttall, *Topics in Several Particle Dynamics* (Holden–Day, San Francisco, 1967).

[3] H. van Haeringen, J. Math. Phys. **17**, 995 (1976).

[4] H. van Haeringen, "The partial-wave projected Coulomb T matrix for all l in closed hypergeometric form," preprint, Delft University of Technology, 1981.

[5] H. van Haeringen, "Long-Range Potentials in Quantum Mechanics," Report 150, University of Groningen, 1979.

[6] L. P. Kok and H. van Haeringen, Phys. Rev. Lett. **46**, 1257 (1981).

[7] J. R. Taylor, Nuovo Cimento B **23**, 313 (1974).

[8] H. van Haeringen and R. van Wageningen, J. Math. Phys. **16**, 1441 (1975).

[9] J. Nuttall and R. W. Stagat, Phys. Rev. A **3**, 1355 (1971).

[10] See Ref. 5, Chap. 9, especially Eq. (9.39).

[11] A. C. Chen and J. C. Y. Chen, Phys. Rev. A **4**, 2226 (1971).

[12] J. C. Y. Chen and A. C. Chen, in "Advances in Atomic and Molecular Physics," edited by D. R. Bates and I. Esterman (Academic, New York, 1972), Vol. 8, p. 71.

Existence and uniqueness of bound-state eigenvectors for some channel coupling Hamiltonians

J. W. Evans
Ames Laboratory[a] and Department of Chemistry, Iowa State University, Ames, Iowa 50011

(Received 9 August 1982; accepted for publication 18 November 1982)

For the three-particle, two-cluster, 2×2 channel coupling Hamiltonians used, e.g., in H_2^+ and He bound-state calculations, we demonstrate that typically there exist unique eigenvectors for all bound states. This result also holds, with some technical assumptions on the potentials, for the corresponding 3×3 case provided there are no spurious eigenvectors with bound-state eigenvalues. The proofs use the analogous results for the corresponding Faddeev-type Hamiltonians together with spurious multiplier relationships.

PACS numbers: 03.65.Nk, 03.65.Ge

I. INTRODUCTION

Inherent difficulties with the standard Lippmann–Schwinger equation approach to many-body scattering theory[1] have lead to the development of a variety of alternative approaches. These are often based on decomposition of the wave function or T-matrices into components associated with various clusterings (arrangement channels) of the particles.[2] The "arrangement channel quantum mechanics" approach[3] for a system of N nonrelativistic particles is characterized by a non-Hermitian Hamiltonian H with components $H_{\alpha\beta}$ labeled by some subset of the N particle clusterings $\alpha, \beta, \cdots$. These satisfy[4]

$$\sum_{\alpha} H_{\alpha\beta} = H \quad \text{for all } \beta, \tag{1}$$

where H is the N particle Hilbert space Hamiltonian. From Eq. (1), any eigenvector $\boldsymbol{\psi}$ of H with components $|\psi_\alpha\rangle$ and eigenvalue λ is either (a) *physical* satisfying $\Sigma_\alpha |\psi_\alpha\rangle = |\psi\rangle \neq 0$ and $\lambda = E$, where $H|\psi\rangle = E|\psi\rangle$ or (b) *spurious* satisfying $\Sigma_\alpha |\psi_\alpha\rangle = 0$.

Despite a recent analysis of the structure of general H,[5] there are still many unresolved questions. For example, if there exists a physical eigenvector of H for each one of H and if the spurious eigenvectors span $\{\boldsymbol{\psi}: \Sigma_\alpha |\psi_\alpha\rangle = 0\}$, then H is scalar spectral (its eigenvectors and their biorthogonal duals provide a spectral resolution of the identity). However, at present, the only nontrivial cases for which this has been proved are some three-particle Faddeev-like choices H_F.[6,7] Should H contain the appropriate physics, then clearly the existence of a "representation" where the wave function naturally decomposes into arrangement channel components is extremely useful.[4] The first application outside of scattering theory involved molecular bound-state calculations,[8] despite the lack of a proof of existence of these solutions for the Baer–Kouri–Levin–Tobocman H_B used. Dramatic early successes with simple wave function component approximations suggested existence of the solutions considered and also a rigorous basis for atoms/molecules-in-molecules pictures. These results are supported by more recent finite-element method calculations.[9] The significance of a rigorous proof of bound-state existence and uniqueness should be clear from the above discussion and is given here for three-particle, two-cluster, 2×2 and 3×3 H_B using this property for the corresponding H_F.

II. BOUND-STATE EIGENVECTOR EXISTENCE AND UNIQUENESS

For a system of three distinguishable particles labeled $i = 1, 2, 3$, we denote the two cluster channels $i = (i)(jk)$, where $\{i, j, k\} = \{1, 2, 3\}$. Let T be the kinetic energy (with center-of-mass part removed) and $V_i = V_{jk}$ the potential internal to channel i, so $H_i = T + V_i$. We assume the particles act through pairwise potentials, so $H = H_i + V^i$ for all i, where $V^i = V_j + V_k$.

Consider first the 2×2 channel coupling Hamiltonians

$$\mathsf{H}_B = \begin{pmatrix} H_1 & V^2 \\ V^1 & H_2 \end{pmatrix}, \quad \mathsf{H}_F = \begin{pmatrix} H_1 + V_3 & V_1 + V_3 \\ V_2 & H_2 \end{pmatrix} \tag{2}$$

that are related through the identity

$$(\lambda - \mathsf{H}_B) = (\lambda - \mathsf{H}_F)(1 + \mathsf{M}(\lambda)), \tag{3}$$

where the "spurious multiplier" $\mathsf{M}(\lambda) = G_0(\lambda)(\mathsf{H}_F - \mathsf{H}_B)$ and $G_0(\lambda) = (\lambda - T)^{-1}$. An "integral" form of Eq. (3) can be obtained by multiplying from the left by $\mathsf{G}_0(\lambda) = (\lambda - \mathsf{H}_0)^{-1}$, where $(\mathsf{H}_0)_{ij} = \delta_{ij} H_i$. If $|\psi\rangle$ is a bound state of H with eigenvalues $E < 0$, then the corresponding H_F eigenvector is given by[7]

$$\boldsymbol{\psi}_F = \begin{pmatrix} G_0(E)(V_1 + V_3)|\psi\rangle \\ G_0(E)V_2|\psi\rangle \end{pmatrix}. \tag{4}$$

From Eq. (3) and the summation condition (a) for physical eigenvectors, any corresponding H_B eigenvector $\boldsymbol{\psi}_B$ must satisfy

$$(\mathsf{I} + \mathsf{M}(E))\boldsymbol{\psi}_B = \boldsymbol{\psi}_F. \tag{5}$$

If $E > 0$, then it is readily verified that the expressions of Eqs. (4) and (5) still hold with $G_0(E) = P(E - T)^{-1}$, where P represents the Cauchy principal value integral. Equations (4) and (5) motivate the following result:

Theorem 1: Suppose that the eigenvalue E of a bound state $|\psi\rangle$ of H is not in the spectrum of $\hat{H}_3 = T - V_3$. Then if $\hat{G}_3(\lambda) = (\lambda - \hat{H}_3)^{-1}$,

$$\boldsymbol{\psi}_B = \begin{pmatrix} \hat{G}_3(E)(V_1 + V_3)|\psi\rangle \\ |\psi\rangle - \hat{G}_3(E)(V_1 + V_3)|\psi\rangle \end{pmatrix} \tag{6}$$

---

[a] Operated for the U.S. Department of Energy under Contract No. W-7405-ENG-82. This work was supported by the Office of Basic Energy Sciences.

is the corresponding H_B eigenvector. If E is in the spectrum of $\hat{H}_3$ but does *not* correspond to the threshold energy of some $\hat{H}_3$ partially bound state or the complete breakup, then Eq. (6) still holds with $\hat{G}_3(E) = P(E - \hat{H}_3)^{-1}$.

Proof: The simplest proof of Eq. (6) is via direct substitution into the H_B eigenvalue equation. For motivational purposes, we remark that this form of ψ_B can be obtained from Eq. (5) by noting that, formally,

$$(1 + \mathsf{M}(E))^{-1} = \begin{pmatrix} (1 + G_0(E)V_3)^{-1} & 0 \\ 1 - (1 + G_0(E)V_3)^{-1} & 1 \end{pmatrix} = \begin{pmatrix} \hat{G}_3(E)(E - T) & 0 \\ 1 - \hat{G}_3(E)(E - T) & 1 \end{pmatrix} \tag{7}$$

and substituting from Eq. (4) for ψ_F. ■

Starting with a modified choice of H_F, where V_3 now appears on the second row, one similarly obtains

$$\psi_B = \begin{pmatrix} |\psi\rangle - \hat{G}_3(E)(V_2 + V_3)|\psi\rangle \\ \hat{G}_3(E)(V_2 + V_3)|\psi\rangle \end{pmatrix}. \tag{8}$$

The consistency of Eqs. (6) and (8) is readily verified. Suppose that particles 1 and 2 are identical so that $|\psi\rangle$ is either gerade $|\psi^+\rangle$ or ungerade $|\psi^-\rangle$. Let $|\psi_{Bi}^{\pm}\rangle$, $i = 1, 2$ denote the channel components of the corresponding $\psi_B^{\pm}$ and let P_{12} be the operator which interchanges particles 1 and 2; so $|\psi^{\pm}\rangle = \pm P_{12}|\psi^{\pm}\rangle$. Then using the expression for $|\psi_{B1}^{\pm}\rangle$ from Eq. (6) and $|\psi_{B2}^{\pm}\rangle$ from Eq. (8), one recovers the previously observed result[8]

$$|\psi_{B1}^{\pm}\rangle = \pm P_{12}|\psi_{B2}^{\pm}\rangle. \tag{9}$$

We remark that the $\psi_F(\psi_B)$ are strictly only eigenvectors if their components lie in the three-particle Hilbert space [T-$(\hat{H}_{3-})$ boundedness of the V_i is sufficient[6]]. Then the results of Ref. 7 show that since the physical H_F and H_B (weak) eigenvectors include all scattering solutions and their spurious (weak) eigenvectors span $\{\psi: \Sigma_i |\psi_i\rangle = 0\}$, these Hamiltonians are scalar spectral. If the components of any $\psi_F(\psi_B)$ be outside the three-particle Hilbert space, then the corresponding E is in the residual (rather than point) spectrum of $\mathsf{H}_F(\mathsf{H}_B)$.[5]

The Hamiltonian H_B was first used for $H + e$ scattering[10] and since for H_2^+, He, H^- bound-state calculations.[8] A natural assignment of particles 1, 2, and 3 is made so that no (1 2) pair bound states exist. The analysis above applies where all degrees of freedom are retained, as well as to the Born–Oppenheimer (BO) case, where the nuclear kinetic energies are ignored. Except for the BO H_2^+ case, an infinite number of $\hat{H}_3$ partially bound states exist.

We now consider 3×3 choices of H_F and H_B corresponding to the above system where the existence of bound-state solutions is of considerable theoretical and possible practical interest. Here we have

$$\mathsf{H}'_B = \begin{bmatrix} H_1 & 0 & V^3 \\ V^1 & H_2 & 0 \\ 0 & V^2 & H_3 \end{bmatrix}, \quad \mathsf{H}'_F = \begin{bmatrix} H_1 & V_1 & V_1 \\ V_2 & H_2 & V_2 \\ V_3 & V_3 & H_3 \end{bmatrix} \tag{10}$$

that are related through the identity

$$(\lambda - \mathsf{H}'_B) = (\lambda - \mathsf{H}'_F)(1 + \mathsf{M}'(\lambda)), \tag{11}$$

where the "spurious multiplier" $\mathsf{M}'(\lambda) = G_0(\lambda)(\mathsf{H}'_F - \mathsf{H}'_B)$. The familiar integral form of Eq. (11)[11] can be obtained by multiplying from the left by $\mathsf{G}'_0(\lambda) = (\lambda - \mathsf{H}'_0)^{-1}$, where $(\mathsf{H}'_0)_{ij} = \delta_{ij}H_i$. For a bound state $|\psi\rangle$ with $E<0$, the corresponding unique H'_F eigenvector is given by $(\psi'_F)_k = G_0(E)V_k|\psi\rangle$ for $k = 1, 2, 3$.[6] From Eq. (11) and the summation condition (a), any corresponding H_B eigenvector ψ'_B must satisfy

$$(1 + \mathsf{M}'(E))\psi'_B = \psi'_F. \tag{12}$$

If $E>0$, the expression for ψ'_F and Eq. (12) still hold with $G_0(E) = P(E - T)^{-1}$ (cf. above). Equation (12) motivates the following result:

Theorem 2: Let $E \neq 0$ be the eigenvalue of some bound state $|\psi\rangle$ of H. Suppose that V_i satisfy the conditions of Hunziker's theorem[12] and guarantee that $\mathsf{M}'(E)$ is bounded. (We assume T-boundedness for the latter.) Then either there exists a unique eigenvector ψ'_B of H'_B satisfying Eq. (12) or H'_B has a spurious eigenvector with eigenvalue E. In the latter case (nonunique), ψ'_B exist only if certain biorthogonality conditions are (accidentally) met.

Proof: A simple calculation shows that $(\mathsf{M}'(\lambda))^2$ is *connected*.[2] Consequently, since $\mathsf{M}'(\lambda)$ involves only the free Green's function and given the assumptions on V_i, Hunziker's theorem may be applied to prove that $(\mathsf{M}'(\lambda))^2$ is *compact*.[12] Then from Fredholm theory[13] and noting that $(1 - \mathsf{M}'(E))\psi'_F$ is normalizable, it follows that either

$$(1 - (\mathsf{M}'(E))^2)\psi'_B = (1 - \mathsf{M}'(E))\psi'_F \tag{13}$$

has a unique solution ψ'_B, which also satisfies Eq. (12), or

$$(1 - (\mathsf{M}'(E))^2)\psi = \mathbf{0} \tag{14}$$

has a nontrivial solution. In the former case, to show that ψ'_B in Eq. (13) satisfies Eq. (12), one simply notes that

$$(1 - (\mathsf{M}'(E))^2)\{(1 + \mathsf{M}'(E))\psi'_B - \psi'_F\} = \mathbf{0}. \tag{15}$$

Let $\mathsf{P}(\lambda, \mathsf{M})$ denote the M-invariant projection operator onto the eigenvectors of the operator M with eigenvalue λ. Then a simple calculation, e.g., using a Dunford contour integral representation for the P's,[14] shows that

$$\mathsf{P}(\pm 1, \mathsf{M}'(E)) = \tfrac{1}{2}(1 \pm \mathsf{M}'(E))\mathsf{P}(1, (\mathsf{M}'(E))^2), \tag{16}$$

which is bounded, since $\mathsf{M}'(E)$ is bounded. Thus, if ψ satisfies Eq. (14), then either $(1 + \mathsf{M}'(E))\psi = \mathbf{0}$ or $(1 - \mathsf{M}'(E))\psi = \mathbf{0}$. Suppose first that the former is satisfied for *some* solution of Eq. (14). This is just the familiar condition for ψ to be a spurious H'_B eigenvector with eigenvalue E.[11] Second, suppose that the latter is satisfied for *all* solutions of Eq. (14). If ζ' denotes a three-component dual-channel space vector, then from Eq. (16) it is clear that all solutions of $\zeta'(1 - (\mathsf{M}'(E))^2) = \mathbf{0}'$ satisfy $\zeta'(1 - \mathsf{M}'(E)) = \mathbf{0}'$. Thus, from Fredholm theory, Eq. (13) still has (nonunique) solutions ψ'_B. Furthermore, the choice of ψ'_B biorthogonal to all the above ζ' is the unique solution of Eq. (12).

Finally, we note that if spurious eigenvectors exist, then the corresponding solutions of $\zeta'(1 + \mathsf{M}'(E)) = \mathbf{0}'$ must be biorthogonal to ψ'_F for there to exist a solution ψ'_B of Eq. (13). ■

The possibility that physical bound states can be replaced by spurious solutions has been anticipated from general spectral theoretic arguments. These show that the point spectrum of H is contained in the union of the point and residual spectra of H (Ref. 5), from which this replacement strictly follows only if the eigenvalue of a "missing" bound state is physically nondegenerate and not contained in the residual spectrum of H. An explicit example of this replacement phenomenon has been given in Ref. 15. Of course, a similar analysis to that given above follows for the other channel-coupling choice of 3×3 H_B'.[3]

Finally, we remark that if a true three-body potential V_{123} is included diagonally in the Hamiltonians of Eq. (2), then the analysis goes through the minor modifications. For those of Eq. (10), the same is true provided the analog of Hunziker's theorem with $G_0(\lambda)$ replaced by $G_{123}(\lambda) = (\lambda - T - V_{123})^{-1}$ is valid, and the V_i are $T + V_{123}$ bounded. A treatment regarding spurious multipliers as intertwining operators for pairs of channel space Hamiltonians, extending the analysis presented here, is given in Ref. 16.

[1]S. Weinberg, Phys. Rev. **133**, B 232 (1964).

[2]K. L. Kowalski, in *Few Body Systems and Nuclear Forces, Lecture Notes in Physics*, Vol. 87 (Springer, New York, 1978).

[3]D. J. Kouri, H. Kruger, and F. S. Levin, Phys. Rev. D **15**, 1156 (1977).

[4]D. K. Hoffman, D. J. Kouri, and Z. H. Top, J. Chem. Phys. **70**, 4640 (1979).

[5]J. W. Evans, J. Math. Phys. **22**, 1672 (1981).

[6]J. W. Evans and D. K. Hoffman, J. Math. Phys. **22**, 2858 (1981).

[7]J. W. Evans, D. K. Hoffman, and D. J. Kouri, "Scattering theory in arrangement channel quantum mechanics," J. Math. Phys. (in press).

[8]F. S. Levin, "Some Recent Developments in n-Particle Scattering Theory" in *The Few Body Problem*, edited by F. S. Levin (North-Holland, Amsterdam, 1981); see also Nucl. Phys. A **353**, 1436 (1981) and references therein.

[9]W. K. Ford and F. S. Levin, Phys. Lett. B **109**, 155 (1982).

[10]M. Baer and D. J. Kouri, J. Math. Phys. **14**, 1637 (1973).

[11]S. K. Adhikari and W. Glockle, Phys. Rev. C **19**, 616, 2121 (1979).

[12]M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators* (Academic, New York, 1978).

[13]K. Yosida, *Functional Analysis*, 4th ed. (Springer, New York, 1974).

[14]Page 11 of Ref. 12.

[15]Appendix C of Ref. 6.

[16]J. W. Evans, "Spectral Theory for A.C.Q.M. Hamiltonians: Perturbation and Spurious Multiplier Techniques," preprint.

The multipole structure of stationary space-times[a)]

Walter Simon and Robert Beig
Institut für Theoretische Physik, Universität Wien, Vienna, Austria

(Received 11 February 1982; accepted for publication 23 April 1982)

A definition of multipole moments for stationary asymptotically flat solutions of Einstein's equations is proposed. It is shown that these moments characterize a given space-time uniquely. Conversely, they can be arbitrarily prescribed, i.e., they generate power series for the field variables which satisfy the field equations to all orders. Despite their apparently rather different origin, they are shown to be identical with the Geroch–Hansen ones.

PACS numbers: 04.20. – q

1. INTRODUCTION

This paper studies the spatial asymptotics of stationary vacuum space-times in sufficient detail to draw conclusions relevant to their physical interpretation in terms of multipole moments.

Multipole moments at spatial infinity were introduced into general relativity by Geroch[1] and Hansen,[2] using ideas of conformal compactification of 3-space. These works left unproved two conjectures, namely:

Geroch's first conjecture: A given (stationary) space-time is uniquely characterized by its moments.

Geroch's second conjecture: Given a set of moments, there always exists, modulo convergence questions, a space-time corresponding to them.

In the present work we describe an independent approach to multipole moments which works in the physical, rather than a conformally compactified ("unphysical") space, and in the framework of which we are able to settle the above two conjectures affirmatively. Unfortunately, the nongeometric nature of our approach prevents us from successfully tackling convergence questions of the multipole series. In the compactified picture and for the first conjecture (i.e., given a solution which satisfies the boundary conditions) this has recently been accomplished by Beig and Simon in the static[3] and stationary[4,5] cases. The method in Ref. 3 was taken up by Kundu,[6] who also treated the stationary case. Our present approach, although in this respect outdated by these papers, has interest in its own right for the following reasons:

1. It will often be convenient to be able to read off the moments purely from the physical metric.

2. In our setting we can prove the analog of Conjecture 2, which is still open in the conformally compactified version.

3. Our assumptions on asymptotic behavior are substantially weaker than the ones required in a treatment of the conformally transformed (unphysical) field equations. Whereas we use assumptions of the type $\Phi = O(1/r^k)$, $k = 1,2,\ldots$, for large radii on the various quantities, one in most cases needs differentiability conditions at the point at infinity in the other approach. For example, the conformal metric has to be at least C^4 in order for the theorems in Refs. 3–6 to apply. This worry about differentiability at infinity is only at first sight a pedantry. To illustrate this, consider a smooth function $\Phi(x^i)$ in Euclidean 3-space of the form

$$\Phi(x^i) = \frac{\phi_0(\Omega)}{r} + \frac{\phi_1(\Omega)}{r^2} + \cdots + \frac{\phi_{k-1}(\Omega)}{r^k} + O^\infty\left(\frac{1}{r^{k+1}}\right),$$

where (r, Ω) are spherical coordinates centered at some origin and the ϕ's are smooth on S^2. A quantity $\Phi(x^i)$ is said to be $O^\infty(f(r))$ if there is a C^∞ function $f(r)$ such that $|\Phi| \leqslant f(r)$, $|\partial_i \Phi| \leqslant |\partial f/\partial r|$, $|\partial_i\partial_j \Phi| \leqslant |\partial^2 f/\partial r^2|, \ldots$. Consider the Kelvin transform Ψ of the function Φ defined by

$$\Psi(\bar{x}^i) = (1/\bar{r})\Phi(\bar{x}^i/\bar{r}^2), \quad \bar{r}^2 = \bar{x}^i\bar{x}_i.$$

For Ψ to be C^0 at the origin it is necessary and sufficient that ϕ_0 is constant on the 2-sphere. Ψ will also be C^1 iff ϕ_1 is of the form $\phi_1(\Omega) = a_i n^i$, $n^i \in S^2$. More generally, for Ψ to be C^k at the origin there will be restrictions on ϕ_r $(0 \leqslant r \leqslant k)$ which, in particular, imply that ϕ_r contains spherical harmonics only up to order $l \leqslant r$. Therefore, differentiability conditions at spatial infinity, rather than being merely technical assumptions on which the study of the asymptotics of the field equations can be based, are an integral part of such an enterprise, which, as this paper shows, is itself strongly related to the structure of Einstein's equations.

Something must be said about previous work on $(1/r)$ expansion of metrics. O'Murchadha[7] has developed a method for expanding static metrics which is closely related to our approach.

Thorne's review article[8] contains a list of references on multipole expansions. It also shows how complicated expansions of a (nonstationary) 4-metric really are. Our work gives a rigorous treatment of the stationary case and shows the crucial advantages of using Hansen's potentials instead of the metric variables. Nevertheless, there is the same underlying idea in Thorne's concept of "asymptotically Cartesian and mass centered to order N" (ACMC-N)-coordinates and in the transformations which we perform in Sec. 3.

Tanabe[9] gave expansions of Hansen's potentials for some special metrics in harmonic coordinates.

Finally, there is our own work on the static[10] and stationary[11] field equations up to order $1/r^2$, of which the present paper is a systematic extension and improvement.

In Sec. 2 we introduce the basic field variables and state our assumptions on their asymptotic behavior.

In Sec. 3 we determine the structure of the general r^{-k} term of these field variables for a given solution. In order to eliminate spurious degrees of freedom we employ a coordinate condition which is similar to the deDonder gauge. (In the context of equations of motion in general relativity the

[a)] This work was supported by Fonds zur Förderung der wissenschaftlichen Forschung in Österreich, Project No. 4069.

same condition has been used before by Synge.[12]) There arises a set of trace-free symmetric tensors (i.e., tensor fields constant in this gauge) which we call multipole moments and the first k of which determine everything else in the solution up to terms of order r^{-k}. (Theorem 1.) In proving Theorem 1 we develop an algorithm which, in principle, explicitly expresses the metric in terms of the given moments. It seems, at first, that we have at the same time established the analog of Conjecture 2 in our setting. However, this is not *a priori* obvious since the equations which the metric obeys are not really the original field equations, but Einstein's equations truncated with our gauge condition. To fulfill our task we have to show that the metric we have obtained satisfies this gauge condition identically (i.e., for arbitrary moments). This is done in Theorem 2. Whereas the algorithm of Theorem 1 could be used for generating $(1/r)$-expansions for a large class of fields, there are some special features of Einstein's equations and Bianchi's identities that lead to Theorem 2.

In Sec. 5 we make contact with the Geroch–Hansen formulation of the multipole problem. We are able to show that, for arbitrary k, there exists a conformal compactification which renders the unphysical variables C^k. (After have reached C^4 we could, as Refs. 3 and 4 show, go over to an unphysical harmonic chart, thereby making everything even analytic in one stroke). With a C^k-conformal metric we can write down the first k Geroch–Hansen multipole moments. We prove in Theorem 3 that they are identical with ours.

In Sec. 6 we discuss questions left open by the preceding paragraphs. These are: the problem of convergence of the multipole expansion, the uniqueness of coordinates satisfying the gauge condition, the possible use of other potentials for multipole expansions, and the independence of the moments under such a change of potentials.

2. THE STATIONARY FIELD EQUATIONS

We consider a stationary space-time which is represented by a manifold X and coordinated by $\{t, x^i\}$. X is topologically $I\times N$ where I is the t axis and N is diffeomorphic to $\mathbb{R}^3$ minus a ball. On X we are given a metric

$$ds^2 = \lambda\,(dt + \sigma_i\, dx^i)^2 - \lambda^{-1}\gamma_{ij}\, dx^i\, dx^j, \tag{2.1}$$

which satisfies Einstein's vacuum field equations. λ, σ_i, and γ_{ij} (the metric on N) are functions of x^i. Tensor indices will be moved with γ_{ij} and its inverse γ^{ij}. This does not apply to indices on coordinates and on constants, i.e., we will write $x^i = x_i$ and $C_{ij\cdots} = C^{ij\cdots}$. The covariant derivative with respect to γ_{ij} will be denoted by D, the covariant Laplace operator by $\Delta(\gamma) = D_iD^i$. We use the following symbols for the connection, the Riemann tensor, and their contractions (α_k is an arbitrary vector field on N):

$$D_i\alpha_k = \partial_i\alpha_k - \Gamma^j{}_{ik}\alpha_j, \quad \Gamma^i := \gamma^{jk}\Gamma^i{}_{jk}, \tag{2.2}$$

$$D_{[i}D_{j]}\alpha_k = \tfrac{1}{2}\mathscr{R}_{ijkl}{}^{h}\alpha_h, \quad \mathscr{R}_{ij} = \mathscr{R}^k{}_{ikj}, \quad \mathscr{R} = \mathscr{R}^i{}_i. \tag{2.3}$$

Consider the vector field

$$\omega_i = -\lambda^2\epsilon_{ijk}D^j\sigma^k, \tag{2.4}$$

where $\epsilon_{ijk} = \epsilon_{[ijk]}$ is the permutation symbol ($\epsilon_{123} = |\det\gamma_{ij}|^{1/2}$). For static metrics, ω_i vanishes identically. The vacuum field equations imply (see e.g., Ref. 13)

$$D_{[i}\omega_{j]} = 0. \tag{2.5}$$

Since N is simply connected, there exists a scalar field ω (twist of the timelike Killing vector) such that $D_i\omega = \omega_i$. We shall write down the field equations in terms of a set of variables introduced by Hansen[2]

$$\Phi_M = \tfrac{1}{4}\lambda^{-1}(\lambda^2 + \omega^2 - 1), \quad \Phi_S = \tfrac{1}{2}\lambda^{-1}\omega,$$

$$\Phi_K = \tfrac{1}{4}\lambda^{-1}(\lambda^2 + \omega^2 + 1), \tag{2.6}$$

$$\tau_{ij} = 2[D_i\Phi_M D_j\Phi_M + D_i\Phi_S D_j\Phi_S - D_i\Phi_K D_j\Phi_K],$$

$$\tau := \tau^i{}_i. \tag{2.7}$$

One obtains

$$\Delta(\gamma)\Phi = 2\tau\Phi \tag{2.8}$$

for each of the Φ's and

$$\mathscr{R}_{ij}(\gamma) = \tau_{ij}. \tag{2.9}$$

For static vacuum solutions in the presence of an electric field there is a similar set of variables and equations (see Hoenselaers[14]). In what follows we consider the structure of stationary fields. The treatment of the electrostatic case is similar.

Our asymptotic conditions are as follows:

Definition: A solution of (2.8) and (2.9) is called a SV solution (stationary, asymptotically flat vacuum solution) iff there exists a coordinate system on N such that

$$\Phi_M = O^\infty(r^{-1}), \quad \Phi_S = O^\infty(r^{-1}), \tag{2.10}$$

$$\gamma_{ij} = \delta_{ij} + O^\infty(r^{-1}). \tag{2.11}$$

As in our previous papers,[10,11] we could equally well require only finite degrees of differentiability and Hölder continuity for the metric variables. Since we do not want to obscure our results by having to count degrees of differentiability throughout the paper we start with O^∞ fields.

In the following paragraphs we shall frequently employ a change of the coordinate system x^i. The corresponding maps will not necessarily be diffeomorphisms or even defined in all of N. It will be clear, however, that at least a neighborhood of infinity is mapped diffeomorphically into itself and this, for the purpose of asymptotics, is all that is needed (compare also Ref. 10).

3. THE MULTIPOLE EXPANSION

Theorem 1: For all SV solutions on N there is a coordinate system x^i and there are sets of constants $A\cdots$, $B\cdots$, $\cdots$ $G\cdots$, such that for all nonnegative integers m

$$\Phi_M = \sum_{l=0}^{m-1}\frac{E_{a_1\cdots a_l}x^{a_1}\cdots x^{a_l}}{l!\,r^{2l+1}} + O^\infty(r^{-(m+1)}) := \Phi_M^{(m)} + O^\infty(r^{-(m+1)}), \tag{3.1}$$

$$\Phi_S = \sum_{l=0}^{m-1}\frac{F_{a_1\cdots a_l}x^{a_1}\cdots x^{a_l}}{l!\,r^{2l+1}} + O^\infty(r^{-(m+1)}) := \Phi_S^{(m)} + O^\infty(r^{-(m+1)}), \tag{3.2}$$

$$\Phi_K = \frac{1}{2} + \sum_{l=1}^{m-1}\frac{G_{a_1\cdots a_{l-1}}x^{a_1}\cdots x^{a_{l-1}}}{l!\,r^{2l}} + O^\infty(r^{-(m+1)}), \tag{3.3}$$

$$\gamma_{ij} = \delta_{ij} + \sum_{l=2}^{m}\left(\frac{x^i x^j A_{a_1\cdots a_{l-2}} x^{a_1}\cdots x^{a_{l-2}}}{r^{2l}} + \frac{\delta_{ij} B_{a_1\cdots a_{l-2}} x^{a_1}\cdots x^{a_{l-2}}}{r^{2l-2}} + \frac{x_{(i} C_{j)a_1\cdots a_{l-3}} x^{a_1}\cdots x^{a_{l-3}}}{r^{2l-2}} + \frac{D_{ija_1\cdots a_{l-4}} x^{a_1}\cdots x^{a_{l-4}}}{r^{2l-4}}\right) + O^{\infty}(r^{-(m+1)}):$$

$$= \delta_{ij} + h_{ij}^{(m)} + O^{\infty}(r^{-(m+1)}). \tag{3.4}$$

The constants $C\cdots$ appear only for $m \geqslant 3$, the constants $D\cdots$ for $m \geqslant 4$. All constants are symmetric in all their a_j indices, and $D\cdots$ is also symmetric in i and j. $A_{a_1\cdots a_{m-2}}$, $B_{a_1\cdots a_{m-2}}$, $C_{ja_1\cdots a_{m-3}}$, $D_{ija_1\cdots a_{m-4}}$, $G_{a_1\cdots a_{m-2}}$, and the trace parts of $E_{a_1\cdots a_{m-1}}$ and of $F_{a_1\cdots a_{m-1}}$ depend on the set

$$\mathscr{P}_m := \{M_{a_1\cdots a_{l-1}} := \mathscr{C}[E_{a_1\cdots a_{l-1}}]$$

and

$$S_{a_1\cdots a_{l-1}} := \mathscr{C}[F_{a_1\cdots a_{l-1}}], \quad 1 \leqslant l \leqslant m\}, \tag{3.5}$$

where $\mathscr{C}$ denotes the trace-free part with respect to δ_{ij} (for an explicit form see Ref. 15) by algebraic relations (linear combinations of products $M_{a_1\cdots a_l}$ and $S_{a_1\cdots a_l}$). These algebraic relations are the same for all SV solutions.

Proof: The proof will be carried out by induction with respect to m. For $m = 0$, the theorem is true by the SV assumption. The cases $m = 1$ and $m = 2$ are treated in Refs. 10 and 11. We introduce the following functions:

$$\gamma_{ij} = \delta_{ij} + h_{ij}, \quad \gamma^{ij} = \delta_{ij} - k^{ij},$$
$$(k^{ij} = h_{ik}\gamma^{kj}), \quad h_{kk} := \delta_{ik} h_{ik}, \tag{3.6}$$

$$\Lambda_i := (h_{ij} - \tfrac{1}{2}\delta_{ij} h_{kk})_{,j}. \tag{3.7}$$

Using the Laplace operators of flat space on the left-hand sides, the field equations (2.8) and (2.9) take on the following form:

$$\Delta\Phi = k^{ij}\Phi_{,ij} + \Gamma^i\Phi_i + 2\tau\Phi, \tag{3.8}$$

$$\Delta h_{ij} = 2\Lambda_{(i,j)} + k^{lk}(h_{ij,kl} + h_{kl,ij} - h_{lj,ik} - h_{ik,lj}) + 2\gamma^{ml}\gamma^{hk}(\Gamma_{mij}\Gamma_{lhk} - \Gamma_{mik}\Gamma_{lhj}) - 2\tau_{ij}. \tag{3.9}$$

In order to perform the induction process, we must insert the series (3.1)–(3.4) into the right-hand sides of (3.8) and (3.9) and invert the Laplacians with the help of Lemmas A and B of the appendix. The essential condition for our success is that the right-hand sides are known up to and including order $r^{-(m+3)}$ if the Φ and h_{ij} series up to order r^{-m} are inserted, since Δ increases the order by two. The term $2\Lambda_{(i,j)}$ in (3.6) apparently does not meet this requirement:

$$\Lambda_{(i,j)} = (\text{known terms of order 3 up to } m+2) + O^{\infty}(r^{-(m+3)}). \tag{3.10}$$

However, the bad $O^{\infty}(r^{-(m+3)})$ remainder does not appear if the coordinate system is adapted suitably before each step of induction. We will show that the required coordinate transformations exist, that they do not destroy the forms of Φ and h_{ij} up to the known order m, and that the Laplace operators in (3.8) and (3.9) can then be inverted. In doing so, we obtain expressions of Φ and h_{ij} which hold up to order $r^{-(m+1)}$. Since we are interested in the general solution, we must admit a contribution from the homogeneous solution in Φ_M, Φ_S, and h_{ij} of order $r^{-(m+1)}$ (Φ_K is determined by Φ_M and Φ_S. In Φ_M and Φ_S these contributions are generated by the 2^m multipole moments $M_{a_1\cdots a_m}$ and $S_{a_1\cdots a_m}$; in h_{ij} they will be shown to be "pure gauge." In order not to make the computations hopelessly complicated, we make use of a special technique (s-algebra) which will be used frequently throughout this paper to manipulate expressions of the form (3.11). The analysis could also be carried out in terms of spherical harmonics, but we choose to work in the Cartesian approach.

The s algebra

Let K... be a constant and δ... the Kronecker symbol. The expression

$$\alpha_{i_1\cdots i_a k_1\cdots k_d}: = \frac{K_{i_1\cdots i_a j_1\cdots j_b}\,\delta_{k_1k_2}\cdots\delta_{k_{c-1}k_c}\, x^{k_{c+1}}\cdots x^{k_d} x^{j_1}\cdots x^{j_b}}{r^{e-c}} \tag{3.11}$$

is called a *term* of order r^{-m} (order m), $m = e - d - b$. It has $f = a + d$ "free indices" and a "characteristic number" s, an integer which satisfies

the congruence $s \equiv (m - (a+b))$ modulo 2, (3.12)

the inequality $s \leqslant m - (a+b)$, (3.13)

and is otherwise chosen arbitrarily.

In characterizing a term we always try to choose the largest possible s, which reflects our information about the "δ content" of the term α. The dependence of s on the order m has proved to be useful. Using these definitions, we will perform addition, multiplication, and differentiation of terms of the form (3.11) entirely in terms of the corresponding simple manipulations of their m, f, and s values. The behavior of m and f under these operations is obvious. The (best) s value for the sum of two terms with s_1 and s_2, $s_1 \equiv s_2$ (modulo 2) is $\min(s_1, s_2)$. We will never add terms with $s_1 \not\equiv s_2$ (mod 2)!

We can assign $s + p$ to the pth derivative of a term with respect to x^i.

Multiplication (with arbitrary contraction of free indices) of terms with s_1 and s_2 produces an $s_1 + s_2$ term. If contraction over indices on constants $\neq \delta_{ij}$ occurs, an even number of indices disappears and hence s can be chosen greater than $s_1 + s_2$.

We consider some examples which are needed in what follows.

Φ_M and Φ_S, as they appear in (3.1,2), consist of terms with $f = 0$ and $s = 1$; h_{ij} contains only $f = 2$, $s = 2$ terms. Moreover, it is easily seen that the series for Φ_M, Φ_S, and h_{ij} display the *general form* of terms with these f and s values, i.e., *every $f = 0$, $s = 1$ term* is of the form (Z... constant)

$$\frac{Z_{a_1\cdots a_k} x^{a_1}\cdots x^{a_k}}{r^{2k+1}}$$

and *every $f = 2$, $s = 2$ term* is of the $A\cdots$, $B\cdots$, $C\cdots$, or $D\cdots$ type (arbitrary constants) which appears in (3.4).

$k^{ij} = \delta_{ij} - \gamma^{ij}$ must be of the same form as h_{ij}. This can be seen as follows. Assume that k^{ij} has $s = 2$ terms up to order m, and compute order $m+1$ with the recursion formula

$$k^{ij} = h_{ij} - h_{il} k^{lj}. \tag{3.14}$$

Because h_{ik} is $O(r^{-2})$, terms from k^{ij} of at most $O(r^{-m+1})$ contribute to the second term on the right-hand side. The s value of the right-hand side is $\min(2, 2+2) = 2$. Therefore k^{ij} can also be written in the form (3.4) with constants $A'\ldots, \ldots D'\ldots$.

We begin the induction proof. Our assumptions read

1. Up to order m, Φ_M, Φ_S, *and* h_{ij}, assume the forms (3.1), (3.2), and (3.4).

2. The coordinates can be chosen such that $\Lambda_i = O^\infty(r^{-(m+2)})$.

Our task is to show 1) and 2) for $m+1$. Consider the equation

$$\Delta f_i = \Lambda_i = O^\infty(r^{-(m+2)}). \tag{3.15}$$

From the proof of Lemma 2 of Ref. 10 for $m=0$ and our Lemma A (Appendix) for $m>0$ there exist solutions

$$\begin{aligned} f_i &= O^\infty(\ln^2 r) \qquad \text{for } m=0,\\ f_i &= O^\infty(r^{-m}\ln r) \quad \text{for } m>0. \end{aligned} \tag{3.16}$$

(Note that for $m=0$ we can add to f_i an arbitrary constant a^i, thereby fixing the origin to which the multipole moments will be related.) The logarithmic form of f_i will in general lead to corresponding logarithmic terms in the O^∞ remainders of the field variables. These, however, will all disappear finally. So, in what follows, we simply write $\ln^* r$ for some finite power of $\ln r$ in order to allow a unified treatment for all $m \geqslant 0$. Under an arbitrary coordinate transformation

$$\bar{x}^i = p^i(x^j), \quad \frac{\partial \bar{x}^i}{\partial x^j} = p^i{}_j(x), \quad \frac{\partial^2 \bar{x}^i}{\partial x^j \partial x^k} = p^i{}_{jk}(x), \tag{3.17}$$

we have

$$\begin{aligned} \Lambda_i = {}& \bar{\gamma}_{jk,\bar{l}} p^j{}_i p^k{}_m p^l{}_m \\ & - \tfrac{1}{2}\bar{\gamma}_{jk,\bar{l}} p^j{}_i p^k{}_m p^l{}_m + \bar{\gamma}_{ik} p^j{}_i p^k{}_{mm}. \end{aligned} \tag{3.18}$$

Hence, Λ_i behaves as follows under $\bar{x}^i = x^i + f^i(x^j)$ with f_i as above:

$$\begin{aligned} \bar{\Lambda}_i(\bar{x}) &= \Lambda_i(x) - \Delta f_i(x) + O^\infty(\ln^* r / r^{m+3}) \\ &= O^\infty(\ln^* \bar{r} / \bar{r}^{m+3}). \end{aligned} \tag{3.19}$$

In the new coordinates Assumption 1 reads (omitting bars)

$$\Phi_M(x) = \Phi_M^{(m)}(x) + O^\infty(r^{-(m+1)}), \tag{3.20}$$

$$\Phi_S(x) = \Phi_S^{(m)}(x) + O^\infty(r^{-(m+1)}), \tag{3.21}$$

$$h_{ij}(x) = h_{ij}^{(m)}(x) + O^\infty((\ln^* r) r^{-(m+1)}). \tag{3.22}$$

The quantities on the right-hand sides of (3.8) and (3.9) take on the following forms:

$$\begin{aligned} \Gamma_{ijk} = {}& [\text{known terms with } s=3 \text{ of order } r^{-3} \text{ up to } r^{-(m+1)}] \\ & + O^\infty((\ln^* r) r^{-(m+2)}) \end{aligned} \tag{3.23}$$

$$\Lambda_i = O^\infty((\ln^* r) r^{-(m+3)}) \tag{3.24}$$

$$\Gamma_i = [s=5; r^{-5} \ldots r^{-(m+2)}] + O^\infty((\ln^* r) r^{-(m+3)}). \tag{3.25}$$

(The definition of s was explained above; the brackets do not appear for low values of m.)

Expanding $\Phi_K := \frac{1}{2}(1 + 4\Phi_M^2 + 4\Phi_M^2)^{1/2}$ one obtains (3.3). According to the definition (2.7) of τ_{ij}, we have

$$\tau_{ij} = [s=4; r^{-4} \ldots r^{-(m+3)}] + O^\infty(r^{-(m+4)}). \tag{3.26}$$

We insert in (3.8) and (3.9) and use a little "s algebra." We obtain

$$\Delta \Phi_M = [s=5; r^{-4} \ldots r^{-(m+3)}] + O^\infty((\ln^* r) r^{-(m+4)}), \tag{3.27}$$

$$\Delta \Phi_S = [s=5; r^{-4} \ldots r^{-(m+3)}] + O^\infty((\ln^* r) r^{-(m+4)}), \tag{3.28}$$

$$\Delta h_{ij} = [s=4; r^{-4} \ldots r^{-(m+3)}] + O^\infty((\ln^* r) r^{-(m+4)}). \tag{3.29}$$

In the brackets in (3.27)–(3.29) we now use the induction hypothesis for terms up to order $r^{-(m+2)}$ and s algebra for order $r^{-(m+3)}$. This yields

$$\begin{aligned} \Delta \Phi_M = {}& \Delta \Phi_M^{(m)} + \left[\frac{H_{a_1 \cdots a_{m+2}} x^{a_1} \cdots x^{a_{m+2}}}{r^{2m+1}} \right] \\ & + O^\infty((\ln^* r) r^{-(m+4)}), \end{aligned} \tag{3.30}$$

$$\begin{aligned} \Delta \Phi_S = {}& \Delta \Phi_S^{(m)} + \left[\frac{I_{a_1 \cdots a_{m+2}} x^{a_1} \cdots x^{a_{m+2}}}{r^{2m+1}} \right] \\ & + O^\infty((\ln^* r) r^{-(m+4)}), \end{aligned} \tag{3.31}$$

$$\begin{aligned} \Delta h_{ij} = {}& \Delta h_{ij}^{(m)} + \left[\frac{x^i x^j J_{a_1 \cdots a_{m+1}} x^{a_1} \cdots x^{a_{m+1}}}{r^{2m+4}} \right. \\ & + \frac{\delta_{ij} K_{a_1 \cdots a_{m+1}} x^{a_1} \cdots x^{a_{m+1}}}{r^{2m+2}} + \frac{x_{(i} L_{j)a_1 \cdots a_{m+2}} x^{a_1} \cdots x^{a_{m+2}}}{r^{2m+2}} \\ & \left. + \frac{N_{ija_1 \cdots a_{m+1}} x^{a_1} \cdots x^{a_{m+1}}}{r^{2m}} \right] + O^\infty\left(\frac{\ln^* r}{r^{m+4}} \right). \end{aligned} \tag{3.32}$$

The constants appearing in (3.30)–(3.32) are certain linear combinations of products of the constants in $\Phi_M^{(m)}$, $\Phi_S^{(m)}$, and $h_{ij}^{(m)}$, which, in turn, depend only on $M, M_i, \ldots, M_{a_1 \cdots a_{m+1}}$ and $S, S_i, \ldots, S_{a_1 \cdots a_{m+1}}$ by the induction hypothesis. Again, the equations simplify for small m.

In order to determine Δ^{-1} of the expressions in brackets, we use Lemma B; for the O^∞ remainders, Lemma A.

Finally, by comparison with the induction hypothesis, we obtain: There are constants

$$P_{a_1 \cdots a_{m+2}}, \quad Q_{a_1 \cdots a_{m+2}}, \quad A_{a_1 \cdots a_{m+1}}, \\ B_{a_1 \cdots a_{m+1}}, \quad C_{ia_1 \cdots a_{m+2}}, \quad D_{ija_1 \cdots a_{m+1}},$$

which are symmetric in the a_k as well as in i and j and there are constants

$$M_{a_1 \cdots a_m}, \quad S_{a_1 \cdots a_m}, \quad T_{ija_1 \cdots a_m},$$

which are, in addition to the properties above, trace-free in a_k, such that

$$\begin{aligned} \Phi_M = {}& \Phi_M^{(m)} + \frac{(P_{(a_1 \cdots a_{m+2}} \delta_{a_m a_{m+1})} + M_{a_1 \cdots a_m}) x^{a_1} \cdots x^{a_m}}{r^{2m+1}} \\ & + O^\infty((\ln^* r) r^{-(m+2)}), \end{aligned} \tag{3.33}$$

$$\begin{aligned} \Phi_S = {}& \Phi_S^{(m)} + \frac{(Q_{(a_1 \cdots a_{m+2}} \delta_{a_m a_{m+1})} + S_{a_1 \cdots a_m}) x^{a_1} \cdots x^{a_m}}{r^{2m+1}} \\ & + O^\infty((\ln^* r) r^{-(m+2)}), \end{aligned} \tag{3.34}$$

$$h_{ij} = h_{ij}^{(m)} + \frac{x^i x^j A_{a_1 \cdots a_{m-1}} x^{a_1} \cdots x^{a_{m-1}}}{r^{2m+2}}$$
$$+ \text{terms with } B\cdots, C\cdots, D\cdots,$$
$$+ \frac{T_{ija_1\cdots a_m} x^{a_1}\cdots x^{a_m}}{r^{2m+1}} + O^{\infty}((\ln^* r) r^{-(m+2)}). \quad (3.35)$$

Now we define

$$E_{a_1\cdots a_m} := P_{(a_1\cdots a_{m-2}}\delta_{a_{m-1}a_m)} + M_{a_1\cdots a_m},$$
$$F_{a_1\cdots a_m} := Q_{(a_1\cdots a_{m-2}}\delta_{a_{m-1}a_m)} + S_{a_1\cdots a_m}, \quad (3.36)$$
$$t_{ij}^{(m+1)} := \frac{T_{ija_1\cdots a_m} x^{a_1}\cdots x^{a_m}}{r^{2m+1}} \quad (3.37)$$

and write

$$\Phi_M = \Phi_M^{(m+1)} + O^{\infty}(\ln^* r) r^{-(m+2)}), \quad (3.38)$$
$$\Phi_S = \Phi_S^{(m+1)} + O^{\infty}((\ln^* r) r^{-(m+2)}), \quad (3.39)$$
$$h_{ij} = h_{ij}^{(m+1)} + t_{ij}^{(m+1)} + O^{\infty}((\ln^* r) r^{-(m+2)}). \quad (3.40)$$

Because of the gauge condition (3.24),

$$(h_{ij}^{(m+1)} - \tfrac{1}{2}\delta_{ij} h_{kk}^{(m+1)})_{,j} = -(t_{ij}^{(m+1)} - \tfrac{1}{2}\delta_{ij} t_{kk}^{(m+1)})_{,j} \quad (3.41)$$

must hold. Compare the $O^{\infty}(r^{-(m+2)})$ expressions in this equation! The terms on the left-hand side have $s_L = 3$, the terms on the right-hand side, however, $s_R = 0$. Since $s_L - s_R$ is odd, or, since both sides have different parity under space reflection $x^i \to -x^i$, both sides must vanish identically. Therefore, $t_{ij}^{(m+1)}$ satisfies the requirements of Lemma C; accordingly there is a function g_i ($g_i = O^{\infty}(r^{-m})$, $\Delta g_i \equiv 0$) such that

$$t_{ij} = g_{i,j} + g_{j,i}. \quad (3.42)$$

We now introduce new coordinates $\bar{x}^i = x^i + g^i(x^j)$, thus disposing of t_{ij} without affecting anything else in (3.38), (3.39), (3.40), and (3.24). In order to get rid of the logarithmic terms, too, we carry out all steps of the present proof once more, starting with the new Φ_M, Φ_S, h_{ij}, and A_i. Nothing happens, except that we must put $m+1$ instead of m and higher powers of $\ln r$ in the O^{∞} remainders everywhere. We obtain, in a new coordinate system, which we call x^i again,

$$\Phi_M = \Phi_M^{(m+2)} + O^{\infty}((\ln^* r) r^{-(m+3)})$$
$$= \Phi_M^{(m+1)} + O^{\infty}(r^{-(m+2)}), \quad (3.43)$$
$$\Phi_S = \Phi_S^{(m+2)} + O^{\infty}((\ln^* r) r^{-(m+3)})$$
$$= \Phi_S^{(m+1)} + O^{\infty}(r^{-(m+2)}), \quad (3.44)$$
$$h_{ij} = h_{ij}^{(m+2)} + O^{\infty}((\ln^* r) r^{-(m+3)})$$
$$= h_{ij}^{(m+1)} + O^{\infty}(r^{-(m+2)}), \quad (3.45)$$
$$A_i = O^{\infty}((\ln^* r) r^{-(m+4)}) = O^{\infty}(r^{-(m+3)}). \quad (3.46)$$

It is clear that Φ_K behaves as it is quoted in the theorem, too. This ends the proof. □

Choosing $m = 3$ in the preceding theorem, we obtain

$$\Phi_M = \frac{M}{r} + \frac{M_i x^i}{r^3} + \frac{1}{2}\frac{M_{ij}x^i x^j}{r^5} + \frac{M(M^2+S^2)}{r^3} + O^{\infty}\left(\frac{1}{r^4}\right), \quad (3.47)$$

$$\Phi_S = \frac{S}{r} + \frac{S_i x^i}{r^3} + \frac{1}{2}\frac{S_{ij}x^i x^j}{r^5} + \frac{S(M^2+S^2)}{r^3} + O^{\infty}\left(\frac{1}{r^4}\right), \quad (3.48)$$

$$\Phi_K = \frac{1}{2} + \frac{M^2+S^2}{r^2} + \frac{2MM_i x^i}{r^4} + \frac{2SS_i x^i}{r^4} + O^{\infty}\left(\frac{1}{r^4}\right), \quad (3.49)$$

$$\gamma_{ij} = \delta_{ij} - \frac{M^2}{r^4}(\delta_{ij} r^2 - x^i x^j) - \frac{2MM_{(i}x_{j)}}{r^4} - \frac{2MM_k x^k \delta_{ij}}{r^4} + \frac{4MM_k x^k x^i x^j}{r^6}$$
$$- (\text{corresponding terms with } S, S_i) + O^{\infty}\left(\frac{1}{r^4}\right). \quad (3.50)$$

Of course, one would like to proceed with this explicit presentation of the expansions to arbitrary high orders. The theorem does not display the constants in detail since we have intentionally and successfully avoided cumbersome computations. Following this principle, we shall only give a simple example: the Taub-NUT metric[16] given by

$$ds^2 = \frac{r^2 - 2mr - l^2}{r^2 + l^2}(dt + 2l\cos\theta\, d\phi)^2$$
$$- \frac{r^2 + l^2}{r^2 - 2mr - l^2} dr^2 - (r^2 + l^2)d\theta^2 + \sin^2\theta\, d\phi^2). \quad (3.51)$$

We restrict ourselves to the stationary domain N, $r > m + (m^2 + l^2)^{1/2}$ and introduce the new coordinate $\rho = r - m$ (t and the angles remain unchanged) and the abbreviation $c^2 := m^2 + l^2$. Then Hansen's potentials are

$$\Phi_M = -\frac{m}{\rho(1 - c^2/\rho^2)} = -\frac{m}{\rho}\sum_{p=0}^{\infty}\left(\frac{c}{\rho}\right)^{2p}, \quad (3.52)$$
$$\Phi_S = -\frac{l}{\rho(1 - c^2/\rho^2)} = -\frac{l}{\rho}\sum_{p=0}^{\infty}\left(\frac{c}{\rho}\right)^{2p}. \quad (3.53)$$

Using Cartesian coordinates ($x = \rho\sin\theta\cos\phi$, $y = \rho\sin\theta\sin\phi$, $z = \rho\cos\theta$) the spatial part of the metric simply reads

$$\gamma_{ij} = \delta_{ij} - (c^2/\rho^4)(\delta_{ij}\rho^2 - x^i x^j). \quad (3.54)$$

By comparison with (3.50) it is obvious that $A_i \equiv 0$, i.e., we have found "good" coordinates at one stroke in this case. The parameters m and l can readily be identified with $-M$ and $-S$, respectively. Note, that (3.52) and (3.53) are convergent precisely on the stationary domain $\rho > c$. We remark that a similar treatment of the Reissner–Nordström metric in the electrostatic case is possible.

4. ON GENERATING SOLUTIONS WITH ARBITRARY MOMENTS

Theorem 2: To every given set $\mathscr{P}_m$ (3.5) of symmetric, trace-free matrices there are corresponding fields $\Phi_M^{(m)}$, $\Phi_S^{(m)}$, and $\gamma_{ij}^{(m)}$ satisfying the asymptotic conditions (2.10) and (2.11) and obeying the field equations up to terms of higher order in the sense that

$$\Delta(\gamma^{(m)})\Phi_M^{(m)} = 2\tau^{(m)}\Phi_M^{(m)} + O^\infty(r^{-(m+3)}), \tag{4.1}$$

$$\Delta(\gamma^{(m)})\Phi_S^{(m)} = 2\tau^{(m)}\Phi_S^{(m)} + O^\infty(r^{-(m+3)}), \tag{4.2}$$

$$\mathscr{R}_{ij}^{(m)} = \tau_{ij}^{(m)} + O^\infty(r^{-(m+3)}) \tag{4.3}$$

($\mathscr{R}_{ij}^{(m)}$ and $\tau_{ij}^{(m)}$ are formed from $\gamma_{ij}^{(m)} = \delta_{ij} + h_{ij}^{(m)}$, $\Phi_M^{(m)}$ and $\Phi_S^{(m)}$.)

Proof: By the procedure described in the preceding paragraph, one can, order by order, find functions Φ_M, Φ_S, and γ_{ij} which satisfy (4.1) and (4.2) but, instead of (4.3), its "truncated" form

$$\mathscr{R}_{ij}^{(m)} = \tau_{ij}^{(m)} + \Lambda_{(i,j)}^{(m)} + O^\infty(r^{-(m+3)}). \tag{4.4}$$

For this purpose, one inserts, in the mth step, (3.1), (3.2), and (3.4) into the right-hand sides of (3.8) and (3.9) with "$2\Lambda_{(i,j)}$" omitted in the latter. Then, one inverts the Laplacians with Lemma B, adds

$$\frac{M_{a_1\cdots a_m}x^{a_1}\cdots x^{a_m}}{r^{2m+1}}$$

to the mass potential and

$$\frac{S_{a_1\cdots a_m}x^{a_1}\cdots x^{a_m}}{r^{2m+1}}$$

to the angular-momentum potential.

We cannot solve the full equation (4.3) by this method because of (3.10) and the remarks made there. But now we cannot just impose the gauge condition

$$\Lambda_i^{(m)} := (h_{ij}^{(m)} - \tfrac{1}{2}\delta_{ij}h_{kk}^{(m)})_{,j} = 0 \tag{4.5}$$

since the generation process described above is already well defined without it.

Fortunately, $\Lambda_i^{(m)}$ which results from this process vanishes automatically for every m. In order to show this, observe that $\tau_{ij}^{(m)}$ satisfies

$$D^{j(m)}\tau_{ij}^{(m)} = \tfrac{1}{2}D_i^{(m)}\tau^{(m)} + O^\infty(r^{-(m+4)}) \tag{4.6}$$

by virtue of (4.1) and (4.2) and $\Phi_M^2 + \Phi_S^2 - \Phi_K^2 = -\frac{1}{4}$.

On the other hand, the contracted Bianchi identity must hold "up to higher order" for $\mathscr{R}_{ij}^{(m)}$:

$$D^{j(m)}\mathscr{R}_{ij}^{(m)} = \tfrac{1}{2}D_i^{(m)}\mathscr{R}^{(m)} + O^\infty(r^{-(m+4)}). \tag{4.7}$$

All this gives

$$D^{j(m)}\Lambda_{(i,j)}^{(m)} = \tfrac{1}{2}D_i^{(m)}(\gamma^{kl(m)}\Lambda_{k,l}^{(m)}) + O^\infty(r^{-(m+4)}) \tag{4.8}$$

which, in turn, implies (Δ is the flat Laplacian)

$$\Delta\Lambda_i^{(m)} = k^{jl(m)}\Lambda_{i,jl}^{(m)} + 2\Gamma^{k(m)}\Lambda_{(i,k)}^{(m)} + O^\infty(r^{-(m+4)}). \tag{4.9}$$

We know that $k^{jl(m)} = O(r^{-1})$, $\Gamma^{k(m)} = O(r^{-2})$, and $\Lambda_i^{(1)} \equiv 0$. Assume, as induction hypothesis, that $\Lambda_i^{(p)} \equiv 0$ for some p. Then, using the definition of $\Lambda_i^{(p+1)}$ in terms of $\gamma_{ij}^{(p+1)}$ (and "s algebra") we have,

$$\Lambda_i^{(p+1)} = \frac{U_{ia_1\cdots a_{p-2}}x^{a_1}\cdots x^{a_{p-2}}}{r^{2p}} + \frac{V_{a_1\cdots a_{p-1}}x^i x^{a_1}\cdots x^{a_{p-1}}}{r^{2p+2}} \tag{4.10}$$

for some constants $U\cdots$ and $V\cdots$.

On the other hand, (4.9) and $\Lambda_i^{(p+1)} = O^\infty(r^{-(p+2)})$ imply $\Delta\Lambda_i^{(p+1)} = O^\infty(r^{-(p+5)})$. Using Lemma A, we obtain

$$\Lambda_i^{(p+1)} = \frac{W_{ia_1\cdots a_{p+1}}x^{a_1}\cdots x^{a_{p+1}}}{r^{2p+3}} + O^\infty((\ln^* r)r^{-(p+3)}) \tag{4.11}$$

for some constants $W\cdots$. Since the terms in (4.10) and (4.11) have different parity under space reflection $x^i \to -x^i$, $U\cdots$, $V\cdots$, and $W\cdots$ are 0. Hence $\Lambda_i^{(m)} \equiv 0$ for every m. □

5. THE CONNECTION WITH THE GEROCH-HANSEN MOMENTS

In slight extension of the original definition due to Geroch[1] we adopt the following

Definition:

1. A space–time $\widetilde{N}$ is called C^k-compactification of N, if $\widetilde{N}$ consists of N plus a point Λ and if there is a C^∞-function Ω and a C^k-metric $\tilde{\gamma}_{ij}$ on $\widetilde{N}$ such that

$$\tilde{\gamma}_{ij} = \Omega^2\gamma_{ij}, \tag{5.1}$$

$$\Omega|_\Lambda = \widetilde{D}_i\Omega|_\Lambda = (\widetilde{D}_i\widetilde{D}_j\Omega - 2\tilde{\gamma}_{ij})|_\Lambda = 0. \tag{5.2}$$

2. Assume in addition that $\widetilde{\Phi}_M := \Omega^{-1/2}\Phi_M$ and $\widetilde{\Phi}_S := \Omega^{-1/2}\Phi_S$, both extend to C^k-functions on $\widetilde{N}$.

The following set of symmetric, trace-free tensor fields can now be defined ($\mathscr{C}$ denotes the trace-free part with respect to $\tilde{\gamma}_{ij}$).

$$P_M = \widetilde{\Phi}_M,$$

$$P_{Ma_1\cdots a_{r+1}} = \mathscr{C}[\widetilde{D}_{a_1}P_{Ma_2\cdots a_{r+1}} - \tfrac{1}{2}r(2r-1)\widetilde{\mathscr{R}}_{a_1a_2}P_{Ma_3\cdots a_{r+1}}], \tag{5.3}$$

and analogously, $P_{Sa_1\cdots a_{r+1}}$ can be constructed from $\widetilde{\Phi}_S$. $P_{Ma_1\cdots a_r}$, and $P_{Sa_1\cdots a_r}$ are defined provided $r \leqslant k$. Their values at Λ, $\overline{M}_{a_1\cdots a_r}$ and $\overline{S}_{a_1\cdots a_r}$, are called 2^r mass- and angular-momentum moments, respectively.

Theorem 3: Our manifold N with metric (3.4) has a C^m-compactification. $\widetilde{\Phi}_M = \Omega^{-1/2}\Phi_M$ and $\widetilde{\Phi}_S = \Omega^{-1/2}\Phi_S$ (where Φ_M and Φ_S are the fields of Theorem 1 or Theorem 2) extend to C^{m-1} functions on $\widetilde{N}$. For all $r \leqslant m-1$, we have

$$\overline{M}_{a_1\cdots a_r} = M_{a_1\cdots a_r}, \quad \overline{S}_{a_1\cdots a_r} = S_{a_1\cdots a_r}. \tag{5.4}$$

Proof: Define a differentiable structure by calling $\bar{x}^i = r^{-2}x^i$ good coordinates on $\widetilde{N} = N \cup (\bar{r} = 0)$ and choose $\Omega = \bar{r}^2$. We obtain

$$\widetilde{\Phi}_M = \sum_{l=0}^{m-1}\frac{1}{l!}E_{i_1\cdots i_l}\bar{x}^{i_1}\cdots\bar{x}^{i_l} + O^\infty(\bar{r}^m), \tag{5.5}$$

$$\widetilde{\Phi}_S = \sum_{l=0}^{m-1}\frac{1}{l!}F_{i_1\cdots i_l}\bar{x}^{i_1}\cdots\bar{x}^{i_l} + O^\infty(\bar{r}^m), \tag{5.6}$$

and, using $\partial\bar{x}^i/\partial x^j = \bar{r}^2(\delta_{ij} - 2\bar{x}_i\bar{x}_j/\bar{r}^2)$,

$$\begin{aligned}\tilde{\gamma}_{ij} = \delta_{ij} + \sum_{l=2}^{M}(&\bar{x}^i\bar{x}^j\overline{A}_{i_1\cdots i_{l-2}}\bar{x}^{i_1}\cdots\bar{x}^{i_{l-2}} \\ &+ \bar{r}^2\delta_{ij}\overline{B}_{i_1\cdots i_{l-2}}\bar{x}^{i_1}\cdots\bar{x}^{i_{l-2}} + \bar{r}^2\bar{x}_{(i}\overline{C}_{j)i_1\cdots i_{l-3}}\bar{x}^{i_1}\cdots\bar{x}^{i_{l-3}} \\ &+ \bar{r}^4\overline{D}_{iji_1\cdots i_{l-4}}\bar{x}^{i_1}\cdots\bar{x}^{i_{l-4}}) + O^\infty(\bar{r}^{m+1}),\end{aligned} \tag{5.7}$$

where

$$\begin{aligned}\overline{A}\cdots &= A\cdots + 4D\cdots, & \overline{B}\cdots &= B\cdots,\\ \overline{C}\cdots &= C\cdots + 4D\cdots, & \overline{D}\cdots &= D\cdots.\end{aligned} \tag{5.8}$$

The conformal factor $\bar{r}^2$ is easily seen to satisfy the Geroch conditions (5.2).

It remains to show (5.4). Once again, we use "s algebra," which will first be adapted to the new coordinates in order to have always positive values for m. The expression

$$\beta_{i_1\cdots i_a k_1\cdots k_d} = B_{i_1\cdots i_a j_1\cdots j_b}\delta_{k_1k_2}\cdots\delta_{k_{c-1}k_c}\bar{r}^{c-e}\bar{x}^{k_{c+1}}\cdots\bar{x}^{k_d}\bar{x}^{j_1}\cdots\bar{x}^{j_b}, \tag{5.9}$$

where $c-e\equiv 0 \pmod 2$ and $c-e\geqslant 0$, is called term of order $m=b+d-e$ with $f=a+d$ "free indices" and "characteristic number" s, where

$$s\equiv m-(a+b) \pmod 2, \tag{5.10}$$

$$s\leqslant m-(a+b). \tag{5.11}$$

Under arithmetic operations, m and s behave as before but differentiation with respect to $\bar{x}^i$ now reduces the values of m and s.

According to this definition, $\widetilde{\Phi}_M$ and $\widetilde{\Phi}_S$ have $s=0$-terms, $\bar{h}_{ij}=\tilde{\gamma}_{ij}-\delta_{ij}$ has $s=2$-terms, the connection $\widetilde{\Gamma}^i{}_{jk}$ has $s=1$, and $\widetilde{\mathscr{R}}_{ij}$ consists of $s=0$-terms. We prove by induction that the tensor fields $P\cdots$, which appear in the Geroch definition (5.3) can be written as follows

$$\begin{aligned} P_{a_1\cdots a_k} &= \mathscr{C}[\bar{\partial}_{a_1}\cdots\bar{\partial}_{a_k}\widetilde{\Phi}] \\ &\quad + [\text{trace-free terms with } s=-k+2] \\ &\quad + O^\infty(\bar{r}^2). \end{aligned} \tag{5.12}$$

The first term on the right-hand side has $s=-k$. Note that the trace-free part in (5.12) is taken with respect to $\delta_{ij}(\mathscr{C})$ rather than $\tilde{\gamma}_{ij}(\widetilde{\mathscr{C}})$. For $k=0$ and $k=1$ the statement (5.12) is trivial; for $k=2$ it is easily shown. We may assume that (5.12) is valid for $P_{a_1\cdots a_{k-1}}$ as well as for $P_{a_1\cdots a_k}$. For $k+1$ indices we obtain from (5.3) (using $\mathscr{C}$ instead of $\widetilde{\mathscr{C}}$) and from the induction hypothesis,

$$\begin{aligned} P_{a_1\cdots a_{k+1}} &= \mathscr{C}[\widetilde{D}_{a_1}P_{a_2\cdots a_{k+1}} - \tfrac{1}{2}k(2k-1)\widetilde{\mathscr{R}}_{a_1a_2}P_{a_3\cdots a_{k+1}}] \\ &\quad + O^\infty(\bar{r}^2) \\ &= \mathscr{C}[\bar{\partial}_{a_1}P_{a_2\cdots a_{k+1}} - \widetilde{\Gamma}^j{}_{a_1a_2}P_{ja_3\cdots a_{k+1}} - \cdots \\ &\quad - \tfrac{1}{2}k(2k-1)\widetilde{\mathscr{R}}_{a_1a_2}P_{a_3\cdots a_{k+1}}] + O^\infty(\bar{r}^2) \\ &= \mathscr{C}[\bar{\partial}_{a_1}\mathscr{C}[\bar{\partial}_{a_2}\cdots\bar{\partial}_{a_{k+1}}\widetilde{\Phi}] \\ &\quad + \text{terms with } s=-k+1] + O^\infty(\bar{r}^2) \\ &= \mathscr{C}[\bar{\partial}_{a_1}\bar{\partial}_{a_2}\cdots\bar{\partial}_{a_{k+1}}\widetilde{\Phi}] \\ &\quad + [\text{tracefree terms with } s=-k+1] + O^\infty(\bar{r}^2). \end{aligned} \tag{5.13}$$

This proves (5.12). Hence we have, for Geroch's 2^k-moments,

$$P_{a_1\cdots a_k}|_\Lambda = \mathscr{C}[\bar{\partial}_{a_1}\cdots\bar{\partial}_{a_k}\widetilde{\Phi}]|_\Lambda + Y_{a_1\cdots a_k}. \tag{5.14}$$

From (5.5) and (5.6) we see that the first term on the right is equal to $M_{a_1\cdots a_k}$ for $\widetilde{\Phi}=\widetilde{\Phi}_M$ and $S_{a_1\cdots a_k}$ for $\widetilde{\Phi}=\widetilde{\Phi}_S$. The constant $Y_{a_1\cdots a_k}$ has k indices, but $s=-k+2$. This is only possible if $Y_{a_1\cdots a_k}$ contains at least one unit matrix:

$$Y_{a_1\cdots a_k} = \delta_{(a_1a_2}X_{a_3\cdots a_k)}. \tag{5.15}$$

Moreover, $Y_{a_1\cdots a_k}$ must be trace-free; therefore, it vanishes. We have then proved that our constants agree with the multipole moments defined by Geroch and Hansen. □

6. DISCUSSION

Our treatment so far has avoided certain questions to which we now turn.

Most pressing is, of course, the question as to whether our asymptotic series converge. This problem splits into two parts.

The first part consists in asking whether, for sufficiently large radius, the series in (3.1)–(3.4) converge when Φ_M, Φ_S, γ_{ij} belong to a given solution of the field equations. As we have seen in the preceding paragraph, there always exists for such a solution a C^k-Geroch compactification, when Ω is chosen to be $\bar{r}^2$. It is easily seen by the same methods that we could just as well choose

$$\Omega = \tfrac{1}{2}B^{-2}[(1+4\Phi_M^2+4\Phi_S^2)^{1/2}-1], \tag{6.1}$$

where $B^2=M^2+S^2$ (assuming $M^2+S^2\neq 0$). Using the results of Beig and Simon[4] it follows that, by a change of the unphysical chart, all unphysical variables can be made analytic. This proves that "some" multipole expansion converges, but not necessarily the one given in (3.1)–(3.4). Still we expect this to be true.

Even harder is the second part where we just assume that we are given the moments (3.5) in the expansion (3.1)–(3.4) and ask for conditions on the 2^k-moments for large k such that these series converge (in which case, they would also form a solution to Einstein's equations, by Theorem 2). The corresponding result in Newtonian theory suggests that

$$|M_{a_1\cdots a_k}|\leqslant AC^k,\quad |S_{a_1\cdots a_k}|\leqslant AC^k \quad (A, C \text{ positive numbers}) \tag{6.2}$$

might be the appropriate convergence criteria. However, convergence is in general far from being obvious, even if only a finite number of moments is nonvanishing. Although there are only a few types of terms in the expansions of Φ_M, Φ_S, and γ_{ij}, the number of terms explodes rather quickly when the fields are expressed in terms of multipole moments [(3.47)–(3.50)]. So the desired result can presumably not be inferred from mere counting of the number of these terms.

Another question is raised by our coordinate-dependent approach to multipole theory. Given a space-time of the form (3.1)–(3.4), satisfying the gauge condition (3.19), is there still a remaining coordinate freedom by which these moments could be changed? It is well known that the most general transformation which preserves the Euclidean metric δ_{ij} is

$$\bar{x}^i = R^i{}_jx^j + a^i, \tag{6.3}$$

where $R^i{}_j$ and a^i are a rigid rotation and a translation. Similarly, we expect all transformations $\bar{x}^i=p^i(x^k)$ which preserve asymptotically Euclidean metrics (2.10) and (2.11) to be of the form

$$p^i(x^k) = R^i{}_jx^j + g^i(x^k), \tag{6.4}$$

where $g^i(x^k)=O^\infty(\ln^* r)$. Assuming this and also $\Lambda_i(x^k)=O^\infty(r^{-(m+3)})$ and $\bar{\Lambda}_i(\bar{x}^k)=O^\infty(\bar{r}^{-(m+3)})$, essentially the same methods as the ones used throughout this paper enable us to show[17] that $g^i=a^i+O^\infty(r^{-m})$. So, if (6.4) holds, the mth order solution [which satisfies $\Lambda_i=O^\infty(r^{-(m+3)})$] is affected only by a rigid rotation and a translation. It is easily seen that these change the multipole moments in precisely the same way as Euclidean motions in flat space change the Newtonian moments. Of course, this can be shown in the geometric approaches, too (see Geroch[1] and Beig[5]).

A third question comes from the observation that the Hansen potentials Φ_M and Φ_S are particularly well suited for our task, but not necessarily singled out. There is the following result.

Theorem 4: Let $F_A(X_B)$ $(A, B = 1, 2)$ be two smooth functions of X_1, X_2 in a neighborhood of (0, 0) which satisfy

$$\left.\frac{\partial F_A}{\partial X_B}\right|_{(0,0)} = \delta_{AB} \tag{6.5}$$

and

$$F_A(X_1, X_2) = -F_A(-X_1, -X_2). \tag{6.6}$$

Then $\Phi'_M := F_1(\Phi_M, \Phi_S)$ and $\Phi'_S := F_2(\Phi_M, \Phi_S)$ have multipole expansions of the same form and give rise to the same multipole moments as Φ_M and Φ_S.

Proof: Because of the requirements on F_1, F_2 there are the following convergent Taylor expansions for Φ'_M, Φ'_S:

$$\Phi'_M = \Phi_M + C_{30}\Phi_M^3 + C_{21}\Phi_M^2\Phi_S + C_{12}\Phi_M\Phi_S^2 + C_{03}\Phi_S^3 + \cdots, \tag{6.7}$$

$$\Phi'_S = \Phi_S + D_{30}\Phi_M^3 + D_{21}\Phi_M^2\Phi_S + D_{12}\Phi_M\Phi_S^2 + D_{03}\Phi_S^3 + \cdots, \tag{6.8}$$

where $C_{\alpha\beta}$ and $D_{\alpha\beta}$ are constants.

Inserting the series (3.1) and (3.2) for Φ_M and Φ_S it is easily seen (i.e., by using "s-algebra") that Φ'_M and Φ'_S have expansions of the same form and that the nonlinear terms above get absorbed into the trace terms of Φ'_M and Φ'_S. Hence their multipole moments cannot be affected. □

As an application of Theorem 4 consider, in the static case, the one-parameter family of potentials U_a defined by

$$U_a = (1/4a)(\lambda^a - \lambda^{-a}), \quad a > 0. \tag{6.9}$$

The limit $a \to 0$ exists and is equal to $\frac{1}{2}\ln\lambda$. Hence, we have

$$U_a = (1/2a)\sinh(2aU_0), \quad a \geqslant 0. \tag{6.10}$$

U_a satisfies the field equations

$$\Delta(\gamma)U_a = 2a^2\mathscr{R}U_a, \tag{6.11}$$

$$\mathscr{R}_{ij}(\gamma) = 2(1 + 4a^2U_a^2)^{-1}D_iU_aD_jU_a. \tag{6.12}$$

Obviously, they take on their simplest form if U_0 is used. The potentials used by Geroch,[1] Hoenselaers,[14] and Hansen[2] reduce, in the static case, to $U_{1/4}$, $U_{1/2}$, and U_1, respectively.

Since $f(X) = (1/2a)\sinh 2aX$ is an analytic, antisymmetric function and $df/dX|_{X=0} = 1$, Theorem 4 can be applied: All members U_a of this family have multipole expansions producing the same set of moments.

Finally we should emphasize that our asymptotic conditions have been stated in terms of Φ_M, Φ_S, and γ_{ij}. The full metric $g_{\mu\nu}$ can be reconstructed from them if Eq. (2.4) can be solved for σ_i. This, as shown in Ref. 11, can be done iff the angular-momentum monopole ("dual mass") S vanishes. This is not the case for our Taub-NUT example: To write down $g_{\mu\nu}$ in this case one has the choice of, either, admitting angular ("wire") singularities in the metric or dispensing with asymptotic flatness. All this is, of course, well known.

ACKNOWLEDGMENT

We are greatly indebted to H. Urbantke for supplying the essential part of Lemma C. Helpful correspondence with A. Martin on the proof of Lemma A is also gratefuly acknowledged.

APPENDIX

Lemma A: On N, we are given a function

$$\rho(x^i) = O^\infty\left(\frac{\ln^q r}{r^{p+3}}\right), \quad p, q = 0, 1, 2\ldots. \tag{A1}$$

Then Poisson's equation (Δ is the Flat Laplace operator)

$$\Delta\Phi = 4\pi\rho \tag{A2}$$

has a solution $\Phi(x)$ on N which vanishes for $r \to \infty$, and every such $\Phi(x)$ has a multipole expansion up to order p, i.e., there are symmetric, trace-free constants $Q, Q_i, \ldots, Q_{a_1\cdots a_l}$ such that

$$\Phi(x) = O^\infty(r^{-1}\ln^{q+1}r) \quad \text{for } p = 0, \tag{A3}$$

$$\Phi(x) = \sum_{l=0}^{p-1}\frac{Q_{a_1\cdots a_l}x^{a_1}\cdots x^{a_l}}{r^{2l+1}l!} + O^\infty\left(\frac{\ln^{q+1}r}{r^{p+1}}\right) \quad \text{for } p \geqslant 1. \tag{A4}$$

Proof: Every solution of $\Delta\overline{\Phi} = 0$, which vanishes for $r \to \infty$, has a multipole expansion of the form (A4) up to arbitrary orders.[18] It remains to show that the Poisson equation above admits some solution Φ_{special} of the form (A3), (A4). This is a special case of Lemma 5 of Meyers.[19] There is also a way to prove this lemma by generalizing the method given in Refs. 10 and 11 in the cases $p = 1$ and $p = 2$. In this case, one proceeds by showing that

$$\Phi_{\text{special}}(\mathbf{x}) = -\int_N \frac{\rho(\mathbf{x})}{|\mathbf{x} - \mathbf{x}'|}d^3x' \tag{A5}$$

has the desired properties, and this can be done by using the inequality[17,20]

$$\left|\frac{1}{(1 + w^2 - 2\omega\xi)^{1/2}} - \sum_{l=0}^{p-1}w^lP_l(\xi)\right| \leqslant \frac{2p+1}{(1 + w^2 - 2w\xi)^{1/2}}w^p, \tag{A6}$$

where $w \geqslant 0$, $|\xi| \leqslant 1$, $p \geqslant 1$, and P_l is the Legendre polynomial of degree l. □

Lemma B: The equation

$$\Delta\Phi = \frac{x_{a_1}\cdots x_{a_k}}{r^{n+2}}, \quad k = 0,1,2\ldots, \quad n \geqslant 2k + 2 \tag{A7}$$

has a solution of the form

$$\Phi = \lambda_k\frac{x_{a_1}\cdots x_{a_k}}{r^n} + \lambda_{k-2}\frac{\delta_{(a_1a_2}x_{a_3}\cdots x_{a_k)}}{r^{n-2}} + \cdots + \begin{cases}\lambda_0\dfrac{\delta_{(a_1a_2}\delta_{a_3a_4}\cdots\delta_{a_{k-1}}\delta_{a_k)}}{r^{n-k}} & \text{if } k \text{ is even,}\\[2ex] \lambda_1\dfrac{\delta_{(a_1a_2}\cdots\delta_{a_{k-2}a_{k-1}}x_{a_k)}}{r^{n-k+1}} & \text{if } k \text{ is odd.}\end{cases} \tag{A8}$$

Proof: Let Δ act on Φ given above, use

$$\Delta\frac{x_{a_1}\cdots x_{a_k}}{r^n} = k(k-1)\frac{\delta_{(a_1a_2}x_{a_3}\cdots x_{a_k)}}{r^n} + n(n - 2k - 1)\frac{x_{a_1}\cdots x_{a_k}}{r^{n+2}}, \tag{A9}$$

and compare the coefficients with (A7).

There results a system of equations in λ_l, n, and k which has a solution in the case $n \geqslant 2k+2$. □

Lemma C: Consider a symmetric tensor field t_{ij} on a simply connected region of $\mathbb{R}^3$ satisfying

$$\Delta t_{ij} = 0, \tag{A10}$$

$$(t_{ij} - \tfrac{1}{2}\delta_{ij} t_{kk})_{,j} = 0. \tag{A11}$$

Then there exists a field g_i (with $\Delta g_i = 0$) such that

$$t_{ij} = g_{i,j} + g_{j,i}. \tag{A12}$$

If $t_{ij} = O^{\infty}(r^{-(m+1)})$, $m \geqslant 0$, g_i can be chosen to be $O^{\infty}(r^{-m})$.

Proof: Because of the symmetry of t_{jk}, the field $\Gamma_{ijkl}: = \partial_{[i} t_{j][k,l]}$ has the same symmetry properties as a Riemann tensor. Moreover, from (A10) and (A11), we have

$$\Gamma_{ijil} = \tfrac{1}{4}(\gamma_{il} - \tfrac{1}{2}\delta_{il}\gamma_{mm})_{,ij} + \tfrac{1}{4}(\gamma_{ij} - \tfrac{1}{2}\delta_{ij}\gamma_{mm})_{,il} - \tfrac{1}{4}\Delta\gamma_{jl} = 0. \tag{A13}$$

Since Γ_{ijkl} has all the symmetries of a Weyl tensor and the dimension of space is 3, it must vanish identically. This means that the integrability conditions for the existence of a field u_{ij} are satisfied such that

$$u_{ij,k} = \tfrac{1}{2}(t_{ij,k} - t_{jk,i} + t_{ki,j}). \tag{A14}$$

It follows that

$$(u_{ij} + u_{ji})_{,k} = t_{ij,k}, \tag{A15}$$

$$t_{ij} = u_{ij} + u_{ji} + \text{const.} \tag{A16}$$

But we can always set the constant equal to 0 by a redefinition of u_{ij}. From (A14) we also infer that

$$u_{i[j,k]} = 0 \tag{A17}$$

which gives rise to the existence of g_i satisfying

$$u_{ij} = g_{i,j}. \tag{A18}$$

Together with (A16) this gives (A12), and (A11) implies $\Delta g_i = 0$. We have only integrated functions and one-forms in this whole argument which we can do globally since our domain is simply connected. If no fall-off conditions on t_{ij} are imposed it is easy to show that g_i is determined uniquely up to addition of $c_{ij}x^j + d_i$, where $c_{ij} = -c_{ji}$ and d_i are constants. If t_{ij} vanishes for $r \to \infty$, then, from Lemma A, we have

$$t_{ij} = \frac{T_{ij}}{r} + O^{\infty}\left(\frac{1}{r^2}\right), \tag{A19}$$

but the constants T_{ij} must vanish by virtue of (A11), so $t_{ij} = O^{\infty}(r^{-2})$. In this case there is available the following explicit formula for g_i:

$$g_i(x) = -x_j x_k \int_1^{\infty} \lambda u_{jk,i}(x\lambda)\, d\lambda = t_{ij}x_j - \frac{1}{2}x_j x_k \int_1^{\infty} \lambda t_{jk,i}(x\lambda)\, d\lambda. \tag{A20}$$

Hence, $g_i = O^{\infty}(r^{-m})$ if $t_{ij} = O^{\infty}(r^{-(m+1)})$ for all $m \geqslant 0$. □

[1] R. Geroch, J. Math. Phys. **11**, 2580 (1970).
[2] R. O. Hansen, J. Math. Phys. **15**, 46 (1974).
[3] R. Beig and W. Simon, Commun. Math. Phys. **78**, 75 (1980).
[4] R. Beig and W. Simon, Proc. R. Soc. London Sec. A **376**, 333 (1981).
[5] R. Beig, Acta Phys. Austr. **53**, 249 (1981).
[6] P. Kundu, J. Math. Phys. **22**, 2006 (1981). For an earlier attempt see P. Kundu, J. Math. Phys. **22**, 1236 (1981).
[7] N. O'Murchadha, Preprint 19, University College, Cardiff, Wales (1976).
[8] K. S. Thorne, Rev. Mod. Phys. **52**, 299 (1980).
[9] Y. Tanabe, Prog. Theor. Phys. **55**, 106 (1976).
[10] R. Beig, Gen. Relativ. Gravit. **12**, 439 (1980).
[11] R. Beig and W. Simon, Gen. Relativ. Gravit. **12**, 1003 (1980).
[12] J. L. Synge, Proc. R. Ir. Acad. Sect. A **69**, 1 (1970).
[13] R. Geroch, J. Math. Phys. **12**, 918 (1971).
[14] C. Hoenselaers, Prog. Theor. Phys. **55**, 406 (1976).
[15] F. A. E. Pirani, A. Trautman, F. A. E. Pirani, and H. Bondi, in *Lectures on General Relativity* (Prentice-Hall, Englewood Cliffs, NJ, 1964), p. 288.
[16] A. M. Taub, Ann. Math. **53**, 472 (1951).
[17] W. Simon, thesis, Universität Wien, Vienna, 1980.
[18] O. D. Kellogg, *Foundations of Potential Theory* (Springer, Berlin, 1926).
[19] N. Meyers, J. Math. Mech. **12**, 247 (1963).
[20] A. Martin (private communication).

Compatibility of rigid motions with the relativistic incompressibility condition

V. Del Olmo and J. Olivert
Departamento de Mecánica y Astronomía, Facultad de Matemáticas, Burjassot (Valencia), Spain

(Received 27 December 1981; accepted for publication 23 April 1982)

The compatibility of the hypoelastic-Synge (isotropic case) and hypoelastic-Carter and Quintana (fully isotropic case) almost-thermodynamic material schemes with the relativistic incompressibility condition, given by Ferrando and Olivert, is analyzed. The behavior of those schemes in Born-rigid motion is also studied and an additional stipulation is joined to the Born rigidity concept. This requisite leads to the vanishing of the relativistic stress tensor spatial variation and, in hypoelastic-Carter and Quintana and hypoelastic-Maugin almost-thermodynamic material schemes, leads to the absence of mechanical power originated by internal rotation. The rigidity definition that is proposed remains valid for more general almost-thermodynamic material schemes, as far as the compatibility with the incompressibility condition quoted above is concerned.

PACS numbers: 04.20. − q, 04.40. + c

1. INTRODUCTION

This paper is a continuation of the work of Ferrando and Olivert[1] in which a definition of incompressibility in general relativity is given for almost-thermodynamic material schemes. That concept reproduces the known one for perfect fluids. In this direction, a new step is to study the concept of rigidity. This notion is treated in the present work.

Here we establish a study of the rigid motions in the almost-thermodynamic material schemes in order to verify the absence of perturbations in them and to fulfill the incompressibility condition suggested in Ref. 1.

First, one works in hypoelastic schemes inasmuch as, at first sight, it looks as though the application of the Born-rigidity definition to a Hooke's law of general type must lead to the vanishing of some kind of variation of the relativistic stress tensor and thereby to the incompressibility condition. Subsequently, the general case is considered.

In Sec. 2 we introduce three classes of hypoelastic almost-thermodynamic material schemes: the Synge's case, the Maugin's and the Carter and Quintana's case. Following the line suggested in Ref. 1, we study the associated principal shock waves speeds. The corresponding results are given in Theorem 2.1, for the Synge's case, and in Theorem 2.2, for the Carter and Quintana's one. Section 3 is devoted to proving the compatibility of the hypoelastic almost-thermodynamic material schemes with incompressibility. In every case, we get satisfactory results.

The Secs. 3 and 4 analyze the behavior of the Born rigidity and its possible insufficiencies. Theorem 4.1 shows us that this hypothesis is enough to prove the absence of shock waves due to the infinitesimal discontinuities of the 4-velocity of the scheme. Hence, the Born rigidity satisfies our purposes. Therefore, according to Theorem 4.2, the hypoelastic-Synge, Born-rigid, almost-thermodynamic material schemes, are incompressible. This is not true, however, for the hypoelastic-Carter and Quintana almost-thermodynamic material schemes or for the Maugin ones in Born-rigid motion (Propositions 4.3 and 4.4).

In Sec. 5, one introduces the concept of rigid almost-thermodynamic schemes, which includes the Born-rigidity one and adds the vanishing of the spatial change of the relativistic stress tensor. With this concept, one attempts to solve the problems which appeared in Sec. 4. The work concludes with an analysis of the Born-rigidity and irrotationality considered together in an almost-thermodynamic material scheme. We prove a theorem which characterizes them geometrically.

Before going into the subject, we'll explain some notations.

According to Lichnerowicz[2] and Sachs and Wu,[3] the 4-dimensional space-time manifold M will be pseudo-Riemannian and hyperbolic, i.e., endowed with a metric tensor field g of signature (3,1), connected, of Hausdorff type, and orientable. The linear connection ∇ is the unique one that M possesses, compatible with g, and without torsion. The coordinate x^4 of a point, belonging to a coordinate neighborhood of a given local chart F, coincides with the coordinate time in the system described by F.

The tangent space at $p\in M$ is represented by T_pM. The measure units have been chosen so that the light velocity in the vacuum has the constant value 1.

Everywhere, the indices represented by Latin letters take values from 1 to 4; the Greek ones are restricted to the values 1, 2, 3, unless explicitly noted.

When we use the concept of spatial tensor in a material scheme D, we are hinting that it is orthogonal to the 4-velocity of D.

Finally, when we write Born-rigid almost-thermodynamic material schemes we want to indicate that we are dealing with rigid motions, which are completely compatible with elasticity hypotheses.

2. HOOKE'S LAWS: PRINCIPAL ELASTIC WAVES

We present in this section some definitions of elastic schemes that have been proposed in the relativistic literature. In our study, we'll consider an almost-thermodynamic material scheme D in the space-time manifold M as it appears in Ref. 1; i.e., D is a domain of the space-time in which is defined a second-order tensor field T (energy–momentum

 0022-2488/83/051172-10$02.50

tensor) that is normal and such that if u is its 4-velocity vector and $-\rho$ the associated eigenvalue, $\rho > 0$. We'll admit besides the Taub hypothesis,[2,4]

$$\rho = r(1+\epsilon) \tag{2.1}$$

where ρ is called the proper mass-energy density of the scheme, r its matter density, and ϵ its specific internal energy.

As is known, in every elastic scheme a relation is established between the relativistic stress tensor t—spatial projection by means of Eckart tensor $\gamma = g + u \otimes u$ of the energy-momentum tensor—and the strain rate tensor of the scheme

$$d = \tfrac{1}{2}\pounds_u \gamma, \tag{2.2}$$

where $\pounds_u$ symbolizes the Lie derivative with respect to the 4-velocity u of the tensor that follows it.

Henceforth, that relation is considered linear and the scheme is called hypoelastic.

We show, at first, the Hooke's law suggested by Synge,[5] which can be expressed as

$$\nabla_u\, t_{ij} = C_{ijkl}\, d^{\,kl}, \tag{2.3}$$

where ∇_u is the covariant derivative with respect to the 4-velocity of the scheme and is given by the linear connection ∇ of the space-time manifold. The t_{ij} and $d^{\,kl}$ are the components of the relativistic stress tensor and the strain rate tensor, respectively.

The elastic tensor C_{ijkl} presents the symmetries

$$C_{(ij)(kl)} = C_{ijkl} = C_{klij} \tag{2.4}$$

and, in the isotropic case, has the form

$$C_{ijkl} = -\lambda g_{ij} g_{kl} - \mu (g_{ij} g_{jl} + g_{il} g_{jk}), \tag{2.5}$$

where g is the metric tensor in M, λ, and μ the Lamé coefficients.

Maugin,[6] suggests a hypoelastic almost-thermodynamic material scheme in which the Hooke's law is given by

$$\gamma^k{}_i \gamma^l{}_j \pounds_u t_{kl} + t_{ij} d^{\,k}{}_k = H_{ij}{}^{kl}(t,r) d_{kl}, \tag{2.6}$$

$d^{\,k}{}_k$ the expansion rate and $H_{ij}{}^{kl}$ the mixed components of a spatial tensor (elastic tensor) linear function of r. Hereafter, we consider the H tensor in the form $H = r\bar{C}$ without any dependence of the relativistic stress tensor, and in which

$$\bar{C}_{ijkl} = -\lambda \gamma_{ij} \gamma_{kl} - \mu (\gamma_{ik} \gamma_{jl} + \gamma_{il} \gamma_{jk}) \tag{2.7}$$

(zero-order hypoelastic schemes).

Lastly, we refer to the hypoelastic-Carter and Quintana schemes.[7] The Hooke's law is expressed in those schemes as

$$\gamma^k{}_i \gamma^l{}_j \pounds_u t^{\,ij} + t^{\,kl} d^{\,c}{}_c = -E^{\,klab} d_{ab} \tag{2.8}$$

and the elastic tensor E presents, in the fully isotropic case, the local form

$$E^{\,klab} = \tfrac{1}{9} A \gamma^{kl} \gamma^{ab} + \tfrac{1}{2} B \left(\gamma^{k(a} \gamma^{b)l} - \tfrac{1}{3} \gamma^{kl} \gamma^{ab}\right), \tag{2.9}$$

A and B being the Carter and Quintana elastic scalars.

By setting, in Eq. (2.9),

$$\lambda = A/9 - B/15 \text{ and } \mu = B/10, \tag{2.10}$$

we can see that the Carter and Quintana elastic tensor $-E$ coincides formally, but for the choice of coefficients, with the Maugin one.

Now, in order to study subsequently the incompressibility and the rigidity in hypoelastic schemes, our purpose is to obtain the principal shock wave speeds[1] in hypoelastic almost-thermodynamic material schemes. We'll use the Hadamard discontinuities method as it is used in Ref. 1. For other questions about the relativistic waves in almost-thermodynamic material schemes, the same reference may be consulted.

Besides the suitable Hooke's law in each case, we often use the conservation equations and the continuity one

$$\nabla_i T^{\,ij} = 0, \quad \nabla_i (ru^i) = 0 \tag{2.11}$$

in the following form, that may be derived easily from Eqs. (2.11):

$$ru^i \nabla_i \epsilon + t^{\,ij} d_{ij} = 0, \tag{2.12}$$

$$rf^r{}_j u^k \nabla_k u^j + \gamma^r{}_i \gamma^k{}_j \nabla_k t^{\,ij} = 0, \tag{2.13}$$

and

$$u^i \nabla_i r + r d^{\,i}{}_i = 0, \tag{2.14}$$

where f is the tensor index of the scheme

$$f = (1+\epsilon)\gamma + t/r, \tag{2.15}$$

defined in Ref. 1.

Now we can prove the result that follows.

Theorem 2.1: In a hypoelastic-Synge almost-thermodynamic material scheme, for the isotropic case, the longitudinal and transverse shock wave speeds have the expressions

$$U_L^2 = \frac{\lambda + 2\mu}{r(1+\epsilon) + t_\parallel}, \tag{2.16}$$

$$U_{T_\alpha}^2 = \frac{\mu}{r(1+\epsilon) + t_\alpha} \quad \alpha = 2,3, \tag{2.17}$$

where $t_\parallel$ and t_α (α:2,3) are the tensor t spatial eigenvalues.

Proof: By applying Hadamard discontinuities to Eq. (2.3) one obtains

$$-U\delta t^{\,ij} = C^{\,ijkl} \lambda_k \delta u_l, \tag{2.18}$$

λ^k being the shock wave spatial propagation direction vector,[6] δ the infinitesimal discontinuity of the tensor to which it is applied, and U the shock wave propagation speed.[17]

Contraction with $\gamma^r{}_i \lambda_j$ in Eq. (2.18) gives

$$-U\gamma^r{}_i \lambda_j \delta t^{\,ij} = \gamma^r{}_i \lambda_j C^{\,ijkl} \lambda_k \delta u_l, \tag{2.19}$$

and using the expression

$$rf^r{}_j U \delta u^j = \gamma^r{}_i \lambda_j \delta t^{\,ij} \tag{2.20}$$

this may be obtained by applying Hadamard discontinuities in Eq. (2.13), we can write:

$$Q_{ri} \delta u^i = 0 \tag{2.21}$$

where Q_{ri} is the 2-covariant symmetric tensor defined as

$$Q_{ri} = rU^2 f_{ri} + \gamma_r{}^l \lambda^j \lambda^k C_{ljki}. \tag{2.22}$$

In the isotropic case, Eqs. (2.5) and (2.22) lead to

$$Q_{ri} = rU^2 f_{ri} - (\lambda + \mu) \lambda_r \lambda_i - \mu \gamma_{ri} \tag{2.23}$$

after using the spatiality of λ_i.

It is easy to check that Q_{ri} is a spatial and symmetric tensor and the system (2.22) may be written[1] in the form

$$Q_{\perp_{ri}} \delta u^i_\perp + q_r \delta u_\parallel = 0, \tag{2.24}$$

$$q_i \delta u^i_\perp + Q_\parallel \delta u_\parallel = 0, \tag{2.25}$$

where

$$Q_{\perp_{ri}} = S^k{}_r S^l{}_i Q_{kl}, \tag{2.26}$$

$$Q_{\parallel} = Q_{ri}\lambda^r\lambda^i, \tag{2.27}$$

$$q_i = \lambda^k Q_{kl} S^l{}_i, \tag{2.28}$$

and

$$S_{ij} = \gamma_{ij} - \lambda_i\lambda_j \tag{2.29}$$

is the Maugin projector.[5] Here Maugin also arrives to the decomposition

$$\delta u^i = \delta u^i_\perp + \lambda^i \delta u_\parallel . \tag{2.30}$$

It is trivial to verify that in our case $q_i = 0$, given the spatiality of λ_i and the principal shock wave hypothesis

$$t_{ij}\lambda^j = t_\parallel \lambda_i . \tag{2.31}$$

Moreover,

$$Q_{\perp_{ri}} = S^s{}_r S^j{}_i (rU^2 f_{sj} - \mu\gamma_{sj}) \tag{2.32}$$

and

$$Q_\parallel = rU^2 f_{ri}\lambda^r\lambda^i - \lambda - 2\mu . \tag{2.33}$$

If we now consider the longitudinal shock waves, $\delta u_\perp = 0$, $\delta u_\parallel \neq 0$, and therefore, Eq. (2.16) follows from Eq. (2.25).

For the transverse shock waves $\delta u_\parallel = 0$, $\delta u_\perp \neq 0$, and, from Eqs. (2.24) and (2.32),

$$S^s{}_r S^j{}_i (rU^2 f_{sj} - \mu\gamma_{sj})\delta u^i_\perp = 0 \tag{2.34}$$

By contraction with the t eigenvectors (principal directions) d^r_α (α:2,3)—it must be remembered that $t_1 = t_\cdot$ and $d^r_1 = \lambda^r$—one gets easily Eq. (2.17), keeping in mind the expressions

$$t_{ri}\delta u^i_\perp d^r_\alpha = a_\alpha t_\alpha,\ \alpha:2,3, \tag{2.35}$$

and

$$\gamma_{ri}\delta u^i_\perp d^r_\alpha = a_\alpha,\ \alpha:2,3, \tag{2.36}$$

where

$$\delta u_\perp = a_2 d_2 + a_3 d_3 \tag{2.37}$$

and the fact that, for each transverse direction, $a_\alpha \neq 0$, α:2,3.

We want to indicate here that Synge obtains[5] the classics results in Relativity:

$$U^2_L = \frac{\lambda + 2\mu}{\rho}, \quad U^2_T = \frac{\mu}{\rho}. \tag{2.38}$$

It must be realized that the difference between these relations and the ones we got in Theorem 2.1, is due to the use, in this work, of principal shock waves.

Equations (2.16) and (2.17) generalize the classic results, as may be seen when one divides by c^2 (the constant light velocity in the vacuum, squared) in order to get physical homogeneity, and performs the classical limit process ($c \to \infty$).

We reproduce here the results that Maugin[6] obtains in zero-order hypoelastic-Maugin, almost-thermodynamic schemes for the principal shock waves

$$U^2_L = \frac{r(\lambda + 2\mu) + 3t_\parallel}{r(1+\epsilon) + t_\parallel} \tag{2.39}$$

and

$$U^2_{T_\alpha} = \frac{r\mu + t_\alpha}{r(1+\epsilon) + t_\alpha};\ \alpha:2,3, \tag{2.40}$$

which also generalize the classical results. We want to say, moreover, that the Lamé coefficients λ, μ in hypoelastic-Synge and hypoelastic-Carter and Quintana schemes may depend linearly upon the matter density but this dependence does not occur in the hypoelastic-Maugin schemes.

We conclude the present section by using a method similar to the one we followed in the proof of Theorem 2.1 for the hypoelastic-Carter and Quintana schemes case. The results are found in Theorem 2.2.

Theorem 2.2: The longitudinal and transverse principal shock wave speeds in a hypoelastic-Carter and Quintana almost-thermodynamic material scheme, have the expressions

$$U^2_L = \frac{\lambda + 2\mu - t_\parallel}{r(1+\epsilon) + t_\parallel}, \tag{2.41}$$

$$U^2_T = \frac{\mu - t_\parallel}{r(1+\epsilon) + t_\alpha}, \alpha:2,3. \tag{2.42}$$

Proof: Let us take Hadamard discontinuities in Eq. (2.8). After developing the Lie derivative $\pounds_u t^{ij}$, we get

$$- U|L|\gamma^k{}_i\gamma^l{}_j\delta t^{ij} - t^{cl}l_c\delta u^k - t^{kc}l_c\delta u^l + t^{kl}\lambda_i|L|\delta u_i$$
$$= - E^{kl}{}_{ab}\lambda^a\delta u^b|L|, \tag{2.43}$$

where L is the vector

$$L = l + g(u,l)u \tag{2.44}$$

and l the orthogonal vector, at $p\in D$, to the timelike characteristic hypersurface associated to the shock wave $H_p{}^1$.

Taking into account that L has the same direction as λ^i, we can write

$$l^i = |L|\lambda^i - g(u,l)u^i. \tag{2.45}$$

By substituting Eq. (2.45) in Eq. (2.43), after contraction with λ_1, use of Eq. (2.20), and due to the spatial character of δu, one obtains

$$\bar{Q}_{ri}\delta u^i = 0, \tag{2.46}$$

where the tensor $\bar{Q}$ has the local form

$$\bar{Q}_{ri} = - rU^2 f_{ri} - t_\parallel\gamma_{ri} + E_{rlji}\lambda^j\lambda^l \tag{2.47}$$

for principal shock waves. $\bar{Q}_{ri}$ is, obviously, spatial and symmetric.

However, if we make use of the decomposition given by Eqs. (2.24) and (2.25) we can't guarantee, in this case, the vanishing of $\bar{q}_i$. If we restrict ourselves to the fully isotropic case, i.e., the one for which the elastic tensor E has the form given by Eq. (2.9), the use of Eq. (2.10) leads to

$$\bar{Q}_{ri} = - rU^2 f_{ri} + (\mu - t_\cdot)\gamma_{ri} + (\lambda + \mu)\lambda_r\lambda_i, \tag{2.48}$$

and application of the decomposition quoted above as well as the vanishing of $\bar{q}_i$ allow us to write

$$\bar{Q}_{\perp_{ri}}\delta u^i_\perp = 0 \tag{2.49}$$

and

$$\bar{Q}_\parallel \delta u_\parallel = 0, \tag{2.50}$$

since

$$\bar{Q}_{\perp ri} = S^s{}_r S^j{}_i(-rU^2 f_{sj} + (\mu - t_{\|})\gamma_{sj}) \tag{2.51}$$

and

$$\bar{Q}_{ri} = -(r(1+\epsilon) + t_{\|})U^2 + (\lambda + 2\mu) - t_{\|}, \tag{2.52}$$

as usual.

Now, consider longitudinal shock waves; $\bar{Q}_{\|} = 0$ and we derive Eq. (2.41) from Eqs. (2.50) and (2.52).

Furthermore, for transverse shock waves, Eq. (2.49) gives

$$S^s{}_r S^j{}_i(-rU^2 f_{sj} + (\mu - t_{\|})\gamma_{sj})\delta u^i_\perp = 0, \tag{2.53}$$

and, hence, the same procedure we used in Theorem 2.1, leads us to Eq. (2.42).

We want to comment briefly, in order to conclude this section, on the differences between the expressions we have obtained in Theorem 2.2 and the Eqs. (2.39) and (2.40), found by Maugin.[6] Besides the explicit absence of the matter density in the Carter and Quintana's case, as we noted above, the differences related to the eigenvalues $t_{\|}$ and t_α (α:2,3), are due to the covariant or contravariant form of the tensor Lie derivative in the respective Hooke's law.

The results of Theorem 2.2 generalize, of course, the classical ones.

3. INCOMPRESSIBILITY IN THE HYPOELASTIC ALMOST-THERMODYNAMIC SCHEMES

For the time being, our purpose is to survey the consequences of the application of the relativistic incompressibility hypothesis[1] to hypoelastic almost-thermodynamic schemes, in order to obtain results which are compatible with those shown there. In these conditions, we can use anyone of the Hooke's law defined in Sec. 2 and consider rigidity hypotheses which lead us to generalize classical results.

According to Ref. 1, an almost-thermodynamic material scheme is incompressible if the spatial change of the dynamic volume tensor K takes zero value across the stream lines, i.e.,

$$\gamma^k{}_i \gamma^l{}_j \nabla_u K^{ij} = 0, \tag{3.1}$$

with

$$K = f/r, \tag{3.2}$$

and f is the tensor index, given by Eq. (2.15).

As is known,[1] Eq. (3.1) becomes

$$\gamma^r{}_i \gamma^l{}_j \nabla_u t^{rl} + (2t^{rl} + r(1+\epsilon)\gamma^{rl})\gamma^k{}_i \nabla_k u^i = \gamma^{rl} t^{ki} d_{ki}, \tag{3.3}$$

which leads, after applying Hadamard discontinuities, to

$$\gamma^r{}_i \gamma^l{}_j U \delta t^{ij} = -\gamma^{rl} t^k{}_i \lambda_k \delta u^i + (2t^{rl} + r(1+\epsilon)\gamma^{rl})\lambda_i \delta u^i. \tag{3.4}$$

In order to study the compatibility of elastic almost-thermodynamic material schemes with the incompressibility, we'll jointly consider both hypotheses, on purpose that the corresponding principal shock waves to be longitudinal and satisfy:

$$U_L^2 = 1. \tag{3.5}$$

For the hypoelastic-Maugin, almost-thermodynamic material schemes, the compatibility claimed has already been discussed[1] and here we'll limit ourselves to studying it in some detail in the hypoelastic-Synge and hypoelastic-Carter and Quintana almost-thermodynamic material schemes. The results are:

Theorem 3.1: The isotropic hypoelastic-Synge almost-thermodynamic material schemes are compatible with the incompressibility condition (3.1).

Proof: Let us consider Eq. (2.18) contracted with $\gamma^r{}_i \gamma^l{}_j$ together with Eq. (3.4), derived from incompressibility equation (3.1). After adding the two equations and contracting with λ_l, we arrive at

$$R_{ri}\delta u^i = 0, \tag{3.6}$$

where

$$R_{ri} = \gamma_r{}^a \lambda^b C_{abij}\lambda^j - \lambda_r t^k{}_i \lambda_k + (2t_r{}^l + r(1+\epsilon)\lambda_r)\lambda_i, \tag{3.7}$$

and inasmuch as we are considering principal shock waves, Eq. (3.7) reduces to

$$R_{ri} = (t_{\|} + r(1+\epsilon) - (\lambda + \mu))\lambda_r \lambda_i - \mu\gamma_{ri}, \tag{3.8}$$

taking into account the elastic-Synge tensor C_{abji} in the form given by Eq. (2.5).

From Eq. (3.8) we deduce that R_{ri} is spatial and symmetric. Hence, we can write

$$R_{\perp ri}\delta u^i_\perp + r_r \delta u^i_{\|} = 0,$$

$$r_i \delta u^i_\perp + R_{\|}\delta u_{\|} = 0,$$

where, obviously, $r_i = 0$, by virtue of its definition in Eq. (2.28). By carrying out the suitable calculation, we obtain

$$R_{\perp ri}\delta u^i_\perp = 0, \tag{3.9}$$

$$R_{\|}\delta u_{\|} = 0, \tag{3.10}$$

with

$$R_{\perp ri} = -\mu S^s{}_r S^j{}_i \gamma_{sj} \tag{3.11}$$

and

$$R_{\|} = t_{\|} + r(1+\epsilon) - (\lambda + 2\mu). \tag{3.12}$$

As may be observed, the shock waves' speed U, has been eliminated in Eqs. (3.11) and (3.12). On the other hand, for a longitudinal shock wave, $R_{\|} = 0$, i.e.,

$$t_{\|} + r(1+\epsilon) = \lambda + 2\mu \tag{3.13}$$

and, for transverse shock waves,

$$-\mu S^s{}_r S^j{}_i \gamma_{sj}\delta u^i_\perp = 0. \tag{3.14}$$

Taking into account the properties of the Maugin projector and due to the spatiality of $\delta u_\perp$ and its orthogonality to the spatial propagation direction vector λ^i, one can write

$$-\mu\delta u_{\perp r} = 0. \tag{3.15}$$

From the expressions obtained in Theorem 2.1, one gets

$U_L^2 = 1.$

Furthermore, either transverse shock waves don't exist or their speeds satisfy

$U^2_{T_\alpha} = 0$ $\quad\alpha$:2,3, which may be considered equivalent.

The outcome relative to the hypoelastic-Carter and Quintana almost-thermodynamic schemes is also satisfactory.

Theorem 3.2: On the assumption of fully isotropy, the hypoelastic-Carter and Quintana almost-thermodynamic material schemes, are compatible with the relativistic incompressibility.

Proof: Again consider the expression (3.4) which, contracted with λ_l, yields

$$U\gamma_{ri}\lambda_j\delta t^{ij} = -t_{\|}\lambda_r\lambda_i\delta u^i + (2t_{ri} + r(1+\epsilon))\lambda_r\lambda_i\delta u^i, \quad (3.16)$$

and the one obtained from Eqs. (2.43) and (2.44) for principal shock waves:

$$-U\gamma_{ri}\lambda_j\delta t^{ij} = (t_{\|}\gamma_{ri} - E_{rlji}\lambda^j\lambda^l)\delta u^i. \quad (3.17)$$

By adding Eq. (3.16) to Eq. (3.17), one arrives at

$$\bar{R}_{ri}\delta u^i = 0, \quad (3.18)$$

with

$$\bar{R}_{ri} = t_{\|}\gamma_{ri} - E_{rlji}\lambda^j\lambda^l - t_{\|}\lambda_r\lambda_i + (2t_{\|} + r(1+\epsilon))\lambda_r\lambda_i. \quad (3.19)$$

If our hypoelastic almost-thermodynamic material scheme is fully isotropic, Eqs. (2.8) and (2.9) lead us to

$$\bar{R}_{ri} = (r(1+\epsilon) + t_{\|} - (\lambda+\mu))\lambda_r\lambda_i + (t_{\|} - \mu)\gamma_{ri} \quad (3.20)$$

after using the unitary spatial character of λ^i.

With these conditions, $\bar{R}_{ri}$ is spatial and symmetric. The usual argument carries us to the system

$$\bar{R}_{\perp ri}\delta u^i_{\perp} = 0, \quad (3.21)$$

$$\bar{R}_{\|}\delta u_{\|} = 0, \quad (3.22)$$

with

$$\bar{R}_{\perp ri} = (t_{\|} - \mu)S^s{}_r S^j{}_i\gamma_{sj} \quad (3.23)$$

and

$$\bar{R}_{\|} = r(1+\epsilon) + 2t_{\|} - (\lambda + 2\mu). \quad (3.24)$$

As $\bar{R}_{\|} = 0$ for the longitudinal shock waves, it follows, in this case, that

$$r(1+\epsilon) + 2t_{\|} = (\lambda + 2\mu); \quad (3.25)$$

the same method carried out for Theorem 3.1, leads to

$$\mu = t_{\|} \quad (3.26)$$

for the transverse ones. Besides, there is not transverse shock waves.

Keeping in mind the expressions (2.41) and (2.42), obtained at Theorem 2.2, the expected result immediately follows.

Finally, let us hint that the possible nullity of the Lamé coefficient μ at Theorem 3.1, is consistent with the classical outcomes in incompressibility according to which, μ is negligible compared with λ.

4. BEHAVIOR OF HYPOELASTIC ALMOST-THERMODYNAMIC MATERIAL SCHEMES IN BORN-RIGID MOTION

The rigid-motion definition in relativity given by several authors is well known. From the one suggested by Born in Special Relativity, that is expressed in analytical form by Synge,[8] to its generalization in General Relativity based in the norm constancy of the separation 4-vector between observers[9] which results equivalent[9] to

$$d = 0, \quad (4.1)$$

where d is the strain rate tensor, for our purposes, in an almost-thermodynamic material scheme and which definition was given by Eq. (2.2).

Authors such as Carter and Quintana[7] or Maugin[10] arrive to the same result; in the former case, after using the class of materially constant spatial tensors (their definition of rigid motion is presented by the material constancy of Eckart tensor) and, in the later one, accepting Eq. (4.1) as definition of Born–Herglotz rigid motion without any more discussion.

Hereafter, we'll adopt the name of Born-rigid almost-thermodynamic material scheme if it satisfies Eq. (4.1). Our purpose is to study the extent to which this definition suffices to lead us to the absence of perturbations as well as the incompressibility condition. In the opposite case, we'll propose a suitable modification.

The answer to the former question is found in next theorem.

Theorem 4.1: In every Born-rigid, almost-thermodynamic material scheme, the shock waves derived from the infinitesimal discontinuities of its 4-velocity covariant derivatives, don't exist.

Proof: According to Eq. (4.1) and taking into account the local expression of Eq. (2.2), we can write

$$\gamma^k{}_i\nabla_k u_j + \gamma^k{}_j\nabla_k u_i = 0, \quad (4.2)$$

which leads us, after applying Hadamard discontinuities, contracting with λ^j, and given the spatial character of the vector δu^i, to

$$(\gamma_{ri} + \lambda_r\lambda_i)\delta u^i = 0. \quad (4.3)$$

The tensor with components

$$P_{ri} = \gamma_{ri} + \lambda_r\lambda_i \quad (4.4)$$

is spatial and symmetric. After checking that

$$P_{\perp ri} = S^s{}_r S^j{}_i\gamma_{sj}, \quad (4.5)$$

$$P_{\|} = 2, \quad (4.6)$$

and

$$p_k = 0, \quad (4.7)$$

the method we have already used in the above proofs, allows us to ensure that

$$\delta u_{\|} = 0; \quad (4.8)$$

moreover, due to the spatial character of $\delta u_{\perp}$,

$$\delta u_{\perp} = 0. \quad (4.9)$$

Hence, the result we looked for follows from Eq. (2.30).

As far as the second point is concerned, keeping in mind the incompressibility condition [Eq. (3.1)], we think that isn't possible, making use of Eq. (4.1) only, to obtain any type of relativistic stress tensor variation—at least when the scheme D doesn't itself present any additional structure.

The close relation that is established, for every Hooke's law, between the strain-rate tensor and the relativistic stress one, will lead us to add these hypotheses in our Born-rigid scheme. The work carried out in Secs. 2 and 3 allows us to use any hypoelastic scheme introduced there.

To begin with, the hypoelastic-Synge material schemes offer us sensible results.

Theorem 4.2: Every hypoelastic-Synge, almost thermodynamic material scheme is incompressible and consistent with the Born-rigidity hypothesis.

Proof: Note that the expression of the Synge–Hooke law, in its general form (2.3), the application of Eq. (4.1), and the Eckart tensor contraction, allow us to write

$$\gamma^k{}_i\gamma^l{}_j\nabla_u t^{ij}=0. \tag{4.10}$$

Hence, the incompressibility condition (3.1), follows from Eq. (2.15), considering the structure of (3.1) and given that, in Born-rigid motion, Eqs. (2.12) and (2.14) reduce to

$$\nabla_u\epsilon=0, \tag{4.11}$$

$$\nabla_u r=0, \tag{4.12}$$

and, furthermore,

$$\gamma^k{}_i\gamma^l{}_j\nabla_u\gamma^{ij}=0, \tag{4.13}$$

as may be checked easily.

On the other hand, if we take Hadamard discontinuities in Eq. (4.10) and consider Eq. (2.20), contract with λ^j leads us to

$$rU^2f_{ri}\delta u^i=0. \tag{4.14}$$

Now, by working in a local inertial proper system at each point[11] $p\in D$ and due to the spatial character of f_{ri} and δu^i, besides the inequality

$$r(1+\epsilon)+t_\alpha\neq 0 \tag{4.15}$$

that is verified by virtue of the material scheme 4-velocity uniqueness, we can write either

$$\delta u^i=0 \tag{4.16}$$

or

$$U=0, \tag{4.17}$$

which allows us to obtain the same results we arrived in Theorem 4.1 and consequently, to establish the consistence we claimed.

As far as the hypoelastic-Maugin and hypoelastic-Carter and Quintana material schemes are concerned, provided that they satisfy the Born-rigidity hypothesis, they lead us to a common expression that we get in next proposition.

Proposition 4.3: The application of the Born-rigidity condition (4.1) to the Maugin and Carter and Quintana–Hooke laws, gives us, in both cases,

$$\gamma^k{}_i\gamma^l{}_j\nabla_u t_{kl}+t^a{}_i\Omega_{aj}+t^a{}_j\Omega_{ai}=0, \tag{4.18}$$

where Ω_{ij} is the rotation tensor[7,10] defined by the tensorial skew-symmetrization of

$$e_{ij}=\gamma^k{}_i\gamma^l{}_j\nabla_k u_l. \tag{4.19}$$

Proof: For the hypoelastic-Maugin almost-thermodynamic material schemes, from using Eq. (4.1) in Eq. (2.6), one derives

$$\gamma^k{}_i\gamma^l{}_j\pounds_u t_{kl}=0. \tag{4.20}$$

If we develop the Lie derivative and take into account the Ehlers decomposition,[7,10] expression is

$$\nabla_j u_i=d_{ij}+\Omega_{ij}-u_j\nabla_u u_i, \tag{4.21}$$

reduced to the Born-rigid material schemes, we obtain easily Eq. (4.18) by virtue of the spatial character of Ω_{ij}, γ_{ij} and the symmetry of the relativistic stress tensor t_{ij}.

If the almost-thermodynamic material scheme D is elastic-Carter and Quintana, Born-rigid, from Eq. (2.8), one gets

$$\gamma^k{}_i\gamma^l{}_j\pounds_u t^{ij}=0, \tag{4.22}$$

and the same argument we used above leads us, again, to Eq. (4.18), thanks to the Ω skew-symmetry.

The expression (4.18) we have obtained in Proposition 4.3, does not lead to the incompressibility condition. However, we'll calculate the respective principal shock wave speeds in order to study whether the hypoelastic-Carter and Quintana and hypoelastic-Maugin almost-thermodynamic material schemes are compatible with Born-rigid motions.

We get the following results:

Proposition 4.4: In a hypoelastic-Carter and Quintana or hypoelastic-Maugin, Born-rigid almost-thermodynamic scheme, the longitudinal and transverse principal shock wave speeds are given by

$$U_L^2=\frac{2t_\parallel}{r(1+\epsilon)+t_\parallel}, \tag{4.23}$$

$$U_{T_\alpha}^2=\frac{t_\alpha}{r(1+\epsilon)+t_\alpha},\quad \alpha{:}2,3. \tag{4.24}$$

Proof: It is enough to write Eq. (4.18), using Eq. (4.21), with $d=0$ to eliminate the rotation tensor, to contract with λ^i, and to apply Hadamard discontinuities in order to obtain, in a way similar to the proof of Theorem 2.2,

$$\bar P_{ri}\delta u^i=0, \tag{4.25}$$

with

$$\bar P_{ri}=-rU^2f_{ri}+t_{ij}\lambda^j\lambda_r+t_{ri}, \tag{4.26}$$

according the expression[1]

$$-g(u,l)=|L|U. \tag{4.27}$$

From here, in principal shock waves

$$\bar P_{ri}=-rU^2f_{ri}+t_\parallel\lambda_r\lambda_i+t_{ri} \tag{4.28}$$

is a spatial and symmetric tensor with $\bar p_i=0$; consequently, one gets

$$\bar P_{\perp ri}\delta u^i_\perp=0 \tag{4.29}$$

and

$$\bar P_\parallel\delta u_\parallel=0, \tag{4.30}$$

where

$$\bar P_{\perp ri}=S^s{}_rS^j{}_i(-rU^2f_{sj}+t_{sj}) \tag{4.31}$$

and

$$\bar P_\parallel=-U^2(r(1+\epsilon)+t_\parallel)+2t_\parallel. \tag{4.32}$$

Finally, for longitudinal shock waves, $\bar P_\parallel=0$ and we obtain Eq. (4.23) from Eq. (4.32), whereas for the transverse ones, the method we carried in the proofs of Theorems 2.1 and 2.2, leads us to Eq. (4.24), from Eqs. (4.29) and (4.31).

In view of the results we have obtained in hypoelastic-Synge material schemes, the present situation is somewhat strange since, at first sight, the hypoelastic-Maugin and the hypoelastic-Carter and Quintana material schemes aren't consistent with the Born-rigidity hypothesis and, as a matter of fact, they don't lead us to the relativistic incompressibility condition (3.1).

At present, we face two possible explanations for the anomalies observes: Either the application of the Born-rigidity hypothesis in Propositions 4.3 and 4.4 hasn't been fully carried out, or this definition is by itself insufficient, even in the presence of hypoelastic schemes, for our purposes. In this direction, we'll move into the next section.

5. RIGID ALMOST-THERMODYNAMIC MATERIAL SCHEMES. A DEFINITION AND SOME CONSEQUENCES

Following Sec. 4, we propose to give a definition that will solve the paradox that appeared in the hypoelastic Carter and Quintana and hypoelastic-Maugin almost-thermodynamic material schemes.

Definition 5.1: An almost-thermodynamic material scheme D in the space-time manifold M is called rigid if it verifies the two following conditions: (i) D is Born-rigid; i.e., satisfies Eq. (4.1). (ii) D fulfills Eq. (4.10).

This definition doesn't present internal contradictions since the results in Theorem 4.2 show us the consistence of (i) and (ii).

In the hypoelastic-Maugin and hypoelastic-Carter and Quintana almost-thermodynamic material schemes we can prove that application of Definition 5.1 to Eq. (4.18) eliminates the problem which arose after the results obtained in Proposition 4.4 and allows us to check the coherence between the results we derived in a general case from Eq. (4.10) and the ones given in Proposition 4.3 in the present case.

Theorem 5.2: The application of condition (4.10) to Eq. (4.18) we obtained in Proposition 4.3 for hypoelastic-Maugin and hypoelastic-Carter and Quintana almost-thermodynamic material schemes, leads to the absence of principal shock waves due to infinitesimal discontinuities of their 4-velocity.

Proof: Direct use of Eq. (4.10), by virtue of Eq. (4.18), allows us to write

$$t^a{}_i\Omega_{aj} + t^a{}_j\Omega_{ai} = 0. \tag{5.1}$$

If we consider Eq. (4.21) besides Eq. (4.1), take Hadamard discontinuities in the resulting expression, and contract with λ^i, λ^j, we obtain straightforwardly

$$\delta u_{\shortparallel} = 0. \tag{5.2}$$

On the other hand, by contracting with λ^i and after using Eqs. (2.45) and (4.27), it is possible to get, for principal shock waves,

$$(t_{\|}\lambda_r\lambda_i + t_{ri})\delta u^i_{\perp} = 0, \tag{5.3}$$

where the tensor in brackets is spatial and symmetric. The method we used in the proof of Theorem 2.1, leads to

$$S^s{}_r S^j{}_i(t_{\|}\lambda_s\lambda_j + t_{ri})\delta u^i_{\perp} = 0. \tag{5.4}$$

The spatial character of $\delta u_{\perp}$ and latter contraction with d^r_α (α:2,3) gives us, from Eq. (2.35),

$$a_\alpha t_\alpha = 0 \quad \alpha\text{:}2,3, \tag{5.5}$$

which implies

$$a_2 = a_3 = 0 \tag{5.6}$$

and, from Eq. (2.37), we obtain

$$\delta u_{\perp} = 0. \tag{5.7}$$

Equations (5.2) and (5.7) lead to the result claimed.

We want to indicate here that if the scheme D hasn't any additional structure, as maybe a Hooke's law, Definition 5.1 shows us that every rigid scheme is, obviously, Born-rigid, but also incompressible according the process followed in the proof of Theorem 4.2 from Eq. (4.10). Hence, the rigidity concept given here avoids the insufficiencies we block out in Sec. 4 for the Born-rigidity when the scheme D isn't hypoelastic or, in the general case, when it hasn't any additional structure that includes the relativistic stress tensor variation.

After proving Theorem 5.2, is interesting to study physically the meaning of the relation (5.1), derived from Definition 5.1 in hypoelastic-Maugin and hypoelastic-Carter and Quintana almost-thermodynamic material schemes.

Let us consider the tensor A with covariant components

$$A_{ij} = t^a{}_i\Omega_{aj} + t^a{}_j\Omega_{ai} \tag{5.8}$$

that is spatial since the rotation tensor and the relativistic stress tensor are spatial. Consequently, as is known, the vanishing of A_{ij} at a point is equivalent to the verification of

$$A(X,Y) = 0 \tag{5.9}$$

for all vectors X, Y in the physical space (hypersurface orthogonal to the 4-velocity u in D) at that point.

We define n_X, n_Y as the unit vectors parallel to X and Y, respectively.

Hence,

$$A(X,Y) = |X|\,|Y|\,A(n_X, n_Y) \tag{5.10}$$

and locally, by virtue of Eq. (5.8),

$$A(X,Y) = |X|\,|Y|\,(t^a{}_i n^i_X \Omega_{aj} n^j_Y + t^a{}_j n^j_Y \Omega_{ai} n^i_X). \tag{5.11}$$

Given that contraction of the relativistic stress tensor with a vector is a stress in its parallel direction, if we denote by t_X and t_Y the stress vectors in the respective directions of n_X and n_Y, Eq. (5.11) can be written

$$A(X,Y) = |X|\, t^a_X \Omega_{aj}(Y^j) + |Y|\, t^a_Y \Omega_{ai}(X^i). \tag{5.12}$$

Note that the rotation tensor can be expressed in the form

$$\Omega_{ai} = -{}^3\eta_{ajk}w^k, \tag{5.13}$$

in which w is the vorticity 4-vector defined[10] by

$$w^i = \tfrac{1}{2}\,{}^3\eta^{ijk}\, e_{kj}, \tag{5.14}$$

where e_{kj} is given by Eq. (4.19) and ${}^3\eta^{ijk}$ is the spatial volume element

$${}^3\eta = i_u\eta, \tag{5.15}$$

i.e., ${}^3\eta$ is the tensor contraction (inner product) with respect to the scheme 4-velocity u, of the volume element in the manifold space-time M.

Consequently, applying Eq. (5.13) in Eq. (5.12), we write

$$A(X,Y) = -|X|\, t^a_X\, {}^3\eta_{ajk}\, w^k\, Y^j - |Y|\, t^a_Y\, {}^3\eta_{aik}\, w^k\, X^i. \tag{5.16}$$

Moreover, the term

$$-{}^3\eta_{ajk}\, w^k\, Y^j \tag{5.17}$$

denotes a vertical product in the physical space of the 4-vectors w and Y. The same may be asserted, from

$$- {}^3\eta_{aik}\, w^k X^i, \tag{5.18}$$

for w and X. Correspondingly, it must be noted that, in a local inertial proper system, one gets

$$w^4 = X^4 = Y^4 = 0 \tag{5.19}$$

and

$${}^3\eta_{4ij} = 0. \tag{5.20}$$

Lastly, Eq. (5.16) yields

$$A(X,Y) = |X|\, t^a_X\, (w \wedge Y)_a + |Y|\, t^a_Y\, (w \wedge X)_a, \tag{5.21}$$

where the vectorial products are linear velocities and their contractions with the stress lead us to a power due to the internal rotation. This power vanishes from Eq. (5.1).

In order to conclude this section, we'll analyze some consequences of the Born-rigidity (4.1) and irrotationality hypotheses, the latter[12] given by

$$\Omega = 0, \tag{5.22}$$

considered together in our almost-thermodynamic material scheme D.

We start by saying that the irrotational hypoelastic-Maugin and hypoelastic-Carter and Quintana almost-thermodynamic material schemes, will fulfill Definition 5.1, as may be seen from results in Proposition 4.3. It is obvious, also, that Eq. (5.1) is verified, with the physical meaning we have given it above.

A more detailed analysis has led us to propose a theorem which gives a condition for the orthogonality of the Landau manifolds[13] L_p $(p\in D)$ to the scheme 4-velocity u, in every point $q \in L_p$, $p\in D$, provided that Eqs. (4.1) and (5.22) are satisfied. This theorem can be stated as follows.

Theorem 5.3: The 4-velocity u of a material scheme $D \subset M$ moves parallelwise and orthogonally in every Landau manifold L_p $(p\in D)$ if and only if the rotation tensor and the strain rate tensor vanish in D.

Proof: To start, we'll assume that u moves parallel and orthogonally to L_p $(p\in D)$. Let v be an arbitrary vector in $T_p\, L_p$. The former condition of the hypothesis allows us to write

$$\nabla_v\, u = 0. \tag{5.23}$$

If now use the Ehlers decomposition (4.21), contraction with v^j leads us to

$$d_{ij}\, v^j + \Omega_{ij}\, v^j - u_j\, v^j\, \nabla_u\, u_i = 0, \tag{5.24}$$

from Eq. (5.23).

The term $u_j\, v^j$ is null according to the Landau manifold properties as they are presented by Olivert.[13] Hence, Eq. (5.24) reduces to

$$(d_{ij} + \Omega_{ij})\, v^j = 0 \tag{5.25}$$

for every $v \in T_p\, L_p$, and $p\in D$ arbitrary. The spatial character of d_{ij} and Ω_{ij} gives, after contracting Eq. (5.25) with arbitrary $\bar{v} \in T_p\, L_p$,

$$d_{ij} + \Omega_{ij} = 0, \quad \forall\, p\in D \tag{5.26}$$

and, from this, we infer the vanishing of both tensors in D if we keep in mind that the strain rate tensor d is symmetric and the rotation tensor is skew-symmetric.

In the preceding reasoning we have used implicitly the existence, for every $p\in D$, of a Landau manifold L_p orthogonal to u at p, which is proved in Ref. 13. We also note that it hasn't been necessary to utilize the orthogonality of u to L_p in every $q \in L_p$.

In order to prove the remaining condition, we'll assume that the rotation and the strain rate tensors are null in D. Let us consider the Pfaff system[11]:

$$u_i\, dx^i = 0, \tag{5.27}$$

constituted by an only linear 1-form and, consequently of rank 1. If we recall the integrability criterion for this type of systems given in Choquet-Bruhat,[11] which uses the differential systems integrability condition (Frobenius) as far as may be seen in Sternberg,[14] we see that the system represented by Eq. (5.27) will be fully integrable if and only if,

$$\theta \wedge d\theta = 0, \tag{5.28}$$

where θ is the 1-form $u_i\, dx^i$.

If we develop the left-hand side of Eq. (5.28), write the outcome in strict components, and take into account that from Eq. (4.21), in the irrotational case,

$$\nabla_i u_j - \nabla_j\, u_i = \partial_i u_j - \partial_j\, u_i = u_j\, \nabla_u u_i - u_i \nabla_u\, u_j \tag{5.29}$$

can be derived due to the symmetry of the strain rate tensor and of the Christoffel symbols Γ^k_{ij} in the i,j indices, since the connection ∇ has null torsion, we easily check Eq. (5.28). Note that the vanishing of d hasn't been used in this step.

Thus, the irrotationality hypothesis is enough to find, for every $p\in D$, an integral manifold of the system given by Eq. (5.27) which contains p. This manifold will be a 3-dimensional submanifold of D, orthogonal to the 4-velocity u in it, at every point, as may be seen from the definition of integral manifold of the exterior differential systems[11] and from the structure of Eq. (5.27).

Now, we propose to apply the Born-rigidity hypothesis (4.1) to find the results we sought.

Consider a local inertial proper system at a point $p \in N$, where N is one of those integral manifolds. Then Eq. (5.27) reduces to

$$dx^4 = 0. \tag{5.30}$$

Therefore, the points of N are simultaneous in a local inertial proper system at p. Since $T_p N$ is the physical space at p by virtue of the orthogonality of the 4-velocity u to N at every point $q \in N$, the theorem of existence and uniqueness of Landau manifolds[13] leads to

$$N = L_p, \tag{5.31}$$

so u is orthogonal to L_p at arbitrary $q \in L_p$.

Moreover, by virtue of Eq. (4.21), with $d = \Omega = 0$, and for $v \in T_q L_p$, $q \in L_p$,

$$\nabla_v\, u = -(\nabla_u\, u)\, g(u,v) = 0 \tag{5.32}$$

after using the orthogonality of u to L_p. This result concludes the proof.

We'll now prove, to finish this section, a corollary of Theorem 5.3 that at present seems to be too restrictive, given its strong hypotheses, but leads to a sufficient condition for a material scheme to be Born-rigid and irrotational.

Corollary 5.4: If the 4-velocity u of a material scheme D is a Killing vector orthogonal to L_p for every $p\in D$, the rotation tensor and the strain rate tensor vanish in D.

Proof: Let us recall the fact that the points $q \in L_p$, $p \in D$ verify[13]

$$g\,(u_p, \exp_p^{-1}(q)) = 0, \tag{5.33}$$

where $\exp_p$ is the diffeomorphism exponential application[15]

$$\exp_p : M_0 \to B_p, \tag{5.34}$$

with M_0 and B_p neighborhoods of zero and p in $T_p M$ and M, respectively. Equation (5.33) is expressed, in a local inertial proper system at p, by

$$u_{1p}\, x^1 + u_{2p}\, x^2 + u_{3p}\, x^3 + u_{4p}\, x^4 = 0 \tag{5.35}$$

for $x = \exp_p^{-1}(q)$. Consequently, the orthogonal vector to L_p will be, in every point, u_{i_p}. Thus,

$$u_{i_p} = u_{i_q} \quad \forall q \in L_p \tag{5.36}$$

by virtue of the orthogonality hypothesis of u to L_p.

From Eq. (5.36), follows

$$v^i \partial_i\, u_j = 0, \tag{5.37}$$

i.e., for every vector $v \in T_q M$,

$$\nabla_v u_i = v^j\, \nabla_j\, u_i = -\, v^j\, \Gamma^k_{ji}\, u_k. \tag{5.38}$$

Morevoer, the Killing hypothesis for u

$$\nabla_i\, u_j + \nabla_j u_i = 0 \tag{5.39}$$

yields, after applying Eq. (5.38) and by the symmetry of the Γ^k_{ij} in the i, j indices,

$$\Gamma^k_{ij}\, u_k = 0. \tag{5.40}$$

If v is restricted to be tangent to L_p, one obtains, from Eqs. (5.38) and (5.40), the parallel-moving condition of u in the Landau manifolds.

Hence, use of Eq. (4.21) and orthogonality hypothesis of u to L_p lead us, using the same method we followed in the first part of the proof in Theorem 5.3, to Eq. (5.25) and thus, to the expected result.

6. DISCUSSION

In this paper we have restricted the definition of Born-rigidity and a concept has been proposed which encompasses it and is coherent, on the one hand, with classical requirements about the absence of internal perturbations in almost-thermodynamic material schemes and, on the other, verifies the relativistic incompressibility condition.[1]

After analyzing some types of hypoelastic schemes, it has been checked that the hypoelastic-Synge, almost-thermodynamic material schemes fulfill those demands, while the hypoelastic-Carter and Quintana and hypoelastic-Maugin ones have given us unsatisfactory results. In order to find a fitting explanation, we have added an extra condition, Eq. (4.10), to the Born-rigidity one. This condition has led us directly to Eq. (5.1)

$$t^a{}_i \Omega_{aj} + t^a{}_j \Omega_{ai} = 0$$

in the hypoelastic-Carter and Quintana and hypoelastic-Maugin Born-rigid cases. Equation (5.1) has also been interpreted physically in Sec. 5 as the nullity of mechanical power due to internal rotation.

We think that (5.1) must be true for every rigid-Born almost-thermodynamic material scheme regardless of its internal structure. This assertion is supported by the fact that, upon the Born-rigidity condition, both the matter density and the specifical internal energy are conserved across the stream lines, as may be seen from Eqs. (4.11) and (4.12). Consequently, the internal energy density

$$\mathrm{E} = r\epsilon \tag{6.1}$$

is also conserved across the scheme stream lines. A more detailed study in this direction may be the subject of later work.

Anyway, if Eq. (5.1) is admitted as the only hypothesis upon the Born rigidity, it can be checked that, by virtue of results in Proposition 4.3, the hypoelastic-Carter and Quintana and hypoelastic-Maugin material schemes verify rigidity Definition 5.1. Nevertheless, the rigidity definition must be retained in the general case since it ensures the incompressibility condition in almost-thermodynamic material schemes even without elastic restrictions.

The hypotheses of Born-rigidity and irrotationality in a material scheme D have been also analyzed altogether. Thus, we have proved Theorem 5.3, which relates kinematical and geometrical concepts. This result seems to us more complete that the ones which can be found in the relativistic literature; for instance, Synge[12] states the existence, upon irrotationality conditions, of 3-dimensional submanifolds of the space-time M, orthogonal to u.

On the other hand, there is the definition[3] of 4-vector locally synchronizable, which coincides with the integrability condition (5.28) relative to the Pfaff system (5.27). For our purposes, this concept is interpreted by saying that the points of every Landau manifold are simultaneous, which isn't evident since the simultaneity condition in each L_p refers to an observer, and his motion changes it.

ACKNOWLEDGMENT

We would like to thank Professor J. Bernabeu for his care in reading the manuscript and suggestions about its physical content.

[1] J. J. Ferrando and J. Olivert, "Incompressibility in relativistic continuous media," J. Math Phys. **22**, 2223 (1981).
[2] A. Lichnerowicz, *Théories rélativistes de la gravitation et l'électromagnétisme* (Masson, Paris, 1955).
[3] R. K. Sachs and H. Wu, *General Relativity for Mathematicians* (Springer-Verlag, New York, 1977).
[4] A. H. Taub, in *Hydrodynamics and general relativity*, edited by R. Wassermann and Ch. P. Wells (Academic, New York, 1963).
[5] J. L. Synge, "A theory of elasticity in general relativity," Math. Z. **72**, 82 (1959).
[6] G. A. Maugin, "Ondes infinitésimales principales dans les milieux élastiques précontraints in relativité," C. R. Acad. Sci. Paris A **284**, 393 (1977).
[7] B. Carter and H. Quintana, "Foundations of general relativistic high-pressure elasticity theory," Proc. R. Soc. London, Ser. A **331**, 57 (1972).
[8] J. L. Synge, *Relativity: The Special Theory* (North-Holland, Amsterdam, 1965).
[9] A. P. Lightman, W. H. Press, R. H. Price, and S. A. Teukolsky, *Problem book in relativity and gravitation*(Princeton U. P., Princeton, 1975).
[10] G. A. Maugin, "On the covariant equations of the relativistic electrodynamics of continua. I. General equations," J. Math Phys. **19**, 1198 (1978).
[11] Y. Choquet-Bruhat, *Geométrie differéntielle et systèmes extérieurs* (Dunod, Paris, 1968).

[12] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1966).
[13] J. Olivert, "On the local simultaneity in general relativity," J. Math. Phys. **21**, 1783 (1980).
[14] S. Sternberg, *Lectures on differential geometry* (Prentice Hall, Englewood Cliffs, NJ, 1964).
[15] F. Brickell and R. S. Clark, *Differentiable Manifolds* (Van Nostrand, New Jersey, 1970).

Rotating hollow cylinders: General solution and Machian effects

Franz Embacher
Institut für Theoretische Physik der Universität Wien, Boltzmanngasse 5, 1090 Wien, Austria

(Received 18 November 1981; accepted for publication 13 May 1982)

The general solution of the exact gravitational field of a rotating hollow cylinder is derived. It is more general than the solution found by E. Frehland. Machian effects are considered.

PACS numbers: 04.20.Jb

1. INTRODUCTION

In 1972, Frehland published a paper[1] concerning the exact gravitational field of an infinitely long rotating hollow cylinder.[2] After a fitting procedure of interior and exterior vacuum metrics he obtained a solution dependent on three freely available parameters (surface density of matter, rotational velocity, and radius). These solutions satisfy

(i) stress in radial direction $= 0$,

(ii) $T^{\mu}{}_{\mu} = 0$.

Frehland's solution is not the general one. (i) is an immediate consequence of the field equations (for every infinitely thin shell), but (ii) is not. Integrating the field equations in a vacuum domain, one obtains a metric with only one physical parameter (a kind of mass enclosed in the interior); all other integration constants can be eliminated by coordinate transformations (Kasner metric[3,4]). However, for the fitting problem of the infinitely thin hollow cylinder it is necessary to use the general form of the vacuum metric as there are more integration constants related to the matter variables. Simply spoken, Frehland has taken too special a form of the exterior metric.

In this paper I will derive the general solution, depending on four free parameters (e.g., matter density, velocity, radius, and stress in the direction parallel to the cylinder axis). In order to avoid the usual difficulties of the identification of interior and exterior coordinates, I will use a coordinate system defined uniquely *a priori*. The metric components will be continuous.

2. COORDINATE SYSTEM, FIELD EQUATIONS, AND VACUUM METRIC

Properties and physical justification of the coordinate system used have been discussed in another paper[5]; here I shall only list the most important results. Under weak assumptions there exists a coordinate system $(x^0, x^1, x^2, x^3) := (t, \rho, z, \phi)$ with the following properties:

$$\begin{aligned} ds^2 &= -A^2(\rho)\,dt^2 + B^2(\rho)(d\rho^2 + dz^2) \\ &\quad + E^2(\rho)(d\phi - \Omega(\rho)\,dt)^2, \\ &A(0) = B(0) = E_{,\rho}(0) = 1, \\ &A_{,\rho}(0) = B_{,\rho} = E(0) = 0, \\ &\Omega(0) \text{ finite}, \quad \Omega_{,\rho}(0) \text{ finite}, \\ &\lim_{\rho\to\infty} \Omega(\rho) = 0, \\ &\rho \geqslant 0, \ 0 \leqslant \phi < 2\pi; \quad t, z \in \mathbb{R}. \end{aligned} \tag{1}$$

Up to transformations $t \to \pm t + \text{const}$, $z \to \pm z + \text{const}$, $\phi \to \pm \phi + \text{const}$, this coordinate system (i.e., the functions x^{μ} from the manifold into $\mathbb{R}$) is unique.[6] It becomes singular on the axis like usual cylindrical coordinates in $\mathbb{R}^3$.

The matter variables are defined by $T^{\mu}{}_{\nu}$ (satisfying convenient energy conditions)

$$\begin{aligned} &T^{\mu}{}_{\nu} = \rho_M\, u^{\mu}u_{\nu} + p_{\rho}\, \delta^{\mu}_1\, \delta^1_{\nu} + p_z\, \delta^{\mu}_2\, \delta^2_{\nu} + p_{\phi} w^{\mu} w_{\nu}, \\ &u^{\mu} = u^0(1,0,0,\Lambda), \quad \Lambda = \left.\frac{d\phi}{dt}\right|_{\text{matter}}, \\ &w^{\mu} = (w^0,0,0,w^3), \\ &u^{\mu}w_{\mu} = 0, \quad u^{\mu}u_{\mu} = -w^{\mu}w_{\mu} = -1. \end{aligned} \tag{2}$$

In the orthonormal basis

$$\hat{e}^{\mu} := (A\,dt, B\,d\rho, B\,dz, E\,(d\phi - \Omega\,dt))$$

we write (tensor components denoted with $\hat{\ }$)

$$\hat{u}^{\mu} = (1 - v^2)^{-1/2}(1,0,0,v), \quad v \in (-1,1). \tag{3}$$

The field equations read[5,7]

$$\begin{aligned} 8\pi B^2 \hat{T}^{00} &= -(B^{-1}B_{,\rho})_{,\rho} - E^{-1}E_{,\rho\rho} \\ &\quad - \tfrac{1}{4}(A^{-1}E\Omega_{,\rho})^2, \\ 8\pi B^2 \hat{T}^{11} &= B^{-1}B_{,\rho}(A^{-1}A_{,\rho} + E^{-1}E_{,\rho}) \\ &\quad + (AE)^{-1}A_{,\rho}E_{,\rho} + \tfrac{1}{4}(A^{-1}E\Omega_{,\rho})^2, \\ 8\pi B^2 \hat{T}^{22} &= -B^{-1}B_{,\rho}(A^{-1}A_{,\rho} + E^{-1}E_{,\rho}) + E^{-1}E_{,\rho\rho} \\ &\quad + (AE)^{-1}A_{,\rho}E_{,\rho} + A^{-1}A_{,\rho\rho} - \tfrac{1}{4}(A^{-1}E\Omega_{,\rho})^2, \\ 8\pi B^2 \hat{T}^{33} &= (B^{-1}B_{,\rho})_{,\rho} + A^{-1}A_{,\rho\rho} - \tfrac{3}{4}(A^{-1}E\Omega_{,\rho})^2, \\ 8\pi B^2 \hat{T}^{03} &= -\tfrac{1}{2}E^{-2}(A^{-1}E^3\Omega_{,\rho})_{,\rho}. \end{aligned} \tag{4}$$

Written in the natural basis,[5] they can be integrated easily within a domain where $T^{\mu}{}_{\nu} = 0$. In such a domain, we obtain the general vacuum solution (equivalent to the Kasner metric,[4] α is the only physical parameter, the other constants can be eliminated by coordinate transformations[6])

$$ds^2 = -\left(\frac{\rho-\rho_{AE}}{\rho_A}\right)^{2\alpha}\left\{1-\left(\frac{\rho-\rho_{AE}}{\rho_S}\right)^{4\alpha-2}\right\}^{-1} dt^2 + \left(\frac{\rho-\rho_{AE}}{\rho_B}\right)^{-2\alpha(1-\alpha)}(d\rho^2+dz^2)$$
$$+(\rho-\rho_{AE})^2\left(\frac{\rho-\rho_{AE}}{\rho_E}\right)^{-2\alpha}\left\{1-\left(\frac{\rho-\rho_{AE}}{\rho_S}\right)^{4\alpha-2}\right\}$$
$$\times\left(d\phi - \left[\sigma\frac{\left(\frac{\rho-\rho_{AE}}{\rho_A}\right)^{\alpha}\left(\frac{\rho-\rho_{AE}}{\rho_E}\right)^{\alpha}\left(\frac{\rho-\rho_{AE}}{\rho_S}\right)^{2\alpha-1}}{(\rho-\rho_{AE})\left\{1-\left(\frac{\rho-\rho_{AE}}{\rho_S}\right)^{4\alpha-2}\right\}}+h\right]dt\right)^2, \tag{5}$$

$\alpha, \rho_{AE}, \rho_A, \rho_E, \rho_B, \rho_S$ and h are integration constants, $\sigma = \pm 1$. The most general solution allows complex α. Then the metric coefficients may be transformed into real functions by an appropriate (formally complex) coordinate transformation that leads to terms behaving like $\sin \ln \rho$. These solutions have been found by van Stockum.[8] As Tipler pointed out,[9] they produce causality violation. Furthermore, when used as exterior metric, they cannot be expressed in a coordinate system (1) because $\lim_{\rho\to\infty} \Omega(\rho) = 0$ cannot be achieved. Thus we will restrict ourselves to the solutions with real α. We assume $0<\alpha<\frac{1}{2}$, which is a necessary condition for the existence of closed geodesic orbits (which should be given at least in the exterior field). In the exterior domain we have to set $h=0$ due to condition (1), $\lim_{\rho\to\infty}\Omega(\rho)=0$.

As a consequence of our special choice of the coordinate system [($\int_0^\rho B(\rho)\,d\rho$ is the proper distance from the axis $\rho=0$], the metric components are continuous. The discontinuities of their derivatives give the surface energy momentum tensor of infinitely thin matter shells.

3. THE GENERAL SOLUTION FOR THE ROTATING HOLLOW CYLINDER

We define

$$\hat{T}^{\mu\nu} = \hat{\tau}^{\mu\nu}\delta(\rho - R)$$

(R = radius of the hollow cylinder). The region $0\leqslant\rho<R$ is denoted by a subscript $-$, the region $\rho>R$ by a subscript $+$. ds^2_- and ds^2_+ are vacuum metrics, so they have the form (5). Energy density and stresses are given by

$$\mathbf{\rho}_M = Q_M\,\delta(\rho-R),$$
$$\mathbf{p}_i = P_i\,\delta(\rho-R), \quad i=\rho,z,\phi.$$

$\hat{u}^\mu$ is defined only for $\rho=R$, v is given by (3). Without loss of generality, we set $v>0$ and

$$\epsilon := (1-v^2)^{-1} > 1,$$
$$\hat{u}^\mu = (\epsilon^{1/2},0,0,(\epsilon-1)^{1/2}).$$

Consider first the interior metric. The conditions (1) for $\rho=0$, applied to (5), imply $\alpha_- = \rho_{AE_-} = \rho_{S_-} = 0$, $\Omega_-(\rho) = h_- =: \Omega_- =$ const. The interior metric is flat; in the coordinate system (1) the interior global inertial frame rotates with angular velocity Ω_-:

$$ds^2_- = -dt^2 + d\rho^2 + dz^2 + \rho^2(d\phi - \Omega_-\,dt)^2.$$

This result is well known.[1] Now R and v have physical meaning: R is the proper radius of the hollow cylinder (i.e., measured by an observer in the interior region), and with

$$\bar{\phi} := \phi - \Omega_- t$$

and

$$\left.\frac{d\bar{\phi}}{dt}\right|_{\text{matter}} = \Lambda - \Omega_- = R^{-1}v$$

we see that v is the matter velocity, measured by the same observer.

Now consider the exterior metric. The conditions (1) imply $h_+ = 0$; the other constants are to be calculated from the fitting procedure. $\sigma_+ = 1$ is a consequence of $v>0$. From now on the subscript $+$ of the constants is suppressed. We define

$$\mu := R - \rho_{AE}, \quad x := \left(\frac{\mu}{\rho_S}\right)^{4\alpha-2},$$
$$y_A := \left(\frac{\mu}{\rho_A}\right)^{\alpha}, \quad y_E := \left(\frac{\mu}{\rho_E}\right)^{\alpha}.$$

The continuity conditions for the metric components become

$$\begin{aligned} y_A &= (1-x)^{1/2},\\ y_E &= R^{-1}\mu(1-x)^{1/2},\\ \rho_B &= \mu,\\ y_A\, y_E\, x^{1/2}\mu^{-1}(1-x)^{-1} &= \Omega_- \end{aligned} \tag{6}$$

($\Omega_- > 0$ is a consequence of $v>0$).

Applying the operation $\lim_{\Delta\downarrow 0}\int_{R-\Delta}^{R+\Delta}$ on the field equations, one obtains

$$\begin{aligned} &8\pi[Q_M\,\epsilon + P_\phi(\epsilon-1)]\\ &\quad = R^{-1} - \mu^{-1}(1-\alpha)^2 - \mu^{-1}(1-x)^{-1}x(1-2\alpha),\\ &8\pi[Q_M(\epsilon-1) + P_\phi\,\epsilon]\\ &\quad = \mu^{-1}\alpha^2 - \mu^{-1}(1-x)^{-1}x(1-2\alpha), \end{aligned} \tag{7}$$

$P_\rho = 0$ (no stress in ρ direction),

$$8\pi(Q_M + P_\phi)\,\epsilon^{1/2}(\epsilon-1)^{1/2} = \mu^{-1}(1-x)^{-1}x^{1/2}(1-2\alpha),$$
$$8\pi P_z = \mu^{-1} - R^{-1}.$$

Now we have to specify the independent variables. We define

$$\kappa := R^{-1}\mu = 1 - R^{-1}\rho_{AE} = y_A^{-1}y_E = (1+8\pi R P_z)^{-1}$$
$$= \left(1 + 8\pi\int_0^\infty \mathbf{p}_z(-g)^{1/2}\,d\rho\right)^{-1} > 0. \tag{8}$$

As a consequence of (7) we obtain

$$\alpha(1-\alpha) = 4\pi(Q_M - P_\phi + P_z)\mu$$
$$= 4\pi\kappa^{-1}\int_0^\infty (\rho_M - p_\phi + p_z)(-g)^{1/2}\,d\rho. \qquad (9)$$

Equations (8) and (9) give the physical meaning of α and κ. α is something like an energy; it can be measured from geodesic orbits[6]:

$$\left(\frac{\text{proper length of orbit}}{\text{proper time of orbit}}\right)^2 = \alpha(1-2\alpha)^{-1}.$$

κ is related to the stress in z-direction (it can be measured by the difference of gyroscope dragging in the two regions near the matter shell).

We use $\alpha \in (0,\frac{1}{2})$, $\kappa > 0$, $\epsilon > 1$, and $R > 0$ as independent variables. The other quantities can be expressed by them. For each choice of these parameters we obtain a singularity-free solution (singularity-free means, no horizon outside R, the functions $g_{\mu\nu}(\rho)$ are regular for $\rho > R$). The results of the calculations shall be listed here.

Two defined quantities:

$$a(\alpha,\kappa,\epsilon) := 2\kappa(\epsilon - \tfrac{1}{2}) - [(1-2\alpha)^2 + 4\kappa^2\epsilon(\epsilon-1)]^{1/2}$$
$$= 8\pi\kappa\int_0^\infty (\rho_M + p_\phi)(-g)^{1/2}\,d\rho,$$
$$b(\alpha,\kappa,\epsilon) := (1-2\alpha)\,[\tfrac{1}{2}(\kappa-1) + \alpha - (\epsilon - \tfrac{1}{2})\,a(\alpha,\kappa,\epsilon)]^{-1} > 0.$$

Interior metric:

$\Omega_- = R^{-1}(1+b)^{-1/2}$,

(so $g_{00} = -1 + \rho^2\Omega_- < 0$ if $\rho \leqslant R$).

Exterior metric:

$h = 0, \quad \sigma = 1, \quad \rho_B = \kappa R, \quad \rho_{AE} = (1-\kappa)R,$

$\rho_A = \kappa R\,(1 + b^{-1})^{1/2\alpha}$,

$\rho_E = \kappa^{1-(1/\alpha)}R\,(1 + b^{-1})^{1/2\alpha}$,

$\rho_S = \kappa R\,(1+b)^{1/(4\alpha-2)}$.

Energy momentum tensor:

$v = (1-\epsilon^{-1})^{1/2} > 0$,

$\Lambda = \Omega_- + R^{-1}v$
$= R^{-1}[(1+b)^{-1/2} + (1-\epsilon^{-1})^{1/2}]$,

$16\pi Q_M R\kappa = a + \kappa - 1 + 2\alpha(1-\alpha)$,

$16\pi P_\phi R\kappa = a - \kappa + 1 - 2\alpha(1-\alpha)$,

$16\pi P_z R\kappa = 2(1-\kappa)$,

$P_\rho = 0$.

Consequences:

Komar mass[10] (per unit length in z-direction)

$$M := -4\pi\int_0^\infty (T^0_0 - \tfrac{1}{2}T^\mu{}_\mu)(-g)^{1/2}\,d\rho = \tfrac{1}{2}\kappa^{-1}\alpha;$$

angular momentum[10] (per unit length in z-direction)

$$J := 2\pi\int_0^\infty T^0_3\,(-g)^{1/2}\,d\rho$$
$$= \tfrac{1}{4}\kappa^{-1}R\,(1-2\alpha)\,b^{-1}(1+b)^{1/2};$$

usual mass–energy (per unit length in z-direction)

$$-2\pi\int_0^\infty T^0_0(-g)^{1/2}\,d\rho = \tfrac{1}{4}[1 - \kappa^{-1}(1-\alpha)^2];$$

trace of energy momentum tensor

$$4\pi(-Q_M + P_\phi + P_z)\kappa R = 1 - \kappa - \alpha(1-\alpha).$$

Frehland's solution satisfies $T^\mu{}_\mu = 0$; thus $\kappa + \alpha(1-\alpha) = 1$. For large velocities, this implies unphysical mass densities. His conclusions, the impossibility of a very fast ($v \approx 1$) rotating hollow cylinder of matter satisfying $\rho_M > |p_i|$ and the impossibility of rotating dust hollow cylinders, are wrong.

We will restrict the parameters to a domain giving solutions that satisfy

$$Q_M > |P_i|, \quad i = \rho, z, \phi. \qquad (10)$$

For dust solutions ($P_\rho = P_z = P_\phi = 0$), we have to set

$$\kappa = 1,$$
$$\epsilon = \tfrac{1}{2}(1+\alpha)(2-\alpha). \qquad (11)$$

The second equation has a solution $\alpha \in (0,\frac{1}{2})$ if $\epsilon \in (1,\frac{9}{8})$, i.e., $v \in (0,\frac{1}{3})$. R can be chosen arbitrarily. The dust hollow cylinders form a subsolution with two parameters (R and ϵ or R and v or R and α). An infinitely long hollow cylinder of dust can never rotate with velocity (measured in the inertial frame inside) larger than $\frac{1}{3}$ [otherwise the exterior metric will become unphysical and will not satisfy (1) [9]]. King has constructed the most general dust solution; thus the hollow dust cylinders should be included as a special case.[11]

Long calculations give the domain of solutions satisfying (10) [6]:

$$Q_M + P_\phi > 0 \Leftrightarrow \kappa > 1 - 2\alpha \Leftrightarrow a(\alpha,\kappa,\epsilon) < 0,$$
$$Q_M - P_\phi > 0 \Leftrightarrow \kappa > 1 - 2\alpha(1-\alpha),$$
$$Q_M + P_z > 0 \Leftrightarrow \kappa < f(\alpha,\epsilon),$$
$$Q_M - P_z > 0 \Leftrightarrow \kappa > g(\alpha,\epsilon),$$

with

$$f(\alpha,\epsilon) := \tfrac{1}{2}[1 + 2\alpha(1-\alpha)] + (\tfrac{1}{4}[1 + 2\alpha(1-\alpha)]^2 + (\epsilon-1)^{-1}\alpha(1-\alpha)(1+\alpha)(2-\alpha))^{1/2}$$

and

$$g(\alpha,\epsilon) := (3\epsilon+1)^{-1}(\epsilon+1)[\tfrac{3}{2} - \alpha(1-\alpha)] + ((3\epsilon+1)^{-2}(\epsilon+1)^2[\tfrac{3}{2} - \alpha(1-\alpha)]^2 - (3\epsilon+1)^{-1}[1 + \{1 - \alpha(1-\alpha)\}^2])^{1/2}.$$

Moreover, we have

$$Q_M > 0 \Leftrightarrow \kappa > l(\alpha,\epsilon),$$
$$P_\phi > 0 \Leftrightarrow \kappa < k(\alpha,\epsilon),$$
$$P_z > 0 \Leftrightarrow \kappa < 1,$$

with

$$l(\alpha,\epsilon) := \tfrac{1}{2}[1 - 2\alpha(1-\alpha)] + (\tfrac{1}{4}[1 - 2\alpha(1-\alpha)]^2 - \epsilon^{-1}\alpha^2(1-\alpha)^2)^{1/2}$$

and

$$k(\alpha,\epsilon) := \tfrac{1}{2}[1-2\alpha(1-\alpha)] + (\tfrac{1}{4}[1-2\alpha(1-\alpha)]^2 - (\epsilon-1)^{-1}\alpha^2(1-\alpha)^2)^{1/2}.$$

As a result of these calculations we see that the domain in which (10) is valid is given by

$$D := \{(R,\alpha,\kappa,\epsilon)\in\mathbb{R}^4 \mid R>0,\ 0<\alpha<\tfrac{1}{2},\ g(\alpha,\epsilon)<\kappa<f(\alpha,\epsilon),\ \epsilon>1\}.$$

If $R>0$ and $\epsilon>1$ are given, the domain of α and κ can be drawn in an (α,κ)-diagram (see Fig. 1).

A detailed discussion of D gives the following results: For each $\epsilon>1$ there exist solutions with positive and negative P_z. For $1<\epsilon<\frac{4}{3}$, P_ϕ can take both signs; for $\epsilon>\frac{4}{3}$, P_ϕ is negative. In the domain $1<\epsilon<\frac{9}{8}$, P_z and P_ϕ can take all four sign combinations; for $\frac{9}{8}<\epsilon<\frac{4}{3}$ the combination $P_\phi>0$, $P_z<0$ does not exist. The point $P_\phi=P_z=0$ only exists for $\epsilon<\frac{9}{8}$ (dust).

The behavior of P_ϕ can be interpreted partly by the Newtonian theory: If mass–energy is constant and v increases, the stability condition on gravitational force, centripedal force, and force produced by P_ϕ implies a decrease of P_ϕ. For large v, only negative P_ϕ can hold the matter together. The Newtonian stability condition is

$$2\pi Q_M = R^{-1}v^2 + 2\pi(RQ_M)^{-1}P_\phi$$

[the gravitational force on a mass element Δm is $2\pi Q_M\,\Delta m$; centripedal force, $R^{-1}v^2\,\Delta m$; force produced by P_ϕ: $2\pi P_\phi(RQ_M)^{-1}\,\Delta m$].

For small RP_ϕ and RP_z we can deduce this equation easily from our results. If masses, stresses, and velocity are small, Newton's theory is valid asymptotically.

The domain for the static limit ($\lim_{\epsilon\to 1} D$) is given by

$$\epsilon = 1,$$

$$\lim_{\epsilon\to 1} g(\alpha,\epsilon) = 1-\tfrac{1}{2}\alpha(2-\alpha) < \kappa < \infty = \lim_{\epsilon\to 1} f(\alpha,\epsilon).$$

κ has no upper bound now. This limit is physically meaningful. P_ϕ is always positive. A dust solution would only exist for $\alpha\to 0$, but here matter vanishes (of course, there cannot exist a static stressless hollow cylinder).

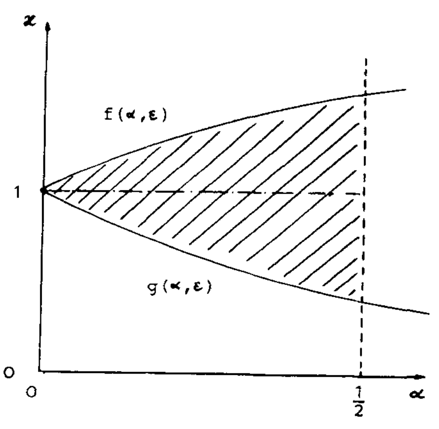

FIG. 1. The domain of solutions satisfying (10).

The limit $\alpha\to\frac{1}{2}$ means something like "maximum mass." The main difference to the axisymmetric case is the absence of an event horizon.

4. MACHIAN EFFECTS

Our solution consists of two separate, static parts of space–time. In the exterior field, a torqueless gyroscope tied to an observer at constant ρ, z, and ϕ will always point towards the same direction (e.g., to the axis). This is a consequence of the condition (1), $\lim_{\rho\to\infty}\Omega(\rho)=0$. Nothing seems to rotate in the exterior field. The interior global inertial frame rotates with angular velocity $\Omega_- = \Omega(0)$ relative to "infinity" (relative to the fixed stars). The dragging coefficient[5] of the interior part of space–time is defined by

$$\chi := \frac{\Omega_-}{\Lambda} = \left.\frac{d\phi}{dt}\right|_{\text{interior inertial frame}} \Big/ \left.\frac{d\phi}{dt}\right|_{\text{matter}}$$

This coefficient corresponds to the coefficients given by Thirring[12,13] and Cohen and Brill[7,14] for the axisymmetric case.

We are interested in the dependence of χ on the matter variables. χ does not depend on R. We find

$$\chi(\alpha,\kappa,\epsilon) = \Omega_-[\Omega_- + R^{-1}(1-\epsilon^{-1})^{1/2}]^{-1} = [1+(1+b)^{1/2}(1-\epsilon^{-1})^{1/2}]^{-1}.$$

For $0<b<\infty$ and $\epsilon>1$, we have $0<\chi<1$. If we don't restrict the domain for the matter variables α, κ, and ϵ, we might get $\chi>1$. We shall not deal with such unphysical situations here.[1]

χ is defined physically meaningfully on the domain D. Moreover, we can consider χ defined on the domain

$$G := \{(\alpha,\kappa,\epsilon)\in\mathbb{R}^3 \mid 0<\alpha<\tfrac{1}{2},\ \kappa>1-2\alpha,\ \epsilon>1\}$$

(here $Q_M+P_\phi>0$). χ is continuous and can be extended to a unique function on the closure $\bar{G}$ of G. Some values of χ on the boundary of $\bar{G}$ might have no physical meaning.

On $\bar{G}$ we have $\chi(\alpha,\kappa,1)\neq 0$; the limit $\epsilon\to 1$ corresponds to slow rotation. In the axisymmetric case and the slow motion limit, the dragging coefficient has been considered by Cohen and Brill.

The qualitative behavior of χ is given by its derivatives. On G we obtain[6]

$$\frac{\partial\chi}{\partial\alpha}>0,\quad \frac{\partial\chi}{\partial\kappa}>0,\quad \frac{\partial\chi}{\partial\epsilon}<0.$$

The dragging becomes better if α and κ increase and ϵ (or v) decreases. Moreover, $0\leqslant\chi\leqslant 1$ on $\bar{G}$.

Finally, we list up some interesting values of χ:

(i)

$$\chi(\alpha,\kappa,1) = 1-\kappa^{-1}(1-2\alpha) = 8\pi\int_0^\infty (T^3_3 - T^0_0)(-g)^{1/2}\,d\rho \neq 0$$

(slow rotation, cf. Refs. 7 and 14).

(ii)

$$\chi(\alpha,1-2\alpha,\epsilon)=0$$

($P_\phi = -Q_M$ boundary of $\bar{G}$).

(iii)

$$\lim_{\epsilon\to\infty} \chi(\alpha,\kappa,\epsilon) = \tfrac{1}{2}[1-\kappa^{-1}(1-2\alpha)] = \tfrac{1}{2}\chi(\alpha,\kappa,1)$$

(fast rotation; thus: dragging coefficient for fast rotation $=\frac{1}{2}$ dragging coefficient for slow rotation).

(iv)

$\chi(0,1,\epsilon)=0$

(vanishing mass).

(v)

$$\sup_{\substack{\alpha\in(0,1/2)\\ \kappa\in[1-2\alpha,\infty)}} \chi(\alpha,\kappa,\epsilon) = \chi(\tfrac{1}{2},\kappa,\epsilon) = (1+v)^{-1}$$

(best dragging for given v, κ arbitrary).

(vi)

$\chi(\frac{1}{2},\kappa,\epsilon) = [1+(1-\epsilon^{-1})^{1/2}]^{-1} = (1+v)^{-1}$

(maximum mass).

For a given velocity, the dragging coefficient is never larger than $(1+v)^{-1}$; its largest value is reached for maximum mass ($\alpha\to\frac{1}{2}$).

(vii)

$$\begin{aligned}\chi_{\text{dust}} &= \chi(\alpha,1,\tfrac{1}{2}(1+\alpha)(2-\alpha)) = \alpha(2-\alpha)\\ &= -8\pi\int_0^\infty T^0_0\,(-g)^{1/2}\,d\rho\\ &= 4\times\text{usual mass energy}\end{aligned}$$

(cf. Ref. 12).

For dust, χ is never larger than $\frac{3}{4}$; this value is reached for $\alpha\to\frac{1}{2}$. $\chi=1$ is only reached for $\alpha\to\frac{1}{2}$ and $v\to 0$ (κ arbitrary), i.e., in the limit of large masses and slow rotation ("perfect dragging").

The results reproduce many effects that are expected from rotating hollow cylinders. The two main differences to the spherical shell are the absence of a Schwarzschild radius (here χ becomes 1 if the Schwarzschild radius is reached for slow rotation) and the absence of purely local Thirring–Lense effects[15] (on planet orbits).

ACKNOWLEDGMENTS

I wish to thank Professor W. Thirring and R. Wallner for corrections and discussions.

[1]E. Frehland, Comm. Math. Phys. **26**, 307 (1972).

[2]General reference to rotating cylinders: D. Kramer, H. Stephani, E. Herlt, and M. MacCallum, *Exact Solutions of Einstein's Field Equations* (Cambridge U.P., Cambridge, 1980), p. 220.

[3]E. Frehland, Comm. Math. Phys. **23**, 127 (1971).

[4]M. M. Som, A. F. F. Teixeira, and I. Wolk, Gen. Relativ. Gravit. **7**, 263 (1967).

[5]F. Embacher, "Global description of Machian effects in stationary space-times of cylindrical symmetry" (Wien, 1981) (to be published).

[6]F. Embacher, "Mach'sche Effekte und Lösungen der Einstein'schen Feldgleichungen im Fall zylindersymmetrischer, stationär rotierender Materie," Ph.D. thesis, University of Vienna, 1980 (unpublished).

[7]J. M. Cohen and D. R. Brill, Nuovo Cimento **56**, 209 (1968).

[8]J. W. van Stockum, Proc. R. Soc. Edinburgh **57**, 135 (1937).

[9]F. J. Tipler, Phys. Rev. **9**, 2203 (1974).

[10]R. O. Hansen and J. Winicour, J. Math. Phys. **16**, 804 (1975).

[11]A. R. King, Comm. Math. Phys. **38**, 157 (1974).

[12]H. Thirring, Phys. Z. **19**, 33 (1918).

[13]H. Thirring, Phys. Z. **22**, 29 (1921).

[14]D. R. Brill and J. M. Cohen, Phys. Rev. **143**, 1011 (1966).

[15]J. Lense and H. Thirring, Phys. Z. **19**, 156 (1918).

A simplified derivation of the Geroch group in two-dimensional reduced gravity

Yong-Shi Wu[a)]
Institute for Advanced Study, Princeton, New Jersey 08540
Department of Physics, FM-15, University of Washington, Seattle, Washington 98195
Mo-Lin Ge
Department of Physics, Lanzhou University, Gansu, People's Republic of China

(Received 23 November 1982; accepted for publication 6 January 1983)

By generalizing our previous treatment of hidden symmetry in two-dimensional chiral models, a simple and explicit approach is proposed to the Geroch group for the vacuum Einstein equations with the metric tensor depending only on two variables ("two-dimensional reduced gravity"). An infinite number of infinitesimal transformations for the metric tensor preserving the equations of motion are summarized by an explicit parametric transformation and the commutators among them are calculated in a simple and straightforward way. These transformations and their group structure are further identified to be those of the well-known Geroch group.

PACS numbers: 04.20.Jb

I. INTRODUCTION

Ten years ago Geroch, in his pioneer work,[1] found an infinite-dimensional symmetry group for the solutions of vacuum Einstein equations which admit two commuting Killing vectors. His treatment was simplified by Kinnersley and Chitre,[2,3] who introduced an infinite hierarchy of potentials with two integer indices and considered the action of the symmetry group on these potentials. Later their treatment was further simplified by Hauser and Ernst.[4,5] They succeeded in deriving a linearization system for the problem in question which does not presuppose the K–C formalism. Moreover, they also proposed to use the Riemann–Hilbert transform, acting on the solutions to their linearization system (i.e., generating functions for part of the K–C hierarchy of potentials), for effecting the K–C transformations. However, when Ueno discussed[6] how to work out the commutation relations for the Geroch group in the H–E framework, he still had to resort to the complete K–C infinite hierarchy of potentials. (For a recent elegant review of the Geroch group, see also Ref. 7.) Because of the importance of the Geroch group it is worthwhile to have a new, simplified derivation for it. This is the object of the present paper.

The following approach of ours is a generalization of our previous explicit approach[8,9] to the hidden symmetry algebra in two-dimensional chiral models. For the latter case, in the light-cone coordinates, the equations of motion are

$$\partial_\xi(g^{-1}\partial_\eta g)+\partial_\eta(g^{-1}\partial_\xi g)=0, \tag{1.1}$$

where $g(\xi,\eta)\in G$, a matrix Lie group. According to Ref. 8 and 9, the following infinitesimal transformation of g,

$$\delta_\alpha(l)g=-gU(l;x)T_\alpha U(l;x)^{-1}\quad (x=\xi,\eta), \tag{1.2}$$

leaves the equations of motion (1.1) invariant. [In (1.2), l is a finite parameter, $T_\alpha=\alpha^a T_a$ is an infinitesimal combination of G's generators, and $U(l;x)$ is a solution to the Zakhrov–Mikhailov linearization system associated with Eq. (1.1).]

Upon expanding in powers of l, (1.2) leads to an infinite set of infinitesimal transformations,

$$\delta_\alpha(l)g=\sum_{m=0}^{\infty} l^m\delta_\alpha^{(m)}g,$$

all of which leave Eq. (1.1) invariant. The commutators between parametric transformations (1.2) can be explicitly calculated with the help of the knowledge of the linearization equations for $U(l;x)$. Again, by expanding in powers of finite parameter, the infinite-dimensional Lie algebra can be easily identified from the parametric commutators. This approach has the following advantages:

(1) It is explicit in that we have an explicit expression for the parametric transformation which summarizes the infinite set of symmetry transformations.

(2) It is direct in that we give the infinitesimal transformations for the basic field $g(x)$, and the calculation of commutators does not assume any knowledge of final results as some induction proof does.

(3) It is simple in that what we mainly need in this approach is just the linearization system (Lax pair) for the equations of motion.

As is well known, for space-time admitting two commuting Killing vectors, the four-dimensional vacuum Einstein equations can be reduced to a two-dimensional problem, for which the equations of motion [see below, Eq. (2.2)] are very similar to Eq. (1.1). This similarity triggered us to try to generalize our approach sketched above to the present case. The generalization, of course, will also have the advantages stated above, of which simplicity is most remarkable. In contrast to the very heavy Kinnersley–Chitre formalism, our approach will avoid completely the introduction of the double infinite hierarchy of potentials.

This paper is organized as follows. Section II is devoted to a brief review of the Hauser–Ernst linearization equations which is the necessary ingredient of our approach. In Sec. III, an explicit parametric transformation which is similar to Eq. (1.2) is proposed and verified to be hidden symmetry leaving the equations of motion under consideration invar-

[a)] Permanent address: Institute of Theoretical Physics, Academia Sinica, Beijing, China.

iant. The infinite-dimensional Lie algebra for the infinite set of transformations obtained by expanding in powers of the parameters is calculated out, in Sec. IV, to be the half Kac–Moody algebra $SL(2,R)\otimes R[t]$ (where $R[t]$ is the ring of polynomials in t with real coefficients). In Sec. V we identify our infinite set of transformations with those of the Geroch group in the following way. We find out the explicit expressions for infinitesimal Riemann–Hilbert transformations in the framework of the R–H transform for the Geroch group.[5,6] Comparing them with ours we find that the half set of the infinitesimal R–H transformations, i.e., those with negative integer indices, are gauge transformations in the sense that they do not change the basic metric field, and the remaining nontrivial half set, those with positive and zero indices, coincide with ours. (This is consistent with the well-known fact[3] that a half set of the Geroch group consists of gauge transformations.)

It is obvious that our approach can be straightforwardly generalized to the electrovacuum with two commuting Killing vectors.[2–4] In the following we will give the presentation of our formulation for the case in which both Killing vectors are spacelike (e.g., for the gravitational plane waves). It is easy to change the notation to adopt the case in which one Killing vector is spacelike and the other is timelike (e.g., for the axially symmetric stationary vacuum).

II. THE HAUSER–ERNST LINEARIZATION EQUATIONS

First of all let us briefly review the H–E formulation of the linearization equations[5] (see also Ref. 6) for the problem in question, and establish our notation.

In the case of gravitational plane waves, the metric of space-time is of the form

$$-ds^2=f(t,z)(-dt^2+dz^2)+g_{ab}(t,z)\,dx^a dx^b, \tag{2.1}$$

where $a,b=1,2$ and $(x^1,x^2)=(x,y)$. From the vacuum Einstein equations it follows that the equations of motion for g_{ab} are just

$$\partial_\xi(\alpha g^{-1}\partial_\eta g)+\partial_\eta(\alpha g^{-1}\partial_\xi g)=0, \tag{2.2}$$

where $g=(g_{ab})$ is a 2×2 symmetric matrix, $\alpha^2=\det g$ and $\xi=\frac{1}{2}(t+z)$, $\eta=\frac{1}{2}(t-z)$ are light-cone variables. The function $f(t,z)$ can then be obtained by solving[10]

$$\begin{aligned}\partial_\xi \log f&=\frac{\partial_\xi^2\log\alpha}{\partial_\xi\log\alpha}+\frac{1}{4\alpha\partial_\xi\alpha}\,\mathrm{tr}(\alpha g^{-1}\partial_\xi g)^2,\\ \partial_\eta \log f&=\frac{\partial_\eta^2\log\alpha}{\partial_\eta\log\alpha}+\frac{1}{4\alpha\partial_\eta\alpha}\,\mathrm{tr}(\alpha g^{-1}\partial_\eta g)^2.\end{aligned} \tag{2.3}$$

Thus the problem is reduced to solving the two-dimensional Eq. (2.2) which is similar to the equations of motion (1.1) in principal chiral models. It is just this similarity that motivates us to generalize our treatment[7,8] of the infinite-parameter hidden symmetry algebra for the latter to the present case.

Denote

$$\epsilon=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

Equation (2.2) is easily shown to be equivalent to

$$\partial_\xi(\alpha^{-1}g\epsilon\partial_\eta g)+\partial_\eta(\alpha^{-1}g\epsilon\partial_\xi g)=0. \tag{2.4}$$

In view of this equation, there exists a twist potential ψ such that

$$\partial_\xi\psi=\alpha^{-1}g\epsilon\partial_\xi g,\quad \partial_\eta\psi=-\alpha^{-1}g\epsilon\partial_\eta g. \tag{2.5}$$

Then we define the generalized (matrix) Ernst potential as follows[2]:

$$E=g+i\psi. \tag{2.6}$$

Using

$$g\epsilon g=(\det g)\epsilon=\alpha^2\epsilon, \tag{2.7}$$

it can be verified that E satisfies

$$\partial_\xi E=i\alpha^{-1}g\epsilon\partial_\xi E,\quad \partial_\eta E=-i\alpha^{-1}g\epsilon\partial_\eta E, \tag{2.8}$$

with $g=\mathrm{Re}\,E$. Instead of g, the Ernst potential E can be thought of as the basic field variables in the problem under consideration for which the equations of motion are Eq. (2.8).

For further development we need more knowledge about ψ and E. Because $\psi-\tilde\psi$ ($\sim$ means transpose) is an antisymmetric matrix, we can write it as

$$\psi-\tilde\psi=2\beta\epsilon, \tag{2.9}$$

where β is a function (not a matrix) of (z,t). From Eq. (2.5) it follows

$$\partial_\xi\beta=\partial_\xi\alpha,\quad \partial_\eta\beta=-\partial_\eta\alpha, \tag{2.10}$$

so that β and α are a pair of conjugate solutions of the free-wave equation: $\partial_\xi\partial_\eta\beta=\partial_\eta\partial_\xi\alpha=0$.

Noting that $\frac{1}{2}(E+E^+)=g+i\beta\epsilon$ ($^+$ means Hermitian conjugate) Eq. (2.8) can be cast into the form

$$\begin{aligned}2(\beta+\alpha)\partial_\xi E&=(E+E^+)(i\epsilon)\partial_\xi E,\\ 2(\beta-\alpha)\partial_\eta E&=(E+E^+)(i\epsilon)\partial_\eta E.\end{aligned} \tag{2.11}$$

From these equations it was shown in Ref. 5 that E satisfies

$$\partial_\xi\widetilde{E}(i\epsilon)\partial_\xi E=\partial_\eta\widetilde{E}(i\epsilon)\partial_\eta E=0, \tag{2.12}$$

$$\partial_\xi E^+(i\epsilon)\partial_\eta E=\partial_\eta E^+(i\epsilon)\partial_\xi E=0. \tag{2.13}$$

Using $E-\widetilde{E}=2i\beta\epsilon$, Eq. (2.12) can also be written as

$$\begin{aligned}\partial_\xi E(i\epsilon)\partial_\xi E&=\partial_\xi(\beta+\alpha)\cdot\partial_\xi E,\\ \partial_\eta E(i\epsilon)\partial_\eta E&=\partial_\eta(\beta-\alpha)\cdot\partial_\eta E.\end{aligned} \tag{2.14}$$

Equations (2.12)–(2.14) will be useful later.

Following Hauser and Ernst,[5] it is convenient to use the notation of differential forms here. Denoting the dual operation on 1-forms in ξ and η by $*$, i.e.,

$$*d\xi=d\xi,\quad *d\eta=-d\eta, \tag{2.15}$$

Eqs. (2.11) and (2.13) can be written, respectively, in a more compact form:

$$2(\beta+\alpha*)dE=(E+E^+)(i\epsilon)dE, \tag{2.11$'$}$$

$$dE^+(i\epsilon)_\wedge dE=dE^+(i\epsilon)_\wedge *dE=0. \tag{2.13$'$}$$

Now we are in a position to derive the linearization equations, which are the necessary ingredient of our later treatment. To this end, we introduce a (complex) parameter t and rewrite the equations of motion, Eq. (2.11$'$), as follows:

$$t[1-2t(\beta+\alpha*)]\,dE=t[1-t(E+E^+)(i\epsilon)]\,dE,$$

or, equivalently,

$$t\,dE = A(t)\Gamma(t), \tag{2.16}$$

where we have defined

$$A(t) = 1 - t(E + E^{+})(i\epsilon), \tag{2.17}$$

$$\begin{aligned}\Gamma(t) &= [1 - 2t(\beta + \alpha^*)]^{-1}\, dE \\ &= \frac{t}{1 - 2(\beta + \alpha)t}\,\partial_\xi E\, d\xi + \frac{t}{1 - 2(\beta - \alpha)t}\,\partial_\eta E\, d\eta.\end{aligned} \tag{2.18}$$

Applying d to Eq. (2.16), we have

$$dA(t)_\Lambda \Gamma(t) + A(t)\, d\Gamma(t) = 0. \tag{2.19}$$

Using Eqs. (2.17), (2.13′), and (2.16), it can be easily shown that

$$dA(t)_\Lambda \Gamma(t) = -A(t)\Gamma(t)(i\epsilon)_\Lambda \Gamma(t). \tag{2.20}$$

Substituting it into Eq. (2.19) leads to

$$d\Gamma(t)(i\epsilon) = \Gamma(t)(i\epsilon)_\Lambda \Gamma(t)(i\epsilon). \tag{2.21}$$

It is obvious that this equation is just the integrability of the following linear differential system,

$$dF(t) = \Gamma(t)(i\epsilon)F(t), \tag{2.22}$$

where $F(t) \equiv F(\xi,\eta;t)$ is a 2×2 matrix function. Equation (2.22) is none other than the H–E linearization system.[5] Its component form is

$$\begin{aligned}\partial_\xi F(t) &= \frac{t}{1 - 2t(\beta + \alpha)}\,\partial_\xi E(i\epsilon)F(t), \\ \partial_\eta F(t) &= \frac{t}{1 - 2t(\beta - \alpha)}\,\partial_\eta E(i\epsilon)F(t).\end{aligned} \tag{2.23}$$

In Ref. 5 the following relations for $F(t)$ have been derived from Eq. (2.22) with the help of Eqs. (2.10)–(2.14):

$$dF(0) = d[\dot{F}(0) - E(i\epsilon)F(0)] = 0, \tag{2.24}$$

$$d[\lambda(t)\det F(t)] = 0, \tag{2.25}$$

$$d[F(t)^x(i\epsilon)A(t)F(t)] = 0, \tag{2.26}$$

where $\dot{F}(t) = \partial F(t)/\partial t$, $F(t)^x = F(\bar{t})^+$ ($\bar{t}$ means complex conjugate of t), and

$$\lambda(t) = [(1 - 2\beta t)^2 - (2\alpha t)^2]^{1/2}. \tag{2.27}$$

Because the linear system (2.22) determines $F(t)$ only up to right-multiplication by an arbitrary matrix function of t, we have freedom to choose the value of $F(0)$. In Refs. 3 and 5 they choose $F(0) = i\epsilon$. For later convenience we would rather choose

$$F(t=0) = 1, \tag{2.28}$$

and, consequently,

$$\dot{F}(0) = E(i\epsilon), \tag{2.29}$$

$$\lambda(t)\det F(t) = 1, \tag{2.30}$$

$$F(t)^x(i\epsilon)A(t)F(t) = i\epsilon. \tag{2.31}$$

Here we have required that the value of the left-hand sides of Eqs. (2.30) and (2.31), which have no dependence on ξ, η according to Eqs. (2.25) and (2.26), are independent of t, too. This requirement fixes the dependence of $F(t)$ on t to a large extent.

III. EXPLICIT PARAMETRIC HIDDEN SYMMETRY TRANSFORMATION

Motivated by the explicit hidden symmetry transformations (1.2) proposed by us[8,9] for 2D principal chiral models, we propose to consider the following one-parameter infinitesimal transformation of the matrix Ernst potential,

$$\delta E = -(1/l)[F(l)TF(l)^{-1} - T](i\epsilon), \tag{3.1}$$

where l is a real parameter, $F(l)$ the solution to the linearization system (2.22) with t replaced by l, and T an infinitesimal, real 2×2 matrix. Below we will show that if we choose T such that

$$\operatorname{tr} T = 0 \quad \text{or} \quad T \in \text{Lie algebra } \mathrm{SL}(2,R), \tag{3.2}$$

then the transformation (3.1) is a hidden symmetry transformation, i.e., one preserving the equations of motion (2.8), for the problem under consideration.[11]

First we note that from Eq. (3.1) it follows that

$$\delta E - \delta\tilde{E} = -(1/l)[\operatorname{tr} F(l)TF(l)^{-1} - \operatorname{tr} T](i\epsilon) = 0, \tag{3.3}$$

where we have used $\Lambda\epsilon + \epsilon\tilde{\Lambda} = (\operatorname{Tr}\Lambda)\epsilon$ (for any 2×2 matrix Λ) and Eq. (3.2). Hence δE is symmetric. Therefore $\delta g \equiv \operatorname{Re}(\delta E)$ is symmetric, as it should be. (This is necessary for $E + \delta E$ to be also a matrix Ernst potential for some space-time.) Moreover, Eq.(3.3) leads to

$$(E + \delta E) - (E + \delta E)^{\sim} = 2i\beta\epsilon \quad \text{or} \quad \delta\beta = 0. \tag{3.4}$$

Now let us further prove that under the transformation (3.1) α is invariant, too, i.e.,

$$\delta\alpha = 0 \quad \text{or} \quad \det(g + \delta g) = \alpha^2. \tag{3.5}$$

Indeed, using $\det\Lambda = -\frac{1}{2}\operatorname{tr}(\tilde{\Lambda}\epsilon\Lambda\epsilon)$ (for any 2×2 matrix Λ) and Eq. (3.4), we have

$$\begin{aligned}\delta(\det g) &= -\operatorname{tr}(\delta g\epsilon g\epsilon) \\ &= \tfrac{1}{4}\operatorname{tr}\{(\delta E + \delta E^{+})(i\epsilon)(E + E^{+} - 2i\beta\epsilon)(i\epsilon)\}.\end{aligned} \tag{3.6}$$

However, Eqs. (3.1) and (3.2) lead to

$$\begin{aligned}\delta E + \delta E^{+} = &-(1/l)[F(l)TF(l)^{-1}(i\epsilon) \\ &+ (i\epsilon)F(l)^{x-1}\tilde{T}F(l)^x],\end{aligned} \tag{3.7}$$

where we have used $T^{+} = \tilde{T}$, $T\epsilon + \epsilon T^{+} = 0$. Substituting Eq. (3.7) into Eqs. (3.6) and taking Eqs. (3.2) and (2.31) into account, we obtain

$$\begin{aligned}\delta(\det g) &= (-1/4l^2)\operatorname{tr}\{[F(l)TF(l)^{-1} \\ &\quad + (i\epsilon)F(l)^{x-1}\tilde{T}F(l)^x(i\epsilon)][1 - A(l)]\} \\ &= (1/4l^2)\operatorname{tr}\{F(l)^{x-1}(i\epsilon T + i\tilde{T}\epsilon)F(l)^{-1}(i\epsilon)\} = 0.\end{aligned}$$

Thus, our transformation (3.1) does not change either α or β.

Finally, we need to verify that $E + \delta E$ satisfies the equations of motion (2.11′), too; i.e., δE satisfies

$$\begin{aligned}2(\beta + \alpha^*)d(\delta E) = &(\delta E + \delta E^{+})(i\epsilon)\, dE \\ &+ (E + E^{+})(i\epsilon)d(\delta E).\end{aligned} \tag{3.8}$$

From Eq. (3.1), using the linearization equation (2.22) for $F(l)$, it is easy to obtain

$$\begin{aligned}d(\delta E) = &-[1 - 2l(\beta + \alpha^*)]^{-1} \\ &\times [dE(i\epsilon), F(l)TF(l)^{-1}](i\epsilon).\end{aligned} \tag{3.9}$$

Multiplying it from the left by $2(\beta + \alpha^*)$, using Eq. (2.11′) and the identity $[AB,C] = A[B,C] + [A,C]B$, we have

$$2(\beta+\alpha^*)d(\delta E) = (E+E^+)(i\epsilon)d(\delta E) - [(E+E^+)(i\epsilon), F(l)TF(l)^{-1}]\Gamma(l). \tag{3.10}$$

On the other hand, starting from Eqs. (3.7) and (2.16) and using Eq. (2.31), we obtain

$$(\delta E+\delta E^+)(i\epsilon)\,dE = -(1/l)\{F(l)TF(l)^{-1}A(l) + (i\epsilon)F(l)^{x-1}\widetilde{T}(i\epsilon)F(l)^{-1}\}\Gamma(l). \tag{3.11}$$

Making use of $\widetilde{T}\epsilon + \epsilon T = (\mathrm{tr}\,T)\epsilon = 0$ and Eq. (2.31) once more,

$$\begin{aligned}(\delta E+\delta E^+)(i\epsilon)dE &= -(1/l)[F(l)TF(l)^{-1}, A(l)]\Gamma(l)\\ &= -[(E+E^+)(i\epsilon), F(l)TF(l)^{-1}]\Gamma(l).\end{aligned} \tag{3.12}$$

In the last step, Eq. (2.17) has been used. Upon substitution of Eq. (3.12) into (3.10), we arrive at Eq. (3.8), as desired.

In summary, Eqs. (3.3), (3.4), (3.5), and (3.8) indicate that $E+\delta E$, with δE given by Eq. (3.1), is really a matrix Ernst potential for some space-time. In other words, the following transformation,

$$\delta g = \mathrm{Re}\{-(1/l)[F(l)TF(l)^{-1} - T](i\epsilon)\}, \tag{3.13}$$

with l real, T an infinitesimal element of SL(2,R), and $F(l)$ satisfying Eqs. (2.22) and (2.28), preserves the determinant and the property of g being symmetric and gives rise to new solutions of the equations of motion (2.2) under consideration. Note that a simpler formula for δg is actually

$$\delta g = (-1/l)\,\mathrm{Re}\{F(l)TF(l)^{-1}(i\epsilon)\}. \tag{3.14}$$

IV. HALF KAC–MOODY ALGEBRA

Since $F(l)$ is analytic around $l=0$, we can expand the right-hand side of Eq. (3.1) in powers of l. In this way we are able to obtain an infinite set of infinitesimal hidden symmetry transformations indexed with zero and positive integers:

$$\delta E = \sum_{n=0}^{\infty} l^n \delta^{(n)} E. \tag{4.1}$$

Each of them satisfies Eqs. (3.3)–(3.5) and (3.8) so that all of them are infinitesimal hidden symmetry transformations. As is well known, the existence of an infinite set of hidden symmetry transformations in the case discussed was first pointed out by Geroch.[1] Here we give explicit expressions for these infinitesimal transformations, for the first time in the literature to our knowledge.

In this section we will calculate the commutators between these infinite numbers of transformations. When doing so, we need first to know what is the variation of $F(t)$ induced by δE. From Eq. (2.22) the differential equation satisfied by $\delta F(t)$ is

$$d\delta F(t) = t[1-2t(\beta+\alpha)^*]^{-1} \times \{d(\delta E)(i\epsilon)F(t) + dE(i\epsilon)\delta F(t)\}, \tag{4.2}$$

and Eqs. (2.28)–(2.31) put the following constraints on δF:

$$\delta F(0) = 0, \quad \delta\dot{F}(0) = (\delta E)(i\epsilon), \tag{4.3}$$

$$\det(F(t)+\delta F(t)) = \det F(t), \tag{4.4}$$

$$\delta(F(t)^x(i\epsilon)A(t)F(t)) = 0. \tag{4.5}$$

By direct calculation it can be verified that the expression

$$\delta F(t) = [t/(t-l)][F(l)TF(l)^{-1} - F(t)TF(t)^{-1}]F(t), \tag{4.6}$$

which is identical to the corresponding one in our treatment of hidden symmetry for 2D principal chiral models [Eq. (4.2) in Ref. 12], indeed satisfies all the defining equations (4.2)–(4.5). As an example, we present here the proof of Eq. (4.5). Using Eqs. (2.17), (2.31), and (3.7), it follows from Eq. (4.6) that

$$\begin{aligned}&\delta(F(t)^x(i\epsilon)A(t)F(t))\\ &= [t/(t-l)]\,[F(t)^xF(l)^{x-1}\widetilde{T}F(l)^xF(t)^{x-1}(i\epsilon) + (i\epsilon)F(t)^{-1}F(l)TF(l)^{-1}F(t)]\\ &\quad + (t/l)F(t)^x[(i\epsilon)F(l)TF(l)^{-1} + F(l)^{x-1}\widetilde{T}F(l)^x(i\epsilon)]F(t).\end{aligned} \tag{4.7}$$

However, using Eq. (2.31) and $A(t) = 1-(t/l)[1-A(l)]$, we have

$$\begin{aligned}(i\epsilon)F(t)^{-1}F(l) &= F(t)^x(i\epsilon)A(t)F(l)\\ &= [(l-t)/l]F(t)^x(i\epsilon)F(l) + (t/l)F(t)^xF(l)^{x-1}(i\epsilon),\end{aligned} \tag{4.8}$$

and, similarly,

$$F(l)^xF(t)^{x-1}(i\epsilon) = [(l-t)/l]F(l)^x(i\epsilon)F(t) + (t/l)(i\epsilon)F(l)^{-1}F(t). \tag{4.8'}$$

Substituting Eqs. (4.8) and (4.8′) into Eq. (4.7) and making use of $\epsilon T + \widetilde{T}\epsilon = 0$, we can check that the right-hand side of Eq. (4.7) vanishes.Q.E.D.

Now let us consider two transformations of the type (3.1):

$$\delta_\alpha(l)E = -(1/l)[F(l)T_\alpha F(l)^{-1} - T_\alpha](i\epsilon) \quad (T_\alpha = \alpha^a T_a) \tag{4.9}$$

$$\delta_\beta(l')E = -(1/l')[F(l')T_\beta F(l')^{-1} - T_\beta](i\epsilon) \quad (T_\beta = \beta^a T_a), \tag{4.9'}$$

where T_a are generators of the Lie algebra SL(2,R), α^a, β^a infinitesimal real constants, and l, l' real parameters. According to (4.6),

$$\delta_\alpha(l)F(l') = [l'/(l'-l)]\,[F(l)T_\alpha F(l)^{-1} - F(l')T_\alpha F(l')^{-1}]F(l'), \tag{4.10}$$

$$\delta_\beta(l')F(l) = [l/(l-l')]\,[F(l')T_\beta F(l')^{-1} - F(l)T_\beta F(l)^{-1}]F(l). \tag{4.10'}$$

Therefore, the commutator between (4.9) and (4.9′) can be computed as follows:

$$\begin{aligned}&[\delta_\alpha(l), \delta_\beta(l')]E\\ &= \delta_\alpha(E+\delta'_\beta E) - \delta_\alpha E - \delta'_\beta(E+\delta_\alpha E) + \delta'_\beta E\\ &= -(1/l)[\delta_\beta(l')F(l)\cdot F(l)^{-1}, F(l)T_\alpha F(l)^{-1}](i\epsilon)\\ &\quad + (1/l')[\delta_\alpha(l)F(l')\cdot F(l')^{-1}, F(l')T_\beta F(l')^{-1}](i\epsilon)\\ &= [1/(l'-l)]\{F(l)[T_\alpha, T_\beta]F(l)^{-1} - F(l')[T_\alpha, T_\beta]F(l')^{-1}\}(i\epsilon)\\ &= \alpha^a\beta^b C^c_{ab}\left(\frac{l}{l-l'}\delta_c(l)E - \frac{l'}{l-l'}\delta_c(l')E\right),\end{aligned} \tag{4.11}$$

where $\delta_\alpha(l)E = \alpha^c\cdot\delta_c(l)E$. Expanding $\delta_\alpha(l)E$ and $\delta_\beta(l')E$ as Eq. (4.1),

$$\delta_\alpha(l)E = \sum_{n=0}^{\infty} l^n \alpha^a \delta_a^{(n)} E,$$

$$\delta_\beta(l')E = \sum_{m=0}^{\infty} l'^m \beta^b \delta_b^{(m)} E,$$

we obtain the commutators

$$[\delta_a^{(n)},\delta_b^{(m)}]E = C^c_{ab}\delta_c^{(n+m)}E \quad (n,m\geqslant 0), \tag{4.12}$$

where $[T_a,T_b] = C^c_{ab}T_c$, i.e., C^c_{ab} are structure constants of SL(2,R). The result shows that the infinite set of our infinitesimal hidden symmetry transformations $\delta_a^{(n)}E$ form the half Kac–Moody algebra SL(2,R) $\otimes R[l]$, where $R[l]$ is the ring of polynomials in l with real coefficients.

V. IDENTIFICATION WITH TRANSFORMATIONS IN THE GEROCH GROUP

In order to establish the relationship between our transformations proposed above and those in the Geroch group, let us consider the Riemann–Hilbert problem approach[4] to the Kinnersley–Chitre transformations of the Geroch group. According to Hauser and Ernst,[4] by introducing an appropriate Riemann–Hilbert problem, the KC transformations $F(t)\to\mathscr{F}(t)$ can be effected by solving the following integral equation of the Cauchy type:

$$\int_C \frac{\mathscr{F}(s)u(s)F(s)^{-1}}{s(s-t)}\,ds = 0, \tag{5.1}$$

subject to the boundary condition

$$\mathscr{F}(0) = 1. \tag{5.2}$$

The complex variable s lies on a circle C surrounding the origin ($t=0$) in the complex t-plane, whereas t is within C. $u(t)$ is a 2×2 complex matrix analytic function such that $u(t)$ is holomorphic on C and

$$\det u(t) = 1, \quad u(t)^\times(i\epsilon)u(t) = i\epsilon. \tag{5.3}$$

It can be shown[5] that if $F(t)$ is a solution to the H–E linearization equation (2.22) which is holomorphic within and on C and satisfies Eqs. (2.28)–(2.31), then the solution $\mathscr{F}(t)$ to Eqs. (5.1) and (5.2) gives a new solution to the H–E linearization equation (2.22) and has the above properties which the original $F(t)$ has. As proved in Ref. 5, this in turn implies that $\mathscr{E} = \mathscr{F}(0)(i\epsilon)$ will be a new matrix Ernst potential satisfying the relations (3.3)–(3.5) and Eq. (2.11).

Following Ueno,[6] let us consider the infinitesimal case. Suppose

$$u(t) = 1 + v(t) \quad \text{with } v(t) \text{ infinitesimal}. \tag{5.4}$$

The conditions (5.3) require that

$$\operatorname{Tr} v(t) = 0, \quad v(t)^\times(i\epsilon) + (i\epsilon)v(t) = 0. \tag{5.5}$$

Without loss of generality, we can take

$$v(t) = t^{-k}T_\alpha, \tag{5.6}$$

where T_α is a 2×2 infinitesimal constant matrix, k can be any integer ($-\infty < k < +\infty$). From Eq. (5.5) it follows that tr $T_\alpha = 0$, $T^*_\alpha = T_\alpha$, i.e., T_α is an infinitesimal element of SL(2,R). Writing the corresponding $\mathscr{F}(t)$ as $F(t) + \delta_\alpha^{(k)}F(t)$, from Eqs. (5.1) and (5.2) we obtain

$$\delta_\alpha^{(k)}F(t)F(t)^{-1} = -\frac{1}{2\pi i}\int_C ds\,\frac{t}{s(s-t)}F(s)v(s)F(s)^{-1}, \tag{5.7}$$

where we have used

$$\delta_\alpha^{(k)}F(t)\cdot F(t)^{-1} = \frac{t}{2\pi i}\int_C \frac{\delta_\alpha^{(k)}F(s)F(s)^{-1}}{s(s-t)}\,ds. \tag{5.8}$$

Using the expansions

$$F(t;\xi,\eta)T_\alpha F(t;\xi,\eta)^{-1} = \sum_{m=0}^{\infty} t^m T_\alpha^{(m)}(\xi,\eta), \tag{5.9}$$

$$\frac{t}{s(s-t)} = \sum_{n=0}^{\infty}\frac{t^{n+1}}{s^{n+2}} \quad \left(\text{for } \left|\frac{t}{s}\right| < 1\right), \tag{5.10}$$

the right-hand side of Eq. (5.7) can be integrated out to be

$$\delta_\alpha^{(k)}F(t)\cdot F(t)^{-1} = -\sum_{n=0}^{\infty} t^{n+1}T_\alpha^{(n+k+1)}(\xi,\eta). \tag{5.11}$$

Note that from Eq. (5.9), $T_\alpha^{(m)}(\xi,\eta)\equiv 0$ for $m<0$. Thus it is easy to see that

$$\delta_\alpha^{(k)}F(t) = F(t)\cdot(-T_\alpha t^{|k|}) \quad \text{for } k<0. \tag{5.12}$$

Moreover, if we define

$$\delta_\alpha(l)F(t) = \sum_{k=0}^{\infty} l^k\delta_\alpha^{(k)}F(t), \tag{5.13}$$

we can obtain from Eq. (5.11)

$$\delta_\alpha(l)F(t)\cdot F(t)^{-1} = \frac{t}{t-l}[F(l)T_\alpha F(l)^{-1} - F(t)T_\alpha F(t)^{-1}], \tag{5.14}$$

which is identical to Eq. (4.6). The results (5.12) and (5.14) indicate that half of the set of generators (with $k<0$) for the Geroch group acting on $F(t)$ are gauge transformations which do not lead to nontrivial transformations for E or g: $\delta_\alpha^{(k)}E = 0$ ($k<0$); and the other half (nontrivial) set of generators can be summarized by our explicit transformations (4.1) with δE given by Eq. (3.1).

Incidentally, we would like to point out that starting from the explicit expressions (5.11) and (5.14) one can easily show that

$$[\delta_a^{(m)},\delta_b^{(n)}]F(t) = C^c_{ab}\delta_c^{(m+n)}F(t) \quad (-\infty < m,n < +\infty), \tag{5.15}$$

where $\delta_\alpha^{(k)}F(t) = \alpha^a\delta_a^{(k)}F(t)$ if $T_\alpha = \alpha^a T_a$. This is the full Kac–Moody algebra SL(2,R) $\otimes R(t,t^{-1})$ for the Geroch group. Thus the presentation in this section can also be viewed as an alternative simplified derivation of the Geroch group which starts with the Riemann–Hilbert problem approach. However, as is well known, only a half set of the generators gives rise to nontrivial transformations for the basic (metric) field g_{ab}, and their algebra is only half Kac-Moody algebra.

ACKNOWLEDGMENTS

One of the authors (Y. S. Wu) is grateful to H. Woolf, S. Adler, and R. Dashen for the warm hospitality extended to him at the Institute for Advanced Study, where most of the work was done. This work was supported in part by the U. S.

Department of Energy under Contract No. DE-AC06-81ER40048.

[1]R. Geroch, J. Math. Phys. **12**, 918 (1971); **13**, 394 (1972).
[2]W. Kinnersley, J. Math. Phys. **18**, 1529 (1977).
[3]W. Kinnersley and D. Chitre, J. Math. Phys. **18**, 1538 (1977); **19**, 1926, 2037 (1978).
[4]I. Hauser and F. J. Ernst, Phys. Rev. D **20**, 362, 1783 (1979).
[5]I. Hauser and F. J. Ernst, J. Math. Phys. **21**, 1126, 1418 (1980).
[6]K. Ueno, RIMS Preprint 374 (1981).
[7]B. Julia, an invited talk at the Johns Hopkins Workshop on Particle Theory, May 1981 (LEPTENS 81-14); lectures given at International School of Cosmology and Gravitation held at Erice, May 1982 (LEPTENS 82-18).
[8]B. Y. Hou, Yale Preprint 80-29 (1980) (unpublished); B. Y. Hou, M. L. Ge, and Y. S. Wu, Phys. Rev. D **24**, 2238 (1981).
[9]M. L. Ge and Y. S. Wu, Phys. Lett. B **108**, 411 (1982).
[10]See, e.g., V. A. Belinskii and V. E. Zakharov, Zh. Eksp. Teor. Fiz. **75**, 1860 (1978) [Sov. Phys. JETP **48**, 935 (1978)].
[11]The key ingredient in Eq. (3.1) is $F(l)TF(l)^{-1}$ [see Eq. (1.2)]. The factor $1/l$ and the second constant term in the bracket are just for later convenience.
[12]Y. S. Wu, Nucl. Phys. B **211**, 160 (1983).

Newtonian gravity on the null cone

J. Winicour[a)]
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 17 September 1982; accepted for publication 12 November 1982)

For a general relativistic ideal fluid, we analyze the Newtonian limit of the initial value problem set on a family of null cones. The underlying Newtonian structure is described using Cartan's elegant space–time version of Newtonian theory and a limiting process rigorously based upon the velocity of light approaching infinity. We find that the existence of a Newtonian limit imposes a strikingly simple relationship between the gravitational null data (i.e., the shear of the null cones) and the Newtonian gravitational potential. This result has immediate application to numerical evolution programs for calculating gravitational radiation and might serve as the basis for a post-Newtonian approximation scheme.

PACS numbers: 04.30. + x

I. INTRODUCTION

As a matter of practice, general relativity is quite often replaced by one of two other theories. In weak-gravity applications, such as atomic or particle physics, it is standard to ignore gravitational effects and use (quantized) special relativistic field theory. In low velocity applications, such as planetary or stellar astronomy, Newtonian gravitational theory gives excellent accuracy. In their appropriate domains, special relativity and Newtonian physics are both complete physical theories in their own right. It would appear fitting that each should provide a background theory for calculating small general relativistic effects.

To this end, in regard to special relativity, linearized theory is standardly adopted in the weak field case. This is formaly based upon a one-parameter family of metrics $g_{ab}(\lambda)$, each satisfying Einstein's equation, with $g_{ab}(0)$ being the Minkowski metric.[1] Geometrically this is a natural approach since the Minkowski metric then provides a space–time background metric. (One might expect that the parameter λ is related to Newton's gravitational constant, in some suitable way. However, a general perturbation theory of this sort has never been rigorously formulated, and, furthermore, Newton's constant would appear to play no role in the description of weak, vacuum gravitational waves.) In this approach, the matter fields satisfy their corresponding special relativistic field equations, to leading order. For example, a system of dust particles, with energy–momentum tensor $T_{\mu\nu} = \rho u_\mu u_\nu$, moves along the straight lines in Minkowski space–time, to leading order in λ. Thus, in the weak field approach, Newtonian gravitational effects appear as perturbations. Post-Newtonian general relativistic corrections involve higher order perturbations, and, at this stage, a slow-motion approximation is often further introduced to facilitate calculations.

From a physical standpoint, it would appear more natural to obtain post-Newtonian, slow-motion corrections by using a perturbation scheme in which Newtonian gravitation theory supplies the background, in the limit of infinite light velocity. At first sight, such a scheme might appear awkward geometrically since Newtonian physics is based upon a three-dimensional Euclidean geometry plus absolute time. However, there exists[2–6] a geometrically elegant space–time version of Newtonian theory based upon an integrable 1-form $t_\mu = t_{,\mu}$, which determines absolute time, a symmetric contravariant "Euclidean metric" of rank three $g^{\mu\nu}$, which satisfies $g^{\mu\nu}\, t_\mu = 0$, and a symmetric connection $\Gamma^\rho_{\mu\nu}$ (nonflat in the presence of gravity), whose geodesics are the free-fall trajectories of Newtonian gravitational theory. This theory reduces to its simplest form in Cartesian inertial coordinates $x^\alpha = (t,x^i) = (t,x,y,z)$, for which $g^{ij} = \delta^{ij}$, $g^{\mu 0} = 0$, and

$$\Gamma^\rho_{\mu\nu} = t_\mu t_\nu g^{\rho\sigma}\Phi_{,\sigma}. \tag{1.1}$$

The Poisson equation for the Newtonian gravitational potential Φ

$$\nabla^2\Phi = 4\pi\rho \tag{1.2}$$

arises from the more geometrical field equation $R_{\mu\nu} = 4\pi\rho t_\mu t_\nu$, which relates the Ricci tensor of the connection to the matter density.

There also exists criteria for the Newtonian limit of a general relativistic space–time.[7–10] As an example of such a limit, consider the one-parameter family of Minkowski metrics described by $ds^2 = dt^2 - \lambda^2\delta_{ij}\, dx^i dx^i$. Here λ is a dimensionless parameter, and we use units for which the actual experimental value of the velocity of light is 1. Thus the limit $\lambda \to 0$ corresponds to a sequence of space–times whose velocity of light (in the above coordinate system) goes to infinity. For $\lambda = 0$,

$$g_{\mu\nu} = t_\mu t_\nu \quad \text{and} \quad \lambda^2 g^{\mu\nu} = -\delta^\mu_i \delta^\nu_j \delta^{ij}; \tag{1.3}$$

i.e., we obtain the ingredients of a Newtonian space–time structure, which in this case is empty of matter and free of gravitation. More generally, if a one-parameter family of space–time with metrics $g_{\mu\nu}(\lambda)$ has a Newtonian limit with matter density ρ, then it can be arranged that, for $\lambda = 0$, (1.1)–(1.3) hold, where $\Gamma^\rho_{\mu\nu}(\lambda)$ is the Riemannian connection associated with $g_{\mu\nu}(\lambda)$.

Despite its attractiveness, this elegant formalism has apparently not been developed as the basis of a perturbation theory for post-Newtonian corrections.[11] The results of this paper contribute to this end. We show that the characteristic initial value formulation of general relativity may be formulated in a way such that the conditions for a Newtonian limit take on some remarkable simplicity. (This suggests analogous results might hold for the Cauchy problem, based upon

 0022-2488/83/051193-06$02.50

a spacelike, initial data hypersurface.)

Actually the motivation for our work comes from a different direction. It stems from a program[12] to study numerically the generation of gravitational waves by an ideal fluid, using a modification of Bondi's version[13-15] of the characteristic initial value problem. As described in Sec. II, in this approach, the initial data for the gravitational field is the shear of an initial outgoing null cone. Although this initial shear can be specified in a constraint-free manner, there is special interest in those choices corresponding to the absence of incoming waves. In the vacuum case, the appropriate data is zero shear, which evolves to produce a flat space–time. However, in the presence of matter, space–times with shear-free initial data contain incoming radiation, except in special cases such as a null cone with spherically symmetric matter distribution. This can be understood by considering a collapsing sphere of dust, with Schwarzschild exterior. Here, although gravitational radiation is clearly absent, the bending of light by matter introduces shear on any outgoing null cone whose vertex is displaced from the center of symmetry. Thus resetting the initial shear to zero is equivalent to introducing incoming radiation.

The underlying physical consideration here is the desire to be able to study the generation of gravitational waves from systems which, at least initially, are in some sense quasi-Newtonian. While it is easy to prescribe exotic initial data, the issue of quasi-Newtonian initial data is quite subtle. We solve this problem in Sec. III by applying the technique of Newtonian limits to the characteristic initial value problem. In Sec. IV, we conclude with a discussion of how the general relativistic system evolves relative to its Newtonian counterpart.

Our conventions are adopted to agree, as closely as possible, with those of Refs. 13–15. We use signature $+ - - -$; units for which Newton's constant and the experimental value of the velocity of light are both unity; Greek letters run from 0 to 3 for space–time indices; a comma to denote partial differentiation; a semicolon to denote space–time covariant differentiation; and Eisenhart's[16] curvature conventions, for which $v_{\mu;\alpha\beta} - v_{\mu;\beta\alpha} = v_\nu R^\nu{}_{\mu\alpha\beta}$, $R_{\mu\nu} = R^\alpha{}_{\mu\nu\alpha}$, and $R = R^\alpha{}_\alpha$. We will deal extensively with quantities defined on a topologically spherical subspace with intrinsic metric h_{AB}, with capital Latin letters running from 2 to 3. We denote the curvature scalar of h_{AB} by $\mathscr{R}$. According to the above conventions, $\mathscr{R} = -2$ for the unit-sphere metric $h_{AB} = q_{AB}$. Two-dimensional covariant differentiation with respect to h_{AB} is denoted by a colon and is often rewritten in $\eth$ notation, based upon a dyad position $h_{AB} = 2m_{(A}\bar m_{B)}$. The numerical conventions for $\eth$ are chosen to agree with Ref. 17. As examples, $v_{A:B}\, m^A \bar m^B = \bar\eth\,(v_A m^A)/\sqrt2 = \bar\eth v$, where $v = v_A m^A/\sqrt2$ is a spin-weight-1 quantity; $f^{:A}{}_A = 2f_{:AB}\, m^A \bar m^B = \eth\bar\eth f$; and, for a quantity η of spin weight s, $(\bar\eth\eth - \eth\bar\eth)\eta = 2s\eta$. Lower case Latin letters, running from 1 to 3, denote the spatial components of the Newtonian limit of tensor fields. As in the previous Minkowski space example, we use a dimensionless parameter λ, corresponding to 1/(velocity of light), to describe the Newtonian limit. The symbol "$\hat=$" stands for "equal to, for $\lambda = 0$"; e.g., $\lambda \hat= 0$.

II. THE INITIAL VALUE PROBLEM

We consider a one-parameter family of metrics $g_{\mu\nu}(\lambda)$ in terms of the characteristic initial value problem based upon a system of outgoing null cones. For an introductory feel for where powers of λ should appear, consider the flat-space case obtained by transforming the Minkowski metrics $ds^2 = dt^2 - \lambda^2\delta_{ij}dx^i dx^j$ into the null polar coordinate system $x^\alpha = (u,r,\theta,\phi)$, with

$$u = t - \lambda r,\quad x = r\sin\theta\cos\phi,$$
$$y = r\sin\theta\sin\phi,\quad z = r\cos\theta. \tag{2.1}$$

We obtain

$$ds^2 = du^2 + 2\lambda\, du\, dr - \lambda^2 r^2 q_{AB}\, dx^A dx^B, \tag{2.2}$$

where $q_{AB}dx^A dx^B = d\theta^2 + \sin^2\theta\, d\phi^2$ is the unit-sphere metric. The contravariant components are given by

$$g^{\mu\nu} f_{,\mu} f_{,\nu} = 2 f_{,0} f_{,1} - \lambda^{-2}(f_{,1} f_{,1} + r^{-2} q^{AB} f_{,A} f_{,B}).$$

We then have[18]

$$g_{\mu\nu} \hat= t_\mu t_\nu \quad\text{and}\quad \lambda^2 g^{\mu\nu} \hat= -\delta^\mu_i \delta^\nu_j e^{ij}, \tag{2.3}$$

where $t_\mu = t_{,\mu} \hat= u_{,\mu}$ and e^{ij} is the polar coordinate version of the contravariant Euclidean metric. Thus in the limit $\lambda \to 0$, we again obtain a Newtonian structure, the difference now being that the hypersurfaces $u =$ const are outgoing null cones for $\lambda \neq 0$ and absolute time slices for $\lambda = 0$. Note that this introduces odd powers of λ whereas only even powers typically occur in Newtonian limits based upon spacelike hypersurfaces.[7-10,19]

This Minkowski space–time example serves as a model for our general relativistic treatment. We introduce a null-polar version of Fermi coordinates based upon the outgoing null cones emanating from a timelike geodesic with proper time u. We let u label these null cones, r be the luminosity distance along these null cones, and $X^A = (\theta,\phi)$ be labels for the null rays. In the $x^\alpha = (u,r,x^A)$ coordinates, the line element for each member of our one-parameter family of metrics takes the Bondi–Sachs form[13,14]

$$ds^2 = g_{00}\, du^2 + 2g_{01}\, du\, dr + 2g_{0A}\, du\, dx^A + g_{AB}\, dx^A dx^B, \tag{2.4}$$

where $\det|g_{AB}| = \lambda^4 r^4 \sin^2\theta$. [Note that diffeomorphism freedom has been used to make (u,r,x^A) a common Bondi–Sachs coordinate system for all metrics of the λ-dependent family.] The boundary and smoothness conditions at the vertex $r = 0$ follow from requiring that the transformation (2.1) lead to a smooth Fermi coordinate system, with a Minkowski metric at the origin. We shall assume asymptotic flatness and that the coordinate system may be extended to $r = \infty$, although for some purposes this is not necessary. These additional assumptions limit the bending of the central null rays and thus qualitatively restrict the size of the system's quadrupole moments.

We set

$$\begin{aligned} g_{AB} &= -r^2\lambda^2 h_{AB},\\ g_{01} &= \lambda e^{2\lambda^2\beta},\\ g_{0A} &= \lambda^3 r^2 U_A,\\ g_{00} &= e^{2\lambda^2\beta} V/r - \lambda^4 r^2 h^{AB} U_A U_B, \end{aligned} \tag{2.5}$$

with $h^{AB}h_{BC} = \delta^A_C$, $\det|h_{AB}| = \sin^2\theta$,

$$h_{AB} = q_{AB} + \lambda^2\gamma_{AB} \tag{2.6}$$

(where q_{AB} is the unit sphere metric) and $V = r + \lambda^2 W$. We raise and lower capital Latin indices with h_{AB}. Here some of the λ factors are introduced so that (2.4) reduces to (2.2) in the flat-space case. The remaining factors are chosen so that it is consistent with Einstein's equation to require that γ_{AB}, β, U_A, and W be smooth fields for $\lambda = 0$ (see Sec. III). The non-vanishing contravariant components of the metric are

$$\begin{aligned} g^{AB} &= -\lambda^{-2}r^{-2}h^{AB}, \\ g^{01} &= \lambda^{-1}e^{-2\lambda^2\beta}, \\ g^{1A} &= e^{-2\lambda^2\beta}U^A, \\ g^{11} &= -\lambda^{-2}e^{-2\lambda^2\beta}V/r. \end{aligned} \tag{2.7}$$

We now consider Einstein's equation $G_{\mu\nu} = -8\pi T_{\mu\nu}$, for a one-parameter family of ideal fluid sources described by

$$T_{\mu\nu} = (\rho + \lambda^2 p)w_\mu w_\nu - \lambda^2 p g_{\mu\nu}, \tag{2.8}$$

with $p = p(\rho)$ and with 4-velocity w_μ of the form

$$w_\mu = t_\mu + \lambda^2 v_\mu, \tag{2.9}$$

where again $t_\mu = t_{,\mu} = (u + \lambda r)_{,\mu}$ and v_μ is a smooth field for $\lambda = 0$. This form for the 4-velocity is chosen so that its contravariant components satsify

$$w^\alpha \mathrel{\hat{=}} (1, V^i) \mathrel{\hat{=}} (1, -v_1, -r^{-2}q^{AB}v_B), \tag{2.10}$$

where $V^i = (V^1, V^A)$ are the velocity components (in polar coordinates) of an ideal fluid in the Newtonian background. In space–time notation, $V^\alpha \mathrel{\hat{=}} v^\alpha$.

For the time being, our primary concern is with the null hypersurface constraint equations.[20] They take the form[18]

$$-8\pi\lambda^{-2}T_{11} = -4\beta_{,1}/r + \lambda^2 c^{AB}c_{AB}/4, \tag{2.11a}$$

$$\begin{aligned} -8\pi\lambda^{-2}T_{1A} = &-(r^4 e^{-2\lambda^2\beta}h_{AB}U^B{}_{,1})_{,1}/2r^2 \\ &-2\beta_{,A}/r + \beta_{,1A} - c_{AB}{}^{:B}/2, \end{aligned} \tag{2.11b}$$

$$\begin{aligned} -8\pi r^2(T - g^{AB}T_{AB}) = &\,2\lambda^{-2}e^{-2\lambda^2\beta}V_{,1} + \lambda^{-2}\mathscr{R} \\ &+ 2\beta^{:A}{}_A + 2\lambda^2\beta^{:A}\beta_{:A} \\ &- e^{-2\lambda^2\beta}r^{-2}(r^4U^A)_{,1:A} \\ &+ \lambda^2 r^4 e^{-4\lambda^2\beta}h_{AB}U^A{}_{,1}U^B{}_{,1}/2, \end{aligned} \tag{2.11c}$$

where we have introduced the shear tensor $c_{AB} = \gamma_{AB,1} = h_{AB,1}/\lambda^2$.

The unconstrained data on an initial null hypersurface consists of the matter variables ρ, v_1, and v_A and the gravitational data c_{AB} (or either γ_{AB} or h_{AB}, since smoothness requires that γ_{AB} vanish at the origin).[21] In terms of this data, radial integration of the hypersurface equations (2.11) determine β, U, and $W = (V - r)/\lambda^2$, and thus the entire initial metric. All integration constants are fixed by the smoothness conditions, which require that β, U, and W vanish at the origin. The normalization condition $w^\mu w_\mu = 1$ determines v_0. The general relativistic matter data induces, for $\lambda = 0$, initial data for a Newtonian field. Our primary question is: What is the appropriate gravitational data such that the limit $\lambda \to 0$ gives the Newtonian gravitational structure for this fluid?

III. NEWTONIAN LIMIT ON THE NULL CONE

In order for a Newtonian limit to exist, the analogs of (1.1)–(1.3) must hold, in a form adapted to spherical coordinates with a freely falling origin. We begin our examination of these conditions by setting $\lambda = 0$ in the hypersurface equations (2.11). This gives, using (2.6) to describe the gravitational degrees of freedom,

$$\beta_{,1} \mathrel{\hat{=}} 2\pi\rho r, \tag{3.1a}$$

$$(r^4U^A{}_{,1})_{,1} \mathrel{\hat{=}} -4r\beta^{:A} + 2r^2\beta_{,1}{}^{:A} - r^2c^{AB}{}_{:B}, \tag{3.1b}$$

$$\begin{aligned} W_{,1} \mathrel{\hat{=}} &\,2\beta - \beta^{:A}{}_A + (r^4U^B)_{,1:B}/2r^2 \\ &+ \gamma^{AB}{}_{:AB}/2 - 4\pi\rho r^2. \end{aligned} \tag{3.1c}$$

Here we have used $(\mathscr{R} + 2)/\lambda^2 \mathrel{\hat{=}} -\gamma_{AB}{}^{:AB}$ in obtaining (3.1c).[18] Also note that all colon-derivatives appearing in (3.1) reduce to covariant derivatives with respect to the unit sphere metric.

From (3.1), it is evident that if γ_{AB}, ρ, p, and v_μ have smooth limits as $\lambda \to 0$, then so do the derived fields β, U^A, and W. As an immediate consequence, it follows from referring to (2.5) and (2.7) that (2.3) holds. Thus, on any null cone for which the gravitational and matter degrees of freedom have the λ dependence (2.6), (2.8), and (2.9) and for which the hypersurface equations hold, the limit $\lambda \to 0$ gives this cone the structure of an absolute-time hypersurface with Euclidean metric.

Now, under these assumptions, consider the limit of the connection. The analog of (1.1) in the present case is

$$\Gamma^\rho_{\mu\nu} \mathrel{\hat{=}} \mathring{\Gamma}^\rho_{\mu\nu} - \lambda^2 t_\mu t_\nu g^{\rho\sigma}\Phi^*{}_{,\sigma}, \tag{3.2}$$

where $\mathring{\Gamma}^\rho_{\mu\nu}$ are the inertial connection coefficients associated with a spherical coordinate system and Φ^* is the solution of Poisson's equation with the boundary condition that it and its spatial gradient vanish at the origin.[22] Thus, because of the free-fall property of the origin, Φ^* differs from the usual Φ, which vanishes at infinity by a monopole–dipole solution of Laplace's equation

$$\Phi^* = \Phi + a + r\sum_m a_m Y_{1m}. \tag{3.3}$$

A straightforward investigation shows that the connection satisfies (3.2) and the boundary conditions if and only if

$$\Phi^* \mathrel{\hat{=}} W/2r + \beta. \tag{3.4}$$

In addition to our previous result concerning absolute time and Euclidean metric, we now have: If the hypersurface equations hold on our system of null cones, for unconstrained variables having the λ dependence (2.6), (2.8), and (2.9), then the connection satisfies (3.2), with Φ^* given by (3.4).

For this Φ^* to satisfy the Poisson equation, the gravitational degrees of freedom γ_{AB} (for $\lambda = 0$) can no longer be arbitrary. Just as γ_{AB} leads uniquely to W via the hypersurface equations, any condition on W restricts γ_{AB}. However, while simple radial integrations lead from γ_{AB} to W, the reverse relationship involves quite complicated partial differential equations. Thus the ($\lambda = 0$) condition that $W/2r + \beta$ satisfy the Poisson equation might seem intractably complex to reexpress in terms of the otherwise unconstrained variable γ_{AB}. Remarkably, the following analysis translates this into a simple condition on γ_{AB}.

We begin this analysis by rewriting (3.1) in terms of spin-weighted functions.[18] We introduce the spin-weight zero quantities α and Z by

$$c_{AB} m^A m^B = \eth^2 \alpha \quad \text{and} \quad (2)^{1/2} U_A m^A = \eth Z. \tag{3.5}$$

These equations define Z up to a term with monopole angular dependence and α up to a monopole-dipole term. (This freedom will be removed below, for $\lambda = 0$.) Also, from $c_{AB} = \gamma_{AB,1}$, we have $\gamma_{AB} m^A m^B \triangleq \eth^2 \int \alpha$, where we introduce the shorthand notation

$$\int f = \int_0^r f(r')\, dr'.$$

The hypersurface equations (3.1) can now be put in the form

$$\beta_{,1} \triangleq 2\pi \rho r, \tag{3.6a}$$

$$\eth(r^4 Z_{,1})_{,1} \triangleq \eth(-4r\beta + 2r^2\beta_{,1} - 2r^2\alpha - r^2\eth\bar{\eth}\alpha), \tag{3.6b}$$

$$W_{,1} \triangleq (2 - \eth\bar{\eth})\beta + \eth\bar{\eth}[r^4(Z + \bar{Z})]_{,1}/4r^2 + \eth^2\bar{\eth}^2 \int (\alpha + \bar{\alpha})/4 - 4\pi\rho r^2. \tag{3.6c}$$

Note that the matter appears as a source in (3.6) only through the density ρ. To leading order in λ, the velocity and pressure do not contribute directly to the gravitational field, as in Newtonian theory. Just as a single scalar suffices to describe the Newtonian gravitational field on an absolute time slice, it seems physically reasonable that a single degree of freedom of the Einsteinian gravitation field suffice to produce this Newtonian limit. With this in mind, let us tentatively proceed on the assumption that (the unconstrained data) α is real for $\lambda = 0$. In the language of $\eth$ calculus,[17] this means that the shear of the null cone c_{AB} is pure electric, for $\lambda = 0$. Also, without further loss of generality (see 3.6b), we assume Z is also real.

With this assumption, we analyze the implications of the Poisson equation (in spherical coordinates)

$$\nabla^2 \Phi^* = (r^2 \Phi^*_{,1})_{,1}/r^2 + \eth\bar{\eth}\Phi^*/r^2 = 4\pi\rho. \tag{3.7}$$

Using (3.4) to substitute for Φ^* and (3.6) to eliminate, as far as possible, radial derivatives of β, Z, and W, (3.7) leads to

$$\eth\bar{\eth}[W + r(r^2 Z)_{,1}] \triangleq 0.$$

We now use this to fix the monopole freedom in Z by setting

$$W + r(r^2 Z)_{,1} \triangleq 0. \tag{3.8}$$

We also now fix the monopole freedom in α by rewriting (3.6b) as

$$(r^4 Z_{,1})_{,1} \triangleq -4r\beta + 2r^2\beta_{,1} - 2r^2\alpha - r^2\eth\bar{\eth}\alpha. \tag{3.6b'}$$

Substituting W from (3.8) into (3.6c) and using (3.6a) and (3.6b′) leads (after some manipulation) to

$$(1 + \eth\bar{\eth}/2)\left[(r^4 Z)_{,1}/r^3 - 2\beta/r - 2\alpha + r^{-1}\eth\bar{\eth} \int \alpha\right] \triangleq 0.$$

We use this to fix the dipole freedom in α by setting

$$(r^4 Z)_{,1}/r^3 - 2\beta/r - 2\alpha + r^{-1}\eth\bar{\eth}\int \alpha \triangleq 0. \tag{3.9}$$

Next we apply the operator $(\partial/\partial r) r^4 (\partial/\partial r)$ to (3.9), which allows us to eliminate Z by using (3.6b′). After extensive manipulation, there results

$$4\beta + 4r\alpha + 5r^2\alpha_{,1} + r^3\alpha_{,11} + r\eth\bar{\eth}\alpha + \eth\bar{\eth}\int \alpha \triangleq 0. \tag{3.10}$$

Finally, by taking the radial derivative of (3.10), β may be eliminated using (3.6a). This leads to the Poisson equation

$$\nabla^2 (r^2\alpha)_{,1} \triangleq -8\pi\rho. \tag{3.11}$$

In fact, a comparison of (3.7) and (3.11) implies that

$$(r^2\alpha)_{,1} \triangleq -2\Phi^*, \tag{3.12}$$

since both sides of this equation share the same smoothness condition at the origin and have the same asymptotic behavior [as may be checked by straightforward expansions using (3.3), (3.4), (3.8), (3.6b′), and (3.9)].[23] Thus, from (3.3), α itself has asymptotic behavior $\alpha \triangleq -\Sigma a_m Y_{1m} - 2a/r + 2M \log r/r^2 + \cdots$, where M is the total Newtonian mass. Although a $\log r$ term, associated with M, occurs in the expansion of α, no such terms occur in $\eth^2\alpha$, so that the shear and all other geometric quantities have asymptotic expansions in $1/r$. This is essential for the smoothness of null infinity, which lies at the heart of a rigorous definition of asymptotic flatness.[24] The form in which α enters (3.12) appears less mysterious when one notes its relation to the Weyl tensor

$$C_{1AB1} m^A m^B \triangleq -\eth^2 (r^2\alpha)_{,1}/2 \triangleq \eth^2 \Phi^*. \tag{3.13}$$

Thus the Newtonian potential also plays the role here of a potential for the Weyl data on a null cone.

We have at this stage attained one major goal, namely the formulation of a simple algorithm for prescribing the null data appropriate for quasi-Newtonian initial conditions. First solve for the potential Φ of the Newtonian system. From this Φ, determine the initial shear from (3.5) and (3.12).[25] This data guarantees, at the initial time, that the λ-dependent family of general relativistic systems has the Newtonian system as a strict limit. In practice, this is an easy scheme to implement, say, in a numerical evolution program. After determining the initial shear, simply set $\lambda = 1$ everywhere. The resulting system will initially approximate the Newtonian system to the extent that p/ρ, v_α, and Φ are small compared to 1.

We note that it is not necessary to make the pure electric assumption $\alpha \triangleq \bar{\alpha}$ in order to satisfy the Poisson equation (3.7). With a complex α, the foregoing analysis proceeds in the same fashion, except for the substitutions $\alpha \to (\alpha + \bar{\alpha})/2$ and $Z \to (Z + \bar{Z})/2$ in (3.8)–(3.12). Thus, in the final result

$$(r^2(\alpha + \bar{\alpha}))_{,1} \triangleq -4\Phi^*, \tag{3.14}$$

only the real part of α is determined by the Newtonian potential and the imaginary part remains free of constraints. However, as far as quasi-Newtonian initial conditions are concerned, it might be expected that the imaginary part of α should vanish. This is in fact necessary for the λ dependence of the gravitational null data, indicated in (2.6), to be preserved by the dynamics, as discussed in the next section.

IV. CONCLUSION

So far we have not considered any dynamical equations but have established, at a kinematic level, the following result for an ideal fluid source. A λ-dependent family of Lorentzian space–times which satisfy the hypersurface equations on a geodesic system of null cones, with null data having the λ dependence indicated in (2.6), (2.8), and (2.9) and satisfying (3.14), has a limit which satisfies the kinematic conditions for a Newtonian space–time, i.e., (2.3), (3.2), and (3.7). We now examine the remaining components of Einstein's equation, which govern the evolution of this system. These are equivalent to the matter evolution equation $T_\mu{}^\nu{}_{;\nu} = 0$ and the gravitational evolution equation $(G_{\alpha\beta} + 8\pi T_{\alpha\beta})m^\alpha m^\beta = 0$.[12]

We first examine the matter evolution equation. To leading order in λ, its component can be reduced to

$$\rho_{,0} - (r^2\rho v_1)_{,1}/r^2 - (\rho v_B)^{:B}/r^2 \hat{=} 0, \tag{4.1a}$$

$$(\rho v_A)_{,0} - (r^2\rho v_A v_1)_{,1}/r^2 - (\rho v_A v_B)^{:B}/r^2 - p_{,A} - \rho\Phi^*_{,A} \hat{=} 0, \tag{4.1b}$$

$$(\rho v_1)_{,0} - (r^2\rho v_1 v_1)_{,1}/r^2 - (\rho v_1 v_A)^{:A}/r^2 + \rho v_A v^A/r^3 - p_{,1} - \rho\Phi^*_{,1} \hat{=} 0. \tag{4.1c}$$

After the substitution $v_i \to V^i$ indicated in (2.10), these are exactly the Euler equations for an ideal fluid,[26] adapted to a spherical coordinate system with freely falling origin. Thus if the above λ-dependent family also satisfies the matter evolution equations, then its limit has all the Newtonian structure of a self-gravitating fluid.

We now investigate to what extent the dynamics of the Einsteinian gravitational field preserves the λ dependence of the null data (2.6) and the Poisson condition (3.14), necessary for the Newtonian limit. The gravitational evolution equation takes the form

$$\begin{aligned} 8\pi\lambda^2(\rho + \lambda^2 p)(v_A m^A)^2 = e^{-2\lambda^2\beta} m^A m^B \big[& r(rh_{AB,0})_{,1}/\lambda \\ & - (rVc_{AB})_{,1}/2 - 2e^{2\lambda^2\beta}(\beta_{:AB} + \lambda^2\beta_{:A}\beta_{:B}) \\ & + r^2 h_{AC} U^C{}_{,1:B} - \lambda^2 r^4 e^{-2\lambda^2\beta} h_{AC} h_{BD} U^C{}_{,1} U^D{}_{,1}/2 \\ & + 2rU_{A:B} + \lambda^2 r^2 c_{AB} U^D{}_{:D}/2 + \lambda^2 r^2 c_{AB:D} U^D \\ & - \lambda^2 r^2 c_A{}^D(U_{B:D} - U_{D:B})\big]. \end{aligned} \tag{4.2}$$

By evaluating (4.2) at $\lambda = 0$, we find that the evolution preserves (2.6) only if

$$\eth^2[(r^2\alpha)_{,1} + 2\beta - (r^2 Z)_{,1}] \hat{=} 0,$$

or only if the term in the bracket vanishes up to an irrelevant monopole–dipole term. The real part of this bracket vanishes as a result of (3.4), (3.14), and (3.8) [with $Z \to (Z + \bar{Z})/2$; see the last paragraph in Sec. III]. Thus we must require $\mathrm{Im}(\alpha - Z) \hat{=} 0$ or, substituting into (3.6b'),

$$\mathrm{Im}[(r^4\alpha_{,1})_{,1}/r^2 + 2\alpha + \eth\bar{\eth}\alpha] \hat{=} 0. \tag{4.3}$$

Taking the radial derivative of (4.3) and setting $f \hat{=} \mathrm{Im}(r^2\alpha_{,1})$, we find $\nabla^2 f = 0$, which implies $f = 0$. Thus, for $\lambda = 0$, α must indeed be real, and the shear must be pure electric, in order for the λ dependence of the gravitational null data to evolve properly.

Consider now a Newtonian fluid, satisfying the Euler and Poisson equations, during some time interval $u_0 \leqslant u < u_1$. Let ρ_N, v_{N^i}, and α_N be some choice of null data, for a general relativistic fluid, on each of the corresponding null cones $u_0 \leqslant u < u_1$ such that, for $\lambda = 0$, ρ_N, v_{N^i}, and $-(r^2\alpha_N)_{,1}/2$ are the density, velocity, and gravitational potential [as in (3.12)] of this Newtonian fluid. Let ρ, v_i, and α be the density, velocity, and shear potential of a fluid satisfying Einstein's equation such that, for $u = u_0$,

$$\rho \hat{=} \rho_N, \quad v_i \hat{=} v_{N^i}, \quad \text{and} \quad \alpha \hat{=} \alpha_N. \tag{4.4}$$

Does (4.4) remain valid throughout this time interval[27] or does the $\lambda = 0$ limit of the Einsteinian fluid evolve away from its Newtonian counterpart?

The (limit of the) matter evolution equation (4.1) guarantees that the first two conditions of (4.4) remains satisfied. In order to investigate the third condition, we must further analyze (4.2). We expand all λ-dependent quantities in the form

$$F(\lambda) = \sum_0^\infty F_n \lambda^n.$$

Then to next order, (4.2) implies

$$0 = \eth^2\Big[-r\Big(r\int \alpha_{0,0}\Big)_{,1} + (r^2\alpha_1)_{,1}/2 + \beta_1 - (r^2 Z_1)_{,1}/2\Big]. \tag{4.5}$$

In order to satisfy (4.4), α_0 must remain equal to α_N (which comprises two conditions since, in general, α is complex). But the unconstrained data α_1, which occurs in (4.5), has just the right degree of freedom to be able to set $\alpha_{0,0} = \alpha_{N,0}$ for $u = u_0$. Proceeding further $\alpha_{0,00}$ depends upon $\alpha_{1,0}$, which in turn depends upon α_2 according to the next order in the expansion of (4.2). Thus again there is the right freedom in the initial data α_2 to set $\alpha_{0,00} = \alpha_{N,00}$, initially. In this fashion, we are led to a sequence of initial conditions on the α_n, which are formally necessary for the equality $\alpha_0 = \alpha_N$ throughout $u_0 \leqslant u < u_1$. The order n to which these conditions are satisfied determine the degree of tangency between the evolution of the $\lambda = 0$ Einsteinian system and its Newtonian counterpart. From a physical point of view, it seems plausible that the effect of these conditions is to eliminate the presence of incoming radiation in the initial data for the λ-dependent system.[28] The algorithm for prescribing quasi-Newtonian initial data, given in Sec. III, may thus be sharpened by extending it to some α_n level.

Our results warrant further investigations to determine whether this approach can be developed into a practical and rigorous scheme for post-Newtonian calculations. One attractive feature is its intimate connection with the curved space null cone, which allows a simple identification of gravitational radiation.

ACKNOWLEDGMENT

I would like to thank A. I. Janis for many important conversations and for a careful reading of the manuscript.

[1]R. K. Sachs, in *Relativity, Groups, and Topology*, edited by C. DeWitt and B. DeWitt (Gordon and Breach, New York, 1964).
[2]E. Cartan, Ann. Ec. Norm. Sup. **40**, 325 (1923).

[3]K. Friedrichs, Math. Ann. **98**, 566 (1962).
[4]A. Trautman, C. R. Acad. Sci. Paris **257**, 617 (1963).
[5]P. Havas, Rev. Mod. Phys. **36**, 938 (1964).
[6]A Trautman, in *Perspectives in Geometry and Relativity*, edited by B. Hoffman (Indiana U.P., Bloomington, 1966).
[7]G. Dautcourt, Acta Phys. Polon. **25**, 637 (1964).
[8]H. P. Kunzle, Ann. Inst. H. Poincaré **42**, 337 (1972).
[9]H. P. Kunzle, Gen. Rel. Grav. **7**, 445 (1976).
[10]J. Ehlers, in *Grundlagen probleme der Physik* (Festschrift for P. Mittelstaedt), edited by J. Nitsch *et al.* (BI Hochschultaschenbücher, Mannheim, 1981).
[11]There are various hybrid, but physically appealing, treatments of post-Newtonian effects which retain the Minkowski background metric but take slow motion limits in such a way that Newtonian gravity appears to leading order in the equations of motion; cf. the masterful work of S. Chandrasekhar, Astrophs. J. **142**, 1488 (1965). It seems that such treatments might be reformulated, as an expansion in a single parameter λ, with a strict Newtonian limit.
[12]R. A. Isaacson, J. S. Welling, and J. Winicour, "Null cone computation of gravitational radiation," J. Math. Phys. (to be published).
[13]H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Proc. R. Soc. London Ser. A **269**, 21 (1962).
[14]R. K. Sachs, Proc. R. Soc. London Ser. A **270**, 103 (1962).
[15]L. Tamburino and J. Winicour, Phys. Rev. **150**, 1039 (1966).
[16]L. P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, NJ, 1925).
[17]E. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[18]See conventions at the end of Sec. I.
[19]It thus appears that odd powers of λ necessarily arise in order to distinguish between incoming and outgoing radiation. A similar feature involving even and odd powers of "v/c" occurs in other expansion formalisms; cf. S. Chandrasekhar and P. Esposito, Astrophys. J. **160**, 153 (1970).
[20]These are the components of Einstein's equation not involving u derivatives.
[21]This version of the characteristic initial value problem is identical to that of Ref. 15, where the inner world tube is now taken to be the vertex worldline. Details for the axisymmetric case are given in Ref. 12.
[22]It is not necessary, but convenient, to fix the zero of Φ^* at the origin. The spatial gradient of Φ^* must vanish because the origin is freely falling.
[23]Here we assume ρ has compact support. A different asymptotic analysis is necessary for cosmological situations; cf. J. Ehlers and W. Rienstra, Astrophys. J. **155**, 105 (1969).
[24]R. Penrose, Phys. Rev. Lett. **10**, 66 (1963).
[25]The distinction between Φ and Φ^* is irrelevant in determining the shear.
[26]This might have been anticipated on the basis of Ref. 9.
[27]We, of course, assume a common equation of state.
[28]Some justification for this interpretation follows from considering the analogous Newtonian limit of the special relativistic wave equation.

Formal state determination

I. D. Ivanović
Department of Physics and Meteorology, Faculty of Sciences, POB 550, 11000 Belgrade, Yugoslavia

(Received 21 May 1981; accepted for publication 29 April 1982)

Following the general results of Park and Band it is shown that any state of a quantal ensemble allowed by the Hilbert space formulation of nonrelativistic quantum mechanics may be determined from the results of measurements, under the assumption that all measurements, those of energy, position, and spin projection, are in accordance with the projection postulate.

PACS numbers: 05.30.Ch, 03.65.Bz

I. INTRODUCTION

In the standard formulation of nonrelativistic quantum mechanics all properties of a collection of quantal systems are usually described by a statistical operator (state) W, $W \geqslant 0$, $\operatorname{tr} W = 1$. Two distinct behaviors of W are adopted[1]: the evolution, governed by the Hamiltonian H as

$$W(t) = \exp(-iHt)W(0)\exp(iHt); \tag{1}$$

the other occurs in the measurement of an observable $A = \Sigma a_i P_i$, and the premeasurement state W_{PM} changes into an aftermeasurement state

$$W_{AM} = \sum_i P_i W_{PM} W_i = \sum_i w_i W_i, \tag{2}$$

where $w_i = \operatorname{tr}(P_i W_{PM})$ and $W_i = P_i W_{PM} P_i / \operatorname{tr}(P_i W_{PM})$.

In fact, a possibility to assign some particular state W to the collection of quantal systems is a consequence of one's ability to prepare it by means of (2) and (1). A description by means of W is applicable to any number of quantal systems forming a collection. Hence, knowing the state W of a particular collection one is able to explain the future behavior of the inspected collection, using (1) and (2). However, if one wants to predict the future behavior of the collection described by W, or what is equivalent, to test the validity of (2) in some other measurements, e.g., that of $B = \Sigma b_i Q_i$ ($[A,B] \neq 0$), the original number of systems in a collection may become insufficient. Having this in mind, we will use the term ensemble for such collections of quantal systems described by W that will satisfy (2) (with a reasonable accuracy) for any allowed measurement.

The first consequence of the adoption of an ensemble is that a probabilistic behavior of a single quantal system in measurement (2) may be replaced by an almost certain behavior of an ensemble in measurement. The second consequence occurs when it is possible to obtain a collection of ensemble's replicas, each replica in the same state W. Performing different measurements, one measurement on one replica, it is sometimes possible to determine the state W from the results of the measurements.

If we assume that the vector space associated with some quantal system is an n-dimensional complex space $E_n(C)$, a statistical operator (state) W of an ensemble is some $n \times n$, Hermitian complex, non-negative matrix, having the property $\operatorname{tr} W = 1$. Denoting by $V_h^{(n)}$ the real n^2-dimensional vector space of all Hermitian operators acting on $E_n(C)$, the set of all states $V_W^{(n)}$ is a convex subset in $V_h^{(n)}$.

In $V_h^{(n)}$, the scalar product of two elements $A,B \in V_h^{(n)}$ is given by $(A,B) \equiv \operatorname{tr}(AB)$ and if $\{A_k\}_1^{n^2}$ is an operator basis in $V_h^{(n)}$, determination of some unknown state W_u, which is now an n^2-component vector from $V_h^{(n)}$, requires its n^2 components $\{\operatorname{tr}(W_u A_k)\}_{k=1}^{n^2}$.

This was a general idea proposed by Fano[2]; later in the work of Park and Band,[3] a basis of irreducible tensorial operators in $V_h^{(n)}$ was given. In particular, they have shown examples of the state determination for spin $j = \frac{1}{2}, 1$[4,5] and for the one-dimensional spinless particle.[6] They have introduced the term "quorum" of observables for a set of observables constituting a basis in $V_h^{(n)}$.

In order to place the aim of this note properly one should notice what was not completely specified in the works mentioned. If $\{A_k\}_1^{n^2}$ is some particular quorum in $V_h^{(n)}$, an expected value of some observable $A_k = \Sigma_i a_{ki} P_{ki}$ is assumed to be the result of an appropriate measurement. Generally, there are two ways to obtain some $\langle A_k \rangle = \operatorname{tr}(W_u A_k)$. The first one is when $A_k = f(A)$ and $A = \Sigma a_i P_i$ is, say a nondegenerate observable ($\Sigma_i P_i = 1$, $P_i P_j = \delta_{ij} P_i$, $\operatorname{tr}(P_i) = 1, a_i \neq a_j$). The measurement of A will give

$$W_{AM} = \sum_{i=1}^{n} P_i W_{PM} P_i = \sum_{i=1}^{n} w_i P_i \tag{2'}$$

and $\langle A_k \rangle = \Sigma_i f(a_i) w_i$. The second way is when an evolution governed by a precisely known Hamiltonian H [$U(t) = \exp(-iHt)$] may be imposed on the premeasurement state $W_{PM}(t_0)$. Then the measurement of $A = \Sigma a_i P_i$ at some later moment $t_1 > t_0$ will give

$$W_{AM}(t_1) = \sum_{i=1}^{n} P_i \exp(-iHt_1) W_{PM}(t_0) \exp(iHt_1) P_i$$

$$= \sum w_i(t_1) P_i \tag{3}$$

and it may be interpreted as a measurement of $A(t_1) = \Sigma a_i U(t_1) P_i U^+(t_1)$ performed on $W_{PM}(t_0)$. Therefore, if $A_k = f(A(t_1))$, $\langle A_k \rangle_{t_0} = \Sigma f(a_i) w_i(t_1)$. Furthermore, a measurement (2′) or (3) gives $(n-1)$ data about the premeasurement state, and it may occur that only $n' < n-1$ data are useful to calculate the quorum means $\langle A_k \rangle$, making the rest of $n - 1 - n'$ data useless. Thus we arrive at the first question which this note considers, namely, to identify and classify a state-determination procedure in accordance with (2′) and (3) and to identify a minimal set of different measurements, i.e., a minimal quorum, and this is the content of Sec. II.

 0022-2488/83/051199-07$02.50

So far we have assumed that any $A\in V_h^{(n)}$ possesses a measurement procedure [by means of (2′) or (3)] i.e., that all operators are observables, but the situation is usually different and only some of the operators possess a clear measurement procedure, while the status of other operators is usually unclear. This occurs in the set of all states $V_W^{(n)}$, and while for some states a clear preparational procedure exists, at the same time for other, nonpreparable states even a physical interpretation may be unclear. Besides, there is no experimental evidence that such states exist and one may question the validity of (2′) or (3) for such states.

In order to justify a state determination for such cases one should proffer some plausible extension of (2′) and (3) on all states from $V_W^{(n)}$, if it is possible, and then to identify a quorum among the allowed set of measurements, and perhaps a minimal one. This is a second point which this note considers, namely, what are the consequences and what assumptions should be addopted in order to justify a state determination in the case when all operators are not effective observables; this will be discussed in Sec. III.

All that has been mentioned occurs in the examples of the spin systems. For $j=\frac{1}{2}$ all operators are observables, for $j>\frac{1}{2}$, the spin operators are usually assumed to be observables, while the standard formulation allows nonspin operators and nonspin states.[7] The state determination for the $j=1$ case is given by means of (2′) and a continual application of (3),[4] but a very important fact is that recently an interpretation for the nonspin operators and states was given with a suggestion for generalized Stern–Gerlach type measurements,[8] that will make them observable. A natural extension of these results is an explicit proof given in Sec. IV that the state determination for a spin j ensemble can be obtained from the results of not more than $4j+1$ standard Stern–Gerlach type measurements.

To include the case of state determination in an infinite-dimensional space, Sec. V starts with an explicit state determination for a one-dimensional linear harmonic oscillator which is followed by the general solution[6] for one-dimensional spinless particle.

The next aim, to establish a composite system's state determination, is fulfilled in Sec. VI. This section also containes a proof that the state determination is fundamentally dependent on the validity of (2).

Finally, the main conclusion of this note, i.e., that any state allowed by the Hilbert space formulation of nonrelativistic quantum mechanics may be determined from measurements, is briefly discussed in Sec. VII.

II. STATE DETERMINATION: ALL OPERATORS ARE OBSERVABLES

In this section we shall give a brief description of some formal properties of $V_h^{(n)}, V_W^{(n)}$, and the state determination assuming, that all operators are observables.

As was mentioned in Sec. I, the real vector space of all Hermitian operators $V_h^{(n)}=\{A\,|A^+=A\}$ is an n^2-dimensional space. For $A\in V_h^{(n)}$ we will define by P_A the projector $(P^2=P)$ on the range(A). Then $V_h^{(n)}(P)=\{A\,|P_A P=P_A\}$ is a subspace of $V_h^{(n)}$ in which P is the unit operator. If $\operatorname{tr}P=k$ then $\dim(V_h^{(n)}(P))=k^2$; in particular, $V_h^{(n)}(1)=V_h^{(n)}$.

Furthermore, we will denote by $V_h^{(n)}(\{P_i\}_1^n)=\Sigma\oplus V_h^{(n)}(P_i)$, where $P_i^2=P_i$, $\Sigma P_i=1$, and $P_iP_j=\delta_{ij}P_i$, a maximal (n-dimensional) commutative subspace in $V_h^{(n)}$. The set $\{P_i\}_1^n$ is an orthogonal basis set in $V_h^{(n)}(\{P_i\}_1^n)$. Sometimes instead of $\{P_i\}_1^n$, a basis set $\{A^r\}_{r=0}^{n-1}$ $[A=\Sigma_{i=1}^n a_iP_i \in V_h^{(n)}(\{P_i\}_1^n)$ is a nondegenerate observable] will be used.

For $A\in V_h^{(n)}$, $A'=UAU^+(UU^+=1)$ is an orthogonal transformation, and for $P=P^2\in V_h^{(n)}$, $A'=PAP$ is the orthogonal projection of A into $V_h^{(n)}(P)$. Hence, the orthogonal projection of A into $V_h^{(n)}(\{P_i\}_1^n)$ is given by $A'=\Sigma_{i=1}^n P_iAP_i$.

The set of all states $V_W^{(n)}=\{W\,|\,W\geqslant 0,\ \operatorname{tr}W=1\}$ is a convex set in $V_h^{(n)}$. Strictly speaking, it is enough to consider only the convex set of points of $W\in V_W^{(n)}$, lying in the hyperplane $H_0=\{A\,|\operatorname{tr}A=1\}$, where every point represents $\overline{A}=A-(1/n)1$.[9]

If we assume that the convex dimension d of a particular convex set V is equal to the maximal dimension of a simplex[10] $S_{(d)}\subset V$, $V_W^{(n)}$ has the convex dimension n^2-1.

For $W\in V_W^{(n)}$, $\operatorname{tr}P_W=k$ is called rank[11] of W. If $k=1$, W is an extremal element in $V_W^{(n)}$ (pure state), i.e., $W^2=W$. If $k>1$, W is a mixed state. If $k<n$ W is a boundary element in $V_W^{(n)}$, $W\in\partial V_W^{(n)}$. If $k=n$, W is an interior element in $V_W^{(n)}$, $W\in\operatorname{int}V_W^{(n)}$.

The set of all $W\in V_h^{(n)}(\{P_i\}_1^n)$ is a maximal convex set of commutative states, $V_W^{(n)}(\{P_i\}_1^n)$ and it is an $n-1$ dimensional simplex. As in the case of $V_W^{(n)}$, the relevant part is a convex set of points $W\in V_W^{(n)}(\{P_i\}_1^n)$, forming a segment for $n=2$, an equilateral triangle for $n=3$, etc. Hence, $V_W^{(n)}(\{P_i\}_1^n)=\operatorname{conv}(\{P_i\}_1^n)$, being the minimal convex hull of $\{P_i\}_1^n$.

If $W_{PM}\in V_W^{(n)}(\{Q_i\}_1^n)$, and $A\in V_h^{(n)}(\{P_i\}_1^n)$ is a nondegenerate observable, Eqs. (2′), e.g., $W_{AM}=\Sigma_{i=1}^n P_iW_{PM}P_i \in VW^{(n)}(\{P_i\}_1^n)$ is the orthostochastic projection of W_{PM} into $V_h^{(n)}(\{P_i\}_1^n)$. Regarding the eigenvalues of W_{PM} and W_{AM} as two n-dimensional vectors $\{w_{(PM)i}\}_1^n,\{w_{(AM)i}\}_1^n$, then $C\{w_{(PM)i}\}=\{w_{(AM)i}\}$ where C is an orthostochastic matrix[12] and $[C]_{ij}=\operatorname{tr}(P_iQ_j)$.

In these circumstances determination of some $W_u\in V_W^{(n)}$ is possible if one is able to find a set of maximal commutative subspaces $\{V_h^{(n)}(\{P_{(k)i}\}_{i=1}^n)\}_{k=1}^N$ such that $V_h^{(n)}=\cup_{k=1}^N V_h^{(n)}(\{P_{(k)i}\}_{i=1}^n)$. Performing N appropriate measurements, i.e., obtaining N orthogonal projections of W_u into $V_W^{(n)}(\{P_{(k)i}\}_{i=1}^n)$, $W_{AM}^{(k)}=\Sigma_{i=1}^n P_{(k)i}W_uP_{(k)i}$, the unknown state may be expressed as $W_u=\Sigma_{k,i}f(\operatorname{tr}(P_{(k)i}W_u)P_{(k)i}$.

The set of measured nondegenerate observables $\{A_k\}_{k=1}^N$ $(A_k\in V_h^{(n)}(\{P_{(k)i}\}_{i=1}^n))$ will form a quorum for $V_h^{(n)}$. It is easy to see, having in mind that $1\in V_h^{(n)}\{P_i\}_1^n)\cap V^{(n)}(\{P_i'\}_1^n)$ that $N\geqslant n+1$. When $N=n+1$ we say that $\{A_{(k)}\}_1^{n+1}$ is a minimal quorum and in that case eigenprojectors of $\{A_k\}_1^{n+1}$ will form a projector basis set in $V_h^{(n)}$. Some aspects of minimal quorums are discussed in Ref. 13.

Assuming that all A_k possess a direct measurement procedure by means of (2′) one obtains a "kinematical quorum". Another case, i.e, a "dynamical quorum" occurs when an observable $A=\Sigma a_iP_i$ is measured at different moments $t_0,t_1,\ldots,t_N$ assuming that from the moment t_0 the unknown state $W_u(t_0)$ undergoes the evolution as $W_u(t)=\exp(-iH(t-t_0))W_u(t_0)\exp(iH(t-t_0))$ governed by a precisely known $H = \Sigma_k h_k Q_k$. It is equivalent to the measurement of N observables $A(t_k) = U(t_k)AU^+(t_k)$ performed on $W_u(t_0)$, one measurement on one replica of the inspected ensemble.

It is easy to show what conditions should be satisfied by A and H in order that the set $\{A(t_k)\}_{k=1}^N$ becomes a quorum in $V_h^{(n)}$, i.e., in terms of eigenprojectors of $A(t_k)$, that

$$\{1,\{P_i(t_0)\}_{i=1}^{n-1},\{P_i(t_1)\}_{i=1}^{n-1},...,\{P_i(t_N)\}_{i=1}^{n-1}\} \tag{4}$$

contains a projector basis set in $V_h^{(n)}$. Introducing a basis set of elementary $n\times n$, complex Hermitian matrices[9] in $V_h^{(n)}$

$$\sigma^r_{mk} = 2^{-1/2}(e_{km} + e_{mk}),$$
$$\sigma^i_{mk} = i2^{-1/2}(e_{km} - e_{mk}), \quad 1\leqslant m<k\leqslant n,$$
$$\text{and } e_{kk} = Q_k \quad 1\leqslant k\leqslant n, \tag{5}$$

where $[e_{mk}]_{ij} = \delta_{mi}\delta_{kj}$ is an elementary $n\times n$ matrix, one may express all relevant operators in terms of (5) and $H = \Sigma h_k Q_k$

$$P_s(t_0) = \sum_{k=1}^n \alpha^s_m Q_k + \sum_{m,k} (\alpha^m_s \alpha^s_k)^{1/2}(\cos(\varphi_{mk})\sigma^r_{mk} + \sin(\varphi_{mk})\sigma^i_{mk})$$

and $\tag{4'}$

$$P_s(t) = \sum_{k=1}^n \alpha^s_k Q_k + \sum_{m,k} (\alpha^s_m \alpha^s_k)^{1/2}[(\cos(\varphi_{mk} - (h_m - h_k)t)\sigma^r_{mk} + \sin(\varphi_{mk} - (h_m - h_k)t)\sigma^i_{mk}].$$

Therefore, (4) may contain a basis set for $V_h^{(n)}$ only if the $(n-1)\times(n-1)$ matrix $[\alpha^s_k]$ of coefficients from (4') is a nonsingular one, which is equivalent to the nonsingularity of the orthostochastic matrix $[\mathrm{tr}(P_i(t_0)Q_j)]$, so that

$$\det|\mathrm{tr}(P_i(t_0)Q_j)| \neq 0. \tag{6}$$

In fact, we have proven that for any nondegenerate $A\in V_h^{(n)}(\{P_i\}_1^n)$ there exists a nondegenerate $H\in V_h^{(n)}(\{Q_i\}_1^n)$ such that (4) contains a quorum if and only if (6) is satisfied. For example, if eigenvalues of H, $\{h_k\}_1^n$ satisfy $h_m - h_k \neq h_{m'} - h_{k'}$ one may select moments $t_0,t_1,...,t_n$ so that $\{A(t_k)\}_{k=0}^n$ is a minimal dynamical quorum in $V_h^{(n)}$ if (6) is satisfied.

If all operators $A\in V_h^{(n)}$ are observables (as we have assumed in this section) a state determinational procedure may be chosen either as a kinematical, a dynamical, or a mixed quorum. It should be noticed that in any real measurement one is usually unable to make a clear distinction between these cases and, what is more important, that only for $n = 2$, i.e., for the spin $j = \frac{1}{2}$ case, are all operators observables.

III. STATE DETERMINATION: SOME OPERATORS ARE OBSERVABLES

In this section we shall discuss a situation that occurs when only some of the operators from $V_h^{(n)}$ are effective observables with a clear measurement procedure, while the status of other operators is, so to say, unclear.

By $V_{h0}^{(n)} \subseteq V_h^{(n)}$ we will denote the set of observables in $V_h^{(n)}$. Accordingly, we will denote by $V_{WP}^{(n)} = \mathrm{conv}(P^2 = P\in V_{h0}^{(n)})$ a convex set of preparable states. Therefore $V_{WP}^{(n)}$ is given as the minimal convex hull of all eigenprojectors of all observables $A\in V_{h0}^{(n)}$ so that $W\in V_{WP}^{(n)}$ if there exists at least one decomposition $W = w_i P_i, w_i\geqslant 0$, $\Sigma w_i = 1$ such that $P_i\in V_{h0}^{(n)}$ for $i = 1,...,n'$. If $P_iP_j = \delta_{ij}P_i$ W is an orthogonal mixture of preparable states; otherwise it is a nonorthogonal mixture of preparable states. In particular, if $A = \Sigma_{i=1}^n a_i P_i$ is a nondegenerate observable then $V_W^{(n)}(\{P_i\}_1^n)$ is a preparable commutative simplex.

If $W\in V_W^{(n)} - V_{WP}^{(n)} = V_{WNP}^{(n)}$ then it is a nonpreparable state. The usual situation is that one lacks the evidence that such states exist; even a physical interpretation may be unclear. Concerning the state determination, there is no evidence that say (1) or (2') are valid for $W\in V_{WNP}^{(n)}$. In particular, for an assumed premeasurement state $W_{PM}\in V_{WNP}^{(n)}$ one can imagine that the measurement of $A = \Sigma a_i P_i$ will result in an after-measurement state

$$W_{AM} = \sum_i f_i(P_i, W_{PM})P_i, \tag{7}$$

where $f_i(P_i,W_{PM})\geqslant 0$, $\Sigma f_i = 1$, but $f_i(P_i,W_{PM})\neq \mathrm{tr}(P_i W_{PM})$. Hence, an extension of (2') on $V_{WNP}^{(n)}$ will mean finding some formal reasons so that $f_i(P_i,W_{PM}) = \mathrm{tr}(P_i W_{PM})$ for all $W\in V_{WNP}^{(n)}$.

One such way is the following: If $W_1,W_2\in V_{WP}^{(n)}$ and $W_3\in V_{WNP}^{(n)}$ and if

$$W_{PM} = W_1 = \alpha W_3 + (1-\alpha)W_2\in V_{WP}^{(n)}, \tag{8}$$

then in the assumed measurement of $A = \Sigma a_i P_i$,

$$W_{AM} = \sum \mathrm{tr}(P_i W_1)P_i = \alpha\sum f_i(P_i,W_3)P_i + (1-\alpha)\sum \mathrm{tr}(P_i W_2)P_i, \tag{9}$$

and consequently $f_i(P_i,W_3) = \mathrm{tr}(P_i,W_3)$ and (7) is equivalent to (2'), i.e., (2') is extended on $W_3\in V_{WNP}^{(n)}$.

If (8) should be valid for all $W\in V_{WNP}^{(n)}$ it will suffice to prove that it is valid for any pure state $P\in V_{WNP}^{(n)}$. In that case any $W\in V_{WNP}^{(n)}$ will be an affine combination of preparable states [as W_3 is in (8)] and consequently the convex dimension of $V_{WP}^{(n)}$ must be equal to the convex dimension of $V_W^{(n)}$ so that a preparable noncommutative (n^2-1)-dimensional simplex $S_{(n^2-1)}\subset V_{WP}^{(n)}$ exists.

This is formally equivalent to the existence of a quorum $\{A_k\}_1^N, A_k\in V_{h0}^{(n)}$, and then a polytope $\mathrm{conv}(\{P_{(k)i}\}_{k=1,i=1}^{N,n=1})$ will contain an $S_{(n^2-1)}\subset V_{WP}^{(n)}$. If a quorum is a minimal one, the given polytope will become the simplex we need. Therefore, if some $S_{(n^2-1)}\subset V_{WP}^{(n)}$ exists, for any $P\in V_{WNP}^{(n)}$ there exist $W\in S_{(n^2-1)}$ and $\alpha>0$ so that

$$W' = \alpha P + (1-\alpha)W\in S_{(n^2-1)}\subset V_{WP}^{(n)}. \tag{8'}$$

It is also clear that if conv. dim. $V_{WP}^{(n)} < n^2-1$ then there exists a $P\in V_{WNP}^{(n)}$ such that any line through P contains at most only one point from $V_{WP}^{(n)}$ and P is not an affine combination of preparable states.

It is only left to give an interpretation for (9), and it may be justified through an assumption which is foundamental for quantum mechanics:

(A) The result of a particular measurement W_{AM} is independent of the way in which W_{PM} was obtained, i.e., all mixtures resulting in the same W_{PM} will give the same result

W_{AM} in accordance with (2′).

Clearly, (A) may be tested on the set of preparable states and, in particular, the state $W_0 = (1/n)1$ may always serve as a test for (A). If $S_{(n^2-1)} \subset V_{WP}^{(n)}$ then a nontrivial subset of states with inequivalent preparational procedures exists in $V_{WP}^{(n)}$.

First, one must notice that if some $S_{(n^2-1)} \subset V_{WP}^{(n)}$ one may always choose an $S'_{(n^2-1)} \subset V_{WP}^{(n)}$ so that $W_0 \in \operatorname{int} S'_{(n^2-1)}$; on the other hand, in $V_{WP}^{(n)}$ one may always identify a preparable commutative simplex $V_W^{(n)}(\{P_i\}_1^n)$ so that $W_0 \in \operatorname{int}(V_W^{(n)}(\{P_i\}_1^n)$ and $\dim[S'_{(n^2-1)} \cap V_W^{(n)}(\{P_i\}_1^n)] = n-1$. This intersection presents a set of states with inequivalent preparational procedures on which (A) may be tested.

One should be aware of the fact that the assumption (A) may be questioned, and for appropriately choosen W_{PM} and observable A (to be measured), there is a finite probability that inequivalent preparations resulting in the same state W_{PM} may cause different results of measurements.[14] However, in this context these remarks may be neglected.

We will clarify the above-mentioned properties on the example of the spin $j=1$ ensembles.

Assuming that the vector space for the spin $j=1$ description is $E_3(C)$, its real space of Hermitian operators is the nine-dimensional space $V_h^{(3)}$. The set of operators

$$\mathbf{J}\cdot\mathbf{n} = \cos(b)J_z + \sin(b)\cos(a)J_x + \sin(b)\sin(a)J_y = \sum_{m=-1}^{1} P_m(a,b), \tag{10}$$

where $n = (\cos a \sin b, \sin a \sin b, \cos b)$ is a unit vector in the real physical space and $J_i, i = x,y,z$ are standard angular momentum operators (spin operators in this case) is the subset of $V_h^{(3)}$ with the unquestionable physical meaning. This subset is invariant under the standard rotations, the matrix representation of which is given by

$$D(\alpha,\beta,\gamma) = \exp(-i\alpha J_z)\begin{bmatrix} \cos^2(\beta/2) & -\dfrac{\sin\beta}{2^{1/2}} & \sin^2(\beta/2) \\ \dfrac{\sin\beta}{2^{1/2}} & \cos\beta & -\dfrac{\sin\beta}{2^{1/2}} \\ \sin^2(\beta/2) & \dfrac{\sin\beta}{2^{1/2}} & \cos^2(\beta/2) \end{bmatrix} \times\exp(-i\gamma J_z), \tag{11}$$

written in the basis in which J_z is a diagonal and α, β, and γ are the Euler angles of a rotation. There is no loss of generality in taking $\gamma = 0$ and every $\mathbf{n}$ will define a three-dimensional commutative subspace $V_h^{(3)}(a,b)$ and $\{P_m(a,b)\}_{-1}^{1}$ as its basis elements.

The unquestionable physical meaning of the spin operators (10) originated from the fact that any orientation $\mathbf{n} = (a,b)$ of a standard Stern–Gerlach set-up will decompose a premeasurement ensemble of the spin $j=1$ particles into the after-measurement subensembles, each described by $P_m(a,b)$ $(m = -1,0,1)$. Therefore, we will denote by $V_{h0}^{(3)} \subset V_h^{(3)}$ the set of observables, i.e., spin operators.

Consequently, in $V_W^{(3)}$ we will identify the subset of preparable states

$$V_{WP}^{(3)} = \left\{ W \in V_W^{(3)} \mid W = \sum_i w_i W_i ; W_i \in V_W^{(3)}(a,b) \right\},$$

where $V_W^{(3)}(a,b) = V_W^{(3)} \cap V_h^{(3)}(a,b)$ is a two-dimensional preparable, commutative simplex (an equilateral triangel in this case). In particular, if some $W \in V_W^{(3)}(a,b)$ it may be prepared by the appropriate mixing of subensembles emerging from the Stern–Gerlach set-up oriented along $\mathbf{n} = (a,b)$. All other $W \in V_{WP}^{(3)}$ are preparable by means of subensembles emerging from several, differently oriented, Stern–Gerlach set-ups.

By $V_{WNP}^{(3)} = V_W^{(3)} - V_{WP}^{(3)}$ we will denote the set of nonpreparable states and, so far, there is no evidence that any $W \in V_{WPN}^{(3)}$ exists.

For any $W \in V_{WP}^{(3)}$, the standard Stern–Gerlach set-up followed by an appropriate detecting device will perform a measurement of $J(a,b)$. If a premeasurement state $W_{PM} \in V_{WP}^{(3)}$ then

$$W_{AM} = \sum_m P_m W_{PM} P_m = \sum_m \operatorname{tr}(P_m W_{PM})P_m = \sum_m w_m P_m(a,b) \in V_W^3(a,b); \tag{12}$$

the $\langle J(a,b)\rangle = w_1 - w_{-1}$ and $\langle J^2(a,b)\rangle = w_1 + w_{-1}$ are the results of the measurement (one knows that $\Sigma_m w_m = 1$ even before the measurement).

If a premeasurement state is some unknown W_u and if a sufficient number of replicas of the inspected ensemble, each in the same state W_u, are available, the state determination should be achieved from the results of different measurements, i.e, from the orthogonal projections of W_u into a different $V_W^{(3)}(a_i,b_i)$, $i = 1,\dots N$.

Before we identify some simplex $S_{(8)} \subset V_{WP}^{(3)}$ that will allow us to apply (8′) and (9) to this example, we will show one particular relationship between the preparable and unpreparable mixed states from $V_W^{(3)}$.

In $V_W^{(3)}$ every mixed state is a mixture of nonpreparable states.

It suffices to show that any $W \in \partial V_W^{(3)}(a,b)$ belongs to $\operatorname{conv}\{V_{WNP}^{(3)}\}$ [if $W \in V_{WP}^{(3)}$ and $W \notin V_W^{(3)}(a,b)$ then it is already an orthogonal mixture of nonpreparable states].

Any $W \in \partial V_W^{(3)}(a,b)$ may occur as one of the following three cases:

$$W_{1,0}(\alpha,a,b) = \alpha P_1(a,b) + (1-\alpha)P_0(a,b), \tag{13a}$$

$$W_{0,-1}(\alpha,a,b) = \alpha P_0(a,b) + (1-\alpha)P_{-1}(a,b), \tag{13b}$$

$$W_{1,-1}(\alpha,a,b) = \alpha P_1(a,b) + (1-\alpha)P_{-1}(a,b). \tag{13c}$$

If one denotes by $V_W^{(3)}(P_{m+m'} = P_m + P_{m'})$ a four-dimensional subspace in $V_h^{(3)}$ in which $P_{m+m'}$ is the unit operator, it may be represented (as $V_W^{(2)}$ for the spin $j = \frac{1}{2}$ case[11]) as a ball of radius $r = 2^{-1/2}$. Every point on or in the ball represents one $\overline{W} = W - (1/2)P_{m+m'}$ pure states are on the surface, interior points are mixed boundary points of $V_W^{(3)}$ and center represents $P_{m+m'}$. Appropriate subspaces corresponding to (13a)–(13c) are given in Fig. 1. For $V_W^{(3)}(P_{1+0})$ [Fig. 1(a)] the north pole N is $P_1(a,b)$ and the south pole S is $P_0(a,b)$; for $V_W^{(3)}(P_{0,-1})$ [Fig. 1(b)] N is $P_0(a,b)$ and S is $P_{-1}(a,b)$; for $V_W^{(3)}(P_{1,-1})$ [Fig. 1(c)] N is $P_1(a,b)$ and S is $P_{-1}(a,b)$; and points on

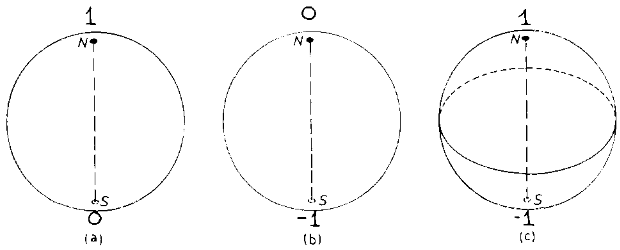

FIG. 1. Boundary states in $V^3_W(a,b)$.

the equator correspond to $P_0(a',b')$ where $(\mathbf{n}(a,b)\cdot\mathbf{n}(a',b'))=0$. Accordingly the set of states [13(a)–13(c)] is represented by the segment $\{N,S\}$ in all three cases. For the intersection $V^{(3)}_{WP}\cap V^3_W(P_{m+m'})$ the situation is slightly different; in Figs. 1(a) and 1(b) this intersection is represented by the segment $\{N,S\}$ and in Fig. 1(c) it is represented by the double cone with tops at N and S, having the equatorial circle as a common base. That all other points correspond to nonpreparable states follows directly from the inspection of (11). Hence, any point from $\{N,S\}$ in Figs. 1(a), 1(b) or from the double cone in Fig. 1(c), corresponding to a mixed state, may be represented as a mixture of nonpreparable pure states lying on the surface of the ball. It is an easy task to apply this result to any mixed preparable state from $V^{(3)}_W$ and we have shown that any preparable mixed state from $V^{(3)}_W$ may be interpreted as a mixture of nonpreparable states.

Finally, in order to extend (12) to all states in $V^{(3)}_W$, adopting (A), it suffices to identify a simplex $S_{(8)}\subset V^{(3)}_{WP}$. In particular, a set of projectors

$$P_1=P_1(0,0),\quad P_2=P_0(0,0),\quad P_3=P_{-1}(0,0),$$
$$P_4=P_0(\pi/4,0),\quad P_5=P_1(\pi/4,0),\quad P_6=P_{-1}(\pi/2,0),$$
$$P_7=P_0(\pi/2,0),\quad P_8=P_0(\pi/4),\ \text{and}\ P_0=P_0(\pi/4,\pi/2)$$

will do, and $S_{(8)}=\operatorname{conv}(\{P_i\}_1^9)$ is an appropriate simplex.

IV. STATE DETERMINATION FOR THE SPIN $j>\frac{1}{2}$ CASE

In this section we will give a particular specification of the general result of Park and Band[5] and it will be shown that a state determination for the spin $j>\frac{1}{2}$ ensembles is possible from the results of not more than $4j+1$ standard Stern–Gerlach type measurements. Because of the adopted type of measurements all operators and mean values will be given in terms of $J_i=J(a_i,b_i)$ and $\langle J_i\rangle$.

We will start with the spin $j=1$ example, introducing an auxiliary operator basis in terms of J_z and $J_\pm=J_x\pm iJ_y$:

$$1, \tag{14a}$$
$$i(J_-+J_+),\quad J_z,\quad J_-+J_+, \tag{14b}$$
$$i(J^2_- - J^2_+);\quad i(J_-J_z-J_zJ_+);$$
$$J_z^2;\ J_-J_z+J_zJ_+;\quad J^2_-+J^2_+, \tag{14c}$$

that allows one to define the following subspaces: $V_0=1$, V_1 will be spanned by (14b) and V_2 by (14c). Denoting by $[J_i^k]_r$ that part of J_i^k that belongs to V_r it is easy to see that only $[J_i^2]_2\neq 0$ and consequently that one needs five different J_i so that $[J_i^2]_2$ may span V_2 and that among $\{J_i\}_1^5$ a span for V_1 exists.

An example of such a set may be

$$J_1=J_z;\ J_2=J_x;\quad J_3=2^{-1/2}(J_z+J_x);$$
$$J_4=2^{-1/2}(J_x+J_y);\quad\text{and}\quad J_5=2^{-1/2}(J_z+J_y), \tag{15}$$

and if $\{\langle J_i\rangle,\langle J_i^2\rangle\}_{i=1}^5$ are the results of the measurements of (15) the unknown premeasurement is represented by the matrix (J_z diagonal)

$$\langle\sigma'_{12}\rangle=2^{-1/2}(\langle J_2\rangle+2\langle J_3^2\rangle-\langle J_1^2\rangle-\langle J_2^2\rangle)$$
$$\langle\sigma'_{13}\rangle=2(\langle J_2^2\rangle+(1/2)\langle J_1^2\rangle-1)$$
$$\langle\sigma'_{23}\rangle=2^{-1/2}(-\langle J_2\rangle+2\langle J_3^2\rangle-\langle J_1^2\rangle\langle J_2^2\rangle)$$
$$\langle\sigma^i_{12}\rangle=2^{-1/2}(\langle J_2\rangle-2^{-1/2}\langle J_4\rangle-2\langle J_5^2\rangle-\langle J_2^2\rangle-2)'$$
$$\langle\sigma^i_{13}\rangle=2(1-\langle J_4^2\rangle-(1/2)\langle J_1^2\rangle)$$
$$\langle\sigma^i_{23}\rangle=2^{-1/2}(\langle J_2\rangle-2^{1/2}\langle J_4^2\rangle+2\langle J_5^2\rangle+\langle J_2^2\rangle-2)$$

$\langle P_{z,1}\rangle=\langle P_1(0,0)\rangle=\frac{1}{2}(\langle J_1^2\rangle+\langle J_1\rangle)$ and $\langle P_0(0,0)\rangle=(1-\langle J_1^2\rangle)$ in terms of (5), assuming, of course, that all results were obtained at the same moment of the ensemble's evolution.

Hence, (15) is a minimal kinematical quorum for $j=1$ ensembles. An interesting consequence is that the measurement of J_x, J_y, and J_z will give a minimal quorum for all $W\in V^3_W(a,b)$ while in general case a minimal quorum cannot contain three mutually orthogonal operators, e.g., J_x, J_y, and J_z, because $P_0(0,0)+P_0(\pi/2,0)+P_0(\pi/2,\pi/2)=1$.

Compared to the general results,[3,5] the measurement of (15) gives, besides five data $\Sigma_m w_m=1$, ten data about the unknown state $\{\langle J_i\rangle,\langle J_i^2\rangle\}$, and it is clear that two data $\langle J_i\rangle$ $i=3,5$ were useless. This excess of two (or seven) unnecessary data is a consequence of the adopted, Stern–Gerlach type measurement.

To obtain a generalization on $V_h^{(2j+1)}$ and $V_W^{(2j+1)}$ we will extend (14) as

$$A_{k0}=J_z^k,\quad 0\leqslant k\leqslant 2j,$$
$$A_{kr}=J^r_-J_z^{(k-r)}+J_z^{(k-r)}J^r_+; \tag{16}$$
$$B_{kr}=i(J^r_-J_z^{(k-r0)}-J_z^{(k-r)}J^r_+)\quad 1<r\leqslant k\leqslant 2j.$$

The subset of (16) for a fixed value of k determines a subspace V_k and $\dim(V_k)=2k+1$. Hence, one should need at least $4j+1$ different J_i in order to span V_{2j}. Denoting by $[J_i^k]_k$ that part of J_i^k that belongs to V_k, a straightforward calculation will give that

$$[J_i^k]_k=\sum_{r=0}^{k}t_{ir}^{(k)}(\cos(ra_i)A_{kr}+\sin(ra_i)B_{kr}), \tag{17}$$

where

$$t_{ir}^{(k)}=(\cos b_i)^k\left(\frac{\tan b_i}{2}\right)^r\binom{k}{r}F\left(\frac{r-k}{2},\frac{r-k+1}{2},r;\frac{\tan^2 b_i}{4}\right);$$

$\binom{k}{r}$ is a binomial coefficient and $F(a,b,c;z)$ is a hypergeometric function.

The state determination is possible if $\{[J_i^{2j}]_{2j}\}_{i=1}^{4j}$ are linearly independent, and that will be the case if the $(4j+1)\times(4j+1)$ determinant of coefficients from (17) is different from zero. Choosing $J_1=J_z$ and $b_i=\text{const}\neq 0,\pi/2$ for $i=1$, that determinant will be proportional to the $4j\times 4j$ determinant

$$\begin{vmatrix} \cos a_2 & \cos a_3 & \cdots & \cos a_{4j+1} \\ \cos 2a_2 & \cdots & & \cos 2a_{4j+1} \\ \vdots & & & \vdots \\ \cos 2ja_2 & \cdots & & \\ \sin a_2 & \sin 2a_3 & \cdots & \sin a_{4j+1} \\ \sin 2a_2 & \cdots & & \sin 2a_{4j+1} \\ \sin 2ja_2 & \cdots & & \sin 2ja_{4j+1} \end{vmatrix} \tag{18}$$

For any j one may select a set $\{a_i\}_2^{4j+1}$ so that (18) is different from zero[15] so that the corresponding set of operators $\{J_i\}_1^{4j+1}$ will become a minimal kinematical quorum in $V_h^{(2j+1)}$. In that case $\text{conv}(\{P_{im}\}_{m=-j,i=1}^{j,\;4j+1})$ will be a convex polytope of states containing a set of preparable noncommutative simplices justifying a state determination procedure for any j.

The existence of fields suggested in Ref. 8 should be tested by means of some state determinational procedure so that the results of this section may occur as not only formally.

V. STATE DETERMINATION FOR THE ONE-DIMENSIONAL SPINLESS PARTICLE

In this section we will reconsider the state determination for a one-dimensional spinless particle[5,6] assuming that the set of observables is x, p and $H = p^2/2 + V(x)$, satisfying $[x,p] = i$. Compared to the spin example, the main differences are that the vector space is infinite-dimensional and that (2) may be valid only for H.

As an "intermediate" case stands the state determination for a linear harmonic oscilator, i.e., $H = (p^2 + x^2)/2 = \Sigma_m (m + \frac{1}{2})P_m$. In the basis where H is diagonal, matrix elements of W_u will be $w_{mn} \exp(\varphi_{mn})$. From the measurement of H one obtains all non-zero diagonal elements $w_{nn} = \text{tr}(P_n W_u)$. An off-diagonal element may be calculated in a following way: one should be able to prepare two replicas of

$$W_u^{(m,n)} = (P_m + P_n) W_u (P_m + P_n), \quad m < n, \tag{19}$$

assuming (19) as an incomplete selective measurement [that of $(P_m + P_n]$. If (19) was prepared at a time t_0 and if x was measured at $t_0 + \Delta t$, then

$$\langle x^k (t_0 + t) \rangle = w_{mm}(x^k)_{mm} + w_{nn}(x^k)_{nn} + 2w_{mn}(x^k)_{mn} \cos(\varphi_{mn} + \Delta t (n - m)).$$

Consequently, from, say, $\langle x^k(t_0)\rangle$ and $\langle x^k(t_0 + \Delta t)\rangle$, w_{mn} and φ_{mn} may be calculated when $k = n - m + 2r$, $r = 0,1,\ldots$.

A more general case $[H \neq (p^2 + x^2)/2]$ will require an entirely different concept, which was given by Park and Band.[6] The set of operators

$$x^m p^n + p^n x^m, \quad i(x^m p^n - p^n x^m) \tag{20}$$

is a particular Hermitian operator basis and appropriate mean values may be calculated from $\langle d^n(x^m)/dt^n \rangle$; hence a state determination is possible from the results of continual measurements of x.

It is obvious that the state determination even in a simple case of an infinite-dimensional space meets two, very serious difficulties. The set of states in not closed and the result of a state determination is dependent on the way in which x is replaced by an actually measured observable with a discrete spectrum. Still, the above-mentioned may be adopted as an approximate state determination. It is also interesting to notice a formal property of coherent states namely, $\langle z | W_u | z \rangle$, will suffice to obtain a state determination.[16]

Compared to the finite-dimensional case we omit the term "preparable state" also the "set of preparable states" due to the fact that (2) is not well defined for a continuous spectrum. The concept of a minimal quorum must also be abandoned so that the conclusion is that a one-dimensional spinless particle possesses an approximate dynamical quorum.

However, these two examples of state determination (for spin and for spinless particle) are very similar in one respect, namely in (16) and (20) one easily recognizes bases for the enveloping algebras of the corresponding algebras [su(2) for spin and Heisenberg algebra for a spinless particle], hence a state determination for a simple system is, in fact, an attempt to identify a basis of observable elements in a relevant subspace of the appropriate representation of the enveloping algebra in question.

At the end of this section we will mention that the recent proposal for a nondemolition quantal measurement[17] is, in fact, a proposal for the state determination of a single one-dimensional spinless particle, not of an ensemble as we have assumed. This is the reason for believing that the problem of state determination may gain some, not only formal, significance, independently of the success of a quantal nondemolition measurement.

VI. STATE DETERMINATION FOR COMPOSITE SYSTEMS

The existence of simple system quorums established in Secs. IV and V allows one to extend a state determination concept to the composite quantal systems. In this section we will consider the case which occurs when it is possible to identify an appropriate subsystem in a unique way.

If two sybsystems S_1 and S_2 are described in e.g., $E_{n_1}(C)$ and $E_{n_2}(C)$, respectively, the composite system $S_1 + S_2$ is usually described in $E_{n_1 n_2} = E_{n_1}(C) \otimes E_{n_2}(C)$. The corresponding space of Hermitian operators $V_h^{(n_1 n_2)}(E_{n_1} \otimes E_{n_2})$ is equivalent to $V_h^{(n_1)}(E_1) \otimes V_h^{(n_2)}(E_2)$, hence any $A \in V_h(E_{n_1 n_2})$ may be decomposed as

$$A = \sum_i a_i A_{(1)i} \otimes A_{(2)i},$$

where

$$\text{tr}(A_{(1)i} A_{(2)j}) = \delta_{ij} = \text{tr}(A_{(2)i} A_{(2)j}).$$

The composite system's set of states $V_W^{(n_1 n_2)}$ has a convex dimension $(n_1 n_2)^2 - 1$ and contains $V_W^{(n_1)} \otimes V_W^{(n_2)}$ as a subset of equal convex dimension. The consequence is that the subsystem quorums, if they exist will constitute a composite-system quorum if the subsystem measurements are performed in coincidence. This means that the measurement of some observable $A_1 \otimes A_{(2)}$ performed on $W_{PM} \in V_W^{(n_1 n_2)}$ will result in

$$W_{AM} = \sum_{k,m} P_{(1)k} \otimes P_{(2)m} W_{PM} P_{(1)k} \otimes P_{(2)m} = \sum_{k,m} w_{km} P_{(1)k} \otimes P_{(2)m},$$

so that $\{A_{(1)i} \otimes A_{(2)j}\}$ will be a composite-system quorum if $\{A_{(1)i}\}$ and $\{A_{(2)j}\}$ are subsystem quorums.

One must also notice that the set of the subsystem's preparable states may become larger, due to the fact that the preparation of a composite-systems state may result in a subsystem state that is a nonpreparable one by means of subsystem procedures. However, this possibility will not affect the set of pure preparable states.

The most important objection against the state determination concept occurs when one tries to replace (2) by a more detailed analysis of a measurement process.[18] Let $|\psi_S\rangle$ be an unknown state of a system and let $|\psi_A\rangle$ be a state of apparatuses. Through an interaction $|\psi_S\rangle \otimes |\psi_a\rangle$ will evolve into $U_{(S+A)}|\psi_S\rangle \otimes |\psi_A\rangle = |\psi_{(S+A)}\rangle$ so that the state of apparatuses will be $W_A = \mathrm{tr}_S(|\psi_{(S+A)}\rangle\langle\psi_{(S+A)}|)$ and the state of the system will be $W_S = \mathrm{tr}_A(|\psi_{(S+A)}\rangle\langle\psi_{(S+A)}|)$. From the measurement one obtains that $W_A = \Sigma w_{(A)i}|\psi_{(A)i}\rangle\langle\psi_{(A)i}|$, and if $U_{(S+A)}$ is exactly known, then

$$U^{+}_{(S+A)}\left(\sum (w_{(a)i})^{1/2}\exp(i\varphi_i)|\psi_{(A)i}\rangle \otimes |\psi_i(\{\varphi_i\})_{(S)}\rangle\right)$$
$$= |\psi_{(A)}\rangle \otimes |\psi(\{\varphi_i\})_{(S)}\rangle$$

will give a set of states $\{|\psi(\{\varphi_i\})_{(S)}\rangle\}$ where the premeasurement state belongs. Repeating a similar procedure for other measurements, a state determination will be effected, but if $U_{(S+A)}$ is unknown the state determination will become an impossible task. Therefore, validity of (2) is the fundamental assumption for this approach to state determination.

We will finish this section with a conclusion that if a subsystem quorum exists the state determination is formally possible for all composite systems from the subsystems measurements performed in coincidence.

VII. CONCLUSIONS

The main conclusion is that any state allowed by a Hilbert space formulation of nonrelativistic quantum mechanics, preparable or nonpreparable, may be determined from the results of measurements. The status of a nonpreparable state, which is either an unimportant consequence of the applied formalism or forbidden by a superselection-type rule, gives two aspects of a state determination. Formally it shows a particular consistency between the applied formalism and the inspected system and in some cases it will give a hint in a search for the "law-breaking" states.

A less important aspect is that the state determination is a quantal counterpart of the term measurement used in a classical sense, which allows us to extract a state determination procedure from any real experiment.

[1]J. von Neumann: *Mathematical Foundations of Quantum Mechanics* (Princeton U. P., Princeton, NJ, 1955), p. 351.
[2]U. Fano, Rev. Mod. Phys. **29**, 74 (1957).
[3]J. L. Park and W. Band, Found. Phys. **1**, 211 (1971).
[4]W. Band and J. L. Park, Found Phys. **1**, 339 (1971).
[5]J. L. Park and W. Band, J. Math. Phys. **14**, 688 (1973).
[6]W. Band and J. L. Park, Am. J. Phys. **47**, 188 (1979).
[7]B. O. Hultgren and A. Shimony, J. Math. Phys. **18**, 381 (1977).
[8]A. R. Swift and R. Wright, J. Math. Phys. **21**, 77 (1980).
[9]J. E. Harriman, Phys. Rev. A **17**, 1249 (1978).
[10]R. T. Rockafellar, *Convex Analysis* (Princeton U. P., Princeton, NJ, 1970) Chap. I.
[11]F. J. Bloore, J. Phys. A **9**, 2059 (1976).
[12]M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities* (Allyn and Bacon, Boston, 1964), Chap. II, Sec. 1.
[13]I. D. Ivanović, J. Phys. A **14**, 3241 (1981).
[14]G. C. Ghirardi, A. Rimini, and T. Weber, Nuovo Cimento **B29**, 135 (1975).
[15]T. Muir, *A Treatise on the Theory of Determinants* (Dover, New York, 1960), p. 358.
[16]J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968), Chap. 7, Sec. 2.
[17]C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, Rev. Mod. Phys. **52**, 341 (1980).
[18]Ref. 1, Chap. VI.

Spinors in two dimensions

Ping Yip
The Enrico Fermi Institute, The University of Chicago, Chicago, Illinois 60637

(Received 1 October 1982; accepted for publication 3 December 1982)

The issue of the meanings of "spinors" in two-dimensional Minkowski spacetime is investigated. Spinors in two dimensions are defined and shown to behave like the familiar spinors in four dimensions. Both of them can be interpreted as square roots of vectors, and both of them describe fermions in a quantum field theory. We also show that an analogous version of the spin-statistics theorem holds in two-dimensional Minkowski spacetime.

PACS numbers: 11.10.Cd, 11.30.Cp

INTRODUCTION

Spinors are important entities in quantum theories in four-dimensional Minkowski space. The mathematical origin of spinors in four dimensions is well known. That the Lorentz group in four dimensions is doubly connected leads us to consider its covering group which is just the SL(2,C) group. Spinors are elements of the two-dimensional complex vector space that forms an irreducible representation of SL(2,C). The role spinors play in quantum physics arises as follows. First, a quantum theory together with a notion of spacetime symmetry as discussed by Wigner[1] leads naturally to considering unitary projective representations of the Poincaré group and naturally of the Lorentz group, as a subgroup of the Poincaré group. In Ref. 1 it was shown that unitary projective representations of the Lorentz group are equivalent to true representations of its covering group SL(2,C). Thus relativistic quantum physics naturally requires SL(2,C), and thus spinors. Wigner[1] further showed that unitary projective representations of the Poincaré group are equivalent to true representations of the inhomogeneous SL(2,C), the covering group of the Poincaré group. It turns out that spinor fields and their tensor products form all the physically relevant unitary representations of the inhomogeneous SL(2,C). Finally, when considering quantum theories of the various fields, we have the well-known spin-statistics theorem: in a reasonable quantum field theory, spinor fields and their odd tensor products describe fermions, while scalar fields and even tensor products of spinor fields describe bosons.

Does an analogous quantity, "spinor," exist in two-dimensional Minkowski spacetime? To attempt to answer the question, we begin by looking at some puzzles. First, the two-dimensional Lorentz group has drastically different structure from its four-dimensional counterpart; namely, it is simply connected (and abelian). Consequently, the usual *raison d'etre* for spinors forming representations of the covering group SL(2,C) of the four-dimensional Lorentz group seems lacking in two dimensions. One might argue, however, that motivated by quantum mechanics, we should consider unitary projective representations of the Lorentz group, and investigate whether we would be led to some suitable notion of spinors, as in the four-dimensional case. But, as we will see in Sec. I, unitary projective representations of two-dimensional Lorentz group are equivalent to true unitary representations. Thus, there does not appear to exist any natural enlargement of the two-dimensional Lorentz group, contrary to the case of four dimensions, where the homogeneous SL(2,C) group forms a natural enlargement, in fact, a universal covering of the Lorentz group.

It is the purpose of this paper to show that, nevertheless, one can define, in two dimensions, "spinors" which have similar properties to those of spinors in four dimensions. They both can be interpreted as square roots of vectors. Most interestingly, in a reasonable quantum field theory, two-dimensional spinor fields do describe fermions. Indeed, more generally, even though the two-dimensional Poincaré group has no rotation subgroup, and hence spin does not exist, we can still prove (Sec. IV) a "spin-statistics" theorem, which is almost identical in content to the usual spin-statistics theorem in four dimensions.

In Sec. I, we will discuss unitary projective representations of the Poincaré and Lorentz groups in two-dimensional Minkowski space. Here and throughout this paper, Poincaré (Lorentz) group means exclusively the *proper* Poincaré (Lorentz) group. In Sec. II, we will define spinor in two dimensions, in connection with the representations of the Lorentz group. We will then, in Sec. III, discuss how the various representations of the Lorentz group give rise to spacetime fields that naturally form unitary representations of the Poincaré group. In Sec. IV, we will state the two-dimensional spin-statistics theorem, and indicate how it is proven.

I. PROJECTIVE REPRESENTATIONS OF THE POINCARÉ GROUP

Spinors are important entities in a quantum theory. So it is satisfying to see how a quantum theory naturally gives rise to spinors. This issue was discussed by Wigner[1] in a broader context. We will only paraphrase his work to make the present discussion self-contained. A quantum theory requires a Hilbert space $\mathscr{H}$ to describe a physical system, and each physical state corresponds to a ray in $\mathscr{H}$. If we further require that this quantum theory be relativistically invariant, then we are naturally led to the requirement that every element of the Poincaré group is to be represented by a unitary operator in $\mathscr{H}$. The composition law, however, of the unitary operators representing the Poincaré group elements need not be such as to render the map between the Poincaré group and the corresponding set of unitary operators a homomorphism. In fact, owing to the ambiguity in phase of

 0022-2488/83/051206-07$02.50

physical states, we have instead, a less stringent requirement—the map between the Poincaré group and its image set of unitary operators should form a projective representation.

More precisely, a projective representation of $\mathscr{P}$ is a map $\mathscr{U}$ from $\mathscr{P}$ into the group of unitary operators of $\mathscr{H}$ such that $\mathscr{U}(\Sigma)\mathscr{U}(\Lambda) = \omega(\Sigma,\Lambda)\mathscr{U}(\Sigma\Lambda)$, where Σ, Λ are elements of the Poincaré group $\mathscr{P}$, $\mathscr{U}(\Sigma)$, and $\mathscr{U}(\Lambda)$ are the corresponding unitary operators in $\mathscr{H}$, and $\omega(\Sigma,\Lambda)$ is a phase factor depending on Σ and Λ. When $\omega = 1$, we have an ordinary, true representation of $\mathscr{P}$. But we emphasize that there is no *a priori* reason for $\omega = 1$, or any particular value for the projective representation to be consistent with the notion of symmetry in a quantum theory.

Wigner[1] first showed that, for the case in a four-dimensional Minkowski space, $\omega(\Sigma,\Lambda)$ can be reduced only to ± 1 by rephasing the unitary operators, ω is 1 only when Σ or Λ or both belong to the translation subgroup; the ambiguity in phase (± 1) essentially comes from the Lorentz group. Thus any unitary projective representation of the four-dimensional Lorentz and Poincaré groups is equivalent to a "double valued" unitary representation of the Lorentz and Poincaré groups. But any double- or single-valued unitary representation of the Lorentz group in four dimensions can be interpreted[2] as a true unitary representation of its covering group SL(2,C). Likewise, any unitary projective representation of the Poincaré group is equivalent to a true unitary representation of its covering group, the inhomogeneous SL(2,C). Thus, although the four-dimensional Minkowski spacetime has $\mathscr{P}$ as a symmetry group, a quantum theory in this spacetime prompts us to consider the unitary representations of the inhomogeneous SL(2,C). Concerning the unitary representations of the inhomogeneous SL(2,C) group, Wigner[1] showed they must be infinite-dimensional. A natural procedure to construct such representations to start from *finite*-dimensional representations of SL(2,C) and construct Hilbert spaces of fields associated with the finite-dimensional representations of SL(2,C). The resulting Hilbert spaces with natural actions of the inhomogeneous SL(2,C) group defined on them will form unitary representation spaces of the inhomogeneous SL(2,C) group. (We will illustrate this process in Secs. II and III, where we construct unitary representations of the Poincaré group in two-dimensional Minkowski space from finite-dimensional representations of the Lorentz group.)

Spinors are elements of a two-dimensional complex vector space which acts as the smallest finite-dimensional nontrivial representation space of the SL(2,C) group. Spinor fields with a natural inner product form a unitary representation of the inhomogeneous SL(2,C) group, as mentioned above. Furthermore, all finite-dimensional representations of SL(2,C) can be interpreted as appropriate tensor product representations of spinor representations. Similarly, all the unitary representations of the inhomogeneous SL(2,C) group by fields associated with the finite-dimensional representation of SL(2,C) (i.e., vector fields, tensor fields, etc.) can be interpreted as product representations by spinor fields. Moreover, all the physically relevant unitary representations of the Poincaré group[1] are just those mentioned, i.e., spacetime fields associated with some finite-dimensional representations of SL(2,C).

Now, we will turn to two-dimensional Minkowski space, and consider unitary projective representations of the two-dimensional Poincaré group. As shown in Ref. 2, all projective representations are equivalent to a one-parameter family of projective representations of the two-dimensional Poincaré group. And this family of projective representations cannot be reduced to true representations of the two-dimensional Poincaré group. However, this one-parameter family arises only in connection with the translation subgroup, while any unitary projective representation of the Lorentz subgroup is equivalent to a true unitary representation of the Lorentz subgroup. This situation is in complete contrast to that in four-dimensions, where the ambiguity in phase ($+$ or -1) comes from the Lorentz and not the translation subgroup. As shown above, we know in four dimensions, spinors arise from the finite-dimensional representations of SL(2,C), which is the natural "enlargement" of the four-dimensional Lorentz group in light of projective representation considerations. Here in two dimensions, since all projective unitary representations of the Lorentz group are equivalent to true representations of the Lorentz group itself, it seems that we should search for "spinors" by considering finite-dimensional representations of the Lorentz group itself. We will do so in the next section, and we will also see how these "spinors" are connected to true unitary representations of the two-dimensional Poincaré group. Although the one-parameter family of projective representations might turn out to be of interest, it seems, however, irrelevant to the question of spinors. We will subsequently concern ourselves only with true unitary representations of the two-dimensional Poincaré group. We wish to point out, nevertheless, that the existence of such a nontrivial family of unitary projective representations indicates that spinors and spinor fields (defined in Secs. II and III) probably would not exhaust all the physically interesting representations of two dimensions.

II. THE FINITE-DIMENSIONAL REPRESENTATION OF THE LORENTZ GROUP

As pointed out in the last section, it is only the true representations of the Lorentz and Poincaré group, not projective representations that concern us henceforth. In this section, we shall investigate the irreducible finite-dimensional representations of the homogeneous subgroup, the Lorentz group, of the Poincaré group in two-dimensional Minkowski spacetime. We will then identify a particular representation as the spinor space for two-dimensional spacetimes. Then, in the next section, we will show how spacetime fields of all the irreducible finite-dimensional representations of the Lorentz group give rise to unitary representations of the Poincaré group. This procedure parallels the case in four dimensions, where, e.g., spinors correspond to a finite-dimensional representation of SL(2,C), and spinor fields form a natural unitary representation of the inhomogeneous SL(2,C).[1]

The Lorentz group $\mathscr{L}$ associated with a two-dimensional Minkowski space, $\mathscr{M}$, is a one-parameter Lie group of isometries of $\mathscr{M}$. Let (x^0,x^1) be the usual coordinate system such that the metric of $\mathscr{M}$ is

$$ds^2 = -dx^{0^2} + dx^{1^2}.$$

Also let $\Lambda(t)$ be a Lorentz transformation $\in\mathscr{L}$ corresponding to a parameter t. Then, we can realize $\Lambda(t)$, an in isometry, $\Lambda(t)$: $\mathscr{M}\to\mathscr{M}$ such that

$$\Lambda(t)\begin{pmatrix}x^0\\x^1\end{pmatrix} = \begin{pmatrix}\cosh t & \sinh t\\ \sinh t & \cosh t\end{pmatrix}\begin{pmatrix}x^0\\x^1\end{pmatrix}. \tag{1}$$

It is obvious that $\Lambda(t)\Lambda(s) = \Lambda(t+s) = \Lambda(s)\Lambda(t)$ and that $\mathscr{L}$ is diffeomorphic to the additive group of the reals $\{\mathbb{R}, +\}$. Thus $\mathscr{L}$ is abelian and simply connected.

To standardize notation, we also define the translation group and the Poincaré group of $\mathscr{M}$. The translation group $\mathscr{J}$ is a two-parameter (a^0,a^1) group of isometries of $\mathscr{M}$. Let $T(a^0,a^1)\in\mathscr{J}$, then $T(a^0,a^1)$: $\mathscr{M}\to\mathscr{M}$ with,

$$T(a^0,a^1)(x^0,x^1) = (x^0 + a^0, x^1 + a^1). \tag{2}$$

The Poincaré group $\mathscr{P}$ is the group of isometries of $\mathscr{M}$ generated by the translation group $\mathscr{J}$ and the Lorentz group $\mathscr{L}$.

We now proceed to consider the finite-dimensional representations of $\mathscr{L}$. Since $\mathscr{L}$ is abelian, all its continuous finite-dimensional complex irreducible representations are one-dimensional. We can then think of the operators representing the group elements in a particularly irreducible representation as continuous complex functions of the group parameter t. Let $f(t)$ be such a function. Then we require $f(t)f(s) = f(s)f(t) = f(t+s)$. The most general solution to this equation is $f(t) = e^{\alpha t}$, where α is a number (complex or real). Thus, each value of α will give rise to an irreducible continuous finite-dimensional representation of by the correspondence $\Lambda(t)\mapsto e^{\alpha t}$. We define $\mathscr{S}^\alpha$ to be the one-dimensional complex vector space, which forms the representation of the α representation of $\mathscr{L}$; i.e., for every t, with $\Lambda(t)\in\mathscr{L}$ as in (1), there corresponds to $\Lambda(t)$ an operator ${}^\alpha\widetilde{\Lambda}(t)$: $\mathscr{S}^\alpha\to\mathscr{S}^\alpha$, such that ${}^\alpha\widetilde{\Lambda}(t)v = e^{\alpha t}v$, $\forall v\in\mathscr{S}^\alpha$.

Now, before identifying what spinors are, we will, instead, first consider the representation connected with vectors. Then we will see how spinors can be viewed as square roots of vectors.

The isometry $\mathscr{L}$ induces an automorphism of the tangent space T_0 at the origin of the coordinate system (x^0,x^1). Thus T_0 forms a finite-dimensional real representation space for $\mathscr{L}$. But T_0 is reducible. In fact, let $r^a\in T_0$ be $(\partial/\partial x^0, \partial/\partial x^1)$, the right-pointing null vector at the origin and $l^a\in T_0$ be $(\partial/\partial x^0, -\partial/\partial x^1)$, the left-pointing one. The mapping $\Lambda(t)\in\mathscr{L}$ will induce an isomorphism $\widetilde{\Lambda}(t)$: $T_0\to T_0$ such that $\widetilde{\Lambda}(t)r^a = e^t r^a$ and $\widetilde{\Lambda}(t)l^a = e^{-t}l^a$. Since any $k^a\in T_0$ can be written as $al^a + br^a$ for some a and b, we see that T_0, as a representation space for $\mathscr{L}$, can be decomposed into a direct sum of two irreducible invariant subspaces generated by r^a and l^a, respectively. Contrary to the four-dimensional case where the tangent space at the origin forms an irreducible representation of the Lorentz group, the representation of by the tangent space T_0 is reducible into the one-dimensional subspaces of right- and left-pointing null vectors. One also easily sees that higher rank tensors at the origin will form representations (generally reducible) corresponding to integral values of α.

Definition: By a spinor space, we will mean either of the vector spaces $S^{1/2}$ or $S^{-1/2}$. By spinors, we mean elements of these spaces.

The relationship between spinor and vector representations in two-dimensional Minkowski space now can be seen. Consider the one-dimensional tensor product space $S^{1/2}\otimes\overline{S}^{1/2}$ where $\overline{S}^{1/2}$ is the complex conjugate vector space of $S^{1/2}$. $\overline{S}^{1/2}$ also forms a $(\alpha = \frac{1}{2})$ representation space of $\mathscr{L}$. Suppose $v\in S^{1/2}\otimes\overline{S}^{1/2}$. Then v can be written as $v = s\otimes\overline{s}'$, where s, $s'\in S^{1/2}$, and $\overline{s}'$, the complex conjugate of s', $\in\overline{S}^{1/2}$. We will see that $S^{1/2}\otimes\overline{S}^{1/2}$ forms a natural product representation space induced by the $S^{1/2}$ representation. Given $\Lambda(t)\in\mathscr{L}$, then we define the operator on $S^{1/2}\otimes\overline{S}^{1/2}$ corresponding to $\Lambda(t)$, as $\widetilde{\Lambda}(t)$ such that $\widetilde{\Lambda}(t)V = ({}^{1/2}\widetilde{\Lambda}(t)s)\otimes({}^{1/2}\widetilde{\Lambda}(t)\overline{s}') = e^{t/2}s\otimes e^{t/2}\overline{s}' = e^t v$, for $v\in S^{1/2}\otimes\overline{S}^{1/2}$ as indicated above. Thus $S^{1/2}\otimes\overline{S}^{1/2}$ forms a representation for $\mathscr{L}$ such that the representation is equivalent to an $\alpha = 1$ representation. In particular, the subspace of $S^{1/2}\otimes\overline{S}^{1/2}$ formed by real multiples of any vector in $S^{1/2}\otimes\overline{S}^{1/2}$ of the form $s\otimes\overline{s}$ for $S\in S^{1/2}$ will form an equivalent representation to the one generated by the space of real multiples of r^a. Similar remarks hold for the space $S^{-1/2}\otimes\overline{S}^{-1/2}$ and the space generated by real multiples of l^a.

The above algebraic property of $S^{1/2}$ and $S^{-1/2}$ makes it reasonable to call them spinor spaces. As is well known in the situation of four-dimensional Minkowski space, if $\mathscr{E}^A$ denotes an SL(2,C) spinor and $\mathscr{E}^{A'}$ its complex conjugate, then $\mathscr{E}^A\mathscr{E}^{A'}$ can be identified as a null vector. Here in two dimensions, if $s\in S^{1/2}$, then $s\otimes\overline{s}\in S^{1/2}\otimes\overline{S}^{1/2}$ can also be identified as r^a, a null vector. Roughly speaking, then, spinors in both two- and four-dimensional Minkowski space are "square roots" of null vectors.

III. UNITARY REPRESENTATIONS OF THE POINCARÉ GROUP

In four-dimensional Minkowski space, an important reason why the finite-dimensional representations associated with the Lorentz group feature so prominently in quantum field theories is that spacetime fields associated with these finite-dimensional representations will naturally form unitary representations of the Poincaré group. The representational spaces formed by these fields will be interpreted as one-particle spaces in a Fock space construction of the Hilbert spaces of the various quantum field theories. Multiparticle spaces correspond to antisymmetric (for fermionic fields) or symmetric (for bosonic fields) tensor products of the one particle Hilbert spaces.

In this section, we will show how various one-dimensional representations of the Lorentz group in two-dimensional Minkowski space corresponding to different α's give rise to spacetime fields which will naturally form unitary representations of the Poincaré group, paralleling the case in four dimensions. But there is a very interesting difference. While in four dimensions, different fields give rise to inequivalent representations, it is not true in two dimensions. We will see that (Lemma 2) all the fields in two dimensions, under suitable restrictions, will form one equivalent representation of the Poincaré group. We will return to this point later after we finish the mathematical preliminaries.

For the following discussion, we will work in momentum space $\mathscr{M}_p$, also a two-dimensional space with Minkowski metric. We will employ a coordinate system (p^0, p^1) such that the metric line element is

$$ds^2 = -(dp^0)^2 + (dp^1)^2. \quad (3)$$

Now we can specify the meaning of the fields associated with the various α representations of $\mathscr{L}$.

Definition: An α field ${}^\alpha\phi$ is a map from $\mathscr{M}_p$ to the vector space S^α.

In particular, when $\alpha = \pm\frac{1}{2}$, ${}^{1/2}\phi$ and ${}^{-1/2}\phi$ are the spinor fields.

Definition: An ${}^\alpha\phi$ has mass μ means that the support of ${}^\alpha\phi$ is restricted to the two mass shells $[(p^0)^2 - (p^1)^2 = \mu^2]$.

The usual spacetime field of $\mathscr{M}$ associated with an ${}^\alpha\phi$ field of mass μ can be obtained by Fourier-transforming the field on $\mathscr{M}_p$:

$${}^\alpha\phi(x) = \int_{\Sigma^+\cup\Sigma^-} e^{ip\cdot x}\,{}^\alpha\phi(p)\frac{dp^2}{|p^0|}, \quad x\in\mathscr{M},\ p\in\mathscr{M}_p, \quad (4)$$

where $p\cdot x = -p^0x^0 + p^1x^1$ and Σ^+ is the positive mass shell $(p^0 \geqslant 0)$ and Σ^- is the negative shell $(p^0 \leqslant 0)$. One readily sees that ${}^\alpha\phi(x)$ satisfies a Klein–Gordon equation of mass μ:

$$(\Box + \mu^2)\,{}^\alpha\phi(x) = 0. \quad (5)$$

Before we proceed further, we will make two remarks. The cases when ${}^\alpha\phi(p)$ has support on timelike surfaces $[(p^0)^2 - (p^2)^2 = -\mu^2 < 0]$ do not appear physically relevant, and we will not consider them here. Also, henceforth, we will only consider fields with real α. As we shall see later, the unitary operator corresponding to a Lorentz group element $\Lambda(t)$, in a representation formed by the ${}^\alpha\phi$ fields, has $e^{\alpha t}$ as a multiplicative factor. Thus if α is complex, we have a phase $e^{i(\mathrm{Im}\,\alpha)t}$ factor appearing in the unitary operator corresponding to the Lorentz group element $\Lambda(t)$. But since, in quantum physics, one has the freedom to rephase the unitary operators representing the Lorentz group,[1] we can thus eliminate the phase $e^{i(\mathrm{Im}\,\alpha)t}$. One might fear that by rephasing the operators, one will end up with a projective representation starting from a true representation. But, it can be seen easily here that, by rephasing as mentioned above, a true representation still remains a true representation. Thus, without loss of generality, we will henceforth deal with real α only.

We now define the appropriate Hilbert spaces formed by the various ${}^\alpha\phi$ fields so as to form unitary representations of the Poincaré group $\mathscr{P}$. We will divide our discussions into two cases, the massive and massless cases. We will first treat the massive case. All μ's are nonzero in the following unless stated.

Definition: Denote by ${}^\alpha_\mu H$ as the collection of ${}^\alpha\phi$ fields of mass $\mu > 0$ such that they are square-integrable with respect to the measure $|(-p^0 + p^1)|^{-2\alpha}dp^1/|p^0|$ on the mass shells Σ^+ and Σ^-, i.e.,

$$\int_{\Sigma^+\cup\Sigma^-} {}^\alpha\phi^*(p)\,{}^\alpha\phi(p)|(-p^0+p^1)|^{-2\alpha}\frac{dp^1}{|p^0|} < \infty,$$

where $*$ denotes complex conjugation.

The inner product on ${}^\alpha_\mu H$ of ${}^\alpha\phi_1$ and ${}^\alpha\phi_2 \in {}^\alpha_\mu H$ is

$$({}^\alpha\phi_1, {}^\alpha\phi_2) = \int_{\Sigma^+\cup\Sigma^-} {}^\alpha\phi_1^*(p)\,{}^\alpha\phi_2(p)|(-p^0+p^1)|^{-2\alpha}\frac{dp^1}{|p^0|}. \quad (6)$$

One sees that with the above inner product ${}^\alpha_\mu H$ is a Hilbert space. We wish to define an action of the Poincaré group on ${}^\alpha_\mu H$, i.e., a map ${}^\alpha_\mu\mathscr{U}$ from the Poincaré group to the set ${}^\alpha_\mu\mathscr{O}$ of linear operators on ${}^\alpha_\mu H$.

We define ${}^\alpha_\mu\mathscr{U}: \mathscr{P}\rightarrow{}^\alpha_\mu\mathscr{O}$ for Lorentz group elements by

$${}^\alpha_\mu\mathscr{U}(\Lambda(t))({}^\alpha\phi(p)) = {}^\alpha\tilde{\Lambda}(t)({}^\alpha\phi(\Lambda(t)(p)), \quad (7)$$

where ${}^\alpha\tilde{\Lambda}(t)$ is as given in Sec. II, $p\in\mathscr{M}_p$, and $\Lambda(t)\in\mathscr{L}\subset\mathscr{P}$. For translations, we define

$${}^\alpha_\mu\mathscr{U}(T(a^0,a^1))({}^\alpha\phi(p)) = e^{ip\cdot a}({}^\alpha\phi(p)), \quad (8)$$

where $p\cdot a = -p^0a^0 + p^1a^1$. In general for any $P\in\mathscr{P}$, if $P = T(a^0,a^1)\Lambda(t)$, then

$${}^\alpha_\mu\mathscr{U}(P) = {}^\alpha_\mu\mathscr{U}(T(a^0,a^1))\,{}^\alpha_\mu\mathscr{U}(\Lambda(t)). \quad (9)$$

Lemma 1: The map ${}^\alpha_\mu\mathscr{U}: \mathscr{P}\rightarrow{}^\alpha_\mu\mathscr{O}$ is a unitary representation of $\mathscr{P}$.

Proof: This is obvious for translations $T(a^0,a^1)$. Let $\Lambda(t)\in\mathscr{L}$. Then we must show that

$$({}^\alpha_\mu\mathscr{U}(\Lambda(t))\,{}^\alpha\phi_1(p), {}^\alpha_\mu\mathscr{U}(\Lambda(t))\,{}^\alpha\phi_2(p)) = ({}^\alpha\phi_1(p), {}^\alpha\phi_2(p)) \quad (10)$$

for ${}^\alpha\phi_1, {}^\alpha\phi_2\in{}^\alpha_\mu H$. First, the left-hand side can be expressed as

$$\int_{\Sigma^+\cup\Sigma^-} ({}^\alpha_\mu\mathscr{U}(\Lambda(t))\,{}^\alpha\phi_1(p))^*({}^\alpha_\mu\mathscr{U}(\Lambda(t))\,{}^\alpha\phi_2(p))|(-p^0+p^1)|^{-2\alpha}\frac{dp^1}{|p^0|}$$
$$= \int_{\Sigma^+\cup\Sigma^-} e^{2\alpha t}\,{}^\alpha\phi_1^*(\Lambda(t)p)\,{}^\alpha\phi_2(\Lambda(t)p)|(-p^0+p^1)|^{-2\alpha}\frac{dp^1}{|p^0|}.$$

Let $q = \Lambda(t)p$; then $-q^0 + q^1 = (-p^0 + p^1)e^{-t}$. The above expression becomes

$$\int_{\Sigma^+\cup\Sigma^-} e^{2\alpha t}\,{}^\alpha\phi_1^*(q)\,{}^\alpha\phi_2(q)|(-q^0+q^1)|^{-2\alpha}e^{-2\alpha t}\frac{dq^1}{|q^0|}$$
$$= \int_{\Sigma^+\cup\Sigma^-} {}^\alpha\phi_1^*(q)\,{}^\alpha\phi_2(q)|(-q^0+q^1)|^{-2\alpha}\frac{dq^1}{|q^0|}$$
$$= ({}^\alpha\phi_1(p), {}^\alpha\phi_2(p)).$$

This proves Lemma 1.

Thus, just as in the four-dimensional case, the various finite-dimensional representations of $\mathscr{L}$ give rise to fields which naturally form unitary representations of $\mathscr{P}$. Interestingly, in contrast to the four-dimensional case, all these unitary representations of the Poincaré group are equivalent.[3] As a reminder, two unitary representations are equivalent if there exists an unitary isomorphism $\mathscr{D}$ from one representation space to another such that the unitary operators corresponding to the Poincaré group on one space are unitarily related by $\mathscr{D}$ to those on the other space. We present the argument in the following lemma.

Lemma 2: Any two representations, ${}^{\alpha}_{\mu}\mathscr{U}: \mathscr{P} \rightarrow {}^{\alpha}_{\mu}\mathscr{O}$, with different α's, are equivalent if the μ's are the same.

Proof: It suffices to show that ${}^{\alpha}_{\mu}\mathscr{U}$ is equivalent to ${}^{\alpha=0}_{\mu}\mathscr{U}$. Let $\mathscr{D}: {}^{\alpha=0}_{\mu}H \rightarrow {}^{\alpha}_{\mu}H$ be such that

$$\mathscr{D}({}^{0}\phi(p)) = {}^{\alpha}\phi(p)|(-p^0+p^1)|^{\alpha}, \tag{11}$$

where ${}^{\alpha}\phi(p) = {}^{0}\phi(p), p \in \mathscr{M}_p$. First we will show that $\mathscr{D}$ is unitary:

$$\begin{aligned}
&(\mathscr{D}({}^{0}\phi_1(p)), \mathscr{D}({}^{0}\phi_2(p))) \\
&\quad = \int_{\Sigma^+ \cup \Sigma^-} [\mathscr{D}({}^{0}\phi_1(p))]^* [\mathscr{D}({}^{0}\phi_2(p))] |(-p^0+p^1)|^{-2\alpha} \frac{dp^1}{|p^0|} \\
&\quad = \int_{\Sigma^+ \cup \Sigma^-} {}^{\alpha}\phi_1^*(p)\, {}^{\alpha}\phi_2(p) \frac{dp^1}{|p^0|} \\
&\quad = ({}^{0}\phi_1(p), {}^{0}\phi_2(p)), \quad {}^{\alpha}\phi_1 \text{ and } {}^{0}\phi_2 \in {}^{0}_{\mu}H.
\end{aligned} \tag{12}$$

Indeed, $\mathscr{D}$ is a unitary map.

Finally, we will see if $\mathscr{D}^{-1}({}^{\alpha}_{\mu}\mathscr{U})\mathscr{D} = {}^{0}_{\mu}\mathscr{U}$. Again, consider $\Lambda(t) \in \mathscr{L}$; the case for translation is trivial. Now

$$\begin{aligned}
&\mathscr{D}^{-1}({}^{\alpha}_{\mu}\mathscr{U}(\Lambda(t)))\mathscr{D}({}^{0}\phi(p)) \\
&\quad = \mathscr{D}^{-1}({}^{\alpha}_{\mu}\mathscr{U}(\Lambda(t)))({}^{\alpha}\phi(p)|(-p^0+p^1)|^{\alpha}) \\
&\quad = \mathscr{D}^{-1} e^{\alpha t} [{}^{\alpha}\phi(\Lambda(t)(p))|(-p^0+p^1)|^{\alpha} e^{-\alpha t}] \\
&\quad = \mathscr{D}^{-1}({}^{\alpha}\phi(\Lambda(t)(p))|(-p^0+p^1)|^{\alpha}) \\
&\quad = {}^{0}\phi(\Lambda(t)(p)) \\
&\quad = {}^{0}_{\mu}\mathscr{U}(\Lambda(t))({}^{0}\phi(p)), \quad \text{with} \quad {}^{0}\phi(p) \in {}^{\alpha=0}_{\mu}H.
\end{aligned} \tag{13}$$

This concludes the proof of Lemma 2.

To complete the discussion, we treat the case where $\mu = 0$. The inner product as used for ${}^{\alpha}_{\mu}H$ when $\mu > 0$ is not well defined where $\mu = 0$. The mass shells for $\mu = 0$ break up into four pieces (see Fig. 1). The new inner product for $\mu = 0$ becomes

$$({}^{\alpha}\phi_1(p), {}^{\alpha}\phi_2(p)) = \int_{\Sigma_R^+ \cup \Sigma_R^-} {}^{\alpha}\phi_1^*(p)\, {}^{\alpha}\phi_2(p)|(p^0+p^1)|^{2\alpha} \frac{dp^1}{|p^0|} + \int_{\Sigma_L^+ \cup \Sigma_L^-} {}^{\alpha}\phi_1^*(p)\, {}^{\alpha}\phi_2(p)|(-p^0+p^1)|^{-2\alpha} \frac{dp^1}{|p^0|}. \tag{14}$$

It can be verified easily that with this inner product, the space ${}^{\alpha}_{\mu=0}H$ of square-integrable functions [i.e., ${}^{\alpha}\phi(p)$ such that $({}^{\alpha}\phi(p), {}^{\alpha}\phi(p)) < \infty$] forms a Hilbert space. Furthermore, by defining the same action of the Poincaré group on ${}^{\alpha}_{\mu=0}H$ as in the massive case, one can show that Lemma 1 and Lemma 2 hold also. Thus, Lemma 1 and 2 hold without restrictions for all μ's.

While the representations of $\mathscr{P}$ by ${}^{\alpha}_{\mu}\mathscr{U}$ are unitary, they are not irreducible. To obtain irreducible representations, we can restrict our map ${}^{\alpha}_{\mu}\mathscr{U}$ as follows. When $\mu > 0$, ${}^{\alpha}_{\mu}H$ is a direct sum of two subspaces ${}^{\alpha}_{\mu^+}H$ and ${}^{\alpha}_{\mu^-}H$, i.e., the subspaces of fields that have support on Σ^+ and that of those having support on Σ^-. It is easy to see that ${}^{\alpha}_{\mu}\mathscr{U}$ acts invariantly on ${}^{\alpha}_{\mu^+}H$ and ${}^{\alpha}_{\mu^-}H$ i.e., ${}^{\alpha}_{\mu}\mathscr{U}(P)({}^{\alpha}_{\mu^+}H) \subseteq {}^{\alpha}_{\mu^+}H$, ${}^{\alpha}_{\mu}\mathscr{U}(P)({}^{\alpha}_{\mu^-}H) \subseteq {}^{\alpha}_{\mu^-}H$, $\forall P \in \mathscr{P}$. In fact, one can easily show that ${}^{\alpha}_{\mu}\mathscr{U}$ does act irreducibly on each ${}^{\alpha}_{\mu^+}H$ and ${}^{\alpha}_{\mu^-}H$. For $\mu = 0$, the space ${}^{\alpha}_{\mu=0}H$ is a direct sum of four subspaces (subspaces of fields having support on Σ_R^+, Σ_R^-, Σ_L^+, and Σ_L^-, respectively). Also in this case, the ${}^{\alpha}_{\mu}\mathscr{U}$ acts irreducibly on each subspace.

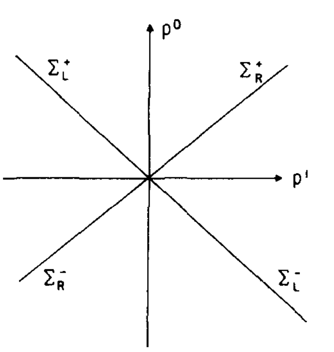

FIG. 1. The four pieces of mass shells when $\mu = 0$, Σ_R^+ ($p^0 = p^1$, $p^0 > 0$), Σ_L^+ ($p^0 = -p^1$, $p^0 > 0$), Σ_R^- ($p^0 = p^1$, $p^0 < 0$), and Σ_L^- ($p^0 = -p^1$, $p^0 < 0$).

We conclude this section with the following observation. The fact spelled out in Lemma 2 marks a sharp contrast between two and four dimensions. In four dimensions, we know that inequivalent finite-dimensional representations of the Lorentz group will induce inequivalent unitary representations of the Poincaré group. But, in two dimensions, all the inequivalent finite-dimensional representations of the Lorentz group give rise to an equivalent—with the proviso of the same mass as in Lemma 2—representation of the Poincaré group. This raises a puzzle of whether it makes sense to note whether there is any field other than the scalar field (i.e., ${}^{\alpha=0}\phi$). In particular, even though spinors as defined in Sec. II make sense in connection with the Lorentz group, a spinor field seems to lose its identity as it may appear as a crooked version of a scalar field. However, this is not the case. In Sec. IV, we will show that scalar fields in a quantum theory describe bosons, while spinor fields describe fermions. We will see that it is the transformation properties of the fields, not to what representations these fields give rise, that determine their statistics, i.e., whether they are fermions or bosons.

This fact that transformation properties, not representational ones, determine the statistics is not at all obvious in the four-dimensional case. There, each inequivalent representation of the Poincaré group induces a different transformation law for the fields; the correspondence between representations and transformation laws is one to one. Thus, it would be difficult, if not impossible, in that case to see that the statistics are determined by the transformation properties, not by representations.

IV. THE "SPIN"-STATISTICS THEOREM IN TWO DIMENSIONS

We will, in this last section, show that spinor fields in a quantum field theory in two-dimensional Minkowski space are indeed fermions. We have, in fact, a more general result: In a quantum field theory of ${}^\alpha\phi(x)$ fields, when α is half-integral, ${}^\alpha\phi(x)$ will be a fermionic field, and, when α is integral, ${}^\alpha\phi(x)$ will be bosonic. Interestingly, when α is neither integral nor half-integral, the ${}^\alpha\phi(x)$ fields are neither fermions nor bosons.

By a quantum field theory of ${}^\alpha\phi(x)$, for any value of α, we mean a theory obeying the Wightman axioms for a field theory,[4] except that Axiom II, describing the transformation properties of fields in four dimensions must be modified to describe instead the transformation properties of fields in two dimensions. Thus, we have, instead, Axiom II′.

Axiom II′: Let $\mathscr{U}(T(a^0,a^1),\Lambda(t))$ be the unitary operator corresponding to the element $T(a^0,a^1)\Lambda(t)$ of the Poincaré group. Then the equation

$$\mathscr{U}(T(a^0,a^1),\Lambda(t))\,{}^\alpha\phi(f)\,\mathscr{U}^{-1}(T(a^0,a^1),\Lambda(t)) = e^{-\alpha t}\,{}^\alpha\phi(\{T(a^0,a^1),\Lambda(t)\}f) \qquad (15)$$

is valid when each side is applied to any vector in D (see Ref. 4). Here ${}^\alpha\phi$ is the quantum field operator and $\{T(a^0,a^1),\Lambda(t)\}f(x) = f(\Lambda^{-1}(t)(x^0-a^0,x^1-a^1)),\ x\in\mathscr{M}$.

One checks, indeed, that all the results leading to the spin-statistics theorem in four dimensions remain true in the two-dimensional context. In fact, we can prove an analogous version of the spin-statistics theorem. We will follow the proof as in Ref. 4 almost verbatim, except for obvious modifications for the context of two dimensions.

Theorem[5,6]: In a quantum field theory of a ${}^\alpha\phi(x)$ field with the quantization rule

$${}^\alpha\phi(x)\,{}^\alpha\phi^*(y) = \theta\,{}^\alpha\phi^*(y)\,{}^\alpha\phi(x), \qquad (16)$$

where $|\theta| = 1$ and $(x-y)^2<0$, i.e., x and y being spacelike separated, if $\theta \neq e^{-2i\alpha\pi}$ then ${}^\alpha\phi(x)\psi_0 = 0$, where ψ_0 is the vacuum state. This further implies ${}^\alpha\phi(x) = 0$.

Proof: The relation (16) gives

$$(\psi_0, {}^\alpha\phi(x)\,{}^\alpha\phi^*(y)\psi_0) - \theta(\psi_0, {}^\alpha\phi^*(y)\,{}^\alpha\phi(x)\psi_0) = 0, \qquad (17)$$

for $(x-y)^2<0$. By the same arguments as in Ref. 4 there are holomorphic functions W and $\widehat{W}$ such that the vacuum expectation values are given by (4-43) of Ref. 4. Generalization of (4-44) of Ref. 4 gives

$$W(\xi) - \theta\widehat{W}(-\xi) = 0. \qquad (18)$$

On the other hand, employing the transformation properties of the fields (Axiom II′), we have

$$\widehat{W}(\xi) = e^{2i\alpha\pi}\widehat{W}(-\xi); \qquad (19)$$

thus 2α essentially replaces the role of J in the proof of the spin-statistics theorem (Theorem 4-10[4]). From (18) and (19), we have

$$W(\xi) - \theta e^{2i\alpha\pi}\widehat{W}(\xi) = 0. \qquad (20)$$

Now, by repeating the same argument as given in the proof of Theorem 4-9[4], we obtain the desired result that ${}^\alpha\phi(x)\psi_0 = 0$. Furthermore, since by definition of a field theory, ψ_0 is a cyclic vector of the polynomial algebra of the ${}^\alpha\phi(x)$ field, this could only be true if the Hilbert space of states consists only of multiples of ψ_0 and ${}^\alpha\phi(x) = 0$. This completes the proof of the theorem.

We remark here that, in general, one would like to consider a theory involving more than one field. In that case, the above theorem still remains valid for the various fields insofar as the conclusion—${}^\alpha\phi(x)\psi_0 = 0$ if the wrong statistics are used—is concerned. To get ${}^\alpha\phi(x) = 0$, we need an extra condition concerning the commutation behaviors between different fields. A sufficient condition is that any two different fields commute via a statistics as in (17) with any value of θ.

Thus one recovers a complete analog of the spin-statistics theorem in two-dimensional Minkowski space, where there is no rotation and hence no spin. For α = integral, the only consistent quantization rule is, from (16), by commutator. For α = half-integral, anticommutator instead should be used. When α is neither half-integral nor integral, the ${}^\alpha\phi(x)$ field does not satisfy either the commutator or the anticommutator; the ${}^\alpha\phi(x)$ field is neither bosonic or fermionic. Instead, it obeys a statistics indicated by (16).

CONCLUSION AND SUMMARY

Even though the Lorentz and the Poincaré group in two-dimensional Minkowski space are simply connected (and for this reason they are drastically different from their four-dimensional counterparts), a representation by "spinors" of the Lorentz group in two dimensions can be defined. Furthermore, spinors in two dimensions do share properties similar to the usual spinors connected with the double-valued representations of the Lorentz group in four dimensions. First, both can be interpreted as "square roots" of vectors. Secondly, just as spinor fields in four dimensions form unitary representations of the inhomogeneous SL(2,C) group, the covering group of the Poincaré group, spinor fields in two dimensions likewise form unitary (and under suitable restrictions, irreducible) representations of the Poincaré group in two dimensions. Most interestingly, spinor fields (in both two and four dimensions) in a quantum field theory, are indeed fermions. In connection with the last point, one in fact has a more general result; there exists an analog of the spin-statistics theorem in two-dimensional Minkowski space, even though the associated Poincaré group does not have a rotation subgroup and hence has no spin.

ACKNOWLEDGMENTS

I would like to thank R. Wald for discussions and guidance. Thanks are also due to S. Fulling for pointing out an error in an earlier draft of the paper. This research was supported in part by NSF Grant PHY80-26043 to the University of Chicago.

[1]E. Wigner, Ann. Math. **40**(1), 149–204 (1939): Proc. Natl. Acad. Sci. USA **34**, 211–23 (1948).

[2]V. Bargmann, Ann. Math. **59**(1), 1–46 (1954).

[3]One can see the equivalence immediately if one employs the Mackey theory of induced representation to construct unitary representations of the Poincaré group. In that case, one finds indeed there is only one representation, namely, that formed by scalar fields. See, e.g., G. W. Mackey, *The Theory of Unitary Group Representations* (The University of Chicago Press, Chicago, 1976); J. M. Jauch, in *Group Theory and Its Applications*, edited by E. M. Loebl (Academic, New York, 1968). We point out that this equivalence implies the absence of a basimir operator for α.

[4]R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Benjamin, Reading, Mass., 1964).

[5]The notation used here follows Ref. 4. Unfortunately, in Ref. 4, α is used as an index for field components, while α here has a totally different meaning.

[6]Though not stated in a general context as the present theorem, similar conclusions were obtained by Klaiber for the case of the Thirring Model. See B. Klaiber, in *Lectures in Theoretical Physics, XA* (Gordon and Breach, New York, 1968).

Canonical structure of the CP_2^{n-1} model in some noncovariant gauges

Ennio Gozzi[a)]

Physics Department, City College of the City University of New York, New York, New York 10031

Arunabha Guha[b)]

Institute for Theoretical Physics, State University of New York, Stony Brook, New York 11794

(Received 20 July 1982; accepted for publication 29 October 1982)

The canonical structure of the CP_2^{n-1} model is derived in four different noncovariant gauges using the Dirac method for constrained Lagrangian systems. A suitable quantization scheme is proposed for all n.

PACS numbers: 11.10Ef, 11.10.Lm, 11.15.Kc

I. INTRODUCTION

The CP_2^{n-1} model[1] is a generalized version of the nonlinear sigma model with a local U(1) invariance. It has aroused the interest of physicists because of its rather remarkable properties. Besides being renormalizable and asymptotically free, it has instanton solutions and allows a systematic $1/n$ expansion.[2] Also, the gauge field, which is a dummy variable at the tree level, acquires a kinetic part (and propagates) due to quantum effects. Classically, like many other two-dimensional models, it has an infinite number of conservation laws.[3] However, what makes it interesting is that, unlike the O(n) nonlinear sigma model[4] or the Gross–Neveu model,[5] pair production is not suppressed since (at least) one of these conservation laws does not survive quantization.[6] All these quantum properties give, in our opinion, a special role to this model and might renew interest in its study. With all this in mind we thought of attempting a Hamiltonian formulation of this model and to study its canonical structure which is the basis from which to start to formulate a canonical quantization.

Having the Hamiltonian might also be useful for a continuous time lattice regularization of the system as has been done for the sigma model.[7] Recent studies[8] of the model on lattice regularization seem to indicate the existence of a new phase and further insight on this phenomenon may come from a lattice Hamiltonian formulation. Our attempt to give a Hamiltonian formulation is not the first; others[9] did it before us. The method of these authors is to solve all the constraints in the system and express everything in terms of the unconstrained variables. However, they succeeded in only one special gauge and the operator ordering problem seems to be rather complicated except in the large n limit.

Our approach is to use the Dirac method[10] and thus not to solve the constraints directly. The advantage is that one gets the canonical structure in different gauges[11] and, in this sense, it is an improvement over the previous attempt. Also, the Hamiltonians in different gauges are rather simple and all the complications due to the constraints are reflected in the Dirac bracket structures.

The authors of Ref. 9 used their unconstrained variable formulation to study the large n limit of the model. What is needed is a consistent quantization scheme for all n. We shall show that in particularly simple gauge (temporal gauge), a quantum Hamiltonian operator can be written down together with commutation relations between the canonical variables. This can be the starting point to derive the Hamiltonian in any other gauge by performing unitary transformations (or quantum canonical transformation). In general, this produces additional terms in the potential.[12] Also, the Hamiltonian and the canonical structure in temporal gauge has been used by the present authors to derive a large n limit by the collective field approach.[13] (See Ref. 14 for further details.)

Worth mentioning is another work[15] where Dirac's method was applied to the supersymmetric version of the CP^{n-1} model. However, the problem of gauge fixing (and therefore, the complete canonical structure) was not discussed in that paper.[15]

The material in this paper is organized as follows. In Sec. II, we introduce the model and derive the constraints by Dirac's procedure. Without fixing the gauge, the primary canonical structure[11] is worked out. We call these "primary Dirac brakcets."

In Sec. III, we derive the final and definitve Dirac brackets by using the primary brackets in three different gauges. Section IV is reserved for a discussion of a temporal gauge (like $A_0 = 0$ gauge in quantum electrodynamics). In the final Sec. V, some conclusions and discussions are presented.

II. PRIMARY DIRAC BRACKET STRUCTURE FOR CP_2^{n-1} MODEL

The model is described by the Lagrangian density

$$\mathscr{L} = (\partial_\mu z_\alpha)(\partial_\mu z_\alpha^*) + (g/2n)(z_\alpha^* \overleftrightarrow{\partial}_\mu z_\alpha)(z_\beta^* \overleftrightarrow{\partial}_\mu z_\beta), \tag{2.1}$$

where n is the number of complex scalar fields z_α and g is the coupling constant. We follow the notation of Ref. 2 which is by now the standard. The fields z satisfy the constraint

$$z^* z = \sum_{\alpha=1}^{n} z_\alpha z_\alpha^* = \frac{n}{2g}. \tag{2.2}$$

The Lagrangian density $\mathscr{L}$ is invariant under a global SU(n) transformation and a local U(1) gauge transformation. In fact, it is easy to see that (2.1) is invariant under the local transformation

---

[a)] Research supported in part by NSF Grant No. PHY 78-24888 and a faculty research award program PSC BHE No. 13735.

[b)] Research supported in part by NSF Grant No. PHY 81-09110.

$$z_\alpha \to e^{i\theta(x,t)} z_\alpha, \quad z^*_\alpha \to z^*_\alpha \, e^{-i\theta(x,t)}. \tag{2.3}$$

We can construct a gauge field A_μ as

$$A_\mu = (ig/n)(z^*_\alpha \overleftrightarrow{\partial}_\mu z_\alpha) \tag{2.4}$$

which transforms under (2.3) as

$$A_\mu \to A_\mu - \partial_\mu \theta, \tag{2.5}$$

and we can write $\mathscr{L}$ as $(D_\mu z)^*(D_\mu z)$ where $D_\mu = \partial_\mu + iA_\mu$. As is well known, the number of real independent field components is $2n-2$. However, we shall enlarge this by writing an equivalent Lagrangian density

$$\begin{aligned}\mathscr{L} = &(\partial_\mu z^*_\alpha)(\partial_\mu z_\alpha) - (2g/n)(z^*_\alpha \partial_\mu z_\alpha)(z_\beta \partial_\mu z^*_\beta)\\ &- \lambda(z^*_\alpha z_\alpha - n/2g),\end{aligned} \tag{2.6}$$

where $\lambda(x)$ is a Lagrange multiplier field that enforces the constraint (2.2) and is to be treated as an independent degree of freedom.

That the Lagrangian in (2.6) is singular can easily be seen by computing the momenta as

$$\begin{aligned}&\pi_\lambda = \frac{\partial \mathscr{L}}{\partial \dot\lambda} = 0,\\ &\pi_\alpha = \frac{\partial \mathscr{L}}{\partial \dot z_\alpha} = M_{\alpha\beta}\, \dot z^*_\beta, \quad \pi^*_\alpha = \frac{\partial \mathscr{L}}{\partial \dot z^*_\alpha} = M^*_{\alpha\beta}\, \dot z_\beta,\\ &M_{\alpha\beta} = \delta_{\alpha\beta} - \frac{2g}{n} z^*_\alpha z_\beta.\end{aligned} \tag{2.7}$$

The matrix M has the properties

$$\begin{aligned}&M = M^+ = M^2,\\ &z_\alpha M_{\alpha\beta} = M_{\alpha\beta} z^*_\beta = 0,\\ &\det M = 0.\end{aligned} \tag{2.8}$$

The last statement follows from the fact that z^* is an eigenvector of M with eigenvalue zero. Thus, the matrix M being noninvertible, the velocities $\dot z_\alpha$ and $\dot z_\alpha$ cannot be completely expressed as functions of momenta and fields. Also, from Eqs. (2.7) and (2.8) one can easily derive the primary[10] constraints.

$$\begin{aligned}&\phi_1 \equiv \pi_\lambda = 0,\\ &\phi_2 \equiv \pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha = 0,\\ &\phi_3 \equiv \pi_\alpha z_\alpha + \pi^*_\alpha z^*_\alpha = 0.\end{aligned} \tag{2.9}$$

The nonvanishing classical Poisson brackets between the various fields and momenta are

$$\begin{aligned}&[z_\alpha(x), \pi_\beta(y)] = [z^*_\alpha(x), \pi^*_\beta(y)] = \delta_{\alpha\beta}\, \delta(x-y),\\ &[\lambda(x), \pi_\lambda(y)] = \delta(x-y).\end{aligned} \tag{2.10}$$

It is easy to see that the constraint ϕ_2 generates the U(1) time-independent gauge transformation as

$$\begin{aligned}&\delta\theta \left[z_\alpha, \int dy\, \phi_2(y) \right] = \delta\theta z_\alpha \equiv \delta z_\alpha,\\ &\delta\theta \left[z^*_\alpha, \int dy\, \phi_2(y) \right] = -\delta z^*_\alpha \equiv \delta z^*_\alpha.\end{aligned} \tag{2.11}$$

From the Lagrangian (2.6), we can derive the canonical Hamiltonian as

$$\begin{aligned}H_c &= \int dx\, (\pi_\alpha(x)\, \dot z_\alpha(x) + \pi^*_\alpha(x)\, \dot z^*_\alpha(x) - \mathscr{L})\\ &= \int dx \Big[\pi_\alpha \pi^*_\alpha + (\partial_1 z^*_\alpha)(\partial_1 z_\alpha)\\ &\quad - \frac{2g}{n}(z_\alpha \partial_1 z^*_\alpha)(z^*_\beta \partial_1 z_\beta) + \lambda\Big(|z|^2 - \frac{n}{2g}\Big)\Big].\end{aligned} \tag{2.12}$$

However, the Hamiltonian H_c is not uniquely determined since one can add the constraints (2.9) with arbitrary multipliers and get a total Hamiltonian H_T (Ref. 10)

$$H_T = H_c + \int dx \sum_{i=1}^{3} v_i \phi_i(x). \tag{2.13}$$

Following Dirac,[10] we now demand that the primary constraints have no time evolution, i.e., the Poisson brackets of ϕ_i with H_T are zero. This gives two new secondary constraints

$$\begin{aligned}&\phi_4 \equiv z^*_\alpha z_\alpha - n/2g = 0,\\ &\phi_5 \equiv \lambda - (2g/n)\,[\, \pi_\alpha \pi^*_\alpha - \partial_1 z^*_\alpha \partial_1 z_\alpha\\ &\qquad + (4g/n)(z_\alpha \partial_1 z^*_\alpha)(z^*_\beta \partial_1 z_\beta)]\,.\end{aligned} \tag{2.14}$$

It should be noted that the constraint ϕ_2 has vanishing Poisson bracket with H_T which is another way of saying that H_T is gauge invariant. The secondary[10] constraints ϕ_4 and ϕ_5 do not generate any further constraints so that the system of constraints is now complete.

The next step is to check which one of these constraints (2.9) and (2.14) is first class[10] and which is second class.[10] From the definition that any first class quantity has vanishing Poisson bracket with all the constraints, we can only see that only ϕ_2 is first class; all the rest are second class. Let us rename the second class constraints χ_i $(i = 1,2,3,4)$ with the identification $\chi_1 = \phi_3, \chi_2 = \phi_5, \chi_3 = \phi_1, \chi_4 = \phi_4$. Then without fixing any gauge we can define the "primary" Dirac brackets[11] as

$$\begin{aligned}&[A,B]' = [A,B]\\ &\qquad - \sum_{i,j=1}^{4} \int dxdy\, [A, \chi_i(x)]\, C^{-1}_{ij}(x,y)[\chi_j(y), B],\\ &C_{ij}(x,y) \equiv [\chi_i(x), \chi_j(y)],\\ &\sum_{j=1}^{4} \int dy\, C_{ij}(x,y)\, C_{jk}(y,z) = \delta_{ik}\, \delta(x-z)\,.\end{aligned} \tag{2.15}$$

From the definition, it is easy to see that only second class constraints will have vanishing primary brackets with any functions of fields and momenta. In order to set all the constraints "strongly zero," we need the final and definitive Dirac brackets[11] which can be obtained iteratively from the primary ones after we have chosen a gauge. This will be done in the next section. For now, we shall only give the primary Dirac bracket structure. The C and C^{-1} matrices can easily be computed as

$$C(x,y)=\begin{bmatrix} 0 & e_{12}(x,y) & 0 & -\dfrac{n}{g}\delta(x-y) \\ -C_{12}(x,y) & 0 & \delta(x-y) & 0 \\ 0 & -\delta(x-y) & 0 & 0 \\ \dfrac{n}{g}\delta(x-y) & 0 & 0 & 0 \end{bmatrix}, \tag{2.16a}$$

where C_{12} has a very complicated structure but we do not need to specify it to calculate [ ]',

$$C^{-1}(x,y)=\begin{bmatrix} 0 & 0 & 0 & \dfrac{g}{n}\delta(x-y) \\ 0 & 0 & -\delta(x-y) & 0 \\ 0 & \delta(x-y) & 0 & \dfrac{g}{n}C_{12}(y,x) \\ \dfrac{-g}{n}\delta(x-y) & 0 & \dfrac{-g}{n}C_{12}(x,y) & 0 \end{bmatrix}, \tag{2.16b}$$

from which the primary Dirac brackets are found to be

$$\begin{aligned}
&[z_\alpha, z_\beta]' = [z_\alpha, z^*_\beta]' = [z^*_\alpha, z^*_\beta]' = 0,\\
&[z_\alpha, \pi_\beta]' = \delta_{\alpha\beta}\,\delta(x-y) - (g/n)\, z_\alpha\, z^*_\beta\, \delta(x-y),\\
&[z_\alpha, \pi^*_\beta]' = -(g/n)\, z_\alpha\, z_\beta\, \delta(x-y),\\
&[\pi_\alpha, \pi_\beta]' = -(g/n)(z^*_\alpha\, \pi_\beta - z^*_\beta \pi_\alpha)\,\delta(x-y),\\
&[\pi_\alpha, \pi^*_\beta]' = (g/n)(\pi_\alpha\, z_\beta - \pi^*_\beta z^*_\alpha)\,\delta(x-y),
\end{aligned} \tag{2.17}$$

where x and y are the arguments of the first and second quantity in the brackets, respectively. We have not written down any bracket with λ or π_λ explicitly just because these are not dynamically significant variables. We also did not bother to calculate the Lagrange multipliers v_i in H_T; they multiply expressions that are "strongly zero" with respect to [ ]' and can be dropped from H_T.

The structure (2.17) is implicit, even if not calculated in detail, in Ref. 15. As we pointed out before, (2.17) does not give the complete canonical structure [see, however, the discussion of temporal gauge in Sec. IV]. The first class constraint ϕ_2 is still to be taken into account. This needs a gauge fixing which we discuss in the next section.

III. GAUGE FIXING

Since we have only one first class constraint ϕ_2, we need only one gauge fixing condition which we call K_1. Defining $K_2 = \phi_2$, we can construct the final and definitive Dirac brackets as

$$[A,B]^* \equiv [A,B]' - \int dx\, dy \times \sum_{i,j=1}^{2} [A,K_i(x)]'\, \tilde{C}_{ij}^{-1}(x,y)\,[K_j(y),B]', \tag{3.1}$$

where $\tilde{C}^{-1}$ is the inverse matrix of

$$\tilde{C}_{ij}(x,y) = [K_i(x), K_j(y)]'. \tag{3.2}$$

Of course, gauge condition K_1 has to be chosen such that $\tilde{C}$ is nonsingular. Various choices of gauge fixing and corresponding Dirac bracket structures are given below.

Case (i). Let us first discuss the choice

$$K_1:\ z_n - z^*_n = 0, \quad \text{unitary gauge.}^9 \tag{3.3}$$

Then $\tilde{C}$ and $\tilde{C}^{-1}$ matrices are

$$\begin{aligned}
\tilde{C}(x,y) &= \begin{bmatrix} 0, & (z_n + z^*_n) \\ -(z_n + z^*_n), & 0 \end{bmatrix} \delta(x-y),\\
\tilde{C}^{-1}(x,y) &= \begin{bmatrix} 0, & -(z_n + z^*_n)^{-1} \\ (z_n + z^*_n)^{-1}, & 0 \end{bmatrix} \delta(x-y).
\end{aligned} \tag{3.4}$$

The final Dirac brackets are

$$\begin{aligned}
&[z_\alpha, z_\beta]^* = [z_\alpha, z^*_\beta]^* = [z^*_\alpha, z^*_\beta]^* = 0,\\
&[z_\alpha, \pi_\beta]^* = (\delta_{\alpha\beta} - (g/n)\, z_\alpha\, z^*_\beta)\,\delta(x-y)\\
&\qquad - (z_n + z^*_n)^{-1} z_\alpha\, \delta_{\beta n}\,\delta(x-y),\\
&[z_\alpha, \pi^*_\beta]^* = (-(g/n)\, z_\alpha\, z_\beta)\,\delta(x-y)\\
&\qquad + (z_n + z^*_n)^{-1} z_\alpha\, \delta_{\beta n}\,\delta(x-y),\\
&[\pi_\alpha, \pi_\beta]^* = (g/n)(\pi_\alpha\, z^*_\beta - z^*_\alpha\, \pi_\beta)\,\delta(x-y)\\
&\qquad - (z_n + z^*_n)^{-1}(\delta_{\alpha n}\, \pi_\beta - \delta_{\beta n}\, \pi_\alpha)\,\delta(x-y),\\
&[\pi_\alpha, \pi^*_\beta]^* = (g/n)(\pi_\alpha\, z_\beta - \pi^*_\beta\, z^*_\alpha)\,\delta(x-y)\\
&\qquad + (z_n + z^*_n)^{-1}(\delta_{\alpha n}\pi^*_\beta - \delta_{\beta n}\, \pi_\alpha)\,\delta(x-y).
\end{aligned} \tag{3.5}$$

On the right-hand side of these brackets we can, of course, set $z_n = z^*_{n*}$ since the structure is consistent with all constraints set "strongly" equal to zero. Finally, we can set the Hamiltonian to be

$$H_T = \int dx \left[\sum_{\alpha=1}^{n} \pi_\alpha\, \pi^*_\alpha + \sum_{\alpha=1}^{n-1} (\partial_1 z_\alpha)(\partial_1 z^*_\alpha) \right] + (\partial_1 z_n)(\partial_1 z_n) - \frac{2g}{n}(z_\beta\, \partial_1 z^*_\beta)(z^*_\gamma\, \partial_1 z_\gamma), \tag{3.6}$$

where we have identified z_n with z_{n*} and set all constraints equal to zero.

Case (ii). Next we shall discuss an axial-like gauge[9]

$$K_1\colon\ z_\alpha\, \partial_1 z_\alpha^* - z_\alpha^*\, \partial_1 z_\alpha = 0. \tag{3.7}$$

The $\tilde{C}$ matrix is now

$$\tilde{C}(x,y) = \begin{pmatrix} 0, & -l(x,y) \\ l(y,x), & 0 \end{pmatrix}, \tag{3.8}$$

where

$$l(x,y) = (z_\alpha^*(x)\, z_\alpha(y) + z_\alpha^*(y)\, z_\alpha(x))\, \partial_x\, \delta(x-y). \tag{3.9}$$

Assuming $\tilde{C}^{-1}$ to be of the form

$$\tilde{C}^{-1}(x,y) = \begin{pmatrix} 0, & f(x,y) \\ -f(y,x), & 0 \end{pmatrix}, \tag{3.10}$$

it is easy to see that f satisfies the equation

$$\partial_x f(z,x) = (g/n)\, \delta(x-z). \tag{3.11}$$

We have to specify the boundary conditions to solve Eq. (3.11). This is not unusual; the same thing also happens in electrodynamics in various noncovariant gauges (see Ref. 11). With appropriate boundary condition the solution is

$$\begin{aligned} &f(y,x) = (g/2n)\, \epsilon(x-y), \\ &\epsilon(x,y) \equiv \mathrm{sign}(x-y). \end{aligned} \tag{3.12}$$

So the definitive Dirac brackets are

$$\begin{aligned}
&[z_\alpha, z_\beta]^* = [z_\alpha, z_\beta^*]^* = [z_\alpha^*, z_\beta^*]^* = 0, \\
&[z_\alpha, \pi_\beta]^* = (\delta_{\alpha\beta} - (2g/n)\, z_\alpha\, z_\beta^*)\, \delta(x-y) \\
&\qquad - (g/n)\, \epsilon(y-x)\, z_\alpha(x)\, \partial_y\, z_\beta^*(y), \\
&[z_\alpha, \pi_\beta^*]^* = (g/n)\, z_\alpha(x)\, \partial_y\, z_\beta(y)\, \epsilon(y-x), \\
&[\pi_\alpha, \pi_\beta]^* = (2g/n)\, (\pi_\alpha\, z_\beta^* - z_\alpha^*\, \pi_\beta)\, \delta(x-y) \\
&\qquad - (g/n)\, \epsilon(x-y)\, [\partial_x\, z_\alpha^*(x)\, \pi_\beta(y) \\
&\qquad + \partial_y\, z_\beta^*(y)\, \pi_\alpha(x)], \\
&[\pi_\alpha, \pi_{\beta *}]^* = (g/n)\, \epsilon(x-y) \\
&\qquad \times [\partial_x\, z_\alpha^*(x)\, \pi_\beta^*(y) + \pi_\alpha(x)\, \partial_y\, z_\beta(y)].
\end{aligned} \tag{3.13}$$

Finally, H_T takes the very simple form

$$H_T = \int dx\, [\pi_\alpha\, \pi_\alpha^* + \partial_1 z_\alpha^*\, \partial_1 z_\alpha]. \tag{3.14}$$

Case (iii). Another interesting case is the radiationlike gauge.

$$K_1 = z_\alpha^*\, \partial_1^2\, z_\alpha - z_\alpha\, \partial_1^2\, z_\alpha^*. \tag{3.15}$$

Proceeding exactly as before, we get

$$\tilde{C}^{-1}(x,y) = \begin{pmatrix} 0, & D(x,y) \\ -D(y,x), & 0 \end{pmatrix}, \tag{3.16}$$

where D satisfies the equation

$$\partial_x^2\, D(z,x) = -(g/n)\, \delta(x-z). \tag{3.17}$$

The solution is

$$D(z,x) = -(g/n)|x-z| + F(x,z), \tag{3.18}$$

where

$$\partial_x^2\, F(x,z) = 0,$$

without specifying any particular boundary condition, we shall indicate the solution of (3.17) as $D(z,x)$. The Dirac brackets are

$$\begin{aligned}
&[z_\alpha, z_\beta]^* = [z_\alpha, z_\beta^*]^* = [z_\alpha^*, z_\beta^*]^* = 0, \\
&[z_\alpha, \pi_\beta]^* = \delta_{\alpha\beta}\, \delta(x-y) - ((g/n)\, \delta(x-y) \\
&\qquad - \partial_y^2\, D(y,x))\, z_\alpha(x)\, z_\beta^*(y) + \{2z_\alpha(x)\, \partial_y\, z_\beta^*(y) \\
&\qquad - (4g/n)\, z_\alpha(x)\, z_\beta^*(y)\, z_\gamma^*(y) + \partial_y\, z_\gamma(y)\}\, \partial_y\, D(y,x), \\
&[z_\alpha, \pi_\beta^*]^* = z_\alpha(x)\, z_\beta(y)\, \{(g/n)\, \delta(x-y) - \partial_y^2\, D(y,x)\} \\
&\qquad\qquad - \{2z_\alpha(x)\, \partial_y\, z_\beta(y)(4g/n)\, z_\alpha(x)\, z_\beta(y)\, z_\gamma(y) \\
&\qquad\qquad \times \partial_y\, z_\gamma^*(y)\}\, \partial_y\, D(y,x), \qquad\qquad (3.19) \\
&[\pi_\alpha(x), \pi_\beta^\pm(y)]^* = (g/n)\, (\pi_\alpha\, z_\beta^\mp - z_\alpha^*\, \pi_\beta^\pm)\, \delta(x-y) \\
&\qquad \pm [z_\alpha^*(x)\, \pi_\beta^\pm(y)\, \partial_x^2\, D(x,y) - \pi_\alpha(x)\, z_\beta^\mp(y)\, \partial_y^2\, D(y,x) \\
&\qquad + 2\pi_\beta^\pm(y)\, \partial_x\, z_\alpha^*(x)\, \partial_x\, D(x,y) \\
&\qquad + 2\pi_\alpha(x)\, \partial_y\, z_\beta \mp (y)\, \partial_y\, D(y,x) \\
&\qquad + (2g/n)\, \pi_\beta^\pm(y)\, z_\alpha(x)(z_\gamma^*\, \overleftrightarrow{\partial} z_\gamma)(x)\, \partial_x\, D(x,y) \\
&\qquad \mp (2g/n)\, \pi_\alpha(x)\, z_\beta \pm (y)(z_\gamma^*\, \overleftrightarrow{\partial} z_\gamma)(y)\, \partial_y\, D(y,x) \\
&\qquad + (2g/n)\, D(x,y)\, \{\pi_\beta^\pm(y)\, \partial_x\, z_\alpha^*(x)(z_\gamma^*\, \partial\, z_\gamma)(x) \\
&\qquad + \pi_\beta^\pm(y)\, \partial_x\, z_\alpha(x)(z_\gamma^*\, \overleftrightarrow{\partial} z_\gamma)(x)\} \\
&\qquad \mp (2g/n)\, D(y,x)\, \{\pi_\alpha(x)\, \partial_y\, z_\beta^\mp(y)(z_\gamma^*\, \overleftrightarrow{\partial} z_\gamma)(y) \\
&\qquad - \pi_\alpha(x)\, \partial_y\, z_\beta^*(y)(z_\gamma^*\, \overleftrightarrow{\partial} z_\gamma(y)\}],
\end{aligned}$$

where we have used the notation $Q \equiv Q^+$ and $Q^* \equiv a^-$ with a any quantity in (3.19). One can see that the Dirac brackets are quite complicated. The Hamiltonian is

$$H_T = \int dx\, \Big[\pi_\alpha\, \pi_\alpha^* + \partial_1 z_\alpha^*\, \partial_1 z_\alpha - \frac{2g}{n}\, (z_\alpha\, \partial_1\, z_\alpha^*)(z_\beta^*\, \partial_1\, z_\beta)\Big]. \tag{3.20}$$

To conclude this section, let us mention that a gauge condition which involves the velocities $\dot{z}$ and $\dot{z}^*$ cannot be directly used in the above formalism. One such example is the temporal gauge which we discuss in the next section.

IV. TEMPORAL GAUGE

Let us now discuss an interesting gauge condition.

$$z_\alpha\, \overleftrightarrow{\partial}_0\, z_\alpha^* = 0 \tag{4.1}$$

which is like the temporal gauge ($A_0 = 0$) in electrodynamics. Using (4.1), the Lagrangian in (2.6) can be written as

$$\begin{aligned} \mathscr{L} &= (\partial_\mu\, z_\alpha)(\partial_\mu\, z_\alpha^*) + (2g/n)\, (z_\alpha^*\, \partial_1\, z_\alpha)(z_\beta\, \partial_1\, z_\beta^*) \\ &\quad - \lambda\, (z_\alpha^*\, z_\alpha - n/2g). \end{aligned} \tag{4.2}$$

This Lagrangian still has time independent U(1) gauge freedom. However, as we shall see, there is no first class constraint since gauge has already been fixed. The momenta are

$$\pi_\alpha = \dot{z}_\alpha^*, \quad \pi_\alpha^* = \dot{z}_\alpha, \quad \pi_\lambda = 0, \tag{4.3}$$

which means that we have only one primary constraint

$$\phi_1 = \pi_\lambda = 0. \tag{4.4}$$

The canonical and total Hamiltonians are

$$H_c = \int dx \Big[\pi_\alpha \pi^*_\alpha - \frac{2g}{n} (z^*_\alpha \partial_1 z_\alpha)(z_\beta \partial_1 z^*_\beta) \Big] + \partial_1 z^*_\alpha \partial_1 z_\alpha + \lambda \Big(|z|^2 - \frac{n}{2g}\Big),$$

$$H_T = H_c + \int dx\, v_\lambda\, \pi_\lambda. \tag{4.5}$$

Proceeding as before, we can derive three secondary constraints:

$$\phi_3 = \pi_\alpha z_\alpha + \pi^*_\alpha z^*_\alpha = 0,$$
$$\phi_4 = z^*_\alpha z_\alpha - n/2g = 0,$$
$$\phi_5 = \lambda - (2g/n) \{ \pi_\alpha \pi^*_\alpha - (\partial_1 z^*_\alpha)(\partial_1 z_\alpha) + (4g/n)(z_\alpha \partial_1 z^*_\alpha)(z^*_\beta \partial_1 z_\beta)\}. \tag{4.6}$$

Thus, $\phi_2 = \pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha = 0$ did not come out as a constraint. Since all four constraints are second class, we immediately see that the Dirac brackets are the same as the primary brackets defined by (2.15) and (2.17), i.e.,

$$[A,B]^* = [A,B]'. \tag{4.7}$$

However, it should be noted that the quantity $\pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha$ has vanishing brackets with all constraints and with H_T

$$[\pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha, \quad \text{any const on } H_T]^* = 0. \tag{4.8}$$

Thus, $(\pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha)$ is a constant of motion. This situation is completely analogous to Gauss' law in quantum electrodynamics in temporal gauge (see Ref. 16). What it means is that it has to be imposed as an initial condition on the physical state. In the quantum theory this would imply that the generator of a time-independent U(1) gauge transformation commutes with the Hamiltonian and so can be simultaneously diagonalized. Thus, a physical state can be written as a simultaneous eigenstate of H and $i(\pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha)$.

$$H|\text{phys}\rangle = E|\text{phys}\rangle,$$
$$i(\pi_\alpha z_\alpha - \pi^*_\alpha z^*_\alpha)|\text{ phys}\rangle = \rho(x)|\text{ phys}\rangle, \tag{4.9}$$

and $\rho(x)$, being time independent, can be interpreted as the "external static U(1) charge." Thus, the temporal gauge is a natural framework for discussing physical states of the CP^{n-1}_2 model with external static charge distributions.

Finally, let us discuss another interesting parametrization of the model in temporal gauge in terms of $2n$ real components of the field Z_α,

$$z_\alpha = (1/\sqrt{2})(q_\alpha + iQ_\alpha), \quad z^*_\alpha = (1/\sqrt{2})(q_\alpha - iQ_\alpha). \tag{4.10}$$

Using the vector $\mathbf{X}$ with $2n$ real components

$$\mathbf{X} = \{ q_1 q_2 \cdots q_n, Q_1 \cdots Q_n \} \tag{4.11}$$

the constraint equation (2.2) becomes

$$\mathbf{X}\cdot\mathbf{X} = \sum_{i=1}^{2n} X_i X_i = \frac{n}{g}. \tag{4.12}$$

Proceeding exactly as before, one can derive the Dirac brackets between X_i and the conjugate momentum P_i

$$[X_i, X_j]^* = 0,$$
$$[X_i, p_j]^* = (\delta_{ij} - (g/n) X_i X_j)\,\delta(x-y),$$
$$[P_i, p_j]^* = -(g/n)(X_i P_j - X_j P_i)\,\delta(x-y). \tag{4.13}$$

However, a more useful variable is the O($2n$) angular momentum J_{ij};

$$J_{ij}(x) = X_i(x) P_j(x) - X_j(x) P_i(x) \tag{4.14}$$

in terms of which the Dirac brackets become

$$[J_{ij}(x), X_k(y)]^* = -(X_i \delta_{jk} - X_j \delta_{ik})\,\delta(x-y),$$
$$[J_{ij}(x), J_{kl}(y)]^* = -[(\delta_{jk} J_{il} - \delta_{ik} J_{jl}) - (k \longleftrightarrow l)]\,\delta(x-y). \tag{4.15}$$

The Hamiltonian in this parametrization is

$$H = \int dx \Big[\frac{g}{4n} \sum_{i,j} J_{ij}(x) J_{ij}(x) + v(x) \Big], \tag{4.16}$$

where the potential energy term is

$$V(x) = \tfrac{1}{2}(\partial X_i)(\partial X_i) - (g/2n)(X_i \sigma_{ij} \partial X_j)^2 \tag{4.17}$$

and σ is a $(2n\times 2n)$ matrix

$$\sigma = \begin{pmatrix} 0 & \mathbb{1}_{n\times n} \\ -\mathbb{1}_{n\times n} & 0 \end{pmatrix}. \tag{4.18}$$

It is easy to see that the kinetic energy term is O($2n$) invariant while the potential energy is only SU(n) invariant. A suitable quantization scheme is to replace

$$[A,B]^* \to -i[A,B]_{\text{commutator}}. \tag{4.19}$$

Since the Hamiltonian operator is unambiguous in this parametrization, one can now perform a quantum canonical transformation to derive the Hamiltonian in other gauges.

We want to discuss, before concluding, two more issues. The first one is the realizability of all these "strange" gauge conditions, that means: given a field configuration $z^\alpha(x,t)$, is it possible to perform a gauge transformation, so to bring it in one of these gauges, for example the transverse one? Remembering that the transverse gauge is $z_\alpha \partial_1 z^*_\alpha - z^*_\alpha \partial_1 z_\alpha = 0$, it might seem quite complicated to find the right transformation. Actually, looking at the definition of gauge field $A_1 \propto z_\alpha \partial_1 z_\alpha{}^* - z^*_\alpha \partial_1 z^\alpha$ and how a gauge field transforms $A'_1 = A_1 - \partial_1\theta$ we see that $z_\alpha \partial_1 z^*_\alpha - z^*_\alpha \partial_1 z_\alpha = 0$ can be satisfied, taking for $\theta(x,t)$,

$$\theta(x,t) = \int_{-\infty}^{x} A_1(y,t)\,dy = \int_{-\infty}^{x} z_\alpha \partial_1 z^*_\alpha - z^*_\alpha \partial_1 z_\alpha)\,dy. \tag{4.20}$$

So we succeed in constructing the gauge transformation that brings our field $z_\alpha(x,t)$ in the transverse gauge. The same can be easily done in all other gauges.

The second point to analyze is to check the Lorentz invariance of our system once the gauge has been chosen. This is very important as all our gauges are noncovariant and the system comes out to be nonmanifestly Lorentz invariant. To make sure that all space-time symmetries are respected, we have to verify the "Schwinger condition."[17] We did not do that in all gauges due to the very complicated structure of our brackets, but we checked that in the $A^0 = 0$ gauge (the only one used in Ref. 14) and there the "Schwinger conditions" are trivially satisfied.

V. CONCLUSIONS AND DISCUSSIONS

We want to point out once more in these conclusions that what we have done: the search for the canonical structure of the CP_2^{n-1} model in different noncovariant gauges, is entirely classical. This is the basis anyway for a consistent canonical quantization in which the Dirac brackets are to be replaced by commutators of different operators. We have also pointed out and used[14] a particularly simple gauge (temporal) for a suitable quantization scheme valid for all n.

Having this structure, an interesting question can be raised which is related to the recently discovered infinite number of classically conserved nonlocal currents. Some of these currents may be anomalous; a direct evaluation of the time dependence of these charges in the framework of a Dirac quantization procedure may give us some insight.

ACKNOWLEDGMENTS

The authors wish to thank Professor B. Sakita for many valuable discussions. One of us (E.G.) is deeply indebted to Professor M. Pauri who introduced him, a long time ago, to the Dirac Hamiltonian method and is also deeply indebted to Dr. G. Rossi for his encouragement.

[1]H. Eichenherr, Nucl. Phys. B **146**, 215 (1978); **155**, 544 (E) (1979); V. Golo and A. Perelomov, Phys. Lett. B **79**, 112 (1978).

[2]A. D'adda, P. DiVecchia, and M. Luscher, Nucl. Phys. B **146**, 63 (1978); **152**, 125 (1978).

[3]M. Luscher and K. Pohlmeyer, Nucl. Phys. B **137**, 46 (1978); E. Brezin, C. Itzykson, J. Zinn-Justin, and J. B. Zuber, Phys. Lett. B **82**, 442 (1979); H. J. De Vega, Phys. Lett. B **87**, 233 (1979); H. Eichenherr and M. Forger, Nucl. Phys. B **155**, 381 (1979); M. K. Prasad, A. Sinha, and L. L. Chau Wang, Phys. Lett. B **87**, 237 (1979); A. T. Ogielski, Phys. Rev. D **21**, 406 (1980); A. T. Ogielski, M. K. Prasad, A. Sinha, and L. L. Chau Wang, Phys. Lett. B **91**, 387 (1980); L. L. Chau Wang, Talk presented at International School of Subnuclear Physics, Erice, Italy, 1980; L. Dolan and A. Roos, Phys. Rev. D **22**, 2018 (1980); B. Y. Hou, M. L. Ge, and Y. Wu, Stony Brook Preprint ITP-SB-81-28; M. L. Ge and Y. Wu, Stony Brook Preprint ITP-SB-81-32.

[4]A. B. Zamolodchikov and A. B. Samolodchikov, Nucl. Phys. B **133**, 525 (1978); M. Luscher, Nucl. Phys. B **135**, 1 (1978).

[5]A. B. Zamolodchikov and A. B. Zamolodchikov, Phys. Lett. B **72**, 481 (1978).

[6]E. Abdalla, M. C. B. Abdalla, and M. Gomes, Phys. Rev. D **23**, 1800 (1981).

[7]C. J. Hamer, J. B. Kogut, and L. Susskind, Phys. Rev. D **19**, 3091 (1979).

[8]E. Rabinovici and S. Samuel, Phys. Lett. B **101**, 323 (1981); P. DiVecchia, A. Holtkamp, R. Musto, F. Nicodemi, and R. Pettorino, Naples preprint, March 1981.

[9]H. E. Haber, J. Hincliffe, and E. Rabinovici, Nucl. Phys. B **172**, 458 (1980).

[10]P. A. M. Dirac, Canad. J. Math. **2**, 129 (1950); **3**, 1 (1951).

[11]A. J. Hanson, T. Regge, and C. Teitelboim, "Constrained Hamiltonian Systems," Accademia Nazionale Dei Lincei, Roma, 1976.

[12]A. Jevicki and J. L. Gervais, Nucl. Phys. B **110**, 93 (1976).

[13]B. Sakita, Phys. Rev. D **21**, 1067 (1980); A. Jevicki and B. Sakita, Nucl. Phys. B **165**, 511 (1980); A. Guha and B. Sakita, Phys. Lett. B **100**, 489 (1981).

[14]"Strings in the CP^{n-1} model," A. Guha, S. C. Lee, and E. Gozzi, Nucl. Phys. B **207**, 111 (1982).

[15]S. Rouhani, Nucl. Phys. B **169**, 430 (1980).

[16]J. L. Gervais and B. Sakita, Phys. Rev. D **18**, 453 (1978).

[17]J. Schwinger, Phys. Rev. **125**, 1043 (1962); **127**, 324 (1962).

Conformal invariance conditions for spinor fields with gauge freedom

Martin Légaré
Centre de Recherche de Mathématiques Appliquées, Université de Montréal, C.P. 6128, Succursale A, Montréal, Canada H3C 3J7

(Received 17 December 1981; accepted for publication 5 March 1982)

The global invariance conditions are given for the (j, j') spinor fields with gauge freedom invariant under subgroups of C(3,1). We formulate two propositions concerning the nonexistence of nontrivial fields with and without compact gauge freedom and apply them to the maximal noncompact subgroups of C(3,1). We also determine explicitly the most general O(4) and $O(2)\times O(4)$ invariant (j, j') spinor fields with compact gauge freedom.

PACS numbers: 11.15. − q

I. INTRODUCTION

Tensor fields and densities invariant under subgroups of the conformal group of space–time C(3,1) were systematically determined by Beckers, Harnad, Perroud, and Winternitz[1] (hereafter noted BHPW). Subsequently, the invariant Dirac spinor fields were investigated with the same methods for subgroups of the Poincaré group P(3,1) and the maximal subgroups of C(3,1) by Beckers, Harnad, and Jasselette[2] (hereafter noted as BHJ). Among these results we recall that BHJ showed there were no nonvanishing Dirac spinor fields invariant under the maximal subgroups of C(3,1).

For physical applications, it is of interest to consider the spinors transforming under some representations of gauge groups. The study of invariant fields was extended to include gauge transformations by Harnad, Shnider, and Vinet[3] and applied to obtain classical solutions to gauge fields equations.[4–6] In the present work, we pursue this line of study to include all spinor representations of the Lorentz group, extended by the definition of a conformal weight (scale factor), and simultaneously all representations of compact gauge groups. The main purpose of finding such invariant spinor fields, knowing that pure Yang–Mills equations are covariant under the full conformal group, is to simplify the equations which describe the coupling of spinor fields with gauge fields. Meetz[7] and Doneux, Saint-Aubin, and Vinet[8] performed this simplification for massless Dirac fields and SU(2) gauge fields invariant under $O(2)\times O(4)$ and some of its subgroups. One important result that emerges from our considerations is that there is an essential distinction to be made between compact and noncompact invariance groups, the latter leading for a wide variety of cases to no nontrivial invariant fields.

In Sec. II, the invariance conditions for spinor fields with gauge freedom are presented in their global (finite) form. The following section contains two propositions which express conditions for the nonexistence of nontrivial invariant (j,j') spinor fields with and without compact gauge transformations. Using these propositions in Sec. IV, we show for any compact gauge group the nonexistence (with some restrictions) of nonvanishing (j,j') spinor fields invariant under any maximal noncompact subgroup of C(3,1) when $j+j'$ is a half-integer and under three maximal noncompact subgroups when $j+j'\in Z$. In Sec. V, we determine explicitly for $O(2)\times O(4)$, the maximal compact subgroup of C(3,1), and for O(4) the most general invariant (j,j') spinor fields with any compact gauge group representation. Finally, in Sec. VI, we summarize our results and suggest future work.

II. INVARIANCE CONDITIONS

In the following, a "gauged spinor field" will mean a spinor field with a gauge freedom, and a "strict invariance," an invariance without any gauge transformation. We make use of the global method described in BHPW and BHJ extended to include gauge freedom. Let M denote Minkowski space with metric g_M of signature $(+\ -\ -\ -)$. M is embedded in a compact space, the so-called compactified Minkowski space $\overline{M}\sim S^1\times S^3/Z_2$, to obtain a global action of C(3,1).[1,9,10] Define on $\overline{M}$ the (j,j') spinor fields Ψ, which on a linear space S (spin space) transforms according to the $D^{(j,j')}$ spinor (finite) representations of the homogeneous Lorentz group. Introducing a gauge freedom, these spinor fields will take values in the tensor product of S with a complex vector space V on which the gauge group H, a compact Lie group, acts by a linear representation $\tilde{D}:H\rightarrow GL(V)$. We then write for a gauge transformation on $\overline{M}$

$$\Psi'(p)=\tilde{D}(h^{-1}(p))\Psi(p), \qquad (2.1)$$

where $p\in\overline{M}$, $h(p)\in H$, and correspondingly $\Psi(p), \Psi'(p)\in S\otimes V$.

Denoting by $x\equiv\{x^\mu(p)\}$, the local coordinates of p, the Jacobian matrix $J_g(x)$ of a conformal transformation $g\in C(3,1)$ relative to any orthogonal frame is given by

$$J_g(x)=e^{f_g(x)}\Lambda_g(x), \qquad (2.2)$$

where $e^{f_g(x)}\in R^+$ (dilatation), and $\Lambda_g(x)\in O(3,1)$. We denote by $\{U_\alpha\}$ an open covering of $\overline{M}$ over which a set of gauges is defined, i.e., trivializations of the principal bundle of the gauge group.

The invariance condition for a (j,j') gauged spinor field is the following[6,8,11,12]:

$$\Psi_\alpha(gx)=e^{lf_g(x)}D^{(j,0)}(\Lambda_g(x)) \otimes D^{(0,j')}(\Lambda_g(x))\otimes\tilde{D}(\rho_\alpha^{-1}(g,x))\Psi_\alpha(x), \qquad (2.3)$$

where $l\in R$ is the scale dimension (conformal weight), $e^{lf_g(x)}\in R^+$, $D^{(j,0)}$, and $D^{(0,j')}$ stand for the finite spinor representations, and the transformation functions, denoted $\rho_\alpha(g,x)$, are functions: $G\times U_\alpha\rightarrow H$, respecting the relation

 0022-2488/83/051219-05$02.50

(composition law) $\rho_\alpha(g_1 g_2, x) = \rho_\alpha(g_2, x)\rho_\alpha(g_1 g_2 x)$ (2.4)

and transforming under a change of gauge on $U_\alpha \cap U_\beta$ with transition functions $k_{\alpha\beta}: U_\alpha \cap U_\beta \to H$ as

$$\rho_\beta^{-1}(g,x) = k_{\alpha\beta}^{-1}(gx)\rho_\alpha^{-1}(g,x)k_{\alpha\beta}(x), \quad (2.5)$$

in order to obtain a global Ψ on $\overline{M}$. More precisely, the domain of definition of ρ_α is only the submanifold $\{(g,x) \in G \times U_\alpha\}$, for which both x and gx belong to $U_\alpha \subset \overline{M}$. For smooth actions this contains a neighborhood of the identity $e \in G$ for all $x \in U_\alpha$.

Two functions ρ satisfying (2.4) and related by (2.5) lead to the same gauged spinor field expressed in different gauges. There exists also a one-to-one correspondence between the nonequivalent transformation functions and the conjugacy classes of homomorphisms $\lambda: G_0 \to H$, obtained by restricting $\rho_\alpha^{-1}(g,x_0)$ at a given reference point x_0 to the isotropy subgroup G_0 (see Ref. 3).

Using (2.3), invariant fields may be uniquely determined on the orbit of a given reference point x_0 from the value $\Psi(x_0)$ which must satisfy the linear isotropy condition:

$$(1 - D({}^{g_0}J(x_0)) \otimes \tilde{D}(\lambda_\alpha(g_0)))\Psi_\alpha(x_0) = 0, \quad (2.6)$$

for all elements $g_0 \in G_0(x_0) \subset G$, where for convenience we have defined

$$D({}^{g_0}J(x_0)) \equiv e^{l f_{g_0}(x_0)} D^{(j,0)}(\Lambda_{g_0}(x_0)) \otimes D^{(0,j')}(\Lambda_{g_0}(x_0)). \quad (2.7)$$

III. PROPOSITIONS

We now formulate two propositions concerning the nonexistence of nontrivial invariant (j,j') spinor fields with and without compact gauge freedom.

Proposition 1: Let Ψ be a (j,j') spinor field with a compact gauge symmetry H. For any orbit and any Lie transformation group G with a noncompact isotropy subgroup G_0 with simple algebra $\tilde{G}_0$, there exist only strictly G-invariant (j,j') spinor fields.

Proof: Since $\tilde{G}_0$ is simple and the kernel of the induced algebra homomorphism $\tilde{G}_0 \to \tilde{H}$ (the gauge algebra) is an ideal, the only algebra homomorphisms are the trivial homorphism or an isomorphism onto a gauge subalgebra. But the latter case is impossible since it would imply that the simple algebra $\tilde{G}_0$ corresponded simultaneously to the noncompact group G_0 and to a compact gauge subgroup. Therefore, the homomorphism λ maps G_0 on $\{e\}$ and (by Proposition 1, Corollary 3 of Ref. 3) $\rho(g,x)$ can always be chosen to be e. The gauge factor disappears in the invariance condition (2.3) and the proposition is proved.

Proposition 2: Assume an orthogonal frame on $\overline{M}$. Let G be a subgroup of C(3,1) and σ any orbit of G in $\overline{M}$. For every one-parameter (t) subgroup of the Jacobian image of the isotropy subgroup of G at $x_0 \in \sigma$ $({}^{g_0}J(x_0): G_0(x_0) \to O(3,1) \times D)$ conjugate to $\exp[(\alpha L_3 + \beta K_3 + \gamma D)t]$ or $\exp[(L_1 - K_2 \pm D)t]$ under $O(3,1) \times D$, where $\alpha, \beta, \gamma, t \in R$ and L_i, K_i, D are, respectively, the rotation, boost, and dilatation generators of $O(3,1) \times D$:

(a) for $\exp[(\alpha L_3 + \beta K_3 + \gamma D)t]$ classes:

(i) If $\alpha \neq 0$ and $j + j' \in$ half-integers, there are no nonvanishing strictly G-invariant (j,j') spinor fields;

(ii) For any α and (j,j'), if $\beta m + \gamma l \neq 0$ [where l is the scale dimension and m ranges over the integers or the half-integers such that $-(j+j') \leqslant m \leqslant (j+j')$], no nonvanishing G-invariant (j,j') gauged spinor fields exist for any compact gauge group.

(b) For $\exp[(L_1 - K_2 \pm D)t]$ classes: If $l \neq 0$, there are no nonvanishing G-invariant (j,j')-gauged spinor fields for all (j,j') for any compact gauge group.

Proof: The isotropy condition (2.6) implies the following condition on Ψ:

$$\Psi = 0, \quad \text{if } \exists\, t \in R \;|\Delta(1 - D(t) \otimes \tilde{D}(t)) \neq 0, \quad (3.1)$$

where $\Delta(\ )$ stands for the determinant and the subscript "α" is dropped for local considerations. Any finite-dimensional (j,j') spinor representations of a one-parameter subgroup of O(3,1) can either be diagonalized (if conjugate to $\alpha L_3 + \beta K_3$) or transformed to a triangular form (if conjugate to $L_1 - K_2$) by a suitable change of basis. The restriction of a linear representation of a compact gauge group acting on V to a one-parameter subgroup is diagonal in a suitable basis and all its eigenvalues are pure phase factors.

Hence with a suitable basis choice we can express (3.1) as

$$\Psi = 0, \quad \text{if } \forall\, i,j \;\; a_i \neq 0 \text{ and/or } (b_i + c_j) \neq 0, \quad (3.2)$$

where $a_i, b_i \in R$ are, respectively, the real and imaginary parts of the ith diagonal element of the one-parameter subgroup generator of the $O(3,1) \times D$ representations within the basis where the (j,j') representation matrices are diagonal or triangular. $c_j \in R$ is the jth eigenvalue of the gauge generator, affecting only the imaginary part of the condition.

For strict invariance: Since $\forall\, j \;\; c_j = 0$, the condition (3.2) reduces to

$$\Psi = 0 \text{ if } \forall\, i \;\; a_i \neq 0 \text{ and/or } b_i \neq 0. \quad (3.3)$$

Looking at the (j,j') representations of the $(\alpha L_3 + \beta K_3 + \gamma D)$ classes: $a_i = \beta m_i + \gamma l$ and $b_i = \alpha n_i$, where $-(j+j') \leqslant m_i \leqslant (j+j')$ and $-(j+j') \leqslant n_i \leqslant (j+j')$ take integer or half-integer values. We thus conclude from (3.3) (i) that $\Psi = 0$ if $\alpha \neq 0$ and $j + j' \in$ half-integers since n_i will never equal zero and that (ii) $\Psi = 0$ for any α and (j,j') if $\beta m_i + \gamma l \neq 0$ for all permissible m_i. In the case of the $(L_1 K_1 \pm D)$ classes we have $a_i = \pm l$ and $b_i = 0$. We deduce from (3.3) that $\Psi = 0$ if $l \neq 0$ for any (j,j').

For invariance up to a gauge transformation: Since we do not know *a priori* the explicit values c_j, we can only conclude from (3.2) that $\Psi = 0$ if $\beta m + \gamma l \neq 0$ for the $(\alpha L_3 + \beta K_3 + \gamma D)$ classes or if $l \neq 0$ for the $(L_1 - K_2 \pm D)$ classes.

By a change of basis, the same results apply for any subgroup G_0 containing in the Jacobian a one-parameter subgroup conjugate to $\exp[(\alpha L_3 + \beta K_3 + gD)t]$ or $\exp[(L_1 - K_2 \pm D)t]$ under $O(3,1) \times D$.

IV. THE MAXIMAL NONCOMPACT SUBGROUPS OF C(3,1)

The maximal noncompact subgroups of C(3,1) are classified by BHPW in eight conjugacy classes: OPT(3,1), SIM(3,1), S(U(2,1)×U(1)), O(3)×O(2,1), O(2,1)×O(2,1), O(3,2), O(4,1) and O(2)×O(2,2). Each class determines only one generic (dense) stratum consisting at most of two open orbits in $\bar{M}$, and all the eight classes possess a noncompact isotropy subgroup on this stratum (for further details, see BHPW).

As an application of our propositions, we investigate the invariant (j,j') spinor fields with compact gauge group and $j+j'$ half-integers for the maximal noncompact subgroups. We also treat the OPT(3,1), SIM(3,1), and $S(U(2,1)\times U(1))$ classes for the $j+j'\in Z$ spinor fields. One point to be stressed is that the exact values attributed to the scale dimension l for the different (j,j') representations can be ignored in the calculus, except for $S(U(2,1)\times U(1))$, if $l\neq m_i\ \forall\ m_i$ permissible and for OPT(3,1), SIM(3,1) if $l\neq 0$.

We summarize the results as follows:

(1) OPT(3,1): At the origin, D is a generator of the isotropy subgroup: $E(2)\times\{3D$ solvable group$\}$. Thus $g_0=e^{tD}$, $J(x_0)\propto e^t 1_4$, and Proposition 2(a) (ii) and (b) imply that there are no nontrivial invariant (j,j') fields (with or without compact gauge freedom) for any (j,j') where $l\neq 0$.

(2) SIM(3,1): At the origin, D is a generator of the isotropy subgroup: $O(1,1)\times O(3,1)$. Thus $g_0=e^{tD}$, $J(x_0)\propto e^t 1_4$, and, like OPT(3,1), Proposition 2(a) (ii) and (b) again imply that there are no nontrivial invariant fields for any (j,j') where $l\neq 0$.

(3) $S(U(2,1)\times U(1))$: At the origin, G_0 and its Jacobian image have a one-parameter subgroup generator $(D-K_3)$ (see BHPW for the explicit form of G_0). In all cases where $m-l\neq 0$, Proposition 2(a) (ii) is applied. Thus there exist no nontrivial invariant fields for any (j,j') where $l\neq m_i\ \forall\ m_i$ permissible.

(4) $SO(3)\times SO(2,1)$: The isotropy subgroup G_0 is the abelian group $SO(2)\times SO(1,1)$. At the reference point $x_0=(0,1,0,0)$, the Jacobian of SO(1,1) (generator P_0-C_0), $J(x_0)\propto e^{K_1 t}$ and Proposition 2(a) (ii) imply that no nontrivial invariant fields exist for $j+j'$ half-integers (since $\exists\ m=0$ if $j+j'\in Z$).

(5) $SO(2,1)\times SO(2,1)$: G_0 is equal to $SO(2)\times SO(1,1)$. At the reference point $x_0=(0,1,0,0)$, the Jacobian of SO(1,1) (generator K_2), $J(x_0)\propto e^{K_2 t}$, and Proposition 2(a) (ii) again imply that there are no nontrivial invariant fields for $j+j'$ half-integers.

(6) $O(2)\times O(2,2)$, O(4,1), and O(3,2): The isotropy subgroups are respectively the simple groups O(2,1), O(3,1), and O(3,1). Using Proposition 1, we conclude that only nonvanishing strictly invariant spinor fields can exist. But at the origin, the simple isotropy subgroups all possess the generator L_3. Thus, applying Proposition 2(a) (i), we find that there are no nontrivial $j+j'$ half-integer invariant fields.

V. THE MAXIMAL COMPACT SUBGROUP OF C(3,1): $O(2)\times O(4)$

We now consider the $O(2)\times O(4)$ case for any (j,j') spinor field with any compact gauge group H. Since $\bar{M}(S^1\times S^3/Z_2)$, up to a Z_2 factor, is identifiable with $U(1)\times SU(2)$, natural group actions on a point $p_0=(e^{i\psi},v)\in\bar{M}$, where $e^{i\psi}\in U(1)$ and $v\in SU(2)$ are:

(a) left action of SU(2):

$$L_{g'}:(e^{i\psi},v)\to(e^{i\psi},g'v);\quad g'\in SU(2) \tag{5.1}$$

(b) right action of SU(2):

$$R_g:(e^{i\psi},v)\to(e^{i\psi},vg);\quad g\in SU(2) \tag{5.2}$$

(c) left and right actions of U(1):

$$L_\phi(e^{i\psi},v)=R_\phi(e^{i\psi},v)=e^{i(\psi+\phi)},v);\quad e^{i\phi}\in U(1) \tag{5.3}$$

(d) left action of $SU(2)_L\times SU(2)_R$:

$$L_{(g',g)}:(e^{i\psi},v)\to(e^{i\psi},g'vg^{-1}); \tag{5.4}$$

(e) diagonal SU(2) subgroup action $SU(2)_D=(SU(2)_L\otimes SU(2)_R)_D$:

$$D_g:(e^{i\psi},v)\to(e^{i\psi},gvg^{-1}). \tag{5.5}$$

Hence, up to a Z_2 factor, $O(2)\times O(4)\sim U(1)\times SU(2)_L\times SU(2)_R$, and the orbit of any $p_0\in\bar{M}$ under $O(2)\times O(4)$ corresponds to $\bar{M}$ itself. The two inequivalent homomorphisms λ of the $O(2)\times O(4)$ isotropy subgroup $G_0=SU(2)_D$ into the gauge group SU(2) can be chosen as: $\lambda((g,g))=e$ and $\lambda'((g,g))=g$. They induce the respective transformation functions $\rho((e^{i\psi}(g,g)),x)=e$ and $\rho'((e^{i\psi},(g',g)),x)=g^{-1}$, where $e^{i\phi}\in U(1)$, $(g',g)\in SU(2)_L\times SU(2)_R$ (see Ref. 3). However, we only need to consider the ρ' case since the ρ case is equivalent to a trivial representation of $\tilde{D}$. For any compact gauge group H, it is sufficient to treat the group $H=SU(2)$ alone since the algebraic constraint (2.6) (the isotropy condition), only involves the isotropy subgroup $SU(2)_D$ and its image under the homomorphism $\lambda:G_0\to H$. This image will either be trivial [i.e., $\lambda(G_0)=e\subset H$] or will determine an SU(2) subgroup of H. In the latter case, we restrict the given representation of H to this SU(2) subgroup and treat each irreducible component separately as for the case $H=SU(2)$. The resulting invariant field obtained from equation (2.3) never mixes these components.

If we choose a left invariant basis on $\bar{M}$ [i.e., the $u(1)\oplus su(2)$ algebra] the Jacobian matrix of the $O(2)\times O(4)$ action is reduced to $SU(2)_R$ action.[3,8] Explicitly, the $O(2)\times O(4)$ transformation: $p=L_\phi L_{(g',g)}p_0$, leads to the Jacobian $J(x)=\Lambda_{(e^{i\phi},(g',g))}(x)=\mathrm{Ad}g\in SO(3)$ and the transformation function $\rho'^{-1}((e^{i\phi},(g',g)),x)=g\in SU(2)$, which, substituted into the global invariance condition (2.3), gives

$$\Psi^{(j,j',i)}(x)=D^{(j,0)}(g)\otimes D^{*(j',0)}(g)\otimes D^{i}(g)\Psi^{(j,j',i)}(x_0), \tag{5.6}$$

where $x=\{x^u(p)\}$ are the local coordinates, D^i is an SU(2) gauge irreducible representation of "isospin i," and $\Psi^{(j,j',i)}$ the corresponding gauged spinor field. In particular, choosing the identity element $p_0=(1,e)$ as origin, we obtain the isotropy condition

$$\Psi^{(j,j',i)}(x_0)=D^{(j,0)}(g)\otimes D^{*(j',0)}(g)\otimes D^{i}(g)\Psi^{(j,j',i)}(x_0) \tag{5.7}$$

for every element of $G_0=SU(2)_D=\{(g,g)\in SU(2)_L\times SU(2)_R\}$. Since $g\in SU(2)$, the unitary transformation $U=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\in SU(2)$ changes $D^{*(j',0)}(g)$ to $D^{(j',0)}(g)$, and (5.7) takes the form

$$\Psi'^{(j,j',i)}(x_0)=D^{(j,0)}(g)\otimes D^{(j',0)}(g)\otimes D^{i}(g)\Psi'^{(j,j',i)}(x_0) \tag{5.8}$$

with the definition

$$\Psi'^{(j,j',i)}(x) \equiv D^{(j',0)}(U)\Psi^{(j,j',i)}(x). \tag{5.9}$$

Next, we introduce a basis change which reduces the tensor product $D^{(j,0)} \otimes D^{(j',0)} \otimes D^{i}$ to SU(2) irreducible parts:

$$\Psi^{J'}_{M'}(x) = \sum_{M,m,m',m''} \langle JiJ'M' | JiMm'' \rangle \times \langle jj'mm' | jj'JM \rangle \Psi'^{(j,j',i)}_{m,m',m''}(x), \tag{5.10}$$

where $\Psi^{J'}_{M'}(x)$ are the components of a mixed spin–isospin field labelled by the projection of its total spin, $-J' \leqslant M' \leqslant J'$, $\langle JiJ'm' | JiMm'' \rangle$ and $\langle jj'mm' | jj'Jm \rangle$ are the Clebsch–Gordan coefficients associated with the transformation and $-j \leqslant m \leqslant j$, $-j' \leqslant m' \leqslant j'$, $-i \leqslant m'' \leqslant i$, $-M \leqslant J \leqslant M$. Then each irreducible part J' satisfies the modified isotropy condition:

$$\Psi^{J'}(x_0) = D^{(J',0)}(g)\Psi^{J'}(x_0). \tag{5.11}$$

Since (5.11) is valid for all $g \in SU(2)$, we have as general solution

$$\Psi^{J'}_{M'}(x_0) = \alpha \delta_{J',0} \delta_{M',0}, \tag{5.12}$$

where $\alpha \in C$ (i.e., a total spin scalar). By means of the inverse basis change, the gauged spinor field $\Psi'^{(j,j',i)}_{m,m',m''}(x_0)$ is recovered:

$$\Psi'^{(j,j',i)}_{m,m',m''}(x_0) = \alpha(-1)^{i+m''} \langle jj'i \; -m'' | jj'mm' \rangle. \tag{5.13}$$

We replace in (5.9) the inverse unitary transformation $U^{\dagger}$ given by

$$D^{(j',0)\dagger}_{m',m'_1}(U) = (-1)^{j'+m'} \delta_{m',-m'_1} \tag{5.14}$$

and derive the invariant fields $\Psi^{(j,j',i)}(x_0)$ in terms of only one Clebsch–Gordan coefficient and one complex constant α:

$$\Psi^{(j,j',i)}_{m,m',m''}(x_0) = \alpha(-1)^{j'+m'+i+m''} \langle jj'i \; -m'' | jj'm \; -m' \rangle. \tag{5.15}$$

Once the (jj')-gauged spinor field is known at the reference point x_0, we may use the invariance condition (5.6) to determine it over $\overline{M}$. But expressed in our left-invariant basis, the $O(2) \times O(4)$ action on the fields corresponds exactly to the $SU(2)_R$ action leaving $\Psi^{(jj',i)}(x_0)$ invariant. Therefore, the most general $O(2) \times O(4)$ invariant gauged spinor field on $\overline{M}$ written in this basis is

$$\Psi^{(jj',i)}_{m,m',m''}(x) = \alpha(-1)^{j'+m'+i+m''} \langle jj'i \; -m'' | jj'm \; -m' \rangle. \tag{5.16}$$

There is only one arbitrary constant for each irreducible SU(2) component. For more general gauge group H, we have as many arbitrary constants as there are irreducible components of the representation obtained by restriction to $\lambda(SU(2)_D)$. We can also determine the O(4)-invariant fields immediately considering that $O(4) \sim SU(2)_L \times SU(2)_R$ has orbits $S^3 \subset \overline{M}$ parametrized by $e^{i\psi} \in U(1) \sim S^1$ on a smooth section. Thus the most general O(4)-invariant (jj') spinor fields with SU(2) gauge freedom are given by (5.16) with α as a complex valued function of the S^1 coordinate ψ.

We now summarize an alternative derivation of (5.16) based upon the infinitesimal invariance condition. In infinitesimal form, the isotropy condition (5.7) becomes:

$$\sum_{m_1} D^{(j,0)}_{mm_1}(\sigma_k)\Psi^{(j,j',i)}_{m_1,m',m''} - \sum_{m'_1} D^{*(j',0)}_{m',m'_1}(\sigma_k)\Psi^{(j,j',i)}_{m,m'_1,m''} + \sum_{m''_1} D^{i}_{m'',m''_1}(\sigma_k)\Psi^{(j,j',i)}_{m,m',m''_1} = 0, \tag{5.17}$$

where the $\{\sigma_k\}$ stand for the Pauli matrices. With σ_3, this relation simplifies to

$$\Psi^{(j,j',i)}_{m,m',m''}(m - m' + m'') = 0; \tag{5.18}$$

this allows us to express the gauged spinor field as

$$\Psi^{(j,j',i)}_{m,m',m''} \equiv \varphi_{m,m'} \delta_{m'-m,m''}. \tag{5.19}$$

Hence, using definition (5.19) in Eq. (5.17), there follow two recurrence relations:

$$a^{+}(j,m-1)\varphi_{m-1,m'} - a^{-}(j',m'+1)\varphi_{m,m'+1} + a^{+}(i,m'-m)\varphi_{m,m'} = 0 \tag{5.20}$$

and

$$a^{-}(j,m+1)\varphi_{m+1,m'} - a^{+}(j',m'-1)\varphi_{m,m'-1} + a^{-}(i,m'-m)\varphi_{m,m'} = 0, \tag{5.21}$$

where

$$a^{\pm}(j,m) \equiv (j(j+1) - m(m \pm 1))^{1/2}, \tag{5.22}$$

which have as solution

$$\varphi_{m,m'} = \alpha(-1)^{j'+2m'+i-m} \langle jj'i(m-m') | jj'm - m' \rangle \tag{5.23}$$

if m and m' satisfy the equation(s), $m' - m = m''$ and $\varphi_{m,m'} = 0$ otherwise. We verify this solution by inserting (5.23) in (5.20) and (5.21) to derive the recurrence relations of the Clebsch–Gordan coefficients.

To show that this is the unique solution, consider a square lattice with perpendicular axis m and m'. We find that all the nonzero $\varphi_{m,m'}$ are located on the lines obeying the equation(s): $m' - m = m''$. But, on the lattice, (5.20) and (5.21) can also be rewritten as horizontal and vertical recurrence relations involving three terms each. Setting $\varphi_{m,m'} = 0$ for $-j > m > j$ and $-j' > m' > j'$, the knowledge of one nonzero $\varphi_{m,m'}$ suffices to determine uniquely all the nonzero coefficients in terms of the known nonzero $\varphi_{m,m'}$ with successive uses of the two recurrence relations. The consistency of the result obtained from various paths to the same point follows from the explicit solution (5.23). Substitution of (5.23) in (5.19) again yields (5.15).

Let us present some specific examples of $\Psi^{(j,j',i)}(x)$, where either j, j' or i equals zero:

$$\text{(a)} \quad j = 0: \Psi^{(0,j',i)}_{0,m',m''} = \alpha \delta_{m'',m'} \delta_{j',i}, \tag{5.24}$$

$$\text{(b)} \quad j' = 0: \Psi^{(j,0,i)}_{m,0,m''} = \alpha(-1)^{i+m''} \delta_{m,-m''} \delta_{i,j}. \tag{5.25}$$

We remark that both cases, (a) $j = \frac{1}{2}$ and (b) $j' = \frac{1}{2}$, agree with invariant spinors found in Ref. 8.

$$\text{(c)} \quad i = 0: \Psi^{(jj',0)}_{m,m',0} = \alpha(-1)^{2j} \delta_{j,j'} \delta_{m,m'}. \tag{5.26}$$

In particular, when $i = 0$, we may compare with results already found in BHPW for vectors and (0,2) tensors. Indeed, since tensors and spinors are related by

$$T_{\mu_1,\ldots,\mu_n}(x) = \sigma^{a_1\dot{a}_1}_{\mu_1} \cdots \sigma^{a_n\dot{a}_n}_{\mu_n} \Psi^{a_1\cdots a_n\dot{a}_1\cdots\dot{a}_n}(x), \tag{5.27}$$

we find by a basis change to irreducible spinors that, in agreement with BHPW,

(1-form) $A = \alpha_{1/2}\omega_L^0$,

(2-form) $F = 0$

(symmetric (0,2) tensor) $G = C\omega_L^0 \otimes \omega_L^0 + \frac{D}{4}\omega_L^i \otimes \omega_L^i$,

where $C = \alpha_0 + 3\alpha_1$ and $D = \alpha_1 - \alpha_0$, $\alpha_0, \alpha_{1/2}, \alpha_1$ denoting, respectively, the constants associated with the nontrivial $O(2)\times O(4)$ invariant spinors $\Psi^{(0,0)}$, $\Psi^{(1/2,1/2)}$, $\Psi^{(1,1)}$, and $\{\omega_L^\mu\}$, the left invariant coframe in $\bar{M}$.

The $O(2)\times O(4)$ invariant spinor fields $\Psi^{(j,j',l)}$ given by (5.17) are particularly simple because they are expressed relative to a left-invariant frame on $\bar{M}$. To rewrite these fields in a Cartesian frame for Minkowski space M, we must make the appropriate change of basis. Let $i:M \hookrightarrow \bar{M}$ denote the conformal immersion of M in $\bar{M}$ and $y = \{y^\mu(i(p))\}$ the corresponding Cartesian coordinates (where defined). Denoting by $\Psi^{(j,j',l)}(y)$ the components of $\Psi^{(j,j',l)}$ pulled back to $\bar{M}$ expressed relative to a Cartesian frame, we have

$$\Psi^{(j,j',l)}_{m,m',m''}(y) = \tau^l \sum_{m_1,m_1'} D^{(j,0)}_{m,m_1}(s) D^{*(j',0)}_{m',m_1'}(s) \Psi^{(j,j',l)}_{m_1,m_1',m''}(x), \tag{5.28}$$

where the transformation matrices define (j,j') spinor representations of the $O(3,1)$ part of the change of spinor basis given by

$$s = \frac{(1+iy^0) + iy^j\sigma_j}{(1+2iy^0 - y^\nu y_\nu)^{1/2}} \in SL(2,C). \tag{5.29}$$

Since i does not preserve the metric g_M, there is also a conformal factor τ^l defined by

$$\tau = [\, y_0^2 + \tfrac{1}{4}(1 - y^\nu y_\nu)^2]^{1/2} \tag{5.30}$$

(see Ref. 8).

VI. SUMMARY

We have formulated two propositions which enable us to deduce rapidly the nonexistence of nonzero invariant (j,j') spinor fields with any compact gauge symmetry if the Jacobian image of the isotropy subgroup contains certain one-parameter subgroups of $O(3,1)\times D$. Applying these two propositions to the eight maximal noncompact subgroups of $C(3,1)$, it follows that only the subgroups $O(3)\times O(2,1)$, $O(2,1)\times O(2,1)$, and $S(U(2,1)\times U(1))$ may lead to nontrivial (j,j') spinor fields which respect invariance up to a compact gauge transformation and that $j + j' \in Z^+$. The three conjugacy classes $O(2)\times O(2,2)$, $O(3,2)$, and $O(4,1)$ require strict invariance (no gauge factor) and $j + j' \in Z$ to permit a nontrivial result.

On the other hand, we have completely determined the $O(4)$ and $O(2)\times O(4)$ invariant (j,j') gauged spinor fields for any compact gauge group. It may be of interest in a further study to explicitly determine the permissible nonzero invariant spinor fields of the maximal noncompact subgroups of $C(3,1)$, or consider lower-dimensional transformation subgroups of $C(3,1)$ with the help of Propositions 1 and 2. A possible application of the $O(4)$ results would be the study of invariant solutions of the coupled system of gauge fields with field representations of the (j,j') type, transforming linearly under some compact gauge groups, thus generalizing the results of Refs. 7 and 8.

ACKNOWLEDGMENTS

The author is sincerely indebted to J. Harnad for guidance and helpful suggestions. Scholarships from the Natural Science and Engineering Research Council of Canada and the Ministère de l'Education du Québec are also gratefully acknowledged.

[1] J. Beckers, J. Harnad, M. Perroud, and P. Winternitz, J. Math. Phys. **19**, 2126 (1978).
[2] J. Beckers, J. Harnad, and P. Jasselette, J. Math. Phys. **21**, 2491 (1980).
[3] J. Harnad, S. Shnider, and L. Vinet, J. Math. Phys. **21**, 2719 (1980).
[4] J. Harnad and L. Vinet, Phys. Lett. B **76**(5), 589 (1978).
[5] J. Harnad, S. Shnider, and L. Vinet, J. Math. Phys. **20**, 931 (1979).
[6] L. Vinet, in *Geometrical and Topological Methods in Gauge Theories, Lecture Notes in Physics*, Vol. 129, edited by J. Harnad and S. Shnider (Springer-Verlag, Berlin, 1980), p. 137.
[7] K. Meetz, Z. Phys. C **6**, 41 (1980).
[8] J. Doneux, Y. Saint-Aubin, and L. Vinet, "The matter-coupled Yang–Mills system in Minkowski space. II. Invariant solutions in the presence of Dirac spinor fields," Phys. Rev. D **25**, 484 (1982).
[9] T. H. Go, Comm. Math. Phys. **41**, 157 (1975).
[10] I. E. Segal, *Mathematical Cosmology and Extragalactic Astronomy* (Academic, New York, 1976).
[11] G. Mack and A. Salam, Ann. Phys. (N.Y.) **53**, 174 (1969).
[12] S. Ferrara, R. Gatto, and A. F. Grillo, *Conformal Algebra in Space-Time, Springer Tracts in Modern Physics*, Vol. 67 (Springer-Verlag, Berlin, 1973).

Self-dual Yang–Mills fields

John R. Porter
Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 9 December 1981; accepted for publication 7 May 1982)

The realization of self-dual Yang–Mills fields as nonlinear superpositions of plane waves is treated in analogy with electromagnetism. Despite the nonlinear nature of the fields, there is a complete correspondence between the two theories. The asymptotic behavior of asymptotically flat fields is examined and the plane wave components are shown to decouple on $\mathscr{I}^-$ and on $\mathscr{I}^+$ for such fields and to exibit the character of linear fields when viewed as propagating from $\mathscr{I}^-$ to $\mathscr{I}^+$.

PACS numbers: 11.15. − q

1. INTRODUCTION

The investigation of self-dual Yang–Mills fields as superpositions of plane waves is carried out in the following sections.

In Sec. 2 the basic ideas and formats used in Minkowski space are introduced with enough detail to allow the reader to proceed. Though a well-developed methodology exists for further exploitation, only the necessary tools are formulated.

In Sec. 3 a brief introduction to plane waves in electromagnetic theory is given, pricipally to illustrate in a known area the techniques to follow and to put these results into the format used in the rest of the paper. Superposition of plane waves is also presented in electromagnetism.

In Sec. 4 the "gauge freedom" that exists in defining plane wave self-dual solutions to the Yang–Mills field equations is carefully discussed as this is crucial to the superposition of such. It turns out that the equivalence class of connections for a single plane wave must be considered in the superposition.

In Sec. 5 the gauge freedom discussed in Sec. 4 is now largely fixed when the superposition of self-dual plane waves is required to again be self-dual. The condition fixing the freedom results in a linear equation giving the representative connections from each equivalence class. The remaining freedom is then discussed.

Having found the connection resulting from the superposition of plane waves, in Sec. 6 the curvature form is calculated and is shown to be explicitly self-dual.

Section 7 concerns asymptotic analysis in electromagnetism and in Sec. 8 a parallel development is given for self-dual Yang–Mills fields. The parallels between the linear and nonlinear cases are remarkable.

In Sec. 9 an example, the field of a single plane wave, is presented in some detail and a brief synopsis is given in the last section, Sec. 10.

A main motivation for the consideration of self-dual Yang–Mills fields is that a critical analysis of the classical solutions of the field equations is a precursor to a quantum description. The self-dual and anti-self-dual fields are eigenstates of helicity and in an asymptotic quantization scheme, it seems reasonable to interpret them as one-particle wave functions in the quantum description.

2. PRELIMINARIES

The analysis takes place on Minkowski space M and its complexification $\mathbb{C}M$ or on an open submanifold of each, though the distinction is not made here. Minkowskian coordinates, $x^a = \{t,x,y,z\}$, are used in which the line element takes the diagonal form, $ds^2 = (dt)^2 - (dx)^2 - (dy^2) - (dz)^2$. On M, the x^a are real while on $\mathbb{C}M$ the coordinates assume complex values. Coordinates $z^a = \{p,q,r,s\}$ are also used for convenience, where the relation between the x^a and z^a is given by

$$p = t + z, \quad q = x + iy, \quad r = t - z, \quad s = -(x - iy).$$

Thus, on real Minkowski space, p and r are real while $\bar{s} = -q$. The null directions, of which there are an S^2's worth, are parametrized by points of the (completed) complex plane $\mathbb{C}$, with coordinates given by Greek letters. A null one-form, $dl(\sigma,\bar{\sigma})$, is given for each point of the sphere, here labeled by σ and $\bar{\sigma}$, which is normalized by demanding that it be future-pointing and have an inner product with dt of 1. Specifically,

$$\begin{aligned} l(\sigma,\bar{\sigma}) &\equiv \frac{1}{1+\sigma\bar{\sigma}}[(1+\sigma\bar{\sigma})t + (\sigma+\bar{\sigma})x + i(\sigma-\bar{\sigma})y \\ &\quad + (1-\sigma\bar{\sigma})z] \\ &= \frac{1}{1+\sigma\bar{\sigma}}[p + \sigma p + \sigma\bar{\sigma}r - \bar{\sigma}s]. \end{aligned}$$

Writing $dl(\sigma,\bar{\sigma}) = l_a(\sigma,\bar{\sigma})dx^a = l_a(\sigma,\bar{\sigma})dz^a$, using $l_a(\sigma,\bar{\sigma})$ for two different sets of objects, the argument $(\sigma,\bar{\sigma})$ will sometimes be given simply as σ with its complex conjugate understood. Thus $l_a(\sigma,\bar{\sigma})$ will be designated by $l_a(\sigma)$ at times. Given $l(\sigma,\bar{\sigma})$, three other functions are defined[1]:

$$\begin{aligned} \eth l(\sigma,\bar{\sigma}) &\equiv (1+\sigma\bar{\sigma})\frac{\partial}{\partial\sigma} l(\sigma,\bar{\sigma}) \\ &= \frac{1}{1+\sigma\bar{\sigma}}\Big(-\bar{\sigma}p + q + \bar{\sigma}r + \bar{\sigma}^2 s\Big), \\ \bar{\eth} l(\sigma,\bar{\sigma}) &= (1+\sigma\bar{\sigma})\frac{\partial}{\partial\bar{\sigma}} l(\sigma,\bar{\sigma}) \\ &= \frac{1}{1+\sigma\bar{\sigma}}\Big(-\sigma p - \sigma^2 q + \sigma r - s\Big), \\ (\eth\bar{\eth}+1) l(\sigma,\bar{\sigma}) &= \left[(1+\sigma\bar{\sigma})^2\frac{\partial^2}{\partial\sigma\partial\bar{\sigma}} + 1\right] l(\sigma,\bar{\sigma}) \\ &= \frac{1}{1+\sigma\bar{\sigma}}(\sigma\bar{\sigma}p - \sigma q + r + \bar{\sigma}s). \end{aligned}$$

Then $dl, d\eth l, d\bar{\eth} l, d(\eth\bar{\eth}+1)l$ form a null tetrad at each point of M with the first and last forms real and the middle two forms complex conjugates of each other. These one-forms provide a convenient basis in which to exibit for each σ a

 0022-2488/83/051224-09$02.50

basis for the self-dual and the anti-self-dual two-forms. Specifically, if the duality operator is denoted by * as usual, for any σ

$$*[dl(\sigma)\wedge d\bar{\eth}l(\sigma)] = i dl(\sigma)\wedge d\bar{\eth}l(\sigma),$$

$$*[d(\eth\bar{\eth}+1)l(\sigma)\wedge d\eth l(\sigma)] = i d(\eth\bar{\eth}+1)l(\sigma)\wedge d\eth l(\sigma),$$

$$*[dl(\sigma)\wedge d(\eth\bar{\eth}+1)l(\sigma) - d\eth l(\sigma)\wedge d\bar{\eth}l(\sigma)]$$
$$= i[dl(\sigma)\wedge d(\eth\bar{\eth}+1)l(\sigma) - d\eth l(\sigma)\wedge d\bar{\eth}l(\sigma)].$$

Thus the three two-forms above are self-dual, while their complex conjugates are anti-self-dual. The conventions used here follow Ref. 2 and thus in the coordinate system x^a the fundamental four-form is $dt\wedge dx\wedge dy\wedge dz$. For complex Minkowski space $\mathbb{C}M$, the coordinates x^a assume complex values and the set of null directions at a point is $S^2\times S^2$ and is realized in the above by complexifying the S^2 of real null directions; that is, by allowing $\bar{\sigma}$ and σ to assume independent complex values. One, in using this representation, must always check that the quantities under consideration are well defined and behaved at $\sigma=\infty$, which will be the case in what follows. Another way of proceeding is to look at the quantities involved under an inversion, $\sigma\to\eta=1/\sigma$, since these two coordinate neighborhoods cover the sphere. There is a well-developed methodology of such and the reader is referred to Refs. 1 and 3 for further details. The point to make here is that the transformation properties will in general be those of the spin-weighted functions which are realized as sections of generically nontrivial bundles over S^2. The operators $\eth$ and $\bar{\eth}$ are maps from sections of a bundle to sections of a different bundle.[4,5]

3. PLANE WAVES IN ELECTROMAGNETISM

The field of a plane wave in electromagnetic theory in Minkowski space propagating in the negative z direction is represented by a two-form (or bivector) as

$$f_1(t+z)d(t-z)\wedge dx + f_2(t+z)d(t-z)\wedge dy,$$

where f_1 and f_2 are arbitrary functions describing the amplitude and phase profile of the plane wave. The plane wave can be decomposed into its self-dual and anti-self-dual parts by writing

$$\tfrac{1}{2}(f_1+if_2)d(t+z)\wedge d(x-iy)$$
$$+\tfrac{1}{2}(f_1-if_2)d(t+z)\wedge d(x+i\,dy)$$

and so the self-dual part is given by

$$F = f(t+z)d(t+z)\wedge d(x-iy)$$
$$\equiv\tfrac{1}{2}(f_1-if_2)d(t+z)\wedge d(x-iy), \qquad (3.1)$$

where f is a complex-valued function of its argument. Note that $l(0,u)=t+z, \bar{\eth}l(0,0)=x-iy$ and so $d(t+z)\wedge d(x-iy) = dl(0,0)\wedge d\bar{\eth}l(0,0)$. If desired, the plane wave given by (3.1) can be given in terms of a superposition of single frequency plane waves by Fourier analyzing the function $f(t+z)$ but this is not done here for the moment.

In what follows only self-dual fields will be considered. To give the field of an arbitrary plane wave with propagation direction determined by $(\sigma,\bar{\sigma})$, consider

$$F^1_\sigma = \ddot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})dl(\sigma,\bar{\sigma})\wedge d\bar{\eth}l(\sigma,\bar{\sigma}), \qquad (3.2)$$

where the "·" refers to the derivative of a with respect to its first argument, which is $l(\sigma,\bar{\sigma})$. Thus, for fixed $(\sigma,\bar{\sigma})$ Eq. (3.2) describes an arbitrary self-dual plane wave whose propagation direction is given by $l(\sigma,\bar{\sigma})$ and whose amplitude and phase profile is given by $\ddot{a}$. A potential for F'_σ is provided by

$$\gamma'_\sigma = \dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})d\bar{\eth}l(\sigma,\bar{\sigma}), \qquad (3.3)$$

which is defined only up to the addition of a gradient. The relation between the connection (potential) and the curvature tensor (field) is provided by

$$F'_\sigma = d\gamma'_\sigma - \gamma'_\sigma\wedge\gamma'_\sigma \qquad (3.4)$$

and the free field equations are satisfied,

$$dF'_\sigma + [F'_\sigma,\gamma'_\sigma] = 0, \qquad (3.5)$$

where

$$[F'_\sigma,\gamma'_\sigma] = F'_\sigma\wedge\gamma'_\sigma + \gamma'_\sigma\wedge F'_\sigma. \qquad (3.6)$$

The quadratic terms in (3.4) and (3.5) vanish identically for electromagnetism. Since the curvature F'_σ is self-dual and satisfies $*F'_\sigma = iF'_\sigma$, Eq. (3.5) may be written as

$$d*F'_\sigma + [*F'_\sigma,\gamma'_\sigma] = 0, \qquad (3.7)$$

which are the Yang–Mills free field equations.

The potential γ'_σ in (3.3) is form invariant under a gauge transformation that takes $\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma}) \to \dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma}) + c(\sigma,\bar{\sigma})$ and use of this condition will be helpful later.

For electromagnetism a general self-dual connection and resulting field can be obtained by linear superposition of plane waves, one for each propagation direction. The connection results from γ'_σ for each $\sigma\in\mathbb{C}$ or alternately $\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})$ on $M\times S^2$ and integrating over the set of propagation directions with respect to the correct measure on S^2, which is induced by the Minkowski space measure pulled back to a section of the null cone at a point consisting of all future null displacements normalized with dt to be 1. The measure is that of a homogeneous unit two-sphere and is

$$d\mu_\sigma = \frac{2}{i}\frac{d\sigma\wedge d\bar{\sigma}}{(1+\sigma\bar{\sigma})^2}. \qquad (3.8)$$

Thus

$$\int_{S^2} d\mu_\sigma = 4\pi,$$

the surface area of a unit sphere. In usual coordinates on S^2 $d\mu_\sigma = \sin^2\theta\, d\theta\, d\phi$. Thus the superposition of the plane waves given by $\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})$ gives rise to a connection

$$\gamma = \int_{S^2} \dot{a}(l(\sigma,\bar{\sigma})\sigma,\bar{\sigma})d\bar{\eth}l(\sigma,\bar{\sigma})d\mu_\sigma$$

and a curvature form

$$F = \int_{S^2} \ddot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})dl(\sigma,\bar{\sigma})\wedge d\bar{\eth}l(\sigma,\bar{\sigma})d\mu_\sigma.$$

Again F is uniquely specified and γ is defined modulo the addition of a gradient. Since each plane wave component is self-dual and satisfies the free field equation and the superposition is linear the resulting field is also self-dual and satisfies the free field equations. If each plane wave is futher decomposed into frequency eigenstates, the integral over S^2 would be replaced with an integral over $S^2\times R$.

4. SELF-DUAL YANG–MILLS PLANE WAVES

The basic setting is a principal fiber bundle over M with structure group $\mathscr{G}$ or an associated vector bundle; the assumption will be made that the analysis is done in a principal fiber bundle since it can then be carried over to any associated vector bundle. With respect to some section of the bundle, (3.2) and (3.3) represent the curvature form and the connection form of a self-dual plane wave where now these forms have values in the Lie algebra $\mathscr{g}$ of $\mathscr{G}$. Equation (3.4), giving the relation between γ'_σ and F'_σ as well as the self-dual field equations (3.6), remains valid since again the quadratic terms are identically zero. It is easy to convince oneself that the same superposition as for electromagnetism does not yield a self-dual curvature form. The reason for this is due to the nonlinearities in the generic Yang–Mills field. It is necessary to be more careful in the superposition in order to obtain a self-dual curvature form.

Given $\gamma'_\sigma, F'_\sigma$ with respect to a section of the bundle and G representing a change of section of the bundle, the transformed connection $\hat{\gamma}_\sigma$ and curvature $\hat{F}_\sigma$ are given by

$$\hat{\gamma}_\sigma = dG\cdot G^{-1} + G\gamma'_\sigma G^{-1}$$

and

$$\hat{F}_\sigma = GF'_\sigma G^{-1}.$$

Instead of using the equivalence classes defined by the above relations for $\hat{\gamma}_\sigma$, $[\hat{\gamma}_\sigma]$ and for $\hat{F}_\sigma$, $[\hat{F}_\sigma]$, here the classes will be $\{\gamma_\sigma\} = (1/4\pi)\,[\hat{\gamma}_\sigma]$ and $\{F_\sigma\} = (1/4\pi)\,[\hat{F}_\sigma]$, where

$$\gamma_\sigma = \frac{1}{4\pi}(dG\cdot G^{-1} + G\gamma'_\sigma G^{-1}) = \frac{1}{4\pi}\hat{\gamma}_\sigma \tag{4.1}$$

and

$$F_\sigma = \frac{1}{4\pi}(GF'_\sigma G^{-1}) = \frac{1}{4\pi}\hat{F}_\sigma. \tag{4.2}$$

This is done because the connections will be integrated on a two-sphere with measure given by (3.8) and with total area 4π (i.e., not 1) and to agree with previous work.[1] Alternately, the surface area can be normalized to one or the defining relations for the field may be taken to be

$$F = d\gamma - \frac{1}{4\pi}\gamma\wedge\gamma$$

and then the Bianchi identities are

$$dF + \frac{1}{4\pi}[F,\gamma] = 0.$$

The function G is a function of $\{x^a,\sigma,\bar{\sigma}\}$ and has values in the Lie group $\mathscr{G}$. It is necessary to consider, instead of $(1/4\pi)\gamma'_\sigma$ and $(1/4\pi)F'_\sigma$, the entire equivalence class of connection forms and associated curvature forms, $\{\gamma_\sigma\}$ and $\{F_\sigma\}$, where the equivalence relation is defined by (4.1) and (4.2). Next consider the question: Given $a(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})$, is it possible to pick from each equivalence class of connections, $\{\gamma_\sigma\}$, a representative so that the superposition of these representatives defines the connection of a self-dual curvature form? In the abelian case such care is irrelevent and the answer is always yes. Of course, if possible, the resulting connection and curvature, γ and F, will not be unique but only up to the equivalence.

$$\gamma\rightarrow dA\cdot A^{-1} + A\gamma A^{-1} \tag{4.3}$$

and

$$F\rightarrow AFA^{-1} \tag{4.4}$$

where A is a $\mathscr{G}$-valued function on M.

To digress and generalize, consider equivalence classes of plane wave connections and curvature forms, $\{\gamma_\sigma\}$ and $\{F_\sigma\}$, and suppose for each σ these satisfy some set of equations invariant under the equivalence relation (e.g., self-dual Yang–Mills', anti-self-dual Yang–Mills', Yang–Mills', Ricci tensor zero if the bundle is the bundle of frames, etc.). Then ask for a representative from each equivalence class, γ_σ and F_σ, so that the superposition of these representatives again satisfies the same set of equations. The resulting connection and curvature forms are

$$\gamma = \int \gamma_\sigma\, d\mu_\sigma,$$

$$\begin{aligned} F = d\gamma - \gamma\wedge\gamma &= \int (d\gamma_\sigma - \gamma_\sigma\wedge\gamma_\sigma)d\mu_\sigma \\ &\quad - \int\!\!\int \gamma_\lambda\wedge\gamma_\sigma(\delta(\lambda,\sigma) - 1)d\mu_\lambda\, d\mu_\sigma \\ &= \int F_\sigma d\mu_\sigma - \int\!\!\int \gamma_\lambda\wedge\gamma_\sigma(\delta(\lambda,\sigma) - 1)d\mu_\lambda\, d\mu_\sigma. \end{aligned}$$

In the case at hand, γ and F should satisfy the self-dual Yang–Mills field equations.

5. SUPERPOSITION OF SELF-DUAL PLANE WAVES

Given a connection γ and the corresponding curvature F, the pair satisfies the free field equation (3.7) automatically if F is self-dual since then the field equations are simply the Bianchi identities (3.6). So one only here requires that F be self-dual. The condition that F is self-dual is equivalent to the statement that the pullback of F to any tangential anti-self-dual two surface of $M\rightarrow\mathbb{C}M$ is zero or that the connection γ pulled back to any such surface is trivial (a connection with zero curvature). This last condition is the most useful one in the present analysis.

In $\mathbb{C}M$ the anti-self-dual two-planes (the points of projective twistor space) are given by the integral submanifolds of the distributions $dl(\lambda,\bar{\lambda})\wedge d\eth l(\lambda,\bar{\lambda})$, where $\lambda\in\mathbb{C}$ is a constant. The submanifolds are given by $l(\lambda,\bar{\lambda}) = c_1$ and $\eth l(\lambda,\bar{\lambda}) = c_2$, where c_1 and c_2 are two (complex) constants. Those with c_1 real intersect M in a null line. Here consider the tangent space to such a two-plane in $\mathbb{C}M$ at a point of M. Let any of the two-planes labelled by λ be given by the injection map

$$i_\lambda : \mathbb{C}^2\rightarrow\mathbb{C}M.$$

Denote the induced map on forms from $\mathbb{C}M$ to $\mathbb{C}^2$ (pullback of forms on $\mathbb{C}M$ to $\mathbb{C}^2$) by i^*_λ as usual. Since $dl(\sigma,\bar{\sigma})\wedge d\bar{\eth} l(\sigma,\bar{\sigma})$ is a self-dual two-form on $\mathbb{C}M$, it pulls-back via i^*_λ to zero.

$$i^*_\lambda(dl(\sigma,\bar{\sigma})\wedge d\bar{\eth} l(\sigma,\bar{\sigma}))\equiv i^*_\lambda dl(\sigma,\bar{\sigma})\wedge i^*_\lambda d\bar{\eth} l(\sigma,\bar{\sigma}) = 0.$$

Thus the two one-forms on $\mathbb{C}^2$, $i^*_\lambda dl(\sigma,\bar{\sigma})$ and $i^*_\lambda d\bar{\eth} l(\sigma,\bar{\sigma})$, are proportional. The factor of proportionality is required and will be a function of $(\lambda,\bar{\lambda})$ and $(\sigma,\bar{\sigma})$. Since the two-plane as imbedded in $\mathbb{C}M$ is given by $l(\lambda,\bar{\lambda}) = C_1$, $l(\lambda,\bar{\lambda}) = C_2$, the equations

$$p - \bar{\lambda} s = C_3,$$
$$q + \bar{\lambda} r = C_4$$

are easily obtained and thus one may take r and s as coordinates on the imbedded submanifold (unless $\bar{\lambda} = 0$). Then calculating $dl(\sigma,\bar{\sigma})$ and $d\bar{\eth} l(\sigma,\bar{\sigma})$ and looking at i^*_λ of these results in

$$i^*_\lambda dl(\sigma,\bar{\sigma}) = (\bar{\lambda} - \bar{\sigma})\left(\frac{ds - \sigma dr}{1 + \sigma\bar{\sigma}}\right),$$
$$i^*_\lambda d\bar{\eth} l(\sigma,\bar{\sigma}) = -(1 + \sigma\bar{\lambda})\left(\frac{ds - \sigma dr}{1 + \sigma\bar{\sigma}}\right).$$

Thus

$$i^*_\lambda d\bar{\eth} l(\sigma,\bar{\sigma}) = -4\pi L(\lambda,\sigma) i^*_\lambda dl(\sigma,\bar{\sigma}), \tag{5.1}$$

where

$$L(\lambda,\sigma) \equiv L(\lambda,\bar{\lambda};\sigma,\bar{\sigma}) \equiv \frac{1}{4\pi}\frac{1 + \sigma\bar{\lambda}}{\bar{\lambda} - \bar{\sigma}}. \tag{5.2}$$

The two point function on the sphere is a singular operator[6] in the sense that

$$\eth_\lambda L(\lambda,\sigma) \equiv (1 + \lambda\bar{\lambda})\frac{\partial}{\partial\lambda} L(\lambda,\sigma) = \delta(\lambda,\sigma), \tag{5.3}$$

that is,

$$\int_{S^2} L(\lambda,\sigma) f(\sigma) d\mu_\sigma = g(\lambda),$$

where

$$\eth_\lambda g(\lambda) = f(\lambda).$$

Also $g(\lambda)$ is orthogonal to the kernel of $\eth_\lambda$, or

$$\int_{S^2} g(\lambda) d\mu_\lambda = \int_{S^2} L(\lambda,\sigma) d\mu_\lambda = 0. \tag{5.4}$$

Now consider $i^*_\lambda \gamma$, where now the coordinates x^a are taken to be real so that the $\mathbb{C}^2$ is examined tangentially to $M \to \mathbb{C}M$.

Then

$$\begin{aligned} i^*_\lambda \gamma &= i^*_\lambda \int \gamma_\sigma d\mu_\sigma \\ &= i^*_\lambda \frac{1}{4\pi}\int dG\cdot G^{-1}(\sigma,\bar{\sigma}) + G\dot{a} d\bar{\eth} l G^{-1}(\sigma,\bar{\sigma}) d\mu_\sigma \\ &= i^*_\lambda \int \frac{1}{4\pi} dG\cdot G^{-1}(\sigma,\bar{\sigma}) - G\, da\, G^{-1}(\sigma,\bar{\sigma}) L(\lambda,\sigma) d\mu_\sigma \end{aligned} \tag{5.5}$$

using (5.1) and the fact that $\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma}) dl(\sigma,\bar{\sigma}) = da(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})$. Triviality of the connection on tangential anti-self-dual two-planes thus demands that $i^*_\lambda \gamma$ be equivalent to a zero connection or that there exist for each λ a section of the bundle, $H(x^a,\lambda)$, such that

$$i^*_\lambda dH\cdot H^{-1}(x^a,\lambda) = i^*_\lambda \gamma. \tag{5.6}$$

Thus $dH\cdot H^{-1}(x_a,\lambda)$ represents a trivial connection for each λ which pulls back via i^*_λ to $i^*_\lambda\gamma$, which is just (5.5). Now using (5.5),

$$\begin{aligned} i^*_\lambda dH\cdot H^{-1} = i^*_\lambda \int \frac{1}{4\pi} & dG\cdot G^{-1}(\sigma,\bar{\sigma}) \\ & - G\, da\, G^{-1}(\sigma,\bar{\sigma}) L(\lambda,\sigma)\, d\mu_\sigma \end{aligned} \tag{5.7}$$

results and since this holds for all x^a and λ,

$$\begin{aligned} dH\cdot H^{-1}(x,\lambda) = \int \frac{1}{4\pi} & dG\cdot G^{-1}(\sigma,\bar{\sigma}) \\ & - G\, da\, G^{-1}(\sigma,\bar{\sigma}) L(\lambda,\sigma)\, d\mu_\sigma. \end{aligned} \tag{5.8}$$

Because of (5.4) and (5.3) H satisfies

$$\int dH\cdot H^{-1}(\lambda,\bar{\lambda}) d\mu_\lambda = \int dG\cdot G^{-1}(\sigma,\bar{\sigma}) d\mu_\sigma \tag{5.9}$$

and

$$\begin{aligned} \eth_\lambda dH\cdot H^{-1}(\lambda,\bar{\lambda}) &\equiv (1 + \lambda\bar{\lambda})\frac{\partial}{\partial\lambda} dH\cdot H^{-1}(\lambda,\bar{\lambda}) \\ &= -G\, da\, G^{-1}(\lambda,\bar{\lambda}), \end{aligned} \tag{5.10}$$

suppressing the coordinate dependence in M via x^a. For (5.9) integrate (5.8) with respect to $d\mu_\lambda$ and remember that $\int d\mu_\lambda = 4\pi$. Now multiplying (5.10) on the left by H^{-1} and on the right H,

$$d(H^{-1}\eth H) = -HG^{-1}\, da\, G^{-1}H = -K\, da\, K^{-1} \tag{5.11}$$

results where $K \equiv H^{-1}G$. The quantity $K\, da\, K^{-1}$ is thus exact, since the left-hand side is exact, and proportional to $dl(\lambda,\bar{\lambda})$ so there exists a g-valued function $b(l(\lambda,\bar{\lambda}),\lambda,\bar{\lambda})$ such that

$$db = K\, da\, K^{-1}.$$

Thus

$$d(H^{-1}\eth H) = -db$$

or

$$H^{-1}\eth H = -b + c(\lambda,\bar{\lambda}),$$

where $dc = 0$. Since c is irrelevent to γ or F, set it equal to zero and

$$H^{-1}\eth H + b = 0 \tag{5.12}$$

Now the connection can be written as

$$\begin{aligned} \gamma &= \frac{1}{4\pi}\int dG\cdot G^{-1} + G\dot{a} G^{-1} d\bar{\eth} l(\sigma,\bar{\sigma}) d\mu_\sigma \\ &= \frac{1}{4\pi}\int dH\cdot H^{-1} + H(dK\cdot K^{-1} \\ &\qquad + K\dot{a} K^{-1} d\bar{\eth} l) H^{-1}(\sigma,\bar{\sigma}) d\mu_\sigma \end{aligned}$$

since

$$\begin{aligned} dG\cdot G^{-1} &= d(HK)\cdot(HK)^{-1} \\ &= (dH\cdot K + H\, dK)(K^{-1}H^{-1}) \\ &= dH\cdot H^{-1} + H dK\cdot K^{-1}H^{-1} \end{aligned}$$

and using (5.9),

$$\int H\, dK\cdot K^{-1}H^{-1}(\sigma,\bar{\sigma}) d\mu_\sigma = 0.$$

Thus $K(x^a,\sigma,\bar{\sigma})$ represents a change of sections for the plane wave components and the plane wave connections change from

$$\dot{a} d\bar{\eth} l \to dK\cdot K^{-1} + K\dot{a}\bar{\eth} l K^{-1}.$$

Given the particular gauge conditions that result in $\dot{a} d\bar{\eth} l$, this freedom can be largely eliminated (remembering $K\, da\, K^{-1}$ is exact) and one can take $K = I$, the identity element of $\mathcal{G}$. Doing that,

$$G^{-1}\eth G + a = 0 \tag{5.13}$$

results. A solution to (5.13) represents a choice of representatives from each equivalence class $\{\gamma_\sigma\}$ for given plane wave data, $a(l(\sigma,\bar\sigma),\sigma,\bar\sigma)$, so that the (nonlinear) superposition of these representatives is again the connection of a self-dual curvature form. Note that the solutions of (5.13) are not unique but defined only up to

$$G \to AG,$$

where A is a section of the bundle over M, and that this is precisely the equivalence class containing γ and F by (4.2) and (4.4).

Assuming that one has solved the linear equation (5.13) then

$$\gamma = \frac{1}{4\pi}\int dG\cdot G^{-1} + G\dot a\, d\bar\eth\, lG^{-1}(\sigma,\bar\sigma)d\mu_\sigma \tag{5.14}$$

and the only part of the integrand that contributes to the integral is the ${}_0Y_{00}(\sigma,\bar\sigma)$ part in a spherical harmonic decomposition. It is possible to isolate this part of the integrand and, in fact,

$$\gamma = dG\cdot G^{-1} + \eth_\sigma h\, dl(\sigma,\bar\sigma) - h\, d\eth_\sigma l(\sigma,\bar\sigma),$$

where $h \equiv \bar\eth(dG\cdot G^{-1})\cdot dl(\sigma,\bar\sigma)$, and this expression for γ is independent of the $(\sigma,\bar\sigma)$. This is the expression given by Newman[1,7] and ties the present work to his previous work. The point here to be made is that the elements, a and G, entering into the field can be identified with plane wave decomposition of the field and a choice of representative for each plane wave component. All the analysis can be done locally and will necessarily be done in this manner if the bundle is not trivial on all of M.

6. THE CURVATURE FORM

It is possible to determine the curvature form and explicitly exhibit its self-dual character. Given the connection γ from (5.14),

$$\begin{aligned} d\gamma &= \frac{1}{4\pi}\int -dG\wedge dG^{-1} + (dG\dot aG^{-1} \\ &\quad + G\ddot aG^{-1}dl + G\dot a dG^{-1})\wedge d\bar\eth l\, d\mu \\ &= \frac{1}{4\pi}\int dG\cdot G^{-1}\wedge dG\cdot G^{-1} + (dG\cdot G^{-1}G\dot aG^{-1} \\ &\quad + G\ddot aG^{-1}dl - G\dot aG^{-1}dG\cdot G^{-1})\wedge d\bar\eth l\, d\mu, \end{aligned} \tag{6.1}$$

since $dG\cdot G^{-1} + GdG^{-1} = 0$ and the variables have been suppressed. The other term in F is $\gamma\wedge\gamma$, which is given by

$$\begin{aligned} \gamma\wedge\gamma = \Big(\frac{1}{4\pi}\Big)^2 \int\int & dG\cdot G^{-1}(\eta)\wedge dG\cdot G^{-1}(\rho) \\ & + dG\cdot G^{-1}(\eta)\wedge G\dot aG^{-1}(\rho)d\bar\eth l(\rho) \\ & + G\dot aG^{-1}(\rho)d\bar\eth l(\rho)\wedge dG\cdot G^{-1}(\eta) \\ & + G\dot aG^{-1}d\bar\eth l(\eta)\wedge G\dot aG^{-1}d\bar\eth l(\rho)\, d\mu_\eta\, d\mu_\rho. \end{aligned} \tag{6.2}$$

The first term in $d\gamma$ and the first term in $\gamma\wedge\gamma$ combine to yield

$$\frac{1}{4\pi}\int\int dG\cdot G^{-1}(\eta)\wedge dG\cdot G^{-1}(\rho)\Big(\delta(\eta,\rho) - \frac{1}{4\pi}\Big)d\mu_\eta\, d\mu_\rho. \tag{6.3}$$

But in a spherical harmonic decomposition $\delta(\eta,\rho) - (1/4\pi) = \Sigma_{l>0,m}\, {}_0Y_{lm(\eta)}\, {}_0\bar Y_{lm(\eta)}$ and since

$$dG\cdot G^{-1}(\eta) = \int \frac{1}{4\pi} dG\cdot G^{-1}(\sigma) - G\, da\, G^{-1}(\sigma)L(\eta,\sigma)d\mu_\sigma, \tag{6.4}$$

where the first term on the right-hand side is spanned by ${}_0Y_{00}(\eta)$ and the second term by ${}_0Y_{lm}(\eta)$ for $l>0$,

$$\begin{aligned} &\int dG\cdot G^{-1}(\eta)\Big(\delta(\eta,\rho) - \frac{1}{4\pi}\Big)d\mu_\eta \\ &\qquad = -\int G\, da\, G^{-1}(\sigma)L(\rho,\sigma)d\mu_\sigma, \end{aligned}$$

and thus (6.3) can be written as

$$-\frac{1}{4\pi}\int\int G\, da\, G^{-1}(\eta)\wedge dG\cdot G^{-1}(\rho)L(\rho,\eta)d\mu_\eta\, d\mu_\rho,$$

but $L(\rho,\eta)$ as a function of ρ is orthogonal to ${}_0Y_{00}(\rho)$ and so the same procedure on the term $dG\cdot G^{-1}(\rho)$ yields for (6.3)

$$\begin{aligned} &\frac{1}{4\pi}\int\int\int G\, da\, G^{-1}(\eta)\wedge G\, da\, G^{-1}(\rho) \\ &\qquad \times L(\sigma,\eta)L(\sigma,\rho)d\mu_\eta\, d\mu_\rho\, d\mu_\sigma. \end{aligned}$$

Integrating this expression with respect to σ first,

$$\int L(\sigma,\eta)L(\sigma,\rho)d\mu_\sigma$$

must be evaluated.

$$\begin{aligned} &\int L(\sigma,\eta)L(\sigma,\rho)d\mu_\sigma \\ &\qquad = \frac{2}{i}\Big(\frac{1}{4\pi}\Big)^2\int \frac{1+\eta\bar\sigma}{\bar\sigma-\bar\eta}\,\frac{1+\rho\bar\sigma}{\bar\sigma-\bar\rho}\,\frac{d\sigma d\bar\sigma}{(1+\sigma\bar\sigma)^2} \\ &\qquad = -\frac{2}{i}\Big(\frac{1}{4\pi}\Big)^2\oint \frac{(1+\eta\bar\sigma)(1+\rho\bar\sigma)}{(\bar\sigma-\bar\eta)(\bar\sigma-\bar\rho)\bar\sigma(1+\sigma\bar\sigma)}\,d\bar\sigma, \end{aligned}$$

where the integration $\oint$ is on contours around each pole of the integrand (by Stoke's theorem). Evaluating the residues yields

$$-\frac{1}{4\pi}\,\frac{\eta-\rho}{\bar\eta-\bar\rho},$$

and thus (6.3) becomes

$$-\Big(\frac{1}{4\pi}\Big)^2\int\int G\, da\, G^{-1}(\eta)\wedge G\, da\, G^{-1}(\rho)\frac{\eta-\rho}{\bar\eta-\bar\rho}d\mu_\eta\, d\mu_\rho.$$

For the second and fourth terms in $d\gamma$ in (6.1) substitute for $dG\cdot G^{-1}$ the right-hand side of (6.4). Two terms of the form

$$\Big(\frac{1}{4\pi}\Big)^2\int\int dG\cdot G^{-1}(\sigma)\wedge G\dot aG^{-1}d\bar\eth l(\eta) - G\dot aG^{-1}(\eta)dG\cdot G^{-1}(\sigma)\wedge d\bar\eth l(\eta)d\mu_\eta\, d\mu_\sigma$$

cancel like terms in $\gamma\wedge\gamma$ in Eq. (6.2). Thus an expression for $F = d\gamma - \gamma\wedge\gamma$ is obtained which is

$$F=\frac{1}{4\pi}\int G\ddot{a}G^{-1}(\eta)dl\wedge d\bar{\delta}l(\eta)d\mu_\eta+\frac{1}{4\pi}\int\int G\dot{a}G^{-1}(\eta)G\dot{a}G^{-1}(\sigma)\Big[-L(\sigma,\eta)dl(\eta)\wedge d\bar{\delta}l(\sigma)$$
$$+L(\eta,\sigma)dl(\sigma)\wedge d\bar{\delta}l(\eta)-\frac{1}{4\pi}d\bar{\delta}l(\eta)\wedge d\bar{\delta}l(\sigma)-\frac{1}{4\pi}\frac{\eta-\sigma}{\bar{\eta}-\bar{\sigma}}dl(\eta)\wedge dl(\sigma)\Big]d\mu_\eta d\mu_\sigma.$$

The term in brackets can be evaluated without difficulty as

$$\frac{1}{4\pi}\left[\frac{-2dp\wedge ds+2(\eta+\sigma)(dp\wedge dr-dq\wedge ds)+2\eta\sigma dq\wedge dr}{\bar{\eta}-\bar{\sigma}}\right].$$

Each term in F is explicitly self-dual and thus F is explicitly self-dual. The integrand in the double integral is skew-symmetric in η and σ so $G\dot{a}G^{-1}(\eta)G\dot{a}G^{-1}(\sigma)$ can be replaced by the commutator of the two terms and

$$F=\frac{1}{4\pi}\int G\ddot{a}G^{-1}dl\wedge d\bar{\delta}l(\sigma)d\mu_\sigma+\left(\frac{1}{4\pi}\right)^2\int\int[G\dot{a}G^{-1}(\eta),G\dot{a}G^{-1}(\sigma)]$$
$$\times\left\{\frac{-dp\wedge ds+(\eta+\sigma)(dp\wedge dr-dq\wedge ds)+\eta\sigma dq\wedge dr}{\bar{\eta}-\bar{\sigma}}\right\}d\mu_\eta d\mu_\sigma. \qquad (6.5)$$

The double integral vanishes identically for abelian groups and is nonsingular in general despite the appearance of the term $\bar{\eta}-\bar{\sigma}$ in the denominator. The commutator is skew-symmetric in η and σ (and thus zero when $\eta=\sigma$), has spin-weight 1 in each variable, and thus must have a factor of $\bar{\eta}-\bar{\sigma}$ cancelling the like term in the denominator and yielding a nonsingular integrand.

7. ASYMPTOTIC ANALYSIS IN ELECTROMAGNETIC THEORY

Given connection and curvature forms in electromagnetic theory, in order to look at the asymptotic behavior, it suffices to examine a single plane wave and then the general follows by linear superposition. The asymptotic analysis referred to is with respect to future (past) null infinity, $\mathscr{I}^+(\mathscr{I}^-)$. A coordinate system for M is used which is adapted to such analysis. The coordinates used are designated by $(u,r,\eta,\bar{\eta})$ and are related to $\{x^a\}$ by the equation

$$l(\zeta,\bar{\zeta})=u+r\frac{(\zeta-\eta)(\bar{\zeta}-\bar{\eta})}{(1+\zeta\bar{\zeta})(1+\eta\bar{\eta})}.$$

The coordinates $(u,r,\eta,\bar{\eta})$ are advanced null polar coordinates and $\mathscr{I}^+$ is given by $r=\infty$, which then has coordinates $(u,\eta,\bar{\eta})$. A single plane wave given by (3.3),

$$\gamma'_\sigma=\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})d\bar{\delta}l(\sigma,\bar{\sigma}),$$

is given in null polar coordinates by

$$\gamma'_\sigma=\dot{a}(l(\sigma),\sigma)\left\{dr\frac{(1+\bar{\eta}\sigma)(\sigma-\eta)}{(1+\eta\bar{\eta})(1+\sigma\bar{\sigma})}+r\left[-\frac{(1+\bar{\eta}\sigma)^2d\eta}{(1+\sigma\bar{\sigma})(1+\eta\bar{\eta})^2}+\frac{(\sigma-\eta)^2d\bar{\eta}}{(1+\sigma\bar{\sigma})(1+\eta\bar{\eta})^2}\right]\right\}. \qquad (7.1)$$

In particular for $\eta=\sigma$

$$\gamma'_\sigma=-\dot{a}(u)\frac{r\,d\sigma}{1+\sigma\bar{\sigma}}, \qquad (7.2)$$

$$F'_\sigma=-\ddot{a}(u)\frac{du\wedge r\,d\sigma}{1+\sigma\bar{\sigma}}. \qquad (7.3)$$

If $\dot{a}\to 0$ as its argument $\to+\infty$ in any manner, then $\gamma'_\sigma\to 0$ as $r\to\infty$ if $\eta\neq\sigma$. To investigate the asymptotic behavior more closely, consider the Minkowskian product of γ'_σ and its complex conjugate, denoted by $\langle\gamma'_\sigma,\bar{\gamma}'_\sigma\rangle$. Remembering that $r\,d\eta/1+\eta\bar{\eta}$ is a complex unit null form, that is,

$$\left\langle\frac{r\,d\eta}{1+\eta\bar{\eta}},\frac{r\,d\bar{\eta}}{1+\eta\bar{\eta}}\right\rangle=1,\quad \langle\gamma'_\sigma,\bar{\gamma}'_\sigma\rangle\to\begin{cases}0 & \eta\neq\sigma\\ |\dot{a}(\sigma)|^2 & \eta=\sigma\end{cases}$$

as $r\to\infty$ with the above growth condition. Provided that a stronger condition is imposed,

$$\dot{a}(z)\sim z^{-1-\epsilon},\quad \epsilon>0 \text{ as } z\to+\infty,$$

then

$$\langle r\gamma'_\sigma,r\bar{\gamma}'_\sigma\rangle\to\begin{cases}0 & \eta\neq\sigma\\ \infty & \eta=\sigma\end{cases}\text{ as } r\to+\infty. \qquad (7.4)$$

The asymptotic limit of $r\gamma'_\sigma$ defines the connection on $\mathscr{I}^+$ in the usual formulation of asymptotic limits. From (7.4) it is clear that $r\gamma'_\sigma$ has an asymptotic limit that has singular support on a generator of $\mathscr{I}^+$ labelled by $\eta=\sigma$. Thus the asymptotic limit of the connection gives rise to a distributional connection on $\mathscr{I}^+$. Thus the asymptotic limit must be interpreted in terms of integrals over the generators of $\mathscr{I}^+$. In (7.1) the coefficient of

$$dr \text{ is } -a,_{\bar{\eta}}\frac{(1+\bar{\eta}\sigma)(1+\eta\bar{\eta})}{r(1+\eta\bar{\sigma})},$$

of

$$d\eta \text{ is } a,_\eta\frac{1+\bar{\eta}\sigma}{\bar{\sigma}-\bar{\eta}},$$

and of

$$d\bar{\eta} \text{ is } -a,_{\bar{\eta}}\frac{\sigma-\eta}{1+\eta\bar{\sigma}}.$$

Thus

$$\gamma'_\sigma=-a,_{\bar{\eta}}\frac{(1+\bar{\eta}\sigma)(1+\eta\bar{\eta})}{r(1+\eta\bar{\sigma})}dr+a,_\eta\frac{1+\bar{\eta}\sigma}{\bar{\sigma}-\bar{\eta}}d\eta-a,_{\bar{\eta}}\frac{\sigma-\eta}{1+\eta\bar{\sigma}}d\bar{\eta}. \qquad (7.5)$$

Given the limiting connection from (7.2), the idea is to integrate the one-form γ'_σ over a two-sphere given by r,u constant and so consider

$$\int_{u,r\ \mathrm{const}} \gamma'_\sigma \wedge d\bar\eta = \int a,\eta \frac{1+\bar\eta\sigma}{\bar\sigma-\bar\eta} d\eta\wedge d\bar\eta = \int d\left(a\frac{1+\bar\eta\sigma}{\bar\sigma-\bar\eta}d\bar\eta\right) = +2\pi i a(u)(1+\sigma\bar\sigma). \tag{7.6}$$

There is another pole at $\eta=\infty$ but with the growth conditions on a this contribution has a zero limit as $r\to\infty$. Note, however, that the expression in (7.6) is independent of the value of r. Also

$$\int_{r,u\ \mathrm{const}} \gamma'_\sigma d\eta = -2\pi i a(u+r)\frac{1+\sigma\bar\sigma}{\bar\sigma^2} \tag{7.7}$$

and this has a zero limit as $r\to\infty$. Thus only (7.6) has a nonzero asymptotic limit which has support on one generator of $\mathscr{I}^+$ and putting the correct measure on the two-sphere

$$\int \gamma'_\sigma \wedge d\eta = \int a,\eta \frac{1+\bar\eta\sigma}{\bar\sigma-\bar\eta}\frac{i}{2}(1+\eta\bar\eta)^2 d\mu_\eta,$$

the asymptotic limit of γ'_σ as $r\to\infty$ is

$$-4\pi a(u)\frac{d\sigma}{1+\sigma\bar\sigma}\delta(\eta;\sigma). \tag{7.8}$$

Again the behavior of a with respect to its argument is given by

$$a(z)\sim z^{-\epsilon} \text{ as } z\to+\infty. \tag{7.9}$$

Assuming that $a(l(\sigma,\bar\sigma),\sigma,\bar\sigma)$ satisfies (7.9) for all $(\sigma,\bar\sigma)$ with respect to its first argument, the general case follows by linearity of the superposition. The connection is given by

$$\gamma = \int \gamma'_\sigma d\mu_\sigma = \int \dot a(l(\sigma,\bar\sigma),\sigma,\bar\sigma)d\bar\delta(\sigma,\bar\sigma)d\mu_\sigma$$

and has an asymptotic limit and defines a connection on $\mathscr{I}^+$ which at the point $(u,\eta,\bar\eta)$ has the form

$$-4\pi a(u,\eta,\bar\eta)\frac{d\eta}{1+\eta\bar\eta}.$$

The asymptotic curvature at the same point is

$$-4\pi\dot a(u,\eta,\bar\eta)\frac{du\wedge d\eta}{1+\eta\bar\eta}.$$

The asymptotic condition (7.9) ensures that the plane wave components decouple asymptotically and is consistent with other estimations of the asymptotic behavior.[1]

8. ASYMPTOTIC ANALYSIS OF SELF-DUAL YANG–MILLS FIELDS

Given the asymptotic analysis of electromagnetic theory in Sec. 7 the correct choice of G, since it is defined only up to $G\to AG$, will lead to similar results for the self-dual Yang–Mills connections and fields. The function a will again be assumed to have the behavior given in (7.9) with respect to its first argument and Eq. (5.13), giving the relationship between a and G, will be written in integral form as

$$G(u,r,\eta,\bar\eta;\zeta,\bar\zeta) = I - \int G(u,r,\eta,\bar\eta;\sigma,\bar\sigma)a\left(u+r\frac{(\eta-\sigma)(\bar\eta-\bar\sigma)}{(1+\eta\bar\eta)(1+\sigma\bar\sigma)},\sigma,\bar\sigma\right)L(\zeta,\sigma)d\mu_\sigma. \tag{8.1}$$

In writing (5.13) as the integral equation given in (8.1) a particular choice has been made to fix the freedom given by A at the beginning of this paragraph, namely the ${}_0Y_{00}$ part of G is I. Examining the limiting solution to (8.1) for fixed $u,\eta,\bar\eta$ as $r\to\infty$, the function a in the integrand has the behavior

$$a\to\begin{cases}0 & \sigma\neq\eta\\ a(u,\eta,\bar\eta) & \sigma=\eta\end{cases}$$

and thus as an integrable function has support on a set of measure zero. The limit of G as $r\to\infty$ is I and so

$$\gamma_\sigma = \frac{1}{4\pi}[dG\cdot G^{-1} + G\dot a d\bar\delta l G^{-1}] \to \frac{1}{4\pi}\dot a d\bar\delta l.$$

As for the electromagnetic field, the asymptotic connection on $\mathscr{I}^+$ is

$$-a(u,\eta,\bar\eta)\frac{d\eta}{1+\eta\bar\eta} \tag{8.2}$$

and on the asymptotic field is

$$-\dot a(u,\eta\bar\eta)\frac{du\wedge d\eta}{1+\eta\bar\eta}. \tag{8.3}$$

The connection on $\mathscr{I}^+$ given by (8.2) is with respect to a section of the bundle that is parallelly propagated along the generators of $\mathscr{I}^+$ [these are given by $(\eta,\bar\eta)$ constant] and such a section is specified by giving it on one cut of $\mathscr{I}^+$, then the requirement of parallel propagation will define the section over all of $\mathscr{I}^+$. It is possible to start with a section at the point I^+ (or I^+) and parallelly propagate this section along the generators of $\mathscr{I}^+$ to obtain a section of the bundle over $\mathscr{I}^+$ with only the ambiguity of the choice of an element of the bundle in the fiber above a single point. This ambiguity is that which appeared in Sec. (5), where the $b=k\dot a k^{-1}$ represented the freedom in the plane wave data in the superposition. This represents an asymptotic change of section and changes the field on $\mathscr{I}^+$ given in (8.3) by

$$-\dot a(u,\eta,\bar\eta)\frac{du\wedge d\eta}{1+\eta\bar\eta} \to -k\dot a k^{-1}(u,\eta,\bar\eta)\frac{du\wedge d\eta}{1+\eta\bar\eta}.$$

Since $a(u,\eta,\bar\eta)$ represents free data for a self-dual solution to the Yang–Mills free field equations, it is possible to define the sum of two such fields (each being asymptotically flat). In order to remove the ambiguity above, choose a point in the fiber over I^+, say, and use the data that results from parallelly propagating each set of free data along the generators of $\mathscr{I}^+$, as discussed above, and then add the resulting free data. The decomposition of the field into positive and negative frequency components can likewise be carried out by Fourier-analyzing the free data on the generators of $\mathscr{I}^+$.

Another feature, first exibited by Newman,[1] is that despite the nonlinearities of the field equations and of the relation between the connection and curvature forms, the asymptotically flat fields have the property that when viewed as propagating from $\mathscr{I}^-$ to $\mathscr{I}^+$ they behave like linear fields (e.g., electromagnetism). This is clearly and easily exibited since in this section, the asymptotic form of G is the identity I whether on $\mathscr{I}^-$ or $\mathscr{I}^+$. Use retarded null polar coordinates $(v,r,\eta,\bar\eta)$ related to the advanced null polar coordinates $(u,r,\eta,\bar\eta)$ of Sec. 7 by $v=u+r$. Now examining integral expressions like (7.6) and (7.7), where the integrals are over two-spheres v,r constant, the expression like (7.6) vanishes and the expression like (7.7) gives a nonzero asymptotic limit for the connection at $(v,\eta,\bar{\eta})$,

$$+\,4\pi a(v)\frac{d\sigma}{1+\sigma\bar{\sigma}}\delta\Big(\eta,\bar{\eta};-\frac{1}{\bar{\sigma}},-\frac{1}{\sigma}\Big),$$

for a single plane wave and for a general asymptotically flat solution the connection on $\mathscr{I}^-$ at the point $(v,\eta,\bar{\eta})$ has the form

$$-\,4\pi\frac{\eta}{\bar{\eta}}a\Big(v,-\frac{1}{\bar{\eta}},-\frac{1}{\eta}\Big)\frac{d\bar{\eta}}{1+\eta\bar{\eta}}.$$

But the map on S^2 from the point labelled by σ to the point labelled by $-1/\bar{\sigma}$ is the antipodal map and thus even for asymptotically flat Yang–Mills fields this linearity of behavior between the data on $\mathscr{I}^-$ and the data on $\mathscr{I}^+$ is exibited, where the connection on $\mathscr{I}^-$ at $(v,\eta,\bar{\eta})$ is

$$-\frac{\eta}{\bar{\eta}}a\Big(v,-\frac{1}{\bar{\eta}},-\frac{1}{\eta}\Big)\frac{d\bar{\eta}}{1+\eta\bar{\eta}}.$$

9. AN EXAMPLE

The simplist example of a self-dual field in the present context would be a single plane wave. Take for the plane wave

$$\gamma'_\sigma=\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})d\bar{\eth}l(\sigma,\bar{\sigma}), \tag{9.1}$$

where as before σ specifies the propagation direction. Then the equation for G is

$$G^{-1}\bar{\eth}_\zeta G(x^a,\zeta,\bar{\zeta})+a(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})\delta(\zeta,\bar{\zeta};\sigma,\bar{\sigma})=0. \tag{9.2}$$

This has a solution

$$G(x^a,\zeta,\bar{\zeta})=e^{-aL(\zeta,\sigma)} \tag{9.3}$$

as can be verified by writing G as a power series and inserting it into (9.2). Thus

$$\begin{aligned}\gamma&=\frac{1}{4\pi}\int dG\cdot G^{-1}+G\gamma'_\lambda G^{-1}(x^a,\lambda,\bar{\lambda})d\mu_\lambda\\&=\frac{1}{4\pi}\int dG\cdot G^{-1}+G(x^a,\lambda)\\&\quad\times[\dot{a}(l(\sigma,\bar{\sigma}),\sigma,\bar{\sigma})d\bar{\eth}l(\lambda,\bar{\lambda})\delta(\lambda,\sigma)]G^{-1}(x^a,\lambda,\bar{\lambda})d\mu_\lambda.\end{aligned} \tag{9.4}$$

Evaluating the integrals in (9.4) involves for the first term evaluating

$$\int L^n(\lambda,\sigma)d\mu_\sigma$$

for $n=1,2\cdots$. Using Stoke's theorem, these can be shown to all be zero. For $n=1$ this has been discussed previously and for $n=2$, the following exhibits the general pattern:

$$\begin{aligned}\int L^2(\lambda,\sigma)d\mu_\lambda&=\Big(\frac{1}{4\pi}\Big)^2\frac{2}{i}\int\Big(\frac{1+\bar{\lambda}\sigma}{\bar{\lambda}-\bar{\sigma}}\Big)^2\frac{d\lambda nd\bar{\lambda}}{(1+\lambda\bar{\lambda})^2}\\&=-\Big(\frac{1}{4\pi}\Big)^2\frac{2}{i}\oint+\frac{(1+\bar{\lambda}\sigma)^2}{\bar{\lambda}-\bar{\sigma})^2\bar{\lambda}(1+\lambda\bar{\lambda})}d\bar{\lambda}\\&=-\Big(\frac{1}{4\pi}\Big)^2\frac{2}{i}(-2\pi i)\Big[\frac{1}{\bar{\sigma}^2}+\frac{2\sigma}{\bar{\sigma}}-\frac{\sigma}{\bar{\sigma}}-\frac{(1+\sigma\bar{\sigma})}{\bar{\sigma}^2}\Big]=0.\end{aligned}$$

The first term in the brackets comes from the residue at $\bar{\lambda}=0$ while the remaining terms come from the residue at $\bar{\lambda}=\bar{\sigma}$. Thus the first term in the integral is zero. The second term involves

$$\int L^n(\lambda,\sigma)\delta(\lambda,\sigma)d\bar{\eth}l(\lambda,\bar{\lambda})d\mu_\lambda$$

for $n=0,1,2,\cdots$. These all vanish with the exception of the term with $n=0$. The reason is that for $n>0$, the evaluation of the integral gives $\bar{\eth}^{n+1}l(\sigma,\bar{\sigma})$, which vanishes for $n>0$. Thus the connection given by (9.4) is just $(1/4\pi)\gamma'_\sigma$ of (9.1) despite the complicated nature of G given by (9.3).

This calculation also lends insight into the asymptotic behavior of G in Sec. 8 since the only difference between the asymptotic behavior of "a" in that section and here is that in Sec. 8 a has nonsingular support on the sphere and in this section has singular support on the sphere but in each case support on the same measure-zero set. In the first instance $G=I$ results and here G is given by (9.3).

10. DISCUSSION

Self-dual solutions to the Yang–Mills free field equations have been constructed in terms of nonlinear superpositions of plane waves. Given the plane wave data, a Lie algebra valued function of three variables, the superposition is mediated via the solution of a linear equation which establishes the correct representation for each of the plane wave components. The superposition is achieved by integrating over the two-sphere of propagation directions and the self-duality of the resulting curvature form is verified by explicitly exibiting that feature. The conditions for the resulting field to be asymptotically flat are established and the asymptotic connection and curvature are given on $\mathscr{I}^+$ and $\mathscr{I}^-$. The linear relationship between the fields on $\mathscr{I}^-$ and on $\mathscr{I}^+$ is observed easily. Finally, an example of the procedure is worked out in some detail: that of a single plane wave.

In conclusion, in another paper the precise mathematical context for the formulation of plane waves and superposition will be given. That formulation is largely ignored in the present paper as the present work is rather complete as presented and involves few ideas of a highly technical nature. The appropriate mathematical context involves ideas from twistor theory and deals with extensions of Atiyah and Ward's work[7] on the self-dual fields.

A similar, but more complicated, version of the methods herein employed yields formulation of the (not necessarily self-dual) Yang–Mills fields but the description of the satisfaction of the free field equations is considerably more awkward. On the other hand a version of this work exists for the self-dual nonlinear graviton.[8] The results in that context parallel very closely those given in the present context.

Finally, it is hoped that analysis of nonlinear fields in the familiar context of "superposition" of plane waves or in terms of momentum eigenstates will lead to a better methodology for quantization, at least for asymptotic quantization, and an S-matrix formulation of interactions between nonlinear fields.

ACKNOWLEDGMENTS

The author would like to thank the relativity group at the University of Pittsburgh, in particular Professor E. T. Newman and Dr. G. A. J. Sparling, for many hours of extremely useful discussions.

[1]E. T. Newman, Phys. Rev. D 22, 3023 (1980).
[2]I. Bialynicki-Birula, B. Lukács, E. T. Newman, Z. Perjés, J. R. Porter, A. Sebestyen, and J. Winicour, J. Math. Phys. 22, 2530 (1981).
[3]E. T. Newman and R. Penrose, J. Math. Phys. 7, 863 (1966).
[4]M. Eastwood and K. P. Tod, Proc. Cambridge Philos. Soc. (to be published).
[5]D. Lerner and J. R. Porter (in preparation).
[6]J. R. Porter, Gen. Relativ. Gravit. (to be published).
[7]M. R. Atiyah and R. Ward, Commun. Math. Phys. 55, 117 (1977).
[8]J. R. Porter "The nonlinear graviton: Superposition of plane waves," Gen. Relativ. Gravit. (to be published).

Self-dual Yang–Mills fields on Minkowski space–time

John R. Porter
Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 2 February 1982; accepted for publication 16 April 1982)

Self-dual Yang–Mills fields are realized as a superposition, but in a nonlinear context, of plane wave fields. The mathematical analysis of such superposition (and the inverse problem of decomposition) is given in terms of connections on the prime spin bundle over Minkowski space–time. This view leads naturally to the twistor construction for such fields given by Ward as well as a CR version of this construction.

PACS numbers: 11.15. − q

I. INTRODUCTION

Given a self-dual connection on a principal bundle over Minkowski space–time, this connection is treated as a distributional connection over the prime spin bundle in Sec. II. The problem of the (nonlinear) superposition of such connections to yield a self-dual curvature tensor is then discussed, and it is found that in the appropriate context such a superposition is possible for general self-dual plane wave connections. The inverse problem of decomposing a self-dual connection into self-dual plane wave connections is considered in Sec. III. Arguments are produced which demonstrate that such a decomposition is possible in quite general circumstances, at least for the radiative portions of the fields. An interpretation of the construction in terms of twistor concepts and CR structures is given in Sec. IV, and this ties the present work in with the previous work of Ward on this topic. A less mathematical treatment of some aspects of this superposition is given in Ref. 1.

II. SUPERPOSITION OF PLANE WAVES

Consider the following spaces and coordinate systems used in defining the relations between the spaces,

M: Minkowski space–time with standard Minkowskian coordinates x^a of real dimension four and signature -2.

P': the projective prime spin bundle over M with coordinates $(x^a, [\pi_{A'}])$ of real dimension six. This is a bundle over M with fibers $P\mathbb{C}^1 \approx S^2$ and $[\pi_{A'}]$ denotes the equivalence class of primed spinors where the equivalence relation is

$$\mu_{A'} \sim \pi_{A'} \quad \text{if } \lambda \in \mathbb{C}^*, \ \mu_{A'} = \lambda \pi_{A'}.$$

$\widetilde{PN}$: the space of null geodesics in M with coordinates $[(x^a, [\pi_{A'}])]$ of real dimension five. Here $\widetilde{PN}$ denotes $PN - I$ (null projective twistor space PN minus the line at infinity I) and $[(x^a, [\pi_{A'}])]$ is the equivalence class of points of P', where the equivalence relation is

$$(y^a, [\mu_{A'}]) \sim (x^a, [\pi_{A'}])$$

if $[\mu_{A'}] = [\pi_{A'}]$ and $\beta \in R$ so that $y^{AA'} - x^{AA'} = \beta \bar{\pi}^A \pi^{A'}$, where $y^{AA'}$ and $x^{AA'}$ denote the spinor transform of y^a and x^a, respectively. P' is a bundle over $\widetilde{PN}$ with fiber R^1.

$\widetilde{NH}$: the space of null hyperplanes in M with coordinates $\{(x^a, [\pi_{A'}])\}$ of real dimension three. $\{(x^a, [\pi_{A'}])\}$ denotes the equivalence class of points of P', where the equivalence relation is

$$(y^a, [\mu_{A'}]) \sim (x^a, [\pi_{A'}])$$

if $[\mu_{A'}] = [\pi_{A'}]$ and $(y^{AA'} - x^{AA'})\bar{\pi}_A \pi_{A'} = 0$. $\widetilde{PN}$ is a bundle over $\widetilde{NH}$ with fibers R^2 and P' is a bundle over $\widetilde{NH}$ with fibers R^3. $\widetilde{NH}$ can be identified with $\mathscr{I}^+$ or with $\mathscr{I}^-$, the two components of the null cone of the point at infinity in a conformal completion of M. All the above spaces can be defined for the conformal completion of M if desired; however, additional structures on these spaces will have to be examined for extendability. Insofar as M is concerned, all these constructions are local. For the complexification, $\mathbb{C}M$, of M like structures exist and will be used at a later stage of this paper.

Schematically, the following diagram is obtained,

$$M \leftarrow P' \rightarrow \widetilde{PN} \rightarrow \widetilde{NH}.$$

Consider now a principal fiber bundle B_M over M (though an associated vector bundle could also be considered), and denote its Lie group by $\mathfrak{G}$ and the corresponding Lie algebra by $\mathfrak{g}$. Then the map $P' \rightarrow M$ induces a bundle $B_{P'}$ over P' (the pullback of the bundle $B_M \rightarrow M$ under $P' \rightarrow M$), which is trivial over the fibers of $P' \rightarrow M$. Since the fibers of $P' \rightarrow \widetilde{PN}$ are R^1, the bundle $B_{P'}$ is trivial over these fibers and a bundle $B_{\widetilde{PN}}$ results over $\widetilde{PN}$. Thus

$$\begin{array}{ccccc} B_M & \rightarrow & B_{P'} & \rightarrow & B_{\widetilde{PN}} \\ \downarrow & & \downarrow & & \downarrow \\ M & \leftarrow & P' & \rightarrow & \widetilde{PN} \end{array}$$

is obtained. The bundle $P' \rightarrow M$ possesses canonical sections labeled by $[\pi_{A'}]$, which form a foliation of $P' \rightarrow M$,

$$S_{[\pi_{A'}]}: M \rightarrow P',$$

$$S_{[\pi_{A'}]}(x^a) = (x^a, [\pi_{A'}]).$$

These sections are formed from horizontal lifts of curves using the spin connection induced from the Minkowskian connection.

Consider a distributional connection on $B_{P'}$ represented with respect to some section $u: P' \rightarrow B_{P'}$ as

$$\gamma_u(x^a, [\pi_{A'}]) = \begin{cases} 0, & [\pi_{A'}] \neq [\alpha_{A'}], \\ \dot{a}\left(\dfrac{x^{AA'}\bar{\alpha}_A \alpha_{A'}}{t^{AA'}\bar{\alpha}_A \alpha_{A'}}\right)\bar{\alpha}_B t^C{}_{B'} \bar{\alpha}_C \, dx^{BB'}, & [\pi_{A'}] = [\alpha_{A'}], \end{cases}$$

 0022-2488/83/051233-07$02.50

where $\dot{a}$ denotes the derivative of a with respect to its argument and a has values in $\mathfrak{g}$. $t^{AA'} = (I)^{AA'}$, when I is the identity matrix. Then consider $S^*_{[\alpha_{A'}]}\gamma_u$, a connection on B_M since $S_{[\alpha_{A'}]}$ is a section of $P' \to M$ so the section u of $B_{p'} \to P'$ pulls back via $S_{[\alpha_{A'}]}$ to a section of $B_M \to M$. Then

$$\gamma = s^*_{[\alpha_{A'}]}\gamma_u = \dot{a}\left(\frac{x^{AA'}\overline{\alpha}_A\alpha_{A'}}{t^{AA'}\overline{\alpha}_A\alpha_{A'}}\right)\overline{\alpha}_B t^C{}_{B'}\overline{\alpha}_C\, dx^{BB'},$$

which represents a connection on B_M with self-dual curvature form, a plane wave with propagation direction determined by $\overline{\alpha}_A\alpha_{A'}$. The curvature form

$$F = d\gamma - \gamma \wedge \gamma$$

is given by

$$F = \ddot{a}\left(\frac{x^{AA'}\overline{\alpha}_A\alpha_{A'}}{t^{AA'}\overline{\alpha}_A\alpha_{A'}}\right)\overline{\alpha}_A\overline{\alpha}_B\epsilon_{A'B'}\, dx^{AA'} \wedge dx^{BB'}$$

since the quadratic terms in F vanish identically. In other words, plane wave self-dual Yang–Mills fields on M are considered as distributional fields on P'. The aim is to exhibit more general fields on M, and the plane wave fields will be used as the basic components in this representation, just as in the case of electromagnetism which is the simplest Yang–Mills field.

Suppose with respect to a section u of $B_{p'} \to P'$ a connection is given on P',

$$\gamma_u(x^a, [\pi_{A'}]) = \dot{a}(x^{AA'}\overline{\pi}_A\pi_{A'}, \overline{\pi}_A, \pi_{A'}) \times \overline{\pi}_B t^C{}_{B'}\overline{\pi}_C\, dx^{BB'}\, t^{DD'}\overline{\pi}_D\pi_{D'}, \quad (2.1)$$

where a is a function on P' homogeneous of degree $(-2,0)$ in $(\overline{\pi}_A, \pi_{A'})$ and $\dot{a}$ denotes the derivative of a with respect to its first argument, $x^{AA'}\overline{\pi}_A\pi_{A'}$. Then $\dot{a}$ is of homogenity degree $(-3,-1)$ in $(\overline{\pi}_A, \pi_{A'})$, and the entire expression is of degree $(0,0)$ in $(\overline{\pi}_A, \pi_{A'})$ and thus defined on P'. For the moment a will be taken to be C^∞ with respect to its arguments on P' or a part of P' that is the pullback of an open subset of M under the map $P' \to M$. The function a could be taken to be distributional valued on the fibers, but C^∞ is easier to deal with. Note that entire fibers are used. This connection pulls back via $S_{[\alpha_{A'}]}$ to a plane wave connection on M with propagation direction determined by $\overline{\alpha}_A\alpha_{A'}$ and phase and amplitude profile given by the function a and its derivatives evaluated for $\pi_{A'} = \alpha_{A'}$. It is desired to construct a self-dual field on M from the plane wave components. While in electromagnetism (or whenever the group $\mathfrak{G}$ is abelian) it is possible to "sum" the individual plane wave connections, in the general nonabelian case this naive procedure doesn't produce a self-dual curvature form. But proceeding more carefully, the same type of superposition is, in fact, possible and will result in a self-dual field on M.

Given the section u of $B_{p'} \to P'$ and the connection γ_u on P', note that the lift of a curve in P' to a curve in $u(P')$ is horizontal if the tangent vector to the lifted curve is annihilated by $\overline{\pi}_B t^C{}_{B'}\overline{\pi}_C\, dx^{BB'}$ at the point in the fiber over $(x^a, [\pi_{A'}])$. Thus in particular the lifts of the fibers of $P' \to \widetilde{PN}$ are horizontal. Since u is horizontal on the fibers of $P' \to \widetilde{PN}$, the bundle $B_{p'} \to P$, *canonically* induces a bundle $B_{\widetilde{PN}} \to \widetilde{PN}$. Tangentially, for $M \to \mathbb{C}M$, the anti-self-dual 2-planes (α-planes) are horizontal. Here $\mathbb{C}M$ denotes complex Minkowski space–time.

Given a section of $B_M \to M$, $s: M \to B_M$, a section s' of $B_{p'} \to P'$, $s': P' \to B_{p'}$ is induced via the commutative diagram:

$$\begin{array}{ccc} B_M & \rightarrow & B_{p'} \\ s \uparrow\downarrow & & \uparrow\downarrow s' \\ M & \leftarrow & P' \end{array}$$

Let G represent the $\mathfrak{G}$-valued function on P' taking the section u to the section s' in $B_{p'} \to P'$. With respect to the section s', the connection form on P' has the representation

$$\gamma' = DG \cdot G^{-1} + G\gamma G^{-1},$$

where D denotes the exterior derivative operator on P'. Given the section s of $B_M \to M$ and the induced section s' of $B_{p'} \to P'$, how is the section u of $B_{p'} \to P'$ chosen (alternately, how is G chosen) so that the fiber integral of γ' over the fibers of $P' \to M$ with respect to the normalized measures on the fibers is the connection of a self-dual curvature form (γ' = self-dual connection) defined on M? That is, it is desired that

$$\Gamma(x) = \frac{1}{4\pi}\int_{S^2}(DG \cdot G^{-1} + G\gamma G^{-1}) \wedge \frac{\Delta\overline{\pi} \wedge \Delta\pi}{t^2(\overline{\pi}, \pi)} \cdot \frac{2}{i} \quad (2.2)$$

is the connection of a self-dual curvature on M. Here

$$\Delta\overline{\pi} \equiv \epsilon^{AB}\overline{\pi}_A d\overline{\pi}_B, \quad \Delta\pi \equiv \epsilon^{A'B'}\pi_{A'} d\pi_{B'}$$

and

$$\frac{1}{4\pi}\int_{S^2}\frac{\Delta\overline{\pi} \wedge \Delta\pi}{(t^{AA'}\overline{\pi}_A\pi_{A'})^2} \cdot \frac{2}{i} \equiv \frac{1}{4\pi}\int_{S^2} d\mu_\pi = 1.$$

In the abelian case G can be chosen arbitrarily, but in the nonabelian or generic case there is nontrivial condition on G (or a nontrivial relationship between a and G) in order that Γ is a self-dual connection. Γ is a 1-form on M with values in the Lie algebra $\mathfrak{g}$ of $\mathfrak{G}$. Given the wedge product in the integrand, the only surviving part of D is just d, the exterior derivative operator on M pulled back to P'.

Instead of deriving the curvature two-form of the connection Γ and demanding that it be self-dual, an equivalent and more useful criterion for Γ to be a self-dual connection is that Γ pulled back tangentially to an anti-self-dual 2-surface be trivial or be the zero connection up to choice of section. More specifically, consider for a moment $\mathbb{C}M$, complexified Minkowski space and the injection map

$$i_{[\lambda]}: \mathbb{C}^2 \to \mathbb{C}M$$

where $i_{[\lambda]}\mathbb{C}^2$ is an anti-self dual 2-plane in $\mathbb{C}M$ (an α-plane) whose normal 2-form, labeled by $\lambda_{A'}$, i s $\epsilon_{AB}\lambda_{A'}\lambda_{B'}dx^{AA'} \wedge dx^{BB'}$, of which there is a two-complex-parameter family, one through each point of $\mathbb{C}M$ (though this labeling doesn't distinguish members of the family). The pullback under $i^*_{[\lambda]}$ of a self-dual 2-form $\epsilon_{A'B'}\overline{\pi}_A\overline{\pi}_B dx^{AA'} \wedge dx^{BB'}$ is zero. But

$$\epsilon_{A'B'}\overline{\pi}_A\overline{\pi}_B dx^{AA'} \wedge dx^{BB'} = \frac{\overline{\pi}_A\pi_{A'}dx^{AA'} \wedge \overline{\pi}_B t^C{}_{B'}\overline{\pi}_C\, dx^{BB'}}{t^{AA'}\overline{\pi}_A\pi_{A'}}$$

and thus the two 1-forms $i^*_{[\lambda]}\overline{\pi}_A\pi_{A'}\, dx^{AA'}$ and $i^*_{[\lambda]}\overline{\pi}_B t^C{}_{B'}\overline{\pi}_C\, dx^{BB'}$ are proportional. To obtain the factor of proportionality, note that $\lambda_{A'}\, dx^{AA'}$ spans the space of normal forms to $i_{[\lambda]}\mathbb{C}^2$. The tangent vectors of the form

$\lambda^{A'}\nabla_{AA'} = \lambda^{A'}\partial/\partial x^{AA'}$ are annihilated by the normal form and thus span the tangent space of $i_{[\lambda]}\mathbb{C}^2$. The 1-forms dual to $\lambda^{A'}\nabla_{AA'}$ are $-t^{B}{}_{A'}\bar{\lambda}_B\, dx^{AA'}/t^{CC'}\bar{\lambda}_C\lambda_{C'}$. Thus given any 1-form $\theta_{AA'}\, dx^{AA'}$, its pullback via $i^{A}_{[\lambda]}$ is the same as $(\theta_{AA'} + \alpha_A\lambda_{A'})\, dx^{AA'}$. Only the primed-part of $\theta_{AA'}$ is in question and so consider $\theta_{A'}$ in a spinor basis $t^{A}{}_{A'}\bar{\lambda}_A$ and $\lambda_{A'}$, and then

$$\theta_{A'} = \frac{\theta^{B'}\lambda_{B'}}{t^{CC'}\bar{\lambda}_C\lambda_{C'}}\, t^{A}{}_{A'}\bar{\lambda}_A + \frac{t^{BB'}\bar{\lambda}_B\theta_{B'}}{t^{CC'}\bar{\lambda}_C\lambda_{C'}}\,\lambda_{A'}.$$

Thus, substituting $\pi_{A'}$ and $t^{C}{}_{A'}\bar{\pi}_C$ for $\theta_{A'}$, $i^*_{[\lambda]}$ of the forms in question are represented by

$$i^*_{[\lambda]}\bar{\pi}_A\pi_{A'}\, dx^{AA'} = i^*_{[\lambda]}\frac{\pi^{B'}\lambda_{B'}}{t^{CC'}\bar{\lambda}_C\lambda_{C'}}\,\bar{\pi}_A t^{D}{}_{A'}\bar{\lambda}_D\, dx^{AA'}$$

and

$$i^*_{[\lambda]}\bar{\pi}_A t^{B}{}_{A'}\bar{\pi}_B\, dx^{AA'} = i^*_{[\lambda]}\frac{t^{BB'}\lambda_{B'}\bar{\pi}_B}{t^{CC'}\bar{\lambda}_C\lambda_{C'}}\,\bar{\pi}_A t^{D}{}_{A'}\bar{\lambda}_D\, dx^{AA'}.$$

Finally,

$$i^*_{[\lambda]}\bar{\pi}_A t^{B}{}_{A'}\bar{\pi}_B\, dx^{AA'} = \frac{t^{BB'}\bar{\pi}_B\lambda_{B'}}{\epsilon^{C'D'}\lambda_{C'}\pi_{D'}}\, i^*_{[\lambda]}\bar{\pi}_A\pi_{A'}\, dx^{AA'}.$$

As a consequence, under $i^*_{[\lambda]}$, Γ pulls back to the same form as $\Sigma(x,[\lambda])$, where

$$\Sigma(x,[\lambda]) = \int_{S^2} [1/4\pi) dG\cdot G^{-1}(x,\bar{\pi},\pi) - L(\lambda,\pi) G\, da\, G^{-1}(x,\bar{\pi},\pi)] \wedge d\mu_\pi. \tag{2.3}$$

Here the function $L(\lambda,\pi)$ is defined to be

$$L(\lambda,\pi) = -\frac{t^{BB'}\bar{\pi}_B\lambda_{B'}}{4\pi\epsilon^{C'D'}\lambda_{C'}\pi_{D'}}\cdot t^{AA'}\bar{\pi}_A\pi_{A'}. \tag{2.4}$$

The homogenities of the entire expression in $(\bar{\pi},\pi)$ and in $(\bar{\lambda},\lambda)$ are both zero so that the 1-form Σ is naturally defined on P' and is pulled back via $S^*_{[\lambda]}$ to M and via $(i_{[\lambda]}\cdot S_{[\lambda]})^*$ to $\mathbb{C}^2$. The function $L(\lambda,\pi)$ defined on the fibers of $P'\to M$ is of homogenity degree (2,0) in $(\bar{\pi},\pi)$ and (0,0) in $(\bar{\lambda},\lambda)$ and has the following properties:

$$\int_{S^2} L(\lambda,\pi)\, d\mu_\lambda = 0$$

and

$$\frac{\partial}{\partial\bar{\lambda}_A} L(\lambda,\pi) = -\bar{\lambda}^A\delta(\lambda,\pi),$$

where $\delta(\lambda,\pi)$ is a distribution with singular support where $[\pi] = [\lambda]$. If $f(\bar{\pi},\pi)$ has homogenity degree $(-2,0)$ and is regular on S^2, then defining $g(\bar{\lambda},\lambda)$ by

$$g(\bar{\lambda},\lambda) = \int_{S^2} L(\lambda,\pi) f(\bar{\pi},\pi)\, d\mu_\pi$$

results in a function of degree (0,0) in $(\bar{\lambda},\lambda)$ which is regular on S^2 and satisfies

$$\frac{\partial g}{\partial\bar{\lambda}_A} = -\bar{\lambda}^A f(\bar{\lambda},\lambda) \quad \text{or} \quad \eth g = f,$$

and

$$\int_{S^2} g(\bar{\lambda},\lambda)\, d\mu_\lambda = 0.$$

The second expression states that g is orthogonal to the kernal of $\partial/\partial\bar{\lambda}_A$ on functions which are regular on S^2 and degree (0,0). The general solution to $\eth g = f$ is given by

$$g(\bar{\lambda},\lambda) = \int_{S^2} \frac{1}{4\pi} g(\bar{\pi},\pi) + L(\lambda,\pi) f(\bar{\pi},\pi)\, d\mu_\pi \tag{2.5}$$

since the first term on the right-hand side is in the kernel of $\eth$.

The condition that Γ be trivial upon pull back to tangential anti-self-dual 2-planes on $M\to\mathbb{C}M$ labeled by $[\lambda]$ can now be expressed as the requirement that $S^*_{[\lambda]}\Sigma(x,[\lambda])$ will trivial for fixed but arbitrary $[\lambda]$ or that $S^*_{[\lambda]}\Sigma(x,[\lambda])$ represent a connection on M with zero curvature. That is to say, there exists a $\mathfrak{G}$-valued function H on P' such that

$$dH\cdot H^{-1}(x,[\lambda]) = \Sigma(x,[\lambda])$$

or

$$dH\cdot H^{-1}(x,[\lambda]) = \int_{S^2} \{(1/4\pi) dG\cdot G^{-1}(x,[\pi]) - L(\lambda,\pi) G\, da\, G^{-1}(x,[\pi])\} \wedge d\mu_\pi. \tag{2.6}$$

Upon fiber integration of (2.6) with respect to $d\mu_\lambda$

$$\int dH\cdot H^{-1}(x,[\lambda]) \wedge d\mu_\lambda = \int dG\cdot G^{-1}(x,[\pi]) \wedge d\mu_\pi \tag{2.7}$$

results, and $\partial/\partial\bar{\lambda}_A$ on (2.6) produces

$$\frac{\partial}{\partial\bar{\lambda}_A} dH\cdot H^{-1}(x,[\lambda]) = \bar{\lambda}^A G\, da\, G^{-1}(x,[\lambda]). \tag{2.8}$$

Multiplying (2.8) on the left by H^{-1} and on the right by H, the equation can be written as

$$d\left(H^{-1}\frac{\partial H}{\partial\bar{\lambda}_A}\right) - \bar{\lambda}^A K\, da\, K^{-1} = 0, \tag{2.9}$$

where $K = H^{-1}G$. The first term in (2.9) is exact, and thus the second term is also exact; thus there exists a g-valued function b on P' so that

$$db = K\, da\, K^{-1}$$

and

$$d\left(H^{-1}\frac{\partial}{\partial\bar{\lambda}_A} H\right) - \bar{\lambda}^A\, db = 0.$$

This equation can be integrated to yield

$$H^{-1}\frac{\partial}{\partial\bar{\lambda}_A} H - \bar{\lambda}^A b = 0$$

modulo an additive function on the fibers, which is ignorable. Multiplying by $d\bar{\lambda}_A$,

$$H^{-1}\bar{\partial}H - b\Delta\bar{\lambda} = 0 \tag{2.10}$$

results where $\bar{\partial}H = (\partial/\partial\bar{\lambda}_A) H\, d\bar{\lambda}_A$. Given (2.7) and $HK = G$.

$$dG\cdot G^{-1} = (dH\cdot K + H dK)(HK)^{-1} = dH\cdot H^{-1} + H(dK\cdot K^{-1})H^{-1},$$

and so

$$\int H(dK{\cdot}K^{-1})H^{-1}\wedge d\mu = 0;$$

as a consequence

$$\begin{aligned}\Gamma(x) &= (1/4\pi)\int [dH{\cdot}H^{-1} + H(dK{\cdot}K^{-1} \\ &\quad + K\gamma K^{-1})H^{-1}](x,[\pi])\wedge d\mu_\pi \\ &= (1/4\pi)\int [DH{\cdot}H^{-1} + H(DK{\cdot}K^{-1} \\ &\quad + K\gamma K^{-1})H^{-1}]\wedge d\mu_\pi .\end{aligned}$$

Thus the $\mathfrak{G}$-valued function K on P' can be viewed as giving the plane wave components γ in a section related to the section u by K. Without loss of generality, K will be taken to be the identity and $H = G$ results. Then

$$\Gamma(x) = (1/4\pi)\int_{S^2}(DG{\cdot}G^{-1} + G\gamma G^{-1})\wedge d\mu,$$

and the equation for G relating the sections u and S' which makes Γ into a self-dual connection on M is

$$dG{\cdot}G^{-1}(x,[\lambda]) = \int_{S^2}[(1/4\pi)dG{\cdot}G^{-1}(x,[\pi] - L(\lambda,\pi)G\,da\,G^{-1}(x,[\pi])]\wedge d\mu_\pi \tag{2.11}$$

or, in equivalent differential form,

$$G^{-1}\bar\partial G - a\Delta\bar\lambda = 0. \tag{2.12}$$

In summary, given plane wave data for a self-dual Yang–Mills connection, the superposition of the data to produce a self-dual connection results in a linear equation for a $\mathfrak{G}$-valued function on P'. Having solved this equation, the solution is then used to construct the self-dual connection on M or a piece thereof.

III. THE INVERSE PROBLEM

Given a self-dual connection $\Gamma(x)$, obtaining the decomposition into plane wave components for an abelian group is simply a matter of using standard procedures involving Fourier analysis of Maxwell theory. If, however, the group is not abelian, the procedure is slightly more complicated.

Since $\Gamma = \Gamma_{AA'}\,dx^{AA'}$ pulled back tangentially by $i^*_{[\pi]}$ to an anti-self-dual 2-plane is trivial, there exists a $\mathfrak{G}$-valued function $G(x,[\pi])$ on P', such that

$$(S_{[\pi]}{\cdot}i_{[\pi]})^*\,dG{\cdot}G^{-1} = i^*_{[\pi]}\Gamma.$$

These two-planes depend only on $[\pi_{A'}]$ and thus

$$\begin{aligned}&\bar\partial_\pi[(S_{[\pi]}{\cdot}i_{[\pi]})^*\,dG{\cdot}G^{-1}] \\ &\quad = (S_{[\pi]}{\cdot}i_{[\pi]})^*(\bar\partial_\pi\,dG{\cdot}G^{-1}) = 0,\end{aligned} \tag{3.1}$$

where $\bar\partial_\pi = (\partial/\partial\bar\pi_A)d\bar\pi_A$. Now $(\partial/\partial\bar\pi_A)(dG{\cdot}G^{-1})$ is at 1-form which pulls back via $(S_{[\pi]}{\cdot}i_{[\pi]})^*$ to zero, and thus it depends only on the normal forms of $\mathbb{C}^2$, that is, $\pi_{A'}\,dx^{AA'}$. And so

$$\frac{\partial}{\partial\bar\pi_B}dG{\cdot}G^{-1} = \bar\pi^B B_A\pi_{A'}\,dx^{AA'}, \tag{3.2}$$

where $\bar\pi^B$ is present on the right-hand side by virtue of the homogenity of $dG{\cdot}G^{-1}$ with respect to $\bar\pi$ which is zero and is expressed by

$$\bar\pi_B\frac{\partial}{\partial\bar\pi_B}dG{\cdot}G^{-1} = 0.$$

Equation (3.2) isn't d-exact, but, multiplying on the left by G^{-1} and on the right by G, gives

$$d\left(G^{-1}\frac{\partial}{\partial\bar\pi_B}G\right) - \bar\pi^B G^{-1}B_A G\pi_{A'}\,dx^{AA'} = 0;$$

the second term is d exact and is d of a function of $x^{AA'}$ of the form $x^{AA'}\pi_{A'}$. Thus there exists a function $a(x^{AA'}\pi_{A'},\bar\pi_A,\pi_{A'})$ on P' with values in $\mathfrak{g}$ such that

$$G^{-1}\frac{\partial}{\partial\bar\pi^B}G - \bar\pi^B a = 0. \tag{3.3}$$

Given Γ, a $\mathfrak{G}$-valued function G on P' satisfying (3.1) or alternately (using the fact that $i_{[\pi]}$ and $S_{[\pi]}$ are local differomorphisms)

$$(\pi^{A'}\nabla_{AA'}G)G^{-1} = \Gamma_{AA'}\pi^{A'} \tag{3.4}$$

is found (where $\nabla_{AA'} = \partial/\partial x^{AA'}$) and then

$$G^{-1}\frac{\partial}{\partial\bar\pi_B}G$$

is constructed. $G = G(x^{AA'},\bar\pi_A,\pi_{A'})$ and is of homogenity (0,0) in $(\bar\pi,\pi)$ and satisfies

$$\bar\pi^B\frac{\partial}{\partial\bar\pi_B}G = 0.$$

Then

$$G^{-1}\frac{\partial}{\partial\bar\pi_B}G = \bar\pi^B a$$

defines $a(x^{AA'},\bar\pi_A,\pi_{A'})$ and upon applying $\pi^{A'}\nabla_{AA'}$ to the above equation, using the fact that $\pi^{A'}\nabla_{AA'}$ and $\partial/\partial\bar\pi_B$ commute, using (3.4), $\pi^{A'}\nabla_{AA'}a = 0$ results or

$$a = a(x^{AA'}\pi_{A'},\bar\pi_A,\pi_{A'}). \tag{3.5}$$

Equation (3.3) may be written as

$$G^{-1}\bar\partial G = a\Delta\bar\pi, \tag{3.6}$$

where a is a $\mathfrak{g}$-valued function on P' satisfying (3.5) and of degree $(-2,0)$ in $(\bar\pi,\pi)$.

If the connection is given with respect to another section related to the first section by the $\mathfrak{G}$-valued function $C(x)$ on M, then the connection 1-forms are related by

$$\Gamma' = dC{\cdot}C^{-1} + C\Gamma C^{-1}.$$

The resulting equations for G' and a' are

$$
\begin{aligned}
(\pi^{A'}\nabla_{AA'}G')G^{-1} &= \Gamma'_{AA'}\pi^{A'} \\
&= (\pi^{A'}\nabla_{AA'}C)C^{-1} + C\Gamma_{AA'}\pi^{A'}C^{-1} \\
&= (\pi^{A'}\nabla_{AA'}C)C^{-1} + C(\pi^{A'}\nabla_{AA'}G)G^{-1}C^{-1} \\
&= (\pi^{A'}\nabla_{AA'}CG)(CG)^{-1},
\end{aligned}
$$

and thus it is possible to choose G' to satisfy

$$G' = CG.$$

Then

$G'^{-1}(\partial/\partial\bar{\pi}_A)G' = (CG)^{-1}(\partial/\partial\bar{\pi}_A)CG = G^{-1}(\partial/\partial\bar{\pi}_B)G$ since $C = C(x)$ and $a' = a$ results. As a consequence, the resulting $\mathfrak{g}$-valued function on P', a, resulting from a solution to (3.3) and (3.4) for a given connection may be chosen to be independent of the section with respect to which the connection 1-form is given.

The freedom in the choice of G is

$$G \to GH = J,$$

where $H(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'})$ is a $\mathfrak{G}$-valued function on P' and so of degree (0,0) in $(\bar{\pi},\pi)$. Then

$$
\begin{aligned}
(\pi^{A'}\nabla_{AA'}J)J^{-1} &= \pi^{A'}\nabla_{AA'}(G\cdot H)(G\cdot H)^{-1} \\
&= (\pi^{A'}\nabla_{AA'}G)G^{-1} + G(\pi^{A'}\nabla_{AA'}H)H^{-1}G^{-1} \\
&= (\pi^{A'}\nabla_{AA'}G)G^{-1}
\end{aligned}
$$

and $dJ\cdot J^{-1}$ pulls back to the same forms as $dG\cdot G^{-1}$ under $(S_{[\lambda]}\cdot i_{[\pi]})^*$. Then proceeding with J in place of G in Eq. (3.6) denote the result by a', where

$$
\begin{aligned}
J^{-1}(\bar{\partial}J) = a'\Delta\bar{\pi} &= (GH)^{-1}\bar{\partial}(GH) \\
&= H^{-1}(G^{-1}JG)H + H^{-1}\bar{\partial}H \\
&= H^{-1}a\Delta\bar{\pi}H + H^{-1}\bar{\partial}H
\end{aligned}
$$

and so

$$a'\Delta\bar{\pi} = H^{-1}a\Delta\bar{\pi}H + H^{-1}\bar{\partial}H. \tag{3.7}$$

Thus (3.7) represents the freedom in the $\mathfrak{g}$-valued function on P' and given a connection Γ on M such a function a is defined up to the equivalence given in (3.7) by the solutions of Eqs. (3.4) and (3.6). Conversely, given a $\mathfrak{g}$-valued function a satisfying (3.5) on P', a self-dual connection Γ results on M from the solutions to (3.4) and (3.6). The connection $\Gamma = \Gamma_{AA'}\,dx^{AA'}$ is in fact given in terms of a and G by

$$
\begin{aligned}
\Gamma(x) = (1/4\pi)\int_{S^2} &(dG\cdot G^{-1} + Ga_{,A}t^{B}{}_{A'}\bar{\lambda}_B \\
&\times t^{CC'}\bar{\lambda}_C\lambda_{C'}\,dx^{AA'}G^{-1})(x,[\lambda])\,d\mu_\lambda,
\end{aligned} \tag{3.8}
$$

where $a_{,A}$ denotes the derivative of a with respect to its first argument, $x^{AA'}\lambda_{A'}$. Equation (3.7) is a slight generalization of Eq. (2.2). Equation (3.7) can be verified by showing that upon contraction of $\Gamma_{AA'}$ with $\pi^{A'}$, Eq. (3.4) results. Thus from (3.7)

$$\pi^{A'}\Gamma_{AA'} = \int_{S^2}[(1/4\pi)(\pi^{A'}\nabla_{AA'}G)G^{-1} + (1/4\pi)Ga_{,A}\pi^{A'}t^{B}{}_{A'}\bar{\lambda}_B t^{CC'}\bar{\lambda}_C\lambda_{C'}G^{-1}](x,[\lambda])d\mu_\lambda. \tag{3.9}$$

Now

$$
\begin{aligned}
\frac{1}{4\pi}a_{,A}\pi^{A'}t^{B}{}_{A'}\bar{\lambda}_B t^{CC'}\bar{\lambda}_C\lambda_{C'}
&= \frac{-\pi^{A'}\nabla_{AA'}a\,t^{BB'}\bar{\lambda}_B\pi_{B'}\,t^{CC'}\bar{\lambda}_C\lambda_{C'}}{4\pi\pi^{D'}\lambda_{D'}} \\
&= -\pi^{A'}\nabla_{AA'}aL(\pi,\lambda),
\end{aligned}
$$

the last using the definition of $L(\pi,\lambda)$ from (2.4). Putting this into (3.8) results in

$$
\begin{aligned}
\pi^{A'}\Gamma_{AA'} = \pi^{A'}\int_{S^2} &\{(1/4\pi)(\nabla_{AA'}G)G^{-1}(x,[\lambda]) \\
&- G\nabla_{AA'}aG^{-1}(x,[\lambda])L(\pi,\lambda)\}\,d\mu_\lambda,
\end{aligned}
$$

where the right-hand side is

$$\int_{S^2}[(1/4\pi)dG\cdot G^{-1} - G\,da\,G^{-1}L(\pi,\lambda)](x,[\lambda])\,d\mu_\lambda \tag{3.10}$$

contracted with $\pi^{A'}$. But from Eq. (3.6) or, equivalently, (3.3)

$$G\,da\,G^{-1}\Delta\bar{\lambda} = Gd(G^{-1}\bar{\partial}G)G^{-1} = \bar{\partial}(dG\cdot G^{-1})$$

or

$$t^{AA'}\bar{\lambda}_A\lambda_{A'}G\,da\,G^{-1} = -t_A{}^{A'}\lambda_{A'}\frac{\partial}{\partial\bar{\lambda}_A}(dG\cdot G^{-1})$$

or

$$\eth(dG\cdot G^{-1}) = -G\,da\,G^{-1},$$

and thus (3.9) may be written as

$$\int_{S^2}[(1/4\pi)dG\cdot G^{-1} + \eth_\lambda(dG\cdot G^{-1})L(\pi,\lambda)](x,[\lambda])\,d\mu_\lambda,$$

which by (3.5) and the fact that $\eth$ and L are inverse operators is just $dG\cdot G^{-1}(x,[\lambda])$. Thus (3.8) may be written as

$$\pi^{A'}\Gamma_{AA'} = (\pi^{A'}\nabla_{AA'}G)G^{-1}(x,[\pi]),$$

which establishes that the connection given by (3.7) is the one given by a form Eqs. (3.4) and (3.6).

Given a self-dual connection Γ on M [modulo the solvability of Eqs. (3.4) and (3.6)] a function a arises of the variables $(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'})$, where a is defined up to the transformations in (3.7). The plane wave decomposition results if an a' can be found such that

$$a' = a'(x^{AA'}\bar{\pi}_A\pi_{A'}, \bar{\pi}_A, \pi_{A'}). \tag{3.11}$$

Given Γ and thus a, one derives to find an $H(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'})$ so that a' from (3.7) satisfies (3.10). Applying $\nabla_{AA'}$ to (3.7) results in

$$\bar{\pi}^B\nabla_{AA'}a' = \bar{\pi}^B\nabla_{AA'}(H^{-1}aH) + \nabla_{AA'}\left(H^{-1}\frac{\partial}{\partial\bar{\pi}_B}H\right),$$

and it is desired that $\nabla_{AA'}a \propto \bar{\pi}_A\pi_{A'}$. Already, because of the functional forms of H, a, and a', $\nabla_{AA'}$ of these is $\propto \pi_{A'}$. Thus

$$\nabla_{AA'}a' = a'_{,A}\pi_{A'},$$

and the requirement that $a' = a'(x^{AA'}\bar{\pi}_A\pi_{A'})$ is simply that

$$a'_{,A} \propto \pi_A \quad \text{or} \quad \bar{\pi}^A a'_{,A} = 0.$$

One can write this as

$$\bar{\pi}^A t^{BA'} \bar{\pi}_B \nabla_{AA'} a' \equiv Y_\pi a' = 0,$$

and thus the condition on H is that

$$Y_\pi \left(H^{-1} \frac{\partial}{\partial \bar{\pi}_A} H \right) + \bar{\pi}^A Y_\pi a = 0 \tag{3.12}$$

since a and H are also functions of $x^{AA'}$ through $x^{AA'}\pi_{A'}$ only. This equation is awkward since, while local solutions are easily obtained, solutions global in the fibers of $P' \to M$ are more difficult to establish and this is what is required. Y_π represents a complex vector field in $\mathbb{C}M$ that is nowhere tangent to M, and the requirement is that a' be constant along this transverse vector field to $M \to \mathbb{C}M$.

Instead of attempting to solve (3.11) globally in the fibers of $P' \to M$, the following considerations may be employed. If the connection is given with respect to a section of $B_M \to M$ so that Γ is asymptotically flat,[1] then a slight extension of the arguments given there [the extension from $a = a(x^{AA'}\bar{\pi}_A \pi_{A'}, \bar{\pi}_A, \pi_{A'})$ to $a = a(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'})$] produces a connection on $\mathscr{I}^+$ of the form $a'(u, \bar{\pi}_A, \pi_{A'}) \Delta \bar{\pi}$ and function $a'(x^{AA'}\bar{\pi}_A \pi_{A'}, \bar{\pi}_A, \pi_{A'})$ provides a solution to (3.11) and gives the same curvature form as a and thus determines H.

Given that the equations and construction are conformally invariant, the null initial-value problem on a null cone of M or on a portion of such can be transformed by a conformal transformation to an initial value problem on $\mathscr{I}^+$ or a position of $\mathscr{I}^+$. And again the connection on $\mathscr{I}^+$ assumes the form

$$a'(u, \bar{\pi}_A, \pi_{A'}) \Delta \pi.$$

In these settings only the radiation part of the field will appear. In fact, for asymptotically flat connections, writing

$$a(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'}) = b(x^{AA'}\bar{\pi}_A \pi_{A'}, x^{AA'} t^{B'}{}_A \pi_{A'} \pi_{B'}, \bar{\pi}_A, \pi_{A'}),$$

reasonable asymptotic conditions[1] would give for the radiation field

$$b(x^{AA'}\bar{\pi}_A \pi_{A'}, 0, \bar{\pi}_A, \pi_{A'}),$$

which would be used in place of a to construct the connection.

IV. INTERPRETATIONS OF THE CONSTRUCTIONS

In Secs. II and III constructions relating a self-dual connection on $B_M \to M$ and a 1-form on $B_{P'} \to P'$

$$A \equiv a(x^{AA'}\pi_{A'}, \bar{\pi}_A, \pi_{A'}) \Delta \pi$$

where carried out. The 1-form A satisfies

$$\pi^{A'} \nabla_{AA'} A = \pi^{A'} \nabla_{AA'} a \Delta \pi = 0$$

and

$$\bar{\partial} A = d\pi_B \wedge \frac{\partial}{\partial \bar{\pi}_B} A = 0.$$

The latter equation follows since a is of degree $(-2,0)$ in $(\bar{\pi}, \pi)$ and thus satisfies

$$\bar{\pi}_B \frac{\partial a}{\partial \bar{\pi}_B} = -\eth a,$$

and thus

$$\bar{\partial} A = \left(\frac{\partial a}{\partial \bar{\pi}_A} \bar{\pi}^B - a \epsilon^{AB} \right) d\bar{\pi}_A \wedge d\bar{\pi}_B,$$

which is a skew-symmetric expression in AB times $d\bar{\pi}_A \wedge d\bar{\pi}_B$, and so proportional to ϵ_{AB}, and upon contraction with ϵ_{AB} yields $+\bar{\pi}_A \partial a / \partial \bar{\pi}_A + 2a = 0$. Also

$$A \wedge A = a \cdot a \Delta \bar{\pi} \wedge \Delta \bar{\pi} = a \cdot a \bar{\pi}^A \bar{\pi}^B d\bar{\pi}_A \wedge d\bar{\pi}_B = 0.$$

Thus A satisfies

$$\bar{\delta} A + A \wedge A = 0,$$

where $\bar{\delta}$ is the operator defined by

$$\bar{\delta} = d\bar{\pi}_A \wedge \frac{\partial}{\partial \bar{\pi}_A} + \theta^A \wedge \pi^{A'} \nabla_{AA'}$$

and $\theta^A = \alpha_{A'} \, dx^{AA'} / \pi^{B'} \alpha_{B'}$, where $\alpha_{A'}$ is a fixed but arbitrary spinor. Because of the functional form of a (namely a is constant on the fibers of $P' \to \widetilde{PN}$ or $\bar{\pi}^A \pi^{A'} \nabla_{AA'} a = 0$), A is the pull back of a 1-form B on $\widetilde{PN}$ with values in. The manifold $\widetilde{PN}$ is a hypersurface of real dimension five in a complex manifold PT of complex dimension three and thus $\widetilde{PN}$ inherits a tangential complex structure. In particular, $\widetilde{PN}$ has a two-complex-dimensional tangent subspace at each point consisting of holomorphic tangent vectors of PT. Also, at each point of $\widetilde{PN}$ is defined a two-complex-dimensional tangent subspace consisting of antiholomorphic tangent vectors of PT. Under $P' \to \widetilde{PN}$ the vector fields $\partial / \partial \bar{\pi}_A, \pi^{A'} \nabla_{AA'}$ on P' are mapped to a basis for antiholomorphic vectors at the image points in PN. These are only essentially two complex vector fields on PN as $\bar{\pi}_A \partial / \partial \bar{\pi}_A$ is the homogenity operator and $\bar{\pi}^A \pi^{A'} \nabla_{AA'}$ maps to zero under $P' \to \widetilde{PN}$ since it is tangent to the fibers. Thus a function on P' of homogenity degree (0,0) in $(\bar{\pi}, \pi)$ is the pullback of the boundary values of a holomorphic function in PT or CR holomorphic function on $\widetilde{PN}$ if the applications of $\partial / \partial \bar{\pi}_A$ and $\pi^{A'} \nabla_{AA'}$ on it produce zero. Thus, if f is a function on P' [of degree (0,0) in $(\bar{\pi}, \pi)$] and if f satisfies $\bar{\pi}^A \pi^{A'} \nabla_{AA'} f = 0$, then

$$\bar{\delta} f = \frac{\partial f}{\partial \bar{\pi}_A} d\bar{\pi}_A + \pi^{A'} \nabla_{AA'} f \theta^A$$

is the pullback of a CR exact (0,1) form in $\widetilde{PN}$ and

$$\bar{\pi}_A \frac{\partial f}{\partial \bar{\pi}_A} = 0, \quad \pi^{A'} \nabla_{AA'} f \propto \bar{\pi}_A$$

and so

$$\bar{\delta} f = A \bar{\pi}^A d\bar{\pi}_A + B \bar{\pi}_A \theta^A,$$

where A is of degree $(-2,0)$ in $(\bar{\pi}, \pi)$ and B is of degree $(-1,1)$ in $(\bar{\pi}, \pi)$. Any CR (0,1) form pulled back from PN has this form and if it is a CR closed one form, it satisfies the integrability conditions for it to be CR exact. That is, if there exists a function f on P' such that

$$A \bar{\pi}^A = \frac{\partial f}{\partial \bar{\pi}_A}, \quad B \bar{\pi}_A = \pi^{A'} \nabla_{AA'} f$$

and the integrability conditions are

$$\pi^{A'}\nabla_{AA'}A = \epsilon_{CA}2B/2\bar{\pi}_C.$$

Equation (4.1) is the appropriate condition that the $\mathfrak{g}$-valued one-form A be $\bar{\delta}$ closed in the CR sense. Thus A is the pullback of a $\bar{\delta}$ closed (0,1) CR form on $\widetilde{PN}$ and represents a cycle in $H^1_{CR}(\widetilde{PN},\mathfrak{G}(0))$. This may be made more explicit by considering the equation

$$G^{-1}(\bar{\delta}G) = A$$

for a $\mathfrak{G}$-valued function G. The integrability conditions for this equation are obtained by applying $\bar{\delta}$ to

$$\bar{\delta}G = GA,$$

whence

$$\begin{aligned} 0 = \delta^2 G &= \bar{\delta}G \wedge A + G\bar{\delta}A \\ &= GA\wedge A + G\bar{\delta}A = G(\bar{\delta}A + A\wedge A), \end{aligned}$$

which are satisfied by virtue of (4.1). Local solutions exist in $\widetilde{PN}$ and there exists a covering V_i of $\widetilde{PN}$ such that on the pullback of V_i, U_i in P', the pulled-back functions G_i satisfy

$$G_i^{-1}\bar{\delta}G_i = A \quad \text{on } U_i.$$

On $U_i \cap U_j$ consider ϕ_{ji} defined by

$$\phi_{ji} = G_j \cdot G_i^{-1}.$$

On the domain of definition applying $\bar{\delta}$ to ϕ_{ji} gives

$$\begin{aligned} \bar{\delta}\phi_{ji} &= \bar{\delta}(G_j \cdot G_i^{-1}) = (\bar{\delta}G_j)G_i^{-1} + G_j\bar{\delta}G_i^{-1} \\ &= G_jAG_i^{-1} - G_jAG_i^{-1} = 0 \end{aligned}$$

and thus $\{\phi_{ji}\}$ are $\bar{\delta}$-closed $\mathfrak{G}$-valued functions defined on $\{U_i \cap U_j\}$. Thus $\{\phi_{ji}\}$ represent the transistion functions for a CR holomorphic $\mathfrak{G}$-bundle over $\widetilde{PN}$ and an element of

$$H^1_{CR}(\widetilde{PN},\mathfrak{G}(0)).$$

A cocycle equivalent to A is provided by A', where

$$A' = H^{-1}AH + H^{-1}\bar{\delta}H,$$

where H is the pullback of a $\mathfrak{G}$-valued function on PN to P', and this freedom is represented in (3.7).

Thus given a self-dual connection Γ on M, a CR (0,1) form on $\widetilde{PN}$ results with values in $\mathfrak{G}$, and this is used to construct a principal $\mathfrak{G}$-bundle over $\widetilde{PN}$ which is CR holomorphic.

The resulting bundle over $\widetilde{PN}$ is trivial on the images of the fibers of $P'\rightarrow M$ under $P'\rightarrow\widetilde{PN}$ which are 2-spheres.

The inverse problem of constructing a self-dual connection on M from a CR holomorphic principal $\mathfrak{G}$-bundle over $\widetilde{PN}$, which is trivial on the images of the fibers of $P'\rightarrow M$ under $P'\rightarrow\widetilde{PN}$ is solved as follows. Represent the bundle over $\widetilde{PN}$ by a $\bar{\partial}$-closed (0,1) CR form on $\widetilde{PN}$ with values in $\mathfrak{G}$. Pulling the bundle back to P' under $P'\rightarrow\widetilde{PN}$ and pulling back the form on $\widetilde{PN}$ gives a $\bar{\delta}$-closed form on P' with values in $\mathfrak{G}$. Pulling this back to a form on a fiber of $P'\rightarrow M$, since the bundle is trivial on this $\mathbb{C}P'$, the (0,1) form on the $\mathbb{C}P'$ is CR exact and thus allows generically the solution of

$$G^{-1}\bar{\partial}G = a\Delta\bar{\pi},$$

where $a\Delta\bar{\pi}$ is the pullback form on the fibers of $P'\rightarrow M$. Solving this for G results in a $\mathfrak{G}$-valued function on P' and

$$(\pi^{A'}\nabla_{AA'}G)G^{-1}$$

is global on the fibers and of degree (0,1) in $(\bar{\pi},\pi)$ and thus

$$(\pi^{A'}\nabla_{AA'}G)G^{-1} = \Gamma_{AA'}(x)\pi^{A'}.$$

the 1-form $\Gamma_{AA'}(x)\,dx^{AA'}$ is the resulting self-dual connection. In this construction alternately it is possible to work with transition functions for the bundle.[2]

If it is assumed that the bundle is defined in $\mathbb{C}M^+$, the forward tube of M in $\mathbb{C}M$, and is holomorphic along with the connection, then following Refs. 2 and 3, and an element of $H^1(PT^+,(0))$ is obtained and a holomorphic principal $\mathfrak{G}$-bundled results on PT^+, which is holomorphically trivial on the $\mathbb{C}P^1$'s which represent points of M^+ in PT^+.

[1]J. R. Porter, "Self-dual Yang-Mills Fields," J. Math. Phys. **24** (1983).

[2]R. S. Ward, Phys. Lett. A **61**, 81–82 (1977).

[3]D. E. Lerner, "The Inverse Twistor Function for Positive Frequency Fields," *Advances in Twistor Theory*, edited by L. P. Hughston and R. S Ward (Pitman, San Francisco, 1979).

Complex scaling of ac Stark Hamiltonians

James S. Howland
Department of Mathematics, University of Virginia, Charlottesville, Virginia 22903

(Received 22 April 1981; accepted for publication 7 May 1982)

For a two-body atom in a temporally periodic, spatially uniform field, it is shown that in an appropriate gauge the essential spectrum of the Floquet Hamiltonian rotates about a certain set of thresholds when subjected to a complex scaling transformation.

PACS numbers: 11.15 − q, 02.30.Hq, 32.60. + i

1. INTRODUCTION

The Hamiltonian for a two-body atom in a uniform electric field $E(t)$ may be written

$$H_1(t) = p^2 + V(x) + E(t) \cdot x, \tag{1.1}$$

where $p = -i\nabla$. The ac Stark Hamiltonian studied by Cerjan, Reinhardt, and Holt[1] (CRH) corresponds to the field

$$E(t) = \epsilon x_1 \cos(\omega t). \tag{1.2}$$

Another case of interest[2] is that of a small oscillating field in a fixed Stark field:

$$E(t) = Ex_1 + \epsilon \cdot x \cos(\omega t), \tag{1.3}$$

where $|\epsilon| \ll E$. We shall be concerned only with the case of $E(t)$ periodic in time, with some period T.

For Hamiltonians periodic in time, the natural object for study is the Floquet Hamiltonian.[1,3] For our purposes, this may be described as the operator

$$K_1 = -i\frac{\partial}{\partial t} + H_1(t) \tag{1.4}$$

on the space $L_2(\mathbb{R} \times [0, T])$, with the periodic boundary condition

$$u(0, x) = u(T, x).$$

One reason for this is that a point eigenvalue λ of K_1 corresponds to a solution $\psi(t)$ of the Schrödinger equation

$$i\frac{\partial \psi}{\partial t}(t) = H_1(t)\psi(t),$$

satisfying

$$\psi(T) = e^{-i\lambda T}\psi(0).$$

For example, if $H_1(t) = H$ were constant in time and $H\phi = \lambda\phi$, then $\psi(t) = e^{-i\lambda t}\phi$ would be such a solution. Thus, point eigenvalues of the unperturbed operator

$$H_0 = p^2 + V(x)$$

correspond to point eigenvalues of the unperturbed Floquet Hamiltonian

$$K_0 = -i\frac{\partial}{\partial t} + p^2 + V(x).$$

Note,[3] however, that the spectrum of a Floquet operator is periodic: If $\lambda \in \sigma_p(K_1)$, then so is $\lambda + 2\pi n T^{-1}$, for any integer n.

In CRH the complex scaling transformation[1,4]

$$x \to e^{i\theta}x, \quad p \to e^{-i\theta}p \tag{1.5}$$

is formally applied to the Floquet Hamiltonian with the field (1.2). [Their form may be obtained from (1.4) by expanding in a Fourier series in t.] A numerical calculation of resonances is then based on the scaled Hamiltonian, in analogy with previous calculations on the ordinary Stark effect[5] and on atomic systems.[6]

The object of this paper is to study the resonance structure of the Floquet Hamiltonian for (1.2) under complex scaling, with a view to providing a justification for the CRH calculations. Apparently, the behavior of (1.1) with the field (1.2) is very bad under complex scaling. We conjecture that the spectrum of $K_1(i\theta)$ is the entire complex plane for $\theta \neq 0$. However, if a more suitable gauge is chosen, one finds that the essential spectrum simply rotates about a certain set of thresholds in the now familiar manner.[4] We shall give a proof of this fact.

The Hamiltonian (1.3), which is better treated by a different method, will be discussed elsewhere.

For the ordinary (dc) Stark effect, complex scaling has been discussed by Herbst and Simon.[7]

2. A GAUGE TRANSFORMATION

If the electromagnetic field $E = E(t)$ and $B \equiv 0$ is described as

$$E = -\nabla\phi - \frac{\partial A}{\partial t}, \quad B = \nabla \times A,$$

where $A \equiv 0$ and

$$\phi(x, t) = E(t) \cdot x,$$

one obtains the Hamiltonian (1.1). However, if

$$\Lambda(x, t) = a(t) \cdot x,$$

where $a'(t) = E(t)$, then the gauge change

$$\tilde{\phi} = \phi - \frac{\partial \Lambda}{\partial t}, \quad \tilde{A} = A + \nabla\Lambda$$

leads to $\tilde{\phi} \equiv 0$ and

$$\tilde{A}(x, t) = a(t),$$

and hence to the Hamiltonian

$$H(t) = (p - a(t))^2 + V(x). \tag{2.1}$$

We shall write

$$K = -i\frac{\partial}{\partial t} + H(t)$$

for the Floquet operator corresponding to (2.1)

 0022-2488/83/051240-05$02.50

3. A SIMPLE DIFFERENTIAL EQUATION

In order to calculate the unperturbed resolvent, it is necessary to have the solution $u(p,t)$ of the equation

$$-i\frac{\partial u}{\partial t}+B(p,t)u=f(p,t) \tag{3.1}$$

satisfying $u(p,0)=u(p,2\pi)$, where $B(p,t)$ and $f(p,t)$ are 2π periodic in t. An elementary integration gives the result:

$$u(p,t)=i\int_0^t e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma + i(e^{i\chi(p)}-1)^{-1}\int_0^{2\pi} e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma, \tag{3.2}$$

where

$$\chi(p)=\int_0^{2\pi}B(p,s)ds$$

and

$$\phi(t,\sigma;p)=\int_\sigma^t B(p,s)ds.$$

Transform the second integral by putting $\sigma=\tau+2\pi$ and use that

$$\phi(t,\tau+2\pi;p)=\phi(t,\tau;p)-\chi(p).$$

Replacing the dummy variable τ by σ in the result gives finally

$$u(p,t)=i\int_0^t e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma + i\Delta^{-1}(p)\int_{-2\pi}^0 e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma, \tag{3.3}$$

where

$$\Delta(p)=1-e^{-i\chi(p)}.$$

For $0\leqslant t\leqslant 2\pi$, one then always has $t\geqslant\sigma$ in (3.3).

4. PARTICLE IN AN ac FIELD

We shall assume for the rest of this paper that $a(t)$ is a real vector-valued measurable function of period $T=2\pi$. Let

$$b^2=\frac{1}{2\pi}\int_0^{2\pi}|a(t)|^2dt \tag{4.1}$$

and assume that

$$\int_0^{2\pi}a(t)dt=0. \tag{4.2}$$

Let $\mathscr{H}=L_2(\mathbb{R}^3)$ and let $\mathscr{K}$ be the space of measurable $\mathscr{H}$-valued functions $u(t)$ with period 2π, and finite norm

$$||u||^2_{\mathscr{K}}=\int_0^{2\pi}|u(t)|^2dt. \tag{4.3}$$

The space $\mathscr{K}$ is equivalent to $L_2(\mathbb{R}^3\times[0,2\pi])$, but it is convenient to extend the functions $u(p,t)$ in this space to be periodic in t on the whole line.

We shall consider the Hamiltonian

$$K_0=-i\frac{\partial}{\partial t}+(p-a(t))^2 \tag{4.4}$$

on $\mathscr{K}$. The equation

$$(K_0-\zeta)u=f$$

reduces to (3.1) with

$$B(p,t)=(p-a(t))^2-\zeta.$$

The resolvent $R_0(\zeta)=(K_0-\zeta)^{-1}$ of K_0 is therefore given by

$$R_0(\zeta)f(p,t)=i\int_0^t e^{i\zeta(t-\sigma)}e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma + i[1-e^{-2\pi i(\chi(p)-\zeta)}]^{-1}\times\int_{-2\pi}^0 e^{i\zeta(t-\sigma)}e^{-i\phi(t,\sigma;p)}f(p,\sigma)d\sigma \tag{4.5}$$

for $\operatorname{Im}\zeta\neq 0$, where

$$\chi(p)=p^2+b^2$$

and

$$\phi(t,\sigma,p)=\int_\sigma^t(p-a(s))^2ds.$$

We have written (4.5) in the momentum representation and will remain in that representation until the next section.

Let $\mathscr{D}$ be the space of C^∞ functions of (p,t) which are 2π periodic in t and of compact support in p.

Theorem 1: The operator K_0 is essentially self-adjoint on $\mathscr{D}$, has spectrum $\mathbb{R}$, and resolvent $R_0(\zeta)$ given by (4.5).

Proof: The two integral operators of (4.5) are bounded for every ζ, since $e^{-i\phi}$ is bounded. Hence, $R_0(\zeta)$ is a bounded operator iff $e^{2\pi i(\chi(p)-\zeta)}$ is bounded away from unity, which is true iff ζ is not real. Thus, it suffices to prove that $R_0(\zeta)$ is actually the resolvent of K_0.

Assume for the moment that $a(t)$ is C^∞. Then for non-real ζ, the vector $u=R_0(\zeta)f$ is in $\mathscr{D}$ whenever f is in $\mathscr{D}$ and satisfies $(K_0-\zeta)u=f$. Since f is arbitrary and $\mathscr{D}$ dense, $(K_0\pm i)\mathscr{D}$ is dense. Hence, K_0 is essentially self-adjoint on $\mathscr{D}$ and its resolvent is $R_0(\zeta)$.

Now choose a sequence $a_n(t)$ of C^∞ vectors converging to $a(t)$ in L_2. Then clearly $R_0^{(n)}(\zeta)\to R_0(\zeta)$ strongly. Fix f in $\mathscr{D}$, and let $u_n=R_0^{(n)}(\zeta)f$. Then $u_n\in\mathscr{D}$, and $u_n\to u=R_0(\zeta)f$ strongly. Hence, $K_0^{(n)}u_n=\zeta u_n+f\to\zeta u+f$, so by closure $u\in D(K_0)$ and $(K_0-\zeta)u=f$. Since f is arbitrary, $(K_0\pm i)\mathscr{D}$ is dense, so K_0 is again self-adjoint. Since $u=R_0(\zeta)f$, it follows that $R_0(\zeta)=(K_0-\zeta)^{-1}$.

For real α, define the unitary scaling transformation

$$U(\alpha)f(p,t)=e^{-3\alpha/2}f(e^{-\alpha}p,t).$$

The resolvent $R_0(\zeta,\alpha)$ of $K_0(\alpha)=U(\alpha)K_0U(-\alpha)$ is then given by

$$R_0(\zeta,\alpha)f(p,t)=i\int_0^t e^{i\zeta(t-\sigma)}e^{-i\phi(t,\sigma,p;\sigma)}f(p,\sigma)d\sigma + i[1-e^{2\pi i(\zeta-\chi(p;\alpha))}]^{-1}\times\int_{-2\pi}^0 e^{i\zeta(t-\sigma)}e^{-i\phi(t,\sigma,p;\sigma)}f(p,\sigma)d\sigma, \tag{4.6}$$

where

$$\phi(t,\sigma,p;\alpha)=\int_\sigma^t(e^{-\alpha}p-a(s))^2ds$$

and

$$\chi(p;\alpha)=e^{-2\alpha}p^2+b^2.$$

Now let $\alpha = \beta + i\theta, 0 < \theta < \pi/2$, and define $R_0(\zeta, \alpha)$ by (4.6). Formally, $R_0(\zeta, \alpha)$ is the resolvent of the differential operator

$$K_0(\alpha) = -i\frac{\partial}{\partial t} + (e^{-\alpha}p - a(t))^2$$
$$= -i\frac{\partial}{\partial t} + e^{-2\alpha}p^2 - 2e^{-\alpha}a(t)\cdot p + a(t)^2.$$

Let $K_0(\alpha)$ denote the closure of the restriction of this operator to $\mathscr{D}$. Let

$$\Sigma(\alpha) = Z + b^2 + e^{-2i\theta}\mathbb{R}^+,$$

where Z is th set of integers and $\mathbb{R}^+ = (0, \infty)$. The set $\Sigma(\alpha)$ consists of rays with end points at $n + b^2, n = 0, \pm 1, \ldots,$ and at an angle -2θ with the real axis.

Theorem 2: (a) The operator $R_0(\zeta, \alpha)$ is holomorphic on

$$\Omega = \{(\zeta, \alpha): \zeta \notin \Sigma(\alpha), 0 < \operatorname{Im} \alpha < \pi/2\}.$$

For fixed ζ, $\operatorname{Im} \zeta > 0$, $R_0(\zeta, \alpha) \to R_0(\zeta, 0)$ in operator norm as $\alpha \to 0$ in Ω.

(b) For fixed α, $R_0(\zeta, \alpha)$ is the resolvent of $K_0(\alpha)$, and $\sigma(K_0(\alpha)) = \Sigma(\alpha)$.

(c) $K_0(\beta + i\theta)$ is unitarily equivalent to $K_0(i\theta)$.

Proof: Let $\alpha = \beta + i\theta, 0 < \theta < \pi/2$. We claim that

$$\operatorname{Im} \phi(t, \sigma, p; \alpha) \leqslant 2\pi b^2 \tan \theta. \tag{4.7}$$

In fact, one computes that

$$\operatorname{Im} \phi(\alpha) = -\sin 2\theta e^{-2\beta}(t - \sigma)p^2 + 2\sin \theta e^{-\beta}p \cdot \int_\sigma^t a(s)ds, \tag{4.8}$$

where $-2\pi < \sigma < t < 2\pi$. If we note the inequality

$$-ap^2 + 2c\cdot p \leqslant c^2/a,$$

which is valid for vectors p, c in $\mathbb{R}^3$ and $a > 0$, we obtain

$$\operatorname{Im} \phi(\alpha) \leqslant [4 \sin 2\theta e^{-2\beta}(t - \sigma)]^{-1} \times 4 \sin^2\theta e^{-2\beta}\left|\int_\sigma^t a(s)ds\right|^2.$$

However, by Schwartz,

$$\left|\int_\sigma^t a(s)ds\right|^2 \leqslant (t - \sigma)\int_\sigma^t |a(s)|^2 ds \leqslant 4\pi b^2(t - \sigma).$$

Combining these estimates leads to (4.7).

The bound (4.7) implies that the integral operators in (4.5) are bounded, holomorphic on Ω, and norm continuous up to the axis $\theta = 0$. The factor $\Delta^{-1}(p, \alpha)$ is clearly bounded and holomorphic for $\zeta \in \Sigma(\alpha)$, so (a) is proved. The proof of (b) follows the proof of Theorem 1 closely, while for (c) we have

$$K_0(\beta + i\theta) = U(\beta)K_0(i\theta)U(-\beta).$$

5. SPECTRUM OF AN ATOM IN AN ac FIELD

We consider the Floquet operator

$$K = -i\frac{\partial}{\partial t} + (p - a(t))^2 + V(x).$$

We say that a function $V(x)$ in $L_p(\mathbb{R}^3), 1 \leqslant p \leqslant \infty$, is L_p *— dilation analytic* in the strip $0 \leqslant \operatorname{Im} \alpha < \theta_0$ if $V(e^\beta x)$ is the strong boundary value for $\operatorname{Im} \alpha = 0$ of an $L_p(\mathbb{R}^3)$ — valued analytic function in $0 < \operatorname{Im} \alpha < \theta_0$.

Theorem 3: Let $V(x)$ be real-valued and L_2-dilation analytic in $0 \leqslant \operatorname{Im} \alpha \leqslant \theta_0 < \pi/2$, and assume that $a(t)$ is bounded. Then

(a) V_α is relatively $K_0(\alpha)$ compact. Hence, $K(\alpha) = K_0(\alpha) + V_\alpha$ is closed with domain $D(K_0(\alpha))$, is self-adjoint for real α, and has essential spectrum

$$\sigma_{\text{ess}}(K(\alpha)) = \Sigma(\theta).$$

(b) $V_\alpha R_0(\zeta, \alpha)$ is holomorphic in $\Omega(\theta_0) = \{(\zeta, \alpha) \notin \Omega: 0 < \operatorname{Im} \alpha < \theta_0\}$, and is continuous up to $\operatorname{Im} \alpha = 0$ for $\operatorname{Im} \zeta > 0$ in operator norm. Hence, $R(\zeta, \alpha)$ is meromorphic on $\Omega(\theta_0)$.

(c) If ϕ is a dilation-entire vector, then the matrix element $\langle R(\zeta)\phi, \phi\rangle$ can be continued meromorphically from $\operatorname{Im} \zeta > 0$ to $\mathbb{C} \setminus \Sigma(\theta)$ for any θ, $0 < \theta < \theta_0$. The poles of the continuation occur only at points of the discrete spectrum of $K(\alpha)$.

(d) The discrete eigenvalues of $K(\alpha)$ are independent of α.

Proof: Let $V(x)$ be a complex-valued L_2 function. We claim that $VR_0(\zeta, \alpha)$ is compact and that

$$\|VR_0(\zeta, \alpha)\| \leqslant C(\alpha)\|V\|_2, \tag{5.1}$$

where $C(\alpha)$ is bounded on compact subsets of $\mathbb{C} \setminus \Sigma(\theta_0)$. For this, we need to estimate

$$Ve^{-i\phi} = Ve^{\operatorname{Im}\phi}e^{-i\operatorname{Re}\phi}.$$

Drop the unitary factor and use (4.8) to obtain

$$V(x)E^{-ap^2 + 2c\cdot p} = V(x)e^{-a(p - c/a)^2}e^{c^2/a}$$
$$= V(x)e^{ix\cdot c/a}e^{-ap^2}e^{-ix\cdot c/a}e^{c^2/a}.$$

Drop the unitaries again to obtain

$$V(x)e^{-ap^2}e^{c^2/a}.$$

But e^{-ap^2}, in configuration space, is convolution by $(2\pi a)^{-3/2}e^{-x^2/4a}$. Therefore, $V(x)e^{-ap^2}$ has finite Hilbert–Schmidt norm equal to

$$(2\pi a)^{-3/2}\|e^{-x^2/4a}\|_2\|V\|_2 = (2\pi a)^{-3/4}\|V\|_2. \tag{5.2}$$

Using (4.7), we find that

$$\|V(x)e^{-i\phi(t, \sigma, p; \alpha)}\|_{HS} \leqslant C(\alpha)(t - \sigma)^{-3/4}\|V\|_2,$$

where

$$C(\alpha) = (2\pi)^{-3/4}e^{3\beta/2}(\csc\theta)^{3/4}e^{2\pi b^2}\tan\theta.$$

For $\delta \geqslant 0$, define

$$T_\delta = \int_0^{\min(t - \delta, 0)} V(x)e^{-i\phi(t, \sigma, p; \alpha)}f(\sigma)d\sigma.$$

If $\delta > 0$, the Hilbert–Schmidt norm of the integral $Ve^{-i\phi}$ is bounded uniformly, so T_δ is Hilbert–Schmidt. For $\delta = 0$, and $0 \leqslant t \leqslant 2\pi$,

$$|T_0 f(t)| \leqslant C(\alpha)\|V\|_2\int_0^t (t - \sigma)^{-3/4}|f(\sigma)|d\sigma.$$

A convolution estimate shows that

$$\|T_0\| \leqslant C(\alpha)\|V\|_2\int_0^{2\pi} s^{-3/4}ds.$$

Similarly,

$$||T_\delta - T_0|| \leqslant C(\alpha)||V||_2 \int_0^\delta s^{-3/4} ds.$$

Thus $T_\delta \rightarrow T_0$ in norm and T_0 is compact. This takes care of the first term of (4.6). The second term is handled similarly, by writing it as

$$\int_{-2\pi}^{0} V(x) e^{-i\phi(t,\sigma,p,\alpha)} [1 - \chi(p,\alpha)]^{-1} f(\sigma) d\sigma.$$

In particular, $V_\alpha R_0(\zeta, \alpha)$ is compact, and (a) follows for $\theta \neq 0$. The analyticity in part (b) also follows from the norm estimate (5.1). Norm continuity is obtained as follows. Let

$$T = -i\frac{\partial}{\partial t} + p^2$$

so that, for $\alpha = 0$,

$$K(0) = T - 2a(t)p + a^2(t) + V(x).$$

Using the operator inequality $|2\,ap| \leqslant \epsilon p^2 + \epsilon^{-1} a^2$, one finds that the last three terms are T-bounded with arbitrarily small T-bound, so that $D(K) = D(K_0) = D(T)$. Similarly, one finds that even for Im $\alpha \neq 0$, $D(K(\alpha)) = D(K_0(\alpha)) = D(T)$. Moreover, $V_\alpha : D(T) \rightarrow L^2$ and $K_0(\alpha) : L^2 \rightarrow D(T)$ are analytic and norm continuous, and it follows that $V_\alpha R_0(\zeta, \alpha)$ is also. Continuation now gives (a) for real α.

The remaining assertions follow by standard dilation-analyticity arguments.[4]

The L_2 assumption in Theorem 3 is much too strong.

Theorem 4: The conclusion of Theorem 3 hold if $V(x)$ is L_p-dilation analytic in $0 \leqslant \text{Im}\,\alpha \leqslant \theta_0 < \pi/2$ for some $p \geqslant 2$. The case $p = \infty$ is permissible if in addition

$$\lim_{|x| \rightarrow \infty} V(e^\alpha x) = 0. \tag{5.3}$$

Proof: We need (5.1) with the L_p-norm rather than the L_2. This can be established by interpolation. For fixed f, the map

$$V \rightarrow V e^{-ap^2} f$$

is bounded from $L_\infty(\mathbb{R}^3)$ onto $\mathscr{H}$, with norm no more than $||f||$, and by (5.2), from $L_2(\mathbb{R}^3)$ into $\mathscr{H}$ with norm no more than $(2\,\pi a)^{-3/4}||f||$. By interpolation,

$$||V e^{-ap^2}|| \leqslant (2\,\pi a)^{-3/2p} ||V||p$$

for $2 \leqslant p \leqslant \infty$, in analogy with (5.2). The bound on $VR_0(\zeta, \alpha)$ is now obtained exactly as before. Compactness is proved by approximating V in L_p-norm by a sequence V_n in L_2. The condition (5.3) permits this for $p = \infty$.

The remainder of the proof is exactly the same.

Remarks:

(1) The same conclusions hold, of course, if $V(x)$ is the *sum* of two L_p — dilation analytic functions, for different p's.

(2) The same proof leads to a generalization of Theorem 4 to $L_p(\mathbb{R}^\nu)$ for $\nu \geqslant 4$. More singular local behavior of $V(x)$ and $a(t)$ can be accommodated by using the factorization method.

6. ANALYTICITY PROPERTIES OF K_1

If one insists on dealing with the original gauge, and the operator

$$K_1 = -i\frac{\partial}{\partial t} + p^2 + a'(t)\cdot x + V(x),$$

there are several ways to proceed. For one thing the gauge change of Sec. 2 can be described as a unitary transformation on $\mathscr{H}$. In configuration space, define the unitary operator

$$Gu(x,t) = e^{ia(t)\cdot x} u(x,t).$$

In momentum space, G acts as a translation

$$G\hat{u}(p,t) = \hat{u}(p + a(t), t).$$

One can easily compute that

$$G\left[p^2 + a'(t)\cdot x - i\frac{\partial}{\partial t}\right]G^* = (p - a(t))^2 - i\frac{\partial}{\partial t}$$

or, what is the same,

$$GK_1G^* = K. \tag{6.1}$$

For real α,

$$G^*K(\alpha)G = W(\alpha)K_1W(-\alpha),$$

where $W(\alpha) = G^*U(\alpha)G$ is a unitary group. Clearly, the spectrum of K_1 behaves regularly for complex α under the new group $W(\alpha)$, rather than the dilation group $U(\alpha)$. Thus, instead of a gauge change, we may think of analyticity with respect to a different group.

The nature of the group $W(\alpha)$ becomes clear if new coordinates

$$\xi = p - a(t), \quad \tau = t \tag{6.2}$$

are introduced in momentum space. In this representation, $W(\alpha)$ is the scale transformation

$$W(\alpha)u(\xi,\tau) = e^{-3\alpha/2} u(e^{-\alpha}\xi, \tau)$$

so that we may think of scaling ξ instead of p. (In classical terms, ξ is essentially the *velocity*.)

For $v(x) = 0$, the resolvent of K_1 can be computed explicitly. If one sets $x = -i(\partial/\partial p)$, one obtains for $(K_1 - \zeta)u = f$, a partial differential equation of first order. The substitution (6.2) leads to an equation of the type (2.1). The resulting formula appears to be very singular under the scaling $p \rightarrow e^{-i\theta} p$, which suggests strongly that the spectrum of $K_1(i\theta)$ is the entire complex plane, although the author has no rigorous proof of this.

If this is true, it need not mean that the CRH calculation is invalid. That calculation is based on matrix elements of K_1 in a scaled basis $\phi_j(\alpha)(j = 1, \ldots, N)$. But, by (6.1),

$$\langle K_1\phi_j(\alpha), \phi_i(\alpha)\rangle = \langle KG\phi_j(\alpha), G\phi_i(\alpha)\rangle$$

so that one also has a matrix of K in a different basis. The choice of basis is clearly of first importance here. However, the author does not feel sufficiently knowledgeable to speculate further.

ACKNOWLEDGMENTS

The author wishes to thank Bill Reinhardt for useful conversations. This work was partly supported by the National Science Foundation, Grant MCS-79-02490.

[1]C. Cerjan, W. Reinhardt, and C. Holt, Phys. Rev. Lett. **39**, 1195 (1977).

[2]R. R. Freeman, N. P. Economou, G. C. Bjorklund, and K. T. Lu., Phys. Rev. Lett. **20**, 1463 (1978).

[3]Ya. B. Zel'dovich, Usp. Fiz. Nauk **110**, 139 (1973) [Sov. Phys. Usp. **16**, 427 (1973)]; J. S. Howland, Indiana J. Math. **28**, 471 (1979).
[4]J. Aguilar and J. M. Combes, Commun. Math. Phys. **22**, 269 (1971); E. Balslev and J. M. Combes, Commun. Math. Phys. **22**, 280 (1971); B. Simon, Int. J. Quantum Chem. **14**, 343 (1978).
[5]W. P. Reinhardt, Int. J. Quantum Chem. Symp. No. 10, 359 (1976); C. Cerjan, W. P. Reinhardt, and J. E. Avron, J. Phys. B **11**, L201 (1978).
[6]The papers cited in Ref. 5 give references to this literature.
[7]I. W. Herbst, Commun. Math. Phys. **64**, 279 (1979); I. W. Herbst and B. Simon, Phys. Rev. Lett. **41**, 67 (1978); Commun. Math. Phys. **80**, 181 (1981).

On a class of synchronized observers attached to the Lorentzian structures

Bartolomé Coll and Carlos Moreno
Chaire de Physique Mathématique, Collège de France, Paris, France

(Received 22 December 1981; accepted for publication 16 April 1982)

The first-order evolution form for the Klein–Gordon equation is characterized by an operator $T(t)$ which, for inertial observers in the Minkowski space–time, is skew-self-adjoint with respect to the energy product. This fact is essential for a rigorous treatment of the equation. We prove here that, in arbitrary Lorentzian structures, there always exists a class of "synchronized observers" (equivalence class of physically admissible local charts), here called *adjoint systems*, for which this property of $T(t)$ remains true. They are completely determined by the Lorentzian structure, and, in this sense, they appear as a suitable generalization of the Killing vector fields. We obtain the definition equations for such observers and state some of their properties. A particular class of them, here called *simple adjoint systems*, has already been introduced by one of us (C.M.) for the study of the Klein–Gordon equation in arbitrary space–times, according with Lichnerowicz's quantization program.

PACS numbers: 11.10.Qr, 04.90. + e

I. INTRODUCTION

With respect to a system of synchronized observers-(timelike congruence of curves endowed with suitable spatial sections), the Klein–Gordon equation on a Lorentzian manifold, $\Delta_{n+1}u - \kappa u = 0$, may be written as a homogeneous first-order evolution equation $dX(t)/dt = T(t)\cdot X(t)$, where $X(t)$ is a curve in the space of Cauchy data and $T(t)$ is a linear operator on this space.

The purpose of this paper is to find *all* the local systems of synchronized observers for which the operator $iT(t)$ is formally self-adjoint with respect to a general energy product. It appears that these systems (which we call "adjoint systems") are completely determined by the Lorentzian structure (i.e., they do not depend on the mass in the Klein–Gordon equation nor on the general volume element used in the evaluation of the energy product).

In Minkowski space–time and with respect to inertial observers, the property of self-adjointness of the operator iT is one of the basic facts which allows us to define the positive and negative part solutions of a real solution of the Klein–Gordon equation.

This property of T has been used by one of us (C.M.) to develop, according to the Lichnerowicz[1] and Segal[2] theories of quantization, a rigorous study of the Klein–Gordon equation in any Lorentzian manifold under suitable global conditions[3] (see also Refs. 4–6 for stationary space–times).

Assuming certain (technical) global conditions, the study of the Klein–Gordon equation may be carried out in the same rigorous way as in Ref. 3 for all our adjoint systems of synchronized observers.

One of the motivations of this paper is to get better knowledge of these observers in view of investigating subsequently the conjecture that, in all Lorentzian structures, there would exist a class of these adjoint systems such that, *relatively to each one of them*, the "natural selection conditions" proposed in Ref. 3 would select a unique decomposition (in the sense of the paragraph 6.c in that reference) into positive- and negative-part solutions of every real solution of the Klein–Gordon equation with positive mass.

By the last time, an alternative rigorous approach to that problem of decomposition into positive and negative parts has been also considered by Paneitz and Segal.[7]

In Sec. II we define what we shall call *systems of synchronized observers*, and we introduce some related notions of interest.

Section III is devoted to the study of the Laplacian operator on functions for a system of synchronized observers (Theorem 1) and to the evaluation, with respect to arbitrary spatial volume elements η_μ, of the adjoints of the two operators M and N which characterize this formulation (Propositions 1 and 2).

In Sec. IV we obtain, relative to the volume η_μ, the total energy of a solution of the Klein–Gordon equation, with explicit dependence on the system of synchronized observers (Proposition 3). We evaluate, with respect to the energy product, the adjoint of the operator T characterizing the first-order evolution form of the Klein–Gordon equation (Proposition 4), and we find the necessary and sufficient conditions for T to be skew-self-adjoint (Proposition 5).

The adjoint systems of synchronized observers are analyzed in Sec. V. We give their definition equations (Theorem 2), determine their time parameter (Proposition 6), and obtain necessary and sufficient conditions for an adjoint system to be orthogonal (Proposition 7). Then, we conclude with the existence of adjoint systems on arbitrary Lorentzian structures (Theorem 3). Two simple results about certain volume elements admitted by adjoint systems are given in Propositions 7 and 8.

In Sec. VI, we seek systems of observers other than the Killing ones, which, like them, generate adjoint systems of synchronized observers for all admissible synchronizations. Their existence implies restrictions on the Lorentzian structures (Proposition 9 and Theorem 4). In fact, we are able to prove that such systems exist only on Lorentzian manifolds conformal to the general Robertson–Walker space–times (Theorem 5).

Finally, in Sec. VII, we indicate how some of the above results may be extended for the study of general second-order hyperbolic linear equations.

 0022-2488/83/051245-11$02.50

Some of the results contained in this paper were announced, without proof, in Ref. 8.

II. SYNCHRONIZED OBSERVERS

(a) Let (V_{n+1},g) be a Lorentzian manifold. By definition, a *system of observers* is a field ξ^* [9] of timelike vectors,

$$g(\xi^*,\xi^*)\equiv(\xi,\xi)>0\,. \tag{1}$$

The integral curves of ξ^* (i.e., orbits of the one-parameter local group of transformations defined by ξ^*) represent the *world lines* of the observers of the system, and the standard parameter along every curve (the local group parameter) defines the *time t* associated with such observers.

For some physical interpretations, one can attach, to every system of observers ξ^*, a system of *unitary* observers u^*, $u^*=|\xi|^{-1}\xi^*$ with $|\xi|^2\equiv(\xi,\xi)$, for which the associated time is the *proper* time. But, in many cases, and, in particular, in ours, the introduction of such unitary observers is not essential.

A *system of synchronized observers* is a system of observers for which the locus of a point $t=$ const is given and defines a one-parameter family of *spatial* hypersurfaces. It is easy to see that if $\phi(x)=$ const is the local equation of the family, then the function ϕ must verify[10]

$$(d\phi,d\phi)>0\,,\quad \mathscr{L}(\xi^*)\phi=1\,, \tag{2}$$

where d and $\mathscr{L}(\xi^*)$ are respectively the *exterior* and *Lie derivatives*. Thus, *a system of synchronized observers is given by a pair* $\{\xi^*;\phi\}$, verifying (1) and (2).

The function ϕ is said to be *a synchronization* for the system of observers ξ^*, and every hypersurface

$$\phi_c\equiv(\phi(x)=\text{const})$$

is called an *instant* of the system of synchronized observers $\{\xi^*;\phi\}$.

(b) Every system of synchronized observers $\{\xi^*;\phi\}$ is locally characterized by the class $\mathscr{C}(\xi^*;\phi)$ of local charts adapted to ξ^* and to ϕ. A chart (U,Ψ) is said to be *adapted* to $\{\xi^*;\phi\}$ if, for every integral curve γ of ξ^* and for every hypersurface ϕ_c, it is such that

$$\Psi(\phi_c\cap U)=\mathbb{R}^n\times\{\tau\}\,,\quad \Psi(\gamma\cap U)=\{\Sigma\}\times\mathbb{R}\,,$$

where τ and Σ are respectively fixed points of $\mathbb{R}$ and $\mathbb{R}^n$. Conversely, every "physically admissible"[11] local chart (U,Ψ) defines locally a system of synchronized observers $\{\xi^*;\phi\}$: If x^0 is the timelike coordinate function induced by (U,Ψ), then $\{\xi^*;\phi\}$ is given by $\xi^*\equiv\partial/\partial x^0$ and $\phi(x)\equiv x^0$. The pseudogroup of transformations of $\mathscr{C}(\xi^*;\phi)$ is called *the inner pseudogroup* of the system of synchronized observers $\{\xi^*;\phi\}$.

The systems of synchronized observers are the basic geometric elements of the evolution formalism of general relativity (sometimes called ADM or 3 + 1 formalism[12]. But ϕ appears there implicitly (because of the explicit use of adapted coordinates), and ξ^* is characterized by its normal and tangential components with respect to the hypersurfaces ϕ_c (called respectively the *lapse* function and the *shift* vector).

(c) From Eq. (2), it is clear that if $\{\xi^*;\phi\}$ and $\{\zeta^*;\phi\}$ are two systems of synchronized observers with the same synchronization ϕ, then $\xi^*-\zeta^*$ belongs to the Lie algebra $L(\phi)$ of the vector fields tangent to the hypersurfaces ϕ_c. Also, if ϕ and ψ are two synchronizations for a system of observers ξ^*, then $\phi-\psi$ belongs to the class $\mathscr{F}(\xi^*)$ of invariant functions with respect to ξ^*, $\mathscr{L}(\xi^*)(\phi-\psi)=0$.

Given two functions ϕ and ψ with timelike gradients, it is easy to see that there always exists a system of observers ξ^* for which ϕ and ψ are synchronizations. The converse is not true, but it can be shown that if two systems of observers ξ^* and ζ^* are such that $[\xi^*,\zeta^*]=\lambda(\xi^*-\zeta^*)$, with λ arbitrary, then they admit a mutual synchronization ϕ.

Let $\{\xi^*;\phi\}$ be a system of synchronized observers; then $\{\hat{\xi}^*;\hat{\phi}\}$, where $\hat{\phi}=F(\phi)$ and $\hat{\xi}^*=(F')^{-1}\xi^*$ with $F'\equiv dF(\phi)/d\phi$, is also a system of synchronized observers. As they have the same integral curves and synchronization, but different parameters, we shall say that they are related by a *pure timelike transformation* (i.e., a change in the definition of time).

(d) With every system of synchronized observers $\{\xi^*;\phi\}$ we can associate an operator π^*_ϕ, with $(\pi^*_\phi)^2=\pi^*_\phi$, which projects vector fields into vector fields tangent to the hypersurfaces ϕ_c. Such an operator is given by

$$\pi^*_\phi\equiv\mathrm{Id}-\xi^*\otimes d\phi\,, \tag{3}$$

where Id is the identity operator. In a similar way, we can define the operator π_ξ, with $(\pi_\xi)^2=\pi_\xi$, which projects 1-forms into 1-forms orthogonal to ξ^*; it must therefore have the form

$$\pi_\xi\equiv\mathrm{Id}-d\phi\otimes\xi^*\,. \tag{4}$$

Thus, if ζ^* and ω are, respectively, a vector field and a 1-form, we have

$$\pi^*_\phi\zeta^*=\zeta^*-i(\zeta^*)d\phi\cdot\xi^*\,,\quad \pi_\xi\omega=\omega-i(\xi^*)\omega\cdot d\phi \tag{5}$$

where $i(\,)$ stands for the *interior product* [in local charts $i(\xi^*)\omega\equiv(\xi^*,\omega)=\xi^\rho\omega_\rho$]. It is clear that, with respect to this product, π^*_ϕ and π_ξ are adjoint: $(\pi^*_\phi\zeta^*,\omega)=(\zeta^*,\pi_\xi\omega)$.

In (V_{n+1},g), a system of synchronized observers $\{\xi^*;\phi\}$ is called an *orthogonal system* if, for every vector field ζ^*, we have

$$g(\pi^*_\phi\zeta^*)=\pi_\xi g(\zeta^*)\,,$$

that is to say, if the hypersurfaces ϕ_c are orthogonal to the integral curves of ξ^*. In this case, by condition (2), $g(\xi^*)\equiv\xi=|d\phi|^{-2}\cdot d\phi$.

Thus, with every function ϕ with timelike gradient, we can associate the 1-form

$$\Phi=|d\phi|^{-2}\cdot d\phi \tag{6}$$

so that the pair $\{\Phi^*;\phi\}$, $\Phi^*\equiv g^*(\Phi)$, is an orthogonal system of synchronized observers.

An orthogonal system for which $|d\phi|=1$ is called a *Gaussian system*.

(e) On one hand, a system of observers ξ^* is usually endowed with the *quotient metric* $g_q\equiv g-|\xi|^{-2}\xi\otimes\xi$. On the other, when the spatial hypersurfaces ϕ_c are considered, it is the *induced metric* g_i which is often used, its (contravariant) spatial inverse g_i^{-1} being given by $g_i^{-1}\equiv g^*-|\Phi|^{-2}\Phi^*\otimes\Phi^*$. But a system of synchronized observers is characterized by the set of these two elements,

ξ^* and ϕ, and, consequently, there is no *a priori* preference, in this case, for any one of the two spatial metrics g_q or g_i. The nature of every particular problem will suggest the best spatial metrics to be considered. In the present paper, we are led to take, as suitable spatial metrics $\tilde{g}$, those conformal to the quotient metric: $\tilde{g} = \lambda g_q$.[13]

If ${}^t Q$ denotes the tensor obtained by matrix transposition of a second-rank tensor Q, we obtain ${}^t\pi_\phi^* = \pi_\xi$, ${}^t\pi_\xi = \pi_\phi^*$, owing to their adjoint character. π_ϕ^* and π_ξ may be extended, in the usual way, to the contravariant and covariant tensors respectively. In particular, for second-rank tensors Q, we have

$$\pi_\phi^* Q^* = \pi_\phi^* \times Q^* \times {}^t\pi_\phi^* = \pi_\phi^* \times Q^* \times \pi_\xi\,,$$
$$\pi_\xi Q = \pi_\xi \times Q \times {}^t\pi_\xi = \pi_\xi \times Q \times \pi_\phi^*\,, \tag{7}$$

where $\times$ is the *cross product* (i.e., contraction over the adjacent indices of the tensorial product). Then, it is easy to see that the contravariant tensor $\pi_\phi^* g^*$ is nothing but the spatial inverse of the quotient metric g_q, while the covariant tensor $\pi_\xi g$ is precisely the induced metric g_i:

$$\pi_\phi^* g^* \times g_q = \pi_\phi^*\,, \quad \pi_\xi g \times g_i^{-1} = \pi_\xi\,. \tag{8}$$

The arbitrariness of the spatial metrics that can be assigned to a system of synchronized observers entails that of the spatial volume elements. The quotient and induced volume n-forms attached to the respective spatial metrics may be written

$$\eta_q = |\xi|^{-1} i(\xi^*)\eta\,, \quad \eta_i = |d\phi|\, i(\xi^*)\eta\,, \tag{9}$$

η being the volume $(n+1)$-form of (V_{n+1}, g). Nevertheless, for our particular problem, it is convenient to associate with every system of synchronized observers $\{\xi^*;\phi\}$ a *spatial volume* n-form $\bar{\eta}$ orthogonal to ξ^* and such that the standard 1-form Φ attached to ϕ by (6) is a Leray's form:

$$i(\xi^*)\bar{\eta} = 0\,, \quad \eta = \Phi \wedge \bar{\eta}\,,$$

where $\wedge$ denotes the *exterior product*; $\bar{\eta}$ is then given by

$$\bar{\eta} = |d\phi|^2 i(\xi^*)\eta\,. \tag{10}$$

All the spatial volume n-forms that can be attached to a system of synchronized observers are, of course, conformal to $\bar{\eta}$; hereafter, we shall write $\eta_\mu \equiv \mu\bar{\eta}$.[14]

(f) From $[\mathscr{L}(\xi^*), d] = 0$ and $\mathscr{L}(\xi^*)g^* = -(\mathscr{L}(\xi^*)g)^*$,[15] it follows that

$$\begin{aligned}\mathscr{L}(\xi^*)|d\phi|^2 &= \mathscr{L}(\xi^*)(g^*(d\phi,d\phi))\\ &= (\mathscr{L}(\xi^*)g^*)(d\phi,d\phi)\\ &= -(\mathscr{L}(\xi^*)g)^*(d\phi,d\phi)\\ &= -i^2(d^*\phi)\mathscr{L}(\xi^*)g\,.\end{aligned} \tag{11}$$

Also, from $[\mathscr{L}(\xi^*), i(\xi^*)] = 0$ and the well-known relation $\mathscr{L}(\xi^*)\eta = -\delta\xi\cdot\eta$, where δ is, except for sign, the *divergence operator* (in local charts, $\delta\xi = -\nabla_\rho\xi^\rho$), we have $\mathscr{L}(\xi^*)i(\xi^*)\eta = -\delta\xi\cdot i(\xi^*)\eta$. Thus, taking the Lie derivative in (10), we get

$$\mathscr{L}(\xi^*)\bar{\eta} = -|d\phi|^{-2} i^2(d^*\phi)L(\xi)\cdot\bar{\eta}\,, \tag{12}$$

where the 2-covariant tensor field $L(\xi)$ is given by

$$L(\xi) \equiv \mathscr{L}(\xi^*)g + \delta\xi\cdot g\,. \tag{13}$$

Now, it follows from (12) that

$$\begin{aligned}\mathscr{L}(\xi^*)\eta_\mu &= \mathscr{L}(\xi^*)(\mu\bar{\eta})\\ &= \{\mathscr{L}(\xi^*)\ln\mu - |d\phi|^{-2} i^2(d^*\phi)L(\xi)\}\cdot\eta_\mu\,,\end{aligned}$$

that is to say,

$$\mathscr{L}(\xi^*)\eta_\mu = -|d\phi|^{-2} i(d^*\phi)(Z_\mu - d\ln\mu\cdot\eta_\mu)\,, \tag{14}$$

Z_μ being the 1-form defined by

$$Z_\mu \equiv \pi_\xi d\ln\mu + i(d^*\phi)L(\xi)\,. \tag{15}$$

The 2-covariant tensor field $L(\xi)$ and the 1-form Z_μ, defined respectively by (13) and (15), will play an important role in the remainder of this paper.

III. DECOMPOSITION OF THE LAPLACE OPERATOR

(a) Let us consider a system of synchronized observers $\{\xi^*;\phi\}$, endowed with the spatial metric $\tilde{g}$ conformal to the quotient metric g_q, $\tilde{g} \equiv \lambda g_q$.

It is well known that, for the class $\mathscr{C}(\xi^*;\phi)$ of local charts adapted to ξ^* and to ϕ, the determinants of the Lorentzian and quotient metrics are related by $\mathscr{g} = |\xi|^2 \mathscr{g}_q$, so that we have

$$\tilde{\mathscr{g}} = (\lambda\Lambda)^2 \mathscr{g}\,, \tag{16}$$

where

$$\Lambda \equiv |\xi|^{-1}\lambda^{(n-2)/2}\,. \tag{17}$$

For the elements v^* of $L(\phi)$, the Lie algebra of the vector fields tangent to the hypersurfaces ϕ_c, we shall consider the operator $\tilde{\delta}$ defined by

$$\tilde{\delta} v^* \equiv \delta v^* - \mathscr{L}(v^*)\ln(\lambda\Lambda)\,. \tag{18}$$

In the local charts of $\mathscr{C}(\xi^*;\phi)$, where the first component of v^* vanishes, (18) may be written, on account of (16),

$$-\tilde{\delta} v^* = \partial_l v^l - v^l \partial_l \ln|\tilde{\mathscr{g}}|^{1/2} \tag{19}$$

so that the operator $\tilde{\delta}$ is, except for sign, the *divergence operator on the hypersurfaces* ϕ_c *associated with the spatial metric* $\tilde{g}$.

Moreover, the operator $\bar{d}$ on the exterior forms of V_{n+1}, defined by

$$\bar{d}\alpha \equiv \pi_\xi d\alpha\,, \tag{20}$$

represents the *exterior derivative on the hypersurfaces* ϕ_c of the corresponding induced exterior forms.

(b) Now we can consider *the Laplace operator* $\tilde{\Delta}_n$ *associated with the spatial metric* $\tilde{g}$ *on the hypersurfaces* ϕ_c. On the functions u, it takes the form

$$\tilde{\Delta}_n u \equiv \tilde{\delta}\bar{d}u = \tilde{\delta}(\tilde{g}^{-1}(\bar{d}u))\,. \tag{21}$$

Let us evaluate $\tilde{\Delta}_n u$ in terms of intrinsic differential operators over (V_{n+1}, g). According to (7) and (8) and taking into account (20), we have

$$\begin{aligned}\tilde{g}^{-1}(\bar{d}u) &= (1/\lambda)g_q^{-1}(\bar{d}u)\\ &= (1/\lambda)(\pi_\phi^*\times g^*\times\pi_\xi)(\bar{d}u)\\ &= (1/\lambda)(\pi_\phi^*\times g^*)(\pi_\xi\bar{d}u)\\ &= (1/\lambda)(\pi_\phi^*\times g^*)(\bar{d}u) = (1/\lambda)\pi_\phi^*\bar{d}^*u\,,\end{aligned}$$

where $\bar{d}^*u \equiv g^*(\bar{d}u)$ is the vector field associated by g with the 1-form $\bar{d}u$. From (21) we thus have

$$\lambda \tilde{\Delta}_n u = \lambda \tilde{\delta}((1/\lambda)\pi_\phi^* \bar{d}^* u) = \tilde{\delta}(\pi_\phi^* \bar{d}^* u) + \mathscr{L}(\pi_\phi^* \bar{d}^* u)\ln \lambda$$

or, using the definition (18) of $\tilde{\delta}$,

$$\lambda \tilde{\Delta}_n u = \delta(\pi_\phi^* \bar{d}^* u) - \mathscr{L}(\pi_\phi^* \bar{d}^* u)\ln \Lambda . \tag{22}$$

From the definition (3) of π_ϕ^*, we have

$$\pi_\phi^* \bar{d}^* u = \bar{d}^* u - (d\phi, \bar{d}^* u)\cdot \xi^* ,$$

with

$$(d\phi, \bar{d}^* u) = (d^*\phi, \bar{d}u) = (d^*\phi, \pi_\xi du) = (\pi_\phi^* d^*\phi, \mathrm{d}u) = \mathscr{L}(\pi_\phi^* d^*\phi)u ,$$

and, then, the first term of the second member of Eq. (22) may be written as

$$\begin{aligned}\delta(\pi_\phi^* \bar{d}^* u) &= \delta(\bar{d}^* u - \mathscr{L}(\pi_\phi^* d^*\phi)u\cdot \xi^*)\\ &= \delta \bar{d}^* u - \mathscr{L}(\pi_\phi^* d^*\phi)u\cdot \delta\xi\\ &\quad + \mathscr{L}(\xi^*)\mathscr{L}(\pi_\phi^* d^*\phi)u ,\end{aligned}$$

that is to say,

$$\delta(\pi_\phi^* \bar{d}^* u) = \delta\bar{d}^* u - \delta\xi\cdot \mathscr{L}(\pi_\phi^* d^*\phi)u + \mathscr{L}(\pi_\phi^* d^*\phi)\mathscr{L}(\xi^*)u + \mathscr{L}([\xi^*, \pi_\phi^* d^*\phi])u. \tag{23}$$

On the other hand, by virtue of the mutual adjoint character of π_ξ and π_ϕ^*, the last term of Eq. (22) is

$$\begin{aligned}\mathscr{L}(\pi_\phi^* \bar{d}^* u)\ln \Lambda &= (\pi_\phi^* \bar{d}^* u, d\ln\Lambda)\\ &= (\bar{d}^* u, \pi_\xi d\ln \Lambda)\\ &= (\bar{d}^* u, \bar{d}\ln\Lambda) = (\bar{d}u, \bar{d}^*\ln\Lambda)\\ &= (\pi_\xi du, \bar{d}^*\ln\Lambda) = (du, \pi_\phi^*\bar{d}^*\ln\Lambda),\end{aligned}$$

that is,

$$\mathscr{L}(\pi_\phi^* \bar{d}^* u)\ln\Lambda = \mathscr{L}(\pi_\phi^* \bar{d}^* \ln\Lambda)u . \tag{24}$$

Let us introduce the 1-form Z_Λ defined by (15),

$$Z_\Lambda \equiv \bar{d}\ln\Lambda + i(d^*\phi)L(\xi),$$

and make use of the definition (13) of $L(\xi)$; we have

$$\begin{aligned}Z_\Lambda &= \bar{d}\ln\Lambda + \delta\xi\cdot d\phi + i(d^*\phi)\mathscr{L}(\xi^*)g\\ &= \bar{d}\ln\Lambda + \delta\xi\cdot d\phi - g(\mathscr{L}(\xi^*)d^*\phi)\end{aligned}$$

or, in contravariant form,

$$Z_\Lambda^* = \bar{d}^*\ln\Lambda + \delta\xi\cdot d^*\phi - \mathscr{L}(\xi^*)d^*\phi .$$

It follows that

$$\pi_\phi^*\bar{d}^*\ln\Lambda = \pi_\phi^* Z_\Lambda^* - \delta\xi\cdot\pi_\phi^* d^*\phi + \mathscr{L}(\xi^*)(\pi_\phi^* d^*\phi) \tag{25}$$

because of $\mathscr{L}(\xi^*)\pi_\phi^* = 0$. On the basis of Eqs. (24) and (25), the last term of Eq. (22) may be written

$$\mathscr{L}(\pi_\phi^*\bar{d}^* u)\ln\Lambda = \mathscr{L}(\pi_\phi^* Z_\Lambda^*)u - \delta\xi\cdot\mathscr{L}(\pi_\phi^* d^*\phi)u + \mathscr{L}([\xi^*, \pi_\phi^* d^*\phi])u . \tag{26}$$

Now, with the aid of (23) and (26), Eq. (22) takes the form

$$\lambda\tilde{\Delta}_n u = \delta\bar{d}^* u - \mathscr{L}(\pi_\phi^* Z_\Lambda^*)u + \mathscr{L}(\pi_\phi^* d^*\phi)\mathscr{L}(\xi^*)u , \tag{27}$$

where

$$\begin{aligned}\delta\bar{d}^* u &= \delta(du - \mathscr{L}(\xi^*)u\cdot d\phi)\\ &= \Delta_{n+1}u - \mathscr{L}(\xi^*)u\cdot\Delta_{n+1}\phi\\ &\quad + \mathscr{L}(d^*\phi)\mathscr{L}(\xi^*)u .\end{aligned}$$

Since $\pi_\phi^* d^*\phi = d^*\phi - |d\phi|^2\cdot\xi^*$, the last term of this relation may still be written as

$$\mathscr{L}(d^*\phi)\mathscr{L}(\xi^*)u = \mathscr{L}(\pi_\phi^* d^*\phi)\mathscr{L}(\xi^*)u + |d\phi|^2\mathscr{L}^2(\xi^*)u ;$$

thus, we finally obtain for $\tilde{\Delta}_n u$ the expression

$$\begin{aligned}\lambda\tilde{\Delta}_n u &= \Delta_{n+1}u - \mathscr{L}(\pi_\phi^* Z_\Lambda^*)u\\ &\quad - \Delta_{n+1}\phi\cdot\mathscr{L}(\xi^*)u\\ &\quad + 2\mathscr{L}(\pi_\phi^* d^*\phi)\mathscr{L}(\xi^*)u + |d\phi|^2\mathscr{L}^2(\xi^*)u .\end{aligned} \tag{28}$$

This result may be summarized in the following form:

Theorem 1: With respect to a system of synchronized observers $\{\xi^*;\phi\}$, endowed with a spatial metric $\tilde{g} = \lambda g_q$, the Laplace operator $\Delta_{n+1}u$ of the functions u over (V_{n+1}, g) admits the decomposition

$$\Delta_{n+1}u = |d\phi|^2(-\mathscr{L}^2(\xi^*)u + N\mathscr{L}(\xi^*)u + Mu)$$

where N and M, which are, respectively, first and second-order differential operators on every hypersurface ϕ_c, are given by

$$\begin{aligned}N &\equiv |d\phi|^{-2}(\Delta_{n+1}\phi - 2\mathscr{L}(\pi_\phi^* d^*\phi)),\\ M &\equiv |d\phi|^{-2}(\lambda\tilde{\Delta}_n + \mathscr{L}(\pi_\phi^* Z_\Lambda^*)).\end{aligned}$$

(c) Consider now our system of synchronized observers endowed with an arbitrary spatial volume element η_μ conformal to the standard one defined by Eq. (10): $\eta_\mu \equiv \mu\bar{\eta}$.

We may then define, for suitable[16] functions α, β, a global product on every hypersurface ϕ_c:

$$\langle\alpha,\beta\rangle_\mu \equiv \int_{\phi_c}\alpha\cdot\beta\eta_\mu . \tag{29}$$

Let v^* be a vector field belonging to the Lie algebra $L(\phi)$: $v^* = \pi_\phi^* v^*$. We shall need the following result:

Lemma 1: The formal adjoint, with respect to the measure η_μ, of the operator $|d\phi|^{-2}\mathscr{L}(v^*)$ on functions, is given by

$$\operatorname{ad}_\mu|d\phi|^{-2}\mathscr{L}(v^*) = -|d\phi|^{-2}(\mathscr{L}(v^*) - \delta v^* + \mathscr{L}(v^*)\ln\mu) .$$

Proof: Let δ_μ be the divergence operator associated with the volume element η_μ, i.e., such that, for suitable[17] $z^*\in L(\phi)$,

$$\int_{\phi_c}\delta_\mu z^*\cdot\eta_\mu = 0 .$$

Starting from the identity $\delta_\mu(fv^*) = f\delta_\mu v^* - \mathscr{L}(v^*)f$, we obtain

$$\begin{aligned}\langle\mathscr{L}(v^*)\alpha,\beta\rangle_\mu &= \int_{\phi_c}\mathscr{L}(v^*)\alpha\cdot\beta\eta_\mu\\ &= \int_{\phi_c}(\mathscr{L}(v^*)(\alpha\beta) - \alpha\mathscr{L}(v^*)\beta)\eta_\mu\\ &= \int_{\phi_c}(\alpha\beta\delta_\mu v^* - \alpha\mathscr{L}(v^*)\beta)\eta_\mu ,\end{aligned}$$

that is

$$\langle \mathscr{L}(v^*)\alpha,\beta\rangle_\mu = \langle \alpha,\delta_\mu v^*\cdot\beta\rangle_\mu - \langle \alpha,\mathscr{L}(v^*)\beta\rangle_\mu . \quad (30)$$

But, according to (10), we have $\eta_\mu = \mu|d\phi|^2 i(\xi^*)\eta$, while for the volume element $\tilde\eta$ associated with $\tilde g$ we have, from (16) and (19), $\tilde\eta = \lambda\Lambda i(\xi^*)\eta$. It follows that, for δ_μ, the relation (18) takes the form

$$\delta_\mu v^* = \delta v^* - \mathscr{L}(v^*)\ln(\mu|d\phi|^2) . \quad (31)$$

The elimination of $\delta_\mu v^*$ by means of (30) and (31) yields

$$\mathrm{ad}_\mu \mathscr{L}(v^*) = -\mathscr{L}(v^*) + \delta v^* - \mathscr{L}(v^*)\ln(\mu|d\phi|^2) ,$$

and, thus, for the operator $|d\phi|^{-2}\mathscr{L}(v^*)$, we have

$$\begin{aligned}\langle |d\phi|^{-2}\mathscr{L}(v^*)\alpha,\beta\rangle_\mu &= \langle \mathscr{L}(v^*)\alpha,|d\phi|^{-2}\beta\rangle_\mu\\ &= \langle \alpha,(\mathrm{ad}_\mu\mathscr{L}(v^*))(|d\phi|^{-2}\beta)\rangle_\mu\\ &= \langle \alpha, -\mathscr{L}(v^*)(|d\phi|^{-2}\beta)\\ &\quad + |d\phi|^{-2}(\delta v^* - \mathscr{L}(v^*)\ln(\mu|d\phi|^2))\beta\rangle_\mu\\ &= \langle \alpha,|d\phi|^{-2}(-\mathscr{L}(v^*) + \delta v^*\\ &\quad - \mathscr{L}(v^*)\ln\mu)\beta\rangle_\mu ,\end{aligned}$$

which proves the lemma.

(d) Now let us consider the operator N of Theorem 1. Its adjoint is given by

$$\begin{aligned}\mathrm{ad}_\mu N &= |d\phi|^{-2}\Delta_{n+1}\phi\\ &\quad - 2\,\mathrm{ad}_\mu |d\phi|^{-2}\mathscr{L}(\pi_\phi^* d{*}\phi) ,\end{aligned} \quad (32)$$

where, according to the preceding lemma,

$$\begin{aligned}\mathrm{ad}_\mu |d\phi|^{-2}\mathscr{L}(\pi_\phi^* d{*}\phi) &= -|d\phi|^{-2}\mathscr{L}(\pi_\phi^* d{*}\phi)\\ &\quad + |d\phi|^{-2}(\delta(\pi_\phi^* d{*}\phi) - \mathscr{L}(\pi_\phi^* d{*}\phi)\ln\mu) .\end{aligned} \quad (33)$$

The second term within the brackets takes the form

$$\begin{aligned}\mathscr{L}(\pi_\phi^* d{*}\phi)\ln\mu &= (\pi_\phi^* d{*}\phi,d\ln\mu)\\ &= (d{*}\phi,\pi_\xi d\ln\mu) \equiv i(d{*}\phi)\pi_\xi d\ln\mu ,\end{aligned}$$

while, with the aid of the relation (11), the divergence term may be written

$$\begin{aligned}\delta(\pi_\phi^* d{*}\phi) &= \delta(d{*}\phi - |d\phi|^2\xi^*)\\ &= \Delta_{n+1}\phi - |d\phi|^2\delta\xi^* + \mathscr{L}(\xi^*)|d\phi|^2\\ &= \Delta_{n+1}\phi - i^2(d{*}\phi)(\delta\xi^*\cdot g) - i^2(d{*}\phi)\mathscr{L}(\xi^*)g\\ &= \Delta_{n+1}\phi - i^2(d{*}\phi)L(\xi) .\end{aligned}$$

Hence, referring back to the definition (15) of the 1-form Z_μ, we have

$$\begin{aligned}\delta(\pi_\phi^* d{*}\phi) - \mathscr{L}(\pi_\phi^* d{*}\phi)\ln\mu\\ = \Delta_{n+1}\phi - i(d{*}\phi)Z_\mu .\end{aligned} \quad (34)$$

From (32), (33), and (34), it follows that

Proposition 1: With respect to the measure η_μ, the formal adjoint of the operator

$$N \equiv |d\phi|^{-2}(\Delta_{n+1}\phi - 2\mathscr{L}(\pi_\phi^* d{*}\phi))$$

on every hypersurface ϕ_c is given by

$$\mathrm{ad}_\mu N = -N + 2|d\phi|^{-2} i(d{*}\phi)Z_\mu .$$

(e) Let us now evaluate the formal adjoint of the operator M of Theorem 1. We have seen that the measure $\tilde\eta$ associated with the spatial metric $\tilde g$ on every hypersurface ϕ_c is given by $\tilde\eta = \lambda\Lambda i(\xi^*)\eta$ so that, according to (10), we have $\tilde\eta = \lambda\Lambda|d\phi|^{-2}\mu^{-1}\eta_\mu$. Moreover, it is evident that $\langle\alpha,\beta\rangle_\mu = \langle\alpha,f^{-1}\beta\rangle_{f\mu}$ for all f such that $f(x)\neq 0$; thus, taking $f\equiv\lambda\Lambda/\mu|d\phi|^2$, we may write, for the operator $|d\phi|^{-2}\lambda\tilde\Delta_n$,

$$\begin{aligned}\langle |d\phi|^{-2}\lambda\tilde\Delta_n\alpha,\beta\rangle_\mu &= \langle\tilde\Delta_n\alpha,|d\phi|^{-2}\lambda\beta\rangle_\mu\\ &= \langle\tilde\Delta_n\alpha,f^{-1}|d\phi|^{-2}\lambda\beta\rangle_{f\mu}\\ &= \langle\tilde\Delta_n\alpha,\mu\Lambda^{-1}\beta\rangle_{f\mu} = \langle\alpha,\tilde\Delta_n(\mu\Lambda^{-1}\beta)\rangle_{f\mu}\\ &= \langle\alpha,f\tilde\Delta_n(\mu\Lambda^{-1}\beta)\rangle_\mu\\ &= \langle\alpha,\mu^{-1}\Lambda(|d\phi|^{-2}\lambda\tilde\Delta_n)(\mu\Lambda^{-1}\beta)\rangle_\mu ,\end{aligned}$$

which shows that the necessary and sufficient condition for $|d\phi|^{-2}\lambda\tilde\Delta_n$ to be self-adjoint for the measure η_μ is

$$\mu^{-1}\Lambda = h(\phi) , \quad (35)$$

where $h(\phi)$ denotes any arbitrary nonnull function on V_{n+1}, constant on every hypersurface ϕ_c. Consider for a moment a λ verifying (35); in such case, we have $Z_\Lambda = Z_\mu$ and, according to Lemma 1,

$$\begin{aligned}\mathrm{ad}_\mu M &= \mathrm{ad}_\mu |d\phi|^{-2}(\lambda\tilde\Delta_n + \mathscr{L}(\pi_\phi^* Z_\mu^*))\\ &= |d\phi|^{-2}\lambda\tilde\Delta_n - |d\phi|^{-2}(\mathscr{L}(\pi_\phi^* Z_\mu^*)\\ &\quad - \delta(\pi_\phi^* Z_\mu^*) + \mathscr{L}(\pi_\phi^* Z_\mu^*)\ln\mu) ,\end{aligned}$$

that is to say,

$$\begin{aligned}\mathrm{ad}_\mu M &= M - |d\phi|^{-2}(2\mathscr{L}(\pi_\phi^* Z_\mu^*) - \delta(\pi_\phi^* Z_\mu^*)\\ &\quad + \mathscr{L}(\pi_\phi^* Z_\mu^*)\ln\mu) .\end{aligned} \quad (36)$$

But Theorem 1 states that M does not depend on λ, and so neither does its adjoint with respect to the measure η_μ: (36) is thus valid for all λ, whether solutions of Eq. (35) or not. We have thus proved:

Proposition 2: With respect to the measure η_μ, the formal adjoint of the operator

$$M \equiv |d\phi|^{-2}(\lambda\tilde\Delta_n + \mathscr{L}(\pi_\phi^* Z_\mu^*))$$

on every hypersurface ϕ_c is given by

$$\begin{aligned}\mathrm{ad}_\mu M &= M - |d\phi|^{-2}(2\mathscr{L}(\pi_\phi^* Z_\mu^*) - \delta(\pi_\phi^* Z_\mu^*)\\ &\quad + \mathscr{L}(\pi_\phi^* Z_\mu^*)\ln\mu) .\end{aligned}$$

IV. FIRST-ORDER FORMULATION

(a) Consider a second-order differential equation of the Klein–Gordon type, namely,

$$\Delta_{n+1}u - \kappa u = 0 , \quad (37)$$

where κ is a positive function on V_{n+1}, $\kappa(x) > 0\ \forall x$. It is well known that Eq. (37) may be obtained as the Euler–Lagrange equations of a standard variational principle, whereby the energy tensor is given by

$$T(u) = du\otimes du - \tfrac{1}{2}[g^*(du,du) - \kappa u^2]g . \quad (38)$$

The fact that $T(u)$ is not divergence-free for $\kappa\neq$ const is not essential here.

According to Theorem 1, with respect to an arbitrary system of synchronized observers $\{\xi^*;\phi\}$, Eq. (37) may be written

$$-\mathscr{L}^2(\xi^*)u + N\mathscr{L}(\xi^*)u + M_\kappa u = 0 , \quad (39)$$

where

$$M_\kappa \equiv M - |d\phi|^{-2}\kappa . \tag{40}$$

But, in any chart of the class $\mathscr{C}(\xi^*;\phi)$, the Lie derivative reduces to the time derivative, $\mathscr{L}(\xi^*) \equiv d/dt$, so that Eq. (39) takes the form

$$\frac{d^2 u}{dt^2} = N\frac{du}{dt} + M_\kappa u ,$$

or, as a first-order system,

$$\frac{dX}{dt} = TX \tag{41}$$

with

$$X \equiv \begin{pmatrix} u \\ v \end{pmatrix}, \quad v \equiv \frac{du}{dt}, \quad T \equiv \begin{pmatrix} 0 & I \\ M_\kappa & N \end{pmatrix}.$$

We are thus led to consider the space $\mathscr{X} \equiv \{X\}$ of pairs of functions at every instant ϕ_c. This is the space of Cauchy data for Eq. (37).

(b) With respect to the system of observers ξ^*, we associate with every solution u of Eq. (37) the *energy–momentum density* $P(u)$ given by

$$P(u) \equiv i(\xi^*)T(u) ,$$

where $T(u)$ is the energy–tensor (38). The *energy density* $e(u)$ relative to the synchronized system $\{\xi^*;\phi\}$ is then defined by

$$e(u) \equiv i(\Phi^*)P(u) , \tag{42}$$

where Φ is the 1-form (6) attached to the synchronization ϕ of ξ^*. Thus, in a domain Ω_ϕ of any instant ϕ_c, the *total energy* $E(u)$ relative to the measure η_μ is

$$E(u) = \int_{\Omega_\phi} e(u)\eta_\mu . \tag{43}$$

In particular, if $\mu = 1$, i.e., if the measure is the one defined by the spatial volume element $\bar\eta$ given by Eq. (10), the corresponding total energy $\bar E(u)$ is

$$\bar E(u) = \int_{\Omega_\phi} e(u)\bar\eta = \int_{\Omega_\phi} i(\Phi^*)i(\xi^*)T(u)\cdot |d\phi|^2 i(\xi^*)\eta$$
$$= \int_{\Omega_\phi} i(d^*\phi)i(\xi^*)T(u)\cdot i(\xi^*)\eta ,$$

or, in local charts of $\mathscr{C}(\xi^*;\phi)$,

$$\bar E(u) = \int_{\Omega_\phi} T^0_{\ 0}(-g)^{1/2}\,dx^n ,$$

which shows that $\bar E(u)$ is the usual total energy in the standard cases.

If Z^*_μ denotes the vector field associated by g with the 1-form Z_μ defined by Eq. (15), we have the following result:

Proposition 3: At every instant ϕ_c of a system of synchronized observers $\{\xi^*;\phi\}$, the total energy $E(u)$ relative to the measure η_μ is given by

$$2E(u) = \langle \mathscr{L}(\xi^*)u, \mathscr{L}(\xi^*)u\rangle_\mu - \langle M_E u, u\rangle_\mu , \tag{44}$$

where

$$M_E \equiv M_\kappa - |d\phi|^{-2}\mathscr{L}(\pi^*_\phi Z^*_\mu) . \tag{45}$$

Proof: From Eqs. (42) and (38), it follows that

$$e(u) = |d\phi|^{-2}(\mathscr{L}(\xi^*)u\cdot(d\phi,du) - \tfrac{1}{2}(du,du) + \tfrac{1}{2}\kappa u^2)$$

or, by the definition (20) of $\bar d$,

$$2e(u) = \mathscr{L}(\xi^*)u\cdot\mathscr{L}(\xi^*)u - |d\phi|^{-2}((\bar d u,\bar d u) - \kappa u^2) .$$

We thus have

$$2E(u) - \langle \mathscr{L}(\xi^*)u, \mathscr{L}(\xi^*)u\rangle_\mu - \langle |d\phi|^{-2}\kappa u, u\rangle_\mu = \int |d\phi|^{-2}(\bar d u,\bar d u)\eta_\mu \equiv I . \tag{46}$$

Now, using the conformal spatial metric $\tilde g$ and integrating by parts, we have

$$I = \int \tilde g^{-1}(\bar d u, \bar d(\mu\Lambda^{-1}u))\tilde\eta - \int \tilde g^{-1}(\bar d u, \bar d(\mu\Lambda^{-1}))u\tilde\eta \equiv I_1 - I_2 .$$

The first integral may thus be written

$$I_1 = \int \tilde\Delta_n u\cdot\mu\Lambda^{-1}u\tilde\eta = \int |d\phi|^{-2}\lambda\tilde\Delta_n u\cdot u\eta_\mu , \tag{47}$$

and, if the scalar product on V_{n+1} and the measure η_μ are restored, the second one takes the form

$$I_2 = \int |d\phi|^{-2}(\bar d u, \bar d\ln(\mu\Lambda^{-1}))u\eta_\mu = \int |d\phi|^{-2}(du, \pi^*_\phi\bar d^*\ln(\mu\Lambda^{-1}))u\eta_\mu ,$$

or, according to the expression (25), valid for all Λ,

$$I_2 = \int |d\phi|^{-2}(\mathscr{L}(\pi^*_\phi Z^*_\mu)u - \mathscr{L}(\pi^*_\phi Z^*_\Lambda)u)u\eta_\mu . \tag{48}$$

Then, the relation (44) follows from Eqs. (46)–(48).

The quadratic form $E(u)$ induces the bilinear form $E(u,u')$ by

$$2E(u,u') \equiv \langle \mathscr{L}(\xi^*)u, \mathscr{L}(\xi^*)u'\rangle_\mu - \langle M_E u, u'\rangle_\mu , \tag{49}$$

which defines a product $\langle\ ,\ \rangle_E$ on the space $\mathscr{X} \equiv \{X\}$ of Cauchy data considered above: if $X \equiv (u, v)$ and $X' \equiv (u', v')$ are two elements of $\mathscr{X}$, then

$$\langle X, X'\rangle_E = \langle v, v'\rangle_\mu - \langle M_E u, u'\rangle_\mu . \tag{50}$$

This we shall call the *energy product* on $\mathscr{X}$.

Let us evaluate the formal adjoint, $\mathrm{ad}_E T$, of the operator T with respect to this energy product. A straightforward calculation gives

$$\mathrm{ad}_E T = -\begin{pmatrix} 0 & A \\ \mathrm{ad}_\mu M_E & -\mathrm{ad}_\mu N \end{pmatrix},$$

where the operator A is defined by the relation

$$\mathrm{ad}_\mu M_\kappa = (\mathrm{ad}_\mu M_E)A .$$

Taking into account Propositions 1 and 2, Lemma 1, and the fact that M_E is self-adjoint with respect to η_μ, and putting $A \equiv \mathrm{Id} - |d\phi|^{-2}H$, one can show that:

Proposition 4: The formal adjoint of the operator T with respect to the energy product is given by

$$\mathrm{ad}_E T = -T + |d\phi|^{-2} V, \tag{51}$$

where

$$V \equiv \begin{pmatrix} 0 & H \\ \mathscr{L}(\pi_\phi^* Z_\mu^*) & 2\mathscr{L}(Z_\mu^*)\phi \end{pmatrix}$$

and H is the operator implicitly defined by

$$M_E H = \mathrm{ad}_\mu\, [\, |d\phi|^{-2} \mathscr{L}(\pi_\phi^* Z_\mu^*)]\ .$$

In the Minkowskian case, and with respect to *inertial* systems of synchronized observers, the operator T, which characterizes the evolution form (41) of the Klein–Gordon-like equation (37), is skew-self-adjoint with respect to the energy product. From Proposition 4, it is clear that this is not true, in general, if we change either the synchronized system of observers or the Lorentzian manifold. For the reasons mentioned in the Introduction, we are interested in the situations under which the skew-self-adjoint character of T is maintained. Taking into account Propositions 1, 2, 4, and Eq. (45), one can easily prove:

Proposition 5: For the operator T to be skew-self-adjoint with respect to the energy product, it is necessary and sufficient that the operators M and N be respectively self-adjoint and skew-self-adjoint with respect to the measure η_μ or, equivalently, that the 1-form Z_μ be zero. In such case, the operators M_E and M_κ coincide.

Note the role of the function κ: It merely insures the definite character of the energy bilinear form (49), and has no influence on the skew-self-adjointness of T, as follows from the preceding proposition. This latter property is related only to the "Laplacian" structure of V_{n+1}, i.e., to the metric structure (V_{n+1}, g).

V. ADJOINT SYSTEMS OF SYNCHRONIZED OBSERVERS

(a) In this section, we shall consider any Lorentzian structure (V_{n+1}, g) on a manifold V_{n+1} as being associated with a set of second-order equations of the Klein–Gordon form parametrized by an arbitrary function κ. As we have seen in the preceding section, for each system of synchronized observers, the structure (V_{n+1}, g) determines the set of operators T defining the first-order evolution form of every equation of the former set.

Definition 1: A system of synchronized observers $\{\xi^*;\phi\}$ will be called an *adjoint system* if it can be endowed with a measure η_μ such that the operators T be skew-self-adjoint with respect to the corresponding energy bilinear form.

By Proposition 5, a system of synchronized observers $\{\xi^*;\phi\}$ will be adjoint iff the 1-form Z_μ is zero, i.e., iff

$$i(\xi^*) Z_\mu = 0\,, \tag{52a}$$

$$\pi_\xi Z_\mu = 0\,. \tag{52b}$$

With the definition (15) of Z_μ, (52b) may be written

$$\bar{d} \ln \mu = -\pi_\xi i(d^*\phi) L(\xi)\,, \tag{53}$$

and thus, it must be verified at every instant ϕ_c. The necessary and sufficient local integrability condition for (53) is then

$$\bar{d}\pi_\xi i(d^*\phi) L(\xi) = 0\,,$$

or, by Eq. (52a) and definition (20) of the derivative $\bar{d}$,

$$di(d^*\phi) L(\xi) + i(\xi^*) di(d^*\phi) L(\xi) \wedge d\phi = 0\,, \tag{54a}$$

which may be written in the equivalent form

$$d\phi \wedge di(d^*\phi) L(\xi) = 0\,. \tag{54b}$$

We thus have

Theorem 2: For a system of synchronized observers $\{\xi^*;\phi\}$ to be adjoint, it is necessary and sufficient that

$$i(d^*\phi) i(\xi^*) L(\xi) = 0\,, \tag{55a}$$

$$d\phi \wedge di(d^*\phi) L(\xi) = 0\,. \tag{55b}$$

It is interesting to remark that the measure η_μ does not appear in the determination of the adjoint systems. Thus, they depend only on the Lorentzian structure (V_{n+1}, g).

(b) Let $\{\hat{\xi}^*;\hat{\phi}\}$ and $\{\xi^*;\phi\}$ be two systems of synchronized observers related by a pure timelike transformation [see Sec. II(c)], i.e., such that $\hat{\phi} = F(\phi)$, $\hat{\xi}^* = (F')^{-1}\xi^*$. A direct calculation shows that

$$\mathscr{L}(\hat{\xi}^*) g = (1/F')\mathscr{L}(\xi^*) g - (F''/F'^2)(d\phi \wedge \xi + \xi \wedge d\phi)$$

and then, as $\mathrm{tr}[\mathscr{L}(\xi^*)g] = -2\delta\xi^*$,

$$\delta\hat{\xi}^* = (1/F')\delta\xi^* + F''/F'^2\,.$$

It follows that

$$L(\hat{\xi}) = (1/F') L(\xi) + (F''/F'^2)(g - d\phi \wedge \xi - \xi \wedge d\phi)\,,$$

and thus

$$i(d^*\hat{\phi}) L(\hat{\xi}) = i(d^*\phi) L(\xi) - (F''/F'^2)|d\phi|^2 \xi\,.$$

If both systems are adjoint, from the condition (54a), it must be that $F'' = 0$ and, for such an F, the condition (54b) is identically satisfied. We then have

Proposition 6: The only pure timelike transformations admitted by the adjoint systems of synchronized observers are linear transformations.

This proposition says that the time parameter is affine for adjoint systems, just as for Killing systems of observers.

(c) Consider now an orthonormal system of synchronized observers $\{\xi^*;\phi\}$ [see Sec. II(d)]. We then have $\xi = |d\phi|^{-2} d\phi$, and it follows that

$$\mathscr{L}(\xi^*) g = |d\phi|^{-2}(\mathscr{L}(d^*\phi) g - d\ln|d\phi|^2 \otimes d\phi - d\phi \otimes d\ln|d\phi|^2)$$

and thus

$$\delta\xi^* = |d\phi|^{-2}(\delta d\phi + \mathscr{L}(d^*\phi)\ln|d\phi|^2)\,.$$

From these expressions, we obtain

$$\begin{aligned} i(d^*\phi) L(\xi) &\equiv i(d^*\phi)\mathscr{L}(\xi^*) g + \delta\xi \cdot d\phi \\ &= |d\phi|^{-2}(\mathscr{L}(d^*\phi) d\phi - |d\phi|^2 d\ln|d\phi|^2 + \delta d\phi \cdot d\phi)\,, \end{aligned}$$

but, as $\mathscr{L}(\xi^*)\phi = 1$, it follows that $\mathscr{L}(d^*\phi) d\phi = d\mathscr{L}(d^*\phi)\phi = d\,|d\phi|^2$, and we may write

$$i(d^*\phi) L(\xi) = \Delta_{n+1}\phi \cdot \xi\,. \tag{56}$$

For adjoint systems, the condition (55a) imposes $\Delta_{n+1}\phi = 0$ and, for such a ϕ, (56) implies that the condition (55b) is identically verified. We have thus proven:

Proposition 7: For an orthogonal system of synchronized observers to be adjoint, it is necessary and sufficient that the synchronization function ϕ be harmonic: $\Delta_{n+1}\phi = 0$.

Since there are always local solutions to the harmonic equation, we have, as a consequence:

Theorem 3: All Lorentzian structures (V_{n+1}, g) admit local adjoint systems of synchronized observers.

(d) A spatial volume element $\eta_\mu \equiv \mu\bar{\eta}$, where $\bar{\eta}$ is given by Eq. (10) and μ satisfies

$$Z_\mu \equiv \pi_\xi d \ln \mu + i(d{}^*\phi) L(\xi) = 0\,, \tag{57}$$

will be said to be *admissible* for the adjoint system of synchronized observers $\{\xi^*;\phi\}$. It is clear from Eq. (57) that if μ and $\hat{\mu}$ define two admissible volumes η_μ and $\eta_{\hat{\mu}}$ then they are related by $\mu - \hat{\mu} = \tau(\phi)$, where $\tau(\phi)$ is a constant function on every instant ϕ_c.

From Eq. (14), we see that an admissible volume element η_μ for an adjoint system $\{\xi^*;\phi\}$ will be ξ^*-invariant iff

$$\mathscr{L}(d{}^*\phi)\mu = 0\,. \tag{58}$$

In such a case, contracting Eq. (57) with $d\phi$, we find

$$\mathscr{L}(\xi^*)\ln \mu = |d\phi|^{-2} i^2(d{}^*\phi) L(\xi)$$

so that Eq. (57) may be written

$$d \ln \mu - i(d{}^*\phi) L(\xi) + |d\phi|^{-2} i^2(d{}^*\phi) L(\xi)\cdot d\phi = 0\,. \tag{59}$$

Equations (58) and (59) are equivalent for adjoint systems. The local integrability condition for Eq. (59) is

$$d i(d{}^*\phi) L(\xi) - d(|d\phi|^{-2} i^2(d{}^*\phi) L(\xi)) \wedge d\phi = 0\,, \tag{60}$$

from which, by contracting with ξ^*, we obtain

$$\mathscr{L}(\xi^*) i(d{}^*\phi) L(\xi) + \bar{d}(|d\phi|^{-2} i^2(d{}^*\phi) L(\xi)) = 0\,. \tag{61}$$

Equations (60) and (61) are equivalent via Eq. (54). We thus have proven

Proposition 7: The adjoint systems of synchronized observers for which an admissible spatial volume element η_μ can be found such that $\mathscr{L}(\xi^*)\eta_\mu = 0$, satisfy the condition

$$\mathscr{L}(\xi^*) i(d{}^*\phi) L(\xi) + \bar{d}(|d\phi|^{-2} i^2(d{}^*\phi) L(\xi)) = 0\,.$$

According to Theorem 2, a sufficient condition for a system of synchronized observers $\{\xi^*;\phi\}$ to be adjoint is

$$i(d{}^*\phi) L(\xi) = 0\,. \tag{62}$$

Such a system is said to be a *simple adjoint system*. From the preceding proposition and Eqs. (15) and (12), it follows that

Proposition 8: Simple adjoint systems are the only adjoint systems for which $\bar{\eta}$ is an admissible volume element. $\bar{\eta}$ is then a ξ^*-invariant volume element.

The class of simple adjoint systems was originally considered in Ref. 3.

VI. ADJOINT SYSTEMS OF OBSERVERS

(a) The adjoint systems of observers are, in a certain sense, an *integrable*[18] generalization of the Killing fields: From Theorem 2, it is clear that the Killing observers (timelike Killing vector fields) always generate adjoint systems of synchronized observers. But if ξ^*_k is a Killing system of observers, all the pairs $\{\xi^*_k;\phi\}$, where ϕ is such that $\mathscr{L}(\xi^*_k)\phi = 1$, are adjoint systems of synchronized observers.

We are thus led to introduce the following definition:

Definition 2: A system of observers ξ^* will be said to be an *adjoint system of observers* if, for every synchronization function ϕ, $\mathscr{L}(\xi^*)\phi = 1$, the pair $\{\xi^*;\phi\}$ is an adjoint system of synchronized observers.

Thus, in particular, the Killing systems of observers are adjoint. But, as we shall see, they are not the only ones to have this property.

(b) Let $\{\xi^*;\phi\}$ be an adjoint system of synchronized observers, $\mathscr{S}(\xi^*)$ the space of synchronization functions for ξ^*, and $\mathscr{F}(\xi^*)$ that of the ξ^*-invariant functions: $\psi\in\mathscr{S}(\xi^*) \Leftrightarrow \mathscr{L}(\xi^*)\psi = 1$, $h\in\mathscr{F}(\xi^*)\Leftrightarrow\mathscr{L}(\xi^*)h = 0$.

Linearizing Eq. (55a) in a neighborhood of ϕ, we have

$$i(d{}^*h) i(\xi^*) L(\xi) = 0\,,$$

where $h\in\mathscr{F}(\xi^*)$; it follows that, for ξ^* to be an adjoint system, we must have $i(\xi^*)L(\xi) = \lambda\xi$, together with Eq. (55a), which implies that

$$i(\xi^*) L(\xi) = 0\,. \tag{63}$$

(63) is a necessary and sufficient condition for (55a) to be verified for all $\psi\in\mathscr{S}(\xi^*)$.

Similarly, linearizing Eq. (54), we have

$$d i(d{}^*h) L(\xi) + i(\xi^*) d i(d{}^*h) L(\xi) \wedge d\phi + i(\xi^*) d i(d{}^*\phi) L(\xi) \wedge dh = 0 \tag{64}$$

with $h\in\mathscr{F}(\xi^*)$. At a point $x\in V_{n+1}$, take $dh|_x = 0$; then, on account of Eq. (63), Eq. (64) reduces to

$$\nabla dh \times L(\xi) - L(\xi)\times\nabla dh + i(\xi^*)(\nabla dh \times L(\xi)) \wedge d\phi = 0\,, \tag{65}$$

where $\times$ is the cross product [see Sec. II(e)]. In an arbitrary local chart, with $L_{\alpha\beta}$ denoting the components of $L(\xi)$, Eq. (65) may be written

$$\delta^{\mu\nu}_{\alpha\beta} L^{\rho}_{\ \nu}(\delta^\sigma_\mu - \xi^\sigma\partial_\mu\phi)\partial_{\sigma\rho} h = 0\,,$$

where $\delta^{\mu\nu}_{\alpha\beta}$ is a Kronecker determinant; and thus, in a local chart of the class $\mathscr{C}(\xi^*;\phi)$, this equation reduces to

$$\delta^{\mu\nu}_{\alpha\beta} L^{i}_{\ \nu}\delta^{j}_{\ \mu}\partial_{ij} h = 0\,, \tag{66}$$

where Latin indices take the values $1,\ldots,n$. For all $h\in\mathscr{F}(\xi^*)$ such that $dh|_x = 0$, Eq. (66) implies

$$L^{i}_{\ k}\delta^{j}_{\ l} + L^{j}_{\ k}\delta^{i}_{l} - L^{i}_{\ l}\delta^{j}_{\ k} - L^{j}_{\ l}\delta^{i}_{k} = 0\,,$$

whose solution is

$$L^{i}_{\ k} = (1/n)\mathrm{tr}[L(\xi)]\cdot\delta^{i}_{k}\,,$$

which may be written in arbitrary local charts in the form

$$L^{\ \beta}_{\alpha} - \xi^\beta\partial_\rho\phi L^{\ \rho}_{\alpha} = (1/n)\mathrm{tr}[L(\xi)](\delta^\beta_\alpha - \partial_\alpha\phi\xi^\beta)$$

or, in intrinsic notation,

$$L(\xi) - i(d{}^*\phi) L(\xi)\otimes\xi = (1/n)\mathrm{tr}[L(\xi)](g - d\phi\otimes\xi)\,. \tag{67}$$

By linearizing (67) again in the neighborhood of ϕ, we obtain

$$i(d{}^*h)(L(\xi) - (1/n)\mathrm{tr}[L(\xi)]\cdot g) = 0\,,$$

and, consequently,

$$L(\xi) - (1/n)\mathrm{tr}[L(\xi)]\cdot g = a\otimes\xi\,,$$

which implies

$$a = -(1/n|\xi|^2)\mathrm{tr}[L(\xi)]\cdot\xi .$$

Thus, we must have

$$L(\xi) = (1/n)\mathrm{tr}[L(\xi)][g-(1/|\xi|^2)\xi\otimes\xi] . \tag{68}$$

(68) is a necessary condition for (54) to be verified for all $\psi\in\mathscr{S}(\xi^*)$, but it is not sufficient: applying operators $di(d^*\phi)$ to (68), we have

$$\begin{aligned} di(d^*\phi)L(\xi) &= (1/n)d\,\mathrm{tr}[L(\xi)]\wedge d\phi \\ &\quad -(1/n)d(|\xi|^{-2}\,\mathrm{tr}[L(\xi)]\cdot\xi) \end{aligned}$$

which, by imposition of (55b) $\forall\ \phi\in\mathscr{S}(\xi^*)$, gives

$$d(|\xi|^{-2}\,\mathrm{tr}[L(\xi)]\cdot\xi) = 0 . \tag{69}$$

Conversely, if Eqs. (68) and (69) are verified, then (55) is identically verified for all $\psi\in\mathscr{S}(\xi^*)$. Then, as $\mathrm{tr}[L(\xi)] = -2\delta\xi + (n+1)\delta\xi = (n-1)\delta\xi$, we have

Proposition 9: A timelike vector field ξ^* defines an adjoint system of observers iff it is such that

$$L(\xi) = \frac{n-1}{n}\delta\xi\,[g-(1/|\xi|^2)\xi\otimes\xi] \tag{70a}$$

and

$$d(|\xi|^{-2}\delta\xi\cdot\xi) = 0 . \tag{70b}$$

(c) From the above proposition it is evident that not all Lorentzian structures admit such a vector field ξ^*. In order to have an idea of the restrictions imposed by Eqs. (70), let us formulate them in terms of geometric congruences or, equivalently, in terms of unitary vector fields.

Let us write $\xi^*\equiv e^{-T}\cdot u^*$ with $|u^*|^2 = 1$. A direct calculation gives

$$\begin{aligned} L(\xi) = e^{-T}[&\mathscr{L}(u^*)g + (\delta u + \dot T)g \\ &- dT\otimes u - u\otimes dT] \end{aligned} \tag{71}$$

with $\dot T\equiv\mathscr{L}(u^*)T$. It follows that the equation $i(\xi^*)L(\xi) = 0$ is equivalent to

$$dT = a + \delta u\cdot u , \tag{72}$$

where $a\equiv\mathscr{L}(u^*)u$ is the proper acceleration of u^*. With the aid of Eqs. (71) and (72), Eq. (70) may be written

$$\mathscr{L}(u^*)(g-u\otimes u) = -(2/n)\delta u(g-u\otimes u) , \tag{73a}$$

$$d(\delta u\cdot u) = 0 . \tag{73b}$$

Actually, it is easy to see that systems $\{(72),(73)\}$ and $\{(70)\}$ are equivalent. The local integrability condition for (72) is, on account of (73b), $da = 0$, but (73b) may also be written

$$\delta u(a\wedge u + du) = 0 . \tag{74}$$

If $\delta u = 0$, (73a) tells us that u^* is a rigid field in the Born sense and then, as is well known, the condition $da = 0$ implies that it is necessarily a Killing field. If $\delta u\neq 0$, by taking the Lie derivative of (74), we get

$$\mathscr{L}(u)a\wedge u + da = 0$$

so that $da = 0\Leftrightarrow\mathscr{L}(u^*)a = 0$. We thus have

Theorem 4: The unitary vector fields u^* associated with the adjoint systems of observers are the conformally rigid fields,

$$\mathscr{L}(u^*)(g-u\otimes u) = -(2/n)\delta u(g-u\otimes u) ,$$

which, in addition, verify

$$d(\delta u\cdot u) = 0 ,\quad \delta u\cdot\mathscr{L}(u^*)a = 0 .$$

Thus, the function T characterizing the adjoint system ξ^*, $\xi^*\equiv e^Tu^*$, is given, apart from an additive constant, by

$$dT = a + \delta u\cdot u .$$

The class of space–times which admit conformally rigid vector fields is very important in general relativity. It contains, in particular, all the Robertson–Walker space–times[19] and the Thompson–Witrow[20] space–times (i.e., nonstationary, spherically symmetric perfect fluids).

(d) It is easy to see, from Eq. (70), that only the adjoint systems of observers ξ^* for which $\delta\xi = 0$ are Killing fields. Let us consider the case $\delta\xi\neq 0$: Eq. (70b) tells us that there exists a family of hypersurfaces orthogonal to ξ^*; taking its local equation to be of the form $\phi(x) = \mathrm{const}$ with $\phi\in\mathscr{S}(\xi^*)$, we have

$$|\xi|^{-2}\delta\xi\cdot\xi = -\mathscr{L}(\xi^*)\ln B\cdot d\phi ,$$

where B is a constant function on every instant $\phi_c\equiv(\phi(x) = const)$, say $B = B(\phi)$. It follows that

$$\delta\xi = -\mathscr{L}(\xi^*)\ln B , \tag{75}$$

and, of course, $\xi = |\xi|^{-2}d\phi$. Moreover, from the definition of $L(\xi)$ and Eq. (70a), we have

$$\mathscr{L}(\xi^*)g = -(\delta\xi/n)[g + (n-1)|\xi|^2 d\phi\otimes d\phi] . \tag{76}$$

Now take a local chart of the class $\mathscr{C}(\xi^*;\phi)$. We there have $\phi = x^0$, $\xi^* = \partial_0\sim\xi^\alpha = \delta^\alpha_0$, and, thus, $\partial_\alpha\phi = \delta^0_\alpha$, $g_{0i} = 0$, $g_{00} = |\xi|^2$. Equations (75) and (76) become

$$\begin{aligned} \partial_0\ln|\varphi|^{1/2} &= \partial_0\ln B(x^0) , \\ \partial_0 g_{\alpha\beta} &= -(1/n)\partial_0\ln B(x^0)[g_{\alpha\beta} \\ &\quad + (n-1)g_{00}\delta^0_\alpha\delta^0_\beta] , \end{aligned}$$

which may be integrated yielding

$$\begin{aligned} \varphi &= B^2(x^0)\cdot h(x^l) , \\ g_{00} &= B(x^0)\cdot\sigma(x^l) , \\ g_{ij} &= -B^{1/n}(x^0)\cdot\sigma(x^l)\cdot f_{ij}(x^l) . \end{aligned}$$

These relations are compatible because, from the identity $\varphi\equiv\det g_{\alpha\beta} = g_{00}\det g_{ij}$ it follows that

$$h(x^l) = -\sigma^{n+1}(x^l)\cdot\det[f_{ij}(x^l)] .$$

Thus, in those charts, we have

$$\begin{aligned} g = \sigma(x^l)[&B(x^0)dx^0\otimes dx^0 \\ &- B^{1/n}(x^0)f_{ij}(x^l)dx^i\otimes dx^j] . \end{aligned}$$

The vector field $\xi^*\equiv\partial_0$ defines an adjoint system of observers both for the metric g and for the (generalized) Robertson–Walker metric

$$g_{\mathrm{RW}}\equiv B(x^0)dx^0\otimes dx^0 - B^{1/n}(x^0)f_{ij}(x^l)dx^i\otimes dx^j ,$$

where the usual "cosmic time" t is related to the "adjoint time" x^0 by $dt = B^{1/2}(x^0)dx^0$. We have thus

Theorem 5: A Lorentzian metric g admits adjoint systems of observers other than the Killing ones iff it is conformal to a Robertson–Walker metric g_{RW} and the conformal factor is invariant by ξ^*: $g = \sigma g_{\mathrm{RW}}$, $\mathscr{L}(\xi^*)\sigma = 0$.

VII. REMARKS ON THE GENERAL CASE

(a) The adjoint systems of synchronized observers considered in Sec. V are, by Proposition 5, those for which the operators M and N characterizing the evolution form of the Laplace operator are respectively self- and skew-self-adjoint with respect to some suitable measure. In Sec. IV, we have seen that this property is intimately related to differential equations of the Klein–Gordon type. Here we shall see how the above method may be extended to the general case.

A general second-order hyperbolic linear equation in a (domain of) $\mathbb{R}^{n+1}$ may be written, in invariant form[21]

$$\Delta_{n+1} u + \mathscr{L}(\Theta^*)u + \kappa u = 0\,, \tag{77}$$

where Δ_{n+1} is the Laplace operator defined by the metric associated with the hyperbolic character of the equation, and Θ^* and κ are respectively a vector field and a scalar function.

From Theorem 1, it follows that *with respect to each synchronized system* $\{\xi^*;\phi\}$, *Eq. (77) may be written in the form*

$$|d\phi|^2(-\mathscr{L}^2(\xi^*)u + N_\Theta \mathscr{L}(\xi^*)u + M_{\kappa,\Theta}u) = 0\,, \tag{78}$$

where the operators N_Θ *and* $M_{\kappa,\Theta}$ *are given by*

$$\begin{aligned} N_\Theta &\equiv N + |d\phi|^{-2}\mathscr{L}(\Theta^*)\phi\,,\\ M_{\kappa,\Theta} &\equiv M_\Theta - |d\phi|^{-2}\kappa\,,\\ M_\Theta &\equiv M + |d\phi|^{-2}\mathscr{L}(\pi_\phi^*\Theta^*)\,. \end{aligned} \tag{79}$$

Thus, Eq. (52) admits the first-order evolution form (41) with

$$T \equiv \begin{pmatrix} 0 & \mathrm{Id} \\ M_{\kappa,\Theta} & N_\Theta \end{pmatrix}. \tag{80}$$

From Propositions 1 and 2 and Lemma 1, it is easy to see that *with respect to the measure* η_μ, *the formal adjoints of the operators* N_Θ *and* $M_{\kappa,\Theta}$ *are given by*

$$\begin{aligned} \mathrm{ad}_\mu N_\Theta &= -N_\Theta + 2|d\phi|^{-2}\mathscr{L}(Y_\mu^*)\phi \\ \mathrm{ad}_\mu M_{\kappa,\Theta} &= M_{\kappa,\Theta} - |d\phi|^{-2}(2\mathscr{L}(\pi_\phi^* Y_\mu^*) \\ &\quad - (1/\mu)\delta(\mu\pi_\phi^* Y_\mu^*))\,, \end{aligned} \tag{81}$$

where

$$Y_\mu^* \equiv Z_\mu^* + \Theta^* \tag{82}$$

In the Klein–Gordon case, we associate with the operator M_κ the self-adjoint operator M_E by Eq. (45), namely,

$$M_E \equiv M_\kappa + |d\phi|^{-2}\mathscr{L}(\pi_\phi^* Z_\mu^*)\,. \tag{83}$$

Analogously, we can now associate with the operator $M_{\kappa,\Theta}$ the self-adjoint operator

$$M_{\kappa,\Theta} - |d\phi|^{-2}\mathscr{L}(\pi_\phi^* Y_\mu^*)\,.$$

Nevertheless, as it is easy to see, this operator is nothing but the original M_E given by Eq. (83). *We are thus led to consider, in the general case, the same energy bilinear form (49)*:

$$2E(X,X') = \langle v, v'\rangle_\mu - \langle M_E u, u'\rangle_\mu\,, \tag{84}$$

which, as in the Klein–Gordon case, is definite iff $\kappa > 0$.

One can then prove that, *for the operator* T *to be skew-self-adjoint with respect to the energy bilinear form (84), it is necessary and sufficient that the operators* M_Θ *and* N_Θ *be respectively self-adjoint and skew-self-adjoint with respect to the measure* η_μ *or, equivalently, that the 1-form* Y_μ *be zero. In such case, the operators* M_E *and* $M_{\kappa,\Theta}$ *coincide.*

(b) As in the Klein–Gordon case, the function κ has no influence on the skew-self-adjointness of T, but now this property is related not only to the Laplacian structure of V_{n+1} but also to the extended structure (V_{n+1}, g, Θ^*).

Thus, a Lorentzian-vectorial structure (V_{n+1}, g, Θ^*) will represent, from our point of view, a set of second-order hyperbolic linear partial differential equations depending on an arbitrary function κ. By Eqs. (79) and (80), with each system of synchronized observers, the structure (V_{n+1}, g, Θ^*) associates the set of operators T characterizing the first-order evolution form of every equation of the former set.

In (V_{n+1}, g, Θ^*), a system of synchronized observers $\{\xi^*;\phi\}$ will be called an *adjoint system* if it can be endowed with a measure η_μ such that the operators T are skew-self-adjoint with respect to the energy bilinear form (84).

By analyzing the equation $Y_\mu^* = 0$, one obtains the analog to Theorem 2: *For a system of synchronized observers* $\{\xi^*;\phi\}$ *to be adjoint with respect to the structure* (V_{n+1}, g, Θ^*), *it is necessary and sufficient that*

$$i(\xi^*)\Lambda = 0\,, \quad d\phi \wedge d\Lambda = 0\,, \tag{85}$$

where the 1-form Λ *is given by*

$$\Lambda \equiv i(d^*\phi)L(\xi) + \Theta\,. \tag{86}$$

We see that, as in the Klein–Gordon case, the measure η_μ does not appear in the determination of the adjoint systems of synchronized observers.

Proposition 6 concerning the pure timelike transformations admitted by adjoint systems remains valid in the present case, but the existence of orthogonal adjoint systems depends upon the particular choice of Θ^*: They are now given by a function ϕ with timelike gradient and such that $\Delta_{n+1}\phi + \mathscr{L}(\Theta^*)\phi = 0$ and $d\Theta \wedge d\phi = 0$.

(c) Finally, it is interesting to remark that iff the 1-form Θ is closed, i.e., $\Theta = d\ln\theta^2$, there always exists, at least locally, a transformation $u \to \hat{u} \equiv \lambda u$, $g \to \hat{g} \equiv (\lambda\theta)^{4/(1-n)}\cdot g$ that reduces, for suitable λ, the general equation (77) to a Klein–Gordon-like equation

$$\hat{\Delta}_{n+1}\hat{u} - \hat{\kappa}\hat{u} = 0$$

with $\hat{\kappa}(x) > 0$. For all these cases, the definite character of the energy bilinear form is insured, and the skew-self-adjointness of T depends only on the metric structure $(V_{n+1}, \hat{g})$.

[1]A. Lichnerowicz, in *Relativity, Groups and Topology Les Houches, 1963*, edited by B. DeWitt and C. DeWitt (Sciences Pub., 1964).

[2]I. Segal, Les variétés des solutions d'une équation hyperbolique non linéaire d'ordre deux, Cours Phys. Math., Collège de France, 1964–1965.

[3]C. Moreno, Rep. Math. Phys. **17**, 171–96 (1980).

[4]C. Moreno, J. Math. Phys. **18**, 2153–61 (1977).

[5]M. Chevalier, J. Math. Pures Appl. **53**, 223–250 (1974).

[6]E. Combet, C. R. Acad. Sci. Paris **262**, 1077–80 (1965).

[7]S. M. Paneitz and I. E. Segal, Proc. Natl. Acad. Sci. USA **77** (12), 6943–7 (1980).

[8]B. Coll and C. Moreno, C. R. Acad. Sci. Paris **292**, I-303–6 (1981).

[9]Covariant and contravariant tensors related by the metric g will be represented by the same symbol. Nevertheless, when a distinction between them will seem convenient, we shall use the superscript * for the contravariant tensors.

[10]In fact, ϕ must be such that $\mathscr{L}(\xi^*)\phi = F(\phi)$, where $F(\phi)$ denotes a constant function on every hypersurface $\phi(x) =$ const. But there always exists a reparametrization of the family such that relation (2) holds.

[11]A local chart $\{x^\alpha\}$ is said to be *physically admissible* if one of the congruences of curves $x^\alpha = \lambda$ is timelike, all the others being spacelike.

[12]The evolution form of the Einstein equations was given, in a Gaussian gauge, by Lichnerowicz ("Sur certains problèmes globaux relatifs au système d'équations d'Einstein," thesis, 1939) and, in an arbitrary gauge, by Choquet-Bruhat [C. R. Acad. Sci. Paris **226**, 1071 (1948)]. In an important work by Arnowitt, Deser, and Misner ["The Dynamics of General Relativity," in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962)], this evolution form was the starting point to obtain the Hamiltonian formalism of general relativity. Due to this fact, the Einstein equations in evolution form are sometimes called ADM equations. We reserve this appellation to situations in which the Hamiltonian formalism plays an important role.

[13]The choice of the spatial metrics $\tilde{g}$ has no influence in the final results, but it allows us to simplify some arguments.

[14]In opposition to what happens with the metric $\tilde{g}$, we will here specify a certain class of spatial volume elements.

[15]According to the notation convention of Ref. 7 $(\mathscr{L}(\xi^*)g)^*$ denotes the contravariant tensor associated with $\mathscr{L}(\xi^*)g$.

[16]It is sufficient that $\alpha,\beta\in C_0^2$ (i.e., C^2 functions with compact support).

[17]I.e., $z^*\in C_0^2$.

[18]*Integrable* in the sense of Theorem 3: Adjoint systems of synchronized observers exist on all Lorentzian structures.

[19]L. Bel, Astrophys. J. **155**, 83 (1969).

[20]I. H. Thompson and G. J. Whitrow, Mon. Not. R. Astron. Soc. **136**, 207–17 (1967).

[21]A. Lichnerowicz, "Propagateurs et commutateurs en Relativité Générale," I. H. E. S., Pub. Math. **10** (1961).

Local S-matrix symmetries in Galilean field theories[a)]

Wolf-Dieter Garber
Institut für Theoretische Physik, Universität Göttingen, Bunsenstrasse 9, D-3400 Göttingen, West Germany

(Received 22 January 1982; accepted for publication 30 April 1982)

A symmetry of the S-matrix S is an operator that commutes with S and is additive, i.e., transforms an incoming n-particle state as a sum of its constituent one-particle states. For field theories with the Galilei group as a symmetry group and with nonvanishing elastic two-particle scattering amplitude, all S-matrix symmetries Q are determined that are local, i.e., transform local asymptotic fields into local ones. For scalar fields, Q is a linear combination of the ten generators of the Galilei group G, internal symmetries, and, possibly, two other generators C and D. By Galilean invariance, the generators of G are S-matrix symmetries in all field theories. In contrast, it depends on the specific field theory if C and D occur. If C is a symmetry, then D is, too, and vice versa. All scattering amplitudes are given that admit C and D as S-matrix symmetries.

PACS numbers: 11.20. – e, 11.30. – j, 11.10. – z

1. INTRODUCTION

The classification of elementary particles by symmetries has a long history. The starting point was the fusion of the isospin and spin groups into SU(6),[1,2] a tentative adaptation of Wigner's successful SU(4) theory of nuclear forces.[3] The coupling of spin and isospin was thought to allow the grouping of particles with different spins into a single supermultiplet. This theory, however, required spin and angular momentum to be separately conserved and therefore could best be understood as a nonrelativistic approximation. It left open the problem of how the SU(6) states should be transformed relativistically or, what is the same, which symmetry group H should be taken that contained P and SU(6). Coleman[4] constructed examples of such groups H which exhibited, however, infinite mass or spin degeneracy. Furthermore, he proved that such unacceptable features should occur, under additional technical assumptions, for any Lie group H which is, locally, not a direct product of the Poincaré group P and some group C of inner symmetries.

To avoid this "no-go" theorem, an infinite parameter group was next suggested.[5,6] The infinitely many conservation laws associated with this group, however, restricted the scattering amplitude so severely that scattering was only allowed in forward or backward directions.[7] This focused attention on scattering in general and the converse problem: If nontrivial scattering is required, which symmetries are still allowed? This question can be asked in any field theory with asymptotic fields $\phi_\alpha^{\text{ex}}(x)$ (where ex stands for either in or out) and corresponding annihilation or creation operators $a^{\text{ex}}(\mathbf{p},\alpha)$, $(a^{\text{ex}}(\mathbf{p},\alpha))^*$. By the term "symmetry" or, more precisely, by "*symmetry of the S-matrix*" is understood an operator that is additive, i.e., bilinear in incoming creation and annihilation operators

$$Q = \int d^3p\, d^3q\, Q^{\alpha\beta}(\mathbf{p},\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta), \tag{1.1}$$

(summation convention), and which commutes with the S-matrix

$$(S^*\psi^{\text{in}}|Q\phi^{\text{in}}) = (Q^*\psi^{\text{in}}|S\phi^{\text{in}}), \tag{1.2}$$

as a form on some suitable domain. In (1.1), $Q^{\alpha\beta}$ is a tempered distribution in $\mathbf{p}$ and $\mathbf{q}$.

For relativistic fields, there is again a no-go theorem for S-matrix symmetries[8]: If the elastic two-particle scattering amplitude is analytic and nonvanishing in the physical domain and if there is a finite mass degeneracy, then any S-matrix symmetry is a linear combination of generators of P and those of some compact group C of inner symmetries. The conclusion is just that of the infinitesimal version of the first no-go theorem and applies to infinite parameter groups as well.

For axiomatic field theories of the Wightman type, however, physical analyticity of the elastic two-particle scattering amplitude has not been derived from the postulates so far. For Wightman field theories, dynamical variables are local fields. Thus it seems natural to consider only *local* S-matrix symmetries, i.e., those Q for which the commutator with a local asymptotic field is again a local field. For local S-matrix symmetries, analogous no-go theorems can be proved without analyticity assumptions:

1. Any local S-matrix symmetry has vanishing matrix elements between states of different mass multiplets. On one-particle states, it acts as a polynomial in momenta and derivatives in momenta.[9–13]

2. The polynomial in 1 can be written as a polynomial in the generators of the Poincaré group P and translation invariant generators.[12,13]

For these statements, no assumption about scattering is required. If, in addition, scattering is assumed to be nontrivial, then the polynomial in 2 must be linear.

3. If the particles can be so ordered (some particles may be counted more than once!) that the elastic two-particle scattering amplitude, for any two consecutive particles, is nonzero in some open subset of the momenta allowed by momentum and energy conservation, then any local S-matrix symmetry is a linear combination of generators of the Poincaré group P and translation invariant generators.

Note that no analyticity assumption has been made. Statement 3 has been proved in Ref. 14 for the case of asymp-

[a)] This work contains parts of the author's "Habilitationsschrift," accepted by the Physics Department, University of Göttingen.

 0022-2488/83/051256-08$02.50

totic *scalar* fields for which the translation invariant generators commute with those of P, i.e., are internal symmetries. Furthermore, one requires the scattering amplitude to be nonzero for any linear combination of creation operators.

In addition, the *ad hoc* assumptions of what should be called an S-matrix symmetry can be justified in a Wightman field theory:

0. Any integral over a conserved, local current density in a Wightman theory can be extended to asymptotic states and is there a symmetry of the S-matrix that is local.

For the proof,[12,15] one requires, apart from the usual assumptions of the Haag–Ruelle scattering theory, the invariance of the vacuum, asymptotic completeness, and, as in Ref. 8, finite multiplicity of mass hyperboloids.

These results clarify the situation for relativistic field theories. It is curious, however, that the original no-go theorem in Ref. 8 does not seem to require the Poincaré group as kinematical invariance group. A result similar to 3 is stated for the Galilei invariant theory of classical point particles in Ref. 16. One can therefore ask if results analogous to 0–3 hold in Galilei invariant theories.

Simple examples that will later be given show that there are operators in Galilei field theories that commute with S and are not additive. It seems, therefore, artificial to look for physical conditions to exclude them. Hence, no proof of a result similar to 0 will be attempted, and S-matrix symmetries will be considered from the start. Now, it is not obvious what results one should expect; the no-go theorems cited so far only applied to relativistic generalizations of a nonrelativistic situation, and the latter escaped those theorems, as was manifest in Wigner's theory. Nevertheless, it will be shown in Sec. 3 that for *local* S-matrix symmetries which transform local asymptotic Galilei fields into local fields, Statement 1 remains valid, with the Galilei group G replacing the Poincaré group P. The analogs of Statements 2 and 3, however, are *wrong*, even if only scalar asymptotic fields are considered for Statement 3, which will always be done in Sec. 4.

To undertand why, consider the double role played by the Poincaré group P in relativistic theories; P is the kinematical invariance group and at the same time the largest group that leaves invariant the differential equation for the free asymptotic fields. For nonrelativistic field theories, the Galilei group G is the kinematical invariance group. The group that leaves invariant the free Schrödinger equation, however, is the 12-parameter Schrödinger group S which contains, in addition to G, two elements D and C reminiscent of a dilation and a conformal inversion.[17,18] Hence, one should expect Statement 3 to be true with P replaced by S, and this will indeed be shown in Sec. 4.

The structure of S-matrix symmetries is thus more complicated in Galilean field theories; any S-matrix symmetry is, apart from some translation invariant symmetries, at most a linear combination of the generators of the Galilei group and, possibly, of C and D, provided scattering is nontrivial. By Galilean invariance, the generators of G are symmetries in any Galilei field theory. It depends, however, on the specific interaction if C or D are symmetries. Section 5 contains the proof that if C leaves the two-particle scattering amplitude invariant, so does D, and vice versa. And this occurs if and only if all phase shifts of the two-particle elastic scattering amplitude are constant. This is the case for the $1/r^2$-potential in quantum mechanics. In contrast, the phase shifts are not constant for any centrally symmetric potential that is less singular locally and drops off faster at infinity.

2. GALILEI INVARIANT FIELD THEORIES

An axiomatic formulation of Galilei invariant field theories has been given in Ref. 19.

The fields of the theory are required to transform under a unitary projective representation of the ten-parameter Galilei group G. Any such representation is unitarily equivalent to a unitary representation of a central extension $\tilde{G}$ of G by, in more than two space dimensions, a single central element M. This operator M is the mass operator; as a central element, it is a superselection rule.[20]

The appearance of the mass superselection rule is the only fact that is more restrictive in Galilean theories than in relativistic ones. For the spectrum condition one should at most require that the spectrum of M is discrete, contains zero as an isolated point, and that the generator of time translations is bounded below in any eigenspace of M. Galilean locality is considerably less restrictive, too, only fields with disjoint support for equal times are required to (anti-) commute.

For Galilean field theories, several of the general theorems known from relativistic theories do *not* hold. In particular, PCT, spin and statistics, and Haag's theorem are not true. Furthermore, the construction of interacting Galilean theories is rather easy; ordinary quantum mechanics in second quantized form is an example which, however, conserves the particle number. But there are also examples which allow particle creation, e.g., a nonrelativistic Lee model.[19,21]

Asymptotic conditions are known mainly for the case of bosons interacting via a pair potential;[22] see however, Ref. 23. For the following, the existence of asymptotic free Galilei fields transforming under the same representation as the interacting fields will simply be assumed. Free Galilei fields are constructed via annihilation and creation operators $a(\mathbf{p},\alpha)$; $a^*(\mathbf{p},\alpha)$ with (anti-) commutation relations

$$[a(\mathbf{p},\alpha),a^*(\mathbf{p},\beta)]_\pm = \delta_{\alpha\beta}\delta(\mathbf{p}-\mathbf{q}). \tag{2.1}$$

The annihilation operators transform as follows:

$$U^{-1}(\mathbf{a})a(\mathbf{p},\alpha)U(\mathbf{a}) = \exp\{-i\mathbf{pa}\}a(\mathbf{p},\alpha), \tag{2.2}$$

$$U^{-1}(b)a(\mathbf{p},\alpha)U(b) = \exp\{ib((2m_\alpha)^{-1}\mathbf{p}^2 + W_\alpha)\}a(\mathbf{p},\alpha), \tag{2.3}$$

$$U^{-1}(\mathbf{v})a(\mathbf{p},\alpha)U(\mathbf{v}) = a(\mathbf{p}-m_\alpha\mathbf{v},\alpha), \tag{2.4}$$

$$U^{-1}(R)a(\mathbf{p},\alpha)U(R) = (D^{s(\alpha)}(R))_{\alpha\beta}a(\mathbf{p},\beta), \tag{2.5}$$

under space translations, time translations, boosts, and rotations, respectively. The projective representation U of the Galilei group G acts irreducibly in the space of one-particle states and is characterized by three constants: the mass m_α, the inner energy W_α, and the spin $s(\alpha)$. Fields in x-space are defined by the Fourier transformation

$$\phi_\alpha(\mathbf{x},t) = (2\pi)^{-3/2}\int d^3p \exp i(\mathbf{px} - E_\alpha(\mathbf{p})t)a(\mathbf{p},\alpha), \tag{2.6}$$

with

$$E_\alpha(\mathbf{p}) = (2m_\alpha)^{-1}\mathbf{p}^2 + W_\alpha. \tag{2.7}$$

The projective nature of the representation (2.2)–(2.5) has an important consequence for any S-matrix symmetry Q of the form (1.1) whether it is local or not.

Lemma 2.1: The kernel $Q^{\alpha\beta}$ of a not necessarily local S-matrix symmetry Q fulfills

$$Q^{\alpha\beta} = 0 \text{ for } m_\alpha = m_\beta.$$

Proof: This is, of course, just the mass superselection rule. Perform, on Q, a translation $\mathbf{a}$, a boost $\mathbf{v}$, a translation $(-\mathbf{a})$, and a boost $(-\mathbf{v})$. Within G, a composition of these four transformations results in the unit element, whereas, by (2.2)–(2.5), $(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta)$ picks up an extra factor $\exp(i\mathbf{av}(m_\alpha - m_\beta))$ so that

$$Q^{\alpha\beta}(\mathbf{p},\mathbf{q}) = \exp(i\mathbf{av}(m_\alpha - m_\beta))Q^{\alpha\beta}(\mathbf{p},\mathbf{q}). \tag{2.8}$$

Differentiation with respect to $\mathbf{av}$ yields the result. □

This lemma makes it possible to consider, from now on, Q on a single mass multiplet m.

3. LOCAL ADDITIVE OPERATORS

In this section, consequences of locality for operators of the form (1.1) will be investigated. It is essential that Q is additive; it will *not* yet be used that Q commutes with S.

An operator Q of the form (1.1) is local, if $[Q,\phi_\alpha^{\text{in}}(\mathbf{x},t)]$ is again a local field. By Galilean locality, the double commutator

$$\begin{aligned}&[[Q,\phi_\alpha^{\text{in}}(\mathbf{x},t)],(\phi_\beta^{\text{in}}(\mathbf{y},t))^*]_\pm \\ &\quad = (2\pi)^{-3}\int d^3p\, d^3q \exp\{-it(E_\alpha(\mathbf{p}) - E_\beta(\mathbf{q}))\} \\ &\qquad \times \exp i(\mathbf{px} - \mathbf{qy})Q^{\alpha\beta}(\mathbf{p},\mathbf{q})\end{aligned} \tag{3.1}$$

can only have support for $\mathbf{x} = \mathbf{y}$,

$$\begin{aligned}&[[Q,\phi_\alpha^{\text{in}}(\mathbf{x},t)],(\phi_\beta^{\text{in}}(\mathbf{y},t))^*]_\pm \\ &\quad = c^{\alpha\beta}_{n_1n_2n_3}(\mathbf{y},t)(\partial_1)^{n_1}(\partial_2)^{n_2}(\partial_3)^{n_3}\delta(\mathbf{x} - \mathbf{y}).\end{aligned} \tag{3.2}$$

Fourier-transforming (3.1) and (3.2) in $\mathbf{x}$ and $\mathbf{y}$ gives

$$\begin{aligned}&\exp\{-it(E_\alpha(\mathbf{p}) - E_\beta(\mathbf{q}))Q^{\alpha\beta}(\mathbf{p},\mathbf{q})\} \\ &\quad = \tilde{c}^{\alpha\beta}_{n_1n_2n_3}(\mathbf{p} - \mathbf{q},t)\prod_{k=1}^{3}(ip_k)^{n_k}.\end{aligned} \tag{3.3}$$

Now use $\mathbf{r} := \mathbf{p} - \mathbf{q}$ and $\mathbf{p}$ as variables and put

$$\hat{Q}^{\alpha\beta}(\mathbf{p} - \mathbf{q},\mathbf{p}) := Q^{\alpha\beta}(\mathbf{p},\mathbf{q}). \tag{3.4}$$

Then, (3.3) reads

$$\exp(-itf(\mathbf{r},\mathbf{p}))\hat{Q}^{\alpha\beta}(\mathbf{r},\mathbf{p}) = \tilde{c}^{\alpha\beta}_{n_1n_2n_3}(\mathbf{r},t)\prod_{k=1}^{3}(ip_k)^{n_k}, \tag{3.5}$$

where, by (2.7),

$$\begin{aligned}f(\mathbf{r},\mathbf{p}) &:= E_\alpha(\mathbf{p}) - E_\beta(\mathbf{p} - \mathbf{r}) \\ &= m^{-1}\mathbf{pr} + (2m)^{-1}\mathbf{r}^2 + W_\alpha - W_\beta\end{aligned} \tag{3.6}$$

since $m_\alpha = m_\beta = m$. This implies

Theorem 3.1: Let Q be an operator of the form (1.1). Then Q is local if and only if

$$Q^{\alpha\beta} = q^{\alpha\beta}\delta(\mathbf{p} - \mathbf{q}), \tag{3.7}$$

where $q^{\alpha\beta}$ is a polynomial in $\mathbf{p}$ and p_k derivatives.

Proof: Assume first that Q is local so that (3.5) holds. Since the right-hand side of (3.5) is a polynomial in $\mathbf{p}$, there exist for each $k = 1,2,3$, natural numbers $N_k \geqslant 1$ so that

$$(\partial/\partial p_k)^{N_k}\{\exp(-itf(\mathbf{r},\mathbf{p}))\hat{Q}^{\alpha\beta}(\mathbf{r},\mathbf{p})\} = 0.$$

If this equation is multiplied by $\exp\{itf(\mathbf{r},\mathbf{p})\}$, a polynomial in t results whose coefficients must vanish. In particular, the coefficient of the highest power in t is

$$\{(\partial/\partial p_k)f(\mathbf{r},\mathbf{p})\}^{N_k}\hat{Q}^{\alpha\beta}(\mathbf{r},\mathbf{p}) = (m^{-1}r_k)^{N_k}\hat{Q}^{\alpha\beta}(\mathbf{r},\mathbf{p}) = 0.$$

This shows that $\hat{Q}^{\alpha\beta}$ has at most support for $\mathbf{r} = 0$. Put $t = 0$ in (3.5) to see that $\hat{Q}^{\alpha\beta}$ is a polynomial in $\mathbf{p}$. By (3.4), this implies (3.7). Conversely, (3.7) implies (3.2). □

Thus, Statement 1 of the introduction is valid for local S-matrix symmetries in Galilei invariant theories, too.

4. *S*-MATRIX SYMMETRIES

By Theorem 3.1, any local additive operator, in particular any local S-matrix symmetry, has the form

$$Q = \sum_n \int d^3p\, d^3q\, q^{\alpha\beta}_{i_1\cdots i_n}(\mathbf{p})\partial^{i_1}\cdots\partial^{i_n}\delta(\mathbf{p} - \mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta), \tag{4.1}$$

with suitable polynomials q, where the sum over n is finite. In this section, consequences of (1.2), i.e., the fact that Q commutes with S, will be investigated. In addition, it will be assumed that all asymptotic fields are scalar. All operators will be considered on the asymptotic spaces $\mathscr{H}^{\text{in}} = \mathscr{H}^{\text{out}}$. Note that Q is defined on D^{ex}, the asymptotic states with finite particle number and test functions in $\mathscr{S}$.

If Q commutes with S, so does Q^*; hence Q can be assumed to be Hermitian. For two particles ρ and τ then,

$$\begin{aligned}&((a^{\text{in}}(\mathbf{p}_1,\rho))^*(a^{\text{in}}(\mathbf{p}_2,\tau))^*\Omega\,|Q(a^{\text{out}}(\mathbf{p}_3,\rho))^*(a^{\text{out}}(\mathbf{p}_4,\tau))^*\Omega) \\ &\quad = (Q(a^{\text{in}}(\mathbf{p}_1,\rho))^*(a^{\text{in}}(\mathbf{p}_2,\tau))^*\Omega\,|(a^{\text{out}}(\mathbf{p}_3,\rho))^*(a^{\text{out}}(\mathbf{p}_4,\tau))^*\Omega),\end{aligned} \tag{4.2}$$

and (1.2) and (4.1) imply

$$\begin{aligned}&q_i^{\alpha\rho}(\mathbf{p}_3)\partial^i_{\mathbf{p}_3}S_{\rho\tau\alpha\tau}(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4) + q_i^{\alpha\tau}(\mathbf{p}_4)\partial^i_{\mathbf{p}_4}S_{\rho\tau\rho\alpha}(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4) \\ &\quad = q_i^{\alpha\rho}(\mathbf{p}_1)\partial^i_{\mathbf{p}_1}S_{\alpha\tau\rho\tau}(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4) + q_i^{\alpha\tau}(\mathbf{p}_2)\partial^i_{\mathbf{p}_2}S_{\rho\alpha\rho\tau}(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4).\end{aligned} \tag{4.3}$$

Here, $i = (i_1,\dots i_n)$, summation in (4.3) is over i_ν, n, and α, and the differential operators carry, as an index, the argument on which they act. As an abbreviation in (4.3),

$$S_{\alpha\beta\gamma\delta}(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3,\mathbf{p}_4) := ((a^{\text{in}}(\mathbf{p}_1,\alpha))^*(a^{\text{in}}(\mathbf{p}_2,\beta))^*\Omega\,|(a^{\text{out}}(\mathbf{p}_3,\gamma))^*(a^{\text{out}}(\mathbf{p}_4,\delta))^*\Omega). \tag{4.4}$$

The difficulty in the following proofs will always be that, though the whole sum (4.1) commutes with S, it is not known *a priori* that single terms in (4.1) do. The following lemmas will serve to isolate terms in (4.1) that commute with S.

Lemma 4.1: (i) Let Q_α be operators fulfilling (4.2), and assume that Q_α converges weakly to Q with Q defined on D^{ex}. Then Q fulfills (4.2).

(ii) Let Q be an operator fulfilling (4.2). Then any finite or infinitesimal Galilei transform of Q fulfills (4.2).

Proof: (i) is just the definition of weak convergence on D^{ex}. (ii): Insert $U^{-1}(g)QU(g)$, for any Galilei group element g, into the left-hand side of (4.2). By the unitarity of U, this is then a matrix element of Q with Galilei transformed two-particle states for which (4.2) is true. Differentiation with respect to group parameters gives the statement for infinitesimal Galilei transformations, too. □

If Q can be split up into parts that transform independently under rotations, then those parts commute separately with S.

Lemma 4.2: Let Q be an operator fulfilling (4.2) with the representation

$$Q=\sum_{l,m} a_{lm}Q_{lm},\quad a_{lm}\in\mathbb{C}$$

where $\{Q_{lm}/m=-l,...,l\}$ forms a basis of a space of operators that transform irreducibly under the representation D^l of the rotation group. Then $a_{lm}Q_{lm}$ fulfills (4.2).

Proof: Under rotations,

$$U^{-1}(R)QU(R)=\sum_{l,m,n} a_{lm}(D^l(R))_{mn}Q_{ln}. \tag{4.5}$$

Multiply (4.5) with $(D^\lambda(R))^*_{\mu\nu}$ and integrate over the rotation group with the Haar measure dR. By the orthogonality relations for the representation matrices

$$\begin{aligned}&\int dR(D^\lambda(R))^*_{\mu\nu}U^{-1}(R)QU(R)\\&=\sum_{l,m,n}\left\{\int dR(D^\lambda(R))^*_{\mu\nu}(D^l_{mn}(R))\right\}a_{lm}Q_{ln}\\&=\sum_{l,m,n}\{(2l+1)^{-1}\delta_{\mu m}\delta_{\nu n}\delta_{\lambda l}\}a_{lm}Q_{ln}\\&=(2\lambda+1)^{-1}a_{\lambda\mu}Q_{\lambda\nu}.\end{aligned}\tag{4.6}$$

The left-hand side is the weak limit of Galilei-transformed operators fulfilling (4.2) and, hence, fulfills (4.2) by Lemma 4.1. For $\mu=\nu$, the assertion follows. □

In applications, some constants a_{lm} may be zero.

Corollary 4.3: Let Q be an operator fulfilling (4.2) with the representation

$$Q=\sum_{i=1}^{N} Q^{l(i)},$$

where the operators $Q^{l(i)}$ transform under the representations $D^{l(i)}$ of the rotation group that are pairwise inequivalent. Then $Q^{l(i)}$ fulfills (4.2).

Call, in (4.1), the largest natural number n for which some polynomial $q^{\alpha\beta}_{i_1\cdots i_n}(\mathbf{p})$ is not identically zero, the degree of Q. Symmetries of degree zero will be analyzed first.

A. Symmetries of degree zero

They are of the form

$$Q=\int d^3p\, d^3q\, q^{\alpha\beta}(\mathbf{p})\delta(\mathbf{p}-\mathbf{q})(a^{\rm in}(\mathbf{p},\alpha))^*a^{\rm in}(\mathbf{q},\beta), \tag{4.7}$$

and, by (2.2), invariant under translations. Consider the special case

$$q^{\alpha\beta}=q^{\alpha\beta}r(\mathbf{p}), \tag{4.8}$$

with a constant matrix $q^{\alpha\beta}$. Since Q is Hermitian, $q^{\alpha\beta}$ is a Hermitian matrix that can be diagonalized by a unitary matrix A

$$A_{\alpha\beta}q^{\beta\gamma}(A^{-1})_{\gamma\epsilon}=\delta_{\alpha\epsilon}q^\alpha;\quad q^\alpha=\overline{q^\alpha}.$$

Put

$$b^{\rm in}(\mathbf{p},\gamma):=A_{\alpha\gamma}a^{\rm in}(\mathbf{p},\alpha).$$

Then $b^{\rm in},(b^{\rm in})^*$ are linear combinations of annihilation and creation operators, and (4.7) reduces to

$$Q=\int d^3p\, d^3q\, q^\alpha r(\mathbf{p})(b^{\rm in}(\mathbf{p},\alpha))^*b^{\rm in}(\mathbf{p},\alpha),$$

so that (4.2) gives

$$\{q^\rho r(\mathbf{p}_1)+q^\tau r(\mathbf{p}_2)-q^\rho r(\mathbf{p}_3)-q^\tau r(\mathbf{p}_4)\}S_{\rho\tau\rho\tau}(\mathbf{p}_1,\mathbf{p}_2;\mathbf{p}_3,\mathbf{p}_4)=0.$$

By the assumption about the nontriviality of scattering (statement 3 of the Introduction), $S_{\rho\tau\rho\tau}\neq0$ in some open set $V_{\rho\tau}$ so that

$$q^\rho r(\mathbf{p}_1)+q^\tau r(\mathbf{p}_2)-q^\rho r(\mathbf{p}_3)-q^\tau r(\mathbf{p}_4)=0, \tag{4.9}$$

in $V_{\rho\tau}$. In fact, (4.9) is true on the whole scattering manifold, i.e., the set of momenta allowed by energy and momentum conservation

$$\mathbf{p}_1^2+\mathbf{p}_2^2=\mathbf{p}_3^2+\mathbf{p}_4^2, \tag{4.10}$$

$$\mathbf{p}_1+\mathbf{p}_2=\mathbf{p}_3+\mathbf{p}_4. \tag{4.11}$$

Consider (4.9) first in the *lab* system $\mathbf{p}_2=0$ which, for $\mathbf{p}_1\neq0$, can be analytically parametrized, see Sec. 17 in Ref. 24. Hence, the left-hand side of (4.9) is real analytic in those parameters so that (4.9) holds in the whole *lab* system and, by Galilei covariance, on the scattering manifold. Now the following lemma has been proven in Ref. 14:

Lemma 4.4: Let $f_i(\mathbf{p})$, $i=1,2$ be two continuously differentiable functions such that

$$f_1(\mathbf{p}_1)+f_2(\mathbf{p}_2)=f_1(\mathbf{p}_3)+f_2(\mathbf{p}_4),$$

whenever (4.10), (4.11) holds. Then

$$f_i(\mathbf{p})=\mathbf{a}\mathbf{p}+b\mathbf{p}^2+c_i,\quad i=1,2,$$

with constants a_k,b,c_i.

This lemma implies

$$q^ir(\mathbf{p})=\mathbf{a}^i\mathbf{p}+b^i\mathbf{p}^2+c^i,\quad i=\rho,\tau,$$

with

$$\mathbf{a}^\rho=\mathbf{a}^\tau;\ b^\rho=b^\tau.$$

By statement 3 of the Introduction, the constants $\mathbf{a}$ and b are not only independent of the two particles just considered, but are the same on the whole mass multiplet. In the old basis (4.8),

$$q^{\alpha\beta}r(\mathbf{p})=\delta^{\alpha\beta}(a_k\,p^k+b\mathbf{p}^2)c^{\alpha\beta}, \tag{4.12}$$

with suitable constants $a_k,b,c^{\alpha\beta}$. This result will now be generalized to arbitrary polynomials $q^{\alpha\beta}(\mathbf{p})$ with the help of Lemma 4.2.

Note first that the polynomial $q(\mathbf{p})$ can be written

$$q(\mathbf{p}) = \sum_{l,m} d_{lm}(\mathbf{p}^2) y_{lm}(\mathbf{p}), \tag{4.13}$$

with suitable polynomials d_{lm} in one variable. In (4.13)

$$y_{lm}(\mathbf{p}) := (\mathbf{p}^2)^{1/2} Y_{lm},$$

where Y_{lm} are the familiar spherical harmonics. Hence, y_{lm} are homogeneous polynomials of degree l that are harmonic, i.e., fulfill $\Delta y_{lm} = 0$. In fact, they form a basis in the space of all harmonic polynomials, (see p. 1270 ff. in Ref. 25). One can prove (4.13) by induction in the degree of q: For constant or linear polynomials, one can obviously choose d_{00} or d_{1m} to be constant. Assume (4.13) to be true for polynomials of degree N, and let q be a polynomial of degree $N+2$. Then Δq is a polynomial of degree N so that

$$\Delta q = \sum_{l,m,n} d_{lmn}(p^2) y_{lm}(\mathbf{p}),$$

with constants d_{lmn}, by the induction assumption. Since

$$\Delta \{ (\mathbf{p}^2)^{n+1} y_{lm} \} = a_{nl} (\mathbf{p}^2)^n y_{ln},$$

with constants $a_{nl} \neq 0$, the polynomial

$$\hat{q}(\mathbf{p}) := \sum_{l,m,n} (a_{nl})^{-1} d_{lmn} (\mathbf{p}^2)^{n+1} y_{lm},$$

fulfills

$$\Delta \hat{q} = \Delta q.$$

Hence, $q - \hat{q}$ is a harmonic polynomial which is a linear combination of y_{lm}. Thus, q has again the representation (4.13).

Hence, the polynomials $q^{\alpha\beta}$ have the form

$$q^{\alpha\beta}(\mathbf{p}) = \sum_{l,m} d^{\alpha\beta}_{lm}(\mathbf{p}^2) y_{lm}(\mathbf{p}) = \sum_{l,m} \left(\sum_{n=0}^{N} d^{\alpha\beta}_{lmn} (\mathbf{p}^2)^n \right) y_{lm}(\mathbf{p}). \tag{4.14}$$

Now, the spherical harmonics form a basis of the subspaces invariant under the rotation group. By Lemma 4.2, $d^{\alpha\beta}_{lm}(\mathbf{p}^2) y_{lm}$ fulfills (4.2) separately. By (4.14), this is still not a monomial of the form (4.8) but it can be reduced to one by the application of boosts.

For an infinitesimal boost in the 3-direction, (2.4) implies

$$\begin{aligned} \partial_3 q^{\alpha\beta} &= \partial_3 (d^{\alpha\beta}_{lm}(\mathbf{p}^2) y_{lm}) \\ &= (d^{\alpha\beta}_{lm}(\mathbf{p}^2))' \cdot 2p_3 y_{lm} + d^{\alpha\beta}_{lm}(\mathbf{p}^2) \partial_3 y_{lm}. \end{aligned} \tag{4.15}$$

Now, $\partial_3 y_{lm}$ is of degree $l-1$ and, hence, a linear combination of $y_{l-1,n}$. By 8.5.3 in Ref. 26, $p_3 y_{lm}$ is a linear combination of $y_{l+1,m}$ and $\mathbf{p}^2 y_{l-1,m}$. By Lemma 4.3, that part of (4.15) which transforms according to D^{l+1}, i.e., $(d^{\alpha\beta}_{lm}(\mathbf{p}^2))' \cdot y_{l+1,m}$ fulfills (4.2) separately. And to this polynomial, the preceding argument can be applied again. By induction, $d^{\alpha\beta}_{lmN} \cdot y_{l+N,m}(\mathbf{p})$ fulfills (4.2). But this is of the form (4.8) so that (4.12) implies

$$d^{\alpha\beta}_{lmN}\, y_{l+N,m}(\mathbf{p}) = \delta^{\alpha\beta}(a_{lmNk}\, p^k + b_{lmN} \mathbf{p}^2) + c^{\alpha\beta}_{lmN} \tag{4.16}$$

with suitable constants $a_{lmNk}, b_{lmN}, c^{\alpha\beta}_{lmN}$. Now, the left-hand side of (4.16) is harmonic whereas the Laplacian, applied to the right-hand side, gives a multiple of b_{lmN}; hence, $b_{lmN} = 0$. Thus, (4.16) is linear which means $l + N \leqslant 1$. Hence, (4.14) contains only terms with $l = 1$ and $l = 0$ that fulfill (4.2) separately. For $l = 1$, $N = 0$ results so that this term is already a monomial of the form (4.8) and hence (4.12). For $l = 0$, (4.14) contains only the terms $(d^{\alpha\beta}_{001} \mathbf{p}^2 + d^{\alpha\beta}_{000}) y_{00}(\mathbf{p})$. But y_{00} is constant, and $d^{\alpha\beta}_{001} = \delta^{\alpha\beta} b_{001}$ by (4.16). Altogether, the polynomial $q^{\alpha\beta}$ fulfills (4.12) again.

Theorem 4.5: Let Q be a S-matrix symmetry of degree zero. In a theory with nontrival scattering,

$$q^{\alpha\beta} = \delta^{\alpha\beta}(a_k\, p^k + b\mathbf{p}^2) + c^{\alpha\beta}. \tag{4.17}$$

Any symmetry of degree zero is thus a linear combination of energy, momentum, and translation-invariant generators with kernels that are constant matrices, cf. (2.2) and (2.3).

However, unlike the situation in relativistic theories, the symmetries given by constant matrices need not be inner ones. For asymptotic particles with spin, Wigner's theory[3] that couples spin with isospin provides a counterexample. Even for scalar asymptotic particles, however, there are symmetries arising from constant matrices that do not commute with the Galilei group G. To see this, start from a theory of asymptotic particles with a symmetry group that *is* a direct product of G and C, with C a compact group of inner symmetries C_i,

$$C_i = \int d^3p\, d^3q\, c^{\alpha\beta}_i\, \delta(\mathbf{p} - \mathbf{q}) (a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta).$$

Assume that C is non-Abelian so that there are two symmetries C_k and C_l in C with

$$[C_k, C_l] \neq 0.$$

Then consider a new theory with all Galilei generators unchanged except the energy H which is replaced by

$$H_k := H + C_k.$$

If C_k is already in diagonal form, this corresponds to a change of inner energies and leads to a new theory of asymptotic free particles that is, in particular, Galilei invariant. However, by construction, $[C_l, C_k] \neq 0$ so that C_l is not an inner symmetry.

B. Symmetries of degree one

They have the form

$$\begin{aligned} Q = \int d^3p\, d^3q \{ q^{\alpha\beta}_i(\mathbf{p}) \partial^i + q^{\alpha\beta}(\mathbf{p}) \} \\ \times \delta(\mathbf{p} - \mathbf{q}) (a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta). \end{aligned} \tag{4.18}$$

By (2.2), an infinitesimal translation results in

$$[iQ, P_k] = \int d^3p\, d^3q\, q^{\alpha\beta}_k(\mathbf{p}) \delta(\mathbf{p} - \mathbf{q}) (a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta),$$

which is, for any $k = 1,2,3$, a symmetry of degree zero. By Theorem 4.5,

$$q^{\alpha\beta}_k(\mathbf{p}) = \delta^{\alpha\beta}(a_{kl}\, p^l + b_k \mathbf{p}^2) + c^{\alpha\beta}_k. \tag{4.19}$$

A double boost in directions l and m of (4.18) gives, by (2.4),

$$\begin{aligned} i^2 [[Q, K_l], K_m] = m^2 \int d^3p\, d^3q \{ \delta^{\alpha\beta} \delta_{lm} 2b_i \partial^i + \partial_l \partial_m q^{\alpha\beta} \} \\ \times \delta(\mathbf{p} - \mathbf{q}) (a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta), \end{aligned} \tag{4.20}$$

The first term on the right-hand side of (4.20) is a linear combination of Galilei boost generators K_i and, hence, fulfills (4.2). Thus, the second term fulfills (4.2), too, and is a symmetry of degree zero

$$\partial_l \partial_m q^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(a_{lmn}\, p^n + b_{lm}\mathbf{p}^2) + c_{lm}^{\alpha\beta}. \tag{4.21}$$

The polynomials $q^{\alpha\beta}$ are at most of degree four. In fact, to any polynomial $q_m^{\alpha\beta}(\mathbf{p})\colon = \partial_m q^{\alpha\beta}(\mathbf{p})$, the following lemma can be applied:

Lemma 4.6: Let $q^{\alpha\beta}$ be a polynomial of degree at most three of the form

$$\partial_l q^{\alpha\beta}(p) = \delta^{\alpha\beta}(a_{ln}\, p^n + b_l \mathbf{p}^2) + c_l^{\alpha\beta}, \tag{4.22}$$

with constants $a_{ln}, b_l, c_l^{\alpha\beta}$. Then $q^{\alpha\beta}$ is, in fact, at most of degree two, and

$$q^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(\hat{a}_{ik}\, p^i p^k) + b_i^{\alpha\beta} p^i + c^{\alpha\beta}. \tag{4.23}$$

Proof: As $q^{\alpha\beta}$ is of degree at most three,

$$q^{\alpha\beta}(\mathbf{p}) = a_{ikl}^{\alpha\beta}\, p^i p^k p^l + b_{ik}^{\alpha\beta} p^i p^k + c_i^{\alpha\beta} p^i + d^{\alpha\beta}.$$

The coefficients a_{ikl}, b_{ik} can be chosen symmetric in i, k, and l and i,k. Hence

$$\partial_l q^{\alpha\beta}(\mathbf{p}) = 3a_{ikl}^{\alpha\beta}\, p^i p^k + 2b_{il}^{\alpha\beta}\, p^i + c_l^{\alpha\beta}. \tag{4.24}$$

Comparing the quadratic terms in (4.24) and (4.23),

$$3a_{ikl}^{\alpha\beta} = \delta^{\alpha\beta}\delta_{ik} b_l, \tag{4.25}$$

where the right-hand side of (4.25) must be symmetric in i,k,l since $a_{ikl}^{\alpha\beta}$ is. For $\alpha = \beta$,

$$\delta_{ik} b_l = \delta_{lk}\, b_i,$$

so that, for arbitrary l and $i = k \neq l$, $b_l = 0$. Hence, $\partial_l q^{\alpha\beta}$ is linear and $q^{\alpha\beta}$ at most quadratic. Comparing the linear terms in (4.24) and (4.23) yields the assertion. □

By this lemma,

$$\partial_m q^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(a_{ikm}\, p^i p^k) + a_{lm}^{\alpha\beta}\, p^l + c_m^{\alpha\beta}. \tag{4.26}$$

Next, perform an infinitesimal time translation on (4.18). By (2.2),

$$\begin{aligned}[Q,H] = \int d^3p\, d^3q \{ & -q_i^{\alpha\beta} p^i m^{-1} \\ & + q_i^{\alpha\beta}(W_\alpha - W_\beta)\partial^i + q^{\alpha\beta}(W_\alpha - W_\beta)\} \\ & \times \delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta).\end{aligned} \tag{4.27}$$

By (4.19) the term in (4.27) that contains ∂^i has the coefficient

$$q_i^{\alpha\beta}(W_\alpha - W_\beta) = c_i^{\alpha\beta}(W_\alpha - W_\beta)$$

with constant $c_i^{\alpha\beta}$ and thus drops out, if (4.27) is boosted:

$$\begin{aligned} & i[[Q,H],K_l] \\ & = \int d^3p\, d^3q \{ -\partial_l(q_i^{\alpha\beta} p^i) + m\partial_l q^{\alpha\beta}(W_\alpha - W_\beta)\}. \\ & \quad \times \delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta).\end{aligned} \tag{4.28}$$

This is a symmetry of degree zero with polynomial

$$\begin{aligned} & -\partial_l(q_i^{\alpha\beta} p^i) + m\partial_l q^{\alpha\beta}(W_\alpha - W_\beta) \\ & \quad = -\delta^{\alpha\beta}\{a_{il} + a_{li})p^i + (2b_i\, p^i p^l + b_l \mathbf{p}^2)\} \\ & \qquad - c_l^{\alpha\beta} + m\partial_l q^{\alpha\beta}(W_\alpha - W_\beta),\end{aligned} \tag{4.29}$$

which must be of the form given in Theorem 4.5. Now, by Lemma 4.6, the last term in (4.29) is linear so that the only terms quadratic in $\mathbf{p}$ are $(2b_i\, p^i p^l + b_l \mathbf{p}^2)$. Because of the first summand, this is of the form (4.17) only for $b_i = 0$.

With $b_i = 0$, one can now show that $q^{\alpha\beta}$ is, in fact, quadratic in $\mathbf{p}$. A single boost in (4.18) gives with $b_i = 0$

$$\begin{aligned} i[Q,K_l] = a_{il} K^i + \int & d^3p\, d^3q (\partial_l q^{\alpha\beta}) \\ & \times \delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta)\end{aligned} \tag{4.30}$$

so that the second term is a symmetry of degree zero,

$$\partial_l q^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(\hat{a}_{ln}\, p^n + \hat{b}_l \mathbf{p}^2) + \hat{c}_l^{\alpha\beta},$$

by Theorem 4.5. Lemma 4.6 now implies

$$q^{\alpha\beta}(\mathbf{p}) = (\delta^{\alpha\beta}\check{a}_{ik}\, p^i p^k) + a_i^{\alpha\beta} p^i + c^{\alpha\beta}. \tag{4.31}$$

Insert (4.31) and (4.19) into (4.27):

$$\begin{aligned}[Q,H] = \int d^3p\, d^3q \{ & -(\delta^{\alpha\beta} a_{ik}\, p^i p^k) m^{-1} - c_i^{\alpha\beta} p^i m^{-1} \\ & + c_i^{\alpha\beta}(W_\alpha - W_\beta)\partial^i \\ & + a_i^{\alpha\beta} p^i (W_\alpha - W_\beta) + c^{\alpha\beta}(W_\alpha - W_\beta)\} \\ & \times \delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^*(\mathbf{q},\beta)\,\cdot\end{aligned} \tag{4.32}$$

Next, an angular momentum analysis will be carried out. The three middle terms in (4.32) belong to $l = 1$, and the representation D^1 does not occur in the first and last terms. By Lemma 4.3, the first and last terms fulfill (4.2), i.e., are a symmetry of degree zero so that, by Theorem 4.5, $a_{ik} = -a_{ki}$ for $i \neq k$. Furthermore, a_{ii} is independent of i and is, say, $b/3$ for a suitable b. Hence,

$$\begin{aligned} q_i^{\alpha\beta}(\mathbf{p})\partial^i + q^{\alpha\beta} = & \tfrac{1}{2}\delta^{\alpha\beta} a_{ik}(p^i\partial^k - p^k\partial^i) \\ & + \delta^{\alpha\beta} b p^i \partial^i + c_i^{\alpha\beta}\partial^i \\ & + \delta^{\alpha\beta}\check{a}_{ik}\, p^i p^k + a_l^{\alpha\beta} p^l + c^{\alpha\beta}.\end{aligned} \tag{4.32'}$$

The first term is, for scalar fields, a linear combination of angular momentum operators, by (2.5), and hence fulfills (4.2); so do the next five terms. Only the summand $\delta^{\alpha\beta} a_{ik}\, p^i p^k$ contains terms that transform under D^2 and hence, by Lemma 4.3, form a symmetry which is then of degree zero and vanishes by Theorem 4.5. Thus, only the D^0 part remains which is $\delta^{\alpha\beta} a\mathbf{p}^2$. Next, only the term $c_i^{\alpha\beta}\partial^i + a_l^{\alpha\beta} p^l$ transforms under D^1. The corresponding S-matrix symmetry transforms under a finite time translation, by (2.3), into a symmetry with kernel

$$\exp(-ib(W_\alpha - W_\beta))\cdot\{m^{-1} ibc_i^{\alpha\beta} p^i + (c_i^{\alpha\beta}\partial^i + a_l^{\alpha\beta} p^l)\}.$$

Fourier transformation of this symmetry in b with a test function f gives, by Lemma 4.1, an S-matrix symmetry with kernel

$$-m^{-1} c_i^{\alpha\beta} p \tilde{f}'(W_\alpha - W_\beta) + \tilde{f}(W_\alpha - W_\beta)(c_i^{\alpha\beta}\partial^i + a_l^{\alpha\beta} p^l).$$

Since $W_\alpha - W_\beta$ runs over a finite number of points only, choose any test function f such that $\tilde{f}$ is zero at those points while $\tilde{f}' = a \neq 0$ there, to see that $c_i^{\alpha\beta} p^i$ is the kernel of an S-matrix symmetry. Thus, by Theorem 4.5, $c_i^{\alpha\beta} = \delta^{\alpha\beta} c_i$ so that, in (4.32'), the third term is a linear combination of boost generators. Hence, $a_l^{\alpha\beta} p^l$ is the kernel of a symmetry, too. It is of degree zero; by Theorem 4.5, $a_l^{\alpha\beta} = \delta^{\alpha\beta} a_l$. In summary,

$$\begin{aligned} q_i^{\alpha\beta}\partial^i + q^{\alpha\beta} = & (\tfrac{1}{2} a_{ik}(p^i\partial^k - p^k\partial^i) + i\check{b}\,[p^i,\partial^i]_+ \\ & + c_i\partial^i + a\mathbf{p}^2 + b_i\, p^i) + \check{c}^{\alpha\beta},\end{aligned} \tag{4.33}$$

which proves

Theorem 4.7. In a theory with nontrivial scattering, any S-matrix symmetry of degree one is a linear combination of angular momentum and boost generators, symmetries of degree zero, and, possibly, the symmetry

$$D: = \frac{i}{2}\int d^3p\, d^3q\, \delta^{\alpha\beta}[\,p^i,\partial^i]_+\delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta) \tag{4.34}$$

of degree one.

This is the first example of an additional symmetry that might occur in Galilean field theories. Additional symmetries will be considered in Sec. 5.

C. Symmetries of degree two

These symmetries have the form

$$Q = \int d^3p\, d^3q\{q_{ik}^{\alpha\beta}(\mathbf{p})\partial^i\partial^k + q_i^{\alpha\beta}(\mathbf{p})\partial^i + q^{\alpha\beta}(\mathbf{p})\} \times\delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta), \tag{4.35}$$

with polynomials $q_{ik}^{\alpha\beta}, q_i^{\alpha\beta}$, and $q^{\alpha\beta}$; $q_{ik}^{\alpha\beta}$ can be chosen symmetric in i,k. By an infinitesimal space translation of Q,

$$i[Q,P_l] = \int d^3p\, d^3q\{2q_{il}^{\alpha\beta}\partial^i + q_l^{\alpha\beta}\} \times\delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta). \tag{4.36}$$

This is a symmetry of degree 1 which, by (4.33), has as polynomials

$$q_{il}^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(a_{ilm}\,p^m + b^l p^i + c_{il}),\quad a_{ilm} = -a_{iml}, \tag{4.37}$$

$$q_l^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}(a_{lm}\,p^m + \check{b}_l\mathbf{p}^2) + c_l^{\alpha\beta}. \tag{4.38}$$

First analyze (4.37), which must be symmetric in i and l; this gives $b_l = 0$. In addition, a_{ilm} is now symmetric in the first and antisymmetric in the last pair of indices. Hence, as a one-dimensional representation of the permutation group [these are known to be completely (anti-) symmetric], $a_{ilm} = 0$. One can also check this easily:

$$a_{ilm} = -a_{iml} = -a_{mil} = a_{mli} = a_{lmi} = -a_{lim} = -a_{ilm}.$$

Thus,

$$q_{ik}^{\alpha\beta}(\mathbf{p}) = \delta^{\alpha\beta}c_{ik},\ c_{ik} = c_{ki}. \tag{4.39}$$

Thus, q_{ik} is constant and hence invariant under boosts so that $i[Q,K_l]$ is a symmetry of degree 1 with kernel

$$(\partial_l q_i^{\alpha\beta})\partial^i + \partial_l q^{\alpha\beta} = \delta^{\alpha\beta}(a_{il} + 2\hat{b}_i\,p^l)\partial^i + \partial_l q^{\alpha\beta}.$$

Comparing this to (4.33) gives $\hat{b}_i = 0$ and

$$\partial_l q^{\alpha\beta} = \delta^{\alpha\beta}(a_{lm}\,p^m + \check{b}_l\mathbf{p}^2) + c_l^{\alpha\beta}$$

so that Lemma 4.6 applies again

$$q^{\alpha\beta} = \delta^{\alpha\beta}(\tilde{a}_{ik}\,p^ip^k) + \tilde{b}_i^{\alpha\beta}p^i + \tilde{c}^{\alpha\beta}. \tag{4.40}$$

Next, because of the factor $\delta^{\alpha\beta}$ in (4.39), an infinitesimal time translation gives a symmetry of degree one with kernel

$$m^{-1}\{\delta^{\alpha\beta}(2c_{il}\,p^i\partial^l + a_{il}\,p^ip^l) + c_l^{\alpha\beta}p^l\} + (W_\alpha - W_\beta)(c_l^{\alpha\beta}\partial^l + \tilde{b}_i^{\alpha\beta}p^i + \tilde{c}^{\alpha\beta}). \tag{4.41}$$

Compare (4.41) to (4.34) and remember that c_{il} is symmetric in i and l

$$c_{il} = c\delta_{il}. \tag{4.42}$$

Moreover,

$$a_{il} = -a_{li},\quad i\neq l$$

$$a_{ii} = \tfrac{1}{3}b.$$

With this, the kernel of Q is

$$\delta^{\alpha\beta}(c\partial^i\partial^i + \tfrac{1}{2}a_{ik}(\,p^i\partial^k - p^k\partial^i) + bp^i\partial^i) + c_i^{\alpha\beta}\partial^i + \delta^{\alpha\beta}\tilde{a}_{ik}\,p^ip^k + \tilde{b}_i^{\alpha\beta}p^i + \tilde{c}^{\alpha\beta}. \tag{4.43}$$

This term differs from the right-hand side of (4.32) only in the additional operator $\delta^{\alpha\beta}c\partial^i\partial^i$ which is a scalar under rotations. Therefore, the D^l-representation analysis following (4.32) remains valid for $l = 2$ and $l = 1$ and leads to the same result (4.33). Hence, the last six terms in (4.43) have the form given in Theorem 4.7. This proves

Theorem 4.8: In a theory with non-trivial scattering, any S-matrix symmetry of degree 2 is a linear combination of

$$C: = \frac{m}{2}\int d^3p\, d^3q\, \delta^{\alpha\beta}\delta_{ik}\partial^i\partial^k\delta(\mathbf{p}-\mathbf{q})(a^{\text{in}}(\mathbf{p},\alpha))^* a^{\text{in}}(\mathbf{q},\beta), \tag{4.44}$$

and symmetries of degree one.

D. Symmetries of higher degree

Fortunately, they do not occur. To see this, consider a symmetry Q of degree 3 with kernel

$$q_{ikl}^{\alpha\beta}(\mathbf{p})\partial^i\partial^k\partial^l + \text{terms of lower degree},$$

where $q_{ikl}^{\alpha\beta}$ is symmetric in i,k,l. Note that the symmetry $i[Q,P_m]$ is of degree 2; its kernel is proportional to $q_{ikm}^{\alpha\beta}(\mathbf{p})\partial^i\partial^k$ + terms of lower degree. By Theorem 4.8, $q_{ikm}^{\alpha\beta} = c_m\delta_{ik}\delta^{\alpha\beta}$ which is symmetric in i,k,m only for $c_m = 0$. Hence, Q is, in fact, of degree 2. Similarly, there are no S-matrix symmetries of higher degree. Altogether, this proves

Theorem 4.9: Assume a theory with asymptotic scalar particles which can be so ordered that the elastic two particle scattering amplitude, for any two consecutive particles, is different from zero in an open subset of the scattering manifold (4.10), (4.11). Then any S-matrix symmetry Q of the form (4.1) is a linear combination of the generators of the Galilei group, translation-invariant symmetries with constant kernel, and the generators D and C of (4.34) and (4.44).

5. THE ADDITIONAL GENERATORS D AND C

So far, it has been proved that any local S-matrix symmetry in a theory with nontrivial interaction is, apart from translation-invariant symmetries, necessarily a linear combination of Galilei group generators and the two generators D and C. Of course, in any given theory, some of the coefficients of these symmetries could be forced to be zero by the form of the elastic two-particle scattering amplitude. In such a case, the corresponding symmetries do not occur.

In a Galilean invariant theory, the generators of the Galilei group are always S-matrix symmetries, so they always occur. It is the purpose of this section to investigate the structure of the elastic two-particle scattering amplitudes that allow D or C as S-matrix symmetries.

First note that, by (4.34) and (4.44), D and C together with the Galilei group generate a twelve-parameter Lie group with Lie algebra

$$i[D,K_l] = K_l;\qquad i[D,P_l] = -P_l;\quad i[D,M_{lm}] = 0, \tag{5.1}$$

$$i[C,K_l] = 0;\qquad i[C,P_l] = -K_l;\quad i[C,M_{lm}] = 0, \tag{5.2}$$

$$i[D,H] = -2H; \quad i[D,C] = 2C; \qquad i[C,H] = D, \tag{5.3}$$

where the Galilei generators P_l, H, K_l, M_{lm} of space and time translations, boosts, and rotations obey the usual Galilei Lie algebra relations. This is the Lie algebra of the twelve-parameter group S, the invariance group of the free Schrödinger equation.[18] By the last relation in (5.3), if C commutes with the scattering matrix, so does D, as H always commutes with the scattering matrix in a Galilean invariant theory. Thus, if C is an S-matrix symmetry, so is D. The converse is true at least on the level of the two-particle scattering amplitudes (see theorem 5.1 below).

Next, Galilean invariance restricts the form of the elastic two-particle scattering amplitude. First, invariance under translations and boosts implies

$$S_{\rho\tau\rho\tau}(\mathbf{p}_1,\mathbf{p}_2;\mathbf{p}_3,\mathbf{p}_4) = \delta(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3 - \mathbf{p}_4)\delta(p^2 - q^2)\hat{S}_{\rho\tau\rho\tau}(\mathbf{p},\mathbf{q}), \tag{5.4}$$

where $\mathbf{p}: = \mathbf{p}_1 - \mathbf{p}_2$ and $\mathbf{q}: = \mathbf{p}_3 - \mathbf{p}_4$ are the relative momenta before and after scattering. Now, invariance under rotations, for scalar states, implies that $\hat{S}_{\rho\tau\rho\tau}$ depends only on the invariants $\mathbf{p}^2, \mathbf{q}^2$, and $\mathbf{p}\cdot\mathbf{q}$. The invariant q^2 can be dropped because of the second δ factor in (5.4), and $S_{\rho\tau\rho\tau}$ can be expanded, in the variable $(\mathbf{pq})p^{-1}q^{-1}$, into spherical harmonics[27]

$$\hat{S}_{\rho\tau\rho\tau}(\mathbf{p},\mathbf{q}) = \sum \bar{Y}_{lm}(\mathbf{p}p^{-1})\tilde{S}^{\,l}_{\rho\tau\rho\tau}(p)Y_{lm}(\mathbf{q}q^{-1}). \tag{5.5}$$

The functions $S^{\,l}_{\rho\tau\rho\tau}: = p\tilde{S}^{\,l}_{\rho\tau\rho\tau}$ are called scattering functions. The operators D, respectively, C are S-matrix symmetries, i.e., fulfill (4.2), if and only if

$$(\mathbf{p}\partial_p + \mathbf{q}\partial_q + 1)\hat{S}_{\rho\tau\rho\tau}(\mathbf{p},\mathbf{q}) = 0, \tag{5.6}$$

respectively,

$$4\delta'(p^2 - q^2)(\mathbf{p}\partial_p + \mathbf{q}\partial_q + 1)\hat{S}_{\rho\tau\rho\tau}(\mathbf{p},\mathbf{q}) + \delta(p^2 - q^2)(\Delta_p - \Delta_q)\hat{S}_{\rho\tau\rho\tau}(\mathbf{p},\mathbf{q}) = 0 \tag{5.7}$$

on the scattering manifold (4.10), (4.11).

Suppose first that D is a symmetry so that (5.6) holds. Insert (5.5) into (5.6) and note that the dilation operator $\mathbf{p}\partial_p + \mathbf{q}\partial_q$ gives zero on the dilation-invariant spherical harmonics while it is simply pd_p on functions of p,

$$(pd_p + 1)\tilde{S}^{\,l}_{\rho\tau\rho\tau}(p) = d_p(p\tilde{S}^{\,l}_{\rho\tau\rho\tau}(p)) = 0, \tag{5.8}$$

so that the scattering functions $S^{\,l}_{\rho\tau\rho\tau}$ are constant. Conversely, if the $S^{\,l}_{\rho\tau\rho\tau}$ are constant, (5.6) is true and hence D is a symmetry. This leads to

Theorem 5.1: Let $S^{\,l}_{\rho\tau\rho\tau}(p)$ be the scattering functions of a Galilean invariant theory. Then the following statements are equivalent:

(i) D fulfills (4.2),

(ii) C fulfills (4.2),

(iii) $S^{\,l}_{\rho\tau\rho\tau}$ is constant.

Proof: (i)⇔(iii) and (ii)⇔(i) has already been shown. To show (iii)⇒(ii), it is enough to prove (5.7). As (iii)⇔(i), the first term in (5.7) vanishes. Thus, only

$$\delta(p - q^2)(\Delta_p - \Delta_q)\hat{S}_{\rho\tau\rho\tau} = 0$$

has to be shown for constant $S^{\,l}_{\rho\tau\rho\tau}$ in (5.5). Since

$$\Delta(f(p)Y_{lm}) = (\Delta f)Y_{lm} - l(l+1)p^{-2}fY_{lm'},$$

(5.5) implies

$$\Delta_p\hat{S}_{\rho\tau\rho\tau} = \Delta_q\hat{S}_{\rho\tau\rho\tau} + \Delta_p(p^{-1})\sum \bar{Y}_{lm}(p^{-1}\mathbf{p})S^{\,l}_{\rho\tau\rho\tau}Y_{lm}(q^{-1}\mathbf{q}).$$

The last term, multiplied by $\delta(p^2 - q^2)$, is zero on test functions in $\mathbf{q}$ as the radial integration contains a factor q^2.

An example of a theory in which C and D occur as symmetries of the scattering amplitude, i.e., fulfill (4.2), is ordinary quantum mechanics in second quantized form for two particles with interaction potential $1/r^2$; the scattering functions S^l in this theory are constant.[28] On the other hand, if the interaction potential $V(r)$ satisfies $\int_0^\infty |V(r)|dr < \infty$, then the scattering functions cannot be constant, by Levinson's theorem.[29] Furthermore, the $1/r^2$ potential allows the construction of a scattering amplitude symmetry that is *not* additive; as C is a symmetry, $(\Delta_1 + \Delta_2)S = (\Delta_3 + \Delta_4)S$. By boosts invariance, $(\partial_1 + \partial_2)S = (\partial_3 + \partial_4)S$ or $(\partial_1 + \partial_2)^2S = (\partial_3 + \partial_4)^2S$ so that $\partial_1\partial_2S = \partial_3\partial_4S$, which cannot be obtained by an additive Q.

ACKNOWLEDGMENTS

The author would like to thank K. Baumann and H. Reeh for constant encouragement and discussions.

[1] F. Gürsey and L. Radicati, Phys. Rev. Lett. **13**, 173–175 (1964).
[2] B. Sakita, Phys. Rev. **136** B, 1756–1760 (1964).
[3] E. P. Wigner, Phys. Rev. **51**, 106–119 (1937).
[4] S. Coleman, Phys. Rev. **138** B, 1262–1267 (1965).
[5] K. T. Mahanthappa and E. C. G. Sudarshan, Phys. Rev. Lett. **14**, 458–461 (1965).
[6] Riazuddin, and L. K. Pandit, Phys. Rev. Lett. **14**, 462–466 (1965).
[7] T. F. Jordan, Phys. Rev. **140** B, 766–768 (1965).
[8] S. Coleman and J. Mandula, Phys. Rev. **159**,1262–1267 (1965).
[9] J. T. Lopuszanski, Commun. Math. Phys. **12**, 80 (1969).
[10] J. T. Lopuszanski, Commun. Math. Phys. **14**, 158–164 (1969).
[11] J. T. Lopuszanski, J. Math. Phys. **12**, 2401–2412 (1971).
[12] W. D. Garber and H. Reeh, J. Math. Phys. **19**, 985–986 (1978).
[13] W. D. Garber and H. Reeh, Commun. Math. Phys. **67**, 179–186 (1979).
[14] W. D. Garber and H. Reeh, Commun. Math. Phys. **70**, 169–180 (1979).
[15] W. D. Garber and H. Reeh, J. Math. Phys. **19**, 59–66 (1978).
[16] H. Reeh, in *Fundamental Problems of Theoretical and Mathematical Physics* (Dubna, 1979), pp. 480–486.
[17] A. O. Barut and R. Rączka, *Theory of Group Representations and Applications* (Polish Scientific Publishers, Warsaw, 1977).
[18] U. Niederer, Helv. Phys. Acta **45**, 802–810 (1972).
[19] J. M. Lévy-Leblond, Commun. Math. Phys. **4**, 157–176 (1967).
[20] V. Bargmann, Ann. Math. **59**, 1–46 (1954).
[21] R. Schrader, Commun. Math. Phys. **10**, 155–178 (1968).
[22] K. Hepp, in *Axiomatic Field Theory* edited by M. Chrétien and S. Deser, Brandeis University Summer Institute, 1965 (Gordon and Breach, New York, 1966), pp. 137–246.
[23] V. Enss, Commun. Math. Phys. **61**, 285–291 (1978).
[24] L. D. Landau and E. M. Lifschitz, *Mechanik* (Akademie-Verlag, Leipzig, 1961).
[25] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics II* (McGraw-Hill, New York, 1953).
[26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Tables* (Dover, New York, 1970).
[27] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
[28] I. I. Goldman, V. D. Krivchenkov, V. I. Kogan, and V. M. Galitskii, *Problems in Quantum Mechanics* (Infosearch Ltd., London, 1960).
[29] K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory* (Springer, New York, 1977).

Finitely many symmetries of the S-matrix

Wolf-Dieter Garber

Institut für Theoretische Physik, Universität Göttingen, Bunsenstrasse 9, D-3400 Göttingen, West Germany

(Received 22 January 1982; accepted for publication 30 April 1982)

S-matrix symmetries are operators that commute with the S-matrix and act additively on incoming n-particle states. For relativistic field theories, the structure of such symmetries is known if they arise from a local, conserved current density. In this paper it is shown that the same structural result is obtained if a relativistic field theory with finitely many but possibly nonlocal S-matrix symmetries is considered.

PACS numbers: 11.20. – e, 11.30. – j, 11.10. – z

1. INTRODUCTION

The structure of S-matrix symmetries has been extensively investigated in the literature. Consider, in a Wightman field theory, a conserved current j^μ that is local and relatively local with respect to other fields. Even for a current that is not translation-covariant, a form Q can be defined that is, in an appropriate sense, the integral over j^0. With the usual assumptions of Haag–Ruelle scattering theory that allow the definition of asymptotic states, Q can be extended to (a dense subset of) those states, is at most bilinear in asymptotic fields, and commutes with the S-matrix.[1] If, in addition, the vacuum is invariant, the linear terms in Q disappear, and Q has the form[1–3]

$$Q = \int dp\, dq\, Q^{\alpha\beta}(p,q)(a^{\rm in}(p,\alpha))^* a^{\rm in}(q,\beta), \tag{1.1}$$

where $a^{\rm in}(p,\alpha)$ is the annihilation operator, with support on the mass hyperboloid $p^2 = m_\alpha$, of the incoming free field of type α. Any form Q that is bilinear in incoming fields and commutes with the S-matrix, even if it does not arise from a local current, will be called an S-matrix symmetry.

If an S-matrix symmetry arises from a local current, the commutator of Q with an asymptotic free field $\phi_\alpha^{\rm ex}(x)$ (where ex stands for either in or out) will again be a local field. Any S-matrix symmetry for which $[Q,\phi_\alpha^{\rm ex}(x)]$ is a local field will itself be called *local* regardless of whether it arises from a local current or not. For a local S-matrix symmetry[2–4]

$$Q^{\alpha\beta} = 0 \text{ for } m_\alpha \neq m_\beta, \tag{1.2}$$

so that it is sufficient to consider only one mass multiplet $m = m_\alpha = m_\beta$, and

$$Q^{\alpha\beta} = q^{\alpha\beta}\delta(p-q) \tag{1.3}$$

where $q^{\alpha\beta}$ is a polynomial both in p and in derivatives with respect to p_ν, $\nu = 0,1,2,3$.

Arbitrary polynomials $q^{\alpha\beta}$ will only occur, in general, for noninteracting fields. If nontrivial interaction is assumed, there are further restrictions on $Q^{\alpha\beta}$. It is sufficient to require that the particles can be ordered in such a way that the elastic two-particle scattering amplitude, for any two consecutive particles, is nonzero in an open set of momenta which fulfill energy and momentum conservation. Then, at least for scalar asymptotic fields,[5]

$$q^{\alpha\beta} = \delta^{\alpha\beta}(a_\nu p^\nu + b_{\mu\nu}(p^\mu\partial^\nu - p^\nu\partial^\mu)) + c^{\alpha\beta}, \tag{1.4}$$

i.e., Q is a linear combination of Poincaré group generators and internal symmetries.

Thus the structure of local S-matrix symmetries is completely known. Now, it is easy to construct nonlocal S-matrix symmetries, at least in free field theories. Consider a single free field of spin zero, and define

$$Q := \int dp\, dq\, \exp\{-(\mathbf{p}-\mathbf{q})^2\} a^*(p) a(q)\,\cdot \tag{1.5}$$

This form Q is obviously additive, but $[Q,\phi,(x)]$ is not local, and (1.3) is manifestly wrong.

For Galilean field theories it has been shown that the analogs of (1.2)–(1.4) can still be proven, even for *a priori* nonlocal S-matrix symmetries, if the space of all such symmetries decomposes into finite-dimensional subspaces that are invariant under the kinematical invariance group (the Galilei group for Galilei theories); under this assumption, the symmetries turned out to be local.[6] It will be shown in this note that the same result holds for relativistic theories with the Poincaré group as kinematical invariance group.

Other, truly nonlocal, conserved quantities have been considered[7] that arise from currents. Those quantities, however, are neither bilinear in asymptotic fields nor do they commute with the S-matrix so that the results of this note do not apply.

2. REDUCTION TO A MASS MULTIPLET

Consider a finite-dimensional space E of S-matrix symmetries of the form (1.1), and let E be invariant under the Poincaré group. With Q, the Poincaré-transformed symmetry $U^{-1}(a,\Lambda)QU(a,\Lambda)$ is thus again in E; for a basis $Q_1,\ldots,Q_n$ of E,

$$U^{-1}(a,\Lambda)Q_i U(a,\Lambda) = D_i^k(a,\Lambda)Q_k\,. \tag{2.1}$$

Hence, $D(a,\Lambda)$ is a finite-dimensional matrix representation of the Poincaré group in E. Recall Lemma 2.2 of Ref. 6.

Lemma 2.1: Let D be a finite-dimensional matrix representation of the d-dimensional translation group

$$D(\mathbf{a})D(\mathbf{b}) = D(\mathbf{a}+\mathbf{b}) \quad \mathbf{a},\mathbf{b}\in\mathbb{R}^d. \tag{2.2}$$

Then there exist d commuting matrices $A_1,\ldots,A_d$ such that

$$D(\mathbf{a}) = \exp\left\{i\sum_{m=1}^{d} a_m A_m\right\}. \tag{2.3}$$

Note that the space translation subgroup of the Poincaré group P and the Galilei group G are the same so that Lemma 2.3 of Ref. 6 can be used without change. This gives, for the kernel $Q^{\alpha\beta}(p,q)$ of any symmetry Q in E, that $Q^{\alpha\beta} = 0$ unless $\mathbf{p}-\mathbf{q}$ takes on a finite number of values. As the rotation subgroups of P and G are also the same, $Q^{\alpha\beta} = 0$ unless $\mathbf{p} = \mathbf{q}$, by the argument following Lemma 2.3 in Ref. 6. Now, in relativistic field theories, there is no mass superselection rule which would immediately imply $Q^{\alpha\beta}=0$ for $m_\alpha \neq m_\beta$. However, one can argue as follows:

Under Lorentz transformations Λ,

$$Q^{\alpha\beta}(p,q) = (D^{s(\alpha),0}(\Lambda))_{\alpha\gamma}(D^{s(\beta),0}(\Lambda))_{\beta\delta} \times D(\Lambda^{-1})Q^{\gamma\delta}(\Lambda^{-1}p,\Lambda^{-1}q). \tag{2.4}$$

Assume that $Q^{\alpha\beta}$ has support for vectors p,q on different mass hyperboloids, i.e., $p^2=m_\alpha^2$, $q^2=m_\beta^2$ with $m_\alpha\neq m_\beta$, and by the argument just given, $\mathbf{p}=\mathbf{q}$. Since $m_\alpha\neq m_\beta$, $p\neq q$ even for $\mathbf{p}=\mathbf{q}$, and there is a Lorentz transformation Λ_0 depending on $\mathbf{p}$, $\mathbf{q}$ such that the space part of $\Lambda_0^{-1}(p-q)$ does not vanish even for $\mathbf{p}=\mathbf{q}$. For such a Lorentz transformation the right-hand side of (2.4) vanishes whereas $Q^{\alpha\beta}\neq 0$. This proves part (ii) of the following lemma:

Lemma 2.2: Assume that the space E of S-matrix symmetries decomposes into finite-dimensional subspaces E_i invariant under the Poncaré group P. Then, the kernel $Q^{\alpha\beta}$ of any Q in E fulfills

(i) $Q^{\alpha\beta}=0$ for $\mathbf{p}\neq\mathbf{q}$,

(ii) $Q^{\alpha\beta}=0$ for $m_\alpha\neq m_\beta$.

Thus, $Q^{\alpha\beta}$ has only support for $p=q$, and

$$Q^{\alpha\beta}(p,q) = q_\nu^{\alpha\beta}(p)\delta^{(\nu)}(p-q), \tag{2.5}$$

where the multi-index $\nu=(\nu_0,\nu_1,\nu_2,\nu_3)$ at the delta function denotes ν_k-fold differentiation with respect to p_k-q_k, $k=0,1,2,3$.

3. POLYNOMIAL CHARACTER OF Q

In this section it will be shown that the distributions $q_\nu^{\alpha\beta}$ are, in fact, polynomials in p. Combine (2.4) and (2.5) and compare terms with the same order in the differentiation operators to get

$$D(\Lambda)q_\nu^{\alpha\beta}(p) = (D^{s(\alpha),0}(\Lambda))_{\alpha\gamma} \times \overline{(D^{s(\beta),0}(\Lambda))_{\beta\delta}}D(\Lambda)_\nu^\mu q_\mu^{\gamma\delta}(\Lambda^{-1}p), \tag{3.1}$$

where D is some tensor product of the vector representation of the Lorentz group, and the sum in μ is only over terms with $\Sigma\mu_k=\Sigma\nu_k$. Now, the boost subgroups act differently in P and in G. Consider the group generated by M_{03} with rapidity ξ, and call D_3 the representation D, restricted to this group. From (3.1),

$$D_3(\xi)q_\nu^{\alpha\beta}(p) = (\tilde{D}_3(\xi))_{\nu,\gamma\delta}^{\mu,\alpha\beta}q_\mu^{\gamma\delta}(\Lambda^{-1}p). \tag{3.2}$$

All representations on the right-hand side of (3.1) have been combined into a single finite-dimensional representation $\tilde{D}_3$. By Lemma 2.1, both D_3 and $\tilde{D}_3$ are generated by finite-dimensional matrices A_3 and Σ_3, say,

$$A_3q_\nu^{\alpha\beta}(p) = (\Sigma_3)_{\nu\gamma\delta}^{\mu\alpha\beta}q_\mu^{\gamma\delta}(p) + (p_3\partial_0 - p_0\partial_3)q_\nu^{\alpha\beta}(p). \tag{3.3}$$

Here A_3 acts only on the (suppressed) index enumerating the elements of E. Put

$$B_3 := A_3 - \Sigma_3,$$

to get

$$(B_3)_{\nu\delta\gamma}^{\mu\alpha\beta}q_\mu^{\gamma\delta}(p) = (p_3\partial_0 - p_0\partial_3)q_\nu^{\alpha\beta}(p). \tag{3.4}$$

Next, on any given mass hyperboloid $p^2=m^2$, one can define

$$\hat{q}_\nu^{\alpha\beta}(\mathbf{p}) = q_\nu^{\alpha\beta}(\mathbf{p},\omega(\mathbf{p}^2,m)),(\omega(\mathbf{p}^2,m))^2 := \mathbf{p}^2+m^2 \tag{3.5}$$

so that (2.3) implies

$$(B_3)_{\nu\gamma\delta}^{\mu\alpha\beta}\hat{q}_\mu^{\gamma\delta}(\mathbf{p}) = \omega(\mathbf{p}^2,m)\partial_3\hat{q}_\nu^{\alpha\beta}(\mathbf{p}).$$

Consider now rotations R. By (3.1) and (3.5), $D(R)$ is a representation of the rotation group on the space F of distributions $\hat{q}_\nu^{\alpha\beta}(\mathbf{p})$. Now, any representation of this group splits up into irreducibles D^l of dimension $(2l+1)$. The corresponding invariant subspace F^l of F is unitarily equivalent to the space H^l of spherical harmonics Y_{lm}, $m=-l,\ldots,1$, i.e. there is a unitary operator $U^l{:}H^l \to F^l$ such that the vectors $\hat{q}_{lm}{:}U^lY_{lm}$ form a basis of F^l. As U^l commutes with D^l, U^l is a multiple of the identity, $U^l=d^l 1$, with a rotation-invariant distribution $d^l=d^l(\mathbf{p}^2)$. Put

$$c_l(\mathbf{p}^2) := (\mathbf{p}^2)^{-1/2}d_l(\mathbf{p}^2), \tag{3.6}$$

to get

$$\hat{q}_{lm} = c_l(\mathbf{p}^2)((\mathbf{p}^2)^{1/2}Y_{lm}) =: c_l(\mathbf{p}^2)y_{lm}(\mathbf{p}). \tag{3.7}$$

The functions $y_{lm}(p)$ are homogeneous polynomials of degree l that are harmonic, i.e., fulfill $\Delta y_{lm}=0$. In fact, they form a basis of harmonic polynomials (see p. 1270 ff. in Ref. 8).

In the new basis $\hat{q}_{lm}$, (3.4) reads

$$(B_3)_{lm}^{ik}\hat{q}_{ik}(\mathbf{p}) = \omega(\mathbf{p}^2,m)\partial_3\hat{q}_{lm}(\mathbf{p}) \tag{3.8}$$

and one has

Lemma 3.1: Let $\{\hat{q}_{lm}(\mathbf{p})\}$ be distributions obeying (3.8) and (3.7). Then there are polynomials $r_{lm}(\mathbf{p})$ and $s_{lm}(\mathbf{p})$ with $\hat{q}_{lm}(\mathbf{p}) = r_{lm}(\mathbf{p}) + \omega(\mathbf{p}^2,m)s_{lm}(\mathbf{p})$.

Proof: (3.7) and (3.8) imply

$$\sum_{i,k}(B_3)_{lm}^{ik}c_i(\mathbf{p}^2)y_{ik}(\mathbf{p}) = \omega(\mathbf{p}^2,m) \times\{2p_3c_l'(\mathbf{p}^2)y_{lm}(\mathbf{p}) + c_l(\mathbf{p}^2)\partial_3y_{lm}(\mathbf{p})\}. \tag{3.9}$$

Now, ∂_3y_{lm} is a harmonic polynomial of degree $l-1$ and hence a linear combination of $y_{l-1,n}$. In polar coordinates

$$y_{lm} = p^lP_l^m(\cos\delta)\exp(im\phi), \tag{3.10}$$

with associated Legendre functions P_l^m. As $p_3=p\cos\delta$, ∂_3y_{lm} is proportional to $\exp(im\phi)$ so that

$$\partial_3y_{lm} = d_my_{l-1,m}. \tag{3.11}$$

In the first term on the right-hand side of (3.9), p_3y_{lm} is not harmonic. However, by the identity

$$(2l+1)zP_l^m(z) = (l-m+1)P_{l+1}^m(z) + (l+m)P_{l+1}^m(z), \tag{3.12}$$

(see 8.5.3 in Ref. 9), one has that p_3y_{lm} is proportional to $\exp(im\phi)$ too. By the orthogonality of the exponentials, the matrix B_3 has only nonvanishing elements for $k=m$, and (3.9) implies

$$\sum_i(B_3)_{lm}^{im}c_i(\mathbf{p}^2)p^iP_i^m = \omega(\mathbf{p}^2,m)\{2c_l'(\mathbf{p}^2) \times\{(l-m+1)p^{l+1}P_{l+1}^m + (l+m)p^{l-1}P_{l-1}^m\} + c_l(\mathbf{p}^2)d_mp^{l-1}P_{l-1}^m\}. \tag{3.13}$$

In (3.13), only associated Legendre functions with the same upper index m occur. These functions (by 8.14.*ii* in Ref. 9) are orthogonal for different lower indices. Hence, B_3 has only nonvanishing elements for $i=l+1$ and $i=l-1$. For $i=l+1$, (3.13) implies

$$(B_3)_{lm}^{l+1,m} c_{l+1}(\mathbf{p}^2) = \omega(\mathbf{p}^2,m) d_{lm} c_l'(\mathbf{p}^2), \tag{3.14}$$

with suitable numeral factors $d_{lm} \neq 0$. As B_3 is a finite-dimensional matrix, the index l runs up to some largest value L, and (3.14) is only valid for $l+1 \leqslant L$. For $l = L$, $c_L'(p^2) = 0$, so that c_L is constant. In general, (3.14) implies

$$c_l(x) = r_l(x) + \omega(x,m) s_l(x), \tag{3.15}$$

with suitable polynomials r_l and s_l. For $l = L$ (3.15) has just been shown. Assume (3.15) for l. To show it for $l-1$, observe that (3.15) and (3.14) give

$$c_{l-1}'(x) = a_{lm}\, \omega(x,m)^{-1} \{ r_{l+1}(x) + \omega(x,m) s_{l+1}(x) \},$$

with suitable numerical factors a_{lm}. Integrating the polynomial $a_{lm} s_{l+1}$ results again in a polynomial, say r_l. In integrating $\omega(x,m)^{-1} r_{l+1}$, put $y := x + m^2$; this gives $\omega(x,m) s_l$ with a polynomial s_l. This proves (3.15), and the assertion follows with (3.7).□

By (3.5), Lemma 3.1 implies that $q_\nu^{\alpha\beta}$ is a polynomial in p. Thus, the following theorem has been proved.

Theorem 3.2: Suppose that the space E of all S-matrix symmetries of the form (1.1) splits up into finite-dimensional subspaces E_i that are invariant under the Poincaré group P. Then the kernel $Q^{\alpha\beta}(p,q)$ of any Q in E fulfills

(i) $Q^{\alpha\beta}(p,q) = 0$ for $m_\alpha \neq m_\beta$,

(ii) $Q^{\alpha\beta}(p,q) = q^{\alpha\beta}(p-q)$,

where $q^{\alpha\beta}$ is a polynomial both in p and $\partial/\partial p_\nu$, $\nu = 0,1,2,3$.

Hence, Q is in fact local.

Nowhere in this note has the fact been used that Q commutes with the S-matrix; only the bilinear character of Q was essential. If the S-matrix in a given theory is nontrivial, then Q has an even simpler form, i.e., (1.4) is valid at least for scalar asympotic fields.[5]

ACKNOWLEDGMENTS

I should like to thank K. Baumann and H. Reeh for their constant encouragement and the many helpful discussions from which many useful clarifications resulted.

[1] W. D. Garber, J. Math. Phys. **20**, 2596–2604 (1979).

[2] W. D. Garber and H. Reeh, J. Math. Phys. **19**, 59–66 (1978).

[3] W. D. Garber and H. Reeh, J. Math. Phys. **19**, 985–986 (1978).

[4] W. D. Garber and H. Reeh, Commun. Math. Phys. **67**, 179–186 (1979).

[5] W. D. Garber and H. Reeh, Commun. Math. Phys. **70**, 169–180 (1979).

[6] W. D. Garber, Galilean Field Theories and Nonlocal S-Matrix Symmetries, preprint, Göttingen, 1981.

[7] D. Buchholz and J. T. Lopuszanski, Lett. Math. Phys. **3**, 175–180 (1979).

[8] P. M. Morse and H. Feshbach, *Methods of Mathematical Physics II* (McGraw-Hill, New York, 1953).

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Tables* (Dover, New York, 1970).

The partial-wave projected Coulomb T matrix for all l in closed hypergeometric form[a)]

H. van Haeringen
Department of Mathematics, Delft University of Technology, Delft, The Netherlands

(Received 4 December 1981; accepted for publication 16 April 1982)

We obtain relatively simple closed expressions for the partial-wave projections of the off-shell Coulomb T matrix in the momentum representation, $\langle p|T_{cl}(k^2)|p'\rangle$, for all $l = 0,1,2,\cdots$. These exact analytic expressions consist of three parts: (i) $\mathscr{F}_l$, simple combinations of the Jacobi polynomial $P_l^{(i\gamma,-i\gamma)}$ and the hypergeometric function ${}_2F_1(1,i\gamma;1+i\gamma;\cdot)$, with different arguments [γ is Sommerfeld's parameter; the Coulomb potential is $V_c(r) = 2k\gamma/r$]; (ii) a polynomial $\mathscr{E}_l$; and (iii) $\mathscr{L}_l$ = a polynomial times $\ln[(p+p')^2/(p-p')^2]$. The polynomials under (ii) and (iii) are given in terms of the Jacobi polynomials $P_l^{(m,-m)}$, $m = 0,1,\ldots,l$. We derive interesting relations, especially valuable for the theory of off-shell Coulomb scattering, and we present simple closed expressions for the special cases $l = 0$, 1, and 2; $p'\to k$, $p'\to p$, $p'\to\infty$, and $\gamma\to 0$.

PACS numbers: 11.20.Dj, 02.30. + g, 03.65. − w

1. INTRODUCTION

An old and intriguing problem in scattering theory is the incorporation of the Coulomb interaction between charged particles in the description of n-particle reactions. On the one hand, the Coulomb interaction is known with great accuracy and it is described by a very simple mathematical formula. On the other hand, its exact incorporation in the theoretical quantum-mechanical description of scattering reactions is complicated. The complications are due to the so-called long range of the Coulomb potential.

For two-particle reactions one needs quantities *on* the energy shell only. Two-particle Coulomb difficulties and peculiarities have been studied extensively and are well understood.

For n-particle reactions with $n > 2$ *off*-shell quantities also play a role. In this case the problems associated with the Coulomb potential are not completely solved.[1,2] Basic in the description of n-particle reactions is the off-shell two-body T matrix. When the total interaction consists of the sum of the Coulomb potential and a short-range potential V_s, the total T matrix can be expressed as $T = T_c + T_{cs}$. Here T_c is the pure Coulomb T matrix and T_{cs} can be obtained by solving an integral equation. It follows that T_c and its partial-wave projections T_{cl} $(l = 0,1,\cdots)$ play a prominent part in off-shell charged-particle scattering.[3]

For the three-dimensional Coulomb T matrix T_c various expressions are given in the literature (see, e.g., Refs. 4 and 5). For T_{cl} we have obtained an integral representation[6] which turned out to be useful to carry out numerical calculations (see Ref. 7). However, the complete analytic structure of T_{cl} is not easily read off from this integral representation. We have also published hypergeometric-function expressions for T_{cl}, in the cases $l = 0$ and $l = 1$ only.[6]

The purpose of this paper is to continue this study of T_{cl}. We shall derive exact analytic expressions for T_{cl} for all $l = 0,1,\cdots$. These are useful (i) to check (numerical) approximations in general, and (ii) to derive analytic properties. We have reported preliminary results in Ref. 8. The closed expressions we shall obtain are especially useful for the investigation and approximation of T_{cl} near the half-shell and on-shell points where it has branch-point singularities. Another singularity of T_{cl}, at zero energy, can also be treated by using these closed forms. Indeed, we shall use only one hypergeometric function, ${}_2F_1(1,i\gamma;1+i\gamma;)\equiv F_{i\gamma}(\ )$, which we have studied extensively before, in particular at the zero-energy singularity.[9]

In Sec. 2 we give the notations we shall use in this paper. These are consistent with notations in previous work (see especially Ref. 10). In Sec. 3 we shall give the aforementioned integral representation for T_{cl} (Ref. 6) and derive an interesting and useful connection with the Coulomb Jost state in the momentum representation, $\langle p|kl\uparrow\rangle_c$.

In Sec. 4 we split T_{cl} into three parts; a hypergeometric, a rational, and a logarithmic part. Here we introduce three functions representing these parts: $\mathscr{F}_l$, $\mathscr{E}_l$ and $\mathscr{L}_l$. Closed expressions for $\mathscr{F}_l$, $\mathscr{E}_l$, and $\mathscr{L}_l$ will be derived in Secs. 5, 6, and 7, respectively.

In Sec. 8 we give a summary of the most important formulas derived in this paper.

2. NOTATIONS

The notations and conventions we shall use in this paper are in conformity with those in previous work on the Coulomb potential, see especially Ref. 10. We choose units such that $\hbar = 2m = 1$ where m is the reduced mass and we denote the energy variable by $E \equiv k^2$. The Coulomb potential is given in the coordinate representation by

$$V_c(r) = 2k\gamma/r,$$

where γ is Sommerfeld's parameter. Since $k\gamma$ is constant, γ is k dependent. We shall work in the momentum representation. The momentum variables $\mathbf{p}$, $\mathbf{p}'$ are real, p and p' are real positive whereas k (and hence γ) is a complex variable. In some derivations k is assumed to be real positive, too, as will be clear from the context. In such cases it is customary to let k approach the real-positive k axis from above. For instance, the (off-shell) partial-wave Coulomb T matrix for positive energy, $k^2 > 0$, is defined by

[a)] This work is part of the project "The Coulomb Potential in Quantum Mechanics and Related Topics."

 0022-2488/83/051267-07$02.50

$$\langle p|T_{cl}(k^2)|p'\rangle = \lim_{\epsilon\downarrow 0}\langle p|T_{cl}((k+i\epsilon)^2)|p'\rangle, \quad k>0,$$

where $k\neq p$, $k\neq p'$. Since this quantity occurs frequently we shall denote it by T_{cl} for brevity. The formulas for T_{cl} that we shall derive are valid for *complex* k, which follows by analytic continuation.

The Coulomb Jost state is denoted in abstract notation by $|kl\uparrow\rangle_c$. Its momentum representation $\langle p|kl\uparrow\rangle_c$ is the Hankel transform of its coordinate representation $\langle r|kl\uparrow\rangle_c$. We have

$$\langle p|kl\uparrow\rangle_c = \int_0^\infty \langle pl|r\rangle\,\langle r|kl\uparrow\rangle_c r^2\,dr,$$

$$\langle pl|r\rangle = (2/\pi)^{1/2} i^{-l} j_l(pr),$$

$$\langle r|kl\uparrow\rangle_c = (2/\pi)^{1/2} e^{\pi\gamma/2}(kr)^{-1} W_{-i\gamma,l+1/2}(-2ikr),$$

where j_l is the spherical Bessel function and $W_{.,.}$ is Whittaker's function. The asymptotic behavior of $\langle r|kl\uparrow\rangle_c$ is determined by

$$\lim_{kr\to\infty}\langle r|kl\uparrow\rangle_c kr\exp(-ikr+i\gamma\ln 2kr) = (2/\pi)^{1/2}.$$

We shall use the following abbreviations:

$$a=(p-k)/(p+k),\quad a'=(p'-k)/(p'+k),$$

$$u=(p^2+k^2)/2pk,\quad u'=(p'^2+k^2)/2p'k,$$

$$v=(p^2-k^2)/2pk,\quad v'=(p'^2-k^2)/2p'k,$$

$$w=(p^2+p'^2)/2pp',$$

$$u^2-v^2=1,$$

$$w=uu'-vv',$$

$$\epsilon_m=\begin{cases}1, & m=0\\ 2, & m=1,2,\cdots,\end{cases}$$

$$f_{cl}=e^{\pi\gamma/2}\, l!/\Gamma(l+1+i\gamma),$$

$$c_{l\gamma}=(l!)^2\Gamma(1+i\gamma)\Gamma(1-i\gamma)/[\Gamma(l+1+i\gamma)\times\Gamma(l+1-i\gamma)],$$

$$c_{l\gamma}^{-1}=\binom{l+i\gamma}{l}\binom{l-i\gamma}{l}=\prod_{n=1}^{l}(1+\gamma^2/n^2),$$

$$c^{lm}=(-l)_m/(l+1)_m=(-)^m(l!)^2/[(l-m)!(l+m)!].$$

Note that $c_{l0}=c^{l0}=1$ and $c^{lm}=c^{l(-m)}$. Furthermore,

$F_{i\gamma}(\cdot)={}_2F_1(1,i\gamma;1+i\gamma;\cdot)$,

$\mathfrak{Q}_l$, $\mathfrak{P}^{\mu}_{\nu}$: Legendre functions,

$P_l^{(\cdot,\cdot)}$: Jacobi's polynomial,

P_l: Legendre's polynomial.

In Secs. 3 and 4 we shall use

$$q=|\mathbf{p}-\mathbf{p}'|,\quad \cos\theta=\hat{\mathbf{p}}\cdot\hat{\mathbf{p}}',$$

$$y=(x+1)/(x-1),$$

$$x=-(1+4q^{-2}pp'vv')^{1/2},$$

$$\alpha=\alpha(y)=\cos\theta=uu'-\tfrac{1}{2}vv'(y+1/y).$$

It is sometimes convenient to use the abbreviations $p(t)$ and p_m:

$$p(t)=P_l(uu'-\tfrac{1}{2}vv'(t+t^{-1})),$$

$$p_m=c^{lm}(aa')^m P_l^{(m,-m)}(u)P_l^{(m,-m)}(u').$$

It can be shown that $p_m=p_{-m}$, $m=0,1,\ldots,l$.

Often variables will be suppressed, e.g.,

$$\mathscr{F}_l=\mathscr{F}_l(p,p')=\mathscr{F}_l(p,p';k;\gamma),$$

$$\mathscr{E}_l=\mathscr{E}_l(p,p')=\mathscr{E}_l(p,p';k;\gamma),$$

$$\mathscr{L}_l=\mathscr{L}_l(p,p')=\mathscr{L}_l(p,p';k;\gamma).$$

3. AN INTEGRAL REPRESENTATION FOR T_{cl} AND ITS RELATION WITH $\langle p|kl\uparrow\rangle_c$

It is well known that

$$\langle\mathbf{p}|V_c|\mathbf{p}'\rangle = k\gamma\pi^{-2}q^{-2}, \tag{3.1}$$

where $q=|\mathbf{p}-\mathbf{p}'|$ and γ is Sommerfeld's parameter (Sec. 2). Furthermore (Im $k\downarrow 0$),

$$\langle\mathbf{p}|T_c(k^2)|\mathbf{p}'\rangle = \frac{-k\gamma}{\pi^2 pp'vv'}\int_0^1 \frac{t^{i\gamma}(1-t^2)\,dt}{[1+t^2-t(y+1/y)]^2}. \tag{3.2}$$

The partial-wave projection of T_c is defined by

$$\langle p|T_{cl}(k^2)|p'\rangle = 2\pi\int_{-1}^{1}\langle\mathbf{p}|T_c(k^2)|\mathbf{p}'\rangle P_l(\cos\theta)\,d(\cos\theta). \tag{3.3}$$

Henceforth the left-hand side of Eq. (3.3) will be denoted by T_{cl} for brevity (Sec. 2). From (3.2) and (3.3) we have

$$T_{cl}=\frac{-2k\gamma}{\pi pp'vv'}\int_{-1}^{1}P_l(\alpha)\,d\alpha \times\int_0^1\frac{t^{i\gamma}(1-t^2)\,dt}{[1+t^2-t(y+1/y)]^2}. \tag{3.4}$$

In Sec. 4 we shall derive an expression for T_{cl} in terms of the hypergeometric function $F_{i\gamma}$ and elementary functions, by starting from Eq. (3.4). In this section we shall briefly consider an integral representation for T_{cl} and its relation with $\langle p|kl\uparrow\rangle_c$.

By using

$$\mathfrak{Q}_l(z)=\frac{1}{2}\int_{-1}^{1}(z-\alpha)^{-1}P_l(\alpha)\,d\alpha,$$

one easily deduces from Eq. (3.4) a useful integral representation for T_{cl} that has been given in Ref. 6, Eq. (24):

$$T_{cl}=\frac{2k\gamma}{\pi pp'}\int_0^1 t^{i\gamma}\frac{d}{dt}\mathfrak{Q}_l(z)\,dt, \tag{3.5}$$

where [cf. Eq. (23) of Ref. 6]

$$z=uu'-\tfrac{1}{2}vv'(t+1/t), \tag{3.6}$$

cf.

$$\alpha=uu'-\tfrac{1}{2}vv'(y+1/y).$$

Equation (3.5) is useful, e.g., for numerical calculations,[7] and for the derivation of the equality (cf. Ref. 11)

$$\lim_{p'\to\infty}p'^{l+2}T_{cl} = \tfrac{1}{2}k\gamma(4k)^{l+1}(l!)^2[(2l+1)!]^{-1} \times(p^2-k^2)\langle p|kl\uparrow\rangle_c f_{cl}^{-1}. \tag{3.7}$$

Here $\langle p|kl\uparrow\rangle_c$ is the Coulomb Jost state in the momentum representation, for which we have obtained the following integral representation[11]

$\langle p|kl\uparrow\rangle_c$

$$=f_{cl}\frac{(l+1)p^l}{2\pi k^{l+3}}\int_0^1 t^{l+i\gamma}(t^2-1)\times\left[t-(1-t)^2\frac{p^2-k^2}{4k^2}\right]^{-l-2}dt. \tag{3.8}$$

Equation (3.7) is obtained by noting that

$$\lim_{p'\to\infty} 2kz/p' = u-\tfrac{1}{2}v(t+t^{-1}),$$

$$\lim_{z\to\infty} z^{l+1}\mathfrak{Q}_l(z) = l!/(2l+1)!!,$$

$$\frac{dz}{dt} = -\tfrac{1}{2}vv'(1-t^{-2}).$$

Hence,

$$\lim_{p'\to\infty} p'^{l+2}T_{cl} = \frac{k\gamma}{\pi p^2}\frac{l!(l+1)!}{(2l+1)!}\times(p^2-k^2)(4k)^l\int_0^1\frac{t^{i\gamma}(1-t^{-2})\,dt}{[u-\frac{1}{2}v(t+t^{-1})]^{l+2}}$$

$$=\frac{k\gamma}{\pi p^2}\frac{l!(l+1)!}{(2l+1)!}(p^2-k^2)(4k)^l\times\left(\frac{p}{k}\right)^{l+2}\int_0^1\frac{t^{i\gamma+l}(t^2-1)\,dt}{[t-(1-t)^2(p^2-k^2)/4k^2]^{l+2}}.$$

4. THE REDUCTION OF T_{cl} TO ITS MAIN PARTS: $\mathscr{F}_l$, $\mathscr{E}_l$, $\mathscr{L}_l$

In this section we shall prove

$$T_{cl} = -\frac{k\gamma}{\pi pp'}c_{l\gamma}\left[(i\gamma)^{-1}\mathscr{F}_l+\mathscr{E}_l+\mathscr{L}_l\ln\left(\frac{p+p'}{p-p'}\right)^2\right]. \tag{4.1}$$

Here $c_{l\gamma}$ has been given in Sec. 2, $\mathscr{F}_l$ contains $F_{i\gamma}$ and Jacobi polynomials, and $\mathscr{E}_l$ and $\mathscr{L}_l$ are simple rational functions (in certain variables even polynomials). Hence $\mathscr{F}_l$ contains the "hypergeometric part" and $\mathscr{L}_l\ln[(p+p')/(p-p')]^2$ the "logarithmic part."

In Secs. 5–7 we shall derive various closed expressions for $\mathscr{F}_l$, $\mathscr{E}_l$, and $\mathscr{L}_l$, respectively. These derivations are lengthy and somewhat complicated. The reader who is interested in the results only is advised to proceed to Sec. 8, where a summary of the most interesting formulas will be given, including explicit formulas for the cases $l=0$, 1, and 2.

We note that we have obtained Eq. (4.1) for $l=0$ and $l=1$ in previous publications.

By inserting $d\alpha = -\tfrac{1}{2}vv'(1-y^{-2})\,dy$ into Eq. (3.4) we obtain

$$T_{cl} = \frac{k\gamma}{\pi pp'}\int_{\alpha=-1}^{1}dy(1-y^{-2})P_l(\alpha)\times\int_{t=0}^{1}\frac{t^{i\gamma}(1-t^2)\,dt}{[1+t^2-t(y+1/y)]^2}. \tag{4.2}$$

Our main task is, to split off the $\mathscr{F}_l$ part from T_{cl}. To this end we shall express the integral $\int dt$ in Eq. (4.2) in terms of $F_{i\gamma}$. From the well-known integral representation

$$F_{i\gamma}(y) = i\gamma\int_0^1 t^{i\gamma-1}(1-ty)^{-1}\,dt, \tag{4.3}$$

one easily verifies that

$$\frac{d}{dy}[F_{i\gamma}(y)+F_{i\gamma}(1/y)] = (1-y^{-2})i\gamma\int_0^1 t^{i\gamma}(1-t^2)(1-ty)^{-2}(1-t/y)^{-2}\,dt. \tag{4.4}$$

By inserting this into Eq. (4.2) we get

$$T_{cl} = \frac{-ik}{\pi pp'}\int_{\alpha=-1}^{\alpha=1}P_l(\alpha)\,d\,[F_{i\gamma}(y)+F_{i\gamma}(1/y)]. \tag{4.5}$$

When $\alpha=-1$, we have (cf. Sec. 2)

$$x = -\frac{pp'+k^2}{k(p+p')}\Rightarrow y = \frac{x+1}{x-1} = a'a. \tag{4.6}$$

Similarly, when $\alpha=+1$, $y=a'/a$. Since α is invariant for $y\to 1/y$ (see Sec. 2), $\alpha = uu'-\tfrac{1}{2}vv'(y+1/y)$, we have [cf. (4.6)]

$$\int_{\alpha=-1}^{1}P_l(\alpha)\,dF_{i\gamma}(1/y) = \int_{1/y=a'a}^{a'/a}P_l(\alpha)\,dF_{i\gamma}(y).$$

In this way we obtain from Eq. (4.5),

$$T_{cl} = \frac{-ik}{\pi pp'}\left(\int_{a'a}^{a'/a}+\int_{(a'a)^{-1}}^{a/a'}\right)P_l(\alpha)\,dF_{i\gamma}(y), \tag{4.7}$$

where the indicated limits are the values for y. Clearly it is convenient to split the expression for T_{cl} into four similar parts. Defining

$$f(z) = (i\gamma)^{-1}\int_{y=c}^{z}P_l(\alpha)\,dF_{i\gamma}(y), \tag{4.8}$$

we rewrite Eq. (4.7) as

$$T_{cl} = k\gamma(\pi pp')^{-1}\times[f(a'/a)+f(a/a')-f(a'a)-f((a'a)^{-1})]. \tag{4.9}$$

The constant c ($c>0$) occurring in (4.8) is of no importance as is easily seen from (4.9). In fact, f is just a primitive function. Introducing for convenience the function p,

$$p(y) = P_l(\alpha), \tag{4.10}$$

where we recall $\alpha = uu'-\tfrac{1}{2}vv'(y+1/y)$, we have from (4.8),

$$f(z) = (i\gamma)^{-1}\int_{y=c}^{z}p(y)\,dF_{i\gamma}(y). \tag{4.11}$$

Now we are in a position to split off the hypergeometric part in a suitable way. We shall prove

$$f(z) = (i\gamma)^{-1}\int_c^z F'_{i\gamma}(t)p(t)\,dt = -F_{i\gamma}(z)\int_0^1 t^{-i\gamma-1}p(zt)\,dt + \int_0^1 dt\,t^{-i\gamma}\int_c^z(1-\tau)^{-1}p'(\tau t)\,d\tau + c_1. \tag{4.12}$$

Here we assume that $\mathrm{Re}(-i\gamma)>l$ in order to ensure convergence of the integrals. In the final formulas this condition can be relaxed by means of analytic continuation.

The step from (4.11) to (4.12) is an important one. Indeed, in Eq. (4.12) the $F_{i\gamma}$ part of $f(z)$ is separated off. In view of the nature of the function p [see Eq. (4.10)] we have to evaluate integrals of elementary functions only. The constant c_1 plays no role in T_{cl} as is easily seen from Eq. (4.9).

In order to prove (the second equation of) (4.12), we differentiate with respect to z. Then we get

$$f'(z) = -F'_{i\gamma}(z)\int_0^1 t^{-i\gamma-1}p(zt)\,dt - F_{i\gamma}(z)\int_0^1 t^{-i\gamma}p'(zt)\,dt + \frac{1}{1-z}\int_0^1 t^{-i\gamma}p'(zt)\,dt.$$

By inserting the well-known equality

$$-F_{i\gamma}(z) + (1-z)^{-1} = (i\gamma)^{-1}z\frac{d}{dz}F_{i\gamma}(z),$$

we get

$$f'(z) = (i\gamma)^{-1}F'_{i\gamma}(z) \times\left[-i\gamma\int_0^1 t^{-i\gamma-1}p(zt)\,dt + z\int_0^1 t^{-i\gamma}p'(zt)dt\right],$$

which gives, after integration by parts,

$$f'(z) = (i\gamma)^{-1}F'_{i\gamma}(z)p(z). \tag{4.13}$$

This completes the proof of Eq. (4.12).

It may be noted that c and c_1 are connected by

$$f(c) = 0 = c_1 - F_{i\gamma}(c)\int_0^1 t^{-i\gamma-1}p(ct)\,dt. \tag{4.14}$$

We shall indicate the $F_{i\gamma}$ part of $f(z)$ by $f_1(z)$. Furthermore, we substitute in Eq. (4.12):

$$p'(\tau t) = \frac{1}{\tau}\frac{d}{dt}\{p(\tau t) - p(t)\} + \frac{1}{\tau}p'(t),$$

and denote the corresponding parts by $f_2(z)$ and $f_3(z)$. Then we have

$$f(z) = f_1(z) + f_2(z) + f_3(z) + c_1, \tag{4.15}$$

where

$$f_1(z) = -F_{i\gamma}(z)\int_0^1 t^{-i\gamma-1}p(zt)\,dt, \tag{4.16}$$

$$f_2(z) = \int_0^1 dt\, t^{-i\gamma}\int_c^z \frac{d\tau}{\tau(1-\tau)}\frac{d}{dt}\{p(\tau t) - p(t)\}, \tag{4.17}$$

$$f_3(z) = \int_0^1 dt\, t^{-i\gamma}\int_c^z \frac{d\tau}{\tau(1-\tau)}\frac{d}{dt}p(t). \tag{4.18}$$

From f_1 we shall derive an expression for $\mathcal{F}_l$, in Sec. 5. Similarly, from f_2 we shall obtain the simple rational function $\mathcal{E}_l$ in Sec. 6, and f_3 will give us the logarithmic part, $\mathcal{L}_l \ln[(p+p')/(p-p')]^2$ (see Sec. 7).

5. THE DERIVATION OF $\mathcal{F}_l$

In this section we shall derive an elegant explicit expression for $\mathcal{F}_l$, the hypergeometric part of T_{cl}. From (4.1), (4.9), and (4.15) we have

$$\mathcal{F}_l = -i\gamma c_{l\gamma}^{-1}[f_1(a'/a) + f_1(a/a') - f_1(a'a) - f_1((a'a)^{-1})], \tag{5.1}$$

where f_1 is defined by (4.16).

We shall use the addition theorem for the Legendre functions $\mathfrak{P}_\nu^\mu$ (cf. Ref. 12, p. 178);

$$P_l(uu' - \tfrac{1}{2}vv'(t+1/t)) = \frac{1}{2}\sum_{m=0}^{\infty}(-)^m\epsilon_m\mathfrak{P}_l^m(u)\mathfrak{P}_l^{-m}(u')(t^m + t^{-m}). \tag{5.2}$$

Because of

$$\mathfrak{P}_l^m(u) = 0, \quad m = l+1,\ l+2,\ \cdots \quad (l\in\mathbb{N}), \tag{5.3}$$

the sum in (5.2) terminates with $m = l$. With the help of the equalities $(-l\leqslant m\leqslant l)$

$$\mathfrak{P}_l^{-m}(u) = \frac{(l-m)!}{(l+m)!}\mathfrak{P}_l^m(u) = a^m\frac{l!}{(l+m)!}P_l^{(m,-m)}(u), \tag{5.4}$$

Eq. (5.2) can be rewritten as

$$P_l(uu' - \tfrac{1}{2}vv'(t+1/t)) = \sum_{m=-l}^{l} c^{lm}(taa')^m P_l^{(m,-m)}(u)P_l^{(m,-m)}(u'). \tag{5.5}$$

Recalling (4.10) and $\alpha(t) = uu' - \frac{1}{2}vv'(t+1/t)$ we have

$$p(zt) = P_l(uu' - \tfrac{1}{2}vv'(zt + (zt)^{-1})), \tag{5.6}$$

and hence

$$\int_0^1 t^{-i\gamma-1}p(zt)\,dt = \sum_{m=-l}^{l} c^{lm}(zaa')^m P_l^{(m,-m)}(u)P_l^{(m,-m)}(u') \times\int_0^1 t^{-i\gamma-1}t^m\,dt = \sum_{m=-l}^{l}\frac{1}{m-i\gamma}c^{lm}(zaa')^m P_l^{(m,-m)}(u)P_l^{(m,-m)}(u'). \tag{5.7}$$

In the four cases $z = a'/a,\ a/a',\ aa',\ (aa')^{-1}$, we have obtained a simple expression for the sum in Eq. (5.7).

By using the basic equality

$$\sum_{m=-l}^{l}\frac{-i\gamma}{m-i\gamma}c^{lm}P_l^{(m,-m)}(u)P_l^{(m,-m)}(u') = c_{l\gamma}P_l^{(i\gamma,-i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u'), \tag{5.8}$$

which we have proved in Ref. 13, we have obained[14]

$$\begin{aligned}\mathcal{F}_l = {} & F_{i\gamma}(aa')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u') \\ & + F_{i\gamma}((aa')^{-1})P_l^{(i\gamma,-i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u') \\ & - F_{i\gamma}(a'/a)P_l^{(i\gamma,-i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u') \\ & - F_{i\gamma}(a/a')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u').\end{aligned} \tag{5.9}$$

It is easy to verify that $\mathcal{F}_l$ has parity $(pp')^{l+1}$. We point out that T_{cl} has parity $(pp')^l$, i.e.,

$$(pp')^l\langle p|T_{cl}|p'\rangle \text{ is even in } p \text{ and in } p'. \tag{5.10}$$

6. THE DERIVATION OF $\mathcal{E}_l$

In this section we shall derive an explicit expression for $\mathcal{E}_l$, the function representing the "rational part" of T_{cl}. From this expression we shall derive many interesting properties of $\mathcal{E}_l$. According to Eqs. (4.1) and (4.9) we have to evaluate

$$-c_{l\gamma}\mathscr{E}_l = f_2(a'/a) + f_2(a/a') - f_2(aa') - f_2((aa')^{-1}), \tag{6.1}$$

where f_2 is given by Eq. (4.17),

$$f_2(z) = \int_0^1 dt\, t^{-i\gamma} \int_c^z \frac{d\tau}{\tau(1-\tau)} \frac{d}{dt}\{p(\tau t) - p(t)\}. \tag{6.2}$$

By inserting [cf. (5.5)]

$$p(t) = P_l(uu' - \tfrac{1}{2}vv'(t + 1/t)) = \sum_{m=-l}^{l} p_m t^m, \tag{6.3}$$

$$p_m = c^{lm}(aa')^m P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'), \tag{6.4}$$

into Eq. (6.2) we get (note $p_m = p_{-m}$)

$$\begin{aligned} f_2(z) &= \sum_{m=-l}^{l} p_m \int_0^1 mt^{m-1}t^{-i\gamma} dt \int_c^z \frac{d\tau}{\tau} \frac{\tau^m - 1}{1-\tau} \\ &= -\sum_{m=1}^{l} \frac{mp_m}{m - i\gamma} \int_c^z \sum_{n=0}^{m-1} \tau^{n-1} d\tau \\ &\quad + \sum_{m=1}^{l} \frac{mp_m}{m + i\gamma} \int_c^z \tau^{-m-1} \sum_{n=0}^{m-1} \tau^n d\tau, \end{aligned} \tag{6.5}$$

where $\Sigma_{m=-l}^{l}$ has been split into $\Sigma_{m=1}^{l} + \Sigma_{m=-1}^{-l}$, and the symmetry $p_m = p_{-m}$ has been used. Hence

$$\begin{aligned} f_2(z) = &-\sum_{m=1}^{l} \frac{mp_m}{m - i\gamma}\left[\ln(z/c) + \sum_{n=1}^{m-1}(z^n - c^n)/n\right] \\ &+ \sum_{m=1}^{l} \frac{mp_m}{m + i\gamma} \sum_{n=0}^{m-1}(z^{n-m} - c^{n-m})/(n-m). \end{aligned} \tag{6.6}$$

In view of Eq. (6.1) the terms involving the constant c and the term involving $\ln z$ will vanish when we take the combination of the f_2's with different arguments as indicated. By introducing the new summation variable $\nu = m - n$ in Σ_n on the second line of Eq. (6.6), we get

$$\begin{aligned} f_2(z) = &-\sum_{m=1}^{l} \frac{mp_m}{m - i\gamma} \sum_{n=1}^{m-1} z^n/n - \sum_{m=1}^{l} \frac{mp_m}{m + i\gamma} \\ &\times \sum_{\nu=1}^{m} z^{-\nu}/\nu + \text{irrelevant terms}. \end{aligned} \tag{6.7}$$

By rearranging and separating off the terms with $\nu = m$, we have obtained from this expression,[14]

$$\begin{aligned} c_{l\gamma}\mathscr{E}_l = &-2\sum_{m=2}^{l} \frac{m^2 p_m}{m^2 + \gamma^2} \sum_{n=1}^{m-1} \frac{1}{n}(a^n - a^{-n})(a'^n - a'^{-n}) \\ &- \sum_{m=1}^{l} \frac{m - i\gamma}{m^2 + \gamma^2} p_m (a^m - a^{-m})(a'^m - a'^{-m}), \end{aligned} \tag{6.8}$$

where p_m is given by (6.4). It follows that $\mathscr{E}_l$ has parity p^{l+1} in p.

We assume for the moment that p, p', k, and γ are real. Then $\operatorname{Re}\mathscr{E}_l$ is even in γ as is easily seen from (6.8). It is interesting and useful to have a simple explicit expression for $\operatorname{Im}\mathscr{E}_l$. According to Eq. (6.8),

$$c_{l\gamma}\operatorname{Im}\mathscr{E}_l = \gamma \sum_{m=1}^{l} \frac{p_m}{m^2 + \gamma^2}(a^m - a^{-m})(a'^m - a'^{-m}). \tag{6.9}$$

By using our basic relation (5.8) we have obtained the following simple and elegant result[14]:

$$\operatorname{Im}\mathscr{E}_l = (-2/\gamma)\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u)\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u'). \tag{6.10}$$

Furthermore, we want a simple expression for $\mathscr{E}_l$ in the particular case when $p' = k$. Then $a' = 0$ and $u' = 1$. From Eq. (6.8) we have

$$\begin{aligned} &c_{l\gamma}\mathscr{E}_l(p' = k) \\ &= \sum_{m=1}^{l} \frac{1}{m + i\gamma} c^{lm}(a^{2m} - 1) P_l^{(m,-m)}(u) P_l^{(m,-m)}(1). \end{aligned} \tag{6.11}$$

By applying Eq. (5.8) we have obtained,[14]

$$\mathscr{E}_l(p' = k) = \frac{2}{\gamma}\binom{l - i\gamma}{l}\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u). \tag{6.12}$$

Finally, we conjecture that $(pp'k^{-2})^{l+1}\mathscr{E}_l$ is a *polynomial* of degree l, both in $(p/k)^2$ and in $(p'/k)^2$, separately.

7. THE DERIVATION OF $\mathscr{L}_l$

In this section we shall derive an explicit expression for the function $\mathscr{L}_l$ representing the logarithmic part of T_{cl}, cf. Eq. (4.1). From this expression we shall derive many interesting properties.

According to Eqs. (4.9) and (4.15) we have to evaluate

$$f_3(a'/a) + f_3(a/a') - f_3(aa') - f_3((aa')^{-1}), \tag{7.1}$$

where f_3 is given by Eq. (4.18),

$$f_3(z) = \int_c^z \frac{d\tau}{\tau(1-\tau)} \int_0^1 t^{-i\gamma} p'(t)\, dt. \tag{7.2}$$

Clearly this double integral is equal to the product of the separate integrals. The integration $\int d\tau$ can be carried out by elementary means. It turns out that neither c nor $\ln z$ occurs in the final answer, due to cancellations of corresponding terms in (7.1). In conclusion, the integral $\int^z d\tau$ gives in the final result a factor which is equal to

$$\begin{aligned} &-\ln(1 - a'/a) - \ln(1 - a/a') + \ln(1 - aa') + \ln(1 - (aa')^{-1}) \\ &= \ln[(1 - aa')(a - a')^{-1}]^2 = \ln[(p + p')/(p - p')]^2. \end{aligned} \tag{7.3}$$

Here the cut structure of the logarithm has been taken into account. We assume $\operatorname{Im} k \downarrow 0$, $\operatorname{Re} k > 0$, $p \neq k$, $p' \neq k$, $p \neq p'$.

Hence, on account of Eqs. (4.1), (4.9), (7.2), and (7.3) we have

$$c_{l\gamma}\mathscr{L}_l = -\int_0^1 t^{-i\gamma} \frac{d}{dt} P_l(uu' - \tfrac{1}{2}vv'(t + 1/t))\, dt. \tag{7.4}$$

We substitute Eq. (5.5) into Eq. (7.4). By using (cf. Sec. 2)

$$p_m = p_{-m} = c^{lm}(aa')^m P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'), \tag{7.5}$$

we obtain

$$p(t) = P_l(uu' - \tfrac{1}{2}vv'(t + 1/t)) = \sum_{m=-l}^{l} p_m t^m; \tag{7.6}$$

hence,[14]

$$\begin{aligned} c_{l\gamma}\mathscr{L}_l = &-2\sum_{m=1}^{l} \frac{m^2}{m^2 + \gamma^2} \frac{(-aa')^m (l!)^2}{(l-m)!(l+m)!} \\ &\times P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'). \end{aligned} \tag{7.7}$$

In order to rewrite (7.7), and to derive simple expressions for $\mathscr{L}_l$ in the particular cases (i) $\gamma \to 0$, (ii) $p' \to p$, and (iii) $p' \to \infty$, we shall first give some useful and interesting relations involving Jacobi polynomials.

From Eq. (5.8) we have

$$c_{l\gamma}P_l^{(-i\gamma,i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u')$$
$$= \sum_{m=-l}^{l} \frac{i\gamma}{m+i\gamma} c^{lm} P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'), \tag{7.8}$$

and from Eq. (5.5),

$$P_l(uu' - \tfrac{1}{2}vv'(t+1/t))$$
$$= \sum_{m=-l}^{l} c^{lm}(taa')^m P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'). \tag{7.9}$$

By taking $t = aa'$ and $t = a'/a$, respectively, in (7.9) we get

$$P_l(-1) = (-)^l = \sum_{m=-l}^{l} c^{lm} P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'), \tag{7.10}$$

$$P_l(1) = 1 = \sum_{m=-l}^{l} c^{lm} P_l^{(m,-m)}(u) P_l^{(-m,m)}(u'). \tag{7.11}$$

By taking $t = 1$ in (7.9) we get [recall $w = (p^2 + p'^2)/2pp'$]

$$P_l(w) = \sum_{m=-l}^{l} c^{lm}(aa')^m P_l^{(m,-m)}(u) P_l^{(m,-m)}(u'), \tag{7.12}$$

i.e.,

$$P_l(w) = \sum_{m=-l}^{l} p_m = \sum_{m=0}^{l} \epsilon_m p_m. \tag{7.13}$$

Because of

$$a^m P_l^{(m,-m)}(u) = a^{-m} P_l^{(-m,m)}(u), \tag{7.14}$$

we have

$$P_l(w) = P_l(u)P_l(u') + 2\sum_{m=1}^{l} p_m. \tag{7.15}$$

Therefore we can rewrite Eq. (7.7) by putting $m^2 = m^2 + \gamma^2 - \gamma^2$,

$$c_{l\gamma}\mathscr{L}_l = P_l(u)P_l(u') - P_l(w) + 2\gamma^2 \sum_{m=1}^{l} p_m (m^2+\gamma^2)^{-1}. \tag{7.16}$$

By taking the limit for $\gamma \to 0$ in Eq. (7.7) we obtain

$$\lim_{\gamma\to 0} \mathscr{L}_l = -\sum_{m\neq 0} p_m = p_0 - \sum_{m=-l}^{l} p_m.$$

Hence,

$$\gamma = 0 \Rightarrow \mathscr{L}_l = P_l(u)P_l(u') - P_l(w). \tag{7.17}$$

By taking the limit for $p' \to p$, i.e., $a' \to a$, in Eq. (7.7) we get

$$\lim_{p'\to p} c_{l\gamma}\mathscr{L}_l$$
$$= -\sum_{m=-l}^{l} \frac{ma^{2m}}{m-i\gamma} c^{lm}\{P_l^{(m,-m)}(u)\}^2$$
$$= -1 + (-)^l \sum_{m=-l}^{l} \frac{-i\gamma}{m-i\gamma}$$
$$\times c^{lm} P_l^{(m,-m)}(u) P_l^{(m,-m)}(-u)$$
$$= -1 + (-)^l c_{l\gamma} P_l^{(i\gamma,-i\gamma)}(u) P_l^{(i\gamma,-i\gamma)}(-u)$$
$$= -1 + c_{l\gamma} P_l^{(i\gamma,-i\gamma)}(u) P_l^{(-i\gamma,i\gamma)}(u).$$

Hence,

$$p' = p \Rightarrow \mathscr{L}_l = P_l^{(i\gamma,-i\gamma)}(u) P_l^{(-i\gamma,i\gamma)}(u) - c_{l\gamma}^{-1}. \tag{7.18}$$

Now we shall establish a simple relation between $\lim_{p'\to\infty}(p')^{-l}\mathscr{L}_l$ and the polynomial A_l introduced in Ref. 15.

By using

$$\lim_{z\to\infty} z^{-l} P_l^{(\alpha,-\alpha)}(z) = 2^{-l}\binom{2l}{l},$$

we obtain

$$\lim_{p'\to\infty} (4p/p')^l \mathscr{L}_l$$
$$= -c_{l\gamma}^{-1}(p/k)^l \binom{2l}{l}$$
$$\times \sum_{m=0}^{l} \frac{\epsilon_m m^2}{m^2+\gamma^2} \frac{(-a)^m (l!)^2}{(l-m)!(l+m)!} P_l^{(m,-m)}(u). \tag{7.19}$$

By taking $p' \to \infty$ in both members of Eq. (7.12) we get

$$\sum_{m=-l}^{l} \frac{(-a)^m (l!)^2}{(l-m)!(l+m)!} P_l^{(m,-m)}(u) = (k/p)^l, \tag{7.20}$$

i.e.,

$$\sum_{m=0}^{l} \epsilon_m c^{lm} a^m P_l^{(m,-m)}(u) = (k/p)^l. \tag{7.21}$$

Writing $m^2 = m^2 + \gamma^2 - \gamma^2$ in Eq. (7.19) and using Eq. (7.20) or (7.21) we get

$$\lim_{p'\to\infty} (4p/p')^l \mathscr{L}_l$$
$$= c_{l\gamma}^{-1}\binom{2l}{l}$$
$$\times\Big[-1 + (p/k)^l\gamma^2 \sum_{m=0}^{l} \frac{\epsilon_m}{m^2+\gamma^2}$$
$$\times \frac{(-a)^m (l!)^2}{(l-m)!(l+m)!} P_l^{(m,-m)}(u)\Big].$$

By comparing this equation with Eq. (3.4a) of Ref. 16 we obtain

$$\lim_{p'\to\infty} (4p/p')^l \mathscr{L}_l = \binom{2l}{l}[A_l(p^2k^{-2};\gamma^2) - c_{l\gamma}^{-1}], \tag{7.22}$$

which is the desired relation with A_l.

Finally, we conjecture that $(pp'k^{-2})^l\mathscr{L}_l$ is a *polynomial* of degree l in $(pk^{-1})^2$ and in $(p'k^{-1})^2$ separately.

8. SUMMARY

In this section we shall give a summary of the most interesting formulas that we have derived in Secs. 3–7. For the notation one should consult Sec. 2.

$$\langle p|T_{cl}|p'\rangle$$
$$= \frac{2k\gamma}{\pi pp'} \int_{t=0}^{1} t^{i\gamma}\, d\mathfrak{Q}_l(uu' - \tfrac{1}{2}vv'(t+1/t)). \tag{8.1}$$

$$\lim_{p'\to\infty} p'^{l+2}\langle p|T_{cl}|p'\rangle$$
$$= \tfrac{1}{2}k\gamma(4k)^{l+1}(l!)^2[(2l+1)!]^{-1}$$
$$\times (p^2-k^2)\langle p|kl\uparrow\rangle_c f_{cl}^{-1}. \tag{8.2}$$

$$\langle p|T_{cl}|p'\rangle$$
$$= \frac{-k\gamma}{\pi pp'} c_{l\gamma}\Big[(i\gamma)^{-1}\mathscr{F}_l + \mathscr{C}_l + \mathscr{L}_l \ln\Big(\frac{p+p'}{p-p'}\Big)^2\Big], \tag{8.3}$$

where

$$\mathscr{F}_l = F_{i\gamma}(aa')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u') + F_{i\gamma}((aa')^{-1})P_l^{(i\gamma,-i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u') - F_{i\gamma}(a'/a)P_l^{(i\gamma,-i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u') - F_{i\gamma}(a/a')P_l^{(-i\gamma,i\gamma)}(u)P_l^{(i\gamma,-i\gamma)}(u'); \quad (8.4)$$

$$c_{l\gamma}\mathscr{G}_l = -2\sum_{m=2}^{l}\frac{m^2 p_m}{m^2+\gamma^2}\sum_{n=1}^{m-1}\frac{1}{n}(a^n - a^{-n})(a'^n - a'^{-n}) - \sum_{m=1}^{l}\frac{p_m}{m+i\gamma}(a^m - a^{-m})(a'^m - a'^{-m}), \quad (8.5)$$

where

$$p_m = \frac{(-)^m (l!)^2}{(l-m)!(l+m)!}(aa')^m P_l^{(m,-m)}(u)P_l^{(m,-m)}(u');$$

$$c_{l\gamma}\mathscr{L}_l = -2\sum_{m=1}^{l} m^2(m^2+\gamma^2)^{-1}p_m = P_l(u)P_l(u') - P_l(w) + 2\gamma^2\sum_{m=1}^{l}(m^2+\gamma^2)^{-1}p_m. \quad (8.6)$$

In particular,

$$\mathscr{G}_0 = 0,\quad \mathscr{G}_1 = 2(1-i\gamma),\quad \mathscr{G}_2 = 3(1-i\gamma)[(1+i\gamma)w - \tfrac{3}{2}i\gamma uu'], \quad (8.7)$$

$$\mathscr{L}_0 = 0,\quad \mathscr{L}_1 = vv',\quad \mathscr{L}_2 = \tfrac{3}{4}vv'[(1+\gamma^2)w + 3uu']. \quad (8.8)$$

Furthermore,

$\mathscr{G}_l$ is symmetric in (p,p'),

$(pp')^{l+1}\mathscr{G}_l$ is even in p and in p',

Re $\mathscr{G}_l$ is even in γ [for p, p', k and γ real],

$$\operatorname{Im}\mathscr{G}_l = -2\gamma^{-1}\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u)\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u'),$$

$$p' = k \Rightarrow \mathscr{G}_l = 2\gamma^{-1}\binom{l-i\gamma}{l}\operatorname{Im}P_l^{(i\gamma,-i\gamma)}(u).$$

Similarly, for $\mathscr{L}_l$,

$\mathscr{L}_l$ is symmetric in (p,p'),

$(pp')^l\mathscr{L}_l$ is even in p and in p',

$\mathscr{L}_l$ is even in γ,

Im $\mathscr{L}_l = 0$ [for p, p', k, and γ real],

$\mathscr{L}_l$ contains the factor $(p-k)(p'-k)$,

$$p' = k \Rightarrow \mathscr{L}_l = 0,$$

$$\gamma = 0 \Rightarrow \mathscr{L}_l = P_l(u)P_l(u') - P_l(w),$$

$$p' = p \Rightarrow \mathscr{L}_l = P_l^{(i\gamma,-i\gamma)}(u)P_l^{(-i\gamma,i\gamma)}(u) - c_{l\gamma}^{-1},$$

$$\lim_{p'\to\infty}(4p/p')^l\mathscr{L}_l = \binom{2l}{l}[A_l(p^2/k^2;\gamma^2) - c_{l\gamma}^{-1}].$$

Since

$(pp')^{l+1}\mathscr{F}_l$ is even in p and in p',

we have

$(pp')^l\langle p|T_{cl}|p'\rangle$ is even in p and in p'.

Finally we conjecture that, considered as functions of $(pk^{-1})^2$ and of $(p'k^{-1})^2$ separately, for γ fixed, $(pp'k^{-2})^{l+1}\mathscr{G}_l$ is a *polynomial* of degree l, and $(pp'k^{-2})^l\mathscr{L}_l$ is a *polynomial* of degree l.

[1]C. Chandler, Nucl. Phys. A **353**, 129c–142c (1981).

[2]L. P. Kok, Nucl. Phys. A **353**, 171c–184c (1981).

[3]We work in the momentum representation throughout; hence T_c and T_{cl} stand for $\langle\mathbf{p}|T_c|\mathbf{p}'\rangle$ and $\langle p|T_{cl}|p'\rangle$, respectively.

[4]J. Schwinger, J. Math. Phys. **5**, 1606 (1964).

[5]L. Hostler, J. Math. Phys. **5**, 1235 (1964).

[6]H. van Haeringen and R. van Wageningen, J. Math. Phys. **16**, 1441 (1975); see also H. van Haeringen, J. Math. Phys. **18**, 927 (1977).

[7]L. P. Kok and H. van Haeringen, Phys. Rev. C **21**, 512 (1980).

[8]H. van Haeringen, Phys. Lett. A **86**, 359 (1981).

[9]H. van Haeringen, Nucl. Phys. A **327**, 77 (1979).

[10]H. van Haeringen, "Long-Range Potentials in Quantum Mechanics," Report 150; "Notations and Conventions," Report 150* (University of Groningen, 1979).

[11]H. van Haeringen, "The Coulomb Jost states in the momentum representation for all partial waves in closed form," preprint (Delft, 1981).

[12]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, New York, 1966).

[13]H. van Haeringen, "Some infinite series involving products of Legendre and gamma functions," preprint (Delft, 1981), in particular Eq. (10).

[14]Following the referee's advice, I have omitted some details of derivations. Full derivations of explicit closed expressions for $\mathscr{F}_l$, $\mathscr{G}_l$, and $\mathscr{L}_l$ can be found in H. van Haeringen, "The partial-wave projected Coulomb T matrix for all l in closed hypergeometric form," Report 8105 (Delft University of Technology, Delft, 1981).

[15]H. van Haeringen, J. Math. Phys. **19**, 1379 (1978).

[16]H. van Haeringen, J. Math. Phys. **20**, 1109 (1979).

Off-shell T matrix for Coulomb plus simple separable potentials for all l in closed form

H. van Haeringen
Department of Mathematics, Delft University of Technology, Delft, The Netherlands

(Received 18 November 1981; accepted for publication 19 March 1982)

We derive simple exact analytic expressions for the off-shell T matrices associated with potentials consisting of the sum of the Coulomb potential and any so-called simple separable potential, for all partial waves. Such potentials play an important role in model calculations for the description of the interaction between charged particles, e.g., protons. We also derive some interesting properties of the Coulomb-modified form factors associated with the form factors of the separable potentials. The analytic expressions obtained are useful for theoretical and numerical applications.

PACS numbers: 11.20.Fm, 03.65.Nk

1. INTRODUCTION

The Coulomb interaction plays an important and, especially from a mathematical point of view, interesting role in the theory of scattering by charged particles. In two-particle reactions only *on-shell* quantities are relevant, for instance, two-particle scattering amplitude and wave functions. On the other hand, in $n(>2)$-particle reactions one needs *off-shell* quantities, in particular the off-shell transition (T) matrix. In many realistic potential models the interaction between two charged particles consists of the sum of the Coulomb potential V_c and a short-range potential V_s. In model calculations one often takes for V_s a separable potential.

In this paper we shall derive explicit expressions for the off-shell T matrix associated with a potential $V = V_c + V_s$ where V_s is a so-called simple separable potential. In the particular partial-wave space characterized by l this potential is given by

$$V_{sl} = -\lambda_l |g_{\beta l}\rangle\langle g_{\beta l}|, \tag{1}$$

$$\langle p|g_{\beta l}\rangle = (2/\pi)^{1/2} p^l (p^2+\beta^2)^{-l-1}.$$

Preliminary results have been reported in Ref. 1.

According to Ref. 2 we need explicit expressions for (i) the pure Coulomb T matrix, $\langle p|T_{cl}|p'\rangle$, (ii) the Coulomb-modified form factor

$$\langle p|g^c_{\beta l}\rangle \equiv \langle p|G_{0l}^{-1}G_{cl}|g_{\beta l}\rangle, \tag{2}$$

and (iii)

$$\langle g_{\alpha l}|G_{0l}|g^c_{\beta l}\rangle \equiv \langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle. \tag{3}$$

G_{0l} is the free resolvent and G_{cl} is the Coulomb resolvent. Here and henceforth we suppress the energy dependence of these operators and of the T operator; cf. Sec. 2.

For the rank-one potential given by Eq. (1) we need an expression for (3) only in the special case $\alpha = \beta$. However, when we take a potential of higher rank with the same simple form factors given by Eq. (1) but with different β's, we need an expression for (3) with $\alpha \neq \beta$.

For $\langle p|T_{cl}|p'\rangle$ an integral representation is known[2] which is convenient for numerical calculations in some cases only, see, e.g., Ref. 3. Furthermore, an infinite-series expression is known for $\langle p|T_{cl}|p'\rangle$. We shall use this representation as a starting point for our derivations [see Eq. (12)]. However, this series representation is not convenient for numerical work; see Ref. 4. We note in particular that the series is divergent when the energy is positive. Recently we have obtained an expression for $\langle p|T_{cl}|p'\rangle$ in terms of hypergeometric functions.[5] This expression is suitable for both numerical and analytic work.

The main result of this paper consists of explicit expressions for the Coulomb-modified form factor $\langle p|g^c_{\beta l}\rangle$; see Eqs. (29), (30), and (33) ff. As a byproduct we also obtain a short derivation of a simple expression for $\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle$; see Eq. (22). We note that Maleki and Macek[6] have derived in a different way essentially the same expression as given by Eq. (22).

The expression for $\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle$ is much simpler than the expression for $\langle p|g^c_{\beta l}\rangle$, which in turn is much simpler than the expression for $\langle p|T_{cl}|p'\rangle$ obtained in Ref. 5. We contend that the complexity of these expressions is inherent in these quantities so that at most slight simplifications are possible; cf. Ref. 7.

In Sec. 2 we give the notations we shall use in this paper. Representations in terms of so-called Sturm states will be derived in Sec. 3. In Sec. 4 we shall derive hypergeometric-function expressions and integral representations.

In Eq. (29) we introduce a function Z_l, which is convenient for the derivation of a closed formula for the Coulomb-modified form factor $g^c_{\beta l}$

$$\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle - kp^{-1}\langle k|g_{\beta l}\rangle Z_l(a;B).$$

This function Z_l in turn is expressed in terms of the hypergeometric function ${}_2F_1(1,i\gamma;1+i\gamma;)$, Jacobi polynomials, and a certain rational function X_l (which can also be rewritten as a polynomial); see Eq. (42).

In Sec. 5 we derive a simple expression for $\operatorname{Im} Z_l(a;B)$ and we prove that $X_l(a;B)$ is real. Finally, in Sec. 6 we give a summary of the most important formulas obtained in this paper.

2. NOTATIONS

We shall use the same notations and conventions as in previous related work.[1,2,5,8] In particular, we choose units

 0022-2488/83/051274-08$02.50

such that $\hbar = 1 = 2m$ and we denote the energy by $E \equiv k^2 \equiv -\kappa^2$ where $k \equiv i\kappa$. The energy dependence of the operators G_{0l}, G_{cl}, and T_{cl}, and of the Coulomb-modified form factor $g^c_{\beta l}$ is suppressed. The Coulomb potential is given by

$$V_c(r) = 2k\gamma/r \equiv -2s/r,$$

where γ is Sommerfeld's parameter. The momentum variables p and p' are real positive. We shall often also take k real positive but most of the results are valid for *complex k*, as will be clear from the context.

We shall use the abbreviations

$$a = (p-k)/(p+k), \quad a' = (p'-k)/(p'+k),$$
$$u = (p^2+k^2)/(2pk), \quad u' = (p'^2+k^2)(2p'k),$$
$$v = (p^2-k^2)/(2pk), \quad v' = (p'^2-k^2)(2p'k),$$
$$A = (\alpha+ik)/(\alpha-ik), \quad B = (\beta+ik)/(\beta-ik),$$
$$f_{cl} = f_{cl}(k) = e^{\pi\gamma/2} l!/\Gamma(l+1+i\gamma),$$
$$F_{i\gamma}(z) = {}_2F_1(1, i\gamma; 1+i\gamma; z).$$

Furthermore, C^{α}_n, $P^{(\alpha,\beta)}_n$, and $L^{(\alpha)}_n$ denote the Gegenbauer, Jacobi, and Laguerre polynomials, respectively. In the literature there exist two different normalizations of the Laguerre polynomials. We shall use

$$L^{(\alpha)}_n(z) = \binom{n+\alpha}{n} {}_1F_1(-n; \alpha+1; z).$$

3. STURM SERIES

In this section we shall derive infinite-series expressions for $\langle p | g^c_{\beta l} \rangle$ and $\langle g_{\alpha l} | G_{cl} | g_{\beta l} \rangle$ by using the so-called Sturm states $|\lambda_n l\rangle$, $n = l+1, l+2, \ldots$. These states are defined as eigenstates of $V_l G_{0l}$ (cf. Ref. 8, p. 106)

$$V_l G_{0l}(-\kappa^2) |\lambda_n l\rangle = \lambda_n |\lambda_n l\rangle, \tag{4}$$

where the normalization is given by

$$\langle \lambda_{n'} l | G_{0l} | \lambda_n l \rangle = -\delta_{n'n}. \tag{5}$$

Note that the energy, $-\kappa^2 < 0$, is fixed here. When V_l is the Coulomb potential, the eigenvalues λ_n are given by

$$\lambda_n = s/(n\kappa), \quad n = l+1, l+2, \ldots. \tag{6}$$

There exists a simple relation with the Coulomb bound states $|\kappa_n l\rangle$,

$$|\kappa_n l\rangle = 2^{1/2}\kappa G_{0l} |\lambda_n l\rangle. \tag{7}$$

The bound states are eigenstates of $H_{0l} + V_{cl}$ with eigenvalues $-\kappa^2 = -s^2/n^2$.

Completeness of the Sturm states can be expressed formally by

$$-G_{0l}^{-1} = \sum_{n=l+1}^{\infty} |\lambda_n l\rangle\langle\lambda_n l|. \tag{8}$$

It follows that

$$T_{cl} = \sum_{n=l+1}^{\infty} \frac{1}{1 - n\kappa/s} |\lambda_n l\rangle\langle\lambda_n l|, \tag{9}$$

and

$$G_{cl} = \sum_{n=l+1}^{\infty} \frac{-n}{n - s/\kappa} G_{0l} |\lambda_n l\rangle \langle \lambda_n l | G_{0l}. \tag{10}$$

In the momentum representation we have [cf. Eq. (6.16) of Ref. 8]

$$\begin{aligned}\langle p | \lambda_n l \rangle &= -2^{-1/2}\kappa^{-1}(p^2+\kappa^2)\langle p | \kappa_n l \rangle \\ &= -(n\kappa/\pi)^{1/2} p^{-1} l! [(n-l-1)!/(n+l)!]^{1/2} \\ &\quad \times [4p\kappa/(p^2+\kappa^2)]^{l+1} C^{l+1}_{n-l-1}(u/v).\end{aligned} \tag{11}$$

From Eqs. (9) and (11) we obtain

$$\begin{aligned}&\langle p | T_{cl} | p' \rangle \\ &= \frac{-s(l!)^2}{\pi p p'} \left(\frac{-4}{vv'}\right)^{l+1} \sum_{n=l+1}^{\infty} \frac{n}{n - s/\kappa} \frac{\Gamma(n-l)}{\Gamma(n+l+1)} \\ &\quad \times C^{l+1}_{n-l-1}(u/v) C^{l+1}_{n-l-1}(u'/v'),\end{aligned} \tag{12}$$

which expression is well known.[4] We note that κ is supposed to be positive here. For positive energy the series in Eq. (12) diverges.

Sloan[4] proved that the series in Eq. (12) is, for negative energy, at most conditionally convergent, and sometimes divergent (e.g., for $p = p'$). A useful device to obtain absolute convergence is to multiply the terms of this series by ρ^n, $0 < \rho < 1$. Then we may integrate the series termwise. Afterwards the limit $\rho \uparrow 1$ is taken. When interchanging integration and summation in the following, we shall always tacitly assume that this procedure is followed. According to Ref. 19 of Sloan,[4] we can attach a meaning to the parameter ρ: It can be thought of as the radial distance in the four-dimensional space in which Fock and Schwinger (Ref. 9) embedded ordinary momentum space to reveal the hidden symmetry of the Coulomb Hamiltonian.

From Eqs. (2) and (10) we have

$$\langle p | g^c_{\beta l} \rangle = \sum_{n=l+1}^{\infty} \frac{-n}{n - s/\kappa} \langle p | \lambda_n l \rangle \langle \lambda_n l | G_{0l} | g_{\beta l} \rangle. \tag{13}$$

In order to evaluate $\langle \lambda_n l | G_{0l} | g_{\beta l} \rangle$ in closed form, we use Eq. (7)

$$\begin{aligned}2^{1/2}\kappa \langle \lambda_n l | G_{0l} | g_{\beta l} \rangle &= \langle \kappa_n l | g_{\beta l} \rangle \\ &= \int_0^{\infty} \langle \kappa_n l | r \rangle \langle r | g_{\beta l} \rangle r^2 dr,\end{aligned}$$

and insert[8]

$$\begin{aligned}\langle \kappa_n l | r \rangle = i^{-l} 2\kappa(\kappa/n)^{1/2} [(n-l-1)!/(n+l)!]^{1/2} \\ \times (2\kappa r)^l e^{-\kappa r} L^{(2l+1)}_{n-l-1}(2\kappa r)\end{aligned} \tag{14}$$

and

$$\langle r | g_{\beta l} \rangle = r^{l-1} e^{-\beta r} (i/2)^l / l!. \tag{15}$$

It should be noted that in Eq. (14) κ is assumed to be fixed. By using (Ref. 10, p. 244)

$$\int_0^{\infty} e^{-zt} t^{a} L^{(a)}_m(t) dt = (m+1)_a z^{-m-a-1}(z-1)^m \tag{16}$$

we get

$$\begin{aligned}\langle \lambda_n l | G_{0l} | g_{\beta l} \rangle = (2\kappa/n)^{1/2} [(n+l)!/(n-l-1)!]^{1/2} \\ \times \kappa^l (l!)^{-1} (\beta+\kappa)^{-n-l-1} (\beta-\kappa)^{n-l-1}.\end{aligned} \tag{17}$$

Substitution of (11) and (17) into (13) gives

$$\langle p|g^c_{\beta l}\rangle = (2/\pi)^{1/2}p^{-1}[iv(\beta^2-\kappa^2)/(2\kappa)]^{-l-1} \times \sum_{n=l+1}^{\infty} \frac{nB^n}{n-s/\kappa} C^{l+1}_{n-l-1}(u/v)$$
$$= \frac{(2/\pi)^{1/2}}{p}\left[\frac{4p\kappa^2}{(p^2+\kappa^2)(\beta^2-\kappa^2)}\right]^{l+1} \times \sum_{n=l+1}^{\infty}\frac{n}{n-s/\kappa}\left(\frac{\beta-\kappa}{\beta+\kappa}\right)^n \times C^{l+1}_{n-l-1}\left(\frac{p^2-\kappa^2}{p^2+\kappa^2}\right). \quad (18)$$

One easily verifies from this expression that

$$\lim_{s\to 0}\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle = (2/\pi)^{1/2}p^l(\beta^2+p^2)^{-l-1},$$

by using the equality (cf. Ref. 10, p. 222)

$$\sum_{m=0}^{\infty} z^m C^\lambda_m(x) = (1-2xz+z^2)^{-\lambda}, |z|<1. \quad (19)$$

In a similar way we obtain from Eq. (10)

$$\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle = \sum_{n=l+1}^{\infty}\frac{-n}{n-s/\kappa}\langle g_{\alpha l}|G_{0l}|\lambda_n l\rangle\langle\lambda_n l|G_{0l}|g_{\beta l}\rangle. \quad (20)$$

4. HYPERGEOMETRIC-FUNCTION REPRESENTATIONS

In this section we shall express $\langle p|g^c_{\beta l}\rangle$ and $\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle$ in terms of hypergeometric functions and elementary functions, by using Eqs. (18) and (20), respectively. We shall first discuss $\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle$ because the derivation for this quantity is relatively easy.

By inserting Eq. (17) into Eq. (20) we get

$$\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle = \frac{2}{\kappa(l!)^2}\left[\frac{\kappa^2}{(\alpha^2-\kappa^2)(\beta^2-\kappa^2)}\right]^{l+1} \times \sum_{n=l+1}^{\infty}\frac{-1}{n-s/\kappa}\frac{(n+l)!}{(n-l-1)!}\left(\frac{\alpha-\kappa}{\alpha+\kappa}\frac{\beta-\kappa}{\beta+\kappa}\right)^n. \quad (21)$$

The infinite sum in Eq. (21) can be rewritten in terms of one hypergeometric function. After a few manipulations we obtain

$$\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle = \frac{-2(2l+1)}{l+1-s/\kappa}\frac{(2\alpha+2\beta)^{-2l-1}}{(\alpha+\kappa)(\beta+\kappa)}\binom{2l}{l} \times {}_2F_1(1,-l-s/\kappa;l+2-s/\kappa;AB). \quad (22)$$

This is a remarkably simple and elegant result. Moreover, this ${}_2F_1$-expression may be analytically continued from real positive κ into the complex κ plane. Hence, replacing κ by $-ik$ and $-s/\kappa$ by $i\gamma$ we obtain from Eq. (22)

$$\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle = \frac{(2l+1)!!}{(2l)!!}\frac{(\alpha+\beta)^{-2l-1}}{(\alpha-ik)(\beta-ik)}\frac{-1}{l+1+i\gamma} \times {}_2F_1(1,i\gamma-l;i\gamma+l+2;AB). \quad (23)$$

We have applied this formula in the particular case $\alpha=\beta$ for the derivation of phase shifts and scattering parameters for Coulomb plus simple separable potentials.[11]

Now we are going to evaluate expressions for $\langle p|g^c_{\beta l}\rangle$. We shall derive an integral representation [Eq. (26)] and an expression involving hypergeometric functions; see Eqs. (29), (33), (38), and (39).

From Eq. (18) we have, denoting κ by $-ik$,

$$\langle p|g^c_{\beta l}\rangle = \frac{(2/\pi)^{1/2}}{p}\left(\frac{-4pk^2}{(p^2-k^2)(\beta^2+k^2)}\right)^{l+1}\Sigma, \quad (24)$$

with

$$\Sigma = \sum_{n=0}^{\infty}\frac{n+l+1}{n+l+1+i\gamma}B^{l+1+n}C^{l+1}_n(u/v).$$

By inserting

$$C^{l+1}_n(u/v) = \frac{\frac{1}{2}}{(l!)^2}\sum_{m=0}^{n}\frac{(l+m)!(l+n-m)!}{m!(n-m)!}(a^{n-2m}+a^{2m-n}),$$

resumming, and using

$${}_2F_1(l+1,m+l+1+i\gamma;m+l+2+i\gamma;Ba) = (m+l+1+i\gamma)\int_0^1 t^{m+l+i\gamma}(1-tBa)^{-l-1}dt,$$

we arrive at

$$\Sigma = B\frac{\partial}{\partial B}\int_0^1 t^{l+i\gamma}[B^{-1}+Bt^2-t(a+1/a)]^{-l-1}dt = (l+1)B^{-i\gamma}\int_0^B(1-t^2)t^{l+i\gamma}[1+t^2-t(a+1/a)]^{-l-2}dt. \quad (25)$$

Integration by parts gives

$$\Sigma = B^{-i\gamma}\int_0^B t^{i\gamma}(l+1)(t^{-2}-1)[t+t^{-1}-a-a^{-1}]^{-l-2}dt$$
$$= B^{-i\gamma}[t^{i\gamma}(t+t^{-1}-a-a^{-1})^{-l-1}]_0^B - i\gamma B^{-i\gamma}\int_0^B t^{i\gamma+l}(t^2+1-t(a+a^{-1}))^{-l-1}dt$$
$$= (B+B^{-1}-a-a^{-1})^{-l-1} - i\gamma B^{-i\gamma}\int_0^1 B^{l+1+i\gamma}t^{l+i\gamma}[(1-tBa)(1-tB/a)]^{-l-1}dt$$
$$= \left\{\frac{-4k^2(p^2+\beta^2)}{(\beta^2+k^2)(p^2-k^2)}\right\}^{-l-1} - i\gamma\int_0^1 t^{i\gamma-1}(tB+(tB)^{-1}-a-a^{-1})^{-l-1}dt.$$

By substituting this into Eq. (24) we obtain

$$\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle - i\gamma(2/\pi)^{1/2}p^l\left\{\frac{-4k^2}{(\beta^2+k^2)(p^2-k^2)}\right\}^{l+1} \times\int_0^1 t^{i\gamma-1}(tB+(tB)^{-1}-a-a^{-1})^{-l-1}dt. \quad (26)$$

The behavior for $\beta\to\infty$ of this expression is easily obtained:

$$\lim_{\beta\to\infty}(\pi/2)^{1/2}\beta^{2l+2}\langle p|g^c_{\beta l}\rangle = p^l - i\gamma p^{-1}(-2k/v)^{l+1} \times\int_0^1\frac{t^{l+i\gamma}dt}{[(1-ta)(1-t/a)]^{l+1}}.$$

According to Eqs. (6) and (7) of Ref. 12 we therefore have

$$\lim_{\beta\to\infty} (\pi/2)^{1/2}\beta^{2l+2}\langle p|g^c_{\beta l}\rangle = p^l - \tfrac{1}{2}\pi k^{l+1}\langle p|V_{cl}|kl\uparrow\rangle_c f_{cl}^{-1} = \tfrac{1}{2}\pi k^{l+1}(p^2-k^2)\langle p|kl\uparrow\rangle_c f_{cl}^{-1}, \tag{27}$$

because of

$$\langle p|V_{cl}|kl\uparrow\rangle_c = (k^2-p^2)\langle p|kl\uparrow\rangle_c + 2(\pi k)^{-1}(p/k)^l f_{cl}. \tag{28}$$

We rewrite Eq. (26) as follows:

$$\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle - (2/\pi)^{1/2}p^{-1}\left(\frac{k}{\beta^2+k^2}\right)^{l+1} Z_l(a;B), \tag{29}$$

with

$$Z_l(a;B) = B^{l+1}(a-1/a)^{l+1} i\gamma\int_0^1 \frac{t^{i\gamma+l}dt}{[(1-tBa)(1-tB/a)]^{l+1}}. \tag{30}$$

In order to evaluate the integral in (30) we perform a number of nontrivial manipulations. In the Appendix we shall prove

$$i\gamma\int_0^1 t^{l+i\gamma}[(1-ta)(1-tb)]^{-l-1}dt = f_l(a;b) + f_l(b;a), \tag{31}$$

where

$$\begin{aligned} f_l(a;b) &= (a-b)^{-l-1}F_{i\gamma}(a)P_l^{(-i\gamma,i\gamma)}\left(\frac{b+a}{b-a}\right) \\ &\quad - i\gamma(-ab)^l(a-b)^{-2l-1} \\ &\quad \times \sum_{n=0}^{l} \binom{2l-n}{l}\left(\frac{b-a}{b}\right)^n \frac{\Gamma(l+1+i\gamma)}{n!} \\ &\quad \times\left[\frac{a^{-l}(-)^n}{\Gamma(l+1+i\gamma-n)}\sum_{\nu=0}^{l-n}\frac{a^\nu}{\nu+i\gamma}\right. \\ &\quad \left. -(1-a)^{-n}\sum_{\mu=0}^{n-1}\frac{\Gamma(n-\mu)}{\Gamma(l+1+i\gamma-\mu)}\left(\frac{a-1}{a}\right)^\mu\right]. \end{aligned} \tag{32}$$

It may be noted that $\lim_{\gamma\to 0} f_l(a;b) = 0$. This is verified from Eq. (32) by using the equality

$$P_l\left(\frac{b+a}{b-a}\right) = (-1)^l {}_2F_1\left(-l,l+1;1;\frac{b}{b-a}\right).$$

By inserting (31) and (32) into (30) we find an explicit expression for $Z_l(a;B)$. It turns out to be convenient to reshape the second term in the right member of (32). In particular, we split off the part emerging from $\nu = 0$. After some simple manipulations we find

$$\begin{aligned} Z_l(a;B) &= [F_{i\gamma}(Ba) - 1]P_l^{(-i\gamma,i\gamma)}(u) \\ &\quad - i\gamma\binom{2l}{l}(1-a^2)^{-l}\sum_{\nu=1}^{l-1}(Ba)^\nu/(\nu+i\gamma) \\ &\quad + (-1)^{l-1}i\gamma\Gamma(l+1+i\gamma) \\ &\quad \times\sum_{m=0}^{l-1}\binom{l+m}{l}(a^2-1)^{-m}[(l-m)!]^{-1} \\ &\quad \times\left[\frac{1}{\Gamma(m+1+i\gamma)}\sum_{\nu=1}^{m}\frac{(Ba)^\nu}{\nu+i\gamma}\right. \\ &\quad - (Ba)^m\sum_{\mu=0}^{l-m-1}\frac{\mu!}{\Gamma(\mu+m+2+i\gamma)} \\ &\quad \left.\times\left(\frac{Ba}{Ba-1}\right)^{\mu+1}\right] \\ &\quad -(-1)^l\{\text{Idem},\ a\to a^{-1}\}. \end{aligned} \tag{33}$$

The term $-P_l^{(-i\gamma,i\gamma)}(u)$ stems from the $\nu = 0$ contribution, and the second term on the right-hand side from the $m = l$ contribution. By $\{\text{Idem}, a\to a^{-1}\}$ we mean that, after the replacement of a by a^{-1} everywhere, all the foregoing expressions on the right-hand side should be repeated. It should be noted that the term originating from $\nu = l$ cancels by this operation.

The expression for $\langle p|g^c_{\beta l}\rangle$ given by Eqs. (29) and (33) can be derived in a slightly different way, as we shall show now. We use Eq. (18) again. With the help of [see Ref. 8, Eq. (A.32)]

$$\begin{aligned} (a-a^{-1})^{l+1}C^{l+1}_{n-l-1}\left(\frac{1+a^2}{2a}\right) &= a^nP_l^{(n,-n)}\left(\frac{1+a^2}{1-a^2}\right) - a^{-n}P_l^{(-n,n)}\left(\frac{1+a^2}{1-a^2}\right) \\ &= a^nP_l^{(n,-n)}\left(\frac{1+a^2}{1-a^2}\right) - (-1)^l\{\text{Idem},a\to a^{-1}\}, \end{aligned} \tag{34}$$

we obtain from Eq. (18)

$$\frac{p\langle p|g^c_{\beta l}\rangle}{k\langle k|g_{\beta l}\rangle} = \sum_{n=l+1}^{\infty}\frac{n(Ba)^n}{n+i\gamma}P_l^{(n,-n)}\left(\frac{1+a^2}{1-a^2}\right) - (-1)^l\{\text{Idem},a\to a^{-1}\}. \tag{35}$$

We substitute

$$\frac{n}{n+i\gamma} = 1 - \frac{i\gamma}{n+i\gamma}.$$

Then the γ-dependence is wholly contained in the second term. By noting that

$$\lim_{\gamma\to 0}\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle,$$

the γ-independent part follows easily; cf. Eqs. (18) and (19).

By using the equality

$$\sum_{n=l+1}^{\infty}(n+i\gamma)^{-1}z^nP_l^{(n,-n)}(u) = z^{l+1}2^{-l}(1+u)^l(l+1+i\gamma)^{-1} \times\sum_{n=0}^{l}\binom{2l-n}{l}\left(\frac{2}{1+u}\right)^n {}_2F_1(n+1,l+1+i\gamma;l+2+i\gamma;z), \tag{36}$$

which will be proved in the Appendix, we obtain from Eq. (35)

$$(\pi/2)^{1/2}p\langle p|g^c_{\beta l}\rangle = \left(\frac{p}{\beta^2+p^2}\right)^{l+1} - \left(\frac{k}{\beta^2+k^2}\right)^{l+1}\times Z_l(a;B), \tag{37}$$

where

$$\begin{aligned} Z_l(a;B) &= \frac{i\gamma Ba}{l+1+i\gamma}\left(\frac{Ba}{1-a^2}\right)^l\sum_{n=0}^{l}\binom{2l-n}{l}(1-a^2)^n \\ &\quad \times {}_2F_1(n+1,l+1+i\gamma;l+2+i\gamma;Ba) \\ &\quad -(-1)^l\{\text{Idem},a\to a^{-1}\}. \end{aligned} \tag{38}$$

By using

$${}_2F_1(a,b;c;z) = (1-z)^{c-a-b}\,{}_2F_1(c-a,c-b;c;z),$$

Eq. (38) can be rewritten as

$$Z_l(a;B) = \frac{i\gamma Ba}{l+1+i\gamma}\left(\frac{Ba}{1-Ba}\right)^l \sum_{m=0}^{l} \binom{l+m}{l}\left(\frac{1-Ba}{1-a^2}\right)^m \times {}_2F_1(1,m+1+i\gamma;l+2+i\gamma;Ba) - (-1)^l\{\text{Idem}, a\to a^{-1}\}. \quad (39)$$

It should be noted that Eq. (37) is just Eq. (29) rewritten. The quantity $Z_l(a;B)$ occurring in each of these equations actually denotes the same function, which proves that this notation is justified. The proof follows by comparing Eq. (33) with Eq. (38) and using the equality

$$\begin{aligned}&{}_2F_1(n+1,l+1+i\gamma;l+2+i\gamma;z)\\ &= \sum_{\mu=0}^{n-1}(-1)^n \frac{\Gamma(n-\mu)\Gamma(l+2+i\gamma)}{n!\Gamma(l+1+i\gamma-\mu)} z^{-\mu-1}(z-1)^{\mu-n}\\ &\quad + (-1)^n \frac{\Gamma(l+2+i\gamma)}{i\gamma n!\Gamma(l+1-n+i\gamma)}\\ &\quad \times z^{-l-1}\left[F_{i\gamma}(z) - 1 - \sum_{\nu=1}^{l-n}\frac{i\gamma z^\nu}{\nu+i\gamma}\right],\end{aligned} \quad (40)$$

which will be proved in the Appendix.

5. EVALUATION OF IM $Z_l(a;B)$ AND PROOF OF IM $X_l(a;B) = 0$

We shall investigate the Coulomb-modified form factor $g^c_{\beta l}$ again. We rewrite Eq. (29),

$$\langle p|g^c_{\beta l}\rangle = \langle p|g_{\beta l}\rangle - \langle k|g_{\beta l}\rangle kp^{-1}Z_l(a;B), \quad (41)$$

and introduce the function X_l by putting

$$\begin{aligned}Z_l(a;B) &= X_l(a;B) + i\,\text{Im}P_l^{(i\gamma,-i\gamma)}(u)\\ &\quad + F_{i\gamma}(Ba)P_l^{(-i\gamma,i\gamma)}(u) - F_{i\gamma}(B/a)P_l^{(i\gamma,-i\gamma)}(u).\end{aligned} \quad (42)$$

In view of Eq. (33) we have

$$\begin{aligned}&X_l(a;B) - i\,\text{Im}P_l^{(i\gamma,-i\gamma)}(u)\\ &= (-1)^{l+1} i\gamma\Gamma(l+1+i\gamma)\sum_{m=0}^{l}\binom{l+m}{l}\frac{(a^2-1)^{-m}}{(l-m)!}\\ &\quad \times\left[\frac{1}{\Gamma(m+1+i\gamma)}\sum_{\nu=1}^{m}\frac{(Ba)^\nu}{\nu+i\gamma}\right.\\ &\quad \left. - (Ba)^m \sum_{\mu=0}^{l-m-1}\frac{\mu!}{\Gamma(\mu+m+2+i\gamma)}\left(\frac{Ba}{Ba-1}\right)^{\mu+1}\right]\\ &\quad - (-1)^l\{\text{Idem}, a\to a^{-1}\}.\end{aligned} \quad (43)$$

In this section we shall prove that $X_l(a;B)$ is *real* when p, k, β, and γ are real. It is difficult to derive this from Eq. (43). Instead we use Eq. (30), which we rewrite

$$\begin{aligned}&Z_l(a;B)\\ &= (a-a^{-1})^{l+1} i\gamma\int_0^1 [tB + (tB)^{-1} - a - a^{-1}]^{-l-1}\\ &\quad \times t^{i\gamma-1}dt.\end{aligned} \quad (44)$$

For convenience we shall suppress the arguments of Z_l and X_l.

We are going to prove the following two equations: (i) From Eq. (42)

$$Z_l - Z_l^* = X_l - X_l^* - |\Gamma(1+i\gamma)|^2 \times [(-aB^*)^{i\gamma}P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}], \quad (45)$$

(ii) from Eq. (44)

$$Z_l - Z_l^* = -|\Gamma(1+i\gamma)|^2 \times B^{-i\gamma}[(-a)^{i\gamma}P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}]. \quad (46)$$

By combining Eqs. (45) and (46) we obtain $X_l - X_l^* = 0$ (note that $B^{-i\gamma}$ is *real*).

(i) It is easily verified that one has from Eq. (42),

$$\begin{aligned}Z_l - Z_l^* &= X_l - X_l^* + 2i\,\text{Im}P_l^{(i\gamma,-i\gamma)}(u)\\ &\quad + P_l^{(-i\gamma,i\gamma)}(u)[F_{i\gamma}(Ba) + F_{-i\gamma}(B^*/a^*)]\\ &\quad - P_l^{(i\gamma,-i\gamma)}(u)[F_{-i\gamma}(B^*a^*) + F_{i\gamma}(B/a)].\end{aligned}$$

Because of the branch cut of $F_{i\gamma}$ we have to be careful in distinguishing a from a^* (recall that $\text{Im}k\downarrow 0$). By using $B^* = B^{-1}$, $F_{i\gamma}(Ba) = F_{i\gamma}(Ba^*)$ (for $\text{Im}\,k\downarrow 0$), and

$$F_{i\gamma}(z) + F_{-i\gamma}(z^{-1}) = 1 + \Gamma(1+i\gamma)\Gamma(1-i\gamma)(-z)^{-i\gamma},$$

we get

$$F_{i\gamma}(Ba) + F_{-i\gamma}(B^*/a^*) = 1 + |\Gamma(1+i\gamma)|^2(-a^*B)^{-i\gamma}.$$

In this way Eq. (45) is readily verified.

(ii) In order to prove Eq. (46), we first put $\tau = t^{-1}$ in the integral in Eq. (44). Then we get

$$\begin{aligned}Z_l(a;B) &= (a-a^{-1})^{l+1} i\gamma\\ &\quad \times\int_1^\infty [\tau^{-1}B + \tau B^{-1} - a - a^{-1}]^{-l-1}\tau^{-i\gamma-1}d\tau.\end{aligned} \quad (47)$$

By combining Eq. (47) with Eq. (44) we obtain

$$Z_l(a;B) - Z_l^*(a;B) = (a-a^{-1})^{l+1} i\gamma B^{-i\gamma} \times\int_0^\infty (x^2+2bx+1)^{-l-1}x^{i\gamma+l}dx, \quad (48)$$

where

$$b = (k^2+p^2)/(k^2-p^2) = -\tfrac{1}{2}(a+a^{-1}).$$

Assuming for convenience $k>p$, we have

$$\begin{aligned}&\int_0^\infty (x^2+2bx+1)^{-l-1}x^{i\gamma+l}dx\\ &= 2^{-l}(l!)^{-1}\Gamma(l+1-i\gamma)e^{\pi\gamma}\left(\frac{k^2-p^2}{2pk}\right)^{l+1}Q_l^{i\gamma}(u)\\ &= \frac{i\pi}{\sinh\pi\gamma}\left(\frac{k^2-p^2}{4pk}\right)^{l+1}\\ &\quad \times[(-a)^{i\gamma}P_l^{(i\gamma,-i\gamma)}(u) - (-a)^{-i\gamma}P_l^{(-i\gamma,i\gamma)}(u)].\end{aligned} \quad (49)$$

Since $a - a^{-1} = 4pk/(k^2-p^2)$, the proof of Eq. (46) follows easily from Eqs. (48) and (49). Finally, by comparing Eqs. (45) and (46) we get $X_l = X_l^*$, by noting that

$$(-a/B)^{i\gamma} = B^{-i\gamma}(-a)^{i\gamma}.$$

When $k>p$, we have $-a>0$ so that this equality is obvious. On the other hand, when $k<p$, we have $a>0$. Then we get, by using $k\to k+i\epsilon$, $\epsilon\downarrow 0$ ($\beta>0$, $k>0$),

$$(-a)^{i\gamma} = e^{-\pi\gamma}a^{i\gamma},$$

$$(-B)^{-i\gamma} = e^{-\pi\gamma}B^{-i\gamma},$$

hence

$$(-a/B)^{i\gamma} = a^{i\gamma}(-B)^{-i\gamma} = e^{-\pi\gamma}a^{i\gamma}B^{-i\gamma} = (-a)^{i\gamma}B^{-i\gamma}.$$

This completes the proof of the fact that X_l is real. The imaginary part of Z_l follows from Eq. (46), which can be rewritten as

$$Z_l(a;B) - Z_l^*(a;B) = \frac{-2\pi\gamma}{e^{2\pi\gamma}-1} B^{-i\gamma} \times [a^{i\gamma} P_l^{(i\gamma,-i\gamma)}(u) - \text{c.c.}]. \quad (50)$$

6. SUMMARY

In this section we shall give a summary of the most important results concerning the off-shell T matrix for Coulomb plus simple separable potentials; see Eq. (1). It is sufficient to give explicit expressions for the quantities given by Eq. (2) and (3). In Eq. (23) we have obtained

$$\langle g_{\alpha l}|G_{cl}|g_{\beta l}\rangle = \frac{(2l+1)!!}{(2l)!!} \frac{(\alpha+\beta)^{-2l-1}}{(\alpha-ik)(\beta-ik)} \frac{-1}{l+1+i\gamma} \times {}_2F_1(1,i\gamma-l;i\gamma+l+2;AB). \quad (51)$$

The Coulomb-modified form factors are given by

$$\langle p|g_{\beta l}^c\rangle = \langle p|g_{\beta l}\rangle - \langle k|g_{\beta l}\rangle kp^{-1} Z_l(a;B), \quad (52)$$

$$Z_l(a;B) = X_l(a;B) + i\,\mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u) + F_{i\gamma}(Ba)P_l^{(-i\gamma,i\gamma)}(u) - F_{i\gamma}(B/a)P_l^{(i\gamma,-i\gamma)}(u), \quad (53)$$

where

$$X_l(a;B) - i\,\mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u) = (-1)^{l+1} i\gamma\Gamma(l+1+i\gamma) \times \sum_{m=0}^{l} \binom{l+m}{l} \frac{(a^2-1)^{-m}}{(l-m)!} \times \left[\frac{1}{\Gamma(m+1+i\gamma)} \sum_{\nu=1}^{m} \frac{(Ba)^\nu}{\nu+i\gamma} - (Ba)^m \times \sum_{\mu=0}^{l-m-1} \frac{\mu!}{\Gamma(\mu+m+2+i\gamma)} \left(\frac{Ba}{Ba-1}\right)^{\mu+1}\right] - (-1)^l\{\mathrm{Idem}, a\to a^{-1}\}. \quad (54)$$

Here p, k, β, and γ are supposed to be real. In this case $X_l(a;B)$ is *real*. For $l = 0,1,2$, we have obtained from Eq. (54),

$$X_0 = 0,$$

$$X_1 = i\gamma(B^{-1}-B)D^{-1} = \beta\gamma k^{-1}(k^2-p^2)(\beta^2+p^2)^{-1}, \quad (55)$$

$$X_2 = \tfrac{1}{2}\gamma D^{-1}[\gamma(a-a^{-1}) + i(B-B^{-1}) \times \left\{\frac{1-a^2}{aD} - 3\frac{1+a^2}{1-a^2}\right\}] = \frac{p\gamma^2}{2k}\frac{\beta^2+k^2}{\beta^2+p^2} + \frac{\beta\gamma(k^2-p^2)}{4pk^2(\beta^2+p^2)^2} \times [3(p^4+k^2\beta^2) + 5p^2(\beta^2+k^2)],$$

where

$$D = B + B^{-1} - a - a^{-1}.$$

It is interesting to note that

$$X_l(a;1) \equiv X_l(p/k), \quad (56)$$

where $X_l(p/k)$ has been introduced in Ref. 12. The particular case $B=1$ is obtained by letting $\beta\to\infty$. Then we get from Eq. (55),

$$X_0 = X_1 = 0, \quad X_2 = \tfrac{1}{2}\gamma^2 p/k,$$

in agreement with Eq. (11a) of Ref. 12.

From Eq. (42) we have derived a convenient recursion relation for $X_l = X_l(a;B)$:

$$(l+1)X_{l+1} = (l+1)\frac{1+a^2}{1-a^2}X_l + a\frac{d}{da}X_l + \gamma D^{-1}[(a-a^{-1})I_l + i(B^{-1}-B)R_l], \quad (57)$$

where

$$R_l = \mathrm{Re}P_l^{(i\gamma,-i\gamma)}(u),$$

$$I_l = \mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u);$$

see the Appendix.

In Eq. (30) we have obtained a useful integral representation for Z_l,

$$Z_l(a;B) = B^{l+1}(a-a^{-1})^{l+1} i\gamma \times \int_0^1 [(1-tBa)(1-tB/a)]^{-l-1} t^{i\gamma+l} dt. \quad (58)$$

A representation which is particularly convenient for numerical calculations is given by Eq. (39),

$$Z_l(a;B) = \frac{i\gamma Ba}{l+1+i\gamma}\left(\frac{Ba}{1-Ba}\right)^l \sum_{m=0}^{l} \binom{l+m}{l}\left(\frac{1-Ba}{1-a^2}\right)^m \times {}_2F_1(1,m+1+i\gamma;l+2+i\gamma;Ba) - (-1)^l\{\mathrm{Idem}, a\to a^{-1}\}. \quad (59)$$

Note that the sum of the first two parameters of ${}_2F_1$ minus the third one is equal to a nonpositive integer.

When p, β, γ are taken real, and $\mathrm{Im}\,k\downarrow 0$, the imaginary part of $Z_l(a;B)$ can be evaluated with the help of Eq. (58). We have from Eq. (50)

$$-2iC_0^{-2}B^{i\gamma}\mathrm{Im}Z_l(a;B) = a^{i\gamma}P_l^{(i\gamma,-i\gamma)}(u) - a^{*-i\gamma}P_l^{(-i\gamma,i\gamma)}(u), \quad (60)$$

which can be rewritten as [cf. Eq. (7.7) of Ref. 8]

$$B^{i\gamma}\mathrm{Im}Z_l(a;B) = \tfrac{1}{2}\pi p(f_{cl}^*)^{-1}\langle p|V_{cl}|kl+\rangle_c. \quad (61)$$

Here

$$C_0^2 = 2\pi\gamma(e^{2\pi\gamma}-1)^{-1},$$

$$f_{cl} = e^{\pi\gamma/2} l!/\Gamma(l+i\gamma+1).$$

Finally, in Eq. (27) we have obtained an interesting connection between the Coulomb-modified form factor $|g_{\beta l}^c\rangle$ and the Coulomb Jost state $|kl\uparrow\rangle_c$

$$\lim_{\beta\to\infty} (\pi/2)^{1/2}\beta^{2l+2}\langle p|g_{\beta l}^c\rangle = p^l - \tfrac{1}{2}\pi k^{l+1}\langle p|V_{cl}|kl\uparrow\rangle_c f_{cl}^{-1} = \tfrac{1}{2}\pi k^{l+1}(p^2-k^2)\langle p|kl\uparrow\rangle_c f_{cl}^{-1}. \quad (62)$$

ACKNOWLEDGMENT

This work is part of the project "The Coulomb Potential in Quantum Mechanics and Related Topics."

APPENDIX

In this Appendix we shall prove Eqs. (31), (32), (36), (40), and (57). Denoting the integral in Eq. (31) by J_l, we have

$$J_l = i\gamma\int_0^1 [(1-ta)(1-tb)]^{-l-1}t^{l+i\gamma}dt$$
$$= i\gamma(\sqrt{ab})^{-l-1-i\gamma}\int_0^{\sqrt{ab}}(1+\tau^2-2\tau x)^{-l-1}\tau^{l+i\gamma}d\tau, \tag{A1}$$

where $\tau = t\sqrt{ab}$ and $x = \frac{1}{2}(ab)^{-1/2}(a+b)$. By using

$$(1+\tau^2-2\tau x)^{-l-1} = \sum_{m=0}^{\infty}\tau^m C_m^{l+1}(x), \quad |\tau|<1, \tag{A2}$$

we get, assuming here $0<a<1$, $0<b<1$,

$$J_l = i\gamma\sum_{m=0}^{\infty}(l+1+i\gamma+m)^{-1}(ab)^{m/2}C_m^{l+1}(x). \tag{A3}$$

By putting $p/k = (a^{1/2}+b^{1/2})^2(a-b)^{-1}$ in Eq. (A32) of Ref. 8, we obtain

$$\left(\frac{b-a}{(ab)^{1/2}}\right)^{l+1}C_{n-l-1}^{l+1}\left(\frac{a+b}{2(ab)^{1/2}}\right)$$
$$= \left(\frac{b}{a}\right)^{n/2}P_l^{(n,-n)}\left(\frac{a+b}{a-b}\right)$$
$$- \left(\frac{b}{a}\right)^{-n/2}P_l^{(-n,n)}\left(\frac{a+b}{a-b}\right). \tag{A4}$$

Inserting this equation into Eq. (A3), and using

$$P_l^{(n,-n)}(-z) = (-)^l P_l^{(-n,n)}(z), \tag{A5}$$

we obtain

$$J_l = f_l(a;b) + f_l(b;a),$$

with

$$f_l(a;b)$$
$$= i\gamma(a-b)^{-l-1}\sum_{n=l+1}^{\infty}(n+i\gamma)^{-1}a^n P_l^{(n,-n)}\left(\frac{b+a}{b-a}\right). \tag{A6}$$

In order to reduce this further we use[10]

$$P_l^{(n,-n)}(u)$$
$$= \binom{2l}{l}\left(\frac{1+u}{2}\right)^l \lim_{\epsilon\to 0}{}_2F_1\left(-l, n-l; \epsilon-2l; \frac{2}{1+u}\right). \tag{A7}$$

Substituting furthermore

$$\sum_{m=0}^{\infty}\frac{l+1+i\gamma}{m+l+1+i\gamma}(n+1)_m\frac{z^m}{m!}$$
$$= {}_2F_1(n+1, l+1+i\gamma; l+2+i\gamma; z)$$

and

$$\binom{2l}{l}(-l)_n/(-2l)_n = \binom{2l-n}{l},$$

we get

$$\sum_{n=l+1}^{\infty}(n+i\gamma)^{-1}z^n P_l^{(n,-n)}(u)$$
$$= z^{l+1}2^{-l}(1+u)^l(l+1+i\gamma)^{-1}$$
$$\times\sum_{n=0}^{l}\binom{2l-n}{l}\left(\frac{2}{1+u}\right)^n$$
$$\times{}_2F_1(n+1, l+1+i\gamma; l+2+i\gamma; z), \tag{A8}$$

which is just Eq. (36). Hence from Eq. (A6),

$$f_l(a;b) = \frac{i\gamma}{l+1+i\gamma}\left(\frac{a}{a-b}\right)^{l+1}\sum_{n=0}^{l}\binom{2l-n}{l}\left(\frac{b-a}{b}\right)^{n-l}$$
$$\times{}_2F_1(n+1, l+1+i\gamma; l+2+i\gamma; a). \tag{A9}$$

By applying now Eqs. (A12–A14) and using the equality

$$P_l^{(-i\gamma,i\gamma)}\left(\frac{b+a}{b-a}\right) = (-1)^l\frac{\Gamma(l+1+i\gamma)}{l!\Gamma(1+i\gamma)}$$
$$\times\sum_{m=0}^{l}\frac{(l+1)_m(-l)_m}{(1+i\gamma)_m m!}\left(\frac{b}{b-a}\right)^m, \tag{A10}$$

the proof of Eq. (32) follows.

It may be noted that, according to Eqs. (30) and (31),

$$Z_l(a;B) = (a-b)^{l+1}J_l, \tag{A11}$$

where in the right member a is replaced by Ba and b by B/a.

Furthermore, we have to prove Eq. (40). The reduction of the hypergeometric function ${}_2F_1(n+1, \alpha; \alpha+1; z)$ to $F_{i\gamma}(z)$ is performed in two steps. (We put $l+1+i\gamma = \alpha$ for convenience here). In the first place we have, by integration by parts,

$${}_2F_1(n+1, \alpha; \alpha+1; z) = \alpha\int_0^1\frac{t^{\alpha-1}dt}{(1-tz)^{n+1}}$$
$$= \frac{\alpha}{nz}\left[\frac{t^{\alpha-1}}{(1-tz)^n}\right]_0^1 - \frac{\alpha(\alpha-1)}{nz}\int_0^1\frac{t^{\alpha-2}}{(1-tz)^n}dt$$
$$= \sum_{\mu=0}^{n-1}(-1)^\mu\frac{\Gamma(\alpha+1)\Gamma(n-\mu)}{\Gamma(\alpha-\mu)\Gamma(n+1)}z^{-\mu-1}(1-z)^{\mu-n}$$
$$+ (-1)^n\frac{\Gamma(\alpha+1)}{\Gamma(\alpha-n+1)\Gamma(n+1)}z^{-n}$$
$$\times{}_2F_1(1, \alpha-n; \alpha+1-n; z), \tag{A12}$$

by iteration. The second step consists of reducing the ${}_2F_1$ in the right member of Eq. (A12). By iteration of

$${}_2F_1(1, \beta; \beta+1; z) = 1 + \frac{\beta z}{1+\beta}{}_2F_1(1, \beta+1; \beta+2; z) \tag{A13}$$

we get

$${}_2F_1(1, \beta; \beta+1; z)$$
$$= \sum_{\nu=0}^{n-1}\frac{\beta z^\nu}{\nu+\beta} + \frac{\beta z^n}{n+\beta}{}_2F_1(1, \beta+n; \beta+n+1; z). \tag{A14}$$

By combining Eqs. (A12) and (A14) the proof of Eq. (40) is readily obtained.

Finally we shall prove the recursion relation for X_l, Eq. (57). To this end, we start from Eq. (42). Introducing for convenience

$$R_l = \mathrm{Re}P_l^{(i\gamma,-i\gamma)}(u),$$
$$I_l = \mathrm{Im}P_l^{(i\gamma,-i\gamma)}(u), \tag{A15}$$

we have, suppressing the arguments of Z_l and X_l,

$$Z_l = X_l + iI_l + F_{i\gamma}(Ba)P_l^{(-i\gamma,i\gamma)}(u)$$
$$- F_{i\gamma}(Ba^{-1})P_l^{(i\gamma,-i\gamma)}(u). \tag{A16}$$

From Eq. (30) we have, introducing $h_l(a)$,

$$
\begin{aligned}
&Z_l(a;B)\\
&= (a-a^{-1})^{l+1} i\gamma \int_0^1 [tB + (tB)^{-1} - a - a^{-1}]^{-l-1}\\
&\quad \times t^{i\gamma-1} dt\\
&\equiv (a-a^{-1})^{l+1} h_l(a), \qquad \text{(A17)}
\end{aligned}
$$

from which we derive

$$\frac{d}{da} h_l(a) = (l+1) h_{l+1}(a) \frac{d}{da}(a + a^{-1}),$$

hence

$$\frac{d}{dx} Z_l = \frac{l+1}{x} \frac{1+x^2}{1-x^2} Z_l - \frac{2(l+1)}{1-x^2} Z_{l+1}, \qquad \text{(A18)}$$

where $x = p/k$. We shall use furthermore,

$$\frac{d}{dz} F_{i\gamma}(z) = \frac{i\gamma}{z(1-z)} - \frac{i\gamma}{z} F_{i\gamma}(z) \qquad \text{(A19)}$$

and

$$
\begin{aligned}
&(1-z^2) \frac{d}{dz} P_l^{(i\gamma,-i\gamma)}(z)\\
&\quad = (i\gamma + lz + z) P_l^{(i\gamma,-i\gamma)}(z) - (l+1) P_{l+1}^{(i\gamma,-i\gamma)}(z). \qquad \text{(A20)}
\end{aligned}
$$

Equation (A19) is well known (cf. Ref. 2), and the proof of Eq. (A20) follows by substituting

$$P_l^{(i\gamma,-i\gamma)}(z) = \binom{l+i\gamma}{l} {}_2F_1(-l, l+1; i\gamma+1; \tfrac{1}{2} - \tfrac{1}{2}z), \qquad \text{(A21)}$$

and using the hypergeometric series representation for ${}_2F_1$.

We differentiate both members of Eq. (A16) with respect to $x = p/k$. Then we obtain, using Eqs. (A18–A20)

$$
\begin{aligned}
\frac{d}{dx} Z_l &= \frac{l+1}{x} \frac{1+x^2}{1-x^2} [X_l + iI_l + F_{i\gamma}(Ba) P_l^{(-i\gamma,i\gamma)}(u)\\
&\quad - F_{i\gamma}(Ba^{-1}) P_l^{(i\gamma,-i\gamma)}(u)]\\
&\quad - \frac{2(l+1)}{1-x^2} [X_{l+1} + iI_{l+1} + F_{i\gamma}(Ba) P_{l+1}^{(-i\gamma,i\gamma)}(u)\\
&\quad - F_{i\gamma}(Ba^{-1}) P_{l+1}^{(i\gamma,-i\gamma)}(u)]\\
&= \frac{d}{dx} X_l + i \frac{x^2-1}{2x^2} \frac{d}{du} I_l + \frac{2i\gamma}{x^2-1}\\
&\quad \times \left[\frac{1}{1-Ba} P_l^{(-i\gamma,i\gamma)}(u) + \frac{a}{a-B} P_l^{(i\gamma,-i\gamma)}(u) \right]\\
&\quad + F_{i\gamma}(Ba) \left[\frac{-2i\gamma}{x^2-1} P_l^{(-i\gamma,i\gamma)}(u) \right.\\
&\quad \left. + \frac{x^2-1}{2x^2} \frac{d}{du} P_l^{(-i\gamma,i\gamma)}(u) \right]\\
&\quad - F_{i\gamma}(Ba^{-1}) \left[\frac{2i\gamma}{x^2-1} P_l^{(i\gamma,-i\gamma)}(u) \right.\\
&\quad \left. + \frac{x^2-1}{2x^2} \frac{d}{du} P_l^{(i\gamma,-i\gamma)}(u) \right]. \qquad \text{(A22)}
\end{aligned}
$$

The coefficients of the corresponding hypergeometric functions in the members of Eq. (A22) turn out to be equal, which gives a considerable reduction. Equating the imaginary parts in Eq. (A22) we obtain

$$
\begin{aligned}
&(l+1) I_{l+1}\\
&\quad = (l+1) \frac{x^2+1}{2x} I_l + \frac{(x^2-1)^2}{4x^2} \frac{d}{du} I_l + \gamma R_l, \qquad \text{(A23)}
\end{aligned}
$$

which can be rewritten as

$$(l+1) I_{l+1} = (l+1) u I_l + (u^2-1) \frac{d}{du} I_l + \gamma R_l. \qquad \text{(A24)}$$

The proof of Eq. (A24) follows also from Eq. (A20) by inserting Eq. (A15) and splitting real and imaginary parts. Equating the *real* parts of the members of Eq. (A22) we obtain

$$
\begin{aligned}
&(x^2-1) \frac{d}{dx} X_l + (l+1)(x + x^{-1}) X_l - 2(l+1) X_{l+1}\\
&= 2i\gamma [R_l - (1-Ba)^{-1}(R_l - iI_l) - (1-Ba^{-1})(R_l + iI_l)]\\
&= 2i\gamma D^{-1} [R_l (B - B^{-1}) + iI_l (a - a^{-1})], \qquad \text{(A25)}
\end{aligned}
$$

where

$$D = B + B^{-1} - a - a^{-1}.$$

Note that D is real and $B - B^{-1}$ is imaginary ($B^{-1} = B^*$), which shows that the right members of Eq. (A25) are real.

Finally, we rewrite Eq. (A25) as follows:

$$
\begin{aligned}
(l+1) X_{l+1}(a;B) &= (l+1) \frac{1+a^2}{1-a^2} X_l(a;B) + a \frac{d}{da} X_l(a;B)\\
&\quad + \gamma D^{-1} [(a - a^{-1}) I_l + i(B^{-1} - B) R_l], \qquad \text{(A26)}
\end{aligned}
$$

which is the desired recursion relation, Eq. (57).

[1]H. van Haeringen, Phys. Lett. A **86**, 359 (1981).
[2]H. van Haeringen and R. van Wageningen, J. Math. Phys. **16**, 1441 (1975).
[3]L. P. Kok and H. van Haeringen, Phys. Rev. C **21**, 512 (1980).
[4]I. H. Sloan, Phys. Rev. A **7**, 1016 (1973).
[5]H. van Haeringen, "The partial-wave projected Coulomb T matrix for all l in closed hypergeometric form," preprint, Delft, University of Technology, 1981.
[6]N. Maleki and J. Macek, Phys. Rev. A **21**, 1403 (1980); cf. E. J. Kelsey and J. Macek, J. Math. Phys. **17**, 1182 (1976).
[7]H. van Haeringen, L. P. Kok, and W. Plessas, J. Phys. G. Nucl. Phys. **7**, L193 (1981).
[8]H. van Haeringen, "Long-Range Potentials in Quantum Mechanics," Report 150, University of Groningen, 1979.
[9]J. Schwinger, J. Math. Phys. **5**, 1606 (1964); V. Fock, Z. Phys. **98**, 145 (1935).
[10]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, New York, 1966), p. 244.
[11]H. van Haeringen and L. P. Kok, Phys. Lett. A **82**, 317 (1981).
[12]H. van Haeringen, "The Coulomb Jost states in the momentum representation for all partial waves in closed form," preprint, Delft, 1981).

On a generalized F. Riesz problem

P. Diţă
Institute for Physics and Nuclear Engineering, P. O. Box MG 6, Bucharest, Romania

(Received 1 March 1982; accepted for publication 2 April 1982)

The description of the pion–pion partial waves which are analytic and crossing symmetric is equivalent to solving a generalized Riesz problem. A continuous family of solutions of this problem is constructed.

PACS numbers: 11.20.Fm, 11.50.Jg

I. INTRODUCTION

Let $(h_n)_1^\infty$ be an infinite sequence of complex numbers. We are interested in the following problem.

Problem: Given a positive number $\rho > 0$ and an infinite sequence $(h_n)_1^\infty$ of complex numbers, it is required to find among the functions

$$f(z) = \sum_{n=1}^{\infty} a_n z^{n-1}$$

regular in the disk $|z| < 1$ those for which

$$\sum_{n=1}^{\infty} a_n h_n = 1, \tag{1.1}$$

$$\|f\|_1 \leqslant \rho^{-1}. \tag{1.2}$$

$\|f\|_1$ is the norm of a function in Hardy space H_1 of analytic functions defined as the set of those functions which satisfy

$$\|f\|_1 = \sup_{r<1}\left[(1/2\pi)\int_0^{2\pi} |f(re^{i\theta})|\, d\theta\right] < \infty .$$

This problem is the generalization of a problem by F. Riesz originally stated for a finite sequence $(h_n)_1^N$ and the minimum norm in the H_1 space.[1]

Stated as above, the problem appears in the S-matrix theory of elementary particle physics and is equivalent to the problem of constructing partial waves of a scattering process that are crossing symmetric and analytic in a certain domain.

This is easily seen if we remember that the crossing symmetry takes a simple form when it is expressed on partial waves. For example in the pion–pion scattering the first (and the simplest) crossing relation has the form[2]

$$\int_0^4 (4-s)\,[\,5f_0^2(s) - 2f_0^0(s)\,]\, ds = 0, \tag{1.3}$$

where $f_1^I(s)$ denotes a partial wave of isospin I and orbital momentum 1.

The partial waves are real analytic functions in a domain of the s complex plane,[3] a domain which includes the disk $|z| < 1$, $z = (s-2)/2$. If we make the transformation

$$S_1^I(s) = 1 + 2[(4-s)/s]^{1/2} f_1^I(s),$$

the integral relation (1.3) is equivalent to

$$\sum_{n=1}^{\infty} a_n h_n = 1,$$

where the sequence $(h_n)_1^\infty$ is given by

$$\begin{aligned} h_n &= (2/\pi)\int_{-1}^{1} (1-x^2)^{1/2} x^{n-1}\, dx \\ &= \{[1 + (-1)^{n-1}]/\pi\} B(3/2, n/2) \end{aligned} \tag{1.4}$$

and a_n are the coefficients in the power series expansion of the function

$$f(z) = \frac{5S_0^2(s) - 2S_0^0(s)}{3} = \sum_{n=1}^{\infty} a_n z^{n-1}.$$

In a previous paper[4] we found the solutions of the problem in the case $\rho = s_k$, s_k being the singular numbers of the Hankel operator $H = (h_{i+j-1})_1^\infty$ defined by the sequence $(h_n)_1^\infty$. The singular numbers of a (non-Hermitian) operator A are the eigenvalues of the nonnegative operator $(A^*A)^{1/2}$. For a Hankel operator $A^* = \bar{A}$, the bar denoting complex conjugation.

The solutions have the form

$$f(z) = s_k^{-1} z^{n-1} P(z)\, \overline{P(1/\bar{z})}\, a^2(z) \varphi_e^2(z), \tag{1.5}$$

where n is the multiplicity of the singular number s_k, $P(z)$ an arbitrary polynomial of degree less than n, and $a(z)$ and $\varphi_e(z)$ respectively inner and outer functions uniquely determined by the operator H.

It happens that analyticity and crossing symmetry together give rise to a Hankel operator which is positive, trace class, and with multiplicity of eigenvalues equal to unity.[4] In this case the solutions (1.5) have the simplified form

$$f(z) = s_k^{-1} \varphi_k^2(z),$$

where $\varphi_k(z)$ denotes the function $\varphi_k(z) = \sum_{n=1}^{\infty} \xi_n z^{n-1}$, $\xi = (\xi_1, \xi_2, \cdots) \in \ell_2$ being the orthonormal eigenvector of the equation

$$H\xi = s_k \xi .$$

If $\rho > \|H\|$, the problem has no solution.[5] The case remaining unsolved is that where

$$s_k > \rho > s_{k+1}, \quad k = 1, 2, \cdots .$$

The aim of this paper is to construct a continuous family of solutions for this case. Although, by this construction, we do not entirely solve the problem, we consider the finding of these solutions a decisive step towards the construction of all the solutions of the Riesz problem.

Our procedure of constructing solutions is given in the next section. The paper ends with some remarks concerning the explicit construction of solutions for the pion–pion scattering.

 0022-2488/83/051282-04$02.50

II. CONSTRUCTION OF SOLUTIONS

With the sequence $(h_n)_1^\infty$ we will associate the following objects:

(1) the Hankel matrix $H = (h_{i+j-1})_1^\infty$;

(2) the continuous functional Φ_H defined on H_1 by the expression

$$\Phi_H(f) = \lim_{n\to\infty} \sum_{j=1}^{n} \left(1 - \frac{j-1}{n}\right) h_j a_j, \tag{2.1}$$

where

$$f(z) = \sum_{n=1}^{\infty} a_n z^{n-1} \in H_1.$$

For

$$\xi = (\xi_j)_1^\infty \text{ and } \eta = (\eta_j)_1^\infty,\ \xi,\eta\in\ell_2,$$

we will define the functions

$$\xi_+(\zeta) = \sum_{j=1}^{\infty} \xi_j \zeta^{j-1}, \quad \eta_-(\zeta) = \sum_{j=1}^{\infty} \eta_j \zeta_j^{-j}, \quad \zeta = \exp(i\theta).$$

It is easily shown[5] that for any bounded H we have

$$\Phi_H(\xi_+\bar{\eta}_-\bar{\zeta}) = (H\xi,\eta), \quad \zeta = \exp(i\theta). \tag{2.2}$$

Hence, if $(H\xi,\eta)\neq 0$, the function

$$g(\zeta) = \xi_+(\zeta)\, \overline{\eta_-(\zeta)}\, \bar{\zeta} / (H\xi,\eta)$$

will satisfy the relation (1.1), i.e., it will be a crossing symmetric function analytic in the disk $|z| < 1$.

If the sequence $(h_n)_1^\infty$ is given by

$$h_n = \int_{-1}^{1} x^{n-1}\, d\sigma(x), \quad n = 1,2,\cdots, \tag{2.3}$$

with $\sigma(x)$, in general, a complex measure with bounded variation, the final results take a remarkably simple form. The case (2.3) includes all the cases of physical interest, and we will state our results for the class of Hankel operators generated by sequences of this form, although our method of constructing solutions is applicable to any Hankel operator which has, at least, a finite number of singular numbers s_k as will be evident from the explicit construction of the solutions.

Let $\|H\| = s_1 \geqslant s_2 \geqslant s_3 \geqslant \cdots s_k > s_{k+1} \geqslant \cdots$ be the enumeration of the point spectrum of $(\bar{H}H)^{1/2}$, the eigenvalues being included with their multiplicities.

In the following we shall suppose that

$$s_k > \rho > s_{k+1}, \tag{2.4}$$

which is the case left unsolved.

The basic idea of our procedure is to find a one-dimensional perturbation of the operator $\bar{H}H$ such that the new operator has the number ρ as a singular number. Afterwards, the solutions will be constructed from the Schmidt pairs associated with the singular number ρ of the perturbed operator. A Schmidt pair (p,q), $p,q\in\ell_2$ corresponding to a singular number s of a Hankel operator is a pair of vectors which satisfy the relations

$$Hp = sq, \quad \bar{H}q = sp.$$

Let R_ρ be the resolvent

$$R_\rho = (\rho^2 I - \bar{H}H)^{-1}. \tag{2.5}$$

For $\rho^2 \notin \mathrm{Sp}(\bar{H}H)$ we define an analytic function of two variables in the bidisk $|a| < 1$, $|z| < 1$:

$$\mathscr{P}_\rho(a,z) = \sum_{m,n=1}^{\infty} (\bar{R}_\rho)_{m,n} a^{m-1} z^{n-1}, \tag{2.6}$$

where $(R_\rho)_{m,n}$ are the matrix elements of R_ρ in the orthonormal base $e_j = (\delta_{ij})_{i=1}^\infty$, $j = 1,2,\cdots$. This object will play an essential role in the construction of solutions of our problem. These solutions will satisfy relation (1.2) taken with the sign equal and are given by the following theorem:

Theorem: Let $\rho > 0$ be a positive number satisfying the inequality (2.4). Then the functions

$$f(\zeta) = \|\rho\mathscr{P}_\rho\|_2^{-2}\, \overline{\mathscr{P}_\rho(a,\bar{\zeta})} \int_{-1}^{1} \frac{\mathscr{P}_\rho(a,x)\, d\,\overline{\sigma(x)}}{1 - x\zeta} \tag{2.7}$$

are solutions of the Riesz problem with norm given by $\|f\|_1 = \rho^{-1}$. These solutions depend on a parameter a, $|a| < 1$, and the root of the equation

$$\mathscr{P}_\rho(a,\bar{a}) = 0. \tag{2.8}$$

Proof: Let $\varphi(z)$ be the Blaschke factor $\varphi(z) = (z-a)\times(1-\bar{a}z)^{-1}$, $|a| < 1$. The operator $H' = H\varphi(T)$, where T is the shift operator, is a Hankel operator since it satisfies the relation $T^*H' = H'T$, which is the relation of definition for a Hankel operator. The shift T acts on ℓ_2 as $T\xi = T(\xi_1,\xi_2,\cdots) = (0,\xi_1,\xi_2,\cdots)$ and its adjoint T^* as $T^*(\xi_1,\xi_2,\cdots) = (\xi_2,\cdots)$. It is easily seen that $\overline{\varphi(T)}\overline{\varphi^*(T)}$ is an orthogonal projection onto the subspace $\overline{\varphi(T)}\ell_2$ and $\dim[\overline{\varphi(T)}\ell_2]^\perp = 1$.

Let e be the normed vector $e = (1-|a|^2)^{1/2}(1,a,a^2,\cdots)$ whose span is the one-dimensional subspace $\ell_2 \Theta\, \overline{\varphi(T)}\ell_2 = [\overline{\varphi(T)}\ell_2]^\perp$.

$$\begin{aligned}\bar{H}'H' &= \bar{H}\,\overline{\varphi(T)}H\varphi(T) = \bar{H}\,\overline{\varphi(T)}\,\overline{\varphi^*(T)}H \\ &= \bar{H}[I - (\ ,e)e]H = \bar{H}H - (\ ,\bar{H}e)\bar{H}e.\end{aligned}$$

This relation shows that $\bar{H}'H'$ and $\bar{H}H$ differ each other by a one-dimensional perturbation. Now it is an easy matter to find an eigenvector for the new operator $\bar{H}'H'$. Let p be the vector

$$p = R_\rho \bar{H}e = \bar{H}\bar{R}_\rho e.$$

The action of $\bar{H}'H'$ on p is

$$\begin{aligned}\bar{H}'H'p &= [\bar{H}H - (\ ,\bar{H}e)\bar{H}e]\,p = \bar{H}H\bar{H}\bar{R}_\rho e - (\bar{H}\bar{R}_\rho e,\bar{H}e)\bar{H}e \\ &= \rho^2 p - \rho^2(\bar{R}_\rho e,e)\bar{H}e.\end{aligned}$$

If we choose a, $|a| < 1$ such that it satisfies the equation

$$(\bar{R}_\rho e,e) = 0 \tag{2.9}$$

the vector p will be an eigenvector corresponding to the eigenvalue ρ^2. The existence of such an a, $|a| < 1$, follows from Stenger.[6] In fact, Eq. (2.9) is nothing else than the equation $\mathscr{P}_\rho(a,\bar{a}) = 0$ written in another form. Since $\mathscr{P}_\rho(a,z)$ is an analytic function of two variables, the dimension of the manifold of the zeros of $\mathscr{P}_\rho(a,\bar{a})$ is equal to unity. So, in general, Eq. (2.9) will have a continuum of solutions.

Now let us find the vector q entering the Schmidt pair associated with the eigenvalue ρ:

$$\begin{aligned}q &= \rho^{-1}H'p = \rho^{-1}H\varphi(T)\bar{H}\bar{R}_\rho e = \rho^{-1}\,\overline{\varphi^*(T)}H\bar{H}\bar{R}_\rho e \\ &= -\rho^{-1}\,\overline{\varphi^*(T)}e + \rho\,\overline{\varphi^*(T)}\,\bar{R}_\rho e.\end{aligned}$$

Since by construction $e \perp \overline{\varphi(T)\ell_2}$, it follows that $\overline{\varphi^*(T)}e = 0$. Hence

$$q = \rho\, \overline{\varphi^*(T)}\, \overline{R}_\rho e.$$

We shall denote by s_k' the singular number of H'. Because $\overline{H}'H'$ and $\overline{H}H$ differ each other by a one-dimensional perturbation, the singular numbers s_k and s_k' verify the separation property[7]

$$s_1 \geqslant s_1' \geqslant s_2 \geqslant s_2' \geqslant \cdots .$$

If ρ is chosen such that $s_k > \rho > s_{k+1}$, this property shows that there is only one Schmidt pair (p,q), i.e., the multiplicity of ρ is equal to unity.

With the pair (p,q) we form the function $g(\zeta) \in H_1$

$$\begin{aligned} g(\zeta) &= \rho^{-1} p_+(\zeta)\, \overline{q_-(\zeta)}\, \bar{\zeta} / \| p_+ \|_2^2 \\ &= \rho^{-1} p_+(\zeta) \overline{q_+(\bar{\zeta})} / \|p_+\|_2^2. \end{aligned} \tag{2.10}$$

The function $g(\zeta)$ has the properties

$$\|g\|_1 = \rho^{-1}, \quad \Phi_{H'}(g) = 1.$$

The first property is a consequence of the unimodularity of $q_-(\zeta)/p_+(\zeta)$, a property which holds for any Schmidt pair of a Hankel operator.[7] The second is evident if we use the relation (2.2). From the same relation (2.2) we get

$$\Phi_{H\psi(T)}(h(\zeta)) = \Phi_H(\psi(\zeta) h(\zeta))$$

for any bounded analytic function $\psi(\zeta)$ and any analytic function $h(\zeta) \in H_1$.

This last relation gives the key of construction of solutions to our problem. Indeed, if we take $\psi(z) = \varphi(z) = (z-a) \times (1-\bar{a}z)^{-1}$, the function

$$f(\zeta) = \varphi(\zeta) g(\zeta) = \rho^{-1} \frac{\zeta - a}{1 - \bar{a}\zeta} p_+(\zeta) \overline{q_+(\bar{\zeta})} / \| p_+ \|_2^2 \tag{2.11}$$

will satisfy the relations

$$\Phi_H(f) = 1, \quad \| f \|_1 = \rho^{-1}.$$

Hence $f(\zeta)$ is a solution of Riesz problem. These solutions depend on the parameter a, the root of Eq. (2.8). They take a simpler form when the Hankel operator H is generated by a sequence of the form (2.3) since then both the functions $p_+(\zeta)$ and $q_+(\zeta)$ are expressed directly in terms of $\mathscr{P}_\rho(a,\zeta)$.

By definition we have

$$\begin{aligned} q_+(z) &= \rho \sum_{n=1}^{\infty} (\overline{\varphi^*(T)}\, \overline{R}_\rho e)_n z^{n-1} \\ &= \rho(\overline{\varphi^*(T)}\, \overline{R}_\rho e, \overline{Z}) = \rho(\overline{R}_\rho e, \overline{\varphi(T)}\, \overline{Z}), \end{aligned}$$

where $\overline{Z}$ is the complex conjugate of $Z = (1, z, z^2, \cdots)$. In order to evaluate the last scalar product, we shall make use of the isometric isomorphism between the Hilbert space ℓ_2 and H_2. By this isomorphism we have the correspondence

$$\overline{\varphi(T)}\, \overline{Z} \in \ell_2 \rightarrow \frac{\zeta - \bar{a}}{1 - a\zeta} \frac{1}{1 - \bar{z}\zeta} \in H_2.$$

Thus

$$\begin{aligned} q_+(z) &= \frac{\rho}{2\pi} \int_0^{2\pi} (\overline{R}_\rho e)_+(\zeta)\, \overline{\frac{\zeta - \bar{a}}{1 - a\zeta} \frac{1}{1 - \bar{z}\zeta}}\, d\theta \\ &= \frac{\rho}{2\pi i} (1 - |a|^2)^{1/2} \oint d\zeta\, \mathscr{P}_\rho(a,\zeta) \frac{1 - a\zeta}{\zeta - \bar{a}} \frac{1}{\zeta - z} \\ &= \rho(1 - |a|^2)^{1/2} \Big(\mathscr{P}_\rho(a,z) \frac{1 - az}{z - \bar{a}} \\ &\quad - \mathscr{P}_\rho(a,\bar{a}) \frac{1 - |a|^2}{z - \bar{a}} \Big). \end{aligned}$$

The last term being equal to zero, we get

$$q_+(z) = \rho(1 - |a|^2)^{1/2} \frac{1 - az}{z - \bar{a}} \mathscr{P}_\rho(a,z).$$

On the other hand,

$$\begin{aligned} p_+(z) &= \sum_{n=1}^{\infty} (\overline{H} \overline{R}_\rho e)_n z^{n-1} = \sum_{m,n=1}^{\infty} \overline{H}_{nm} (\overline{R}_\rho e)_m z^{n-1} \\ &= \sum_{m,n=1}^{\infty} \int_{-1}^{1} d\, \overline{\sigma(x)} x^{m+n-2} (\overline{R}_\rho e)_m z^{n-1} \\ &= (1 - |a|^2)^{1/2} \int_{-1}^{1} \frac{d\, \overline{\sigma(x)} \mathscr{P}_\rho(a,x)}{1 - xz}. \end{aligned}$$

Since $\| p \|_2 = \|q\|_2 = \rho(1 - |a|^2)^{1/2} \| \mathscr{P}_\rho(a,\zeta) \|_2$, the relation (2.11) takes the form (2.7), i.e.,

$$f(\zeta) = \| \rho \mathscr{P}_\rho(a,\zeta) \|_2^{-2} \overline{\mathscr{P}_\rho(a,\bar{\zeta})} \int_{-1}^{1} \frac{d\, \overline{\sigma(x)} \mathscr{P}_\rho(a,x)}{1 - x\zeta},$$

which proves the theorem.

By the above construction we do not get all the solutions of the Riesz problem. Taking $\varphi(z)$, a finite Blaschke product, we can obtain an n-dimensional perturbation of $\overline{H}H$, and similarly one can construct other solutions. All the solutions of the problem will be given by the convex hull of the extremal solutions.

Our conjecture is that the solutions (2.7) are the only extremal solutions of the Riesz problem.

III. CONCLUDING REMARKS

In the following we will present some heuristic considerations concerning the construction of pion–pion partial waves by supposing our conjecture is true.

Let $a = a(\epsilon)$ be the solution of Eq. (2.8), where ϵ goes over the frontier of a domain D entirely contained inside the disk $|a| < 1$. Let $\epsilon(w)$ be the conformal mapping of the unit disk $|w| < 1$ onto this domain. So $a = a(w)$.

The norm $\| \rho \mathscr{P}_\rho(a,\zeta) \|_2$ being positive and depending on w, we can construct an analytic outer function

$$G(w) = \exp\Big(\frac{1}{2\pi} \int_0^{2\pi} \frac{w + e^{i\theta}}{w - e^{i\theta}} \ln \| \rho \mathscr{P}_\rho \|_2\, d\theta \Big)$$

such that

$$|G(e^{i\theta})| = \| \rho \mathscr{P}_\rho \|_2.$$

The solutions (2.7) then take the form

$$f(\zeta) = \frac{\mathscr{P}_\rho \overline{(a(1/\bar{w}), \bar{\zeta})} \int_{-1}^{1} d\, \overline{\sigma(x)} \mathscr{P}_\rho(a(w),x)/(1 - x\zeta)}{\overline{G(1/\bar{w})} G(w)},$$

where $w = \exp(i\theta)$.

Since the domain D is entirely contained inside the disk $|a| < 1$, $G(w)$ will be analytic in a bigger domain $|w| \leqslant R$, $R > 1$, the singularities of $G(w)$ appearing for $|a(w)| \rightarrow 1$, as one can see from (2.6). Thus $\overline{G(1/\bar{w})}$ will be analytic in the exterior of a disk of radius $r = 1/R$. Most probably $\overline{G(1/\bar{w})}$ will have a cut $-1 < -r \leqslant w \leqslant r < 1$. Thus $f(\zeta)$ will be also analytic in w in a domain that will contain at least the annulus $1/R < w < R$. Let $w(z)$ be the conformal mapping of this domain onto a standard annulus. Now we may try to construct the convex hull of solutions (2.7) by using the geometric methods of Adamyan, *et al.*[7]

In any case the above method requires the knowledge of the resolvent operator R_ρ. For pion–pion scattering this amounts to solving the eigenvalue equation

$$H\varphi = s\varphi,$$

and, if H is generated by the sequence (1.4), the equation has the form

$$\int_{-1}^{1} \frac{(1-x^2)^{1/2}}{1-xy}\varphi_k(x)\,dx = \tfrac{1}{2}\pi\cdot s_k\varphi_k(y).$$

On this form one does not see explicitly that H is a Hankel operator, but one can easily show that this is the case.

Indeed, let $K(x,y)$ be the kernel of the integral operator $K: R_+\times R_+ \to R$ defined as

$$K(x,y) = \pi^{-1/2}\psi(3/2,3/2;(x+y)/2)$$
$$= (2/\pi)\int_0^\infty \exp[-t(x+y)/2]\,[t^{1/2}/(t+1)]\,dt,$$

where $\psi(a,c;x)$ is the confluent hypergeometric function of second kind.[8]

In the orthonormal base

$$\Phi_n(x) = \exp(-x/2)L_{n-1}(x),\quad n = 1,2,\cdots,$$

where $L_n(x)$ are the Laguerre polynomials, we have

$$(K\Phi_n)(x) = \frac{2}{\pi}\int_0^\infty dt\, e^{-tx/2} \times \frac{t^{1/2}}{t+1}\left(\frac{t-1}{2}\right)^{n-1}\left(\frac{t+1}{2}\right)^{-n}$$

and

$$(K\Phi_n,\Phi_m) = \frac{8}{\pi}\int_0^\infty dt\,\frac{t^{1/2}(t-1)^{m+n-2}}{(t+1)^{m+n+1}}$$
$$= \frac{2}{\pi}\int_{-1}^{1}(1-x^2)^{1/2}x^{m+n-2}\,dx$$
$$= \frac{1+(-1)^{m+n}}{\pi}B\left(\frac{3}{2},\frac{m+n-1}{2}\right),$$

i.e., the operators H and K are similar. Thus another form of the eigenvalue equation is

$$\int_0^\infty dx\,\psi(3/2,3/2;(x+y)/2)\,g(x) = \pi^{1/2}sg(y)$$

Although until now only very few Hankel operators have been diagonalized, we hope that the experts in the field will find the spectrum and the eigenfunctions of this equation, facilitating the solving of the Riesz problem at least for the pion–pion scattering.

[1]F. Riesz, Acta Math. **42**, 145 (1920).
[2]R. Roskies, Nuovo Cimento A **65**, 467 (1970).
[3]A. Martin, *Scattering Theory: Unitarity, Analyticity and Crossing* (Springer-Verlag, Heidelberg, 1969).
[4]P. Diţă, Int. J. Theor. Phys. **16**, 201 (1977).
[5]V. M. Adamyan, D. Z. Arov, and M. G. Krein, Funct. Anal. Appl. **2**, 1 (1968).
[6]W. Stenger, J. Math. Anal. Appl. **28**, 625 (1969).
[7]V. M. Adamyan, D. Z. Arov, and M. G. Krein, Math. USSR-Sbornik **15**, 31 (1971).
[8]A. Erdély, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1.

Minimal electromagnetic coupling schemes. I. Symmetries of potentials and gauge transformations

J. Beckers and V. Hussin[a)]
Université de Liège, Physique théorique et mathématique, Institut de Physique au Sart Tilman, Bâtiment B.5, B-4000 Liège 1, Belgium

(Received 20 January 1982; acccepted for publication 19 March 1982)

Minimal electromagnetic coupling schemes entering into Klein–Gordon or Schrödinger equations are studied in connection with symmetries *inside and outside* the symmetry groups of the corresponding *free* equations. Subsymmetries of relativistic potentials are classified up to conjugacy under the kinematical groups of the associated (constant and uniform) electromagnetic fields. Through invariance conditions on (physical) four-potentials a *maximal* character of the symmetry is obtained leading to the maximal symmetry groups of the corresponding *wave equations with interaction*. Usual and compensating gauge transformations are analyzed within the context of such invariance conditions applied to arbitrary electromagnetic fields.

PACS numbers: 11.30.Cp, 11.10.Qr, 02.20. + b, 11.15. − q

1. INTRODUCTION

The quantum theory of charged particles in classical (external) electromagnetic fields has always been considered as an interesting part of "physics with interaction". For example, the current text books in quantum mechanics present and extend the correspondence principle between classical and quantum physics through electromagnetic couplings; they also describe and discuss, for example, the Dirac equation with *minimal* coupling to the electromagnetic field (in order to get the value of the Landé factor, etc.).

In the last fifteen years, several contributions on such classical electromagnetic interactions have appeared in the specialized literature; they essentially developed *group-theoretical* studies applied to nonrelativistic and (or) relativistic particle physics. Let us mention more particularly the approaches of Bacry, Combe and Richard,[1,2] Janner and Janssen,[3] Schrader,[4] Combe and Richard,[5] Giovannini,[6] Hoogland,[7,8] and Hussin.[9,10]

All these studies deal with *minimal electromagnetic coupling schemes* (m.e.c.s.) or with *symmetries* of associated electromagnetic *fields* or *potentials*. These considerations concern well-known wave equations such as those of Schrödinger, Klein–Gordon, Dirac,···. In fact, the above authors have studied, on the one hand, the nonrelativistic context[2,3,5,7,8] subtended by the Galilei symmetry group[11] and (or), on the other hand, the relativistic context[1,3–9] dealing with the Poincaré symmetry group.[12] They all consider *interactions* which can essentially be studied in connection with two categories of symmetries:

i) *inside* the Galilei or (and) Poincaré symmetry group(s) such as the Bacry–Combe–Richard methods,[1,2] for example;

ii) *outside* these symmetry groups, considering their extensions[11,13–15] and related properties such as the works of Schrader[4] or Hoogland,[7,8] for example.

Let us also notice that if some of the above mentioned contributions do study symmetries of the electromagnetic *fields*,[1,2,4,7,8] others really consider symmetries of the *potentials*.[3,6,9,10] Moreover we remark that the m.e.c.s. enter into the wave equations *either* by construction[4] *or* by deduction[7,8] depending on the purpose the authors follow. In such wave equations with interaction, let us recall that *potentials* and not fields occur; consequently symmetries of the wave equations are effectively correlated to symmetries of the potentials (corresponding to specific m.e.c.s.) although we evidentlyknow that the physical quantity is the electromagnetic field $F\equiv(\mathbf{E},\mathbf{B})$ and that the scalar (V) and vector ($\mathbf{A}$) potentials leading to such a field F fall into equivalence classes (through usual gauge transformations).

In this note, we want to enhance and analyze the symmetries of potentials and their associated m.e.c.s. into wave equations and, more specifically, into Klein–Gordon equations. Our approach is different with respect to others[3,6]; our procedure shall combine invariance conditions[16] on fourvectors[17] *and* physical constraints induced from the Maxwell theory (in the relativistic case). When the Schrödinger equation and its nonrelativistic context are considered, we can also refer to a Galilean electromagnetic field F (after Levy-Leblond and Le Bellac,[18] for example) and to an m.e.c.s. described in terms of the corresponding potentials V and $\mathbf{A}$. These potentials can also be submitted to invariance conditions.

Effectively we will deal with *both* of the above categories called "inside *and* outside the symmetry group" associated with the *free* particle description we refer to. With the study of *interactions*, we will get informations *either* "inside the symmetry group" through its nonequivalent subgroups as will be clear in this paper (hereafter quoted as I), *or* "outside the symmetry group" through its extensions as will be clear in the next paper (hereafter quoted as II). Our study will always be connected with the m.e.c.s. introduced into relativistic Klein–Gordon or nonrelativistic Schrödinger wave equations.

Let us now describe the contents of this paper I and give some informations about II, these two contributions being interdependent in what concerns essentially the point of view of symmetries associated with m.e.c.s. *outside* the (*free*) symmetry group. *This* paper deals with inside and outside

[a)] Boursier I.R.S.I.A. 1981–1982.

symmetry properties and can be considered in the relativistic context as well as in the nonrelativistic one. For brevity we shall limit ourselves to the consideration of the relativistic case but since both (relativistic and nonrelativistic) descriptions can be set on an equal footing, we devote Sec. 2 to generalities on Poincaré and Galilean symmetries. We particularly insist on kinematical groups,[1,2] invariance conditions on potentials[17] which, in the Galilean context, are established for the first time, and on transformation laws of the electromagnetic fields in each context. This summary will be of special interest for the two papers. Symmetries *inside* the Poincaré group are studied in Sec. 3. By taking the explicit case of Klein–Gordon equations with minimal couplings, we first discuss specific realizations of the symmetry operators. We relate them to covariant derivatives (Sec. 3a) given in terms of four-potentials as usual, but more particularly in terms of the so-called symmetric gauge.[1,19] The Poincaré subsymmetries on four-potentials A are then discussed (Sec. 3b) when these potentials lead to a *constant and uniform* electromagnetic field $F^{(0)}$. In correspondence with the only two[1] nonequivalent fields $F^{(0)}$ (i.e., the parallel $F_{\parallel}^{(0)}$ *and* the perpendicular $F_{\perp}^{(0)}$) we determine up to conjugacy the nonequivalent subgroups of the kinematical groups[1] $G_{F_\parallel}^{(0)}$ and $G_{F_\perp}^{(0)}$ in order to get those of *maximal* dimension leading to some potentials $A_{\parallel}$ and $A_{\perp}$, respectively. Here we apply the Patera–Winternitz–Zassenhaus method[20,21] of subgroup classifications combined with our invariance conditions on physical four-vectors.[17] The results are collected in two tables; we notice that the maximal dimensions are $n_{\parallel} = 3$ and $n_{\perp} = 4$, respectively, and we determine some of these invariant four-potentials. Finally (Sec. 3c), we establish some properties and consequences in connection with gauge-symmetrical potentials, realizations of generators and symmetries of the associated Klein–Gordon equations. In Sec. 4, we make a few further remarks on compensating gauge transformations[3,6] and symmetry properties independent of specific forms of potentials. Here we effectively consider *arbitrary* (nonconstant) electromagnetic fields and show other interests of our invariance conditions[16,17] on four-potentials in connection with compensating gauges, for example. Section 5 is devoted to some comments on symmetries *inside* the *Galilei* group when nonrelativistic m.e.c.s. are considered; through a parallel study of Sec. 3 with the elements displayed in Sec. 2b, we obtain the corresponding results. For brevity, we only quote some interesting facts enhancing the differences between nonrelativistic and relativistic considerations.

As already mentioned, the next paper II will concern more effectively m.e.c.s. and their associated symmetry properties *outside* the (free) symmetry group through group extensions of the Poincaré and Galilei groups. The Schrader developments[4] will there be of first importance, but their connections with our considerations (essentially Sec. 3a) will be studied in the relativistic context. The Galilean Maxwell group will also be constructed starting from our considerations given in Sec. 2b.

In order to simplify the equations of papers I and II, we have chosen to use the natural units ($\hbar = 1$, $c = 1$) and the summation convention on repeated indices (Greek indices running from 0 to 3 and Latin ones from 1 to 3). Moreover, let us mention that we did not give further useful information on necessary notions such as group extensions, factor sets, compensating gauges,..., but only refer for brevity to the original contributions.

2. POINCARÉ AND GALILEAN SYMMETRIES—GENERALITIES AND NOTATIONS

In order to study electromagnetic coupling schemes in the context of relativistic and nonrelativistic theories and equations, we have to give a very brief survey of the necessary elements on Poincaré and Galilean invariances. Let us present some useful considerations in the following two subsections.

2a. Poincaré symmetry and its implications

Minkowski space-time characterized by the tensor metric $(+,-,-,-)$ refers to events $x \equiv \{x^\mu(\mu = 0,1,2,3)\} = \{x^0, x^i (i = 1,2,3)\} \equiv \{t,\mathbf{x}\}$ submitted to Poincaré (or restricted inhomogeneous Lorentz) transformations $g \equiv (a, \Lambda)$ where a and Λ are, as usual, associated with space-time translations (generators P^μ) and restricted homogeneous Lorentz transformations (generators $M^{\mu\nu} = -M^{\nu\mu}$), respectively. The set $\{g\}$ is the so-called Poincaré group P with the multiplication law

$$gg' = (a + \Lambda a', \Lambda\Lambda') = (a'', \Lambda'') = g''. \tag{2.1}$$

Its elements act on x in the following way:

$$x \to x' = gx : x^\mu \to x'^\mu = (gx)^\mu = (\delta^\mu{}_\nu + \omega^\mu{}_\nu)x^\nu + a^\mu \tag{2.2}$$

when *infinitesimal* parameters a^μ and $\omega^{\mu\nu} = -\omega^{\nu\mu}$ are considered. In connection with the usual choice

$$\omega = (\boldsymbol{\phi},\boldsymbol{\theta}),\quad \phi^i = \omega^{0i},\quad \theta^i = \tfrac{1}{2}\epsilon^{ijk}\omega_{jk},\quad \epsilon^{123} = 1, \tag{2.3}$$

the six Lorentz generators are defined by

$$\{M^{\mu\nu}\} \equiv (\mathbf{K},\mathbf{J}), M^{0i} = K^i \text{ and } J^i = \tfrac{1}{2}\epsilon^{ijk}M_{jk} \tag{2.4}$$

in correspondence with pure Lorentz transformations (boosts) and spatial rotations, respectively. So the Lie algebra of the Poincaré group can then be written

$$\begin{aligned}
&[J^j,J^k] = i\epsilon^{jk}{}_l J^l,\quad [J^j,K^k] = i\epsilon^{jk}{}_l K^l,\\
&[K^j,K^k] = -i\epsilon^{jk}{}_l J^l,\\
&[J^j,P^k] = i\epsilon^{jk}{}_l P^l,\quad [K^j,P^k] = i\delta^{jk}P^0,\\
&[K^j,P^0] = iP^j,
\end{aligned} \tag{2.5}$$

all the other commutators being equal to zero. A very well-known realization of these generators is given by

$$\begin{aligned}
&P_\mu = -i\partial_\mu = -i\frac{\partial}{\partial x^\mu}\,(P^0 = -i\partial_t, \mathbf{P} = i\nabla),\\
&\mathbf{J} = i(\mathbf{x}\Lambda\nabla),\quad \mathbf{K} = i(t\nabla + \mathbf{x}\partial_t).
\end{aligned} \tag{2.6}$$

Under finite Poincaré transformations, tensors transform very simply. For example, second rank tensors $F(x) \equiv \{F^{\mu\nu}(x)\}$ and 4-vectors $A(x) \equiv \{A^\mu(x)\}$ have the following transformation laws:

$$(gF)^{\mu\nu}(x) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F_{\rho\sigma}(g^{-1}x) \tag{2.7}$$

and

$$(gA)^\mu(x) = \Lambda^\mu{}_\nu A^\nu(g^{-1}x). \tag{2.8}$$

In particular, specific tensors are of special physical interest; so is the (antisymmetric) second rank electromagnetic tensor $F \equiv \{F^{\mu\nu}\}$ given in terms of the electric and magnetic fields by

$$F \equiv \{F^{\mu\nu}\} \equiv (\mathbf{E},\mathbf{B}),\ F^{0i} = E^i,\ B^i = \tfrac{1}{2}\epsilon^{ijk}F_{jk}, \tag{2.9}$$

or the physical 4-potential $A \equiv \{V,\mathbf{A}\}$ where V is the scalar potential and $\mathbf{A}$ the vector potential. It is connected to F (through Maxwell theory) by

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu. \tag{2.10}$$

This corresponds to the usual definitions

$$\mathbf{E} = -\operatorname{grad} V - \partial_t \mathbf{A}, \quad \mathbf{B} = \operatorname{rot}\mathbf{A}. \tag{2.11}$$

Let us also mention the dual of F defined by

$$*F \equiv \{*F^{\mu\nu}\} = (-\mathbf{B},\mathbf{E}), \quad *F^{\mu\nu} = x - \tfrac{1}{2}\epsilon^{\mu\nu}F^{\rho\sigma}, \quad (\epsilon^{0123} = 1) \tag{2.12}$$

and the Maxwell equations

$$\partial_\nu F^{\mu\nu} = j^\mu, \quad \partial_\nu {*F^{\mu\nu}} = 0, \tag{2.13}$$

where $j \equiv \{j^\mu\} \equiv \{\rho,\mathbf{j}\}$ is the current 4-vector. The first set of Eqs. (2.13) corresponds to the Maxwell equations with sources

$$\operatorname{rot}\mathbf{B} - \frac{\partial}{\partial t}\mathbf{E} = \mathbf{j}, \quad \operatorname{div}\mathbf{E} = \rho, \tag{2.14}$$

and the second set is equivalent to the definitions (2.11).

Electromagnetic tensors *invariant* under Poincaré transformations, i.e., such that

$$gF = F \tag{2.15}$$

from Eq. (2.7), have already been studied.[22] Under *infinitesimal* transformations, the invariance conditions are:

$$\begin{aligned} &\boldsymbol{\theta}\Lambda\,\mathbf{E} - \boldsymbol{\phi}\Lambda\,\mathbf{B} + \mathscr{D}\mathbf{E} = 0, \\ &\boldsymbol{\theta}\Lambda\,\mathbf{B} + \boldsymbol{\phi}\Lambda\,\mathbf{E} + \mathscr{D}\mathbf{B} = 0 \end{aligned} \tag{2.16}$$

with

$$\mathscr{D} \equiv (t\boldsymbol{\phi} + \mathbf{x}\Lambda\boldsymbol{\theta})\cdot\frac{\partial}{\partial\mathbf{x}} + (\mathbf{x}\cdot\boldsymbol{\phi})\frac{\partial}{\partial t} - a\cdot\nabla. \tag{2.17}$$

In particular if, after Bacry–Combe–Richard,[1] we are interested in the *constant and uniform* case ($\mathscr{D}\mathbf{E} = \mathscr{D}\mathbf{B} = 0$), it is easy to determine the so-called *kinematical group* $G_F^{(0)}$ of F which is a Poincaré subgroup of dimension six

$$G_F^{(0)} \equiv \{P^\mu, Q, Q^*\}, \tag{2.18}$$

where

$$Q = \mathbf{J}\cdot\mathbf{B} - \mathbf{K}\cdot\mathbf{E}, \quad Q^* = \mathbf{J}\cdot\mathbf{E} + \mathbf{K}\cdot\mathbf{B}. \tag{2.19}$$

In fact, we essentially get two nonisomorphic kinematical subgroups in correspondence with the only two nonequivalent constant and uniform tensors $F_\parallel$ and $F_\perp$. We notice that if

$$\text{i)}\ F_\parallel \equiv \{\mathbf{E} \equiv (0,0,E),\ \mathbf{B} \equiv (0,0,B)\} \tag{2.20}$$

we get

$$G_{F_\parallel}^{(0)} \equiv \{J^3, K^3, P^\mu\}; \tag{2.21}$$

$$\text{ii)}\ F_\perp \equiv \{\mathbf{E} \equiv (E,0,0),\ \mathbf{B} \equiv (0,E,0)\} \tag{2.22}$$

we have

$$G_{F_\perp}^{(0)} \equiv \{\mathscr{A}^1 = J^1 + K^2, \mathscr{A}^2 = J^2 - K^1, P^\mu\}. \tag{2.23}$$

Finally, 4-vectors *invariant* under Poincaré transformations, i.e., such that

$$gA = A \tag{2.24}$$

from Eq. (2.8), have also been considered[16,17] leading to the following invariance conditions:

$$\boldsymbol{\phi}\cdot\mathbf{A} - \mathscr{D}V = 0, \quad \boldsymbol{\theta}\Lambda\mathbf{A} - \boldsymbol{\phi}V + \mathscr{D}\mathbf{A} = 0, \tag{2.25}$$

where $\mathscr{D}$ is given once again by (2.17). The corresponding kinematical group $G_A^{(0)}$ of a (constant and uniform) A is also a Poincaré subgroup, but of dimension seven; its explicit structure has already been discussed.[17]

2b. Galilean symmetry and its implications

Space-time events of the Newtonian world can also be referred to by $x \equiv (t,\mathbf{x})$ but are submitted to Galilean transformations[11] $g \equiv (b,\mathbf{a};\mathbf{v},R)$, where b, $\mathbf{a}$, $\mathbf{v}$, and R are the parameters associated with time and spatial translations, pure Galilean transformations and spatial rotations, respectively. The corresponding generators are then denoted by H, $\mathbf{P}$, $\mathbf{K}$, and $\mathbf{J}$. Such Galilean transformations act on x in the following way:

$$x \to x' = gx: gt = t + b,\ g\mathbf{x} = R\mathbf{x} + \mathbf{v}t + \mathbf{a}, \tag{2.26}$$

and form the so-called Galilei group with the multiplication law

$$\begin{aligned} gg' &= (b + b', \mathbf{a} + R\mathbf{a}' + \mathbf{v}b'; \mathbf{v} + R\mathbf{v}', RR') \\ &= (b'', \mathbf{a}''; \mathbf{v}'', R'') = g''. \end{aligned} \tag{2.27}$$

The corresponding Lie algebra has the following nonzero commutation relations[11]:

$$\begin{aligned} &[J^j, J^k] = i\epsilon^{jk}{}_l J^l, \quad [J^j, K^k] = i\epsilon^{jk}{}_l K^l, \\ &[J^j, P^k] = i\epsilon^{jk}{}_l P^l, \quad [K^j, H] = iP^j. \end{aligned} \tag{2.28}$$

By comparison with the Poincaré algebra (2.5), let us notice the (expected) differences in the commutation relations between the $\mathbf{K}$'s among themselves and between the $\mathbf{K}$'s and $\mathbf{P}$'s. A well known[11] realization of the generators is also given by (2.6) with $P^0 \equiv H$ but when the $\mathbf{K}$'s have the form

$$\mathbf{K} = it\boldsymbol{\nabla}. \tag{2.29}$$

Now as the Galilean context deals with nonrelativistic considerations, it makes no sense to study Maxwell theory without modifications. Different works have already analyzed such questions and, after Le Bellac–Levy-Leblond,[18] we maintain the discussion of an "electromagnetic" field F defined in terms of electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields, satisfying the following equations:

$$\operatorname{rot}\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0, \quad \operatorname{div}\mathbf{B} = 0, \tag{2.30}$$

and

$$\operatorname{rot}\mathbf{B} = \mathbf{j}, \quad \operatorname{div}\mathbf{E} = \rho. \tag{2.31}$$

Such an "electromagnetic" theory refers to the "magnetic limit[18]" and is invariant under the Galilean transformations. Let us notice that Eqs. (2.30) leave the definitions of $\mathbf{E}$ and $\mathbf{B}$ in the same form as in the relativistic case [cf. Eqs. (2.11)] and lead to the introduction of a scalar potential $V(x)$ and a vector potential $\mathbf{A}(x)$. Then it is easy to establish the following transformation laws[18] (the laws analogous to Eqs. (2.7) and (2.8) in the relativistic context):

$$g\mathbf{E}(x) = R\mathbf{E}(g^{-1}x) - \mathbf{v}\Lambda R\mathbf{B}(g^{-1}x), \tag{2.32a}$$

$$g\mathbf{B}(x) = R\mathbf{B}(g^{-1}x), \tag{2.32b}$$

and

$$gV(x) = V(g^{-1}x) + \mathbf{v}\cdot R\mathbf{A}(g^{-1}x), \tag{2.33a}$$

$$g\mathbf{A}(x) = R\mathbf{A}(g^{-1}x). \tag{2.33b}$$

Equations (2.32) ensure the invariance of the theory summarized by Eqs. (2.30) and (2.31).

In the following, we will always refer to *Galilean electromagnetic fields* $F \equiv (\mathbf{E},\mathbf{B})$ when the above considerations are taken into account.

Now let us deal with Galilean electromagnetic fields *invariant* under Galilean transformations in complete analogy with the relativistic considerations. Bacry, Combe, and Richard[2] have determined the possible structures of the kinematical group associated with a *constant and uniform* Galilean electromagnetic field. If $\mathbf{B}\neq 0$, they got a Galilei subgroup of *dimension six*

$$G_F^{(0)\prime} \equiv \{H,\mathbf{P},Q',Q^{*\prime}\}, \tag{2.34}$$

where

$$Q' = \mathbf{B}\cdot\mathbf{J} - \mathbf{E}\cdot\mathbf{K},\ Q^{*\prime} = \mathbf{B}\cdot\mathbf{K}, \tag{2.35}$$

so that if we consider the case $F_{\parallel} \equiv \{\mathbf{E}\equiv(0,0,E),\mathbf{B}\equiv(0,0,B)\}$, the kinematical group $G_{F_{\parallel}}^{(0)\prime}$ is:

$$G_{F_{\parallel}}^{(0)\prime} \equiv \{H,\mathbf{P},J^3,K^3\}. \tag{2.36}$$

It is the "same" subgroup as the one given by Eq. (2.21) in the relativistic case. If $\mathbf{B}\equiv 0$, the corresponding kinematical group is of *dimension eight*

$$G_F^{(0)\prime\prime} \equiv \{H,\mathbf{P},\mathbf{K},Q''\}, \tag{2.37}$$

with

$$Q'' = \mathbf{E}\cdot\mathbf{J}. \tag{2.38}$$

Finally we can determine the invariance conditions on Galilean electromagnetic fields and potentials in the sense of (2.15) and (2.24) of the preceding subsection, but when the transformation laws (2.32) and (2.33) are considered. We get, respectively,

$$\boldsymbol{\theta}\Lambda\mathbf{E} - \mathbf{v}\Lambda\mathbf{B} + \mathscr{D}'\mathbf{E} = 0, \tag{2.39}$$

$$\boldsymbol{\theta}\Lambda\mathbf{B} + \mathscr{D}'\mathbf{B} = 0, \tag{2.40}$$

and

$$\mathbf{v}\cdot\mathbf{A} + \mathscr{D}'V = 0, \tag{2.41}$$

$$\boldsymbol{\theta}\Lambda\mathbf{A} + \mathscr{D}'\mathbf{A} = 0, \tag{2.42}$$

where

$$\mathscr{D}' \equiv (\mathbf{x}\Lambda\boldsymbol{\theta} - \mathbf{v}t - \mathbf{a})\cdot\nabla - b\partial_t. \tag{2.43}$$

The constant and uniform particular case ($\mathscr{D}'\mathbf{E} = \mathscr{D}'\mathbf{B} = 0$) is a starting point which can lead to Bacry, Combe, Richard's results[2] and the discussion of the kinematical groups (2.34) and (2.37). For *invariant* constant and uniform potentials $(V,\mathbf{A})$ we get a kinematical group generated by *seven* operators, as in the relativistic case.

3. SYMMETRIES OF RELATIVISTIC POTENTIALS

We have already discussed[9] m.e.c.s. following Hoogland's developments.[7,8] In fact, starting from kinematical groups such as those given by Eqs. (2.21) or (2.23), their extensions[5] and some of their (equivalent) exponents, we constructed Klein–Gordon-type equations with specific four-potentials issued from realizations of the generators of the extensions.

In this section, let us *first* make a few comments on these specific realizations associated with the Klein–Gordon equations, on the corresponding four-potentials (leading to a *constant and uniform* $F^{(0)}$) and the associated covariant derivatives (Sec. 3a). *Secondly,* let us consider (Sec. 3b) the problem of Poincaré subsymmetries of such four-potentials and their *maximal* character through group theoretical arguments associated with invariance conditions. *Thirdly,* let us mention some properties and consequences (Sec. 3c) directly obtained from the preceding results.

3a. Realizations of generators and associated four-potentials

If the parallel F-case is taken into account with $G_{F_{\parallel}}^{(0)} \equiv$(2.21), the extension of R by $G_{F_{\parallel}}^{(0)}$, denoted by $\bar{G}_{F_{\parallel}}^{(0)}$, has an algebra characterized by the nonzero commutation relations[8]

$$\begin{aligned}
&[J^3,\pi^1] = -i\pi^2,\ [J^3,\pi^2] = i\pi^1,\\
&[K^3,\pi^0] = i\pi^3,\ [K^3,\pi^3] = i\pi^0,\\
&[\pi^1,\pi^2] = ieB,\ [\pi^0,\pi^3] = ieE.
\end{aligned} \tag{3.1}$$

Some realizations of these generators can easily be determined in correspondence with specific exponents.[9] With an exponent called[9] ξ_1 we got the realization of generators π_μ, J^3 and K^3 issued from Eqs. (2.6) but with the substitution of P_μ by π_μ, i.e.,

$$\begin{aligned}
&\pi_\mu = P_\mu + eA_\mu^{(0)},\\
&J^3 = i(x\partial_y - y\partial_x),\quad K^3 = i(t\partial_z + z\partial_t),
\end{aligned} \tag{3.2}$$

where $A_\mu^{(0)}$ is the gauge symmetrical potential

$$A_\mu^{(0)}(x) = \tfrac{1}{2}F_{\mu\nu}x^\nu \tag{3.3}$$

given here by

$$V_{\parallel}^{(0)} = -\tfrac{1}{2}Ez,\quad \mathbf{A}_{\parallel}^{(0)} = \tfrac{1}{2}(-By,Bx,-Et). \tag{3.4}$$

The method for determining the equation of motion is to consider the Casimir operator of the extension $\bar{G}_{F_{\parallel}}^{(0)}$. So with the realization (3.2), we get the equation

$$(D_\mu D^\mu + m^2)\psi(x) = 0, \tag{3.5}$$

where

$$D_\mu = \partial_\mu - ieA_{\mu,\parallel}^{(0)} \equiv D_\mu^{(0)}. \tag{3.6}$$

In terms of these covariant derivatives (3.6), we notice that:

$$\pi_\mu = -iD_\mu^{(0)} + 2eA_\mu^{(0)}, \tag{3.7}$$

a simple fact due to the consideration of the exponent ξ_1.

Now if we take an *equivalent* exponent, called[9] ξ_1' for example, we got another (equivalent) realization

$$\begin{aligned} \pi'_\mu &= P_\mu + eC_\mu, \\ J'^3 &= J^3 - \tfrac{1}{2}eB(x^2 - y^2), \\ K'^3 &= K^3 - \tfrac{1}{2}eE(t^2 + z^2), \end{aligned} \tag{3.8}$$

with

$$C \equiv (C^0, \mathbf{C}): C^0 = 0,\ \mathbf{C} = (-By, 0, -Et). \tag{3.9}$$

We also get Eq. (3.5) with this realization but with new covariant derivatives

$$D'_\mu = \partial_\mu - ieA'_\mu, \tag{3.10}$$

where

$$V' = -Ez, \quad \mathbf{A}' = (0, Bx, 0). \tag{3.11}$$

This four-potential does *not* enter into the π'_μ's in contrast with the preceding case [cf. Eqs. (3.2) with (3.4)]. The correction of such a "defect" can be given in terms of gauge transformations. Indeed it is easy to show that the potentials (3.4) and (3.11) are physically equivalent,

$$A'_\mu = A^{(0)}_\mu + \partial_\mu \lambda, \tag{3.12}$$

where the gauge function λ is given by

$$\lambda = -\tfrac{1}{2}(Ezt + Bxy). \tag{3.13}$$

So we notice that the realization (3.8) can also be rewritten with Eq. (3.10)

$$\begin{aligned} \pi'_\mu &= -iD'_\mu + 2eA^{(0)}_\mu, \\ J'^3 &= [J^3(1 - ie\lambda)], \\ K'^3 &= [K^3(1 - ie\lambda)]. \end{aligned} \tag{3.14}$$

Through these considerations we clearly see the interesting role of the symmetric gauge, the generality of Eq. (3.7), the infinity of realizations we can get with specific gauge functions λ, *and* their interdependence at the level of all the realizations of the generators associated with equivalent exponents.

If the perpendicular F-case is studied with $G^{(0)}_{F_\perp} \equiv (2.23)$, the extension $\bar{G}^{(0)}_{F_\perp}$ is given by the nonzero commutation relations

$$\begin{array}{ll} [\mathscr{A}^1, \pi^0] = i\pi^2, & [\mathscr{A}^2, \pi^0] = -i\pi^1, \\ [\mathscr{A}^1, \pi^1] = 0, & [\mathscr{A}^2, \pi^1] = -i(\pi^0 - \pi^3), \\ [\mathscr{A}^1, \pi^2] = i(\pi^0 - \pi^3), & [\mathscr{A}^2, \pi^2] = 0, \\ [\mathscr{A}^1, \pi^3] = i\pi^2, & [\mathscr{A}^2, \pi^3] = -i\pi^1, \\ [\pi^0, \pi^1] = ieE, & [\pi^3, \pi^1] = ieE. \end{array} \tag{3.15}$$

We can also obtain[9] a specific realization from the gauge symmetrical potential (3.3) which takes the form

$$V^{(0)}_\perp = -\tfrac{1}{2}Ex, \quad \mathbf{A}^{(0)}_\perp = \tfrac{1}{2}E(z - t, 0, -x). \tag{3.16}$$

We then have

$$\begin{aligned} \pi_\mu &= P_\mu + eA^{(0)}_{\mu,\perp}, \\ \mathscr{A}^1 &= i\{ y(\partial_t + \partial_z) + (t - z)\partial_y \}, \\ \mathscr{A}^2 &= -i\{ x(\partial_t + \partial_z) + (t - z)\partial_x \}. \end{aligned} \tag{3.17}$$

It is easy once again to show the validity of Eq. (3.7) with the corresponding covariant derivatives expressed in terms of (3.16). If equivalent exponents are also considered in this "perpendicular" case, the developments analogous to the above ones can be done and the conclusions once again hold.

Let us finally remark that in both cases $F^{(0)}_\parallel$ and $F^{(0)}_\perp$, the symmetric gauge potentials [respectively given by (3.4) and (3.16)] lead to realizations of the extensions where the *homogeneous* Lorentz part is unchanged according to Eqs. (2.6) and where the new *translation* generators π_μ are only submitted to modifications according to Eq. (3.7).

3b. Four-potentials and group theory

The space-time dependent four-potentials have a well-defined *Poincaré* subsymmetry. By the use of the invariance conditions (2.25) we have shown,[9] for example, that the potentials (3.4), (3.11), and (3.16) admit the following *Poincaré* subgroups: $\{J^3, K^3\}$, $\{P^0, P^2\}$, and $\{\mathscr{A}^1, \mathscr{A}^2, P^2, P^0 - P^3\}$, respectively, as symmetry groups. Then let us ask the following question: "among the infinity of equivalent potentials leading to a constant and uniform $F^{(0)}$, what are those which admit a *maximal* Poincaré subsymmetry of the field?" Such informations will give, *inside the Poincaré group*, the corresponding *maximal* symmetry groups for the associated Klein–Gordon equations with minimal couplings.

In order to solve such a problem we consider two steps

i) to determine all the nonequivalent Poincaré subgroups of the kinematical groups $G^{(0)}_{F_\parallel}$ or $G^{(0)}_{F_\perp}$;

ii) to examine among these subgroups those which are of maximal dimension and admit *invariant* four-potentials A leading to the associated field $F^{(0)}$. We then speak of *maximal symmetry* and the corresponding subgroups G_A will be denoted $G^{\max}_A$.

The *first step* can be realized by the application of the Patera–Winternitz–Zassenhaus[20] method on subgroup classifications applied up to conjugacy under $G^{(0)}_F$. Let us only recall the main points of the classification algorithm[20] when we deal with associated algebras and subalgebras. If the Lie algebra (we use the same notations for groups and algebras)

$$G^{(0)}_F = L \oplus N,$$

where L is the homogeneous part of the Lorentz algebra and N is an abelian ideal ($N \equiv \{P^\mu\}$), we search for

a) all conjugacy classes (under the homogeneous part) of subalgebras L_i of L, and

b) all the subspaces of N invariant for each L_i.

Then all the so-called splitting and nonsplitting subalgebras can be determined.

Table I gives all the nonequivalent subalgebras of $G^{(0)}_{F_\parallel} \equiv (2.21)$ but of dimension $\geqslant 3$ and Table II contains those of $G^{(0)}_{F_\perp} \equiv (2.23)$ up to dimension 4, these dimensions being sufficient in connection with our purposes. In these two tables we use the Patera–Winternitz–Zassenhaus notations and refer to splitting algebras $G_{i,j}$ and to nonsplitting algebras $\tilde{G}_{i,j}$. The index i corresponds to the structure of the homogeneous part L_i and the index j refers to the different subspaces N_j of N for each L_i. In these tables, the parameters a, b, α, and ρ are arbitrary real numbers; let us notice that a, b, and α are always nonzero except in the algebras $\tilde{G}^\perp_{1,2}$, $\tilde{G}^\perp_{1,3}$, and $\tilde{G}^\perp_{1,4}$, where a and b cannot be simultaneously equal to zero. The

TABLE I. Subalgebras of $G^{(0)}_{F_\|} \equiv \{J^3, K^3, P^\mu\}$.

Notation	Dimension $n_\|$	Generators	$G^{\max}_{A_\|}$	Particular invariant potentials $A_\|$
$G^{\|\pm}_{1,2}$	$n=5$	$\{J^3,K^3,P^1,P^2,P^0 \pm P^3\}$	no	—
$G^{\|}_{2,1}$	$n=5$	$\{J^3,P^\mu\}$	no	—
$G^{\|}_{3,1}$	$n=5$	$\{K^3,P^\mu\}$	no	—
$G^{\|}_{4,1}$	$n=5$	$\{J^3+aK^3,P^\mu\}$	no	—
$G^{\|}_{1,3}$	$n=4$	$\{J^3,K^3,P^0,P^3\}$	no	—
$G^{\|}_{1,4}$	$n=4$	$\{J^3,K^3,P^1,P^2\}$	no	—
$G^{\|}_{2,2}$	$n=4$	$\{J^3,P^0,P^1,P^2\}$	no	—
$G^{\|}_{2,3}$	$n=4$	$\{J^3,P^1,P^2,P^3\}$	no	—
$G^{\|\pm}_{2,4}$	$n=4$	$\{J^3,P^1,P^2,P^0 \pm P^3\}$	no	—
$G^{\|}_{3,2}$	$n=4$	$\{K^3,P^0,P^1,P^3\}$	no	—
$G^{\|\pm}_{3,3}$	$n=4$	$\{K^3,P^1,P^2,P^0 \pm P^3\}$	no	—
$G^{\|\pm}_{4,2}$	$n=4$	$\{J^3+aK^3,P^1,P^2,P^0 \pm P^3\}$	no	—
$G^{\|}_{5,1}$	$n=4$	$\{P^\mu\}$	no	—
$\tilde{G}^{\|}_{2,2}$	$n=4$	$\{J^3+aP^3,P^0,P^1,P^2\}$	no	—
$\tilde{G}^{\|}_{2,3}$	$n=4$	$\{J^3+aP^0,P^1,P^2,P^3\}$	no	—
$\tilde{G}^{\|\pm}_{2,4}$	$n=4$	$\{J^3+a(P^0 \pm P^3),P^1,P^2,P^0 \mp P^3\}$	no	—
$\tilde{G}^{\|}_{3,2}$	$n=4$	$\{K^3+aP^2,P^0,P^1,P^3\}$	no	—
$G^{\|\pm}_{1,5}$	$n=3$	$\{J^3,K^3,P^0 \pm P^3\}$	yes	$A_\| = \frac{1}{2}(\mp E(t \pm z), -By, Bx, -E(t \pm z))$
$G^{\|}_{2,5}$	$n=3$	$\{J^3,P^0,P^3\}$	no	—
$G^{\|}_{2,6}$	$n=3$	$\{J^3,P^1,P^2\}$	no	—
$G^{\|}_{3,4}$	$n=3$	$\{K^3,P^0,P^3\}$	no	—
$G^{\|}_{3,5}$	$n=3$	$\{K^3,P^1,P^2\}$	no	—
$G^{\|\pm}_{3,6}$	$n=3$	$\{K^3,P^1,P^0 \pm P^3\}$	yes	$A_\| = \frac{1}{2}(\mp E(t \pm z), -2By, 0, -E(t \pm z))$
$G^{\|}_{4,3}$	$n=3$	$\{J^3+aK^3,P^0,P^3\}$	no	—
$G^{\|}_{4,4}$	$n=3$	$\{J^3+aK^3,P^1,P^2\}$	no	—
$G^{\|}_{5,2}$	$n=3$	$\{P^0,P^1,P^2\}$	no	—
$G^{\|}_{5,3}$	$n=3$	$\{P^0,P^1,P^3\}$	no	—
$G^{\|}_{5,4}$	$n=3$	$\{P^1,P^2,P^3\}$	no	—
$G^{\|}_{5,5}$	$n=3$	$\{P^0,P^1,P^2+aP^3\}$	no	—
$G^{\|\pm}_{5,6}$	$n=3$	$\{P^1,P^2,P^0 \pm P^3\}$	no	—
$G^{\|}_{5,7}$	$n=3$	$\{P^2,P^3,P^0+aP^1\}$	no	—
$\tilde{G}^{\|}_{2,61}$	$n=3$	$\{J^3+aP^0,P^1,P^2\}$	no	—
$\tilde{G}^{\|}_{2,62}$	$n=3$	$\{J^3+aP^3,P^1,P^2\}$	no	—
$\tilde{G}^{\|}_{2,63}$	$n=3$	$\{J^3+a(P^0 \pm P^3),P^1,P^2\}$	no	—
$\tilde{G}^{\|}_{3,4}$	$n=3$	$\{K^3+aP^2,P^0,P^3\}$	no	—
$\tilde{G}^{\|\pm}_{3,6}$	$n=3$	$\{K^3+aP^2,P^1,P^0 \pm P^3\}$	no	—

first three columns of Tables I and II collect all these results. For brevity we do not include the corresponding calculations. Let us only notice that we recover all the Patera–Winternitz–Zassenhaus subalgebras of the corresponding dimensions obtained in their classification[21] (up to conjugacy under the Poincaré algebra) *and* supplementary subalgebras as expected (because of the conjugacy under the kinematical algebra).

The *second step* consists in extracting *maximal* symmetries and their associated subalgebras or subgroups $G^{\max}_A$. The method is straightforward (but tedious) through the use of our invariance conditions (2.25). There are a lot of simplifying properties which can be collected when such explicit exercises are realized but the interested reader can find them by working systematically. For brevity we only summarize the results:

α) *the* $F^{(0)}_\|$*-case*: through Table I and its specific algebras we get

a) the subalgebras of dimensions 5 and 4 do not admit invariant four-potentials;

b) among the subalgebras of dimension 3, only the four following ones are of the type $G^{\max}_{A_\|}$, i.e.,

$$G^{\pm,\max}_{1,5} \equiv \{J^3, K^3, P^0 \pm P^3\}, \quad G^{\pm,\max}_{3,6} \equiv \{K^3, P^1, P^0 \pm P^3\}. \tag{3.18}$$

So the maximal dimension is $n_\| = 3$. Explicitly, we get with respect to $G^{\pm,\max}_{1,5}$

$$A_\| = \tfrac{1}{2}\{\mp E(t \pm z), -By, Bx, -E(t \pm z)\}, \tag{3.19}$$

and with respect to $G^{\pm,\max}_{3,6}$;

$$A'_\| = \tfrac{1}{2}\{\mp E(t \pm z), -2By, 0, -E(t \pm z)\}. \tag{3.20}$$

This information is collected in the last two columns of Table

TABLE II. Subalgebras of $G_{F_\perp}^{(0)} \equiv \{\mathscr{A}^1, \mathscr{A}^2, P^\mu\}$.

Notation	Dimension $n_\perp$	Generators	$G_{A_\perp}^{\max}$	Particular invariant potentials $A_\perp$
$G_{1,2}^\perp$	$n=5$	$\{\mathscr{A}^1,\mathscr{A}^2,P^1,P^2,P^0-P^3\}$	no	—
$G_{2,1}^\perp$	$n=5$	$\{\mathscr{A}^1,P^\mu\}$	no	—
$G_{3,1}^\perp$	$n=5$	$\{\mathscr{A}^2,P^\mu\}$	no	—
$G_{4,1}^\perp$	$n=5$	$\{\mathscr{A}^1+a\mathscr{A}^2,P^\mu\}$	no	—
$\tilde{G}_{1,2}^\perp$	$n=5$	$\{\mathscr{A}^1+aP^0,\mathscr{A}^2+bP^0,P^1,P^2,P^0-P^3\}$	no	—
$G_{1,3}^\perp$	$n=4$	$\{\mathscr{A}^1,\mathscr{A}^2,P^1,P^0\text{-}P^3\}$	no	—
$G_{1,4}^\perp$	$n=4$	$\{\mathscr{A}^1,\mathscr{A}^2,P^2,P^0-P^3\}$	yes	$A_\perp^{(0)}=-\frac{1}{2}E(x,t-z,0,x)$
$G_{1,5}^\perp$	$n=4$	$\{\mathscr{A}^1,\mathscr{A}^2,P^0-P^3,P^1+aP^2\}$	no	—
$G_{2,2}^\perp$	$n=4$	$\{\mathscr{A}^1,P^0,P^2,P^3\}$	yes	$A_\perp=-E(x,0,0,x)$
$G_{2,3}^\perp$	$n=4$	$\{\mathscr{A}^1,P^1,P^2,P^0-P^3\}$	yes	$A_\perp=(0,-E(t-z),0,0)$
$G_{3,2}^\perp$	$n=4$	$\{\mathscr{A}^2,P^0,P^1,P^3\}$	no	—
$G_{3,3}^\perp$	$n=4$	$\{\mathscr{A}^2,\mathrm{P}^1,\mathrm{P}^2,\mathrm{P}^0-\mathrm{P}^3\}$	no	—
$G_{4,2}^\perp$	$n=4$	$\{\mathscr{A}^1+a\mathscr{A}^2,P^0,P^3,P^2-aP^1\}$	no	—
$G_{4,3}^\perp$	$n=4$	$\{\mathscr{A}^1+a\mathscr{A}^2,P^1,P^2,P^0-P^3\}$	no	—
$G_{5,1}^\perp$	$n=4$	$\{P^\mu\}$	no	—
$\tilde{G}_{1,3}^\perp$	$n=4$	$\{\mathscr{A}^1+aP^0,\mathscr{A}^2+bP^2,P^1,P^0-P^3\}$	no	—
$\tilde{G}_{1,4}^\perp$	$n=4$	$\{\mathscr{A}^1+aP^1,\mathscr{A}^2+bP^0,P^2,P^0-P^3\}$	yes	$A_\perp=E\left(-x+\frac{1}{2b}(t-z)^2,0,-a,-x+\frac{1}{2b}(t-z)^2\right)$
$\tilde{G}_{1,5}^\perp$	$n=4$	$\{\mathscr{A}^1+\alpha P^0,\mathscr{A}^2+\alpha a P^0+\rho P^1,P^0-P\text{-}{}^3,P^1+aP^2\}$	no	—
$\tilde{G}_{2,2}^\perp$	$n=4$	$\{\mathscr{A}^1+aP^1,P^0,P^2,P^3\}$	yes	$A_\perp=-E(x,0,a,x)$
$\tilde{G}_{2,3}^\perp$	$n=4$	$\{\mathscr{A}^1+aP^0,P^1,P^2,P^0-P^3\}$	no	—
$\tilde{G}_{3,2}^\perp$	$n=4$	$\{\mathscr{A}^2+aP^2,P^0,P^1,P^3\}$	no	—
$\tilde{G}_{3,3}^\perp$	$n=4$	$\{\mathscr{A}^2+aP^0,P^1,P^2,P^0-P^3\}$	yes	$A_\perp=-E\left(\frac{1}{2a}(t-z)^2,(t-z),0,\frac{1}{2a}(t-z)^2+a\right)$
$\tilde{G}_{4,2}^\perp$	$n=4$	$\{\mathscr{A}^1+a\mathscr{A}^2+bP^2,P^0,P^3,P^2-aP^1\}$	no	—
$\tilde{G}_{4,3}^\perp$	$n=4$	$\{\mathscr{A}^1+a\mathscr{A}^2+bP^0,P^1,P^2,P^0-P^3\}$	no	—

I:

β) *the $F_\perp^{(0)}$-case*: through Table II and its specific algebras we get

a) the subalgebras of dimension 5 do not admit invariant four-potentials;

b) there are three subalgebras of dimension 4 which are $G_{A_\perp}^{\max}$, i.e.;

$$G_{1,4}^{\max} \equiv \{\mathscr{A}^1,\mathscr{A}^2,P^2,P^0-P^3\},\quad G_{2,2}^{\max} \equiv \{\mathscr{A}^1,P^0,P^2,P^3\},$$
$$G_{2,3}^{\max} \equiv \{\mathscr{A}^1,P^1,P^2,P^0-P^3\} \tag{3.21}$$

and three infinite families among the nonsplitting subalgebras of dimension 4, i.e.,

$$\tilde{G}_{1,4}^{\max},\ \tilde{G}_{2,2}^{\max},\ \tilde{G}_{3,3}^{\max}. \tag{3.22}$$

So the maximal dimension is $n_\perp = 4$ and the corresponding invariant four-potentials can be determined. This information is collected in the two last columns of Table II. Let us only notice that $G_{1,4}^{\max}$ is, as expected, the one admitting the four-potential (3.16) corresponding to the gauge symmetrical case.

From the above results we get the *maximal* dimensions $n_\parallel = 3$ and $n_\perp = 4$ for invariant potentials. This distinction between the parallel and perpendicular cases can be simply understood; recall that there is a supplementary degree of freedom for choosing the typical perpendicular (electric and magnetic) fields (2.22) with respect to the typical parallel fields (2.20).

As a final remark in this subsection, let us notice that we restricted our results in Tables I and II to the determination of subalgebras of dimensions $n_\parallel \geqslant 3$ and $n_\perp \geqslant 4$, respectively. Our reasons are evident in connection with "maximal symmetry" as defined above. The other subalgebras ($n_\parallel = 1,2$ and $n_\perp = 1,2,3$) can evidently be obtained through the same method. All the subalgebras of our $G_{i,j}^{\max}$ evidently admit invariant four-potentials and consequently are of the type $G_A \subset G_F^{(0)}$. There are also others leading to nontrivial results but their determination has no special interest in connection with our considerations.

3c. Properties and consequences

α) *Gauge symmetrical four-potentials*: The gauge symmetrical four-potentials (3.4) and (3.16) play a more specific role in our considerations. So let us insist on the fact that in the perpendicular case, $A_\perp^{(0)} \equiv (3.16)$ leads to a maximal symmetry associated with $G_{1,4}^{\max} \equiv (3.17)$ (cf. Table II) but in the parallel case, $A_\parallel^{(0)} \equiv (3.4)$ only admits a symmetry of dimension two. These two structures have to contain the homogeneous part of the corresponding $G_F^{(0)}$ because the potentials (3.3) are always such that

$$gA^{(0)} = A^{(0)},\quad \forall g \in G_F^{(0)} \cap L_+^\uparrow \tag{3.23}$$

where $L_+^\uparrow$ refers to the (restricted) homogeneous part of the Poincaré group. In fact we have

$$(gA^{(0)})_\mu(x) = \Lambda_\mu{}^\nu A_\nu^{(0)}(\Lambda^{-1}x) = \tfrac{1}{2}\Lambda_\mu{}^\nu F_{\nu\rho}(\Lambda^{-1}x)^\rho$$
$$= \tfrac{1}{2}\Lambda_\mu{}^\nu \Lambda_\sigma{}^\rho F_{\nu\rho} x^\sigma = \tfrac{1}{2}F_{\mu\sigma}x^\sigma = A_\mu^{(0)}(x). \tag{3.24}$$

β) Realization and unchanged Poincaré generators: We already mentioned that m.e.c.s. corresponding to a constant and uniform tensor $F^{(0)}$ are associated with specific Klein–Gordon equations and realizations of generators. Now in correspondence with the four-potentials (3.19) and (3.20), and those mentioned in Table II, we can also get sets of realizations of generators. Let us apply these considerations to the four-potential (3.19) as an example. We first find that the gauge symmetrical potential (3.4) and the potential (3.19) are physically equivalent up to the gauge function

$$\lambda^* = \pm \frac{E}{4}(z^2 - t^2), \tag{3.25}$$

where the star superscript refers to these specific considerations. Then the corresponding realization is directly obtained through our Eqs. (3.14) and (3.25). We get explicitly

$$\begin{aligned} &\pi_\mu{}^* = P_\mu + eC_\mu{}^*,\ C_0{}^* = \pm \frac{E}{2}(t-z),\\ &\qquad\qquad \mathbf{C}^* = -\tfrac{1}{2}(By, -Bx, E(t \mp z)),\\ &J^{*3} = J^3,\\ &K^{*3} = K^3. \end{aligned} \tag{3.26}$$

From these relations we remark that not only the generators J^3, K^3 are unchanged with respect to the original Poincaré realization (2.6) but *also* the combinations

$$\pi_0{}^* \mp \pi_3{}^* = P^0 \pm P^3. \tag{3.27}$$

So we recover the set $\{P^0 \pm P^3, J^3, K^3\}$ associated with $G_{1,5}^{\parallel,\pm,\max}$ (cf. Table I). This is a general property; for a specific potential and from the associated realization of the $\bar{G}_F^{(0)}$ generators, the *unchanged* operators do form the subsymmetry of this potential.

γ) Symmetries of Klein–Gordon equations: According to Hoogland's developments,[7,8] we can construct Klein–Gordon equations from the realizations [of the type (3.26), for example] through the Casimir (invariant) operators of the extension $\bar{G}_F^{(0)}$. Thus the resulting equation (with interaction) admits the *Poincaré* subsymmetry of the associated potential; in fact the only unchanged Poincaré generators are those belonging to the symmetry group of the potential. Then we recover here the motivation in determining the potential with *maximal* Poincaré symmetry, this maximal character being now translated on the associated Klein–Gordon equations.

As a final remark, let us recall that if the above considerations (maximal symmetry,...) apply to Klein–Gordon equations, they also hold when *Dirac* equations (or others) have to include interaction with a constant and uniform electromagnetic field $F^{(0)}$.

4. REMARKS ON COMPENSATING GAUGES AND SYMMETRIES

When we limit ourselves to the *Poincaré* subsymmetries of potentials and wave equations (with interaction), we have seen in the last section that the symmetry explicitly depends on the specific form of the potentials associated with the field. As already mentioned in the introduction, we are also interested in the study of another kind of symmetry *independent* of specific forms of the potentials but going out the Poincaré context. After Janner and Janssen[3] and Giovannini,[6] we then study the symmetry group of a potential A (leading to an *arbitrary* electromagnetic field F) and its *compensating* gauges.[3]

If $G_F \equiv \{g\}$ is the Poincaré symmetry group of an arbitrary F, we know that gA and A lead to the same field F. We have[3]

$$gA(x) = A(x) + \partial\chi_g(x),\ \ \forall g \in G_F, \tag{4.1}$$

where χ_g is the compensating gauge function. The *symmetry group of A* is then defined as the set $\{(\phi, \chi_g, g), \phi \in R, g \in G_F\}$ with the multiplication law

$$(\phi_1, \chi_{g_1}, g_1)(\phi_2, \chi_{g_2}, g_2) = (\phi_1 + \phi_2 + f(g_1, g_2), \chi_{g_1 g_2}, g_1 g_2), \tag{4.2}$$

where the factor set $f: G_F \times G_F \to R$ is given by

$$f(g_1, g_2) = \chi_{g_1}(x) + \chi_{g_2}(g_1{}^{-1}x) - \chi_{g_1 g_2}(x). \tag{4.3}$$

Complementary information can be found elsewhere[3,6] but here we want to insist on the facts that the symmetry group of A is simply the extension $\bar{G}_F$ of R by G_F *and* that two physically equivalent potentials A and A', i.e.,

$$A'(x) = A(x) + \partial\lambda(x), \tag{4.4}$$

lead to the same factor set and consequently to the same symmetry group. Let us also mention that the compensating gauges $\chi'_g(x)$ and $\chi_g(x)$ are simply related by

$$\chi'_g(x) = \chi_g(x) + (g - 1)\lambda(x). \tag{4.5}$$

Now, having two kinds of symmetries [one inside the Poincaré group (cf. Sec. 3) and another one outside the Poincaré context (cf. $\bar{G}_F$)], let us give some interconnections among them. In this respect the specific results obtained elsewhere[6] are of special interest although we essentially refer to the case of a constant and uniform field $F^{(0)}$.

The essential point concerns the determination of the compensating gauges χ_g through the knowledge of Poincaré subsymmetries on potentials. For example, for each potential A admitting a Poincaré subsymmetry $G_A \subset G_F$, we notice that χ_g is a constant for each $g \in G_A$ and that we are left with the consideration of the elements $g \in G_F$ but $\notin G_A$. More generally, let us notice that our invariance conditions[17] on four-vectors (2.25) do measure here the defect of invariance when compensating gauges are considered. In fact when Eqs. (4.1) are required, we get

$$\begin{aligned} &\boldsymbol{\phi}\cdot\mathbf{A} - \mathscr{D}V = -\partial_t\chi_g,\\ &\boldsymbol{\theta}\Lambda\mathbf{A} - \boldsymbol{\phi}V + \mathscr{D}\mathbf{A} = -\nabla\chi_g, \end{aligned} \tag{4.6}$$

with the differential operator $\mathscr{D}$ given by (2.17). It is then very easy to determine the associated χ_g when $g \in G_F$ but $\notin G_A$.

As a final property enhancing the interest of these invariance conditions on A, let us recall that two physically equivalent potentials A and A' are related by Eq. (4.4) and their associated compensating gauges by Eq. (4.5). So it is easy to show that

$$(g - 1)\lambda(x) = \mathscr{D}\lambda(x) \tag{4.7}$$

so that

$$\chi_g'(x) = \chi_g(x) + \mathscr{D}\lambda(x). \tag{4.8}$$

Explicit examples can be considered from our Tables and permit us to see how all these notions are working. Let us only suggest as an instructive example the consideration of the potential $A_{\parallel}^{(0)} \equiv (3.4)$, the determination of its compensating gauge

$$\chi_g(x) = \tfrac{1}{2}E(a^3t - a^0z) + \tfrac{1}{2}B(a^1y - a^2x), \tag{4.9}$$

and their connections with other potentials A' *and* compensating gauges $\chi_g'(x)$.

5. GALILEAN SYMMETRIES OF POTENTIALS

If the nonrelativistic theory and the associated Schrödinger equations are taken into consideration, all the developments of our Sec. 3 can be applied; we only have to use the information contained in Sec. 2b on Galilean symmetries. For brevity we only want to mention some results in connection with the corresponding maximal symmetry of Galilean electromagnetic potentials when the magnetic limit is studied.[18]

We have seen that the kinematical group $G_{F_{\parallel}}^{(0)\prime} \equiv (2.36)$ of a constant and uniform Galilean electromagnetic field is of dimension six as in the relativistic case. The corresponding subgroup classification can also be obtained through the Patera–Winternitz–Zassenhaus method,[20] for example. Let us only notice that the results are not the same as those contained in Table I [here we have to use the algebra (2.28)]. Then with our invariance conditions (2.41) and (2.42) and the differential operator (2.43), we can select the different subalgebras leading to scalar and vector potentials V and $\mathbf{A}$ and get the interesting information on maximal symmetry. We found that the maximal dimension is $n_{\parallel}' = 2$ (remember that $n_{\parallel} = 3$ in the relativistic context) and that one of the corresponding subalgebras is associated with the gauge symmetrical potential. In particular, here we have

$$V = -\tfrac{1}{2}Ez,\ \mathbf{A} = -\tfrac{1}{2}(By, -Bx, Et) \tag{5.1}$$

as a particular case of the gauge symmetrical potential

$$V = -\tfrac{1}{2}\mathbf{E}\cdot\mathbf{x},\ \mathbf{A} = -\tfrac{1}{2}(\mathbf{E}t + \mathbf{x}\Lambda\mathbf{B}), \tag{5.2}$$

leading to the constant and uniform Galilean electromagnetic field. The Galilean subsymmetry ensuring the invariance of (5.1) is the $n_{\parallel}' = 2$-structure $\{J^3, K^3\}$ as expected (i.e., the homogeneous part of the Galilean group contained in $G_{F_{\parallel}}^{(0)\prime}$).

The above comments apply when $\mathbf{B} \neq 0$ as noted in Sec. 2b, but we can also consider the case $\mathbf{B} = 0$ leading to the eight-dimensional $G_{F}^{(0)\prime\prime} \equiv (2.37)$. Here let us only notice that this case is not of physical interest because we are working in the magnetic limit[18] where magnetic effects have to be very important with respect to the electric effects. If we wanted to consider a null magnetic field then we would have to go to the electric limit,[18] where the usual definitions (2.11) of the electric and magnetic fields in terms of the potentials V and $\mathbf{A}$ are lost.

ACKNOWLEDGMENTS

We want to express our hearty thanks to Dr. P. Winternitz for his enlightening discussion on the Patera–Winternitz–Zassenhaus algorithm in connection with Sec. 3 of this paper. We also want to thank M.-P. Biémont-Gérard for her careful typing of the articles I and II.

[1]H. Bacry, Ph. Combe, and J. L. Richard, Nuovo Cimento A **67**, 267 (1970).
[2]H. Bacry, Ph. Combe, and J. L. Richard, Nuovo Cimento A **70**, 289 (1970).
[3]A. Janner and T. Janssen, Physica **53**, 1 (1971).
[4]R. Schrader, Forthschr. Phys. **20**, 701 (1972).
[5]Ph. Combe and J. L. Richard, Physica **71**, 66 (1974).
[6]N. Giovannini, Physica A **87**, 546 (1977).
[7]H. Hoogland, "On local and global equivalence of wave functions in quantum mechanics," Ph.D Thesis, Nijmegen University (1977).
[8]H. Hoogland, J. Phys. A **11**, 797 (1978).
[9]V. Hussin, Nuovo Cimento A **65**, 39 (1981).
[10]V. Hussin, Lett. Nuovo Cimento **32**, 181 (1981).
[11]J. M. Levy-Leblond, *Group Theory and Its Applications*, Vol. 2, edited by E. M. Loebl (Academic, New York, 1971), p. 221.
[12]E. P. Wigner, Ann. Math. **40**, 149 (1939).
[13]V. Bargmann, Ann. Math. **59**, 1 (1954).
[14]M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Massachusetts, 1962), Sec. 12.7.
[15]J. M. Levy-Leblond, J. Math. Phys. **4**, 776 (1963).
[16]J. Beckers, J. Harnad, M. Perroud, and P. Winternitz, J. Math. Phys. **19**, 2126 (1978).
[17]J. Beckers and V. Hussin, Found. Phys. **11**, 881 (1981).
[18]J. M. Levy-Leblond and M. Le Bellac, Nuovo Cimento B **14**, 217 (1973).
[19]L. Landau and E. Lifshitz, *Théorie des Champs* (Mir, Moscow, 1970).
[20]J. Patera, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **16**, 1597 (1975).
[21]J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **17**, 977 (1976).
[22]Ph. Combe and P. Sorba, Physica A **80**, 271 (1975).

Minimal electromagnetic coupling schemes. II. Relativistic and nonrelativistic Maxwell groups

J. Beckers and V. Hussin[a)]

Université de Liège, Physique théorique et mathématique, Institut de Physique au Sart Tilman, Bâtiment B.5, B-4000 Liège 1, Belgium

(Received 20 January 1982; accepted for publication 19 March 1982)

Minimal electromagnetic coupling schemes entering into Klein–Gordon or Schrödinger equations are studied in connection with symmetries *outside* the symmetry groups of the corresponding *free* equations. The Schrader construction of the so-called (relativistic) Maxwell group is reviewed through group extensions of kinematical groups associated with (constant and uniform) electromagnetic fields. The construction of the Galilean (nonrelativistic) Maxwell group is given.

PACS numbers: 11.10.Qr, 11.30.Cp, 02.20. + b

1. INTRODUCTION

In a previous paper[1] (hereafter denoted I) we analyzed minimal electromagnetic coupling schemes (m.e.c.s.) entering into wave equations such as Klein–Gordon (or Dirac) ones in the relativistic context or such as Schrödinger equations in the nonrelativistic context. In order to get symmetries of the associated equations, we distinguished between "*inside or outside symmetries*" with respect to the symmetry group of the *free* particle descriptions.

In this paper we discuss more particularly "outside symmetries" in the relativistic as well as nonrelativistic cases in connection with group extensions[2,3,4] of the symmetry groups of the *free* equations, i.e., the Poincaré[5] and Galilei[6] groups, respectively. In fact, our study will be directly related to the Schrader considerations[7] and more particularly to the construction of the so-called Maxwell group in the relativistic theory. The Maxwell group is the symmetry group of Klein–Gordon (or Dirac) equations when the corresponding particles *interact* with the classical homogeneous (constant and uniform) electromagnetic field F. By introducing an explicit F-dependence in the wavefunction, Schrader discussed equations corresponding to m.e.c.s. through essentially the replacement of momentum operators by covariant derivatives. Such a point of view is directly connected with our discussion in I and it can be extended to the nonrelativistic context.

With the generalities (notations, transformation laws, algebras,...) on the Poincaré and Galilei groups quoted in I, Sec. 2, we first want to show (Sec. 2) some connections between Schrader's results[7] and more recent contributions due to Combe and Richard,[8] Hoogland,[9] and Hussin[10] when we deal with Poincaré group and Klein–Gordon equations including interaction with an external (constant and uniform) electromagnetic field. In particular, the extensions of some kinematical groups[11] associated with specific fields will play interesting roles. Then, in Sec.3, we construct the *nonrelativistic Maxwell group* from the study of Schrödinger equations describing a charged particle in a constant and uniform *Galilean* electromagnetic field (as described in Sec. 2 of I) when the magnetic limit[12] is under consideration. This magnetic limit has, with respect to the electric limit,[12] the advantage of keeping the electromagnetic field defined, as in the relativistic case, by the usual relations in terms of the scalar and vector potentials V and $\mathbf{A}$ respectively.

Let us notice that the real motivation for constructing such Maxwell groups in both the relativistic and nonrelativistic cases is the search for *true* irreducible unitary representations of the symmetry groups associated with wave equations describing interactions.

Our notations, conventions, and units are those given in I.

2. RELATIVISTIC MAXWELL GROUP AND KLEIN–GORDON EQUATIONS

After Schrader,[7] we can study the symmetry group of Klein-Gordon equation rewritten in terms of the covariant derivatives

$$(D_\mu D^\mu + m^2)\psi(x,F) = 0 \tag{2.1}$$

where

$$D_\mu = \partial_\mu - ieA_\mu^{(0)}(x,F). \tag{2.2}$$

The four-potential $A^{(0)}(x,F)$ is associated with a *constant and uniform* electromagnetic tensor $F^{(0)}$ (the naught superscript will refer to this specific case as in I, Subsec. 2a). In the so-called symmetric gauge,[1,11] we have

$$A^{(0)}(x,F) = \tfrac{1}{2}F_{\mu\nu}x^\nu \tag{2.3a}$$

or, explicitly,

$$V^{(0)}(x,F) = -\tfrac{1}{2}\mathbf{E}\cdot\mathbf{x}, \quad \mathbf{A}^{(0)}(x,F) = -\tfrac{1}{2}(\mathbf{E}t + \mathbf{x}\wedge\mathbf{B}). \tag{2.3b}$$

Under restricted homogeneous Lorentz transformations Λ, Schrader showed the invariance of Eq. (2.1) when

$$A_\mu(\Lambda x,\Lambda F) = \Lambda_\mu{}^\rho A_\rho(x,F), \qquad \Lambda F \rightarrow (\Lambda F)_{\mu\nu} \equiv \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma F_{\rho\sigma}, \tag{2.4}$$

and

$$\psi'(\Lambda x,\Lambda F) = (U(\Lambda)\psi)(x',F') = \psi(x,F). \tag{2.5}$$

He obtained

$$U(\Lambda)U(\Lambda') = U(\Lambda\Lambda') = U(\Lambda''), \tag{2.6}$$

[a)] Boursier I.R.S.I.A. 1981–1982

 0022-2488/83/051295-04$02.50

enhancing the *true* character of the associated representations.

Under space-time translations, the invariance of Eq. (2.1) requires

$$(U(a)\psi)(x',F') = \exp\left(\frac{ie}{2}x^\mu F_{\mu\nu}a^\nu\right)\psi(x'-a,F), \qquad (2.7)$$

leading to

$$U(a)U(a') = \exp\left(-\frac{ie}{2}a^\mu F_{\mu\nu}a'^\nu\right)U(a+a'), \qquad (2.8)$$

i.e., to *projective* representations dealing with a structure other than the usual Poincaré group. Finally, Schrader defined new *phase transformations* on the wave function

$$(U(\alpha)\psi)(x,F) = \exp(-i\alpha_{\mu\nu}F^{\mu\nu})\psi(x,F) \qquad (2.9)$$

in order to get the so-called Maxwell group $\mathscr{M}$ with only *true* representations. It is the set $\mathscr{M} \equiv \{(a,\alpha,\Lambda)\}$ with multiplication law

$$(a,\alpha,\Lambda)(a',\alpha',\Lambda') = \left(a+\Lambda a', \alpha+\Lambda\alpha' + \tfrac{1}{2}e\left(a \underset{\Lambda}{\oplus} \Lambda a'\right), \Lambda\Lambda'\right) \qquad (2.10)$$

where

$$(\Lambda\alpha')_{\mu\nu} = \Lambda_\mu{}^\rho\Lambda_\nu{}^\sigma\alpha'_{\rho\sigma}, \qquad (2.11)$$

$$\left(a \underset{\Lambda}{\oplus} a'\right)_{\mu\nu} = \tfrac{1}{2}(a_\mu a'_\nu - a'_\mu a_\nu) \qquad (2.12)$$

and $\alpha \in \mathscr{A}$, the set of real skewsymmetric 4×4 matrices. These elements lead to

$$U(\alpha)U(\alpha') = \exp\left(\frac{ie}{2}a^\mu F_{\mu\nu}a'^\nu\right)U(\alpha'') \qquad (2.13)$$

such that

$$U(a,\alpha,\Lambda)U(a',\alpha',\Lambda') = U(a'',\alpha'',\Lambda''), \qquad (2.14)$$

as expected.

Now let us come back to Eqs. (2.6) and (2.8), and the corresponding projective representations. Such a step was not considered by Schrader but it appears useful in connection with properties of exponents and extensions of kinematical groups $G_F^{(0)}$ (associated with specific electromagnetic fields F introduced in I Sec. 2a). In fact, by combining (2.5) and (2.7), we get

$$(U(a,\Lambda)\psi)(\Lambda x + a, \Lambda F) = \exp\left(\frac{ie}{2}x_\mu F^{\mu\nu}\widetilde{\Lambda}_\nu{}^\rho a_\rho\right)\psi(x,F), \qquad (2.15)$$

leading to

$$U(a,\Lambda)U(a',\Lambda') = \exp\left[\frac{ie}{2}(\widetilde{\Lambda}'a')_\sigma F^{\sigma\alpha}(\widetilde{\Lambda}''a)_\alpha\right]U(a'',\Lambda''). \qquad (2.16)$$

It is well known [8,10] that an *exponent* of $G_F^{(0)}$ (2.18) of I is

$$\xi^{(0)}(g,g') = \frac{e}{2}(\Lambda a')^\mu F_{\mu\nu}a^\nu \qquad (2.17)$$

(see, for example, Eq. (3.2) of Ref. 10 when $q \equiv e$, $q^* = d = 0$), when $g,g' \in G_F^{(0)}$. In this context, we can rewrite Eq. (2.17) in the following form:

$$\xi^{(0)}(g,g') = \frac{e}{2}(\widetilde{\Lambda}'a')_\sigma F^{\sigma\alpha}(\widetilde{\Lambda}''a)_\alpha, \qquad (2.18)$$

so that Eq. (2.16) becomes

$$U(g)U(g') = \exp(i\xi^{(0)}(g,g'))U(g''), \quad g,g',g'' \in G_F^{(0)}. \qquad (2.19)$$

Furthermore, through the same considerations, Eq. (2.15) can be written

$$(U(g)\psi)(gx,gF) = \exp(i\xi^{(0)}(g,h_x))\psi(x,F), \forall g \in G_F^{(0)}, \qquad (2.20)$$

where h_x is an element of $G_F^{(0)}$ defined by $a = x$ and $\Lambda = \mathbb{1}$.

At this stage we can thus connect the Schrader method and the recent developments presented by Hoogland[9] and Hussin[10] when the *construction* of wave equations with minimal electromagnetic couplings is considered.

With the so-obtained Eqs. (2.19) and (2.20) we clearly see the specific role of the kinematical group of F, which admits an extension and, consequently, a nontrivial exponent.

Now let us come back to Eq. (2.14) and the associated Maxwell algebra in order to show another interesting feature in connection with the Schrader results and the Hoogland–Hussin method, i.e., that the Lie algebra of the extension $\overline{G}_F^{(0)}$ of $G_F^{(0)}$ is nothing else than a subalgebra of the Maxwell algebra.

The Maxwell algebra is generated by sixteen operators denoted by π^μ, $\{M^{\mu\nu}\} \equiv (\mathbf{J},\mathbf{K})$, $\{F^{\mu\nu}\} \equiv (\mathbf{E},\mathbf{B})$ associated with the ten usual Poincaré parameters and with six parameters corresponding to the real skewsymmetric matrices α. From Eq. (2.10) the Maxwell Lie algebra can be written explicitly in the following form (all the other commutators are equal to zero):

$$[J^j,J^k] = i\epsilon^{jk}{}_l J^l, [J^j,K^k] = i\epsilon^{jk}{}_l K^l, \\ [K^j,K^k] = -i\epsilon^{jk}{}_l J^l, \qquad (2.21)$$

$$[J^j,E^k] = i\epsilon^{jk}{}_l E^l, [J^j,B^k] = i\epsilon^{jk}{}_l B^l, \\ [K^j,E^k] = -i\epsilon^{jk}{}_l B^l, [K^j,B^k] = -i\epsilon^{jk}{}_l E^l, \qquad (2.22)$$

$$[J^j,\pi^k] = i\epsilon^{jk}{}_l\pi^l, [J^i,\pi^0] = 0, [K^j,\pi^k] = i\delta^{jk}\pi^0, \\ [K^j,\pi^0] = i\pi^j, \qquad (2.23)$$

$$[\pi^0,\pi^j] = ieE^j, [\pi^j,\pi^k] = -ie\epsilon^{jk}{}_l B^l. \qquad (2.24)$$

Now let us take particular cases for the six generators $\mathbf{E}$, $\mathbf{B}$. For example, if we choose $\mathbf{E} \equiv (0,0,E)$ and $\mathbf{B} \equiv (0,0,B)$, all the commutators (2.22) vanish and the relations (2.24) reduce to the only two nontrivial commutators

$$[\pi^0,\pi^3] = ieE, \quad [\pi^1,\pi^2] = ieB. \qquad (2.25)$$

So, without further information, we get a subalgebra characterized by Eqs. (2.21), (2.23), and (2.25). We also notice that the generators E and B do commute with all the π^μ, $\mathbf{J}$, and $\mathbf{K}$ operators and commute among each other so that they belong to the center of *this* subalgebra. Now if we limit the homogeneous part (2.21) to J^3 and K^3 we get a new subalgebra of the Maxwell algebra which is the algebra of the extension of R by $G_{F_\parallel}^{(0)} \equiv$(2.21) of I when E and B become numbers (multiplying the generator I of R). Such a remark completes our discussion based on Eqs. (2.19) and (2.20).

Analogous developments can evidently be realized when $\mathbf{E} \equiv (E,0,0)$, $\mathbf{B} \equiv (0,E,0)$. Equation (2.24) give

$$[\pi^0,\pi^1] = ieE, \quad [\pi^3,\pi^1] = ieE \tag{2.26}$$

and we can get the algebra associated with the extension of R by $G_{F_1}^{(0)} \equiv$(I.2.23) as a subalgebra of the Maxwell one.

As the last point of this section, let us notice that the Maxwell group constructed by Schrader is the extension of R^6 by the Poincaré group as already noted by Giovannini.[13]

3. NONRELATIVISTIC MAXWELL GROUP AND SCHRÖDINGER EQUATIONS

Let us use a compact notation according to the Galilean transformations (2.26) of I. We define

$$x \to x' = gx = (a,L)x = Lx + a,$$

where

$$L = \begin{pmatrix} 1 & 0 \\ \mathbf{v} & R \end{pmatrix}, \quad a = (b,\mathbf{a}). \tag{3.1}$$

It is well known[4] that the *free* Schrödinger equation

$$i\partial_t \psi(x) + \frac{1}{2m} \nabla \cdot \nabla \psi(x) = 0 \tag{3.2}$$

is invariant under the Galilei group *if* we have

$$(U(L)\psi)(Lx) = \exp(if(x))\psi(x), \tag{3.3}$$

and

$$(U(a)\psi)(x+a) = \psi(x), \tag{3.4}$$

where[6]

$$f(x) = m(\tfrac{1}{2}v^2 t + \mathbf{v} \cdot R\mathbf{x}). \tag{3.5}$$

We then obtain

$$U(L)U(L') = U(LL'), \tag{3.6}$$

but

$$U(a,L)U(a',L') = \exp(i\xi(g,g'))U(a'',L''). \tag{3.7}$$

The group exponent is given by[6]

$$\xi(g,g') = m(\tfrac{1}{2}v^2 b' + \mathbf{v} \cdot R\mathbf{a}') \tag{3.8}$$

so that we observe that the representations (3.7) are *projective*. If we are interested in only *true* representations, we know[4] that we have to consider *the extension of R by the Galilei group*. By introducing

$$(U(\theta)\psi)(x) = [\exp(i\theta)]\psi(x), \quad \theta \in R, \tag{3.9}$$

we get a new group—the set of elements (θ,g)—characterized by the multiplication law

$$(\theta,g)(\theta',g') = (\theta + \theta' + \xi(g,g'),gg') = (\theta'',g''), \tag{3.10}$$

where $\xi(g,g')$ is given by (3.8). The corresponding representations

$$U(\theta,a,L) = U(\theta)U(a,L) \tag{3.11}$$

are *true* representations

$$U(\theta,a,L)U(\theta',a',L') = U(\theta'',a'',L''), \tag{3.12}$$

and the symmetry group of the *free* Schrödinger equation becomes the extension of R by the Galilei group as already mentioned.

Now if we consider the Schrödinger equation for a particle in a constant and uniform *Galilean* electromagnetic field $(\mathbf{E},\mathbf{B})$, we can search for the corresponding Maxwell group with only true representations. This is an application of Schrader's development[7] but in the nonrelativistic context.

Let us consider the following Schrödinger equation:

$$iD_t\psi(x,F) + \frac{1}{2m}\mathbf{D}\cdot\mathbf{D}\psi(x,F) = 0, \tag{3.13}$$

where

$$D_t = \partial_t - ieV(x,F), \quad \mathbf{D} = \nabla + ie\mathbf{A}(x,F) \tag{3.14}$$

and

$$V(x,F) = -\tfrac{1}{2}\mathbf{E}\cdot\mathbf{x}, \quad \mathbf{A}(x,F) = -\tfrac{1}{2}(\mathbf{E}t + \mathbf{x}\Lambda\,\mathbf{B}). \tag{3.15}$$

These potentials are exactly the "same" as those issued from the symmetric gauge in the relativistic case. The homogeneous Galilean transformations do not affect Eq. (3.13) if once again we have Eqs. (3.3) and (3.5) but, under translation, we get

$$\begin{aligned}(U(a)\psi)(x',F') &= (U(a)\psi)(x',F) \\ &= (U(a)\psi)(t+b,\mathbf{x}+\mathbf{a},\mathbf{E},\mathbf{B}) \\ &= [\exp i\beta(x,\mathbf{E},\mathbf{B})]\psi(x,\mathbf{E},\mathbf{B}),\end{aligned} \tag{3.16}$$

where

$$\begin{aligned}\beta &= e[V(a)t - \mathbf{A}(a)\cdot\mathbf{x}] \\ &= -\tfrac{1}{2}e[\mathbf{E}\cdot(t\mathbf{a} - b\mathbf{x}) + (\mathbf{x}\Lambda\mathbf{B})\cdot\mathbf{a}].\end{aligned} \tag{3.17}$$

Then we obtain

$$U(a)U(a') = \exp[i\phi(a,a')]U(a''), \tag{3.18}$$

with the exponent

$$\begin{aligned}\phi(a,a') &= -e[V(a)b' - \mathbf{A}(a)\cdot\mathbf{a}'] \\ &= \tfrac{1}{2}e[\mathbf{E}\cdot(\mathbf{a}b' - \mathbf{a}'b) + \mathbf{B}\cdot(\mathbf{a}\Lambda\mathbf{a}')].\end{aligned} \tag{3.19}$$

The relation (3.19) has to be compared with Eq. (2.8) in the relativistic case.

In order to get *true* representations, let us introduce new phase transformations characterized by six real parameters $r \equiv (\mathbf{j},\mathbf{k})$,

$$(U(r)\psi)(x,\mathbf{E},\mathbf{B}) = \exp[-i(\mathbf{j}\cdot\mathbf{E} + \mathbf{k}\cdot\mathbf{B})]\psi(x,\mathbf{E},\mathbf{B}), \tag{3.20}$$

such that

$$U(r)U(r') = \exp[-i\phi(a,a')]U(r''). \tag{3.21}$$

This corresponds to a multiplication law defined by

$$\begin{aligned}\mathbf{j}'' &= \mathbf{j} + \mathbf{j}' - \tfrac{1}{2}e(\mathbf{a}b' - \mathbf{a}'b), \\ \mathbf{k}'' &= \mathbf{k} + \mathbf{k}' - \tfrac{1}{2}e(\mathbf{a}\Lambda\mathbf{a}').\end{aligned} \tag{3.22}$$

So we finally obtain true representations

$$\begin{aligned}&U(\theta;b,\mathbf{a};\mathbf{j},\mathbf{k};\mathbf{v},R)U(\theta';b',\mathbf{a}';\mathbf{j}',\mathbf{k}';\mathbf{v}',R') \\ &\quad = U(\theta'';b'',\mathbf{a}'';\mathbf{j}'',\mathbf{k}'';\mathbf{v}'',R'')\end{aligned} \tag{3.23}$$

and the symmetry group of elements $(\theta;b,\mathbf{a};\mathbf{j},\mathbf{k};\mathbf{v},R) \equiv (\theta;a;r;L)$ is the (nonrelativistic) *Galilean Maxwell group* with the multiplication law

$$(\theta;a;r;L)(\theta';a';r';L') = (\theta + \theta' + \xi(g,g');a + La';r'';LL'), \tag{3.24}$$

where

$$\begin{aligned}&\xi(g,g') \equiv (3.8), \quad r'' = (\mathbf{j}'',\mathbf{k}''), \\ &\mathbf{j}'' = \mathbf{j} + R\mathbf{j}' - \mathbf{v}\Lambda R\,\mathbf{k}' - \tfrac{1}{2}e[\mathbf{a}b' - (\mathbf{v}b' + R\mathbf{a}')b], \\ &\mathbf{k}'' = \mathbf{k} + R\,\mathbf{k}' - \tfrac{1}{2}e[\mathbf{a}\Lambda(\mathbf{v}b' + R\mathbf{a}')].\end{aligned} \tag{3.25}$$

The relation (3.22) is a particular case of (3.25) when $\mathbf{v} = 0$ and $R = \mathbb{1}$.

Now let us come back to the representations corresponding to (3.3) and (3.16). We can see that the representation $\{U(a,L)\}$ is, in this case, characterized by

$$(U(a,L)\psi)(gx,g\mathbf{E},g\mathbf{B}) = \exp\{ie[f(x) + \gamma(x,\mathbf{E},\mathbf{B})]\}\psi(x,\mathbf{E},\mathbf{B}), \tag{3.26}$$

where

$$f(\mathbf{x},t) \equiv (3.5),$$

and

$$\begin{aligned}\gamma(x,\mathbf{E},\mathbf{B}) &= tV(L^{-1}a) - \mathbf{x}\cdot\mathbf{A}(L^{-1}a) \\ &= -\tfrac{1}{2}\{\mathbf{E}\cdot[t\widetilde{R}(\mathbf{a} - b\mathbf{v}) - b\mathbf{x}] + (\mathbf{x}\Lambda\,\mathbf{B})\cdot\widetilde{R}(\mathbf{a} - b\mathbf{v})\}.\end{aligned} \tag{3.27}$$

Then we have

$$\begin{aligned}U(a,L)U(a',L') &= \exp\{i[\xi(g,g') - \phi(g,g',\mathbf{E},\mathbf{B})]\}U(a'',L''),\end{aligned} \tag{3.28}$$

the exponent $\xi(g,g')$ being given by (3.8) and $\phi(g,g',\mathbf{E},\mathbf{B})$ by

$$\begin{aligned}\phi(g,g',\mathbf{E},\mathbf{B}) = \tfrac{1}{2}e\{&\mathbf{E}\cdot[\widetilde{R}''(\mathbf{a} - b\mathbf{v}'')b' - \widetilde{R}'(\mathbf{a}' - b'\mathbf{v}')b] \\ &+ \mathbf{B}\cdot[\widetilde{R}''(\mathbf{a} - b\mathbf{v}'')\Lambda\widetilde{R}'(\mathbf{a}' - b'\mathbf{v}')]\}.\end{aligned} \tag{3.29}$$

So we do need the introduction of the parameters $\theta, r \equiv (\mathbf{j},\mathbf{k})$ and their associated representations in order to get true representations as noted before. These are the characteristics which led us *outside* the symmetry group.

The Lie algebra of the (nonrelativistic) Galilean Maxwell group is generated by the *seventeen* operators $I(\theta)$, $\pi_0(b)$, $\boldsymbol{\pi}(\mathbf{a})$, $\mathbf{E}(\mathbf{j})$, $\mathbf{B}(\mathbf{k})$, $\mathbf{J}(R)$, and $\mathbf{K}(\mathbf{v})$. It has the following structure:

$$\begin{aligned}&[J_j,J_k] = i\epsilon_{jkl}J_l,\ [J_j,K_k] = i\epsilon_{jkl}K_l,\ [K_i,K_j] = 0,\\ &[J_j,E_k] = i\epsilon_{jkl}E_l,\ [J_j,B_k] = i\epsilon_{jkl}B_l,\\ &[K_j,E_k] = 0,\ [K_j,B_k] = -i\epsilon_{jkl}E_l,\\ &[J_j,\pi_0] = 0,\ [J_j,\pi_k] = i\epsilon_{jkl}\pi_l,\\ &[K_j,\pi_0] = i\pi_j,\ [K_j,\pi_k] = -im\delta_{jk}I,\\ &[E_j,\pi_0] = 0,\ [E_j,\pi_k] = 0,\ [B_j,\pi_0] = 0,\ [B_j,\pi_k] = 0,\\ &[\pi_0,\pi_j] = ieE_j,\ [\pi_j,\pi_k] = -ie\epsilon_{jkl}B_l.\end{aligned} \tag{3.30}$$

Finally, let us notice that all the considerations parallel to those of the relativistic case on Maxwell subalgebras (cf. Sec. 2) can also be discussed in this context, i.e., for example, those related to the corresponding nonrelativistic kinematical subalgebras (2.36) of I and (2.37) of I and their extensions. For brevity, we do not give these elements explicitly here.

[1]J. Beckers and V. Hussin, "Minimal electromagnetic coupling schemes I: Symmetries of potentials and gauge transformations," J. Math. Phys. **24**, 1286 (1983).

[2]V. Bargmann, Ann. Math. **59**, 1 (1954).

[3]M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1962), Sec. 12.7.

[4]J. M. Levy-Leblond, J. Math. Phys. **4**, 776 (1963).

[5]E. P. Wigner, Ann. Math. **40**, 149 (1939).

[6]J. M. Levy-Leblond, *Group Theory and its Applications*, Vol. 2, edited by E. M. Loebl (Academic, New York, 1971), p. 221.

[7]R. Schrader, Fortschr. Phys. **20**, 701 (1972).

[8]Ph. Combe and J. L. Richard, Physica **71**, 66 (1974).

[9]H. Hoogland, "On local and global equivalence of wave functions in quantum mechanics," Ph.D. thesis, Nijmegen University (1977); J. Phys. A **11**, 797 (1978).

[10]V. Hussin, Nuovo Cimento A **65**, 39 (1981); Lett. Nuovo Cimento **32**, 181 (1981).

[11]H. Bacry, Ph. Combe, and J. L. Richard, Nuovo Cimento A **67**, 267 (1970).

[12]J. M. Levy-Leblond and M. Le Bellac, Nuovo Cimento B **14**, 217 (1973).

[13]N. Giovannini, Helv. Phys. Acta **50**, 337 (1977).

Generalized Hooke groups and the mass-spectrum problem

Daniel Beau
Physique—Mathématique, Faculté des Sciences Mirande, 21004—Dijon, Cedex, France

(Received 18 September 1981; accepted for publication 4 April 1982)

We briefly recall the notion of internal structure and the relativistic covariant method [introduced in Beau and Horchani, J. Math. Phys. **20**, 1700 (1979)] of unifying external and internal structures leading to a kinematical Lie algebra. In this framework we propose a concept of dynamical development of the physical systems defined by this Lie algebra. So we obtain some Lie algebras, a generator of which (the Hamiltonian) gives rise to various mass formulas capable of describing the hadron spectrum; we make use of both unitary irreducible global representations and partially integrable, Schur-irreducible, symmetric local representations.

PACS numbers: 12.70. + q, 11.30.Ly, 02.20.Sv

I. INTRODUCTION

A few years ago we proposed,[1] along the conventional lines of the Lie algebras, a relativistic covariant internal formalism and unification method leading to various exact hadrons mass formulas, both within the framework of irreducible unitary global representations and within the framework of not necessarily integrable, Schur-irreducible, symmetric local representations.[2] Contrary to the usual procedure, the attitude we have taken is of reconsidering the concept of "internal structure" rather than the symmetries of this structure and of giving the "mass observable" priority over the "internal observables" since, in our opinion, the essential role of the internal symmetries, at least in a first aproximation, is to reduce the number of degrees of internal freedom and to raise the possible degeneracy of energy, which function can be accomplished by means of an analytical or perturbative method, for instance. Moreover, there is nothing to prevent the mass from depending on other parameters not conserved in the interactions observed at present, i.e., on supplementary quantum numbers that do not appear in the present characterization of the particles, as the quantum numbers have so far been considered chiefly within the framework of the laws of conservation.

In this composite-particles model the internal structure is described by the Heisenberg algebra $\mathfrak{h}_n$ and the unification of the internal and external structures by a Lie subalgebra $\mathfrak{G}$, containing the Poincaré algebra $\mathfrak{p}$, of $\mathscr{U}(\mathfrak{h}_n) \otimes \mathscr{U}(\mathfrak{p})$ ($\mathscr{U}$ denotes the enveloping algebra). We also adopt the following hypothesis:

The generators of $\mathfrak{G}$, other than those of the Poincaré algebra, are relativistic covariant; more precisely, they commute with the translations of $\mathfrak{p}$ and constitute a basis of a real finite-dimensional representation (direct sum of irreductible representations) of $\mathfrak{sl}(2, \mathbb{C})$.

The hypothesis that the (relativistic) internal structure is described by the Lie algebra of the (nonrelativistic) commutation relations might, at first sight, seem ambiguous. Now it often happens that nonrelativistic concepts are used in a relativistic context as is the case, for instance, of the parton model in the infinite momentum frame (cf., for example, Ref. 3). Moreover, to our knowledge, no experiment mentions any observable difference between the relativistic and nonrelativistic internal structures; and these hypotheses, when the Poincaré group would be replaced by the Galilée group could, in our opinion, be considered in the frame of nonrelativistic approximation. The hadrons are interpreted as the excited levels of these composite particles, which interpretation shows how energy (mass) is created on the basis of the internal dynamics thus defined.

The internal degrees of freedom result therefore, in our model, from the Heisenberg Lie algebra $\mathfrak{h}_n$ and its enveloping algebra. Several reasons suggest that this choice is natural and reasonable.

Thus the success of the quark model, the importance of the harmonic oscillator in nuclear physics, and the fact that the group SU (3) is a symmetry group of the latter have led physicists to look for a (nonrelativistic) hadron model based on the harmonic oscillator.[4] Various relativistic generalizations have been made since then.[5]

But one of the difficulties of these approaches is that the very concept of the relativistic (nonquantized) harmonic oscillator is not yet perfectly defined and one of their drawbacks is that the relativistic covariance is not always satisfied.

The most important (formal) contribution in this respect is that of Feynman *et al.*, which is based on Greenberg's nonrelativistic harmonic oscillator symmetric quark model. Nevertheless, several criticisms can be levelled at their analysis, the most important of which (cf. also Ref. 6) bear on the ambiguity of definition and the spectrality of the operators on one hand, and on the fact that they disregard the timelike excitations after having postulated at the outset a relativistic treatment.

Contrary to these models our approach[1] has allowed us a rigorous relativistic treatment of the mass-spectrum problem.

Now in the present paper we extend the concepts introduced in Ref. 1 and assume a dynamical evolution principle which associates some dynamical algebras $\tilde{\mathfrak{G}}_n$ with each kinematical algebra $\mathfrak{G}_n$. So the dynamics will be described by one generator of $\tilde{\mathfrak{G}}_n$ (the Hamiltonian) to which we connect a mass-splitting operator.

An analogous dynamical principle was introduced by Roman *et al.*,[5,7] but we do not follow their algebraic approach[5] used to obtain the mass spectrum.

Finally, we must point out that the presence of internal variables relates our model to the multilocal field approach

 0022-2488/83/051299-12$02.50

(notion introduced by Yukawa and taken up again since by various authors, cf., for instance, Ref. 8) and to the new Dirac equation.[9]

Before working out our model, let us give a short survey of this paper.

In Sec. 2, we consider the unification (Ref. 1, Sec. 6) of the Lie subalgebra $\mathfrak{G}_3$ of $\mathscr{U}(\mathfrak{h}_3)\otimes\mathscr{U}(\mathfrak{p})$ with the minimal isospin Lie algebra $\mathfrak{su}(2)$, so the isospin Lie algebra works on the internal degrees of freedom described by $\mathfrak{h}_3$. We obtain then a Lie subalgebra $\mathfrak{G}$ of $\mathscr{U}(\mathfrak{h}_3 \dotplus \mathfrak{su}(2))\otimes\mathscr{U}(\mathfrak{p})$, the semidirect sum $\mathfrak{h}_3\dotplus\mathfrak{su}(2)$ being defined by the representation $D(1)\oplus D(1)\oplus D(0)$ of $\mathfrak{su}(2)$ [$D(j)$ denotes the irreducible representation of weight j of $\mathfrak{su}(2)$]. Then we briefly study this Lie algebra $\mathfrak{G}$, the connected and simply connected corresponding group G, and some useful unitary irreducible representations (U.I.R.) of G.

Section 3 deals with the cohomology groups $H^n(\mathfrak{G},\mathfrak{G})$ $(n=0,1,2)$, for $H^1(\mathfrak{G},\mathfrak{G})$ allows us to define without ambiguity our dynamical Lie algebras $\tilde{\mathfrak{G}}$ and $H^2(\mathfrak{G},\mathfrak{G})$ gives us some important information about the deformations of $\mathfrak{G}$.

In Sec. 4 we introduced our dynamical principle and the first inferences are drawn from it. We determine then the various dynamical Lie algebras $\tilde{\mathfrak{G}}$ associated with the kinematical Lie algebra $\mathfrak{G}$ and the connected and simply connected dynamical Lie groups $\tilde{G}$ corresponding to each $\tilde{\mathfrak{G}}$.

In Sec. 5 we extend the U.I.R. (obtained in Sec. 2) of the kinematical group G to two of the preceding dynamical Lie groups. The corresponding representations of the associated Lie algebras are given.

Section 6 is entirely devoted to the physical discussion of our model. One hadron mass formula is obtained in each of the two preceding examples. We make use of both unitary irreducible global representations and partially integrable, Schur-irreducible, symmetric local representations.

2. GENERALITIES ON THE GROUP G

A. Definition and structural properties

Let $(M_{\mu\nu},P_\sigma)$ be the canonical basis of the Poincaré algebra $\mathfrak{p}=\mathbb{R}^4\dotplus\mathfrak{sl}(2,\mathbb{C})$ (where $\dotplus$ denotes the semidirect sum of Lie algebras, the semidirect product of the groups will be denoted by a point; the Greek indices vary from 1 to 4) and $(J_{ij};p_i,q_j,I$, with $[p_i,q_j]=\delta_{ij}I)$ the basis of the Lie algebra $\mathfrak{h}_3\dotplus\mathfrak{su}(2)$ (the Latin indices vary from 1 to 3), the semidirect sum being defined by the representation $D(1)\oplus D(1)\oplus D(0)$ of the Lie algebra $\mathfrak{su}(2)$.

Then $\mathfrak{G}$ will be the subalgebra of $\mathscr{U}(\mathfrak{h}_3\dotplus\mathfrak{su}(2))\otimes\mathscr{U}(\mathfrak{p})$ generated by

$$L_{\mu\nu}=I\otimes M_{\mu\nu},\quad T_\mu=I\otimes P_\mu,\quad C_{\mu\nu}=I\otimes P_\mu P_\nu,$$
$$I_{ij}=J_{ij}\otimes 1,\quad Q_{i\mu}=q_i\otimes P_\mu,\quad A_{i\mu}=p_i\otimes P_\mu.$$

So $\mathfrak{G}$ is a Lie algebra of dimension 47 whose (nonzero) commutation relations are given by

$$\begin{aligned}
&[L_{\mu\nu},L_{\rho\sigma}]=-g_{\mu\rho}L_{\nu\sigma}-g_{\nu\sigma}L_{\mu\rho}+g_{\mu\sigma}L_{\nu\rho}+g_{\nu\rho}L_{\mu\sigma},\\
&[L_{\mu\nu},X_\rho]=g_{\nu\rho}X_\mu-g_{\mu\rho}X_\nu, \qquad (2.1)\\
&[L_{\mu\nu},C_{\rho\sigma}]=-g_{\mu\rho}C_{\nu\sigma}+g_{\nu\sigma}C_{\mu\rho}-g_{\mu\sigma}C_{\nu\rho}+g_{\nu\rho}C_{\mu\sigma},\\
&[A_{i\mu},Q_{j\nu}]=\delta_{ij}C_{\mu\nu},\\
&[I_{ij},I_{kl}]=-\delta_{ik}I_{jl}-\delta_{jl}I_{ik}+\delta_{il}I_{jk}+\delta_{jk}I_{il},\\
&[I_{ij},Y_k]=\delta_{jk}Y_i-\delta_{ik}Y_j,
\end{aligned}$$

where X_ρ (respectively, Y_k) denotes $T_\rho, A_{i\rho}$, or $Q_{i\rho}$, for i fixed (respectively, $A_{k\rho}, Q_{k\rho}$, for ρ fixed), δ_{ij} the Kronecker symbol which is equal to 1 if $i=j$ and to zero otherwise, and $g_{\alpha\beta}$ is the usual metric tensor ($g_{ij}=\delta_{ij}$ and $g_{4i}=0$, for $1\leqslant i,j\leqslant 3$; $g_{44}=-1$).

Let $\mathfrak{N}_3$ (respectively $\mathbb{R}^4$) be the subalgebra of $\mathfrak{G}$ generated by $(A_{i\mu},Q_{j\nu},C_{\mu\nu})$ [respectively, (T_μ)]. If we denote by $D(j,j')$ the irreducible representation of dimension $(2j+1)(2j'+1)$ of $\mathfrak{sl}(2,\mathbb{C})$, and by $D(j)$ the one of weight j of $\mathfrak{su}(2)$, we have

Proposition 2.1: $\mathfrak{G}$ is the semidirect sum of $\mathfrak{sl}(2,\mathbb{C})\oplus\mathfrak{su}(2)$ by the nilpotent ideal $\mathbb{R}^4\oplus\mathfrak{N}_3$ relative to the representation

$$\{D(\tfrac12,\tfrac12)\otimes D(0)\}\oplus\{D(\tfrac12,\tfrac12)\otimes D(1)\oplus D(\tfrac12,\tfrac12)\otimes D(1)\}$$
$$\oplus\{[D(1,1)\oplus D(0,0)]\otimes D(0)\}.$$

Proof: Let $\mathbb{R}^{10}_c$ be the subalgebra of $\mathfrak{N}_3$ generated by $(C_{\mu\nu})$; it is the center of $\mathfrak{N}_3$ and the quotient algebra $\mathfrak{N}_3/R^{10}_c$ is isomorphic to $\mathbb{R}^{24}$. Then $\mathfrak{N}_3$ is nilpotent as a central extension of commutative Lie algebras.[10] As for the semidirect sum, the relations (2.1) show that the proposition simply follows from the definition of the tensor product of Lie algebras representations[10] and Proposition 2.1 of Ref. 1. Q.E.D.

Let $\mathbb{R}^{10}_c$ (respectively $\mathbb{R}^{12}_a$; $\mathbb{R}^{12}_q$) be the groups generated by $(C_{\mu\nu})$ [respectively, $(A_{i\mu})$; $(Q_{j\mu})$]; we shall denote by $\underset{\sim}{x}$ (respectively, $\underset{\sim}{c}$; $\underset{\sim}{a}$; $\underset{\sim}{q}$; Λ; R) a 4-vector [respectively, the generic element of the groups $\mathbb{R}^{10}_c$; $\mathbb{R}^{12}_a$; $\mathbb{R}^{12}_q$; SL(2, $\mathbb{C}$); SU(2)], $\theta^{\mu\nu}$ the function equal to 1/2 if $\mu=\nu$ and to 1 if not, and $g=\{\underset{\sim}{t},\underset{\sim}{c},\underset{\sim}{a},\underset{\sim}{q},\Lambda,R\}$ the generic element of the connected and simply connected group G of Lie algebra $\mathfrak{G}$. In what follows we shall besides write Λ [respectively, R; $S(\Lambda)$] instead of $D(\frac12,\frac12)(\Lambda)$ (respectively, $D(1)(R)$; $[D(1,1)\oplus D(0,0)](\Lambda)\otimes D(0)(R)$).

Proposition 2.2: The group law of G is given by

$$\begin{aligned}g_1g_2=\{&\underset{\sim}{t}_1+\Lambda_1\underset{\sim}{t}_2,\underset{\sim}{c}_1+S(\Lambda_1)\underset{\sim}{c}_2-\underset{\sim}{\beta}(\underset{\sim}{q}_1,\Lambda_1\otimes R_1\underset{\sim}{a}_2),\\ &\underset{\sim}{a}_1+\Lambda_1\otimes R_1\underset{\sim}{a}_2,\quad \underset{\sim}{q}_1+\Lambda_1\otimes R_1\underset{\sim}{q}_2,\Lambda_1\Lambda_2,R_1R_2\},\end{aligned}$$

where $\underset{\sim}{\beta}$ is defined by

$$\underset{\sim}{\beta}^{\mu\nu}(\underset{\sim}{q}_1,\underset{\sim}{a}_2)=\theta^{\mu\nu}\delta_{ij}\{q_1^{i\mu}a_2^{j\nu}+q_1^{i\nu}a_2^{j\mu}\}.$$

Proof: At first we have to give the group law of the connected and simply connected Lie group N_3 which has $\mathfrak{N}_3$ as Lie algebra. This group law is obtained by direct exponentiation of the Lie brackets. We denote the generic element of N_3 by

$$(\underset{\sim}{c},\underset{\sim}{a},\underset{\sim}{q})=e^{cC}e^{aA}e^{qQ}.$$

Multiplying together two such elements and reordering the product in the same normal form by repeatedly using the Baker–Hausdorff formula and the Lie algebra properties (2.1) of the infinitesimal generators, we obtain

$$(\underset{\sim}{c}_1,\underset{\sim}{a}_1,\underset{\sim}{q}_1)(\underset{\sim}{c}_2,\underset{\sim}{a}_2,\underset{\sim}{q}_2)=(\underset{\sim}{c}_1+\underset{\sim}{c}_2-\underset{\sim}{\beta}(\underset{\sim}{q}_1,\underset{\sim}{a}_2),\underset{\sim}{a}_1+\underset{\sim}{a}_2,\underset{\sim}{q}_1+\underset{\sim}{q}_2).$$

Finally, Proposition 2.2 is a direct consequence of the preceding one. Q.E.D.

Remarks 2.1: (1) The identity element is $\{\underset{\sim}{0},\underset{\sim}{0},\underset{\sim}{0},\underset{\sim}{0},I,I\}$ and the inverse

$g^{-1} = \{ -\Lambda^{-1}\underset{\sim}{t}, -S(\Lambda)^{-1}[\underset{\sim}{c} + \underset{\sim}{\beta}(\underset{\sim}{a},\underset{\sim}{q})],$
$-\Lambda^{-1} \otimes R^{-1}\underset{\sim}{a}, -\Lambda^{-1} \otimes R^{-1}\underset{\sim}{q}, \Lambda^{-1}, R^{-1}\}.$

(2) The definition of $\underset{\sim}{\beta}$ shows us

$\underset{\sim}{\beta}(\Lambda \otimes R\underset{\sim}{a}, \Lambda \otimes R\underset{\sim}{q}) = \underset{\sim}{\beta}(\Lambda \otimes I\underset{\sim}{a}, \Lambda \otimes I\underset{\sim}{q}) = S(\Lambda)\underset{\sim}{\beta}(\underset{\sim}{a}, \underset{\sim}{q})$ $[\forall R \in \mathrm{SU}(2)]$.

B. Some unitary irreducible representations

Proposition 2.2 shows us that the group G admits the decomposition

$$G = H \cdot K = \{\mathbb{R}^4 \times \mathbb{R}^{10}_c \times \mathbb{R}^{12}_a\} \cdot \{\mathbb{R}^{12}_q \cdot (\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SU}(2))\}.$$

Consequently, the most natural method of determining its strongly continuous unitary irreducible representations (U.I.R.) is the method of induced representations[11] (in stages), provided that it turns out to have the required properties.

At first we have the following result:

Lemma 2.1: Let $\hat{h} = \{\underset{\sim}{p}, \underset{\sim}{d}, \underset{\sim}{b}\}$ be the generic element of $\hat{H} = \mathbb{R}^4_p \times \mathbb{R}^{10}_d \times \mathbb{R}^{12}_b$, the dual of the normal subgroup $H = \mathbb{R}^4 \times \mathbb{R}^{10}_c \times \mathbb{R}^{12}_a$ of G, $d^{[\mu\nu]}$ $(\mu \leqslant \nu)$ the components of $\underset{\sim}{d}$ and DG the matrix with the generic element $d^{\mu}_{\lambda} = g_{\lambda\nu} d^{\mu\nu}$, where $d^{\mu\nu} = d^{\nu\mu} = d^{[\mu\nu]}$ and, finally, B (respectively, Q) the matrix, associated with the tensor $\underset{\sim}{b}$ (respectively, $\underset{\sim}{q}$), defined by $b^{\sigma}_{j} = b^{j\sigma}$ (respectively, $q^{\sigma}_{j} = q^{j\sigma}$).

Then the action of $k = \{0, 0, 0, \underset{\sim}{q}, \Lambda, R\} \in K$ on $\hat{H}$ is given by

$$k(\{\underset{\sim}{p}, \underset{\sim}{d}, \underset{\sim}{b}\}) = \{\underset{\sim}{p}', \underset{\sim}{d}', \underset{\sim}{b}'\},$$

where

$$\underset{\sim}{p}' = \Lambda \underset{\sim}{p},$$
$$D'G = \Lambda DG\Lambda^{-1},$$
$$B' = \Lambda BR^{-1} + D'GQ.$$

Proof: The action of K on $\hat{H}$ is defined by

$$\langle k(\hat{h}), h\rangle = \langle \hat{h}, k^{-1}hk\rangle \quad \forall h \in H,$$

where k (respectively, $\hat{h}$) is an element of K (respectively, $\hat{H}$) and $\langle , \rangle$ denotes the action of $\hat{H}$ on H defined here by

$$\langle \hat{h}, h\rangle = \exp i\,\{\underset{\sim}{p} \cdot \underset{\sim}{t} + (\underset{\sim}{b} \cdot \underset{\sim}{a})_1 + (\underset{\sim}{d} \cdot \underset{\sim}{c})_2\},$$

where

$$\underset{\sim}{p} \cdot \underset{\sim}{t} = g_{\mu\nu} p^{\mu} t^{\nu}, \quad (\underset{\sim}{b} \cdot \underset{\sim}{a})_1 = \delta_{ij} g_{\mu\nu} b^{i\mu} a^{j\nu},$$

$$(\underset{\sim}{d} \cdot \underset{\sim}{c})_2 = g_{\mu\nu} g_{\rho\sigma} d^{\mu\rho} K(\underset{\sim}{c})^{\nu\sigma} = \sum_{\mu \leqslant \nu} d_{\mu\nu} c^{[\mu\nu]},$$

$K(\underset{\sim}{c})$ being the symmetric tensor

$$K(\underset{\sim}{c})^{\mu\mu} = c^{[\mu\mu]}, \quad K(\underset{\sim}{c})^{\mu\nu} = K(\underset{\sim}{c})^{\nu\mu} = \tfrac{1}{2} c^{[\mu\nu]} \quad (\mu < \nu),$$

and $(g_{\mu\nu})$ the metric tensor of Lorentz.

This definition of $(\underset{\sim}{d} \cdot \underset{\sim}{c})_2$ and the use of the fact that $D(1,1) \oplus D(0,0)$ is equivalent to the symmetric component of the representation $D(\frac{1}{2},\frac{1}{2}) \otimes D(\frac{1}{2},\frac{1}{2})$ of $\mathrm{SL}(2, \mathbb{C})$ allow us to give the explicit determination of the action of K on $\hat{H}$ given in Lemma 2.1. Q.E.D.

Now in view of the physical applications we wish to draw from our model, we shall confine ourselves in this chapter to the determination of a class of U.I.R. of G associated with the orbits Ω in $\mathbb{R}^4_p \times \mathbb{R}^{10}_d \times \mathbb{R}^{12}_b$ characterized by the stabilized point $\hat{h}_0 = (\underset{\sim}{p}_0, D^0G, 0)$ and the stabilizer $\mathbb{R}^9 \cdot (\mathrm{SU}(2) \times \mathrm{SU}(2))$, where

$$D^0G = \begin{pmatrix} 0 & 0 \\ 0 & -m_0^2 \end{pmatrix}$$

and $\underset{\sim}{p}_0 = (0, 0, 0, m_0)$. The result is given in the following proposition:

Proposition 2.3: The U.I.R. of G induced, starting from Ω, by the U.I.R. of $\mathbb{R}^9 \cdot (\mathrm{SU}(2) \times \mathrm{SU}(2))$ associated with the trivial orbit of $\mathrm{SU}(2) \times \mathrm{SU}(2)$ in $\mathbb{R}^9$ are

$$[U(\underset{\sim}{t}_0, \underset{\sim}{c}_0, \underset{\sim}{a}_0, \underset{\sim}{q}_0, \Lambda_0, R_0)F](\underset{\sim}{p}, \mathbf{x}) = \exp i[\underset{\sim}{p}\underset{\sim}{t}_0 + (\underset{\sim}{b} \cdot \underset{\sim}{a}_0)_1 + (\underset{\sim}{d} \cdot \underset{\sim}{c}_0)_2]$$
$$\times D^s(A_{\underset{\sim}{p}}^{-1} \Lambda_0 A_{\Lambda_0^{-1}\underset{\sim}{p}}) \otimes D^l(R_0)$$
$$\times F(\Lambda_0^{-1}\underset{\sim}{p}, R_0^{-1}(\mathbf{x} - g_{\mu\nu} p^{\mu} \mathbf{q}_0^{\nu})),$$

where $d^{\mu\nu} = p^{\mu}p^{\nu}$, $b^{i\mu} = x^i p^{\mu}$, D^k is the U.I.R. of SU(2) of weight k operating in $\mathbb{C}^{2k+1}$, $\mathbf{q}_0^{\nu}$ (ν fixed) the 3-vector of components $q_0^{i\nu} (i = 1, 2, 3)$, $F \in \mathscr{L}^2_{\mu}(\Omega^{m_0^2}_{+} \times \mathbb{R}^3; \mathbb{C}^{2s+1} \otimes \mathbb{C}^{2l+1})$, $\Omega^{m_0^2}_{+}$ being the one-sheeted hyperboloid defined by $(p^{\mu} p_{\mu} = -m_0^2, p^4 > 0)$, and $d\mu = (\mathbf{p}^2 + m_0^2)^{-1/2}\, d^3\mathbf{p}\, d^3\mathbf{x}$.

Proof: The U.I.R. of G looked for are obtained (up to unitary equivalence) by choosing an arbitrary function $\hat{h} \to \Gamma_{\hat{h}}$, of Ω in K, such that $\Gamma_{\hat{h}}(\hat{h}_0) = \hat{h}$ and an U.I.R. L of the stabilizer of $\hat{h}_0$ acting in a Hilbert space $\mathscr{H}_L$.

These representations are defined in the Hilbert space $\mathscr{L}^2_{\mu}(\Omega, \mathscr{H}_L)$ of functions F on Ω with values in $\mathscr{H}_L$ such that $\int_{\Omega} \|F(\hat{h})\|^2_{\mathscr{H}_L}\, d\mu(\hat{h}) < \infty$, where $d\mu(\hat{h})$ is an invariant measure under K concentrated on Ω, by the formula

$$\{U(h_0k_0)F\}(\hat{h})$$
$$= \langle \hat{h}, h_0\rangle L(\Gamma_{\hat{h}}^{-1} k_0 \Gamma_{k_0^{-1}(\hat{h})}) F(k_0^{-1}(\hat{h})).$$

Now it is sufficient to give the construction of $\Gamma_{\hat{h}}$. Lemma 2.1 shows us that any point of Ω can be parametrized by the multiple

$$\left(A_{\underset{\sim}{p}}\, \underset{\sim}{p}_0, A_{\underset{\sim}{p}} D^0 G A_{\underset{\sim}{p}}^{-1}, A_{\underset{\sim}{p}} \begin{pmatrix} [0] \\ m_0\, {}^t\mathbf{x} \end{pmatrix}\right),$$

where

$$[0] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad {}^t\mathbf{x} = (x^1, x^2, x^3),$$

with $(x^i) \in \mathbb{R}^3$ and $\underset{\sim}{p} \to A_{\underset{\sim}{p}}$ is the field of Lorentz transformations in the canonical formalism,[12] which associates with every $\underset{\sim}{p} \in \Omega^{m_0^2}_{+}$,

$$A_{\underset{\sim}{p}} = [2m_0(p^4 + m_0)]^{-1/2} \begin{pmatrix} m_0 + p^4 + p^3 & p^1 + ip^2 \\ p^1 - ip^2 & m_0 + p^4 - p^3 \end{pmatrix}$$
$$\in \mathrm{SL}(2, \mathbb{C})$$

such that $A_{\underset{\sim}{p}}\underset{\sim}{p}_0 = \underset{\sim}{p}$.

It follows that Ω depends on six real parameters:

$$((p_{\mu}), \text{ with } p_{\mu}p^{\mu} = -m_0^2, \text{ and } x^i \in \mathbb{R}\ (1 \leqslant i \leqslant 3)).$$

Then we can take as field $\hat{h} \to \Gamma_{\hat{h}}$ [such that $\Gamma_{\hat{h}}(\hat{h}_0) = \hat{h}$]:

$$\Gamma_{\hat{h}} = \left(0, 0, 0, A_{\underset{\sim}{p}} \begin{pmatrix} [0] \\ -m_0^{-1}\, {}^t\mathbf{x} \end{pmatrix}, A_{\underset{\sim}{p}}, 1\right). \quad \text{Q.E.D.}$$

Finally, differentiating this representation we obtain for infinitesimal generators (defined on the space of its $\mathscr{C}^{\infty}$-vectors)

$$T_\mu = p_\mu,$$
$$\mathbf{M} = i\mathbf{p}\wedge\nabla_\mathbf{p} + \mathbf{S}_s \otimes 1,$$
$$\mathbf{N} = -(p^4+m_0)^{-1}(\mathbf{p}\wedge\mathbf{S})\otimes 1 + ip^4\nabla_\mathbf{p},$$
$$\mathbf{J} = i\mathbf{x}\wedge\nabla_\mathbf{x} + 1\otimes\mathbf{S}_i,$$
$$A_{i\mu} = x_i p_\mu,$$
$$Q_{i\mu} = i\left(\frac{\partial}{\partial x^i}\right)p_\mu,$$
$$C_{\mu\nu} = p_\mu p_\nu, \tag{2.2}$$

where

$$\mathbf{M} = (M^1, M^2, M^3),\quad \mathbf{N} = (N^1, N^2, N^3),\quad \mathbf{J} = (J^1, J^2, J^3),$$
$$M^i = \tfrac{1}{2}\epsilon^{ijk}L_{jk},\quad N^i = L_{i4},\quad J^i = \tfrac{1}{2}\epsilon^{ijk}I_{jk},$$
$$\nabla_\mathbf{p} = \left(\frac{\partial}{\partial p^1}, \frac{\partial}{\partial p^2}, \frac{\partial}{\partial p^3}\right),\quad \nabla_\mathbf{x} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right),$$
$$\mathbf{S}_j = (S_j^1, S_j^2, S_j^3);$$

$(S_j^k)_{1<k<3}$ are the infinitesimal generators of the representation D^j of SU(2) and $\wedge$ denotes the vector product.

3. COHOMOLOGY GROUPS $H^n(\mathfrak{G},\mathfrak{G})$ ($n = 0,1,2$)

A. Definition of the cohomology groups

In this paragraph we give the main definitions which we shall use for the study of the cohomology groups $H^n(\mathfrak{G},\mathfrak{G})$ (for an interesting review on this subject see Ref. 13).

Let $\mathfrak{A}$ be a Lie algebra over a field K, and let ρ be a representation of $\mathfrak{A}$ by linear transformations of a vector space V of finite dimension over K. A n-linear alternating mapping of $\mathfrak{A}$ into V will be called a n-cochain (with coefficients in V). These n-cochains form a vector space $C^n(\mathfrak{A}, V)$. By definition $C^0(\mathfrak{A}, V) = V$.

We now define three linear applications:

(a) the first, $f \to \delta f$ of $C^n(\mathfrak{A}, V)$ into $C^{n+1}(\mathfrak{A}, V)$ by the formula

$$\delta f(x_1,\dots,x_{n+1}) = \sum_{i<j}(-1)^{i+j}f([x_i,x_j],x_1,\dots,\hat{x}_i,\dots,\hat{x}_j,\dots,x_{n+1}) + \sum_{i=1}^{n+1}(-1)^{i+1}\rho(x_i)f(x_1,\dots,\hat{x}_i,\dots,x_{n+1}) \quad (\text{if } n>0),$$

where the symbol $\hat{}$ over a variable indicates that the appropriate variable is to be omitted.

If $n = 0$ then $f \in V$ and δf is defined by

$$(\delta f)(x) = \rho(x)f.$$

The $n+1$-cochain δf associated with the n-cochain f is called the coboundary of f. This mapping δ has the important property

$$\delta\delta f = 0,\quad \forall f \in C^n(\mathfrak{A}, V)\quad \forall n \geqslant 0.$$

(b) The second, $d_x: C^n(\mathfrak{A}, V) \to C^n(\mathfrak{A}, V)$, for each $x \in \mathfrak{A}$, is defined by setting

$$(d_x f)(x_1,\dots,x_n) = \rho(x)f(x_1,\dots,x_n) - \sum_{i=1}^{n} f(x_1,\dots,[x,x_i],\dots,x_n) \quad (\text{if } n>0),$$
$$d_x f = \rho(x) f \quad (\text{if } n=0).$$

(c) Finally, for $x \in \mathfrak{A}$ we define a third linear mapping $f \to f_x$ of $C^{n+1}(\mathfrak{A}, V)$ into $C^n(\mathfrak{A}, V)$ by setting

$$f_x(x_1,\dots,x_{n-1}) = f(x,x_1,\dots,x_{n-1}) \quad (\text{if } n>0),$$
$$f_x = 0 \quad (\text{if } n=0).$$

Knowing that $\delta\delta f = 0$ we define the space $Z^n(\mathfrak{A}, V)$ of n-cocycles as the kernel of the transformation $\delta: C^n \to C^{n+1}$, and the $B^n(\mathfrak{A}, V)$ of n-coboundaries as the image δC^{n-1}. By definition $B^0(\mathfrak{A}, V) = \{0\}$. The cohomology groups of $\mathfrak{A}$ with coefficients in V are then defined as the quotient spaces $H^n(\mathfrak{A}, V) = Z^n(\mathfrak{A}, V)/B^n(\mathfrak{A}, V)$; $H^0(\mathfrak{A}, V)$ is the subspace of the invariant elements of V.

If $\mathfrak{A}'$ is a subalgebra of $\mathfrak{A}$, $f \in C^n(\mathfrak{A}, V)$ will be called orthogonal to $\mathfrak{A}'$ if

$$f_x = 0 \text{ and } d_x f = 0 \quad \text{for } x \in \mathfrak{A}'.$$

The relations $\delta d_x = d_x\delta$ and $(\delta f)_x + \delta(f_x) = d_x f$ show us if f is orthogonal to $\mathfrak{A}'$, δf also is.

Consequently, if we denote by $C^n(\mathfrak{A}, \mathfrak{A}', V)$ the n-cochains orthogonal to $\mathfrak{A}'$ and define

$$Z^n(\mathfrak{A}, \mathfrak{A}', V) = Z^n(\mathfrak{A}, V) \cap C^n(\mathfrak{A}, \mathfrak{A}', V),$$
$$B^n(\mathfrak{A}, \mathfrak{A}', V) = \delta C^{n-1}(\mathfrak{A}, \mathfrak{A}', V),\quad B^0(\mathfrak{A}, \mathfrak{A}', V) = \{0\},$$
$$H^n(\mathfrak{A}, \mathfrak{A}', V) = Z^n(\mathfrak{A}, \mathfrak{A}', V)/B^n(\mathfrak{A}, \mathfrak{A}', V),$$

we arrive at the notion of the relative cohomology group. Finally, in the following paragraphs $H^n(\mathfrak{R}, V)^{\mathfrak{A}}$ will denote $H^n(\mathfrak{A}, \mathfrak{S}, V)$ if $\mathfrak{A} = \mathfrak{R} \oplus \mathfrak{S}$ is a Levy decomposition of the Lie algebra $\mathfrak{A}$.

B. The cohomology groups $H^n(\mathfrak{G}, \mathfrak{G})$ ($n = 0, 1, 2$)

The $H^n(\mathfrak{G},\mathfrak{G})$ are defined by remarking that $\mathfrak{G}$ is a $\mathfrak{G}$-module for the adjoint representation of $\mathfrak{G}$. Let us denote by $Z(\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2))$ [respectively, $Z(\mathfrak{G})$] the centralizer of $\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2)$ in $\mathfrak{G}$ (respectively, the center of $\mathfrak{G}$). Then we obtain the following results.

Lemma 3.1: The three vector spaces $H^0(\mathfrak{G},\mathfrak{G})$, $Z(\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2))$, and $Z(\mathfrak{G})$ are one-dimensional and generated by $\mathscr{C} = g^{\mu\nu}C_{\mu\nu}$.

Proof: A simple calculation and the fact that $Z^0(\mathfrak{G},\mathfrak{G}) = Z(\mathfrak{G})$ and $B^0(\mathfrak{G},\mathfrak{G}) = \{0\}$ give the result. Q.E.D.

Lemma 3.2: $H^1(\mathfrak{G},\mathfrak{G})$ [respectively, $H^2(\mathfrak{G},\mathfrak{G})$] is isomorphic to $H^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ [respectively, $H^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$].

Proof: Let us consider $\mathfrak{G} = (\mathbb{R}^4 \oplus \mathfrak{N}_3) \oplus (\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2,\mathbb{C}))$, the Levy decomposition of $\mathfrak{G}$. The Hochschild–Serre theorem[14] gives us

$$H^n(\mathfrak{G},\mathfrak{G}) \approx \sum_{i+j=n} H^i(\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2),\mathbb{R}) \otimes H^j(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}.$$

Then it is sufficient to use Lemma 3.1 and the Whitehead lemmas, that is $H^i(\mathfrak{S},\mathbb{R}) = \mathbb{R}$ if $i = 0$ and $\{0\}$ if $i = 1,2$, when $\mathfrak{S}$ is semisimple. Q.E.D.

Proposition 3.2: (1) $B^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}} = \{0\}$. (2) Each $f \in Z^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ can be parametrized by

$$f(T_\mu) = \alpha T_\mu,$$
$$f(A_{i\mu}) = \beta A_{i\mu} + \gamma Q_{i\mu},$$
$$f(Q_{i\mu}) = \beta' A_{i\mu} + \gamma' Q_{i\mu},$$

$$f(C_{\mu\nu}) = (\beta + \gamma')C_{\mu\nu},$$

with $\alpha, \beta, \gamma, \beta', \gamma' \in \mathbb{R}$.

Proof: At first the equalities

$$C^0(\mathbb{R}^4 \oplus \mathfrak{N}_3, \mathfrak{G})^{\mathfrak{G}} = \{ f \in C^0(\mathfrak{G},\mathfrak{G}) = \mathfrak{G}$$

such that

$$d_x f = f_x = 0 \quad \forall x \in \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2)\}$$

$$= \{\lambda g^{\mu\nu} C_{\mu\nu} \text{ with } \lambda \in \mathbb{R}\}$$

show us that $B^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}} = \delta C^0(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}} = \{0\}$. The second assertion results from the fact that each $f \in Z^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ is an element of $C^1(\mathfrak{G},\mathfrak{G})$, which verifies the relations

$$f(s) = 0,$$

$$[s, f(h)] = f([s,h]),$$

$$f([h_1, h_2]) = [f(h_1), h_2] + [h_1, f(h_2)]$$

for any $s \in \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2)$ and $h, h_1, h_2 \in \mathbb{R}^4 \oplus \mathfrak{N}_3$. Q.E.D.

Lemma 3.2 and the preceding proposition give us

Theorem 3.1: The dimension of the $\mathfrak{G}$-module $H^1(\mathfrak{G},\mathfrak{G})$ is equal to 5.

Proposition 3.3: (1) Each $f \in B^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ is defined by $f(X,Y) = 0$ for any $X,Y \in \{L_{\mu\nu}, T_\mu, C_{\mu\nu}, I_{ij};\ 1 \leqslant i,j \leqslant 3,\ 1 \leqslant \mu,\ \nu \leqslant 4\}$.

$f(A_{i\mu}, Q_{j\nu}) = \alpha\delta_{ij} C_{\mu\nu} + \beta\delta_{ij} g_{\mu\nu} g^{\alpha\beta} C_{\alpha\beta}$ with $\alpha, \beta \in \mathbb{R}$.

So the dimension of $B^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ is equal to 2.

(2) $Z^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}} = B^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$.

Proof: Knowing that each $f \in C^1(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ is defined by

$$f(X) = 0 \text{ for any } X \in \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2),$$

$$f(T_\mu) = \alpha T_\mu,$$

$$f(A_{i\mu}) = \beta A_{i\mu} + \gamma Q_{i\mu},$$

$$f(Q_{i\mu}) = \beta' A_{i\mu} + \gamma' Q_{i\mu},$$

$$f(C_{\mu\nu}) = \delta C_{\mu\nu} + \lambda g_{\mu\nu} g^{\alpha\beta} C_{\alpha\beta},$$

with $\alpha, \beta, \gamma, \beta', \gamma', \delta, \lambda \in \mathbb{R}$, the first assertion easily follows from the definition of $B^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$.

Finally, the last result is obtained by a tedious calculation after having remarked that any $f \in Z^2(\mathbb{R}^4 \oplus \mathfrak{N}_3,\mathfrak{G})^{\mathfrak{G}}$ verifies the following properties:

$$f(x_1, x_2) = 0,$$

$$f(x_1, h_1) = 0,$$

$$[x_1, f(h_1, h_2)] = f([x_1, h_1], h_2) + f(h_1, [x_1, h_2]),$$

$$f([h_1, h_2], h_3) + f([h_3, h_1], h_2) + f([h_2, h_3], h_1)$$

$$= [h_1, f(h_2, h_3)] + [h_3, f(h_1, h_2)] + [h_2, f(h_3, h_1)]$$

for any $x_i \in \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{su}(2)$ and any $h_i \in \mathbb{R}^4 \oplus \mathfrak{N}_3$ $(1 \leqslant i \leqslant 3)$. Q.E.D.

Then Lemma 3.2 and this proposition give us the following results:

Theorem 3.2:

$H^2(\mathfrak{G}, \mathfrak{G}) = \{0\}$.

Corollary 3.1: $\mathfrak{G}$ is a rigid Lie algebra, that is, it cannot be deformed into an inequivalent algebra.[15]

Remark 3.1: The above calculations allow us to find again the deformations of some subalgebras which were given in Ref. 1. For instance, we shall study one of these cases in the following paragraph.

C. The $H^n(\mathfrak{A}, \mathfrak{A})$ ($n = 0, 1, 2$) with $\mathfrak{A} = \mathbb{R}^{12} \oplus (\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{su}(2))$

$\mathfrak{A}$ is the subalgebra of $\mathfrak{G}$ generated by $\{L_{\mu\nu}, A_{i\mu} \text{ or } Q_{i\mu}, I_{ij};\ 1 \leqslant \mu,\ \nu \leqslant 4,\ 1 \leqslant i,j \leqslant 3\}$. We immediately see that $H^0(\mathfrak{A}, \mathfrak{A}) = \{0\}$. Furthermore, we easily obtain

Lemma 3.3: $H^1(\mathfrak{A}, \mathfrak{A})$ [respectively, $H^2(\mathfrak{A},\mathfrak{A})$] is isomorphic to $H^1(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}}$ [respectively, $H^2(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}}$].

Proposition 3.4:

(1) $B^1(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}} = \{0\}$.

(2) Each $f \in Z^1(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}}$ is parametrized by

$f(A_{i\mu}) = \alpha A_{i\mu}$ with real α.

Then the dimension of $H^1(\mathfrak{A}, \mathfrak{A})$ is equal to one.

Proposition 3.5:

(1) $B^2(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}} = \{0\}$.

(2) Each $f \in Z^2(\mathbb{R}^{12}, \mathfrak{A})^{\mathfrak{A}}$ is parametrized by

$f(A_{i\mu}, A_{j\nu}) = \lambda\, \{\delta_{ij} L_{\mu\nu} + g_{\mu\nu} I_{ij}\}$ with real λ.

So the dimension of $H^2(\mathfrak{A}, \mathfrak{A})$ is equal to 1. Finally, we end this paragraph with a last result useful for the study of the first order deformations of this Lie algebra $\mathfrak{A}$.

Proposition 3.6: $H^{2\text{int}}(\mathfrak{A}, \mathfrak{A})$ is the whole space $H^2(\mathfrak{A}, \mathfrak{A})$.

Proof: In any deformation, the semisimple subalgebra $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{su}(2)$ of $\mathfrak{A}$ is stable.[16] Then we interest ourself in the first order deformations which keep invariant each Lie bracket in which one element of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{su}(2)$ at least appears.

If $[\ ,\ ]_t$ denotes the Lie bracket of the deformation $\mathfrak{A}_t$ of the Lie algebra $\mathfrak{A}$, we necessarily have for any $h_1, h_2 \in \mathbb{R}^{12}$

$$[h_1, h_2]_t = [h_1, h_2] + t\psi(h_1, h_2) = t\psi(h_1, h_2),$$

with $\psi \in H^2(\mathfrak{A}, \mathfrak{A})$.

The integrability conditions reduce themselves to the relation

$$\sum_{P(h_1,h_2,h_3)} \psi(\psi(h_1,h_2)h_3) = 0,$$

where $h_i \in \mathbb{R}^{12}$ $(1 \leqslant i \leqslant 3)$ and $P(h_1,h_2,h_3)$ is the circular permutation of h_1, h_2, h_3. This last condition is always satisfied. Q.E.D.

4. EXTENSION OF THE LIE ALGEBRA $\mathbb{R}$ BY THE LIE ALGEBRA $\mathfrak{G}$

In order to introduce the dynamics we assume that there exist some development transformations which generate the one-dimensional real algebra $\mathbb{R}$ (in the sequel K will denote one of its generators). The relation that must exist between our Lie algebra $\mathfrak{G}$ and the development Lie algebra will be defined by an exact sequence,

$$\{0\} \to \mathfrak{G} \xrightarrow{\lambda} \tilde{\mathfrak{G}} \xrightarrow{\mu} \mathbb{R} \to \{0\},$$

where λ is an isomorphism into $\tilde{\mathfrak{G}}$ and μ a homomorphism onto $\mathbb{R}$ with $\lambda(\mathfrak{G}) = \operatorname{Ker} \mu$.

Such an inessential extension,[10] since $\mathbb{R}$ is one-dimensional, defines and is defined by a linear mapping which associates to the generator K of $\mathbb{R}$ a derivation ϕ of $\mathfrak{G}$, such that

$[K,x] = \phi(x)$ for any $x \in \mathfrak{G}$.

Moreover, we ask the relativistic invariance as well as the invariance under $\mathfrak{su}(2)$, the isospin algebra. That is to say, we want the relations

$$[K,x] = \phi(x) = 0 \quad [\forall x \in \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{su}(2)], \tag{4.1a}$$

$$[K, T_\mu] = \phi(T_\mu) = 0 \quad (1 \leqslant \mu \leqslant 4). \tag{4.1b}$$

So contrary to Ref. 7 we do not require any *ad hoc* form for the generator K. This latter will be defined by its action on $\mathfrak{G}$ and his representative in the various integrable or local partially integrable representations of $\tilde{\mathfrak{G}}$. Finally, we must point out that the Lie algebras $\tilde{\mathfrak{G}}$ appear as generalizations of the quantum mechanical Galilée algebra and of the Hooke algebra (cf. Roman *et al.*[5,7]).

Now let us determine the various extensions $\tilde{\mathfrak{G}}$ of the Lie algebra $\mathbb{R}$ by the Lie algebra $\mathfrak{G}$. At first the relation (4.1a) shows us that ϕ belongs to $Z^1(\mathbb{R}^4 \oplus \mathfrak{N}_3, \mathfrak{G})^{\mathfrak{G}}$ and (4.1b), with Proposition 3.2.(2), gives us $\alpha = 0$. So we have obtained the following result.

Proposition 4.1: The various Lie algebras $\tilde{\mathfrak{G}}$ are defined by the following brackets:

$$[K, A_{i\rho}] = \beta A_{i\rho} + \gamma Q_{i\rho} \quad (\beta, \gamma \in \mathbb{R}),$$
$$[K, Q_{i\rho}] = \beta' A_{i\rho} + \gamma' Q_{i\rho} \quad (\beta', \gamma' \in \mathbb{R}),$$
$$[K, C_{\rho\sigma}] = (\beta + \gamma')C_{\rho\sigma},$$

the others being the ones of $\mathfrak{G}$.

A. Classification of the various extensions

Proposition 4.2: Let K be a generator of the Lie algebra $\mathbb{R}$. The different extensions $\tilde{\mathfrak{G}}$(up to isomorphism) of $\mathbb{R}$ by the Lie algebra $\mathfrak{G}$, such that

$$[K, L_{\mu\nu}] = [K, T_\mu] = [K, I_{ij}] = 0$$
$$(1 \leqslant \mu, \nu \leqslant 4;\ 1 \leqslant i,j \leqslant 3),$$

are given by

(1) $[K, A_{i\rho}] = [K, Q_{i\rho}] = [K, C_{\rho\sigma}] = 0,$

(2) $[K, A_{i\rho}] = A_{i\rho}, \quad [K, Q_{i\rho}] = -Q_{i\rho}, \quad [K, C_{\rho\sigma}] = 0,$

(3) $[K, A_{i\rho}] = \xi^2 A_{i\rho}, \quad [K, Q_{i\rho}] = -\xi^{-2} Q_{i\rho},$
$[K, C_{\rho\sigma}] = (\xi^4 - 1)\xi^{-2} C_{\rho\sigma} \quad (\xi^2 \neq 0,1),$

(4) $[K, A_{i\rho}] = \xi^2 A_{i\rho}, \quad [K, Q_{i\rho}] = \xi^{-2} Q_{i\rho},$
$[K, C_{\rho\sigma}] = (\xi^4 + 1)\xi^{-2} C_{\rho\sigma} \quad (\xi \neq 0),$

(5) $[K, A_{i\rho}] = \xi^2 A_{i\rho}, \quad [K, Q_{i\rho}] = 0,$
$[K, C_{\rho\sigma}] = \xi^2 C_{\rho\sigma}, \quad (\xi \neq 0),$
[in brackets (3)–(5) ξ is a real number],

(6) $[K, A_{i\rho}] = -Q_{i\rho}, \quad [K, Q_{i\rho}] = [K, C_{\rho\sigma}] = 0,$

(7) $[K, A_{i\rho}] = -Q_{i\rho}, \quad [K, Q_{i\rho}] = A_{i\rho}, \quad [K, C_{\rho\sigma}] = 0,$

(8) $[K, A_{i\rho}] = \cos\varphi_0 A_{i\rho} - \sin\varphi_0 Q_{i\rho},$
$[K, Q_{i\rho}] = \sin\varphi_0 A_{i\rho} + \cos\varphi_0 Q_{i\rho}$
(with $\varphi_0 \neq 0, \pi/2[k\pi]$),
$[K, C_{\rho\sigma}] = 2\cos\varphi_0 C_{\rho\sigma},$

(9) $[K, A_{i\rho}] = A_{i\rho} + Q_{i\rho}, \quad [K, Q_{i\rho}] = Q_{i\rho},$
$[K, C_{\rho\sigma}] = 2C_{\rho\sigma}.$

The other Lie brackets are the ones of $\mathfrak{G}$.

Proof: Let us consider the matrix

$$\Lambda = \begin{pmatrix} \beta & \beta' \\ \gamma & \gamma' \end{pmatrix}$$

of the restriction of adK (cf. Proposition 4.1) to the subspace generated by the vectors $A_{i\rho}$, $Q_{i\rho}$ (i, ρ fixed). Then we look for the different types of matrix Λ (up to equivalence) by classifying the various Jordan forms of such a matrix. To make this, the characteristic polynomial of Λ being $P(\lambda) = \lambda^2 - (\operatorname{Tr}\Lambda)\lambda + \det\Lambda$, we must discuss the six following cases:

$(\operatorname{Tr}\Lambda = 0$ or $\operatorname{Tr}\Lambda \neq 0)$ and $(\det\Lambda = 1, -1, 0)$.

It is sufficient to consider $\det\Lambda = \pm 1$ since, if $\det\Lambda \neq 0$, the transform $K \to K' = |\det\Lambda|^{-1/2} K$ does not change the Lie brackets of $\mathfrak{G}$ and, Λ' denoting the matrix of ad K' restricted to the subspace generated by $A_{i\rho}$, $Q_{i\rho}$, (i,ρ fixed), $\det\Lambda'$ will be equal to ± 1. Now starting from the Lie brackets of Proposition 4.1 (with some matrix Λ) we carry out a real change of basis in the Lie algebra $\tilde{\mathfrak{G}}$ which gives us Proposition 4.2. Q.E.D.

B. Lie groups associated with the preceding Lie algebras

We interest ourselves with the various Lie groups $\tilde{G}$ associated with the Lie algebras (2)–(9) of Proposition 4.2 [the case (1) being a direct sum].

To construct these group laws we denote the generic element of $\tilde{G}$ by

$$g = (k, \underset{\sim}{t}, \underset{\sim}{c}, \underset{\sim}{a}, \underset{\sim}{q}, \Lambda, R) = e^{kK} e^{tT} e^{cC} e^{aA} e^{qQ} \Lambda R,$$

and we use the same method as for Proposition 2.2.

The cases (2)–(5) having the same general form,

$$[K, A_{i\rho}] = \theta_1 A_{i\rho}, \quad [K, Q_{i\rho}] = \theta_2 Q_{i\rho},$$
$$[K, C_{\rho\sigma}] = (\theta_1 + \theta_2) C_{\rho\sigma},$$

we only give one formula for these four cases. So we have obtained

Proposition 4.3: Let $g = \{k, \underset{\sim}{t}, \underset{\sim}{c}, \underset{\sim}{a}, \underset{\sim}{q}, \Lambda, R\}$ be the generic element of one of the Lie groups $\tilde{G}$. The group laws of the various $\tilde{G}$ are given by

(*a*) *Cases (2)–(5)*:

$$g_1 g_2 = \{k_1 + k_2, \underset{\sim}{t}_1 + \Lambda_1 \underset{\sim}{t}_2, e^{-(\theta_1+\theta_2)k_2} \underset{\sim}{c}_1 + S(\Lambda_1) \underset{\sim}{c}_2 - \underset{\sim}{\beta}(e^{-\theta_2 k_2} \underset{\sim}{q}_1, \Lambda_1 \otimes R_1 \underset{\sim}{a}_2),$$
$$e^{-\theta_1 k_2} \underset{\sim}{a}_1 + \Lambda_1 \otimes R_1 \underset{\sim}{a}_2, e^{-\theta_2 k_2} \underset{\sim}{q}_1 + \Lambda_1 \otimes R_1 \underset{\sim}{q}_2, \Lambda_1 \Lambda_2, R_1 R_2\},$$

(b) Case (6):

$$g_1g_2 = \{k_1 + k_2, t_1 + \Lambda_1 t_2, c_1 - 2^{-1}k_2\beta(a_1,a_1) + S(\Lambda_1)c_2 - \beta(q_1 + k_2 a_1, \Lambda_1 \otimes R_1 a_2), a_1 + \Lambda_1 \otimes R_1 a_2, q_1 + k_2 a_1 + \Lambda_1 \otimes R_1 q_2, \Lambda_1\Lambda_2, R_1R_2\}.$$

(c) Case (7):

$$g_1g_2 = \{k_1 + k_2, t_1 + \Lambda_1 t_2, c_1 + S(\Lambda_1)c_2 - \tfrac{1}{4}\sin 2k_2[\beta(a_1,a_1) - \beta(q_1,q_1)] + (\sin k_2)^2\beta(a_1,q_1) - \beta(q_1\cos k_2 + a_1\sin k_2, \Lambda_1 \otimes R_1 a_2), a_1\cos k_2 - q_1 \sin k_2 + \Lambda_1 \otimes R_1 a_2, q_1\cos k_2 + a_1 \sin k_2 + \Lambda_1 \otimes R_1 q_2, \Lambda_1\Lambda_2, R_1, R_2\}.$$

(d) Case (8):

$$g_1g_2 = \{k_1 + k_2, t_1 + \Lambda_1 t_2, e^{-2k_2\cos\varphi}[c_1 + \tfrac{1}{4}\sin(2k_2\sin\varphi)(\beta(q_1,q_1) - \beta(a_1,a_1)) + (\sin(k_2\sin\varphi))^2\beta(a_1,q_1)] + S(\Lambda_1)c_2 - \beta(e^{-k_2\cos\varphi}[\sin(k_2\sin\varphi)a_1 + \cos(k_2\sin\varphi)q_1], \Lambda_1 \otimes R_1 a_2), e^{-k_2\cos\varphi}[\cos(k_2\sin\varphi)a_1 - \sin(k_2\sin\varphi)q_1] + \Lambda_1 \otimes R_1 a_2, e^{-k_2\cos\varphi}[\sin(k_2\sin\varphi)a_1 + \cos(k_2\sin\varphi)q_1] + \Lambda_1 \otimes R_1 q_2, \Lambda_1\Lambda_2, R_1R_2\}.$$

(e) Case (9):

$$g_1g_2 = \{k_1 + k_2, t_1 + \Lambda_1 t_2, e^{-2k_2}[c_1 + 2^{-1}k_2\beta(a_1,a_1)] + S(\Lambda_1)c_2 - \beta(e^{-k_2}[q_1 - k_2 a_1], \Lambda_1 \otimes R_1 a_2), e^{-k_2}a_1 + \Lambda_1 \otimes R_1 a_2, e^{-k_2}[q_1 - k_2 a_1] + \Lambda_1 \otimes R_1 q_2, \Lambda_1\Lambda_2, R_1R_2\}.$$

Remark 4.1: In the sequel of this paper we shall only consider the two following cases:

- the oscillator group, denoted by H, which is defined by Proposition 4.3(c),
- the Galilean group, denoted by L, which is defined by Proposition 4.3(b).

5. SOME INTERESTING U.I.R. OF THE GROUPS H AND L

Now our group $\tilde{G}$ being given, we must determine its U.I.R. in order to obtain the various representatives of the generator K defining the internal development transformation of the systems described by the group G.

In this chapter we shall study only one type of U.I.R. for each group H or L. But to begin with, we remark that our groups H and L admit the decomposition H (or L) $= K \cdot N$, where K is the one-dimensional Lie group generated by K (notation of the preceding chapter) and N, isomorphic to G, is a normal non-abelian subgroup. The theory of the U.I.R. of such semidirect products was given by Mackey.[11] We also refer the reader to the work[17] which is our main reference.

Now we summarize this method.

A. U.I.R. of the general semidirect products $K \cdot N$

Let $\hat{N}$ be the space of equivalence classes of U.I.R. of N. It is then possible to put a Borel structure on $\hat{N}$ in a natural and unique way. The group K will act in $\hat{N}$ as follows: Let $n \rightarrow U(n)$ be unitary representation, i.e., an element of $\hat{N}$. Then the representation $n \rightarrow U(g^{-1}ng)$ is said to be equivalent using K, $g \in K$, to $n \rightarrow U(n)$. We put two (possibly inequivalent) representations of N in the same orbit if they are equivalent using K. Thus $\hat{N}$ is divided up into orbits. One may define the semidirect product to be "regular" if there exists a countable number of Borel subsets of $\hat{N}$ whose union meets each orbit exactly once. If this is the case then the following construction (which works for any semidirect product but in general does not give all representations) does indeed give all representations.

Let χ be a point in $\hat{N}$. Then the stability group of χ (= little group of χ), K_χ, is the subgroup of K such that $\chi_k \simeq \chi$ for all $k \in K_\chi$, where the action $\chi \rightarrow \chi_k$ of K on $\hat{N}$ is defined by $\chi_k : n \rightarrow \chi(k^{-1}nk)$. If χ_1 and χ_2 are two points on the same orbit then K_{χ_1} is isomorphic to K_{χ_2}, and so we may associate one abstract group with each orbit, the little group for the orbit. One may then obtain a representation of the group $G_\chi = K_\chi \cdot N$ as follows.

Let ρ be a representation of K_χ in a Hilbert space $\mathscr{H}_\rho$ and suppose $\chi \in \hat{N}$ acts in a Hilbert space $\mathscr{H}_\chi$. Since χ_k is unitary equivalent to χ for $k \in K_\chi$ we may identify the carrier space of χ_k with $\mathscr{H}_\chi$ for any $k \in K_\chi$. Then there exists a unique (up to a factor) operator $W(k): \mathscr{H}_\chi \rightarrow \mathscr{H}_\chi$ such that $\chi_k(n) = W^{-1}(k)\chi(n)W(k)$ for all $k \in K_\chi$ and $n \in N$. The map $(k,n) \rightarrow W(k)\chi(n)$ defines a representation (projective in general) of G_χ in $\mathscr{H}_\chi$. A more general representation is then of the form $(k, n) \rightarrow \rho(k) \otimes W(k)\chi(n)$ acting in the Hilbert space $\mathscr{H}_\rho \otimes \mathscr{H}_\chi$. Starting from this last one we can obtain a representation of $K \cdot N$ by induction. If the semidirect product is regular, and K has no nontrivial multipliers (which is our case since K is a one-dimensional abelian Lie group), then all the representations of $K \cdot N$ are of this form.

After this summary of the Mackey method we can pass to the physical examples constructed from the U.I.R. U of G, which are given by Proposition 2.3.

B. Oscillator Lie group model H

Proposition 5.1: The U.I.R. of H induced, starting from the U.I.R. U of G, by the U.I.R. of $\mathbb{R}$ (the little group of U) are

$$[V(k_0,t_0,c_0,a_0,q_0,\Lambda_0,R_0)F](p,x) = [\exp i\, 2^{-1}k_0(I - \Delta + \mathbf{x}^2)U(t_0,c_0,a_0,q_0,\Lambda_0,R_0)F](p,\mathbf{x}),$$

where $F \in \mathscr{L}^2_\mu(\Omega^{m_0^2}_{+} \times \mathbb{R}^3; \mathbb{C}^{2s+1} \otimes \mathbb{C}^{2i+1})$, with $d\mu = (\mathbf{p}^2 + m_0^2)^{-1/2}\, d^3\mathbf{p}\, d^3\mathbf{x}$.

Proof: Let us denote by k [respectively, n] the element $\{k, 0, 0, 0, 0, 1, 1\} \in H$ [respectively, $\{0, t, c, a, q, \Lambda, R\} \in H$].

So we make clear the semidirect product $H = K \cdot N$. Now we must study the representations $U(knk^{-1})$ of N for any $k \in K$. Notice that k commutes with

$$t = \{0, \underset{\sim}{t}, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, 1, 1\},$$

$$c = \{0, \underset{\sim}{0}, \underset{\sim}{c}, \underset{\sim}{0}, \underset{\sim}{0}, 1, 1\},$$

$$\Lambda = \{0, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, \Lambda, 1\},$$

$$R = \{0, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, 1, R\}.$$

For this reason, it will be sufficient to consider $U(kak^{-1})$ and $U(kqk^{-1})$ with any $a = \{0, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{a}, \underset{\sim}{0}, 1, 1\}$ and $q = \{0, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{0}, \underset{\sim}{q}, 1, 1\} \in N$. The definition of U (cf. Proposition 2.3) and of its infinitesimal generators (formulas 2.2) show us that

$$U(kak^{-1}) = \exp i\left\{\cos k a^{j\mu}[x_j p_\mu] + \frac{\sin 2k}{4} \sum_{\mu<\nu} \beta(\underset{\sim}{a},\underset{\sim}{a})^{[\mu\nu]}[p_\mu p_\nu]\right\} \times \exp i\left\{-\sin k a^{j\mu}\left[i\frac{\partial}{\partial x^j} p_\mu\right]\right\},$$

$$U(kqk^{-1}) = \exp i\left\{\sin k q^{j\mu}[x_j p_\mu] - \frac{\sin 2k}{4} \sum_{\mu<\nu} \beta(\underset{\sim}{q},\underset{\sim}{q})^{[\mu\nu]}[p_\mu p_\nu]\right\} \times \exp i\left\{\cos k q^{j\mu}\left[i\frac{\partial}{\partial x^j} p_\mu\right]\right\}.$$

Finally, writing

$$U(a) = \exp i a^{j\mu}[x_j p_\mu],$$

$$U(q) = \exp i q^{j\mu}\left[i\frac{\partial}{\partial x^j} p_\mu\right],$$

the Baker–Hausdorff formula gives us, $\forall k \in K$,

$$U(kak^{-1}) = W(k)U(a)W(k^{-1}),$$

$$U(kqk^{-1}) = W(k)U(q)W(k^{-1}),$$

where $W(k)$ is the unitary operator,

$$W(k) = \exp i\, 2^{-1}k(-\Delta + \mathbf{x}^2) \quad (\forall k \in \mathbb{R} \simeq K).$$

Therefore, we have obtained

$$U(knk^{-1}) = W(k)U(n)W(k^{-1}) \quad (\forall k \in K, \forall n \in N).$$

So the orbit of K in $\hat{N}$, on which lies the U.I.R. U of G, is only reduced to one point and K is its little group. Since K is abelian, its U.I.R. are one-dimensional and of the form

$$\exp ilk \quad (l \in \mathbb{R}, k \in \mathbb{R} \simeq K).$$

Then the application of the Mackey method (outlined in Sec. 5.A) gives us the result. Q.E.D.

Differentiating this U.I.R. V of the Lie group H we obtain for infinitesimal generators (defined on the space of its $\mathscr{C}^\infty$-vectors)

$$K = \tfrac{1}{2}(-\Delta + \mathbf{x}^2 + l), \tag{5.1}$$

the others being given by the formulas (2.2).

Note that the formula (5.1) defines the representative of the element of the enveloping algebra $\mathscr{U}(\mathfrak{H})$ ($\mathfrak{H}$ is the Lie algebra of H),

$$A = -(2m_0^2)^{-1} g^{\mu\nu}\delta^{ij}(Q_{i\mu}Q_{j\nu} + A_{i\mu}A_{j\nu}) + l/2,$$

which is suitable to describe the hadron mass spectrum[1] as we shall recall it in the following chapter.

C. Galilean Lie group model L

Proposition 5.2: The U.I.R. of L induced, starting from the U.I.R. U of G, by the characters of $\mathbb{R}$ (the little group of U) are

$$[W(k_0, \underset{\sim}{t}_0, \underset{\sim}{c}_0, \underset{\sim}{a}_0, \underset{\sim}{q}_0, \Lambda_0, R_0)F](\underset{\sim}{p}, \mathbf{x}) = [\exp ik_0(-2^{-1}\Delta + l)U(\underset{\sim}{t}_0, \underset{\sim}{c}_0, \underset{\sim}{a}_0, \underset{\sim}{q}_0, \Lambda_0, R_0)F](\underset{\sim}{p}, \mathbf{x}),$$

where $F \in \mathscr{L}^2_\mu(\Omega^{m_0^2}_+ \times \mathbb{R}^3; \mathbb{C}^{2s+1} \otimes \mathbb{C}^{2i+1})$, with $d\mu = (\mathbf{p}^2 + m_0^2)^{-1/2}\, d^3\mathbf{p}\, d^3\mathbf{x}$.

Proof: In this case we remark that it is sufficient to consider $U(kak^{-1})$. The definitions of U and of its infinitesimal generators give us

$$U(kak^{-1}) = \exp i\left\{a^{j\mu}[x_j p_\mu] + 2^{-1}k \sum_{\mu<\nu} \beta(\underset{\sim}{a},\underset{\sim}{a})^{[\mu\nu]}[p_\mu p_\nu]\right\} \times \exp i\left\{-k a^{j\mu}\left[i\frac{\partial}{\partial x^j} p_\mu\right]\right\}.$$

With the help of the Baker–Hausdorff formula we show that

$$U(kak^{-1}) = W(k)U(a)W(k^{-1}) \quad (\forall k \in K),$$

where $W(k)$ is the unitary operator,

$$W(k) = \exp(-i2^{-1}k\Delta) \quad (\forall k \in K).$$

$U(a)$ was already defined.

Therefore, we have obtained

$$U(knk^{-1}) = W(k)U(n)W(k^{-1}) \quad (\forall k \in K, n \in N).$$

We end according to the same method as for Proposition 5.1. Q.E.D.

Differentiating this U.I.R. W of the Lie group L we obtain for infinitesimal generators (defined on the space of its $\mathscr{C}^\infty$-vectors)

$$K = -\tfrac{1}{2}\Delta + l, \tag{5.2}$$

the others being the ones given by the formulas (2.2). We remark that the formula (5.2) defines the representative of the element of the enveloping algebra $\mathscr{U}(\mathfrak{L})$ ($\mathfrak{L}$ is the Lie algebra of L),

$$B = -(2m_0^2)^{-1} g^{\mu\nu}\delta^{ij}Q_{i\mu}Q_{j\nu} + l.$$

An analogous operator was used[1] in the framework of a Schur-irreducible Poincaré partially integrable local representation[2] to describe the hadron mass spectrum.

Remark 5.1: Before giving the mass-spectrum interpretation of our model, we must remark that the restrictions, to the subgroup G, of the U.I.R. defined in the two preceding propositions are U.I.R. of this subgroup. This noteworthy feature is not general.

6. REPRESENTATIONS OF THE LIE ALGEBRAS $\mathfrak{H}$ AND $\mathfrak{L}$ AND MASS SPECTRUM

In order for the interpretation that is the outcome of our model to be coherent, we must first point out the physical relevance of the generators of group H (or L) and also explain the construction of the mass operators which we adopt.

Let us point out first of all that all generators of $\mathfrak{H}$ (or $\mathfrak{L}$) are relativistically covariant since they are invariant under translations and transform, under the action of the Lorentz group, like a Lorentz tensor. Furthermore, for each fixed pair (μ,ν), $(A_{i\mu},Q_{j\nu})$ $(1\leqslant i,j\leqslant 3)$ are "canonically conjugate" relatively to the intrinsically internal variables. So it is possible to interpret the $A_{i\mu}$ and the $Q_{j\nu}$ (and the $C_{\mu\nu}$ which depend on them algebraically) as describing the relative internal dynamics of a system of composite particles (the components of which may not exist in the free state). Similarly, we shall assume that the mass operator for the composite system can be written (at least in a sufficient approximation) as $M^2 = M_0^2 + M_I^2$, where $M_0^2 = T_\mu T^\mu$ is the relativistic free mass and M_I^2 the mass-splitting term due to the description of particles as a composite system. This mass-splitting term will be given by some combinations of the internal components $(Q_{i\mu}, A_{i\mu})$ the representatives of which, in the physical U.I.R. of Sec. 5, will be proportional to the ones of the generator K and homogeneous to a square mass.

In what follows, M_I^2 will be called (square) mass-observable as it is this term that, in a suitable representation of $\mathfrak{H}$ (or $\mathfrak{L}$), will give the mass spectrum. The various mass-observables which we consider in the following are symmetric homogeneous polynomials of the second degree in the conjugate canonical variables which describe the internal dynamics. We shall thus have a description of the creation of energy from the internal motion, the particles being the excited states of an energetically more fundamental system.

In a way, this interpretation seems to be a generalization of the standard description of systems of n quantum particles possessing only properties that have a classical analog. As a matter of fact there exist, for such systems, n pairs of canonically conjugate variables $(P_i,Q_i)_{1<i<n}$ representing $\mathfrak{h}_n$ such that the Hamiltonian (and also every other physical observable) is a function of it. In our model $\mathfrak{N}_3$, which characterizes the internal dynamics, in a first approximation, would generalize $\mathfrak{h}_n$.

The subgroup SU(2), commuting with the external symmetry described by the Poincaré subgroup, is interpreted as being the isospin group. As for the relativistic spin, we shall continue in this article to take for the definition the one introduced by Bargmann and Wigner (cf., for example, Ref. 12). In this interpretation, the spin remains linked integrally with the external symmetry; but the mass, while depending on the external symmetry, is linked with the degrees of internal excitation of our composite model. In this paragraph we shall also determine the isospin content of the two interaction operators introduced in Secs. 5.A and 5.B, the associated models of which shall, in what follows, be called, respectively, the harmonic oscillator model and the vibrating sphere model. A glance at these operators K and formulas (2.2) easily shows that the eigenfunctions of M^2 will be classified by a principal quantum number connected with the internal excitation level and a secondary quantum number associated with the weight of a U.I.R. of SU(2) (the isospin group).

This interpretation also has the advantage that none of the generators is superfluous, contrary to the various relativistic generalizations of the SU(6) model (cf., for instance, Ref. 18).

Now let us pass to the physical examples.

A. Harmonic oscillator model

This example is defined by the U.I.R. of H, which is given in Proposition 5.1. In this framework we take for the mass-splitting operator

$$M_I^2 = -2^{-1}g^{\mu\nu}\delta^{ij}(Q_{i\mu}Q_{j\nu} + A_{i\mu}A_{j\nu}) \tag{6.1}$$

(we consider the U.I.R. of H with $l = 0$).

The internal degrees of freedom thus contribute to the mass splitting via a term which is (for each space-time direction) the Hamiltonian of the (three-dimensional) harmonic oscillator. This mass-observable is represented by the operator $2^{-1} m_0^2(-\Delta + \mathbf{x}^2)$, which has a discrete spectrum and leads to the mass formula

$$m_n = m_0(n + \tfrac{3}{2})^{1/2},$$

where n is a non-negative integer.

Let $\mathscr{H}_n$ be the eigensubspace associated with the eigenvalue $m^2 = m_0^2(n + \frac{3}{2})$ of the representative $2^{-1} m_0^2$ $(-\Delta + x^2)$ of M_I^2 defined by (6.1). $\mathscr{H}_n$ is written

$$\mathscr{H}_n = \mathscr{L}_\nu^2(\Omega_+^{m_0^2},\mathbb{C}^{2s+1}) \otimes E(n) \otimes \mathbb{C}^{2i+1}$$
$$\text{with } d\nu = (\mathbf{p}^2 + m_0^2)^{-1/2} d^3\mathbf{p},$$

where $E(n)$ is the eigensubspace associated with the eigenvalue $\lambda_n = (2n + 3)$ of the operator $-\Delta + \mathbf{x}^2$.

As the isospin subgroup SU(2) does not operate in $\mathscr{L}_\nu^2(\Omega_+^{m_0^2},\mathbb{C}^{2s+1})$, all we need to know in order to elucidate the isospin content of $\mathscr{H}_n$ is the decomposition of the unitary representation V_I of SU(2) in $E(n) \otimes \mathbb{C}^{2i+1}$ into irreducible components. Now λ_n represents the energy spectrum of the three-dimensional harmonic oscillator, so the degeneracy of the level λ_n is equal to the dimension of the representation $(n, 0)$ (in the notations of Ref. 19) of the group SU(3), which implies that $\dim E(n) = \frac{1}{2}(n + 1)(n + 2)$. Starting from the decomposition of the restriction of the representation $(n,0)$ from SU(3) to SO(3),[20] we obtain

$$V_I = \sum_{s=0}^{n/2} \sum_{k=|n-2s-i|}^{n-2s+i} D^k \quad \text{if } n \text{ is even,}$$
$$V_I = \sum_{s=0}^{(n-1)/2} \sum_{k=|n-2s-i|}^{n-2s+i} D^k \quad \text{if } n \text{ is odd.}$$

As for the spin states of the eigenvectors of M_I^2, they are all equal to s, as the restriction of the representation V to the Poincaré group decomposes on $\mathscr{H}_n$ according to

$$\sum_{\alpha=1}^{(1/2)(n+1)(n+2)(2i+1)} D_\alpha^+(m_0,s),$$

where $D_\alpha^+(m_0,s) = D^+(m_0,s)$ for all α. $D^+(m_0,s)$ is the U.I.R. of positive energy, mass $m_0 > 0$ and spin s of the Poincaré group.

Let us detail the cases $i = 0, \frac{1}{2}$. To a mass $m^2 = m_0^2(n + \frac{3}{2})$ correspond

(a) if $i = 0$, $(n + 2)/2$ [respectively, $(n + 1)/2$] particles with isospin $n - 2r$, where $r = 0, 1,\dots, n/2$ [respectively,

$r = 0, 1, \ldots, (n-1)/2]$ if n is even (respectively, odd), each one having the same spin s;

(b) if $i = \frac{1}{2}$, $(n+1)$ particles with isospin $(2r+1)/2$, where $r = 0, 1, \ldots, n$ and the same spin s.

B. Vibrating sphere model

This example will be defined by a local representation[2] of the Lie algebra $\mathfrak{L}$ constructed upon the differential of the unitary global representation W of the Lie group L which is given in Proposition 5.2 (with $l = 0$). In this framework we take for mass-splitting operator

$$M_I^2 = -2^{-1} g^{\mu\nu}\delta^{ij} Q_{i\mu} Q_{j\nu}. \tag{6.2}$$

Let us point out first of all that the use of the local (partially integrable) representations in this context is justified by the following general remark (cf. also Ref. 21): if the symmetry "differentiates" (naturally) locally, the dynamics (*a fortiori* the internal dynamics), on the other hand, does not necessarily "integrate" globally; in fact, direct use of the theorem of Noether leads (by differentiation) to observables obeying laws of conservation; conversely, if we start out from an algebra of observables, each of them can generate a one-parameter group, but the dynamics by no means requires that there is global integrability; now, in a situation such as this, it is the dynamics which must take precedence. As for the method of investigation of the local representations of $\mathfrak{L}$, it is analogous to that of Ref. 2, the definitions of which we adopt. They are constructed upon the differential of a unitary global representation of L, carried out in a certain functional space, by truncating the domain of variation of the variables with suitable boundary conditions in order to destroy integrability. These boundary conditions can also be considered as reflecting the internal dynamics (or even the internal geometry) of hadrons.

We are going to study here one irreducible Poincaré partially integrable local representation w of the Lie algebra $\mathfrak{L}$, related to the integrable representation W defined above. By irreducibility of w we understand Schur-irreducibility, namely, that every bounded operator B commuting strongly with some integrable observables (i.e., with their spectral resolutions) and weakly with the others is a multiple of the identity operator. More precisely, we require the commutation with the unitary group corresponding to maximum integrable Lie subalgebra of $w(\mathscr{U}(\mathfrak{L}))$ and commutation with the operators of $w(\mathscr{U}(\mathfrak{L}))$ on the common invariant domain. Note that the integrable Lie subalgebra of $w(\mathscr{U}(\mathfrak{L}))$ need not be the representation of a Lie subalgebra of $\mathscr{U}(\mathfrak{L})$, since some elements of $\mathscr{U}(\mathfrak{L})$ can be trivially represented (i.e., by multiples of the identity). In the example that we consider it will be enough to restrict ourselves to elements of degree $\leqslant 2$ in $w(\mathscr{U}(\mathfrak{L}))$. For this w, we introduce two domains: the domain of definition of the representation and the mass-spectrum domain on which the mass-observable considered is represented by an essentially self-adjoint positive operator with a purely discrete spectrum, consisting of isolated eigenvalues. The former is a subspace of the latter.

Let $\mathscr{H}_w = \mathscr{L}^2_\nu(\Omega^{m_0^2}_+, \mathbb{C}^{2s+1}) \otimes \mathscr{L}^2(S) \otimes \mathbb{C}^{2i+1}$, where S is the sphere $\{\mathbf{x}\in\mathbb{R}^3$ such that $|\mathbf{x}|\leqslant a\}$. $\mathscr{L}^2(S)$ decomposes into

$$\mathscr{L}^2(S) = \mathscr{L}^2([0,a], r^2\,dr) \otimes \sum_{j\in\mathbb{N}} \mathscr{H}_j(\Theta,\varphi),$$

where (r, Θ, φ) are the spherical coordinates and $\mathscr{H}_j(\Theta,\varphi)$ is the carrying space of the U.I.R. of weight j of $\mathfrak{so}(3)$. Let S_0 be the dense subspace of $\mathscr{H}_w$ defined by

$$S_0 = \mathscr{S}_{\mathbf{p}}(\mathbb{R}^3, \mathbb{C}^{2s+1}) \otimes \mathscr{C}^\infty_{0r}([0,a]) \otimes \sum_{j\in\mathbb{N}} \mathscr{H}_j(\Theta,\varphi) \otimes \mathbb{C}^{2i+1},$$

where $\mathscr{C}^\infty_{0r}([0,a])$ is the space of the functions $\mathscr{C}^\infty$ (in r) vanishing together with all derivatives at the endpoints of the interval $[0,a]$. S_0 is a common invariant domain of the infinitesimal generators of the representation W on which they are symmetric.

This defines a local representation w of $\mathfrak{L}$ on $\mathscr{H}_w$ by symmetric operators. A basis of $\mathscr{H}_w$ is defined by the following functions:

$$\psi_n(\mathbf{p}, m_s) r^{-1/2} J_{l+1/2}(k_l\, r) Y^l_{m_l}(\Theta,\varphi) V^i_{m_i},$$

with $n\in\mathbb{N}$, $-s\leqslant m_s\leqslant s$, $-i\leqslant m_i\leqslant i$, $l\in\mathbb{N}$, $-l\leqslant m_l\leqslant l$, $k_l\in S_l = \{\lambda; J_{l+1/2}(\lambda a) = 0\}$. $\psi_n(\mathbf{p}, m_s)$ denotes a basis of $\mathscr{L}^2_\nu(\Omega^{m_0^2}_+, \mathbb{C}^{2s+1})$, $J_{l+2/1}$ the Bessel functions, solution to the Sturm–Liouville problem relative to the vibrating sphere[22]; the $Y^l_{m_l}(\theta,\varphi)$ are the spherical harmonics and $(V^i_{m_i};\ -i\leqslant m_i\leqslant i)$ is the basis of $\mathbb{C}^{2i+1}$ that simultaneously diagonalizes the Casimir operator and the Cartan subalgebra of the Lie algebra $\mathfrak{su}(2)$. Now we use the strong commutation with $w(L_{\mu\nu})$, $w(T_\mu)$, $w(g^{\mu\nu}\delta^{ij}Q_{i\mu}Q_{j\nu})$, which generates a $P\times \mathrm{U}(1)$ group, $w(I_{ij} - L_{ij})$ and $w(L_{ij})$, where $L_{ij} = g^{\mu\nu}(Q_{i\mu}A_{j\nu} - Q_{j\mu}A_{i\nu})$ $(l\leqslant i, j\leqslant 3)$. The L_{ij} do not close to a $\mathfrak{so}(3)$ subalgebra of $\mathscr{U}(\mathfrak{L})$; but, since $w(g^{\mu\nu}T_\mu T_\nu)$ factors out in $w(L_{ij})$ and is a multiple of the identity, the $w(L_{ij})$ generate the quasiregular representation of SO(3) in the $\mathbf{x}$ part of $\mathscr{H}_w$. We therefore require that B commutes with the above unitary representation of $P\times \mathrm{U}(1)\times \mathrm{SU}(2)\times \mathrm{SO}(3)$; weak commutation with $w(g^{\mu\nu}T_\mu Q_{3\nu})$ will then imply that B is a multiple of the identity.

Let

$$S_\pi = \mathscr{L}^2_\nu(\Omega^{m_0^2}_+, \mathbb{C}^{2s+1}) \otimes \mathscr{C}^\infty_{0r}([0,a]) \otimes \sum_{l\in\mathbb{N}} \mathscr{H}_l(\theta,\varphi) \otimes \mathbb{C}^{2i+1},$$

$\mathscr{C}^\infty_{or}([0,a])$ being the space of the functions $\mathscr{C}^\infty$ (in r), be vanishing at the right end point a.

The representative of the mass-splitting operator M_I^2, defined by (6.2), is essentially self-adjoint on the domain S_π. It has a discrete spectrum which leads to the mass formula

$$m_{l,k_l} = \frac{m_0}{\sqrt{2}}|k_l| \quad (l\in\mathbb{N},\ k_l\in S_l).$$

Remark: Using also a three-dimensional "internal space" but a local representation of the 11-parameter Weyl Lie algebra, and taking for total (squared) mass-operator the second order invariant of the (integrable) Poincaré subalgebra, Snellman[23] recently obtained a discrete mass spectrum in terms of squares of the zeros of Bessel functions relative to the vibrating sphere. The technique used for w is similar, but this representation is Schur-irreducible and our interpretation is different.

Let $\mathscr{H}(l, k_l)$ be the eigensubspace associated with the eigenvalue $m^2 = 2^{-1}m_0^2 k_l^2$ of the interaction operator $M_I^2 = -2^{-1} g^{\mu\nu} \delta^{ij} Q_{i\mu} Q_{j\nu}$, which is represented in w by $M^2 = -2^{-1} m_0^2 \Delta$. $\mathscr{H}(l, k_l)$ is written

$$\mathscr{H}(l,k_l) = \mathscr{L}_v^2(\Omega_+^{m_0^2}, \mathbb{C}^{2s+1}) \otimes E(l,k_l) \otimes \mathbb{C}^{2i+1},$$

where $E(l, k_l)$ is the subspace of $\mathscr{L}^2(S)$ generated by the vectors

$$r^{-1/2} J_{l+1/2}(k_l r) Y^l_{m_l}(\theta,\varphi), \quad -l \leqslant m_l \leqslant l.$$

As in the preceding case, in order to elucidate the isospin content of $\mathscr{H}(l, k_l)$, it is enough to know the decomposition of the unitary representation w_I of SU(2) in $E(l, k_l) \otimes \mathbb{C}^{2i+1}$. Now $w_I = \sum_{k=|l-i|}^{l+i} D^k$. So the states of $\mathscr{H}(l, k_l)$ are all of spin s and mass $m = (m_0/\sqrt{2})|k_l|$ (k_l being determinable approximately for large masses using an asymptotic representation of the Bessel functions) and their isospin (for the particular values $i = 0, \frac{1}{2}$, for example) is

(a) $I = l$ if $i = 0$,

(b) $I = \frac{1}{2}$ (respectively, $I = l \pm \frac{1}{2}$) if $l = 0$ (respectively, $l \geqslant 1$) and $i = \frac{1}{2}$.

7. DISCUSSION AND OUTLOOK

After having considered the experimental trajectories $I = f(m^2)$ and $I = f(m)$ (with constant spin) for those particles which are best established at present,[24] it emerges that our model presents a real potential interest, as the observed divergences from the theoretical trajectories (both those associated with the harmonic oscillator model and those associated with the vibrating sphere model) are not significant enough to invalidate it; all the more so, as we have only taken into account a restricted number of quantum numbers. The presence of high isospins which do not correspond to observed particles could be accounted for by the fact that they are associated with particles heavy enough to decay very quickly by means of strong interactions. Furthermore, the fact that this model permits defining a denumerable infinity of fermions (or of bosons) of integer or half-integer isospin (in an irreducible representation) brings up the question of the interest of the "multiquark" theories ("colored", "charmed", ...), especially as there are some very strong presumptions as to the existence of massive hadrons which do not fit into the framework of the present multiquark theories.

We end this work by mentioning briefly a few remarks and suggestions concerning its possible continuations

(1) The relation $\mathbf{J} = 1 \otimes \mathbf{S}_k + i\mathbf{x} \wedge \nabla_\mathbf{x}$ [cf. formulas (2.2)] giving the isospin can be interpreted as follows. $\mathbf{S}_k$ represents the isospin of the basic (nonexcited) state and $i\mathbf{x} \wedge \nabla_\mathbf{x}$ is the orbital momentum of internal excitation creating the real isospin. If a mass formula were desired depending explicitly on the isospin, one could add to M_I^2 an interaction of the (iso) spin–orbit type of the form $I_{ij}\Sigma^{ij}$, where Σ^{ij} is an appropriate function of the internal canonical variables $(A_{i\mu}, Q_{j\nu})$. For instance, if we take the mass-observable $M^2 = M_I^2 - g\, I_{ij}\, \Sigma^{ij}$ (g being a coupling constant), with $\Sigma_{ij} = g^{\mu\nu}(A_{i\mu} Q_{j\nu} - A_{j\mu} Q_{i\nu})$ we obtain, in the harmonic oscillator framework, the following mass formula:

$$m^2 = m_0^2(n + \tfrac{3}{2}) + g m_0^2 [I(I+1) + l(l+1) - k(k+1)],$$

with $l = n - 2s$, $0 \leqslant s \leqslant$ integer part of $n/2$,

$$I \in \{n - 2s + k, n - 2s + k - 1, ..., |n - 2s - k|\}.$$

(2) In another perspective it would be interesting to define (working from internal canonical variables) operators $X^{\pm}_{i\mu}$ of the creation and annihilation type, such that the internal quantum numbers (including the isospin) be functions of them, and to deduce from the latter the internal quantum number content of the various mass formulas considered (as we have done above for the isospin), or mass formulas which depend on them explicitly, as in the preceding paragraph.

(3) It would be interesting to study other (local or global) representations of one of the extensions which are constructed in Proposition 4.3, in order to obtain more general mass-operators K. For some of them, contrary to those studied in this paper, a spin spectrum might be found which would lead to a group theoretical formulation of the Regge trajectories. More simply, in view of the results obtained in this paper, the fact that the group $H_4 \cdot \mathrm{SL}(2, \mathbb{C})$ [25] admits (nonfaithful) U.I.R. such that the Poincaré mass is a strictly positive scalar and such that the spin spectrum is composed of all the integers or all the half-integers presents another possibility for arriving at a formulation of Regge trajectories by building representations of $\mathfrak{H}$ (or $\mathfrak{L}$), according to our analysis (cf. Sec. 2), for which the restriction to Poincaré subalgebra gives a spin spectrum, as in the Flato–Snellman model (Ref. 26, Sec. 2C).

(4) Let us mention in conclusion that the methods employed in this paper could equally be considered in the framework of the algebras of supersymmetries by adjoining anticommutation relations to the commutation relations of $\mathfrak{h}_3$.

ACKNOWLEDGMENTS

Our study of this problem was motivated by ideas of M. Flato. We wish to thank him for friendly help and criticism on the topics of this work. We should like also to thank D. Sternheimer for his careful reading of the manuscript and useful remarks.

[1]D. Beau and S. Horchani, J. Math. Phys. **20**, 1700 (1979).

[2]M. Flato and D. Sternheimer, Phys. Rev. Lett. **16**, 1185 (1966); Commun. Math. Phys. **12**, 296 (1969); D. Sternheimer, "Represention locales d'algèbres de Lie" (Mimeographed Memoir, Collège de France, 1968, unpublished).

[3]J. Kogut and L. Susskind, Phys. Rep. C **8**, 75 (1973).

[4]M. Gell-Mann, Phys. Lett. **8**, 214 (1964); O. W. Greenberg, Phys. Rev. Lett. **13**, 598 (1964); R. H. Dalitz, in *Proceedings of the 13th International Conference on High Energy Physics, Berkeley, 1966* (University of California, Berkeley, 1967); in *Lectures at the 2nd Hawaii Topical Conference on Particle Physics* (University of Hawaii, Honolulu, 1967); D. Faiman and A. W. Hendry, Phys. Rev. **173**, 1720 (1968); **180**, 1572 (1969); L. A. Copley, G. Karl, and E. Obryk, Phys. Lett. B **29**, 117 (1969); D. L. Katyal and A. N. Mitra, Phys. Rev. D **1**, 338 (1970); D. K. Choudhury and A. N. Mitra, Phys. Rev. D **1**, 351 (1970).

[5]R. P. Feynman, M. Kislinger, and F. Ravndal, Phys. Rev. D **3**, 2706 (1971); J. L. Rosner, Phys. Rev. D **6**, 1781 (1972); M. Böhm, H. Joos, and M. Krammer, Nucl. Phys. B **51**, 397 (1973); in *Recent Developments in Mathematical Physics*, edited by P. Urban (Springer, New York, 1973); P. J. Walters, A. M. Thomson, and F. D. Gault, J. Phys. A. **7**, 1681 (1974); P. Roman and J. Haavisto, J. Math. Phys. **22**, 403 (1981).

[6]M. Guenin, in *Mathematical Physics and Physical Mathematics*, edited by K. Maurin and R. Raczka (Reidel, Dordrecht, Holland, 1976).
[7]P. Roman and J. P. Leveille, J. Math. Phys. **15**, 1760 (1974); **15**, 2053 (1974).
[8]Y. Katayama and H. Yukawa, Prog. Theor. Phys. Suppl. **41**, 1 (1968); Y. Katayama, I. Umemura, and H. Yukawa, *ibid.* **41**, 22 (1968); P. Roman and J. Haavisto, Int. J. Theor. Phys. **16**, 915 (1977).
[9]P. A. M. Dirac, Proc. R. Soc. London Ser. A **322**, 435 (1971); **328**, 1 (1972); E. C. G. Sudarshan, N. Mukunda, and C. C. Chiang, preprint of the University of Texas at Austin.
[10]N. Bourbaki, *Groupes et Algèbres de lie* (Hermann, Paris, 1960), Chap. 1.
[11]G. W. Mackey, *The theory of Unitary Group Representations* (University of Chicago, Chicago, 1976).
[12]A. S. Wightman, in *Relations de dispersions et particules élémentaires*, edited by C. de Witt and R. Omnes (Hermann, Paris, 1960).
[13]C. Chevalley, and S. Eilenberg, Trans. Am. Math. Soc. **63**, 85 (1948).
[14]G. Hochschild and J. P. Serre, Ann. Math. **57**, 591 (1953).
[15]M. Gerstenhaber, Ann. Math. **79**, 59 (1964).
[16]S. Page and R. W. Richardson, Trans. Am. Math. Soc. **127**, 302 (1967); R. W. Richardson, Illinois J. Math. **11**, 92 (1967).
[17]R. F. Streater, Commun. Math. Phys. **4**, 217 (1967).
[18]*Proceedings of the Seminar on High-energy Physics and Elementary Particles*, Trieste, 3 May–30 June 1965 (IAEA, Vienna, 1965).
[19]G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[20]J. W. B. Hugues, J. Phys. A **6**, 281 (1973).
[21]M. Flato, in *Mathematical Physics and Physical Mathematics*, edited by K. Maurin and R. Raczka (Reidel, Dordrecht, Holland, 1976).
[22]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1953), Vol. I.
[23]H. Snellman, J. Math. Phys. **19**, 2278 (1978).
[24]N. Barash-Schmidt *et al.*, Rev. Mod. Phys. **42**(2), Part II, April 1976; T. G. Trippe *et al.*, Phys. Lett. B **68**, 1 (1977).
[25]E. Angelopoulos, Ann. Inst. Henri Poincaré **18**, 39 (1973).
[26]M. Flato and H. Snellman, J. Math. Phys. **15**, 857 (1974).

Interaction of a spin-1/2 tachyon with a spin-1/2 bradyon and a tachyon–bradyon model

Sushil K. Srivastava
Government College, Port Blair 744104, India

(Received 6 April 1981; accepted for publication 14 May 1982)

Interaction of a spin-1/2 tachyon with the gravitational field of a spin-1/2 bradyon is investigated in the background of Schwarzschild geometry. A tachyon–bradyon bound system (PTBBS) is proposed. In PTBBS, a tachyon revolves around a bradyon in a circular orbit. It is found that as the tachyon goes away from the bradyon, it emits energy under a certain condition; but if this condition is not satisfied, the tachyon gains energy while going away from the bradyon. Also, PTBBS is compared with Bohr's hydrogen atom model. In the end, the possibility of formation of PTBBS is discussed.

PACS numbers: 14.80.Pb, 12.25. + e, 04.20. – q

1. INTRODUCTION

There are two schools of thought regarding the existence of tachyons. One is that investigations on tachyons are a waste of time, because so far attempts to detect or to produce them have yielded null results. Another thought is that tachyons should exist in our universe and research on them is highly needed for proper understanding of the theory of relativity. Another reason for study of tachyons is that one cannot draw a Feynman diagram without, in some sense, talking about tachyons. Corben[1] asks the question "How could an electron turn the corner and go backwards in time without passing through a state where it was a tachyon?" He says that one could argue that, during that period, it was in a virtual state and that the Feynman diagram is not to be taken literally anyhow.[1]

However, if one believes in the existence of tachyons and does not consider their study a waste of time, as we do here, one can naturally ask a question. "What happens when a tachyon interacts with a bradyon (a particle of ordinary matter)?" An answer to this question is also necessary to draw the attention of scientists of the first school of thought, in favor of tachyons. In this article, we have dealt with this problem, though in recent years many authors such as Narlikar and Sudarshan,[2] Dhurandhar,[3] Honig *et al.*,[4] and Narlikar and Dhurandhar[5] have discussed such problems in different ways.

In the present paper, we consider that a spin-1/2 tachyon, having its metamass μ equal to the mass of an electron, is revolving around a spin-1/2 bradyon (say a proton). Thus, we propose that a spin-1/2 tachyon and a spin-1/2 bradyon form a bound system like Bohr's model of the hydrogen atom in which electron revolves around proton in circular orbits. Here we suppose that the gravitational field of the bradyon is described by the Schwarzschild geometry.

In Sec. 2, we have considered a spacelike state function $\psi(t,R,\theta,\phi)$ for a spin-1/2 tachyon of metamass μ, such that its rest mass[6] $\bar{m} = i\mu$ satisfies the Dirac equation. Also, we have transformed the first order Dirac equation into a second-order equation by applying the operator $(i\hbar\gamma^\mu p_\mu + \bar{m}c)$.

In Sec. 3, we have discussed various aspects of the effective potential of a tachyon in the proposed tachyon–bradyon bound system (PTBBS). In Sec. 4, we have solved the differential equation derived in Sec. 2 by the method of Wentzel–Kramers–Brillouin (WKB) approximation. In this section, we have also computed the probability density of a tachyon. In Sec. 5, we have tested the validity of WKB solutions obtained in Sec. 4.

In Sec. 6, we have computed the energy of a tachyon in PTBBS. In Sec. 7, we have compared PTBBS with Bohr's hydrogen atom. In Sec. 8, we have speculated that PTBBS would have formed after the epoch of the big bang if the assumption of Narlikar and Sudarshan[2] regarding production of primordial tachyons is correct.

Throughout the entire paper we have used cgs system of units. Also everywhere in this paper we have used $G = 10^{38}G_N$ (where G_N is the Newtonian gravitational constant). The reason behind using this value of G is given in Sec. 2.

2. DIRAC EQUATION AROUND BRADYON (PROTON)

The Schwarzschild line element is given by

$$ds^2 = \left(1 - \frac{2mG}{Rc^2}\right)dt^2 - \left(1 - \frac{2mG}{Rc^2}\right)^{-1} dR^2 - R^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{2.1}$$

where $R = 2mG/c^2$ gives the radius of bradyon and m is its mass. Here we choose G as the Gravitational constant of strong gravity which is 10^{38} times of the Newtonian gravitational constant, because according to Sivaram and Sinha[7] the Newtonian gravitational constant yields very unphysical results at the microscopic level, e.g., it yields the radius of a proton on the order of 10^{-52} cm which is not correct. c is the speed of light.

The motion of a tachyon, in the Schwarzschild space–time, is described by the Dirac equation

$$(i\hbar\gamma^\mu p_\mu - \bar{m}c)\psi = 0, \tag{2.2}$$

where p_μ and $\bar{m}$ are the four-momentum and rest mass of tachyon, respectively.

For any function ψ there exists a time-reversed function $\bar{\psi}$ given as

$$T\psi(t,R,\theta,\phi) = u_T\bar{\psi}(-t,R,\theta,\phi), \tag{2.3}$$

where u_T is a unitary matrix satisfying the condition

 0022-2488/83/051311-06$02.50

$$u_T\tilde{\gamma} = -\gamma u_T.$$

Hence, if ψ is the solution of Eq. (2.2), it satisfies

$$\bar{\psi}(i\hbar\gamma^\mu p_\mu + \bar{m}c) = 0. \tag{2.4}$$

We transform the first-order equation (2.2) to a second-order equation by applying the operator $(i\hbar\gamma^\mu p_\mu + \bar{m}c)$ as

$$(p^2 - \bar{m}^2c^2)\psi = 0. \tag{2.5}$$

On substituting $p^\mu = i\hbar\partial/\partial x^\mu$ in Eq. (2.5) we have

$$(\Box^2 + \bar{m}^2c^2/\hbar^2)\psi = 0, \tag{2.6}$$

where

$$\Box^2 = \frac{1}{(-g)^{1/2}}\frac{\partial}{\partial x^i}\left((-g)^{1/2}g^{ij}\frac{\partial}{\partial x^j}\right). \tag{2.7}$$

On substituting g^{ij} from Eq. (2.1) in Eq. (2.6) we have

$$\frac{\partial^2\psi}{\partial t^2} - \frac{c^2}{R^2}\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial}{\partial R}\left[R^2\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial\psi}{\partial R}\right]$$
$$- \frac{c^2}{R^2}\left(1 - \frac{2mG}{Rc^2}\right)\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)$$
$$- \frac{c^2}{R^2\sin^2\theta}\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial}{\partial\phi}\left(\frac{\partial\psi}{\partial\phi}\right)$$
$$- \frac{\mu^2c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right)\psi = 0. \tag{2.8}$$

Since ψ is independent of ϕ due to the azimuthal symmetry of the problem about $\theta = 0$,

$$\frac{\partial\psi}{\partial\phi} = 0. \tag{2.9}$$

Now ψ can be expanded in terms of a complete set of eigenfunctions of the angular momentum operator, that is, we can set

$$\psi = \sum_l\left[\frac{1}{R}\psi_l(R,t)\right]\phi_l(\cos\theta). \tag{2.10}$$

Substituting ψ in Eq. (2.8) we have a partial differential equation for $\psi_l(R,t)$ as

$$\frac{\partial^2\psi_l}{\partial t^2} - \frac{c^2}{R^2}\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial}{\partial R}\left[R^2\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial\psi_l}{\partial R}\right]$$
$$+ c^2\frac{l(l+1)}{R^2}\psi_l - \frac{\mu^2c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right)\psi_l = 0. \tag{2.11}$$

To separate out t from $\psi_l(R,t)$ we can write by Fourier analysis

$$\psi_l(R,t) = \int_{-\infty}^{\infty} A(\Omega)\psi_l^\Omega(R)\exp(-i\Omega t)\,d\Omega. \tag{2.12}$$

Substituting it in (2.11) we have an ordinary differential equation

$$\frac{d^2\psi_l^\Omega}{dR^2} + \frac{2mG}{R(Rc^2 - 2mG)}\frac{d\psi_l^\Omega}{dR} + \frac{R^2c^2}{(Rc^2 - 2mG)^2}$$
$$\times\left[\Omega^2 + \left(1 - \frac{2mG}{Rc^2}\right)\frac{\mu^2c^4}{\hbar^2} - \frac{2mG}{R^3}\left(1 - \frac{2mG}{Rc^2}\right)\right.$$
$$\left.- \frac{c^2l(l+1)}{R^2}\left(1 - \frac{2mG}{Rc^2}\right)\right]\psi_l^\Omega = 0. \tag{2.13}$$

In Eq. (2.14) the coefficient of ψ_l^Ω consists of two parts.

(i) the term involving Ω^2 corresponds to the energy or frequency of the partial wave;

(ii) the other three terms constitute the effective potential given by

$$V^2(R) = \frac{\mu^2c^4R}{\hbar(Rc^2 - 2mG)} - \frac{2mG}{R^2(Rc^2 - 2mG)}$$
$$- \frac{c^2l(l+1)}{R(Rc^2 - 2mG)}. \tag{2.14}$$

3. POTENTIAL OF TACHYON IN PTBBS

The potential of tachyon in PTBBS is given by Eq. (2.14). For V maximum, we have

$$\frac{dV}{dR} = 0 \tag{3.1}$$

which yields

$$R^3 - \frac{\hbar^2l(l+1)}{mG\mu^2}R^2 + \frac{\hbar^2}{\mu^2c^2}[3 + l(l+1)]R + \frac{4mG\hbar^2}{\mu^2c^4} = 0. \tag{3.2}$$

This is a cubic equation having three roots

$$(q - H/q) - b, \quad (\omega q - \omega^2H/q) - b, \quad (\omega^2q - \omega H/q) - b, \tag{3.3}$$

where

$$b = \frac{\hbar^2l(l+1)}{3mG\mu^2},$$
$$H = \frac{\hbar^2}{\mu^2c^2}\left[1 + \frac{l(l+1)}{3} - \frac{\hbar^4l^2(l+1)^2}{9m^2G^2\mu^4}\right],$$
$$F = \frac{4mG\hbar^2}{\mu^2c^4} + \frac{3\hbar^4l(l+1)}{mG\mu^4c^2}$$
$$\times\left[1 + \frac{l(l+1)}{3} - \frac{2h^6l^2(l+1)^2}{m^3G^3\mu^6}\right],$$
$$q = [\tfrac{1}{2}\{-F + (F^2 + 4H^3)^{1/2}\}]^{1/3},$$

and 1, ω, and ω^2 are cube roots of unity. Of the above three roots, the last two roots are complex; hence, we are interested in the first root only, i.e., Eq. (3.2) gives

$$R = q - H/q - b.$$

It shows that V will be maximum at

$$R = q - H/q - b. \tag{3.4}$$

If l is small,

$$b \approx 0,$$
$$H \approx \hbar^2/\mu^2c^2,$$
$$F \approx 4mG\hbar^2/\mu^2c^4,$$
$$q \approx \left[\tfrac{1}{2}\left\{-\frac{4mG\hbar^2}{\mu^2c^4} + \left(\frac{16m^2G^2\hbar^4}{\mu^4c^8} + \frac{4\hbar^6}{\mu^6c^6}\right)^{1/2}\right\}\right]^{1/3}$$

On substituting

$$m = 1.67\times10^{-24}, \quad \mu = 9.11\times10^{-28},$$
$$G = 6.67\times10^{30}, \quad c = 3\times10^{10}, \quad \hbar = 1.05\times10^{-27}$$

(all values in cgs units) we find that

$$\frac{\hbar^4c^2}{4\mu^2m^2G^2} \approx 10^{-38} \ll 1$$

and

$\hbar^2 \ll \hbar.$

Hence, we have

$$q = [2mG\hbar/\mu^2c^4]^{1/3} = 4.12\times 10^{-3},$$
$$H = \hbar^2/\mu^2c^2 = 0.15\times 10^{-21}. \quad (3.5)$$

Now,

$$R \approx 4.12\times 10^{-3}\ \text{cm}. \quad (3.6)$$

This shows that at distance 4.12×10^{-3} cm from the center of the bradyon (proton) the potential of the tachyon would be maximum if l is small.

In the same way we can get the value of R putting values of m, G, μ, c, and l, in case l is large.

In the following table we find the value of potential of the tachyon corresponding to different values of R when l is small.

S.No.	Distance from the center of bradyon (proton) R (in centimeters)	Potential of tachyon (in ergs)
1	2.475×10^{-14} (radius of proton)	$-\infty$
2	4.198×10^{-12}	0
3	4.121×10^{-3}	$V_{\max} > 1.59\times 10^{10}$
4	∞	2.59×10^{10}

From the above table we learn that potential of the tachyon increases as R increases. This is maintained up to $R = 4.12\times 10^{-3}$ cm. But beyond 4.12×10^{-3} cm the potential decreases as R increases, but very slowly. This shows that the field is attractive for a tachyon up to $R = 4.12\times 10^{-3}$ cm, but beyond this value of R the bradyon (proton) field becomes repulsive for tachyons. If a tachyon reaches the surface of bradyon (proton), it will be absorbed in the bradyon (proton). This phenomenon will take place when the distance of the tachyon from the center of the bradyon (proton) is less than 4.198×10^{-12} cm. At $R = 4.198\times 10^{-12}$ cm, the potential of the tachyon vanishes. Hence, when the value of R is 4.198×10^{-12} cm the tachyon can not revolve around the bradyon (proton). From the above discussion, we find that circular orbits of a tachyon revolving around a bradyon (proton) can be found between $R = 4.198\times 10^{-12}$ cm and $R = 4.121\times 10^{-3}$ cm.

4. WKB SOLUTIONS

In the differential equation (2.13) we see that the coefficient of $d\psi_l^\Omega/dR$ and $2mG/R^2(Rc^2 - 2mG)$ are very small and can be neglected except in the shell where $|R - 2mG/c^2| < \epsilon$ (where ϵ is a very small positive quantity). Hence, we have the differential equation (2.13) as

$$\frac{d^2\psi_l^\Omega}{dR^2} + \left[\frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} + \frac{\mu^2c^4}{\hbar^2(c^2 - 2mG/R)} - \frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)}\right]\psi_l^\Omega = 0. \quad (4.1)$$

This equation is written in one of the forms

$$\frac{d^2\psi_l^\Omega}{dR^2} + k^2(R)\psi_l^\Omega = 0 \quad \text{for} \quad k^2(R) > 0, \quad (4.2)$$

where

$$k^2(R) = \frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} + \frac{\mu^2c^4}{\hbar^2(c^2 - 2mG/R)} - \frac{c^2\Omega(l+1)}{R^2(c^2 - 2mG/R)}, \quad (4.3)$$

$$\frac{d^2\psi_l^\Omega}{dR^2} - k'^2(R)\psi_l^\Omega = 0 \quad \text{for} \quad k'^2(R) > 0, \quad (4.4)$$

where

$$k'^2(R) = \frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)} - \frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} - \frac{\mu^2c^4}{\hbar^2(c^2 - 2mG/R)}. \quad (4.5)$$

We restrict our attention for the present to Eq. (4.2); we shall be able to generalize the resulting expression for ψ_l^Ω to obtain solutions of (4.4). We put

$$\psi_l^\Omega(R) = B_1 \exp[iS(R)/\hbar], \quad (4.6)$$

which on substitution into (4.2) gives

$$i\hbar\frac{d^2S}{dR^2} - \left(\frac{dS}{dR}\right)^2 + \hbar^2k^2 = 0. \quad (4.7)$$

We substitute an expansion of S in powers of $\hbar$ into Eq. (4.7) and equate coefficients of equal powers of $\hbar$:

$$S = S_0 + \hbar S_1 + \cdots.$$

We find that

$$\left(\frac{d^2S_0}{dR}\right)^2 + \hbar^2k^2 = 0,$$
$$i\frac{d^2S_0}{dR^2} - 2\frac{dS_0}{dR}\frac{dS_1}{dR} = 0, \quad \text{etc.}$$

Integration of these equations gives

$$S_0(R) = \pm\hbar\int k\,dR, \quad S_1(R) = \tfrac{1}{2}i\log k(R),$$

where arbitrary constants of integration that can be absorbed in the coefficient B_1 have been omitted. We thus obtain to this order of approximation,

$$\psi_l^\Omega(R) = B_1 k^{-1/2}\exp\left[\pm i\int k(R)\,dR\right]. \quad (4.8)$$

In similar fashion, the approximate solution of (4.4) is

$$\psi_l^\Omega(R) = B_2 k'^{-1/2}\exp\left[\pm i\int k'(R)\,dR\right]. \quad (4.9)$$

The probability density is defined as

$$P = \psi_l^\Omega(R)\psi_l^{*\Omega}(R), \quad (4.10)$$

where $\psi_l^{*\Omega}(R)$ is the complex conjugate of $\psi_l^\Omega(R)$. The probability density corresponding to solutions (4.8) and (4.9) are

$$P_1 = B_1^2 \Big/ \left[\frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} + \frac{\mu^2c^4}{\hbar^2(c^2 - 2mG/R)} - \frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)}\right] \quad (4.11)$$

and

$$P_2 = B_2^2 \Big/ \left[\frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)} - \frac{c^2 \Omega^2}{(c^2 - 2mG/R)^2} - \frac{\mu^2 c^4}{\hbar^2 (c^2 - 2mG/R)} \right], \tag{4.12}$$

respectively.

Equations (4.11) and (4.12) show that P_1 decreases with R, whereas P_2 increases. This means that according to Eq. (4.11), the probability of finding a tachyon decreases as R increases, but according to Eq. (4.12) this probability increases with R. In the preceding section, we find that the potential of a tachyon increases with R. The physical laws tell us that a particle wants to be in a lower potential state. Hence, on this ground information given by (4.12) is not correct. Now we are in a position to infer that the differential equation (4.4), which is the second form of the differential equation (4.1) under the condition $k'^2 > 0$ does not suit the problem considered here. Hence the differential equation (4.4) and its solution (4.9) is discarded.

5. TEST OF VALIDITY OF WKB SOLUTION

The accuracy of the WKB solution can be tested by comparing the magnitude of the successive terms S_0 and $\hbar S_1$ in the series of S. The ratio $|\hbar S_1/S_0|$ is expected to be small if $|(\hbar(dS_1/dR))/(dS_0/dR)|$ is small.

$$\left| \frac{\hbar(dS_1/dR)}{dS_0/dR} \right| = \left| \frac{dk/dR}{2k^2} \right| \leqslant \left| \frac{2mGc^2\Omega^2}{R^2(c^2 - 2mG/R)^3 k^3} \right| + \left| \frac{-2\mu^2 c^4 mG}{\hbar^2 R^2 (c^2 - 2mG/R)^2 k^3} \right| + \left| \frac{c^2 l(l+1)(Rc^2 - mG)}{R^2(c^2 - 2mG/R)^2 k^3} \right|$$

$$\leqslant \frac{1}{k^3} \left[\frac{2mGc^2\Omega^2}{R^2(c^2 - 2mG/R)^3} + \frac{2\mu^2 c^4 mG}{\hbar^2 R^2 (c^2 - 2mG/R)^2} + \frac{c^2 l(l+1)(Rc^2 - mG)}{R^2(c^2 - 2mG/R)^2} \right]. \tag{5.1}$$

In Eq. (5.1) the term

$$\left[\frac{2mGc^2\Omega^2}{R^2(c^2 - 2mG/R)^3} + \frac{2\mu^2 c^4 mG}{\hbar^2 R^2 (c^2 - 2mG/R)^2} + \frac{c^2 l(l+1)(Rc^2 - mG)}{R^2(c^2 - 2mG/R)^2} \right]$$

decreases with R. Hence we have

$$\left| \frac{\hbar(dS_1/dR)}{dS_0/dR} \right| \ll 1.$$

This shows that the WKB solution (4.8) is valid.

In Eq. (4.5) if

$$\frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)} > \frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} + \frac{\mu^2 c^4}{\hbar^2 (c^2 - 2mG/R)}$$

then k' will be a decreasing function with R. Under these circumstances, we can not find

$$\left| \frac{\hbar(dS_1/dR)}{dS_0/dR} \right| \ll 1.$$

Hence, if

$$\frac{c^2 l(l+1)}{R^2(c^2 - 2mG/R)} > \frac{c^2\Omega^2}{(c^2 - 2mG/R)^2} + \frac{\mu^2 c^4}{\hbar^2 (c^2 - 2mG/R)}, \tag{5.2}$$

the WKB solution (4.9) is not valid.

It is under this condition (5.2) that the solution (4.9) does not suit the problem undertaken in this paper, as it has been discussed in the preceding section.

6. ENERGY OF TACHYON IN PTBBS

The partial differental equation (2.11) can be rewritten as

$$\frac{\partial^2 \psi_l}{\partial t^2} - c^2\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial^2 \psi_l}{2R^2} - \frac{2c^2}{R^3}\left(R - \frac{2mG}{c^2}\right)\left(R - \frac{mG}{c^2}\right)\frac{\partial \psi_l}{\partial R} + \frac{c^2 l(l+1)}{R^2}\psi_l - \frac{\mu^2 c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right)\psi_l = 0. \tag{6.1}$$

As in Sec. 4, here also the coefficient of $\partial\psi_l/\partial R$ can be neglected except in the shell $|R - qmG/c^2| < \epsilon$. Hence, we have

$$\frac{\partial^2 \psi_l}{\partial t^2} - c^2\left(1 - \frac{2mG}{Rc^2}\right)\frac{\partial^2 \psi_l}{\partial R^2} + \frac{c^2 l(l+1)}{R^2}\psi_l - \frac{\mu^2 c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right)\psi_l = 0. \tag{6.2}$$

Now suppose that

$$\psi_l(R,t) = \int_{-\infty}^{\infty} C(P)\psi_l^p(t)\exp(ipR)\,dp. \tag{6.3}$$

Substituting $\psi_l(R,t)$ from Eq. (6.3) into Eq. (6.2) we get an ordinary differential equation

$$\frac{d^2\psi_l^p(t)}{dt^2} + \left[c^2p^2\left(1 - \frac{2mG}{Rc^2}\right)^2 + \frac{c^2 l(l+1)}{R^2} - \frac{\mu^2 c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right) \right]\psi_l^p(t) = 0. \tag{6.4}$$

This differential equation is integrated into

$$\psi_l^p(t) = D\exp\left[\pm i\left\{ c^2p^2\left(1 - \frac{2mG}{Rc^2}\right)^2 + \frac{c^2 l(l+1)}{R^2} - \frac{\mu^2 c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right) \right\}^{1/2} t \right]. \tag{6.5}$$

Now the magnitude of the energy of the tachyon associated with this tachyon wave ψ_l^p is given by

$$E = \left| \left\{ c^2p^2\left(1 - \frac{2mG}{Rc^2}\right)^2 + \frac{c^2 l(l+1)}{R^2} - \frac{\mu^2c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right)\right\}^{1/2} t \right|, \tag{6.6}$$

which yields

$$E = c^2p^2\left(1 - \frac{2mG}{Rc^2}\right)^2 + \frac{c^2 l(l+1)}{R^2} - \frac{\mu^2c^4}{\hbar^2}\left(1 - \frac{2mG}{Rc^2}\right). \tag{6.7}$$

We know that in the case of a tachyon

$$E^2 - c^2p^2 = -\mu^2c^4. \tag{6.8}$$

Connecting Eqs. (6.7) and (6.8) we have

$$E^2 = -\frac{\mu^2c^4(R - 2mG/c^2)}{2(R - mG/c^2)} - \frac{\mu^2c^6R(R - 2mG/c^2)}{4mG\hbar^2(R - mG/c^2)} + \frac{c^4 l(l+1)}{4mG(R - mG/c^2)}. \tag{6.9}$$

Substituting values of m, G, $\hbar$, μ, and l we see that

$$\mu^2c^6R(R - 2mG/c^2) \gg \frac{\mu^2c^4(R - 2mG/c^2)}{2(R - mG/c^2)}$$

for any value of $R > 2mG/c^2$. Hence, the first term on the right-hand side of Eq. (6.9) is negligible. Now,

$$E^2 = \frac{c^6}{4mG(Rc^2 - mG)}\left[l(l+1) - \frac{\mu^2R(Rc^2 - 2mG)}{\hbar^2}\right]. \tag{6.10}$$

This expression yields

$$E = \left| \left[\frac{c^6}{4mG(Rc^2 - mG)}\left\{ l(l+1) - \frac{\mu^2R(Rc^2 - 2mG)}{\hbar^2}\right\}\right]^{1/2}\right|. \tag{6.11}$$

This expression shows that as R increases, the energy of the tachyon decreases.

But from Eq. (6.11) we also find that if for some value of

$$l(l+1) < \mu^2R(Rc^2 - 2mG)/\hbar^2, \tag{6.12}$$

the energy of a tachyon will increase with R.

From above analysis we note that if condition (6.12) is not satisfied, a tachyon would emit energy on ascending to higher orbits and would gain energy on descending to lower orbits, contrary to Bohr's hydrogen atom. But if the condition (6.12) is satisfied, a tachyon would gain energy on ascending to higher orbits and would emit energy on descending to lower orbits like the electron in the hydrogen atom.

7. COMPARISON OF PTBBS WITH BOHR'S HYDROGEN ATOM

Now we are in a position to compare and contrast PTBBS with Bohr's model of the hydrogen atom as follows:

(1) In an H atom, an electron revolves around a proton under the influence of the Coulomb force and centrifugal force. In PTBBS, we find that a tachyon revolves around a proton (bradyon) in a circular orbit under the influence of gravitational force of attraction and centrifugal force.

(2) In an H atom, there acts a short range repulsive force between the electron and proton due to which, the electron does not fall into proton. But, in PTBBS, there is a possibility of falling down of the tachyon into the proton as has been discussed in Sec. 3, if the tachyon comes very close to proton (bradyon).

(3) In an H atom, beyond a short range (where the force between the electron and proton is repulsive), the force between electron and proton is attractive. In PTBBS, the force between the tachyon and proton (bradyon) is attractive, but beyond a fixed range (where potential of tachyon is maximum) force between tachyon and proton is repulsive.

(4) In an H atom, when an electron ascends to higher orbits it gains energy and on descending to lower orbits it emits energy. In PTBBS, two situations happen: (a) if condition (6.12) is satisfied the tachyon also, like the electron in the H atom, emits energy on descending to lower orbits and gains energy on ascending to higher orbits; (b) but if condition (6.12) is not satisfied, the situation is reversed. In this case, tachyon looses energy on ascending to higher orbits and gains energy on descending to lower orbits. Hence in situation (a) to knock out the tachyon from PTBBS we will have to supply energy as energy is supplied to knock out the electron from an H atom. But in situation (b), to knock out the tachyon from PTBBS, we would have to extract energy from it.

8. POSSIBILITY OF PTBBS

We have discussed above various aspects of PTBBS. However, one thing is very important to know. Can such systems exist in nature or is there any possibility of formation of these systems in the future or could these systems have formed in the extreme past? In this section, we speculate that such systems would have formed in the past on or after the epoch of the big bang.

According to cosmologists, the big bang was the extreme phenomenon of the universe and many fundamental particles were produced during the epoch of the big bang. Narlikar and Sudarshan[2] have assumed that tachyons also would have been produced on or after the epoch of the big bang along with many other fundamental particles of ordinary matter (bradyonic matter). It is the view of both cosmologists and particle physicists[8] that when this universe was 3 minutes old, nucleosynthesis of many atoms such as He, D, Li, etc. took place from primordial neutrons, protons, and electrons. In Sec. 3 as well as the author's work in Ref. 9 we find that a tachyon is attracted by a bradyon when they are at a short distance. Now if primordial tachyons were produced, there is the possibility of formation of PTBBS-like systems from primordial protons and tachyons in the early stages of the universe, because it is expected that at that time tachyons and protons would have been sufficiently close to each other.

ACKNOWLEDGMENTS

I am extremely grateful to the referee for useful comments and helpful suggestions regarding revision of the original manuscript. I also thank Dr. V. B. Johri for stimulating discussions.

[1]H. C. Corben, "Electromagnetic and hadronic properties of tachyons," *Proceedings of Tachyons, Monopoles and Related Topics*, edited by E. Re-

cami (North-Holland, Amsterdam, 1978).
[2]J. V. Narlikar and E. C. G. Sudarshan, Mon. Not. R. Astron. Soc. **175**, 105–16 (1976).
[3]S. V. Dhurandhar, Pramana **8**, 133 (1977).
[4]E. Honig, K. Lake, and R. C. Roeder, Phys. Rev. D **10**, 3155 (1974).
[5]J. V. Narlikar and S. V. Dhurandhar, Pramana **6**, 388 (1976).
[6]O. M. P. Bilanuik, V. K. Deshpande, and E. C. G. Sudarshan, Am. J. Phys. **30**, 718 (1962).
[7]C. Sivaram and K. P. Sinha, Phys. Rep. **51** (3), 111–187 (1979).
[8]M. S. Turner and D. N. Schramm, Phys. Today **3** (9), 42–48 (1979).
[9]S. K. Srivastava, Ph.D. thesis, Gorakhpur University, Gorakhpur (1978).

Production of tachyons from Schwarzschild black holes

Sushil K. Srivastava
Government College, Port Blair 744104, India

(Received 10 November 1981; accepted for publication 14 May 1982)

Inside the Schwarzschild surface the line element becomes spacelike. Hence, in the present paper, the Schwarzschild black hole is considered as a spherical world of tachyons. It is assumed that tachyons are present in this world in a superdense state. In this framework a spacelike line element for this world is derived. Moreover, the scalar wavefunction of a tachyon is obtained which satisfies the Klein–Gordon equation. Expressions for transmission and reflection coefficients of a scalar tachyon wave from a Schwarzschild surface are obtained. It is found that transmission and reflection coefficients are equal. Expressions for energy and momentum of a tachyon, emitted from a Schwarzschild black hole, are also obtained.

PACS numbers: 14.80.Pb, 04.20. – q, 03.65. – w

1. INTRODUCTION

In spite of the failure of experiments to detect and produce tachyons in the laboratory, there has been continuing interest in research on tachyons. Regarding production of these particles many hypothesis have been developed. In 1976, Narlikar and Sudarshan[1] suggested that tachyons might have been produced at or after the epoch of the big bang alongwith elementary particles of bradyonic matter provided that the universe was created in a big bang. In 1979, Cole[2] considered emission of tachyons from hadrons (protons) using recently developed six-dimensional relativity. In 1980, Recami and Maccarrone[3] have considered emission of tachyons from gravitational black holes as well as hadrons and have investigated effects of emission of tachyons on them. By the end of the sixties some papers, authored by some particle physicists such as Arons and Sudarshan,[4] Feinberg,[5] and Dhar and Sudarshan,[6] appeared which have relevance to production and detection of these superluminal particles.

In this paper, we consider production of tachyons from Schwarzschild black holes. The Schwarzschild line element is given as

$$ds^2 = (1 - 2GM/c^2r)c^2\,dt^2 - (1 - 2GM/c^2r)^{-1}\,dr^2 - r^2(d\theta^2 + \sin\theta\, d\varphi^2), \tag{1.1}$$

where G is the Newtonian gravitational constant, M is the mass of the black hole, and c is the speed of light.

$r = r_S = 2GM/c^2$ gives the radius of the black hole which is usually called the Schwarzschild radius. In case $r > r_S$ the world lines of a particle are timelike, i.e., $ds^2 > 0$. But if $\theta = \theta_0$ (constant), $\phi = \phi_0$ (constant), and $r < r_S$, we find that the world lines of a test particle become spacelike, i.e., $ds^2 < 0$. From this observation, we are in a position to note that the speed of particles, inside the surface $r = r_S$ is greater than c. Hence, we find that $r = r_S$ which gives the equation of a spherical world of tachyons in polar coordinates.

In Sec. 2, we derive a line element for this spherical world of tachyons using Einstein's field equations. We assume that tachyons are present inside the surface $r = r_S$ in a superdense state. Hence we use the equation of state $p = \rho$ for $r < r_S$ where p is the pressure and ρ is the energy density of the tachyonic perfect fluid.

In Sec. 3 the Klein–Gordon equation for a free spinless tachyon has been given, which reduces to an ordinary differential equation on substituting

$$\psi(r,t) = \Phi(r)\exp(-ik_1t_x - ik_2t_y - ik_3t_z),$$

where $\psi(r,t)$ is the state function of a spinless tachyon, and $\Phi(r) = (1/r)f(r)$.

In Sec. 4, we have solved the ordinary differential equation by the method of Wentzel–Kramers–Brillouin (hereafter called WKB) approximation.

In Sec. 5, we test the validity of WKB solutions obtained in Sec. 4.

In Sec. 6, we have calculated the transmission and reflection coefficients from the barrier $r = r_S$. It is found that a tachyon wave is half transmitted and half reflected from the Schwarzschild surface.

In Sec. 7, we have got expressions for energy and momentum of tachyon emitted from Schwarzschild black holes.

In the last section, we remark that at $\sigma = \sigma_1$ and $t = t_1$ the energy of the emitted tachyon may vanish.

Throughout the entire paper, we choose $\hbar = 1$ where $\hbar$ is the Planck's constant divided by 2π.

2. SPACELIKE LINE ELEMENT INSIDE THE SCHWARZSCHILD SURFACE

In 1976, Mignani and Recami[7] proposed that in the case of tachyons, time is a vector quantity and all the three components of the time vector (along the x, y, and z axes) will be measurable, whereas the position vector components loose their individual physical meaning. In other words, we can say that for a tachyon only the quantity $|r|^2 = x^2 + y^2 + z^2$ should be observable. Hence, we assume the following line element:

$$ds^2 = e^{\mu}dr^2 - e^{\nu}c^2(dt_x^2 + dt_y^2 + dt_z^2) \tag{2.1}$$

for the spherical world of tachyons given by the surface $r = r_S$. Here μ and ν are functions of r and t: t_x, t_y, and t_z are three components of time t.

Further we assume that $\tau = ct$, i.e., $\tau_1 = ct_x$, $\tau_2 = ct_y$, and $\tau_3 = ct_z$. Hence the line element (2.1) is written as

$$ds^2 = e^{\mu}dr^2 - e^{\nu}(d\tau_1^2 + d\tau_2^2 + d\tau_3^2), \tag{2.2}$$

 0022-2488/83/051317-04$02.50

where $\mu \equiv \mu(r,\tau)$ and $\nu = \nu(r,\tau)$.

Now we choose a timelike observer's coordinate system such that the four velocities are given by

$$u^\mu \equiv (1,0,0,0). \tag{2.3}$$

We also assume isotropic development of time in every direction; hence we have

$$\tau_1 = \tau_2 = \tau_3 = \tau/\sqrt{3}. \tag{2.4}$$

The Einstein field equations are given as

$$R_{ij} - \tfrac{1}{2}Rg_{ij} + \Lambda g_{ij} = -(8\pi G/c^2)T_{ij}, \tag{2.5}$$

where $T^i_j = (\rho + p)u^i u_j + p\delta^i_j$, R_{ij} is the Ricci rotation tensor; g_{ij} is the metric tensor given by the line element (2.2); and Λ is the cosmological constant.

In the above framework, the Einstein field Eqs. (2.5) are computed as

$$\tfrac{3}{4}e^{-\nu}\dot{\nu}^2 + 3e^{\nu}\ddot{\nu} - \tfrac{3}{4}e^{-\mu}\nu'^2 - e^{-\mu}\nu'' + \tfrac{3}{4}e^{-\nu}\dot{\mu}^2 + 3e^{-\nu}\ddot{\mu} + \tfrac{1}{2}e^{-\mu}\mu'\nu' + \tfrac{3}{2}e^{-\nu}\dot{\mu}\nu = -(8\pi G/c^2)p, \tag{2.6a}$$

$$\tfrac{9}{2}e^{-\nu}\ddot{\nu} - \tfrac{9}{4}e^{-\nu}\dot{\nu}^2 - \tfrac{3}{4}e^{-\mu}\nu'^2 = -(8\pi G/c^2)(\rho + 2p), \tag{2.6b}$$

$$\tfrac{3}{4}\dot{\nu}^2 - \tfrac{3}{4}\dot{\mu}^2 - \tfrac{3}{2}\ddot{\nu} - \tfrac{3}{2}\ddot{\mu} + \tfrac{3}{2}\dot{\mu}\dot{\nu} = 0, \tag{2.6c}$$

$$\dot{\nu}' - \dot{\nu}\nu' - \dot{\mu}\nu' = 0, \tag{2.6d}$$

where dot denotes partial differentiation with respect to τ and prime denotes partial differentiation with respect to r.

On integrating Eq. (2.6d) partially with respect to τ we have

$$\nu' = (2R'/R)e^{(\mu+\nu)}, \tag{2.7}$$

where R is function of r only.

Equation (2.6c) yields

$$\mu + \nu = 0. \tag{2.8}$$

Connecting Eqs. (2.7) and (2.8) we have

$$\nu' = 2R'/R. \tag{2.9}$$

On integrating the partial differential equation (2.9) we get

$$\nu = \chi(\tau) + \log R^2. \tag{2.10}$$

Using the equation of state $p = \rho$ we have from Eqs. (2.6a) and (2.6b),

$$\tfrac{3}{2}e^{-\nu}\dot{\nu}^2 - \tfrac{1}{2}e^{-\mu}\nu'^2 - e^{-\mu}\nu'' + \tfrac{3}{2}e^{-\nu}\mu'^2 + 3e^{-\nu}\ddot{\mu} + \tfrac{1}{2}e^{-\mu}\mu'\nu' + \tfrac{3}{2}e^{-\nu}\dot{\mu}\dot{\nu} = 0. \tag{2.11}$$

Connecting Eqs. (2.8) and (2.11) we have

$$e^{-\nu}\nu'^2 - \tfrac{3}{2}e^{-\nu}\dot{\nu}^2 + e^{\nu}\nu'' + 3e^{-\nu}\ddot{\nu} = 0. \tag{2.12}$$

Substitution of ν from Eq. (2.10) into Eq. (2.12) yields two ordinary differential equations

$$\frac{d^2\chi}{d\tau^2} - \frac{1}{2}\left(\frac{d\chi}{d\tau}\right)^2 = 0 \tag{2.13}$$

and

$$R\frac{d^2R}{dr^2} + \left(\frac{dR}{dr}\right)^2 = 0. \tag{2.14}$$

Integration of ordinary differential equation (2.13) yields the solution

$$\chi(\tau) = B - \log(\tau + A)^2.$$

Here A and B are integration constants, hence, on physical grounds there is no harm in putting $A = 1$ and $B = 0$. Thus we have

$$\chi(\tau) = -\log(\tau + 1)^2. \tag{2.15}$$

Integration of the ordinary differential equation (2.14) yields the solution

$$R^2 = 2Cr + D.$$

Here also C and D are integration constants. Hence for our convenience we set $C = 1$ and $D = 0$. Thus we have

$$R^2 = 2r. \tag{2.16}$$

Substituting χ and R^2 from Eqs. (2.15) and (2.16) into Eq. (2.10) we get

$$\nu = \log 2r - \log(\tau + 1)^2. \tag{2.17}$$

From Eqs. (2.8) and (2.17) we have

$$\mu = \log(\tau + 1)^2 - \log 2r. \tag{2.18}$$

Substituting ν and μ from Eqs. (2.17) and (2.18) into the line element (2.2) we have

$$ds^2 = \frac{(\tau+1)^2}{2r}dr^2 - \frac{2r}{(\tau+1)^2}(d\tau_1^2 + d\tau_2^2 + d\tau_3^2). \tag{2.19}$$

Again putting $\tau = ct$ we have the line element (2.19) in the form

$$ds^2 = \frac{(1+ct)^2}{2r}dr^2 - \frac{2rc^2}{(1+ct)^2}(dt_x^2 + dt_y^2 + dt_z^2). \tag{2.20}$$

This is the required line element, where $t^2 = t_x^2 + t_y^2 + t_z^2$.

3. THE KLEIN–GORDON EQUATION INSIDE THE SCHWARZSCHILD SURFACE

Here we consider the spinless tachyon. The spacelike scalar wavefunction $\psi(r,t_x,t_y,t_z)$ satisfies the Klein–Gordon equation

$$(\Box^2 + m^2)\psi = 0, \tag{3.1}$$

where m is the metamass of the tachyon. $\Box^2$ is defined as

$$\Box^2 = (-g)^{-1/2}\frac{\partial}{\partial x^i}\left[(-g)^{1/2}g^{ij}\frac{\partial}{\partial x^j}\right], \tag{3.2}$$

where g_{ij} is given by the metric (2.20) and $x^i \equiv (r,t_x,t_y,t_z)$. From the line element (2.20) we have

$$g = -4r^2c^6/(1+ct)^4$$

and

$$g^{11} = 2r/(1+ct)^2, \quad g^{22} = g^{33} = g^{44} = -(1+ct)^2/2rc^2.$$

Now the Klein–Gordon equation (3.1) is written as

$$\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} - \frac{(1+ct)^4}{4c^2r^2}\left[\frac{\partial^2}{\partial t_x^2} + \frac{\partial^2}{\partial t_y^2} + \frac{\partial^2}{\partial t_z^2}\right]\psi + \frac{m^2(1+ct)^2}{2r}\psi = 0. \tag{3.3}$$

We consider the plane wave solution of Eq. (3.3) as

$$\psi(r,t) = \Phi(r)\exp(-ik_1t_x - ik_2t_y - ik_3t_z), \tag{3.4}$$

where k_1,k_2,k_3 are constants and $k^2 = k_1^2 + k_2^2 + k_3^2$. Substitution of $\psi(r,t)$ from Eq. (3.4) into the partial differential equation (3.3) yields the ordinary differential equation

$$\frac{d^2\Phi(r)}{dr^2}+\frac{2}{r}\frac{d\Phi(r)}{dr}+\left[\frac{(1+ct)^4k^2}{4c^2r^2}+\frac{m^2(1+ct)^2}{2r}\right]\Phi(r)=0. \tag{3.5}$$

Substituting $f(r)=r\Phi(r)$ into Eq. (3.5) we get

$$\frac{d^2f}{dr^2}+\left[\frac{(1+ct)^4k^2}{4c^2r^2}+\frac{m^2(1+ct)^2}{2r}\right]f=0 \tag{3.6}$$

for some value of t.

4. WKB SOLUTIONS

We approximate the coefficient of $f(r)$ in the ordinary differential equation (3.6) near $r=r_S$ as

$$\frac{(1+ct)^4k^2}{4c^2r^2}+\frac{m^2(1+ct)^2}{2r}=\left[\frac{(1+ct)^4k^2}{4c^2r_S^2}+\frac{m^2(1+ct)^2}{2r_S}\right]-\left[\frac{(1+ct)^4k^2}{2c^2r_S^3}+\frac{m^2(1+ct)^2}{2r_S^2}\right](r-r_S). \tag{4.1}$$

Hence, the ordinary differential equation (3.6) is written as

$$\frac{d^2f}{dr^2}+\eta^2f=0 \quad \text{for } \eta^2>0, \tag{4.2}$$

where

$$\eta^2=\left[\frac{(1+ct)^4k^2}{4c^2r_S^2}+\frac{m^2(1+ct)^2}{2r_S}\right]-\left[\frac{(1+ct)^4}{2c^2r_S^3}k^2+\frac{m^2(1+ct)^2}{2r_S^2}\right](r-r_S). \tag{4.3}$$

The WKB solution of the Eq. (4.2) is

$$f(r)=4\eta^{-1/2}\exp\left(\pm i\int\eta\,dr\right). \tag{4.4}$$

If we shift the initial point from $r=0$ to $r=r_S$ the ordinary differential equation (3.6) is written as

$$\frac{d^2f}{d\sigma^2}+\left[\frac{(1+ct)^4k^2}{4c^2(\sigma+r_S)^2}+\frac{m^2(1+ct)^2}{2(\sigma+r_S)}\right]f=0. \tag{4.5}$$

The WKB solution of this differential is

$$f=B\eta'^{-1/2}\exp\left(\pm i\int\eta'\,d\sigma\right), \tag{4.6}$$

where

$$\eta'^2=\frac{(1+ct)^4k^2}{4c^2(\sigma+r_S)^2}+\frac{m^2(1+ct)^2}{2(\sigma+r_S)}. \tag{4.7}$$

Hence, we consider a black hole of one solar mass. Hence, substitution of

$$G=6.67\times10^{-8},\quad M=M_0=9.2\times10^{33},\quad c=3\times10^{10}$$

(all values in cgs units) yields

$$r_S=1.36\times10^6\ \text{cm}. \tag{4.8}$$

5. TEST OF VALIDITY OF WKB SOLUTIONS

For the validity of WKB solutions obtained above, we test where $|d\eta/dr/2\eta^2|\ll1$. For the solution (4.4) we get

$$\left|\frac{d\eta/dr}{2\eta^2}\right|=\left|\frac{(1+ct)^4/2r_S^3c^2k^2+m^2(1+ct)^2/2r_S^2}{4[\{(1+ct)^4/4c^2r_S^2k^2+m^2(1+ct)^2/2r_S\}-\{(1+ct)^4/2r_S^3c^2k^2+m^2(1+ct)^2/2r_S^2\}(r-r_S]^{3/2}}\right|$$
$$\approx\left|\frac{2^{3/2}r_S}{8m(1+ct)[2r_S-r]^{3/2}}\right|$$
$$\approx\left|\frac{2^{3/2}r_S^{-1/2}}{8m(1+ct)}\left[1+\frac{3}{2}\frac{a}{r_S}\right]\right| \quad \text{(where a is very small)}$$
$$\approx\left|\frac{1}{2^{3/2}(1+ct)m}\left[\frac{1}{r_S^{1/2}}+\frac{3}{2}\frac{a}{r_S^{3/2}}\right]\right|\ll1.$$

Similarly for solution (4.6) we have

$$\left|\frac{d\eta'/d\sigma}{2\eta'^2}\right|=\left|\frac{(1+ct)^4k^2/2c^2(\sigma+r_S)^3+m^2(1+ct)^2/2(\sigma+r_S)^2}{[(1+ct)^4k^2/4c^2(\sigma+r_S)^2+m^2(1+ct)^2/2(\sigma+r_S)]^{3/2}}\right|.$$

This shows that as σ is large $|(d\eta'/d\sigma)/2\eta'^2|\ll1$.

6. TRANSMISSION AND REFLECTION COEFFICIENTS FROM THE BARRIER

The incident wave on the barrier $r=r_S$ is

$$\Phi(r)=\frac{A}{r}\eta^{-1/2}\exp\left(+i\int\eta\,dr\right). \tag{6.1}$$

The transmitted as well as the reflected waves from the barrier $r=r_S$ are

$$\Phi(r)=\frac{B}{(\sigma+r_S)}\eta'^{-1/2}\exp\left(+i\int\eta'\,d\sigma\right) \tag{6.2}$$

and

$$\Phi(r)=\frac{B}{(\sigma+r_S)}\eta'^{-1/2}\exp\left(-i\int\eta'\,d\sigma\right), \tag{6.3}$$

respectively.

The wave equations (6.1), (6.2), and (6.3) are written as

$$\Phi(r)=(A/r)\eta^{-1/2}\xi_1,\quad \xi_1=\exp\left(+i\int\eta\,dr\right), \tag{6.4}$$

$$\Phi(r) = \frac{B}{(\sigma + r_S)} \eta'^{-1/2} \xi_2, \quad \xi_2 = \exp\left(+ i \int \eta' \, d\sigma \right), \quad (6.5)$$

$$\Phi(r) = \frac{B}{(\sigma + r_S)} \eta'^{-1/2} \xi_3, \quad \xi_3 = \exp\left(- i \int \eta' \, d\sigma \right), \quad (6.6)$$

respectively.

The probability current density is defined as

$$\mathrm{j} = \frac{i}{2m}\left[\xi \frac{\partial}{\partial r} \xi^* - \xi^* \frac{\partial \xi}{\partial r} \right],$$

where ξ^* is the complex conjugate of ξ. The transmission coefficient

$$D = \frac{\text{Probability current density transmitted wave}}{\text{Probability current density incident wave}}$$

$$= \left| \frac{\eta'}{\eta} \right|$$

$$\simeq \left| \frac{r_S}{(\sigma + r_S)} \left[\frac{(1 + ct)^2 k^2 + 2m^2c^2(\sigma + r_S)}{(1 + ct)^2 + 2m^2 r_S - 2m^2c^2(r - r_S)} \right]^{1/2} \right|. \quad (6.7)$$

Putting $r = r_S + a$ (where a is very small) in Eq. (6.7) we have

$$D = \left| \frac{r_S}{(\sigma + r_S)} \left[\frac{(1 + ct)^2 k^2 + 2m^2c^2(\sigma + r_S)}{(1 + ct)^2 + 2m^2 r_S} \right]^{1/2} \times \left[1 + \frac{m^2c^2a}{(1 + ct)^2 + 2m^2 r_S} \right] \right|. \quad (6.8)$$

Similarly we have the reflection coefficient

$$R = \frac{\text{Probability current density reflected wave}}{\text{Probability current density incident wave}}$$

$$= \left| \frac{\eta'}{\eta} \right|.$$

Thus we have $R = D$. Also $R = 1 - D$. This shows that the incident tachyon wave on the barrier $r = r_S$ is half reflected and half transmitted through the Schwarzschild surface. In other words, the Schwarzschild surface is partially transparent for tachyons.

7. MOMENTUM AND ENERGY OF THE TACHYON BEYOND THE SCHWARZSCHILD SURFACE

From the above investigations we find that the wave associated with the tachyon is half reflected and half transmitted through the Schwarzschild surface. It means that the tachyon is produced from black holes. Therefore, the transmitted wave may be considered the wave associated with the spinless tachyons produced from the black hole. The momentum of this tachyon is calculated as

$$\rho' = \left| \frac{d}{d\sigma} \left[\int \eta' \, d\sigma \right] \right|$$

$$= \left| \frac{(1 + ct)}{2c} \left[\frac{(1 + ct)^2 k^2}{(\sigma + r_S)^2} + \frac{2m^2c^2}{(\sigma + r_S)} \right]^{1/2} \right|. \quad (7.1)$$

This shows that the momentum of the produced tachyon decreases as σ increases.

From the equation

$$c^2 \rho'^2 - E^2 = m^2c^4 \quad (7.2)$$

for tachyons, we calculate the energy of tachyon

$$E = \epsilon \left[\frac{(1 + ct)^2}{4} \left\{ \frac{(1 + ct)^2 r^2}{(\sigma + r_S)^2} + \frac{2m^2c^2}{(\sigma + r_S)} \right\} - m^2c^4 \right]^{1/2}, \quad (7.3)$$

where $\epsilon \pm 1$ ($\epsilon = +1$ corresponds to tachyons, whereas $\epsilon = -1$ corresponds to antitachyons).

8. CONCLUDING REMARKS

From the above investigations, we find that tachyons would be produced from a Schwarzschild black hole. From Eq. (7.3) we also note that at the distance $\sigma = \sigma_1$, from the Schwarzschild surface and at time

$$t_1 = -\frac{1}{c} \left[\frac{(\sigma_1 + r_S) mc}{k^2} \{ (m^2 + 4k^2)^{1/2} - m \} \right]^{1/2}$$

the energy of the produced tachyon will vanish.

[1] J. V. Narlikar and E. C. G. Sudarshan, Mon. Not. R. Astron. Soc. **175**, 105 (1976).
[2] E. A. B. Cole, Phys. Lett. A **75**, 29 (1979).
[3] E. Recami and G. D. Maccarrone, Lett. Nuovo Cimento **28**, 151 (1980).
[4] M. E. Arons and E. C. G. Sudarshan, Phys. Rev. **173**, 1622 (1968).
[5] G. Feinberg, Phys. Rev. **159**, 1089 (1967).
[6] J. Dhar and E. C. G. Sudarshan, Phys. Rev. **174**, 1808 (1968).
[7] R. Mignani and E. Recami, Lett. Nuovo Cimento **16**, 449 (1976).

Stability analysis of the neutron transport equation with temperature feedback

C. V. Pao
Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27650

(Received 12 January 1982; accepted for publication 9 April 1982)

This paper discusses the stability and instability of neutron transport in a reactor system where positive temperature feedback is taken into consideration. The feedback effect is considered only through the multiplication factor in the neutron transport equation. This consideration leads to a coupled system of nonlinear partial integrodifferential equations of neutron transport and heat conduction. The investigation of this coupled nonlinear initial boundary-value problem is based on the existence-comparison theorem established in an earlier paper in which an iterative scheme for the determination of the solution is given. It is shown in this paper that by constructing a suitable pair of the comparison functions, called upper and lower solutions, as two distinct initial iterations, the corresponding sequences converge monotonically from above and below, respectively, to a unique solution of the system. Under certain conditions on the multiplication factor, the property of upper and lower solutions determines the asymptotic behavior of the solution and leads to stability or instability of the system. A stability region for the zero steady state and an instability region of the system are also given.

PACS numbers: 28.20.He, 28.40. – f, 02.30.Jr

1. INTRODUCTION

The stability problem of neutron transport with temperature feedback is one of the important concerns in reactor dynamics. When the spatial effect for the temperature distribution is taken into consideration the time-dependent neutron transport equation is suplemented by a temperature equation which is a partial differential equation of parabolic type. This consideration leads to a coupled system of nonlinear partial integrodifferential equations. Suppose that the transport medium is in slab geometry with length l and that the temperature feedback is only through the multiplication factor γ. Then the equations governing the neutron density $N \equiv N(t, x, \mu)$ and temperature differential $u(t, x) \equiv T(t, x) - T_c$ are given by (cf. Refs. 1 and 2).

$$v^{-1}N_t + \mu N_x + \sigma N = (\gamma(u)/2)\int_{-1}^{1} N(t, x, \mu')\,d\mu',$$
$$u_t - Du_{xx} + \beta u = (h(u)/2l)\int_{-1}^{1}\int_0^l N(t, x', \mu')\,dx'\,d\mu'$$
$$\times (t > 0, 0 < x < l, -1 \leqslant \mu \leqslant 1), \quad (1.1)$$

where v, σ, D, β are physical constants, T_c is the coolant temperature, and $h(u)$ is the energy generation coefficient. Assume that there is no neutron entering the slab from the slab faces where the temperature is kept at the coolant temperature T_c. Then the boundary condition for (N,u) is given by

$$N(t,0,\mu) = 0 \quad \text{for } 0 < \mu \leqslant 1,$$
$$N(t,l,\mu) = 0 \quad \text{for } -1 \leqslant \mu < 0, \quad (1.2)$$
$$u(t,0) = u(t,l) = 0,$$

where $t > 0$.

The initial condition is in the usual form

$$N(0,x,\mu) = N(x,\mu), \quad u(0,x) = u_0(x) \quad (1.3)$$

where $u_0(x) \equiv T_0(x) - T_c$ and $T_0(x)$ is the initial temperature.

The problem (1.1)–(1.3) forms a coupled system of nonlinear initial boundary-value problem of transport-parabolic type. This sytem has been investigated in a recent paper by Pao[3] where an existence-comparison theorem is established using a monotone iterative scheme and the notion of upper and lower solutions. This existence-comparison theorem is then used to study the global asymptotic stability of the steady state $(0,T_c)$ when γ, h are both uniformly bounded on $[0,\infty)$. Special attention has been given to the model where γ, h are governed by the equations (cf. Refs. 3 and 4).

$$\frac{d\gamma}{du} = a_1(u + T_c)^{-m}, \quad \gamma(0) = \gamma_0 \geqslant 0 \quad (a_1 \geqslant 0),$$
$$\frac{dh}{du} = a_2(u + T_c)^{-n}, \quad h(0) = h_0 \geqslant 0 \quad (a_2 \geqslant 0).$$

It has been shown in Ref. 3 that when $m > 1$, $n > 1$, the steady state $(0,T_c)$ is globally asymptotically stable. Here the definition of stability is in the usual sense of Lyapunov, and the global asymptotic stability is with respect to nonnegative initial perturbations. However, if $\gamma(u)$, $h(u)$ are not uniformly bounded there can be no global stability. In fact, if the heat-conduction term Du_{xx} is neglected then under certain conditions on the physical parameters the solution grows unbounded and its rate of growth is in the order no less than $\exp(\exp \eta t)$ for some $\eta > 0$ (cf. Ref. 5). An interesting question is then under what conditions on the physical parameters and for what class of initial functions the time-dependent solution does converge to a steady state if the effect of the heat conduction term Du_{xx} is taken into consideration. This question involves not only the stability condition but also the stability region of the system. The purpose of this paper is to investigate this question in the framework of the system (1.1)–(1.3) for some general functions $\gamma(u)$, $h(u)$. For this purpose, we make use of the existence-comparison theorem established in Ref. 3 and then construct suitable comparison functions so that the stability or instability property of the system can be determined.

The neutron transport equation with temperature feedback has been investigated by many researchers in the field. For some related work to this problem we refer the reader to the references in the earlier papers in Refs. 3 and 5.

2. MONOTONE SEQUENCES AND EXISTENCE-COMPARISON THEOREM

Let $\Omega_1 = (0,t_1)\times(0,l)$, $\Omega_2 = (0,l)\times[-1,1]$. $Q = (0,t_1]\times(0,l)\times[-1,1]$ and let $\bar{\Omega}_1,\bar{Q}$ be the closure of Ω_1,Q, where $t_1>0$ is finite but can be arbitrarily large. In order to obtain an existence-comparison theorem for the system (1.1)–(1.3) we need the following hypothesis on γ, h.

(H_0) The functions γ, h are positive continuous on $[0,\infty)$ and for each $r>0$ there exist positive constants K_1, K_2, which may depend on r, such that

$$\begin{aligned}&0\leqslant\gamma(u_2)-\gamma(u_1)\leqslant K_1(u_2-u_1),\\ &|h(u_2)-h(u_1)|\leqslant K_2(u_2-u_1),\end{aligned}\tag{2.1}$$

where $0\leqslant u_1\leqslant u_2\leqslant r$. The above hypothesis on γ implies that $\gamma(u)$ is a nondecreasing function of $u\geqslant 0$ which corresponds to positive temperature feedback. The requirement on h is the usual Lipschitz condition on the finite interval $[0,r]$. Under these conditions it is possible to construct two monotone sequences which converge from above and below, respectively, to a unique solution of (1.1)–(1.3). The monotone property of the sequence is based on the notion of upper and lower solutions as defined in the following:

Definition 2.1: A pair of smooth functions $(\tilde{N},\tilde{u})$ is called an upper solution of (1.1)–(1.3) if it satisfies the inequalities

$$\begin{aligned}&v^{-1}\tilde{N}_t+\mu\tilde{N}_x+\sigma\tilde{N}\geqslant(\gamma(\tilde{u})/2)\int_{-1}^{1}\tilde{N}(t,x,\mu')\,d\mu'\\ &[(t,x,\mu)\in Q],\\ &\tilde{u}_t-D\tilde{u}_{xx}+\beta\tilde{u}\geqslant(h(\tilde{u})/2l)\int_{-1}^{1}\int_0^l\tilde{N}(t,x',\mu')\,dx'\,d\mu'\\ &[(t,x)\in\Omega_1],\\ &\tilde{N}(t,0,\mu)\geqslant 0\quad\text{for } 0<\mu\leqslant 1,\\ &\tilde{N}(t,l,\mu)\geqslant 0\quad\text{for } -1\leqslant\mu<0,\\ &\tilde{u}(t,0)\geqslant 0,\quad\tilde{u}(t,l)\geqslant 0\quad[t\in(0,t_1]],\\ &\tilde{N}(0,x,\mu)\geqslant 0,\quad\tilde{u}(0,x)\geqslant 0\quad[(x,\mu)\in\Omega_2].\end{aligned}\tag{2.2}$$

A similar definition holds for a lower solution $(\underset{\sim}{N},\underset{\sim}{u})$ when it satisfies all the reversed inequalities in (2.2).

Here by a smooth function $(\tilde{N},\tilde{u})$ we mean that $(\tilde{N},\tilde{u})$ is continuously differentiable up to the order appearing in (2.2). The same is true for the lower solution $(\underset{\sim}{N},\underset{\sim}{u})$.

Let $(\tilde{N},\tilde{u})$, $(\underset{\sim}{N},\underset{\sim}{u})$ be given upper and lower solutions such that $(\tilde{N},\tilde{u})\geqslant(\underset{\sim}{N},\underset{\sim}{u})\geqslant(0,0)$ (i.e., $\tilde{N}\geqslant\underset{\sim}{N}\geqslant 0,\tilde{u}\geqslant\underset{\sim}{u}\geqslant 0$ pointwise in $\bar{Q}$) and let M be a constant such that

$$M\geqslant(K_2/2l)\int_{-1}^{1}\int_0^l\tilde{N}(t,x',\mu')\,dx'\,d\mu'\quad(t\in[0,t_1]).\tag{2.3}$$

Then by adding the same term Mu on both sides of the second equation in (1.1) and using any continuous function (N^0,u^0) as the initial iteration we can construct a sequence $\{N^{(k)},u^{(k)}\}$ successively from the uncoupled linear systems

$$\begin{aligned}&v^{-1}N_t^{(k)}+\mu N_x^{(k)}+\sigma N^{(k)}\\ &\quad=(\gamma(u^{(k-1)})/2)\int_{-1}^{1}N^{(k-1)}(t,x,\mu')\,d\mu',\\ &\quad N^{(k)}(t,0,\mu)=0\quad\text{for } 0<\mu\leqslant 1,\\ &\quad N^{(k)}(t,l,\mu)=0\quad\text{for } -1\leqslant\mu<0,\\ &\quad N^{(k)}(0,x,\mu)=N_0(x,\mu),\end{aligned}\tag{2.4}$$

$$\begin{aligned}&u_t^{(k)}-Du_{xx}^{(k)}+(\beta+M)u^{(k)}=Mu^{(k-1)},\\ &\quad+(h(u^{(k-1)})/2l)\int_{-1}^{1}\int_0^l N^{(k-1)}(t,x',\mu')\,dx'\,d\mu',\\ &\quad u^{(k)}(t,0)=u^{(k)}(t,l)=0,\\ &\quad u^{(k)}(0,x)=u_0(x)\end{aligned}\tag{2.5}$$

for $k=1,2,\cdots$. Since for each k the above two systems are linear, uncoupled (but interrelated), the sequence $\{N^{(k)},u^{(k)}\}$ is well defined. Clearly the property of the sequence depends on the choice of the initial iteration. In order to construct monotone convergent sequences we choose the initial iterations $(N^{(0)},u^{(0)})=(\tilde{N},\tilde{u})$ and $(N^{(0)},u^{(0)})=(\underset{\sim}{N},\underset{\sim}{u})$ independently, and obtain two sequences from (2.4) and (2.5). Denote these two sequences by $\{\bar{N}^{(k)},\bar{u}^{(k)}\}$ and $\{\underline{N}^{(k)},\underline{u}^{(k)}\}$ which are referred to as maximal and minimal sequence, respectively. It has been shown in Ref. 3 that these two sequences are monotone and converge from above and below, respectively, to a unique solution of (1.1)–(1.3). Specifically, we have the following existence-comparison theorem from Ref. 3.

Theorem 2.1: Let $(\tilde{N},\tilde{u})$, $(\underset{\sim}{N},\underset{\sim}{u})$ be upper and lower solutions such that $(\tilde{N},\tilde{u})\geqslant(\underset{\sim}{N},\underset{\sim}{u})\geqslant(0,0)$ and let the hypothesis (H_0) hold. Then the maximal sequence $\{\bar{N}^{(k)},\bar{u}^{(k)}\}$ converges monotonically from above to a unique solution (N,u) of (1.1)–(1.3) whilst the minimal sequence $\{\underline{N}^{(k)},\underline{u}^{(k)}\}$ converges monotonically from below to the same solution. Moreover

$$\begin{aligned}&(0,0)\leqslant(\underset{\sim}{N},\underset{\sim}{u})\leqslant(\underline{N}^{(1)},\underline{u}^{(1)})\leqslant\cdots\leqslant(\bar{N},\bar{u})\leqslant\cdots\leqslant(\bar{N}^{(1)},\bar{u}^{(1)})\\ &\quad\leqslant(\tilde{N},\tilde{u})\end{aligned}\tag{2.6}$$

In view of Theorem 2.1 the stability and instability problem of the system (1.1)–(1.3) can be investigated through suitable construction of upper and lower solutions or by using any pair of the iterations. In the following discussion we construct upper and lower solutions so that the stability or instability of the system can be determined. We first investigate the global stability problem when $\gamma(u)\leqslant\bar{\gamma}<\infty$ but $h(u)$ is not necessarily uniformly bounded. Motivated by the functions given by (1.4) we assume that $dh/du\equiv h'(u)\leqslant\bar{c}<\infty$ for $u\geqslant 0$. Under this condition we seek an upper solution from the linear equations

$$v^{-1}\mathscr{N}_t+\mu\mathscr{N}_x+\sigma\mathscr{N}=(\bar{\gamma}/2)\int_{-1}^{1}\mathscr{N}(t,x,\mu')\,d\mu',\tag{2.7}$$

$$U_t - DU_{xx} + \beta U = \left[(\bar{c}/2l)\int_{-1}^{1}\int_0^l \mathcal{N}(t,x',\mu')\,dx'd\mu'\right]U + g(t), \quad (2.8)$$

where

$$g(t) = (h(0)/2l)\int_{-1}^{1}\int_0^l \mathcal{N}(t,x',\mu')\,dx'\,d\mu'.$$

The boundary and initial conditions for $\mathcal{N}$ and U are the same as in (1.2) and (1.3). Since (2.7) is the standard linear transport equation, a unique nonnegative solution $\mathcal{N}$ exists and can be constructed by successive iteration the same way as in (2.4) except with $\gamma(u^{(k-1)})$ replaced by $\bar{\gamma}$ (see also Ref. 6). Knowing the function $\mathcal{N}\geqslant 0$, the solution U of (2.8) is nonnegative and can be obtained by standard methods (such as eigenfunction expansion or the method of Green's function, e.g., see Ref. 7).

To obtain a nonnegative lower solution we consider the equations

$$v^{-1}\hat{N}_t + \hat{N}_x + \sigma\hat{N} = 0, \quad (2.9)$$

$$\hat{u}_t - D\hat{u}_{xx} + (\beta + M)\hat{u} = 0 \quad (2.10)$$

under the same boundary and initial conditions as for $\mathcal{N}$, U, where M is the constant in the iteration process (2.5). By the positivity lemmas for transport and parabolic operators the functions $\hat{N},\hat{u}$ are nonnegative and satisfy the relation $(\hat{N},\hat{u})\leqslant(\mathcal{N},U)$ (cf. Refs. 3 and 6). Furthermore, if we define

$$N_0(x,\mu) = 0 \quad \text{for } x\notin[0,l], \quad \mu\in[-1,1], \quad (2.11)$$

the solution $\hat{N}$ of (2.9) is given by

$$\hat{N}(t,x,\mu) = \exp(-v\sigma t)N_0(x - v\mu t,\mu). \quad (2.12)$$

With this construction of the four functions $\mathcal{N},U,\hat{N},\hat{u}$ we have the following conclusion.

Theorem 2.2: Let γ, h satisfy (H_0) and let $\gamma(u)\leqslant\bar{\gamma}<\infty$, $h'(u)\leqslant\bar{c}<\infty$ for $u\geqslant 0$. Then the sequence $\{\bar{N}^{(k)},\bar{u}^{(k)}\}$ obtained from (2.4) and (2.5) with $(\bar{N}^{(0)},\bar{u}^{(0)}) = (\mathcal{N},U)$ converges monotonically from above to a unique solution (N,u) and the sequence $\{\underline{N}^{(k)},\underline{u}^{(k)}\}$ with $(\underline{N}^{(0)},\underline{u}^{(0)}) = (\hat{N},\hat{u})$ converges monotonically from below to the same solution. In particular,

$$\begin{aligned}&\hat{N}(t,x,\mu)\leqslant N(t,x,\mu)\leqslant\mathcal{N}(t,x,\mu),\\ &\hat{u}(t,x)\leqslant u(t,x)\leqslant U(t,x).\end{aligned} \quad (2.13)$$

Proof: Let $(\tilde{N},\tilde{u}) = (\mathcal{N},U)$. Then by (H_0), (2.7), (2.8), and the relation $h(u)\leqslant\bar{c}u + h(0)$ we have

$$v^{-1}\tilde{N}_t + \mu\tilde{N}_x + \sigma\tilde{N} = (\bar{\gamma}/2)\int_{-1}^{1}\tilde{N}\,d\mu'\geqslant\gamma(\tilde{u})\int_{-1}^{1}\tilde{N}\,d\mu',$$

$$\begin{aligned}\tilde{u}_t - D\tilde{u}_{xx} + \beta\tilde{u} &= \left[(\bar{c}/2l)\int_{-1}^{1}\int_0^l\tilde{N}\,dx'\,d\mu'\right]\tilde{u} + g(t)\\ &= ((\bar{c}\tilde{u} + h(0))/2l)\int_{-1}^{1}\int_0^l\tilde{N}\,dx'\,d\mu'\\ &\geqslant(h(\tilde{u})/2l)\int_{-1}^{1}\int_0^l\tilde{N}\,dx'\,d\mu'.\end{aligned}$$

Since $\mathcal{N},U$ satisfy the corresponding boundary and initial conditions in (1.2) and (1.3) the pair $(\mathcal{N},U)$ is a nonnegative upper solution. It is easily seen by the same reasoning that the pair $(N,\underline{u}) = (\hat{N},\hat{u})$ is a lower solution and $(\mathcal{N},U)\geqslant(\hat{N},\hat{u})\geqslant(0,0)$. It follows from Theorem 2.1 that the maximal sequence $\{\bar{N}^{(k)},\bar{u}^{(k)}\}$ amd the minimum sequence $\{\underline{N}^{(k)},\underline{u}^{(k)}\}$ converge monotonically from above and below, respectively, to a unique solution (N,u) which satisfies the relation (2.13). This proves the theorem.

It has been shown in Ref. 6 that if

$$\bar{\gamma}<\sigma[1 - E_2(\sigma l/2)]^{-1}, \quad (2.14)$$

where $E_2(z)$ is the exponential integral of order 2 then the solution $\mathcal{N}$ of the linear equation (2.7) converges to zero exponentially as $t\to\infty$. Knowing the exponential decay of $\mathcal{N}$ the standard argument for linear parabolic systems shows that the solution U of (2.8) also converges exponentially to zero as $t\to\infty$ (e.g., see Ref. 3). Since $\hat{N}$ and $\hat{u}$ both converge to zero in exponential order as $t\to\infty$ we have the following conclusion regarding the global stability of the system.

Theorem 2.3: Let γ, h satisfy (H_0) and let $\gamma(u)\leqslant\bar{\gamma}<\infty$, $h'(u)\leqslant\bar{c}<\infty$ for $u\geqslant 0$. If, in addition, (2.14) holds then for any $N_0\geqslant 0, u_0\geqslant 0$ the corresponding solution (N,u) converges exponentially to $(0,0)$ as $t\to\infty$.

When the functions γ,h are governed by (1.4) with $m>1$, $n\geqslant 0$ the requirements in Theorem 2.3 are all fulfilled with

$$\begin{aligned}\bar{\gamma} &= \sup_{u>0}\{\gamma_0 + a_1(m-1)^{-1}[T_c^{1-m} - (u+T_c)^{1-m}]\}\\ &= \gamma_0 + a_1[(m-1)T_c^{m-1}]^{-1}, \quad (2.15)\\ \bar{c} &= \sup_{u>0}\{a_2(u+T_c)^{-n}\} = a_2T_c^{-n}.\end{aligned}$$

An immediate consequence of Theorem 2.3 is the following:

Corollary: Let γ,h be governed by (1.4) with $m>1$, $n\geqslant 0$ and let

$$\gamma_0 + a_1[(m-1)T_c^{m-1}]^{-1}<\sigma[1 - E_2(\sigma l/2)]^{-1}. \quad (2.16)$$

Then for any $N_0\geqslant 0$, $u_0\geqslant 0$, the corresponding solution (N,u) of (1.1)–(1.3) converges exponentially to $(0,0)$ as $t\to\infty$.

It is to be noted that if $a_1 = 0$, that is, if $\gamma(u) = \gamma_0$ which is independent of u, then the result of the above corollary coincides with a result in Ref. 6. Thus the term $a_1[(m-1)T_c^{m-1}]^{-1}$ may be regarded as the effect of the temperature feedback on the stability of the system.

3. REGIONAL STABILITY AND INSTABILITY

The construction of the function $\mathcal{N}$ in Theorems 2.2 and 2.3 is based on the assumption that $\gamma(u)$ be uniformly bounded. When this condition is not satisfied the value of $\bar{\gamma}$ is no longer finite and condition (2.14) cannot be fulfilled. However, the proof of Theorem 2.2 indicates that if U is bounded by some constant ρ then the function $\mathcal{N}$ can be determined from (2.7) with $\bar{\gamma}$ replaced by $\gamma(\rho)$. In particular, if condition (2.14) holds with respect to $\gamma(\rho)$ then under a suitable condition on h the decay property of the solution (N,u) remains true. In order to establish this conclusion we need to use the function

$$\phi(x,\mu)\equiv\begin{cases}1-\exp(-\sigma x/l) & \text{when } \mu>0,\\ 1-\exp(\sigma(l-x)/\mu) & \text{when } \mu<0\end{cases} \tag{3.1}$$

for the construction of an upper solution. The following theorem gives an existence-stability result when γ, h are not necessarily uniformly bounded.

Theorem 3.1: Let γ,h satisfy (H_0) and let there exist positive constants B,ρ,ν with $\nu\geqslant 1$ such that

$$h(u)\leqslant Bu^{\nu} \quad \text{when } 0\leqslant u\leqslant\rho, \tag{3.2}$$

$$\gamma(\rho)<\sigma[1-E_2(\sigma l/2)]^{-1}. \tag{3.3}$$

Then for any

$$\rho^*<(\beta+D\pi^2/l^2)(B\rho^{\nu-1})^{-1} \tag{3.4}$$

there exists a constant $\alpha>0$ such that a unique global solution (N,u) to (1.1)–(1.3) exists and satisfies the relation

$$\begin{aligned}&\hat{N}(t,x,\mu)\leqslant N(t,x,\mu)<\rho^* e^{-\alpha t}\phi(x,\mu),\\ &\hat{u}(t,x)\leqslant u(t,x)\leqslant\rho e^{-\alpha t}\sin(\pi x/l),\end{aligned} \tag{3.5}$$

for $t>0,(x,\mu)\in\bar{\Omega}_2$ whenever it holds at $t=0$.

Proof: Let $\tilde{N}=\rho^* e^{-\alpha t}\phi(x,\mu),\tilde{u}=\rho e^{-\alpha t}\sin(\pi x/l)$. Since $\tilde{N}$ and $\tilde{u}$ satisfy the respective boundary and initial conditions in (1.2) and (1.3) the pair $(\tilde{N},\tilde{u})$ is an upper solution if it satisfies the differential inequality in (2.2). Thus it suffices to show that

$$\begin{aligned}&\rho^* e^{-\alpha t}[\mu\phi_x+(\sigma-\alpha v^{-1})\phi]\\ &\quad\geqslant(\gamma(\tilde{u})/2)\int_{-1}^{1}\rho^* e^{-\alpha t}\phi(x,\mu')\,d\mu',\\ &\rho e^{-\alpha t}[(\beta-\alpha)\sin(\pi x/l)-D(\sin(\pi x/l))_{xx}]\\ &\quad\geqslant(h(\tilde{u})/2l)\int_{-1}^{1}\int_0^l\rho^* e^{-\alpha t}\phi\,dx'\,d\mu'.\end{aligned} \tag{3.6}$$

Since by (H_0),(3.2), and $\tilde{u}\leqslant\rho$,

$$\gamma(\tilde{u})\leqslant\gamma(\rho),\quad h(\tilde{u})\leqslant B(\rho e^{-\alpha t}\sin(\pi x/l))^{\nu},$$

The inequalities in (3.6) are fulfilled if

$$\begin{aligned}&\mu\phi_x+(\sigma-\alpha v^{-1})\phi\geqslant(\gamma(\rho)/2)\int_{-1}^{1}\phi(x,\mu')d\mu',\\ &(\beta-\alpha)+D\pi^2/l^2\\ &\quad\geqslant(B\rho^*/2l)(\rho\sin(\pi x/l))^{\nu-1}\int_{-1}^{1}\int_0^l\phi\,dx'\,d\mu'.\end{aligned} \tag{3.7}$$

It is easily seen from the definition of ϕ that

$$\begin{aligned}&\mu\phi_x+\sigma\phi=\sigma,\quad\int_{-1}^{1}\int_0^l\phi(x',\mu')\,dx'\,d\mu'\leqslant 2l,\\ &\int_{-1}^{1}\phi(x,\mu')\,d\mu'\\ &\quad=\int_0^1[2-e^{-\sigma x/\mu'}-e^{-\sigma(l-x)/\mu'}]\,d\mu'\\ &\quad\leqslant 2(1-E_2(\sigma l/2)).\end{aligned} \tag{3.8}$$

Hence the relation (3.7) holds if

$$\sigma-\alpha v^{-1}\phi\geqslant\gamma(\rho)(1-E_2(\sigma l/2)),$$

$$(\beta-\alpha)+D\pi^2/l^2\geqslant B\rho^*\rho^{\nu-1}.$$

In view of (3.3) and (3.4) both inequalities are satisfied by a suitable constant $\alpha>0$. This proves that $(\tilde{N},\tilde{u})$ is a nonnegative upper solution. Since $(\underset{\sim}{N}\underset{\sim}{u})=(0,0)$ is clearly a lower solution the existence of a solution (N,u) to (1.1)–(1.3) and the relation $(\tilde{N},\tilde{u})\geqslant(N,u)\geqslant(0,0)$ follow immediately from Theorem 2.1. To improve the lower bound of the solution we use the initial iteration $(\underset{\sim}{N}^{(0)},\underset{\sim}{u}^{(0)})=(0,0)$ and construct the first iteration $(\underset{\sim}{N}^{(1)},\underset{\sim}{u}^{(1)})$ from (2.4) and (2.5). Since $(\underset{\sim}{N}^{(1)},\underset{\sim}{u}^{(1)})$ coincides with the function $(\hat{N},\hat{u})$ given by (2.9) and (2.10) [under the same boundary-initial conditions (1.2) and (1.3)] we conclude that (N,u) satisfies the relation (3.5). This proves the theorem.

The conclusion in Theorem 3.1 implies that under the conditions (3.2) and (3.3) the zero steady-state solution is asymptotically stable. We next give a sufficient condition on γ,h so that the system is unstable. In fact, by following the approach of Ref. 5 we construct a suitable lower solution which ensures that either a unique global solution exists and grows unbounded at $t\to\infty$ or the solution exists locally and blows up in finite time. In stating this result in the following theorem it is convenient to set

$$\hat{\phi}\equiv(2l)^{-1}\int_{-1}^{1}\int_0^l\phi(x,\mu)\,dx\,d\mu.$$

Theorem 3.2: Let γ,h satisfy (H_0) and let

$$h(u)\geqslant bu^{\mu}\quad(u\geqslant 0) \tag{3.9}$$

for some constants $b>0,0\leqslant\mu\leqslant 1$. If

$$\gamma(0)>2\sigma[1-E_2(\sigma l)]^{-1}, \tag{3.10}$$

then for any $N_0\geqslant\rho_1^*\phi$, $u_0\geqslant\rho_1\sin(\pi x/l)$ with

$$\rho_1^*>(\beta+D\pi^2/l^2)\rho_1^{1-\mu}(b\hat{\phi})^{-1} \tag{3.11}$$

either a unique global solution (N,u) exists and satisfies

$$N(t,x,\mu)\geqslant\rho_1 e^{\epsilon t}\phi(x,\mu),\quad u(t,x)\geqslant\rho_1 e^{\epsilon t}\sin(\pi x/l) \tag{3.12}$$

for some $\epsilon>0$ or else the solution blows up in finite time.

Proof: Let M^* be a sufficiently large constant and define modified functions γ^*, h^* by

$$\begin{aligned}&\gamma^*(u)=\gamma(u),\quad h^*(u)=h(u)\quad\text{when } u\leqslant M^*,\\ &\gamma^*(u)=\gamma(M^*),\quad h^*(u)=h(M^*)\quad\text{when } u>M^*.\end{aligned} \tag{3.13}$$

Then γ^*, h^* are uniformly bounded on $[0,\infty)$ and satisfy the hypothesis (H_0). Upon replacing γ,h in (1.1) by γ^*,h^* the corresponding "modified problem" (1.1)–(1.3) has a unique global solution (N^*,u^*) which satisfies the relation $(\underset{\sim}{N}\underset{\sim}{u})\leqslant(N,u)\leqslant(\tilde{N},\tilde{u})$ provided that $(\tilde{N},\tilde{u})$ and $(\underset{\sim}{N}\underset{\sim}{u})$ are upper and lower solutions of the modified problem with $(\underset{\sim}{N}\underset{\sim}{u})\leqslant(\tilde{N},\tilde{u})$. We first seek a lower solution in the form

$$\underset{\sim}{N}=\rho_1^* e^{\epsilon t}\phi(x,\mu),\quad\underset{\sim}{u}=\rho_1 e^{\epsilon t}\sin(\pi x/l). \tag{3.14}$$

Since the boundary and initial requirements for a lower solution are fulfilled it suffices to varify the reversed inequality in (2.2) which, in the present situation, is equivalent to

$$\rho_1^* e^{\epsilon t}[\mu\phi_x + (\sigma + v^{-1}\epsilon)\phi]$$
$$\leqslant (\gamma^*(\underset{\sim}{u})/2)\int_{-1}^{1}\rho_1^* e^{\epsilon t}\phi\, d\mu',$$
$$\rho_1 e^{\epsilon t}(\beta + \epsilon + D\pi^2/l^2)\sin(\pi x/l)$$
$$\leqslant (h^*(\underset{\sim}{u})/2l)\int_{-1}^{1}\int_0^l \rho_1^* e^{\epsilon t}\phi\, dx'\, d\mu'.$$

By the relation (3.8) and the hypothesis (3.9) these inequalities hold if

$$\sigma + v^{-1}\epsilon\phi \leqslant (\gamma^*(\underset{\sim}{u})/2)\int_{-1}^{1}\phi\, d\mu',$$
$$\beta + \epsilon + D\pi^2/l^2 \leqslant b\rho_1^*\hat{\phi}\,(\rho_1\sin(\pi x/l))^{\mu-1}. \tag{3.15}$$

Since $\gamma^*(u) \geqslant \gamma(0)$, $0 \leqslant \mu \leqslant 1$, and (see Ref. 6)

$$\int_{-1}^{1}\phi(x,\mu')\, d\mu' \geqslant 1 - E_2(\sigma l),$$

the relation (3.15) is satisfied if

$$\sigma + v^{-1}\epsilon\phi \leqslant (\gamma(0)/2)(1 - E_2(\sigma l)),$$
$$\beta + \epsilon + D\pi^2/l^2 \leqslant b\rho_1^*\hat{\phi}\rho_1^{\mu-1}.$$

In view of the conditions (3.10) and (3.11) these two inequalities are both satisfied by a suitable constant $\epsilon > 0$. With this choice of ϵ the function $(\underset{\sim}{N}\underset{\sim}{u})$ given by (3.14) is a lower solution of the modified problem.

To find an upper solution $(\tilde{N},\tilde{u}) \geqslant (\underset{\sim}{N}\underset{\sim}{u})$ we let

$$\tilde{N} = C_1 e^{\lambda t}, \quad \tilde{u} = C_2 e^{\lambda t}, \tag{3.16}$$

where C_1, C_2, λ are positive constants with $C_1 \geqslant N_0, C_2 \geqslant u_0$, and $\lambda \geqslant \epsilon$. It is clear that $(\tilde{N},\tilde{u})$ satisfies the boundary and initial requirements in (2.2) as well as $(\tilde{N},\tilde{u}) \geqslant (\underset{\sim}{N}\underset{\sim}{u})$. Hence it suffices to show that

$$C_1 e^{\lambda t}(\sigma + v^{-1}\lambda) \geqslant (\gamma^*(\tilde{u})/2)\int_{-1}^{1} C_1 e^{\lambda t}\, d\mu',$$
$$C_2 e^{\lambda t}(\beta + \lambda) \geqslant (h^*(\tilde{u})/2l)\int_{-1}^{1}\int_0^l C_1 e^{\lambda t}\, dx'\, d\mu'.$$

The above relation is equivalent to

$$(\sigma + v^{-1}\lambda) \geqslant \gamma^*(C_2 e^{\lambda t}),$$
$$C_2(\beta + \lambda) \geqslant C_1 h^*(C_2 e^{\lambda t}).$$

Since γ^*, h^* are uniformly bounded, both inequalities are satisifed by a sufficiently large λ. The above construction shows that the modified problem (1.1)–(1.3) has a unique global solution (N^*,u^*) which satisfies the relation

$$\rho_1^* e^{\epsilon t}\phi(x,\mu) \leqslant N^*(t,x,\mu) \leqslant C_1 e^{\lambda t},$$
$$\rho_1 e^{\epsilon t}\sin(\pi x/l) \leqslant u^*(t,x) \leqslant C_2 e^{\lambda t}.$$

Since (N^*,u^*) is also a solution of the original system for as long as $N^* \leqslant M^*$, $u^* \leqslant M^*$ and since M^* can be assigned an arbitrarily large value, we conclude by a contradiction argument as in Ref. 5 that either a global solution (N,u) to the original system exists and satisfies the relation (3.12), or else there exists a finite t^* such that (3.12) holds on $[0,t^*)\times[0,l]\times[-1,1]$ and (N,u) blows up at t^*. This completes the proof of the theorem.

The result of Theorem 3.1 implies that under the conditions (3.2) and (3.3) the steady-state solution (0,0) is asymptotically stable and a stability region is given by

$$\Lambda_1 \equiv \{(N_0,u_0);\ 0 \leqslant N_0 \leqslant \rho^*\phi(x,\mu),$$
$$0 \leqslant u_0 \leqslant \rho\sin(\pi x/l)\}, \tag{3.17}$$

where ρ^* and ρ are related by (3.4). On the other hand, the conclusion of Theorem 3.2 states that under the conditions (3.9) and (3.10) the system (1.1)–(1.3) is unstable and an instability region is given by

$$\Lambda_2 \equiv \{(N_0,u_0);\ N_0 \geqslant \rho_1^*\phi(x,\mu),\ u_0 \geqslant \rho_1\sin(\pi x/l)\}, \tag{3.18}$$

where ρ_1^* and ρ_1 are related by (3.11). It is clear that the stability and instability regions depend on the properties of γ, h as well as the diffusion effect which is measured by the thermal conductivity D and slab length l. This result demonstrates the fact that smaller slab length tends to stabilize and increases the domain of the stability region. This can be seen either from conditions (3.3) and (3.4) for stability or from (3.10) and (3.11) for instability. The relations (3.4) and (3.11) also demonstrate that larger thermal conductivity increases the stability region and decreases the instability region. The above observation is valid for at least the present positive temperature feedback system.

[1] C. I. Bell and S. Glasstone, *Nuclear Reactor Theory* (Van Norstrand Reinhold, New York, 1970).

[2] A. Belleni-Morante, Nucl. Sci. Eng. **59**, 56–58 (1976).

[3] C. V. Pao, *Proceedings of the Seventh International Conference on Transport Theory*, edited by P. Nelson and H. D. Victory, Lubbock, Texas, March, 1981, Prog. Nucl. Energy **8**, 191–202 (1982).

[4] D. H. Nguyen, Nucl. Sci. Eng. **50**, 370–381 (1973).

[5] C. V. Pao, "Comparison and stability of solutions for a neutron transport problem with temperature feedback" (to appear in SIAM J. Math. Anal.).

[6] C. V. Pao, J. Integral Eqs. **1**, 131–152 (1979).

[7] I. Stakgold, *Boundary Value Problems on Mathematical Physics* (Macmillan, New York, 1968), Vol. II.

Rayleigh–Taylor instability of a fluid supported against gravity by an oscillating magnetic field

B. B. Chakraborty
Mathematics Section, Department of Chemical Technology, Bombay University, Matunga, Bombay-400 019, India

H. K. S. Iyengar
Department of Mathematics, Patkar College, Bombay-400 062, India

A. R. Nayak
Bhaba Atomic Research Centre, Bombay-400 085, India

(Received 5 March 1982; accepted for publication 7 May 1982)

Rayleigh–Taylor instability of an incompressible, inviscid, and infinitely conducting heavy fluid supported against gravity by a horizontal and uniform magnetic field is studied. The magnetic field consists of a steady part and an unsteady part, of magnitude varying with time, rotating with a fixed period about a vertical axis. The growth rate of instability of any unstable disturbance is obtained analytically when the magnitude of the unsteady part of the field is small. When the unsteady part of the field is not small, the growth rate is calculated numerically. The analytical as well as the numerical results show that, except in some special cases, the steady and the unsteady field components together have a greater stabilizing influence on the fluid than in the case when the unsteady field is acting alone. A physical explanation of the stabilizing influence of the combined presence of the steady and unsteady field is suggested.

PACS numbers: 47.20. + m, 47.65. + a

I. INTRODUCTION

Hydromagnetic Rayleigh–Taylor instability is of importance in magnetic confinement in fusion problems.[1] Two models of Rayleigh–Taylor[2,3] instability of a heavy fluid in the presence of a magnetic field are of special interest. In the one, studied by Chandrasekhar,[4] a heavier fluid is supported against gravity by a lighter fluid, both fluids being infinitely conducting and permeated by a uniform, steady, horizontal magnetic field. In another model, studied by Kruskal and Schwarzchild,[5] an infinitely conducting fluid is supported against gravity by a uniform, steady, horizontal magnetic field in vacuum. The fluid is also permeated by a similar magnetic field of different strength.

The steady magnetic field generally has a stabilizing influence on the flow of a conducting fluid. Drazin[6] has pointed out that, since it is not difficult to produce and maintain an oscillatory field, it is of interest to investigate if such an oscillating field has a greater stabilizing influence on the flow of a conducting fluid than a steady field. He has investigated the Kelvin–Helmholtz instability in a conducting fluid in the presence of an oscillating magnetic field. He has concluded that an oscillating magnetic field is less efficient in stabilizing a vortex sheet than a steady field.

In view of the importance of magnetohydrodynamic Rayleigh–Taylor instability in various applications, it is of interest to examine this instability in the presence of an oscillating field. The effect of an oscillating magnetic field on the Rayleigh–Taylor instability model of Chandrasekhar[4] has been recently studied in Ref. 7. Berkowitz *et al.*[8] made one of the earliest studies of Rayleigh–Taylor instability in a modified Kruskal–Schwarzschild model of a heavy fluid supported against gravity by a uniform, oscillating magnetic field, of constant magnitude, rotating uniformly about a vertical axis.

In the present paper we extend the discussion of Berkowitz *et al.*[8] by reconsidering their model with a more general oscillating magnetic field. The oscillating magnetic field in our problem consists of a uniform, steady part, and an unsteady part, of magnitude varying with time, rotating about a vertical axis with constant period. The rotating field can be resolved into two mutually perpendicular, oscillating components of the same frequency but of different amplitudes and having a phase difference of $\pi/2$. In the special case when the steady part of the magnetic field vanishes and the unsteady part is a field rotating with constant angular velocity and has a fixed magnitude, so that it can be resolved into two mutually perpendicular oscillating components of equal amplitudes but having a phase difference of $\pi/2$, the present problem reduces to the one considered in Ref. 8. In the present study we want to find the effect of the simultaneous presence of a steady component of the magnetic field and of a more general unsteady field than that considered in Ref. 8, on the stability of the fluid supported by the field.

We obtain the dispersion relation as a differential equation for the vertical displacement of the fluid surface from its equilibrium horizontal position with time as the independent variable. This equation turns out to be a Hill's equation, and it reduces to a Mathieu's equation, obtained by Berkowitz *et al.*[8] in the special case when our problem reduces to the one considered by them. Using the theory of Mathieu's equation they showed that for large wave number of disturbances the growth rate of instability has an upper bound independent of the value of the wave number and in this sense the oscillating field can be said to control the instability. In the present paper, when the amplitudes of the oscillating components of the unsteady part of the field are small, we obtain the growth rate of instability analytically by adapting a perturbation method developed by Whittaker,[9]

 0022-2488/83/051326-06$02.50

and discussed in Nayfeh,[10] for finding the characteristic exponent in the Mathieu equation.

When the amplitudes of the oscillating components of the unsteady part of the field are not small, we obtain the growth rate of instability numerically. Here we first note that discussion of stability can be split up into two distinct cases, according to whether θ is different from $\pi/2$ or not, where θ is the angle the direction of propagation for disturbances makes with the steady part of the magnetic field. When $\theta \neq \pi/2$, the stability discussion may be made to depend on the stability analysis for disturbances propagating along the steady component of the magnetic field (so that $\theta = 0$). When $\theta = \pi/2$, we find that the effect of the steady part of the magnetic field vanishes, and the present problem becomes, in effect, identical with that considered in Berkowitz *et al.*[8] We obtain numerical values of the growth rate of instability for different values of wave numbers of disturbances and of the other physical parameters, and compare the results when there is a steady component of the magnetic field with those obtained when there is no such component (as in Ref. 8). The analytical as well as the numerical results show that the steady and the unsteady components of the magnetic field together have a greater stabilizing influence on the fluid than in the case when the unsteady component is acting alone. The present model of Rayleigh–Taylor instability, because of the presence of the additional steady component of the magnetic field, is thus more stable than the model considered in Berkowitz *et al.*[8] A physical explanation for this enhancement of stability is suggested.

We formulate the problem and obtain the dispersion relation in Sec. II. When the amplitudes of the oscillating components of the unsteady part of the magnetic field are small, the dispersion relation is discussed analytically in Sec. III. When these amplitudes are not small, the dispersion relation is discussed numerically in Sec. IV. Some concluding remarks are made in Sec. V.

II. FORMULATION OF THE PROBLEM AND THE DISPERSION RELATION

An infinitely conducting, inviscid, and incompressible fluid of density ρ is at rest and occupies the region $z > 0$. It is supported against gravity by a horizontal magnetic field

$$\mathbf{H} = H\,[a + \epsilon \cos \omega t\,,\, \epsilon' \sin \omega t\,,\, 0]\,, \tag{1}$$

which occupies the region $z < 0$ in vacuum. The uniform magnetic field has thus a steady component along the x axis and an unsteady part having oscillating components $H\epsilon \cos \omega t$ and $H\epsilon' \sin \omega t$ along the x and y axes, respectively. These mutually perpendicular oscillating components have the same frequency but they differ in phase by $\pi/2$. The unsteady part of the field thus rotates with constant period about a vertical axis, but has a magnitude varying with time. The fluid is subjected to a gravitational acceleration $\mathbf{g} = (0,0,-g)$. The stability of the fluid is studied against small disturbances in which perturbation in any physical quantity is taken in the form

$$\tilde{q} = \hat{q}(z,t)\exp(ik_x x + ik_y y)\,. \tag{2}$$

Since the fluid is taken as infinitely conducting and is free from a magnetic field before disturbances, it remains so after disturbances also. The motion in the fluid is thus governed by ordinary hydrodynamic equations. The velocity of the fluid after disturbances is given by

$$\tilde{\mathbf{v}} = \nabla\tilde{\phi}\,, \tag{3}$$

where $\tilde{\phi}$ is perturbation in velocity potential. The equation of continuity takes the form

$$\nabla^2\tilde{\phi} = 0\,. \tag{4}$$

In view of Eq. (2), the solution of (4), bounded as $z \to \infty$, is

$$\tilde{\phi} = A\,(t)\exp[\,-kz + i(k_x x + k_y y)]\,, \tag{5}$$

where

$$k^2 = k_x^2 + k_y^2\,, \tag{6}$$

and $A\,(t)$ is a function of time t to be suitably determined. The fluid motion satisfies Bernoulli's equation[11]

$$\frac{\partial\tilde{\phi}}{\partial t} + \tfrac{1}{2}\tilde{\mathbf{v}}^2 + gz + p/\rho = \text{constant}\,, \tag{7}$$

where p is pressure in the fluid. The magnetic field in vacuum, satisfying Maxwell's equations, can be expressed, in the absence of a displacement current, as

$$\tilde{\mathbf{h}} = \nabla\tilde{\psi}\,, \tag{8}$$

where $\tilde{\psi}$, like $\tilde{\phi}$, satisfies Laplace's equation and is given as

$$\tilde{\psi} = B\,(t)\exp[kz + i(k_x x + k_y y)]\,. \tag{9}$$

$\tilde{\psi}$ is bounded as $z \to -\infty$ and $B\,(t)$ in (9) is a function of t to be suitably determined.

Since disturbances are taken small, we shall neglect products as well as squares and higher powers of small perturbations in the further analysis.

The surface separating the fluid and vacuum is given by

$$z = \eta(t)\exp(ik_x x + ik_y y)\,. \tag{10}$$

The kinematic condition to be satisfied on the boundary is that the normal component of fluid velocity on the boundary should be equal to the velocity of the surface of separation in the normal direction.[11]

Two other boundary conditions involving magnetic field, to be satisfied on the surface are[12]

(i) continuity of normal component of the magnetic field;

(ii) equality of fluid pressure on the boundary to the magnetic pressure in vacuum there.

The kinematic condition determines $A\,(t)$ in term of $d\eta/dt$. Of the two boundary conditions involving the magnetic field, the first condition determines $B\,(t)$ in term of η, while the second condition gives $\tilde{p}$, perturbation of fluid pressure on the boundary, in terms of η. In view of these results, the Bernoulli's equation (7) on being linearized and applied on the disturbed boundary, finally gives the dispersion relation

$$\frac{d^2\eta}{dt^2} + \eta\left\{\frac{\mu(\mathbf{k}\cdot\mathbf{H})^2}{\rho} - kg\right\} = 0\,. \tag{11}$$

Taking θ as the angle between the direction of propagation of a disturbance and the direction of the steady component of the field (x axis), b as a characteristic length, and defining the dimensionless quantities

$$\bar{\eta}=\eta/b\,,\quad X=kb\,,\quad G=\frac{g}{\omega^2 b}\,,\quad \bar{H}=\frac{\mu H_0^2}{\rho b g}\,, \tag{12}$$

we can write the dispersion relation (11) as

$$\frac{d^2\bar{\eta}}{dT^2}+J(T)\bar{\eta}=0\,, \tag{13}$$

where

$$2T=\omega t-\alpha\,, \tag{14}$$

$$\cos\alpha=(\epsilon k_x)/l\,,\quad \sin\alpha=(\epsilon' k_y)/l\,, \tag{15}$$

with

$$l^2=\epsilon^2 k_x^2+\epsilon'^2 k_y^2\,, \tag{16}$$

and

$$J(T)=\theta_0+2\theta_1\cos 2T+2\theta_2\cos 4T\,, \tag{17}$$

with

$$l'^2=\epsilon^2\cos^2\theta+\epsilon'^2\sin^2\theta\,, \tag{18}$$

$$\theta_0=4[\bar{H}GX^2(a^2\cos^2\theta+l'^2/2)-XG]\,, \tag{19}$$

$$\theta_1=4a\bar{H}G\cos\theta l'X^2\,, \tag{20}$$

$$\theta_2=\bar{H}l'^2X^2G\,. \tag{21}$$

Equations (17)–(21) show that, in the final dispersion relation (13), the amplitudes ϵ and ϵ' of the two mutually perpendicular oscillating components of the unsteady part of the equilibrium magnetic field occur only through (18), which defines l'. Equation (18) shows that the given unsteady part of the field is equivalent to one for which amplitudes ϵ,ϵ' of the mutually perpendicular oscillating components are both equal to l', and, therefore, they together represent a rotating field with constant magnitude l' and constant angular velocity ω. Similarly, since a enters the dispersion relation through the expression $a\cos\theta$, the effective magnitude of the steady part of the magnetic field is reduced to $a\cos\theta$ from a. When disturbances travel normal to the steady component of the magnetic field ($\theta=\pi/2$), this component has no effect on the stability problem.

Noting the difference between the notations in Ref. 8 and in the present paper, we find that Eq. (11) above reduces in the limiting case, when the steady part of the magnetic field vanishes ($a\to 0$) and $\epsilon=\epsilon'$, to the dispersion relation obtained by Berkowitz *et al.*[8] [their Eq. (77)].

III. STABILITY ANALYSIS FOR SMALL AMPLITUDE MAGNETIC FIELD OSCILLATIONS

A. General case

Since the dispersion relation (13) is a Hill's equation, we assume a solution in the form[13,14]

$$\bar{\eta}=e^{\lambda T}u(T)\,, \tag{22}$$

where the constant λ is the characteristic exponent and $u(T)$ is a periodic function of period π. We have instability only if λ has a real positive part. For small values of the amplitudes ϵ,ϵ' of the oscillating components of the unsteady part of the equilibrium magnetic field, we take l' as small [cf. Eq. (18)] and approximately determine λ by adapting a perturbation analysis used by Whittaker[9] in determining the characteristic exponent of the Mathieu equation.

Substituting for $\bar{\eta}$, using Eq. (22), we obtain from Eq. (13)

$$\frac{d^2u}{dT^2}+2\lambda\frac{du}{dT}+(\lambda^2+J(T))u=0\,. \tag{23}$$

We use the expansions

$$\lambda=\lambda_0+l'\lambda_1+l'^2\lambda_2+\dots\,, \tag{24}$$

$$u=u_0+l'u_1+l'^2u_2+\dots\,. \tag{25}$$

and expand $J(T)$ in powers of l' in Eq. (23). We use the perturbation theory to obtain in succession $u_0(T)$, $u_1(T)$, $u_2(T)$,... by solving the equations obtained by equating the coefficients of l'^0, l', l'^2,... on the left-hand side of (23) to zero. Thus, equating the term on the left-hand side of (23), that is independent of l', to zero, we have

$$\frac{d^2u_0}{dT^2}+2\lambda_0\frac{du_0}{dT}+(\lambda_0^2-m^2)u_0=0\,, \tag{26}$$

where

$$m^2=4GX(1-Xa^2\cos^2\theta\bar{H})\,. \tag{27}$$

Putting

$$\bar{\eta}_0=u_0\exp(\lambda_0 T)\,, \tag{28}$$

in (26), we find that a solution for (26) may be taken as

$$\bar{\eta}_0=Ce^{mt}\,, \tag{29}$$

where C is an arbitrary constant.

Comparing the solution (29) with (28), we have

$$\lambda_0=m \tag{30}$$

and

$$u_0=C\,. \tag{31}$$

Similarly, from the left-hand side of (23), using (24) and (25) and equating the coefficients of l' to zero, we have, in view of the fact that u_0 is a constant and (30) is true,

$$\frac{d^2u_1}{dT^2}+2\lambda_0\frac{du_1}{dT}=-2\lambda_0\lambda_1u_0-8au_0\bar{H}GX^2\cos\theta\cos 2t\,. \tag{32}$$

Taking

$$\lambda_1=0 \tag{33}$$

in (32) to avoid secular terms in u_1, we obtain, in view of (30),

$$u_1=-2u_0a\bar{H}GX^2\cos\theta\times(m\sin 2T-\cos 2T)(1+m^2)\,. \tag{34}$$

The coefficient of l'^2 on the left-hand side of (23), in view of (24), (25), (33), and (34), when equated to zero, gives

$$\begin{aligned}\frac{d^2u_2}{dT^2}+2\lambda_0\frac{du_2}{dT}&+[\lambda_0^2+4(a^2\cos^2\theta\bar{H}GX^2-GX)]u_2\\&=-2\lambda_0\lambda_2u_0-8a\bar{H}GX^2\cos\theta\cos(2T)u_1\\&\quad-2\bar{H}GX^2(1+\cos 4T)u_0\,.\end{aligned} \tag{35}$$

To avoid secular terms in u_2, we equate the constant term on the right-hand side of (35) to zero and obtain, in view of (30) and (34),

$$\lambda_2=-\frac{\bar{H}GX^2}{m}\left[1+\frac{4a^2\bar{H}GX^2\cos^2\theta}{(1+m^2)}\right]. \tag{36}$$

From (24), (30), (33), and (36), we have

$$\lambda = m - \frac{\overline{H}GX^2}{m}\left[1 + \frac{4a^2\overline{H}GX^2\cos^2\theta}{(1+m^2)}\right]l'^2 + O(l'^3). \quad (37)$$

If, instead of (29), we take

$$\bar{\eta}_0 = D\exp(-mt), \quad (38)$$

where D is a constant, as another solution of (26), we shall then have, in view of (28),

$$u_0 = D \quad (39)$$

and

$$\lambda_0 = -m. \quad (40)$$

We shall also get λ_1, λ_2, and λ as given by (33), (36), and (37), respectively, with m replaced by $-m$, so that

$$\lambda = -m + \frac{\overline{H}GX^2}{m}\left[1 + \frac{4a^2\overline{H}GX^2\cos^2\theta}{(1+m^2)}\right]l'^2 + O(l'^3) \quad (41)$$

is a second value of the characteristic exponent, other than the one given by (37).

In the absence of any oscillating component in the unsteady magnetic field ($l' = 0$), there will be instability with growth rate $\lambda = m$ [cf. Eq. (37)] when m^2, given by Eq. (27), is positive. Equation (37) shows that this growth rate is decreased by the amount

$$\frac{\overline{H}GX^2}{m}\left[1 + \frac{4a^2\overline{H}GX^2\cos^2\theta}{(1+m^2)}\right]l'^2.$$

This decrement in growth rate vanishes when $l' = 0$, but it increases with l', i.e., with the amplitudes of the oscillating components [cf. Eq. (18)]. It also increases with a, the strength of the steady part of the magnetic field, since m decreases with a [cf. Eq. (27)]. Thus the steady and the unsteady fields together have a greater stabilizing influence on the fluid than in the case when either of them is acting alone.

B. Special cases

The expressions for λ, obtained in Eqs. (37) and (41), break down when $m = 0$ or $m = \pm i$. In these cases we have to reconsider the stability problem.

1. Case (i): $m = 0$

In this case, using (24) and (25) in (23), and equating the terms independent of l' on the left-hand side of (23) to zero, we obtain

$$D^2(u_0 e^{\lambda_0 t}) = 0. \quad (42)$$

Solving (42) we have

$$e^{\lambda_0 t}u_0 = A + Bt, \quad (43)$$

where A and B are arbitrary constants. Taking $B = 0$ to avoid secular terms, we have, from (43),

$$\lambda_0 = 0, \quad u_0 = A. \quad (44)$$

The coefficient of l' on the left-hand side of (23), when equated to zero, gives an equation for u_1, the solution for which is

$$u_1 = 2a\overline{H}G\cos(\theta)X^2 u_0 \cos 2T. \quad (45)$$

Similarly, equating the coefficient of l'^2 to zero, we obtain the differential equation for u_2 as

$$\frac{d^2u_2}{dT^2} + 2\lambda_1\frac{du_1}{dT} + (\lambda_1^2 + 2\overline{H}GX^2 + 2\overline{H}X^2G\cos 4T)u_0 = 0. \quad (46)$$

In view of (44) and (45), the constant term on the left-hand side of (46), when equated to zero to avoid secular terms in u_2, gives

$$\lambda_1 = \pm 2^{1/2}i(\overline{H}G)^{1/2}X(1 + 4a^2\overline{H}GX^2\cos^2\theta)^{1/2}. \quad (47)$$

Using (44) and (47), we can express λ, as determined up to $O(l')$, as

$$\lambda = \pm 2^{1/2}i(\overline{H}G)^{1/2}X(1 + 4a^2\overline{H}GX^2\cos^2\theta)^{1/2}l' + O(l'^2). \quad (48)$$

Since we have taken $m = 0$, we have a state of marginal stability in the absence of any oscillating part ($l' = 0$) in the equilibrium magnetic field. The oscillations in the magnetic field make λ, to $O(l'^2)$, purely imaginary [cf. Eq. (48)], and hence have no stabilizing influence [cf. Eq. (22)].

2. Case (ii): $m = \pm i$

Using (24) and (25) in (23), and equating the terms independent of l' on the left-hand side of (23) to zero, we have

$$\frac{d^2}{dT^2}(e^{\lambda_0 T}u_0) + e^{\lambda_0 T}u_0 = 0. \quad (49)$$

Solving (49) we have

$$e^{\lambda_0 T}u_0 = A\cos T + B\sin T, \quad (50)$$

where A and B are arbitrary constants. In view of (50) we make the choice

$$\lambda_0 = 0. \quad (51)$$

Similarly, equating the coefficient of l' on the left-hand side of (23) to zero, we have, in view of (51),

$$\frac{d^2u_1}{dT^2} + u_1 = -2\lambda_1\frac{du_0}{dT} - 8a\overline{H}GX^2\cos\theta\cos(2T)u_0. \quad (52)$$

Using (50) and (51), and equating the coefficients of $\sin T$ and $\cos T$ on the right-hand side of (52) to zero to avoid secular terms in the solution for u_1, we have

$$A\lambda_1 + 2Ba\overline{H}GX^2\cos\theta = 0,$$
$$B\lambda_1 + 2Aa\overline{H}GX^2\cos\theta = 0.$$

Eliminating A and B from the last two equations we have

$$\lambda_1 = \pm 2a\overline{H}GX^2\cos\theta.$$

Using these values of λ_1, and in view of (51), we obtain to $O(l')$,

$$\lambda = \pm 2(a\overline{H}GX^2\cos\theta)l' + O(l'^2). \quad (53)$$

When the upper sign is taken in (53), since λ is real and positive, we have instability [cf. Eq. (22)]. In a steady equilibrium magnetic field $\mathbf{H} = H(a,0,0)$, $m = \pm i\omega$ represent a disturbance of same frequency as that of the unsteady field $\mathbf{H} = H(a + \epsilon\cos\omega t, \epsilon'\sin\omega t, 0)$. The interaction of this disturbance with the unsteady field leads to instability.

IV. NUMERICAL DISCUSSION OF DISPERSION RELATION FOR MAGNETIC FIELD OSCILLATIONS OF ANY AMPLITUDES

From the theory of the Hill's equation[13,14] we know that

$$\sin^2(\pi\lambda i/2) = \Delta\,(0)\sin^2(\pi\sqrt{\theta_0}/2)\,, \tag{54}$$

where $\Delta\,(0)$ is Hill's determinant

$$\Delta\,(0) = \begin{vmatrix} \cdots & 1 & \dfrac{-\theta_1}{4^2-\theta_0} & \dfrac{-\theta_2}{4^2-\theta_0} & 0 & 0 & \cdots \\ \cdots & \dfrac{-\theta_1}{2^2-\theta_0} & 1 & \dfrac{-\theta_1}{2^2-\theta_0} & \dfrac{-\theta_2}{2^2-\theta_0} & 0 & \cdots \\ \cdots & \dfrac{\theta_2}{\theta_0} & \dfrac{\theta_1}{\theta_0} & 1 & \dfrac{\theta_1}{\theta_0} & \dfrac{\theta_2}{\theta_0} & \cdots \\ \cdots & 0 & \dfrac{-\theta_2}{2^2-\theta_0} & \dfrac{-\theta_1}{2^2-\theta_0} & 1 & \dfrac{-\theta_1}{2^2-\theta_0} & \cdots \\ \cdots & 0 & 0 & \dfrac{-\theta_2}{4^2-\theta_0} & \dfrac{-\theta_1}{4^2-\theta_0} & 1 & \cdots \end{vmatrix}\,. \tag{55}$$

The equation (54) is solved numerically for λ, choosing different values of the physical parameters of the problem.

We first note that the stability analysis can be separated into two distinct classes according to whether disturbances, against which stability is being studied, propagate along a direction normal to the steady component of the equilibrium magnetic field or not (i.e., θ is $\pi/2$ or not).

When $\theta \neq \pi/2$, the equations (19)–(21) can be put in the form

$$\theta_0 = 4[\tilde{H}GX^2(1+\bar{l}^2/2) - XG]\,, \tag{56}$$

$$\theta_1 = 4\tilde{H}G\bar{l}X^2\,, \tag{57}$$

$$\theta_2 = \tilde{H}\bar{l}^2X^2G\,, \tag{58}$$

where

$$\tilde{H} = \bar{H}a^2\cos^2\theta \tag{59}$$

and

$$\bar{l} = l'/(a\cos\theta)\,. \tag{60}$$

In view of the basic dispersion equation (13), and the equations (17)–(21) and (56)–(60), it is clear that the stability problem for a disturbance, with a given nondimensional wave number X, moving along a direction making an angle θ ($\neq \pi/2$) with the direction of the steady component of the magnetic field (direction of x axis), is identical with that for a disturbance of same wave number X, but propagating along the steady component of the magnetic field (so that $\theta = 0$), when $\bar{H}$, l', and the magnitude of the steady component of field a are replaced by $\tilde{H}$, $\bar{l}$ and 1, respectively. Thus the determination of the characteristic exponent λ for any disturbance can be made to depend on the determination of the same for a two-dimensional disturbance. We have therefore studied only two-dimensional disturbances ($\theta = 0$) when $a = 1$, and evaluated λ numerically from Eqs. (54) and (55) for different values of $\bar{H}$, l', and X. The results in a few typical cases are displayed in Fig. 1, where variation of λ_R, the real part of λ, with X are shown for different given values of $\bar{H}$ and l'. We find that there are several critical values of X at which λ_R vanishes. Between any two consecutive critical values of X, the value of λ_R first increases with X to a maximum and then gradually diminishes to zero. The largest of these maximum values in λ_R gives the absolute maximum value in λ_R, the growth rate of instability. From Fig. 1 it is clear that this absolute maximum λ_R, for given values of other physical parameters G, $\bar{H}$, increases very slightly with l', the effective amplitude of the oscillating component of the equilibrium magnetic field.

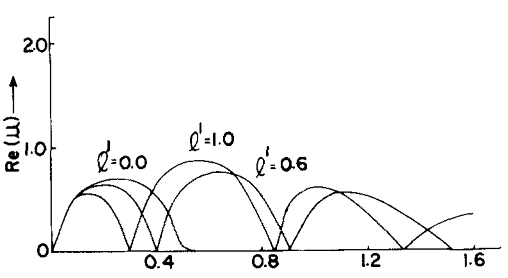

FIG. 1. Variation of Re(μ) with X for different values of l' in the presence of a steady component in the equilibrium magnetic field ($a = 1$) when $H = 2.0$.

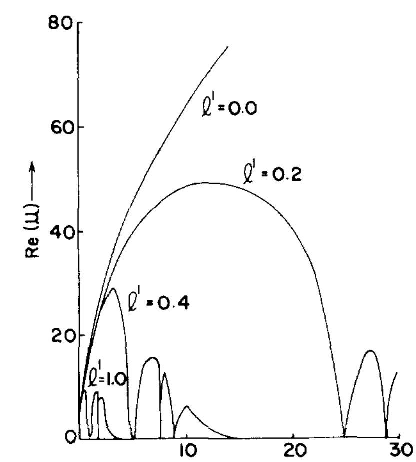

FIG. 2. Variation of Re(μ) with X for different values of l' in the absence of a steady component in the equilibrium magnetic field ($a = 0$) when $H = 2.0$.

We have also computed λ_R when $\theta = \pi/2$, so that the equilibrium magnetic field has no steady effective component ($a \cos\theta = 0$). The variations of λ_R with X, for given values of G, $\overline{H}$, and l' are shown in Fig. 2. It is clear from the graphs in Fig. 2 that the maximum value of λ_R decreases drastically with an increase in l', showing the stabilizing influence of the oscillating magnetic field. A comparison of the values of λ_R from Fig. 1 and Fig. 2 for the same set of values of G, $\overline{H}$ and l' shows that the maximum value of λ_R is smaller when the equilibrium magnetic field has both a steady ($a = 1$), as well as an unsteady part with oscillating components, than when it has no steady part ($a = 0$).

We can give a physical explanation for the stabilizing influence of the oscillating magnetic field in the presence of a steady field. The effective strength of the steady part of the field that influences the stability is $a \cos\theta$, which is the component of this steady field along the propagation direction of disturbances. This has a stabilizing influence except when $\theta = \pi/2$. The unsteady (rotating) part of the field always has a component along the direction of propagation of the disturbance and hence should have a stabilizing influence, except at the moment when the unsteady field is normal to this direction. This stabilizing influence of the unsteady component of the magnetic field will be felt even when $\theta = \pi/2$, so that steady part of the field has no stabilizing influence.

V. CONCLUDING REMARKS

Studying the Rayleigh–Taylor stability of a heavy fluid supported by an oscillating magnetic field we come to the following conclusions.

When l' is small, the steady and the oscillating components of the magnetic field together have a greater stabilizing influence than either of these components has individually. In some special cases the oscillating components of the equilibrium field has either no stabilizing or a destabilizing influence.

When $\epsilon \neq \epsilon'$ the unsteady field, rotating with variable magnitude but constant period about a vertical axis, has the same effect as that of one, discussed in Ref. 8, rotating uniformly with the same period but having the constant magnitude l'.

When $\theta \neq \pi/2$, the stability analysis may be made to depend on the stability discussion for two-dimensional disturbances propagating along the steady component of the equilibrium magnetic field ($\theta = 0$) with $a = 1$ and changed values of H and l'.

When l' is not small and $\theta \neq \pi/2$, numerical determination of λ_R shows that the combined presence of a steady component and oscillating components in the equilibrium magnetic field has a more stabilizing influence than when the steady part is absent ($a = 0$).

ACKNOWLEDGMENTS

The numerical computations reported in this paper were performed in the Computer Center of Tata Institute of Fundamental Research. The authors are thankful to the authorities of the Institute for help. One of us (A.R.N.) is thankful to S. D. Soman, Head, Health Physics Division, Bhaba Atomic Research Centre, for providing facilities and encouragement.

[1]N. A. Krall and A. W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill, New York, 1973).
[2]Lord Rayleigh, Proc. London Math. Soc. **14**, 170 (1883).
[3]G. I. Taylor, Proc. R. Soc. London Ser. A **201**, 192 (1950).
[4]S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability* (Oxford U. P., London, 1961).
[5]M. Kruskal and M. Schwarzschild, Proc. R. Soc. London Ser. A **223**, 348 (1954).
[6]P. G. Drazin, Quart. J. Mech. Appl. Math. **20**, 201 (1967).
[7]B. B. Chakraborty, Phys. Fluids **23**, 464 (1980).
[8]J. Berkowitz, H. Grad, and H. Rubin, in *Proceedings of the Second Conference on Peaceful Uses of Atomic Energy* (United Nations, Geneva, 1959), Vol. 31, p. 177.
[9]E. T. Whittaker, Edinburgh Math. Soc. Proc. **32**, 75 (1914).
[10]A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).
[11]L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, London, 1959).
[12]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, London, 1960).
[13]E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, New York, 1955).
[14]E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).

Symmetry of the complete second-order nonlinear conductivity tensor for an unmagnetized relativistic turbulent plasma

Howard E. Brandt
U.S. Army Electronics Research and Development Command, Harry Diamond Laboratories, Adelphi, Maryland 20783

(Received 21 August 1981; accepted for publication 23 December 1982)

A new exact symmetry is proved for the complete second-order nonlinear conductivity tensor of an unmagnetized relativistic turbulent plasma. The symmetry is not limited to principal parts. If Cerenkov resonance is ignored, the new symmetry reduces to the well-known symmetry related to the Manley–Rowe relations, crossing symmetry, and nondissipation of the principal part of the nonlinear current. Also, a new utilitarian representation for the complete tensor is obtained in which all derivatives are removed and the pole structure is clearly exhibited.

PACS numbers: 52.60. + h, 52.35.Ra, 52.25.Fi, 52.35.Mw

I. INTRODUCTION AND BACKGROUND

The symmetry properties of the nonlinear conductivity tensor are useful in the algebraic reduction of plasmon–plasmon and plasmon–plasma particle interaction probabilities in the theory of turbulent plasmas.[1–5] One of the most well known and useful symmetries involves the second-order nonlinear conductivity tensor $S_{ijl}(k,k_1,k_2)$. For an unmagnetized relativistic turbulent plasma, this tensor is given by[6–10]

$$S_{ijl}(k,k_1,k_2)=e^2\int\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{v_i}{\omega-\mathbf{k}\cdot\mathbf{v}+i\delta}\times\left[(\omega_1-\mathbf{k}_1\cdot\mathbf{v})\frac{\partial}{\partial p_j}+v_jk_{1_m}\frac{\partial}{\partial p_m}\right]\times\left(\frac{\partial}{\partial p_l}+\frac{v_l}{\omega_2-\mathbf{k}_2\cdot\mathbf{v}+i\delta}k_{2_n}\frac{\partial}{\partial p_n}\right)f_p^{R\,(0)},\tag{1}$$

and is related to the Fourier transform of the second-order nonlinear current $\mathbf{j}_k^{(2)}$ by

$$j_{ki}^{(2)}=-e\int\frac{dk_1\,dk_2\,\delta(k-k_1-k_2)}{(\omega_1+i\delta)(\omega_2+i\delta)}\times S_{ijl}(k,k_1,k_2)\,E_{k_1j}\,E_{k_2l}.\tag{2}$$

Here e, $\mathbf{p}$, and $\mathbf{v}$ are particle charge, momentum, and velocity, respectively; $k=(\mathbf{k},\omega)$ is a wave 4-vector; δ is a small imaginary part of the frequency; $f_p^{R\,(0)}$ is the spatially uniform and time-independent background distribution function; $\mathbf{E}_k$ is the Fourier transform of the electric field; and the Einstein sum convention is used throughout. The physical basis for the symmetry is the vanishing dissipation of the nonlinear current in the absence of Cerenkov resonance. The symmetry involves principal parts only, with respect to the wave-particle or Cerenkov resonance denominators in Eq. (1). If the operation of taking principal parts is designated by P, the well-known symmetry can be written as

$$P\,[S_{ijl}(k_1+k_2,k_1,k_2)+S_{ilj}(k_1+k_2,k_2,k_1)]=P\,[S_{jil}(k_1,k_1+k_2,k_2)-S_{jli}(k_1,k_2,k_1+k_2)].\tag{3}$$

To obtain Eq. (3), one first expresses the total energy $W^{(2)}$ dissipated by the nonlinear current $\mathbf{j}^{(2)}$ in the form

$$W^{(2)}=(2\pi)^4\int\mathbf{j}_k^{(2)}\cdot\mathbf{E}_k\,\delta(k+k')\,dk\,dk'.\tag{4}$$

If we substitute Eq. (2) in Eq. (4), perform the integral over k', change the variable k to $-k$, and symmetrize, Eq. (4) becomes

$$W^{(2)}=-\frac{(2\pi)^4}{6}e\int E_{ki}\,E_{k_1j}\,E_{k_2l}\left[\frac{\sigma_{ijl}(-k,k_1,k_2)}{(\omega_1+i\delta)(\omega_2+i\delta)}+\frac{\sigma_{jil}(-k_1,k,k_2)}{(\omega+i\delta)(\omega_2+i\delta)}+\frac{\sigma_{lij}(-k_2,k,k_1)}{(\omega+i\delta)(\omega_1+i\delta)}\right]\times\delta(k+k_1+k_2)\,dk_1\,dk_2\,dk,\tag{5}$$

where $\sigma_{ijl}(k,k_1,k_2)$ denotes the symmetrized form

$$\sigma_{ijl}(k,k_1,k_2)=S_{ijl}(k,k_1,k_2)+S_{ilj}(k,k_2,k_1).\tag{6}$$

If we use the properties of the delta function and simplify, then Eq. (5) becomes[6]

$$W^{(2)}=\frac{(2\pi)^4}{6}e\int dk\,dk_1\,dk_2\times\frac{E_{ki}E_{k_1j}\,E_{k_2l}\delta(k+k_1+k_2)}{(\omega+i\delta)(\omega_1+i\delta)(\omega_2+i\delta)}\times[\omega_1O_{ijl}(k_1,k_2)+\omega_2O_{ilj}(k_2,k_1)],\tag{7}$$

where

$$O_{ijl}(k_1,k_2)=\sigma_{ijl}(k_1+k_2,k_1,k_2)-\sigma_{jil}(-k_1,-k_1-k_2,k_2).\tag{8}$$

According to Eq. (7), the second-order nonlinear current is nondissipative or, equivalently, $W^{(2)}$ is vanishing, if the tensor $O_{ijl}(k_1,k_2)$ given by Eq. (8) is vanishing. However, as will be made clear below, $O_{ijl}(k_1,k_2)$ is in general vanishing only if the small imaginary part δ of the frequency is ignored, or equivalently if the principal values are taken with respect to the resonance denominators of Eq. (1). This is valid only if single wave-particle resonance is ignorable. Thus the second-order nonlinear current is nondissipative provided the principal part P of $O_{ijl}(k_1,k_2)$ in Eq. (8) is vanishing, namely

$$PO_{ijl}(k_1,k_2)=0,\tag{9}$$

and also provided the fields and particle distribution are such that Cerenkov resonance is ignorable. Substituting Eq. (8) in Eq. (9), we obtain

$$P\sigma_{ijl}(k_1+k_2,k_1,k_2)=P\sigma_{jil}(-k_1,-k_1-k_2,k_2), \quad (10)$$

or, equivalently, using Eq. (6) in Eq. (10), we obtain

$$P\left[S_{ijl}(k_1+k_2,k_1,k_2)+S_{ilj}(k_1+k_2,k_2,k_1)\right]$$
$$=P\left[S_{jil}(-k_1,-k_1-k_2,k_2)+S_{jli}(-k_1,k_2,-k_1-k_2)\right]. \quad (11)$$

Furthermore, by Eq. (1), it immediately follows that

$$PS_{jil}(-k_1,-k_1-k_2,k_2)=PS_{jil}(k_1,k_1+k_2,k_2) \quad (12)$$

and

$$PS_{jli}(-k_1,k_2,-k_1-k_2)=-PS_{jli}(k_1,k_2,k_1+k_2). \quad (13)$$

Substituting Eqs. (12) and (13) in Eq. (11), we obtain Eq. (3). Equation (3) is the well-known symmetry, involving principal parts only, of the second-order nonlinear conductivity tensor. Because it ignores wave-particle resonance which results in dissipation, it is frequently at best an approximation to use this symmetry to describe a real physical nonlinear current.

To compare the symmetry Eq. (3) with the current literature, Eqs. (12) and (13) may be used to write Eq. (3) in the form given by Eq. (10). The latter is equivalent to Eq. (10.86) of Ref. 5, namely,

$$P\alpha_{ijl}(k_1+k_2,k_1,k_2)=P\alpha_{jil}(-k_1,-k_1-k_2,k_2). \quad (14)$$

Note that in this form the first written argument of the nonlinear conductivity tensor is equal to the sum of the other arguments. For purposes of later comparison with the new result in this paper, we note also the following: Since by Eq. (1)

$$PS_{jil}(k,k_1,k_2)=PS_{jil}(k,k_1,-k_2) \quad (15)$$

and

$$PS_{jli}(k_1,k_2,k_1+k_2)=-PS_{jli}(k_1,-k_2,k_1+k_2), \quad (16)$$

then by substituting Eqs. (15) and (16) in Eq. (3), we obtain

$$P\sigma_{ijl}(k_1+k_2,k_1,k_2)=P\sigma_{jil}(k_1,k_1+k_2,-k_2). \quad (17)$$

Equations (3), (10), (14), and (17) are clearly all equivalent forms for the symmetry involving principal parts ony.

Expressed in terms of the pure longitudinal nonrelativistic conductivity S_{k,k_1,k_2}, the symmetry Eq. (17) is given by Eq. (2.83) of Ref. 1, namely,

$$P\left[(1/\omega_2)(S_{k_2,k_1+k_2,-k_1}+S_{k_2,-k_1,k_1+k_2})\right]$$
$$=P\{-[1/(\omega_1+\omega_2)](S_{-k_1-k_2,-k_1,-k_2}+S_{-k_1-k_2,-k_2,-k_1})\}, \quad (18)$$

where the implicit principal value operation P with respect to resonance denominators is here made explicit. Equation (18) follows from Eq. (17) as follows. First,

$$S_{k,k_1,k_2}=-e\frac{k_i}{|\mathbf{k}|}\frac{k_{1j}}{|\mathbf{k}_1|}\frac{k_{2l}}{|\mathbf{k}_2|}\frac{1}{\omega_1\omega_2}\times S_{ijl}(k,k_1,k_2). \quad (19)$$

From Eqs. (6) and (19) it follows that

$$(1/\omega_2)(S_{k_2,k_1+k_2,-k_1}+S_{k_2,-k_1,k_1+k_2})=-e\frac{(k_{1j}+k_{2j})\,k_{1l}k_{2i}\,\sigma_{ijl}(k_2,k_1+k_2,-k_1)}{|\mathbf{k}_1+\mathbf{k}_2|\,|\mathbf{k}_1|\,|\mathbf{k}_2|(\omega_1+\omega_2)\,\omega_1\omega_2} \quad (20)$$

and

$$-\frac{1}{\omega_1+\omega_2}(S_{-k_1-k_2,-k_1,-k_2}+S_{-k_1-k_2,-k_2,-k_1})=-e\frac{(k_{1j}+k_{2j})\,k_{1l}k_{2i}\,\sigma_{jil}(-k_1-k_2,-k_2,-k_1)}{|\mathbf{k}_1+\mathbf{k}_2|\,|\mathbf{k}_1|\,|\mathbf{k}_2|(\omega_1+\omega_2)\,\omega_1\omega_2}. \quad (21)$$

Also by Eqs. (1) and (6)

$$P\sigma_{ijl}(-k,-k_2,-k_1)=P\sigma_{ijl}(k,k_2,k_1). \quad (22)$$

Next writing Eq. (22) for $k=k_1+k_2$ and then substituting Eq. (17) in the right-hand side, we conclude that

$$P\sigma_{ijl}(-k_1-k_2,-k_2,-k_1)=P\sigma_{jil}(k_2,k_1+k_2,-k_1). \quad (23)$$

Finally, substituting Eq. (23) in the right-hand side of Eq. (21) and comparing with Eq. (20), we obtain Eq. (18).

Tsytovich and Akopyan used the relativistic symmetry Eq. (3) in obtaining an expression for the collective bremsstrahlung probability in a relativistic weakly turbulent beam-plasma system.[3] A nonlinear bremsstrahlung associated with the three-plasmon dynamic polarization vertex was shown to make an important contribution to the bremsstrahlung probability.[2,3] The connection of the three-plasmon dynamic polarization vertex to the collective bremsstrahlung process had been recognized[11] for some time previous to the difficult and explicit calculations of Tsytovich and Akopyan. In vacuum quantum electrodynamics, because of Furry's theorem,[12] three-photon vertices are vanishing; however, the three-plasmon vertex in a medium is not. Near the plasma frequency, the standard Bethe–Heitler cross section[13] is substantially modified.[2,3] Tsytovich and Akopyan used the collective bremsstrahlung probability to establish conditions for the occurrence of a collective bremsstrahlung instability.[2,3,14–18] The three-plasmon dynamic polarization vertex can be expressed in terms of the second-order nonlinear conductivity tensor. The usefulness of the nonlinear conductivity tensor in calculating the conversion of plasma waves into electromagnetic waves had also been recognized for some time.[19–23] The second-order nonlinear conductivity enters the theory through the bremsstrahlung recoil force $F_{\rm dp}^{(1)}$ on the lowest order nonlinear dynamic polarization charge of a relativistic test particle. This force can be shown to be given by[3,24,25]

$$\mathbf{F}_{\rm dp}^{(1)}=-\lim_{t\to\infty}\frac{(2\pi)^4}{t}\sum_s e_s\times\int\frac{dk\,dk_1\,dk_2\,\delta(k+k_1+k_2)}{(\omega+i\delta)(\omega_1+i\delta)(\omega_2+i\delta)}\times E_{ki}\,E_{k_1j}\,E_{k_2l}\,\mathbf{k}\,S_{ijl}^{(s)}(-k,k_1,k_2), \quad (24)$$

where t denotes time and s is a species label. The second-order nonlinear conductivity tensor also enters through the lowest-order field $E_{{\rm dp}k}^{(1)}$ produced by the dynamic polarization current which is induced by the test particle participating in the bremsstrahlung process This field can be shown to be given by[3,26]

$$E^{(1)}_{\mathrm{dp}kn} = \frac{i}{2(\omega + i\delta)} G_{nm}(k) \sum_s e_s \times \int \frac{dk_1\, dk_2}{(\omega_1 - i\delta)(\omega_2 - i\delta)} \delta(k + k_1 + k_2) \times [S^{(s)}_{mjl}(k, -k_1, -k_2) + S^{(s)}_{mlj}(k, -k_2, -k_1)] \times E^*_{k_1 j} E^*_{k_2 l}, \tag{25}$$

where $G_{nm}(k)$ is the linear photon Green's function. Other than the omission of an overall factor of $-2e_s$ and complex conjugation signs in the integrand in Eq. (22) of Ref. 3, Eqs. (24) and (25) differ from Eqs.(18) and (22) there only because of differing Fourier transform and normalization conventions and choice of units. Also, in Ref. 3 Cerenkov resonance is ignored and principal parts are implicitly taken, whereas in Eqs. (24) and (25) they are not.

Akopyan and Tsytovich stated that they used the property that "the nonlinear current in a plasma is not of a dissipative nature" to algebraically reduce the bremsstrahlung recoil force.[3] They then obtained the collective bremsstrahlung probability by comparison with another expression for this force obtained from the particle balance equations.[3,27,28] They ignored wave-particle Cerenkov resonance and used the symmetry property given by Eqs. (3), (10), (14), and (17) above. For example, in obtaining Eq. (28) of Ref. 3 and in the process of algebraically reducing the lowest-order recoil force on the dynamic polarization charge of a test particle, there occur, among others, terms involving $P\sigma_{ilj}(k,k+k_1,-k_1)$ and ones involving $P\sigma_{lij}(k+k_1,k,k_1)$. By using the symmetry equation (17), these terms can be conveniently combined. This results in considerable simplification and leads to a simple set of diagrammatic rules for calculating collective bremsstrahlung matrix elements. As already stated, the use of this symmetry is only physically justified if the field and particle distributions are such that resonant single wave-particle interactions may be ignored, or equivalently if the nonprincipal parts of the resonance denominators have vanishing support.

Tsytovich ignored Cerenkov resonance on the ground that, when it occurs, Cerenkov radiation predominates over bremsstrahlung to such an extent that the bremsstrahlung is of no interest.[2] However, it is desirable to have a more quantitative evaluation of the effect of Cerenkov resonance on the photonic growth rate due to the collective bremsstrahlung instability. It is not quantitatively known to what extent the Cerenkov and bremsstrahlung radiation mechanisms can interfere with each other. More speculatively, the feasibility of Cerenkov–bremsstrahlung radiative synergisms might be explored.

In the present work a new exact symmetry not limited to the principal parts is established for the complete second-order nonlinear conductivity tensor for an unmagnetized relativistic turbulent plasma. The symmetry is[8,10]

$$S_{ijl}(k_1+k_2,k_1,k_2) + S_{ilj}(k_1+k_2,k_2,k_1) = S_{jil}(k_1,k_1+k_2,k_2) - S_{jli}(k_1,k_2,k_1+k_2). \tag{26}$$

The antisymmetrization in (i,k_1+k_2) and (l,k_2) appearing on the right-hand side is noteworthy. If principal parts only are taken, then Eq. (26) becomes the well-known symmetry equation (3) or equivalently Eqs. (10), (14), and (17).

The part of this symmetry that involves principal parts only expresses the fact that there is no nonlinear absorption from the principal part of the second-order nonlinear conductivity. The physical significance of the nonprincipal part of the new symmetry equation (26) is obscure and is still under investigation. However, its mathematical basis is indisputable, as will be shown here. One possible use of the new symmetry will be in the algebraic reduction of the dynamic polarization force on a test particle and of associated plasmon–plasma particle and plasmon–plasmon interaction probabilities including Cerenkov resonance.

It is shown below that the new symmetry can be rewritten in a more convenient form, namely,

$$\sigma_{ijl}(k_1+k_2,k_1,k_2) = \sigma_{jil}(k_1,k_1+k_2,-k_2) + \Delta_{lji}(-k_2,-k_1-k_2,k_1), \tag{27}$$

where

$$\Delta_{lji}(k,k_1,k_2) = -2\pi i e \int [d^3\mathbf{p}/(2\pi)^3]\, v_l \times \delta(\omega - \mathbf{k}\cdot\mathbf{v})\, \Lambda_{ij}(k_2,k_1)\, \mathbf{k}\cdot\nabla_p f^{R(0)}_p. \tag{28}$$

Here, $\Lambda_{ij}(k_2,k_1)$ is the matrix element for the Compton conversion of a photon with wave vector k_1 into one with k_2 that occurs when the photon scatters off a particle with velocity v, namely[3,25,28]

$$\Lambda_{ij}(k_2,k_1) = \frac{e}{\gamma m}\left[\delta_{ij} + \frac{v_i k_{1j} - v_j k_{2i}}{\omega_1 - \mathbf{k}_1\cdot\mathbf{v} - i\delta} - \frac{v_i v_j}{(\omega_1 - \mathbf{k}_1\cdot\mathbf{v} - i\delta)^2}\left(\mathbf{k}_1\cdot\mathbf{k}_2 - \frac{\omega_1\omega_2}{c^2}\right)\right]. \tag{29}$$

Comparing Eq. (27) with Eq. (17), the modification due to the nonprincipal part is manifest. Note also that in Eq. (27) the new symmetry has been written so that the first written argument in each term is equal to the sum of the other arguments. Since, as will be shown, $\Delta_{lji}(-k_2,-k_1-k_2,k_1)$ arises from nonprincipal parts of resonance denominators, it is to be understood implicitly that

$$P\Delta_{lji}(-k_2,-k_1-k_2,k_1) = 0. \tag{30}$$

Thus, if we take the principal part of the new symmetry Eq. (27) and use Eq. (30), Eq. (27) becomes the well-known symmetry equation (17), or, equivalently, Eqs. (3), (10), and (14). If the exact symmetry equation (26), or, equivalently, Eq. (27), is used in algebraic reductions rather than Eq. (17) (which was used in Ref. 3), then additional terms will appear, involving Compton conversion through the matrix element Δ_{lji}. This must take account of wave-particle resonance and the fact that the complete nonlinear current is not in general nondissipative. The use of the new symmetry will hopefully facilitate deeper understanding of collective radiative instability and the dissipative properties of the complete nonlinear current. It is also hoped that the new symmetry will find applications in other areas of plasma turbulence theory and nonequilibrium plasma kinetic theory.

The symmetry of the nonlinear conductivity tensor, involving principal parts only, has been investigated also in other work. The nonlinear conductivity is simply related to the nonlinear response or wave coupling coefficients, the symmetries of which are often discussed in the literature. For nonrelativistic weakly turbulent plasma and waves of arbitrary polarization, the symmetry involving principal parts was established long ago.[29,30] In the nonlinear response, half residues due to Cerenkov resonance were neglected. This was justified on the grounds that such resonances lead to dominating contributions to the ordinary linear response. However, for purposes of some applications this argument is weak, such as in the analysis of radiative instability or under conditions of slowly converging perturbation series. For relativistic magnetized plasmas, the symmetry involving principal parts has been demonstrated to all orders.[31–34] Recently an elegant representation of the nonlinear current response, in which the symmetry is explicitly evident in each order, has been constructed by means of a powerful coordinate-free differential geometric formulation.[32–34]

Others have also investigated the relationship between the principal value symmetry and the vanishing total energy dissipated by the nonlinear current.[4,6,34] The generalization of the symmetry to arbitrary order, including magnetization, but ignoring Cerenkov resonance, was shown to imply that there is no nonlinear absorption from the principal part of the nonlinear conductivity to any order, and therefore that all absorption is due to wave-particle interactions.[34] It is, of course, well known that plasmon damping does occur and that because of this the very concept of elementary excitation can only be introduced approximately.[4] The resonating particles must be small in number. This is equivalent to the condition of negligible imaginary part in resonance denominators, or equivalently that $\operatorname{Re}\omega \gg \operatorname{Im}\omega = \delta$. Of course, the very concept of quanta of the electromagnetic field is meaningful only for almost monochromatic waves, for which the latter inequality is satisfied. Tsytovich also used the nondissipative argument to obtain the symmetry, which he used in the approximate reduction of a dispersion relation for a wave produced in a three-plasmon interaction.[4] The nondissipative condition made it possible to establish relationships not only between moduli of the nonlinear response, as determined by the equality between the probabilities of various processes, but also between the real and imaginary parts. The approximate character of this procedure was emphasized in a footnote on p. 53 of Ref. 4.

The relationship of the principal value symmetry to crossing symmetry in three-plasmon interactions has also been investigated.[4,31,35] For example, in the case of an isotropic nonrelativistic plasma, the symmetry can be used to relate the second-order nonlinear response describing the coalescence of a Langmuir wave and an ion sound wave to form a transverse wave to that describing the coalescence of a transverse wave and an ion sound wave to form a Langmuir wave. The inverse processes can also be related.

The principal value symmetry has also been related to symmetries of Poisson brackets in a perturbation-theoretic Hamiltonian formulation of the nonlinear response.[29,31] The symmetry follows from the antisymmetry property and the Jacobi identity satisfied by Poisson brackets involving the Hamiltonian.

In coherent three-wave interactions and the weak turbulence equations, it follows, from the symmetry of the principal part of the coupling coefficients, that wave energy and momentum are approximately conserved, and the Manley–Rowe relations of action conservation obtain.[29,32–34,36–51] Some effects of background inhomogeneities on the symmetry property have been addressed.[46,47] The Manley–Rowe relations for action conservation have been shown to be maintained even in the presence of a high-frequency external electric field strongly modifying the nonlinear response.[36,48] To a limited extent, symmetry-breaking effects associated with violation of the Manley–Rowe relations have also been addressed.[52,53] Nonlinear Landau damping[43,54–56] can be greatly modified when the Manley–Rowe relations are violated, and rapid dissipation of total energy in electron plasma waves into thermal energy can occur.[53]

In Sec. II of this paper, a new representation of the second-order nonlinear conductivity tensor of an unmagnetized relativistic turbulent plasma is presented. All derivatives are removed and the pole structure is clearly exhibited. In Sec. III, the new symmetry equation (26) is proved using the new representation. In Sec. IV, the alternative form equation (27) for the new symmetry is obtained. In Sec. V, the results are briefly summarized.

II. NEW REPRESENTATION OF SECOND-ORDER NONLINEAR CONDUCTIVITY TENSOR

In this section a new representation is obtained for the second-order nonlinear conductivity tensor of an unmagnetized relativistic turbulent plasma. Integrating equation (1) by parts and dropping surface terms, then using the relativistic kinematic relations, performing the differentiations, and combining terms, we obtain[6,8,57]

$$\begin{aligned} S_{ijl}(k,k_1,k_2) &= e^2c^2\int\frac{d^3\mathbf{p}}{(2\pi)^3}f_p^{R(0)}\frac{1}{\epsilon^2} \\ &\times\Big[\frac{\alpha_1+\overline{\alpha}_1\Omega_1}{\Omega+i\delta}+\frac{\alpha_2+\overline{\alpha}_2\Omega_1}{(\Omega+i\delta)(\Omega_2+i\delta)} \\ &+\frac{\alpha_3+\overline{\alpha}_3\Omega_1}{(\Omega+i\delta)(\Omega_2+i\delta)^2}+\frac{\alpha_4+\overline{\alpha}_4\Omega_1}{(\Omega+i\delta)^2} \\ &+\frac{\alpha_5+\overline{\alpha}_5\Omega_1}{(\Omega+i\delta)^2(\Omega_2+i\delta)}+\frac{\alpha_6+\overline{\alpha}_6\Omega_1}{(\Omega+i\delta)^2(\Omega_2+i\delta)^2} \\ &+\frac{\alpha_7+\overline{\alpha}_7\Omega_1}{(\Omega+i\delta)^3}+\frac{\alpha_8+\overline{\alpha}_8\Omega_1}{(\Omega+i\delta)^3(\Omega_2+i\delta)}\Big]. \end{aligned} \tag{31}$$

Here the relativistic single-particle energy is ϵ, given by

$$\epsilon=(m^2c^4+p^2c^2)^{1/2}, \tag{32}$$

and the quantities Ω, Ω_1, and Ω_2 are given by

$$\{\Omega,\Omega_1,\Omega_2\}=\{\omega-\mu,\omega_1-\mu_1,\omega_2-\mu_2\}, \tag{33}$$

where

$$\{\mu,\mu_1,\mu_2\} = \{\mathbf{k}\cdot\mathbf{v},\mathbf{k}_1\cdot\mathbf{v},\mathbf{k}_2\cdot\mathbf{v}\}. \tag{34}$$

Also $\{\alpha_n, n=1,8\}$ and $\{\overline{\alpha}_n, n=1,8\}$ are complicated tensor polynomials in the components of $\mathbf{v}$, $\mathbf{k}$, $\mathbf{k}_1$, and $\mathbf{k}_2$. They are simply a renaming of the coefficients C_n of Ref. 6. For notational convenience, the tensor indices of the $\alpha_n \equiv \alpha_{nijl}(k,k_1,k_2)$ and $\overline{\alpha}_n \equiv \overline{\alpha}_{nijl}(k,k_1,k_2)$ are suppressed. Explicitly, the α_n are given by

$$\alpha_1 = (c^2k_{1i} - \mu_1 v_i)\,\delta_{jl} - (c^2k_{1l} - u_1v_l)\,\delta_{ij} - \mu_1 v_j\,\delta_{il} - 2k_{1i}v_jv_l + 3c^{-2}\mu_1v_iv_jv_l, \tag{35}$$

$$\alpha_2 = (\mu_1\mu_2 - c^2\mathbf{k}_1\cdot\mathbf{k}_2)\,v_l\delta_{ij} + 4c^{-2}\mu_1\mu_2v_iv_jv_l + c^2k_{1i}v_jk_{2l} + c^2k_{1i}k_{2j}v_l - \mu_1k_{2i}v_jv_l - \mu_1v_ik_{2j}v_l - \mu_1v_iv_jk_{2l} - 3\mu_2k_{1i}v_jv_l, \tag{36}$$

$$\alpha_3 = (c^2k_2^2 - \mu_2^2)\,k_{1i}v_jv_l + (c^{-2}\mu_1\mu_2^2 - \mu_1k_2^2)\,v_iv_jv_l, \tag{37}$$

$$\alpha_4 = (c^2\mathbf{k}\cdot\mathbf{k}_1 - \mu\mu_1)(v_i\,\delta_{jl} + v_j\,\delta_{il}) - c^2v_ik_jk_{1l} + \mu_1v_ik_jv_l - 2\mu_1v_iv_jk_l - \mu k_{1i}v_jv_l + c^2k_{1i}v_jk_l + (5c^{-2}\mu\mu_1 - 3\mathbf{k}\cdot\mathbf{k}_1)\,v_iv_jv_l, \tag{38}$$

$$\alpha_5 = (-c^2\mathbf{k}_1\cdot\mathbf{k}_2 + \mu_1\mu_2)\,v_ik_jv_l + (c^2\mathbf{k}\cdot\mathbf{k}_1 - \mu\mu_1)\,v_iv_jk_{2l} + (c^2\mathbf{k}\cdot\mathbf{k}_1 - \mu\mu_1)\,v_ik_{2j}v_l + (c^2\mathbf{k}\cdot\mathbf{k}_1 - \mu\mu_1)\,k_{2i}v_jv_l + (c^2\mathbf{k}\cdot\mathbf{k}_2 - \mu\mu_2)\,k_{1i}v_jv_l + 2(3c^{-2}\mu\mu_1\mu_2 - 2\mathbf{k}\cdot\mathbf{k}_1\mu_2 - \mathbf{k}\cdot\mathbf{k}_2\,\mu_1)\,v_iv_jv_l, \tag{39}$$

$$\alpha_6 = (c^2\mathbf{k}\cdot\mathbf{k}_1k_2^2 - \mu\mu_1\,k_2^2 - \mathbf{k}\cdot\mathbf{k}_1\,\mu_2^2 + c^{-2}\mu\mu_1\,\mu_2^2)\,v_iv_jv_l, \tag{40}$$

$$\alpha_7 = 2(c^2\mathbf{k}\cdot\mathbf{k}_1 - \mu\mu_1)\,v_iv_jk_l - 2(\mathbf{k}\cdot\mathbf{k}_1\,\mu - c^{-2}\mu^2\mu_1)\,v_iv_jv_l, \tag{41}$$

$$\alpha_8 = 2(c^2\mathbf{k}\cdot\mathbf{k}_1\mathbf{k}\cdot\mathbf{k}_2 - \mathbf{k}\cdot\mathbf{k}_2\,\mu\mu_1 - \mathbf{k}\cdot\mathbf{k}_1\,\mu\mu_2 - c^{-2}\mu^2\mu_1\,\mu_2)\,v_iv_jv_l. \tag{42}$$

The $\overline{\alpha}_n$ are given by

$$\overline{\alpha}_1 = -v_i\,\delta_{jl} - v_j\,\delta_{il} - v_l\,\delta_{ij} + 3c^{-2}v_iv_jv_l, \tag{43}$$

$$\overline{\alpha}_2 = (c^2k_{2l} - 2\mu_2v_l)\,\delta_{ij} - k_{2i}v_jv_l - v_ik_{2j}v_l - v_iv_jk_{2l} + 4c^{-2}\mu_2v_iv_jv_l, \tag{44}$$

$$\overline{\alpha}_3 = (c^2k_2^2 - \mu_2^2)\,v_l\,\delta_{ij} + (c^{-2}\mu_2^2 - k_2^2)\,v_iv_jv_l, \tag{45}$$

$$\overline{\alpha}_4 = (c^2k_l - \mu v_l)\,\delta_{ij} + (c^2k_j - \mu v_j)\,\delta_{il} - \mu v_i\delta_{jl} - 2v_iv_jk_l - 2v_ik_jv_l + 5c^{-2}\mu v_iv_jv_l, \tag{46}$$

$$\overline{\alpha}_5 = (c^2\mathbf{k}\cdot\mathbf{k}_2 - \mu\mu_2)\,v_l\,\delta_{ij} + 2(3c^{-2}\mu\mu_2 - \mathbf{k}\cdot\mathbf{k}_2)\,v_iv_jv_l + c^2k_{2i}k_jv_l + c^2v_ik_jk_{2l} - \mu v_ik_{2j}v_l - \mu v_iv_jk_{2l} - \mu k_{2i}v_jv_l - 3\mu_2\,v_ik_jv_l, \tag{47}$$

$$\overline{\alpha}_6 = (c^2k_2^2 - \mu_2^2)\,v_ik_jv_l + (c^{-2}\mu\mu_2^2 - \mu k_2^2)\,v_iv_jv_l, \tag{48}$$

$$\overline{\alpha}_7 = 2c^2v_ik_jk_l - 2\mu v_ik_jv_l - 2\mu v_iv_jk_l + 2c^{-2}\mu^2v_iv_jv_l, \tag{49}$$

$$\overline{\alpha}_8 = 2(c^2\mathbf{k}\cdot\mathbf{k}_2 - \mu\mu_2)\,v_ik_jv_l - 2(\mathbf{k}\cdot\mathbf{k}_2\,\mu - c^{-2}\mu^2\mu_2)\,v_iv_jv_l. \tag{50}$$

The new representation given by Eq. (31) contains no derivatives, and the explicit pole structure is clearly exhibited. The coefficients are numerous; however, they are simple polynomials. This representation should be useful for numerical and analytical computations in nonlinear plasma kinetic theory and plasma turbulence theory. By comparison, the representation of Refs. 33 and 34 is mathematically more elegant but less explicit and less directly utilitarian. In Sec. III, the new representation equation (31) is used to prove the new symmetry equation (26).

III. THE NEW SYMMETRY

In this section, the new representation equation (31) of the nonlinear conductivity tensor of an unmagnetized relativistic turbulent plasma is employed to prove the new exact symmetry equation (26) of the complete tensor. Another closely related symmetry has also been obtained by this method[6,8,9,57]; however, the symmetry holds outside the domain of wave vector space in which the nonlinear current equation (2) is defined.[10]

First replacing k by $k_1 + k_2$ in Eq. (31) and then combining terms, we obtain

$$S_{ijl}(k_1+k_2,k_1,k_2) + S_{ilj}(k_1+k_2,k_2,k_1) = e^2c^2\int\frac{d^3\mathbf{p}}{(2\pi)^3}\,\frac{f_p^{R(0)}\sum_{n=1}^{24}\beta_n\Pi_n}{\epsilon^2(\Omega_1+\Omega_2+i\delta)^3(\Omega_1+i\delta)^3(\Omega_2+i\delta)^3}, \tag{51}$$

where

$$\{\Pi_n,\ n=1,24\} = \{\Omega_1\Omega_2^4,\Omega_1\Omega_2^5,\Omega_1\Omega_2^6,\Omega_1^2\Omega_2^3,\Omega_1^2\Omega_2^4,\Omega_1^2\Omega_2^5, \Omega_1^2\Omega_2^6,\Omega_1^3\Omega_2^2,\Omega_1^3\Omega_2^3,\Omega_1^3\Omega_2^4,\Omega_1^3\Omega_2^5,\Omega_1^3\Omega_2^6, \Omega_1^4\Omega_2,\Omega_1^4\Omega_2^2,\Omega_1^4\Omega_2^3,\Omega_1^4\Omega_2^4,\Omega_1^4\Omega_2^5,\Omega_1^5\Omega_2, \Omega_1^5\Omega_2^2,\Omega_1^5\Omega_2^3,\Omega_1^5\Omega_2^4,\Omega_1^6\Omega_2,\Omega_1^6\Omega_2^2,\Omega_1^6\Omega_2^3\} \tag{52}$$

and where $\{\beta_n, n=1,24\}$ are tensor polynomials in the components of $\mathbf{v}$, $\mathbf{k}_1$, and $\mathbf{k}_2$. The expressions for the β_n are complicated and are given by Eqs. (27)–(50) of Ref. 8. By similarly reducing $S_{jil}(k_1,k_1+k_2,k_2) - S_{jli}(k_1,k_2,k_1+k_2)$ and comparing the result with Eq. (51) term by term, the new symmetry relation given by Eq. (26) follows.

IV. CONVENIENT FORM FOR THE NEW SYMMETRY

In this section the new symmetry equation (26) is rewritten in the form (27), which is both convenient for ready comparison with the well-known symmetry involving principal parts only, Eq. (17), and directly useful for purposes of algebraic reduction of mathematical expressions occurring in problems in plasma turbulence theory and nonlinear plasma physics.

First from Eq. (1) it is clear that

$$S_{ijl}(k,k_1,k_2) = -S_{ijl}(k,-k_1,k_2). \tag{53}$$

If Eq. (53) is used in Eq. (26), the latter can be rewritten as

$$\sigma_{ijl}(k_1+k_2,k_1,k_2) = S_{jil}(k_1,k_1+k_2,k_2) + S_{jli}(k_1,-k_2,k_1+k_2). \tag{54}$$

Next adding and subtracting $S_{jil}(k_1,k_1+k_2,-k_2)$ to Eq. (54), and using Eqs. (1) and (6) and simplifying, we obtain

$$\sigma_{ijl}(k_1+k_2,k_1,k_2) = \sigma_{jil}(k_1,k_1+k_2,-k_2) + \bar{\Delta}_{ijl}(k_1,k_2), \tag{55}$$

where

$$\bar{\Delta}_{ijl}(k_1,k_2) = -2\pi i e^2 \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{v_j}{(\omega_1 - \mathbf{k}_1\cdot\mathbf{v} + i\delta)} \times\{[\omega_1+\omega_2-(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{v}]\,\delta_{im} + v_i(k_{1m} + k_{2m})\}\nabla_{p_m}\,\delta(\omega_2-\mathbf{k}_2\cdot\mathbf{v})\,v_l\mathbf{k}_2\cdot\nabla_p f_p^{R(0)}. \tag{56}$$

Integrating Eq. (56) by parts with respect to the first gradient and simplifying, we reduce Eq. (56) to

$$\bar{\Delta}_{ijl}(k_1,k_2) = 2\pi i e \int \frac{d^3\mathbf{p}}{(2\pi)^3} v_l\,\delta(\omega_2-\mathbf{k}_2\cdot\mathbf{v}) \times \Lambda_{ij}(k_1,-k_1-k_2)\,\mathbf{k}_2\cdot\nabla_p f_p^{R(0)}, \tag{57}$$

where $\Lambda_{ij}(k_1,k_2)$ is defined by Eq. (29) and is the matrix element for Compton conversion. Using Eqs. (55) and (57), we can rewrite Eq. (55) as Eq. (27) with $\Delta_{lji}(k,k_1,k_2)$ defined by Eq. (28), in terms of the Compton conversion matrix element. Equation (27) is a convenient and potentially useful form for the new symmetry, as has already been discussed in the Introduction.

V. CONCLUSION

In conclusion, then, a new exact symmetry equation (26) has been proved for the complete second-order nonlinear conductivity tensor of an unmagnetized relativistic weakly turbulent plasma. A new utilitarian representation (31) for the tensor, in which all derivatives are removed and the pole structure is clearly exhibited, was employed in the proof. The symmetry is not limited to principal parts and includes Cerenkov resonance. Noteworthy is the fact that the new symmetry relates in a natural way the customary symmetrization of $S_{ijl}(k_1+k_2,k_1,k_2)$ in (j,k_1) and (l,k_2) to a curious antisymmetrization of $S_{jil}(k_1,k_1+k_2,k_2)$ in (i,k_1+k_2) and (l,k_2). The symmetry can also be written in a convenient equivalent form, Eq. (27), in which the modifications due to nonprincipal parts of wave-particle resonance denominators are manifest. An explicit nonprincipal part of the symmetry is expressed simply in terms of the matrix element for Compton conversion. The principal part of the symmetry reduces immediately to the well-known symmetry equation (17) that applies when resonant wave-particle interactions are negligible, the Manley–Rowe relations obtain, and the nonlinear current is nondissipative. Although the physical meaning for the principal part of the new symmetry is well understood, that for the complete symmetry remains obscure. The possible relevance of the symmetry to the algebraic reduction and interpretation of transition probabilities for collective radiation processes in nonequilibrium relativistic beam-plasma systems is currently being investigated.[24,26,28] It is hoped that the new symmetry will find other applications in nonlinear plasma physics and the theory of plasma turbulence.

1. V. N. Tsytovich, *Theory of Turbulent Plasma* (Consultants Bureau) (Plenum, New York, 1977), pp. 70–84.
2. V. N. Tsytovich, "Bremsstrahlung of a Relativistic Plasma," Tr. Fiz. Inst. Akad. Nauk SSSR **66**, 191 (1973) [Proc. P. N. Lebedev Phys. Inst. **66**, 199 (1975)].
3. A. V. Akopyan and V. N. Tsytovich, "Bremsstrahlung in a Nonequilibrium Plasma," Fiz. Plazmy **1**, 673 (1975) [Sov. J. Plasma Phys. **1**, 371 (1975)].
4. V. N. Tsytovich, "Kinetic Equations for Elementary Excitations and Nonlinear Interactions Between Waves in a Plasma," Preprint No. 12, Fiz. Inst. Akad. Nauk SSSR, 1968 (CTO/616 Culham Translations Office, 1969).
5. D. B. Melrose, *Plasma Astrophysics* (Gordon and Breach, New York, 1980), Vol. 2, pp. 156–173.
6. H. E. Brandt, "Symmetries of the Nonlinear Conductivity of a Relativistic Turbulent Plasma," HDL-TR-1927, Harry Diamond Laboratories, March 1981.
7. V. N. Tsytovich, "Nonlinear Absorption of Electromagnetic Waves During Resonant Plasma Heating," Fiz. Plazmy **6**, 1105 (1980) [Sov. J. Plasma Phys. **6**, 608 (1980)].
8. H. E. Brandt, "On the Nonlinear Conductivity Tensor for an Unmagnetized Relativistic Turbulent Plasma," HDL-TR-1970, Harry Diamond Laboratories, February 1982.
9. H. E. Brandt, "Exact symmetry of the second-order nonlinear conductivity for a relativistic turbulent plasma," Phys. Fluids **24**, 1760 (1981).
10. H. E. Brandt, "Comment on 'Exact symmetry of the second-order nonlinear conductivity for a relativistic turbulent plasma,' " Phys. Fluids **25**, 1922 (1982).
11. D. F. Dubois, V. Gilinsky, and M. G. Kivelson, "Propagation of Electromagnetic Waves in Plasmas," Phys. Rev. **129**, 2376 (1963).
12. W. Furry, "Symmetry Theorem in the Positron Theory," Phys. Rev. **51**, 125 (1937).
13. H. Bethe and W. Heitler, "On the Stopping of Fast Particles and on the Creation of Positive Electrons," Proc. Roy Soc. London A **146**, 83 (1934).
14. A. V. Akopyan and V. N. Tsytovich, "Transition Bremsstrahlung of Relativistic Particles," Zh. Eksp. Teor. Fiz. **71**, 166 (1976) [Sov. Phys. JETP **44**, 87 (1976)].
15. A. V. Akopyan and V. N. Tsytovich, "Bremsstrahlung of Relativistic Electrons in a Plasma in a Strong Magnetic Field," Zh. Eksp. Teor. Fiz. **72**, 1824 (1977) [Sov. Phys. JETP **45**, 957 (1977)].
16. A. V. Akopyan and V. N. Tsytovich, "Bremsstrahlung of Relativistic Electrons in Plasmas," in *Proceedings of the XIIIth International Conference on Phenomena in Ionized Gases* (Physical Society of German Democratic Republic, Berlin, 1977), p. 917.
17. A. V. Akopyan and V. N. Tsytovich, "Bremsstrahlung Instability of Relativistic Electrons in a Plasma," Astrofizika **13**, 717 (1977) [Astrophysics **13**, 423 (1977)].
18. V. Tsytovich, "Collective Effects in Bremsstrahlung of Fast Particles in Plasmas," Comments, Plasma Phys. Conf. Fusion **4**, 73 (1978).
19. S. A. Kaplan and V. N. Tsytovich, "Plasma Radiation Mechanisms in Astrophysics," Usp. Fiz. Nauk **97**, 77 (1969) [Sov. Phys. Usp. **12**, 42 (1969)].
20. V. N. Tsytovich, "Nonlinear Effects in a Plasma," Usp. Fiz. Nauk **90**, 435 (1966) [Sov. Phys. Usp. **9**, 805 (1967)].
21. V. N. Tsytovich, *Nonlinear Effects in Plasma* (Plenum, New York, 1970).
22. V. N. Tsytovich and A. B. Shvartsburg, "Contribution to the Theory of Nonlinear Interaction of Waves in a Magnetoactive Anisotropic Plasma," Zh. Eksp. Teor. Fiz. **49**, 797 (1965) [Sov. Phys. JETP **22**, 554 (1966)].
23. V. N. Tsytovich and A. B. Shvartsburg, "Nonlinear Wave Interaction in a Plasma in a Strong Magnetic Field," Zh. Tekh. Fiz. **36**, 1431 (1966) [Sov. Phys. Tech. Phys. **11**, 1431 (1967)].
24. H. E. Brandt, "Nonlinear Dynamic Polarization Force on a Relativistic Test Particle in a Nonequilibrium Beam-Plasma System," HDL-TR-1994, Harry Diamond Laboratories, 1982.
25. H. E. Brandt, "Nonlinear Force on an Unpolarized Relativistic Test Particle to Second-Order in the Total Field in a Nonequilibrium Beam-Plasma System," HDL-TR-1995, Harry Diamond Laboratories, 1982.
26. H. E. Brandt, "The Total Field in Collective Bremsstrahlung in a Nonequilibrium Relativistic Beam-Plasma System," HDL-TR-1996, Harry Diamond Laboratories, 1982.
27. H. E. Brandt, "On the Form of the Collective Bremsstrahlung Recoil Force in a Nonequilibrium Relativistic Beam Plasma System," preprint, Harry Diamond Laboratories, 1982 (to be published).
28. H. E. Brandt, "Theoretical Methods in the Calculation of the Bremsstrah-

lung Recoil Force in a Nonequilibrium Relativistic Beam-Plasma System," preprint, Harry Diamond Laboratories, 1982 (to be published).
[29]L. M. Al'tshul' and V. I. Karpman, "The Kinetics of Waves in a Weakly Turbulent Plasma," Zh. Eksp. Teor. Fiz. **47**, 1552 (1964) [Sov. Phys. JETP **20**, 1043 (1965)].
[30]A. A. Galeev, V. I. Karpman, and R. Z. Sagdeev, "Multiparticle Aspects of Turbulent Plasma Theory," Nucl. Fusion **5**, 20 (1965).
[31]D. B. Melrose, "Symmetry Properties of Nonlinear Responses in a Plasma," Plasma Phys. **14**, 1035 (1972).
[32]J. Larsson, "Current responses of all orders in a collisionless plasma. I. General theory," J. Math. Phys. **20**, 1321 (1979).
[33]J. Larsson, "Current responses of all orders in a collisionless plasma. II. Homogeneous plasma," J. Math. Phys. **20**, 1331 (1979).
[34]J. Larsson, "Conductivity Tensors of All Orders in a Collisionless Plasma," J. Plasma Phys. **21**, 519 (1979).
[35]L. M. Kovrizhnykh, "On the Theory of Turbulent Plasma," Zh. Eksp. Teor. Fiz. **48**, 1114 (1965) [Sov. Phys. JETP **21**, 744 (1965)].
[36]A. N. Kaufman and L. Stenflo, "Action Conservation in the Presence of a High-Frequency Field," Plasma Phys. **17**, 403 (1975).
[37]J. M. Manley and H. E. Rowe, "Some General Properties of Nonlinear Elements—Part I. General Energy Relations," Proc. IRE **44**, 904 (1956).
[38]P. A. Sturrock, "Action-Transfer and Frequency-Shift Relations in the Nonlinear Theory of Waves and Oscillations," Ann. Phys. (N.Y.) **9**, 422 (1960).
[39]P. A. Sturrock, "Non-linear Effects in Electron Plasmas," Proc. R. Soc. London A **242**, 277 (1957).
[40]S. Johnston and A. N. Kaufman, "Generalized Ponderomotive Forces and Three-Wave Interaction," in *Plasma Physics, Nonlinear Theory and Experiments*, edited by H. Wilhelmsson (Plenum, New York, 1977), p. 159.
[41]J. Dougherty, "Lagrangian Methods in Plasma Dynamics, I. General Theory of the Method of the Averaged Lagrangian," J. Plasma Phys. **4**, 761 (1970).
[42]J. Galloway and H. Kim, "Lagrangian Approach to Nonlinear Wave Interactions in a Warm Plasma," J. Plasma Phys. **6**, 53 (1971).
[43]R. Z. Sagdeev and A. A. Galeev, *Nonlinear Plasma Theory* (Benjamin, New York, 1969).
[44]B. Coppi, M. N. Rosenbluth, and R. N. Sudan, "Nonlinear Interactions of Positive and Negative Energy Modes in Rarefied Plasmas I and II," Ann. Phys. (N.Y.) **55**, 207, 248 (1969).
[45]J. Weiland and H. Wilhelmsson, *Coherent Non-Linear Interaction of Waves in Plasmas* (Pergamon, New York, 1977).
[46]J. Larsson, "Current responses of first and second order in a collisionless plasma. I. Stationary plasma," J. Math. Phys. **23**, 176 (1982).
[47]J. Larsson, "Current responses of first and second order in a collisionless plasma. II. Three-wave interaction," J. Math. Phys. **23**, 183 (1982).
[48]J. Larsson and L. Stenflo, "Three-Wave Interaction in the Presence of an Oscillating Electric Field," Phys. Lett. A **76**, 396 (1980).
[49]L. Stenflo and J. Larsson, "Three-Wave Coupling Coefficients for Magnetized Plasmas," in *Plasma Physics, Nonlinear Theory and Experiments*, edited by H. Wilhelmsson (Plenum, New York, 1977), p. 152.
[50]L. Stenflo, "Higher-Order Nonlinearities in the Theory of Parametric Decay Instabilities," Phys. Scr. **18**, 5 (1978).
[51]S. Johnston and A. N. Kaufman, "Lie-operator approach to mode coupling in nonuniform plasma," Phys. Rev. Lett. **40**, 1266 (1978).
[52]K. Mima, "Modification of Weak Turbulence Theory Due to Perturbed Orbit Effects, I. General Formulation," J. Phys. Soc. Jpn. **34**, 1620 (1973).
[53]K. Mima, "Modification of Weak Turbulence Theory Due to Perturbed Orbit Effects, II. Nonlinear Landau Damping of Electron Plasma Waves," J. Phys. Soc. Jpn. **35**, 261 (1973).
[54]B. B. Kadomtsev, *Plasma Turbulence* (Academic, New York, 1965).
[55]R. C. Davidson, *Methods in Nonlinear Plasma Theory* (Academic, New York, 1972).
[56]A. I. Akhiezer, I. A. Akhiezer, R. V. Polovin, A. G. Sitenko, and K. N. Stepanov, *Plasma Electrodynamics, Vol. 2. Nonlinear Theory and Fluctuations* (Pergamon, New York, 1975).
[57]H. E. Brandt, "Second-Order Nonlinear Conductivity Tensor for an Unmagnetized Relativistic Turbulent Plasma," in *Plasma Astrophysics, Course and Workshop Organized by the International School of Plasma Physics*, 27 August–7 September 1981, Varenna (Como), Italy (European Space Agency ESA SP-161, Nov. 1981) (also to be published by Pergamon), p. 361.

Shift operator techniques for the classification of multipole-phonon states. XI. Properties of mixed type quadratic product operators in R(7)

J. Van der Jeugt,[a] H. E. De Meyer,[b] and G. Vanden Berghe
Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit-Gent Krijgslaan 281-S9, B-9000 Gent, Belgium

(Received 10 February 1982; accepted for publication 26 March 1982)

Expressions connecting R(3) scalar and nonscalar product operators of the type $P^j_{l+k}O^k_l$ and $O^k_{l+j}P^j_l$ are constructed within the group R(7).

PACS numbers: 63.20.Dj, 02.20.Nq, 02.30.Tb

1. INTRODUCTION

In a set of previous papers[1–10] (to be referred to as I–X) it was evident that operators shifting the eigenvalues l of the R(3) Casimir operator L^2 could play an important role in the classification of multipole-phonon states. The quadrupole-phonon state labelling problem could be completely solved by means of the shift operator technique.[7]

It is known that the symmetry group of the octupole-phonon Hamiltonian is the unitary group in seven dimensions, U(7). Considering the Casimir operators of the subgroups of U(7), only four independent labels specifying its symmetric irreducible representations, which are connected to the considered phonon states, could be deduced in (IV). In IV and V G_2 and R(7) shift operators, P^k_l $(-5\leqslant k\leqslant +5)$ and O^k_l $(-3\leqslant k\leqslant +3)$, respectively, have been introduced. Either the operator P^0_l or O^0_l can be used as a fifth label generating operator. However, since they do not commute, we cannot diagonalize them simultaneously, and thus no set of orthogonal phonon states can be generally constructed such that they are always eigenstates of both O^0_l and P^0_l.

For the G_2 group, expressions connecting R(3) scalar and nonscalar quadratic product operators have been reported (IV,IX). The application of those expressions allowed us to calculate a part of the eigenvalue spectrum of P^0_l (X). For the R(7) group the relations between quadratic product operators of the R(3) scalar and nonscalar type $O^j_{l+k}O^k_l$ (V and VI) also contained the P^k_l operators. Therefore the P^0_l-eigenvalues and the P^k_l action on states had to be known before one could start with calculating the O^0_l eigenvalues. To evaluate the O^0_l eigenvalues when l-degeneracy occurs, the relations between product operators of the type $P^k_{l+j}O^j_l$ and $O^k_{l+j}P^j_l$ have to be considered.

2. THE R(3) SCALAR AND NONSCALAR PRODUCT OPERATORS OF THE TYPE $P^k_{l+j}O^j_l$ AND $O^j_{l+k}P^k_l$ AND THEIR MUTUAL RELATIONS

The quadratic operators $P^k_{l+j}O^j_l$ or $O^j_{l+k}P^k_l$ $(-3\leqslant j\leqslant +3,\ -5\leqslant k\leqslant 5$, and $0\leqslant|j+k|\leqslant 8)$ shift the l values of the state upon which they act by $s=j+k$. With the available shift operators (I.2.1–I.2.4 and IV.2.4–IV.2.9) and the properties (I.2.5) and (IV.2.3), one can construct 14 product operators with $s=j+k=0$, 14 with $s=\pm 1$, 14 with $s=\pm 2$, 12 with $s=\pm 3$, 10 with $s=\pm 4$, eight with $s=\pm 5$, six with $s=\pm 6$, four with $s=\pm 7$, and two with $s=\pm 8$. It must be noted that the mentioned expressions are only valid when they act to the right upon states with angular momentum projection $m=0$. It has been remarked previously that this condition does not seriously detract from the generality of the presented calculations.

The considered quadratic product operators consist of terms composed of one p_μ, one q_μ (the p_μ, respectively, q_μ, are defined in IV.1.3, respectively, V.1.1), and eight or less R(3) generators l_i $(i=0,\pm)$. In order to obtain relations between them, we have chosen to transform the terms of the operators into the standard form $q_\mu p_\nu l_-^{\mu+\nu}$ (if $\mu+\nu>0$), $q_\mu p_\nu l_+^{-\mu-\nu}$ (if $\mu+\nu<0$), or $q_\mu p_\nu$ (if $\mu+\nu=0$). On account of the commutator relations between p- and q-generators which are summarized in the Appendix, it is evident that operators, where q-generators appear in expressions between $P^k_{l+j}O^j_l$ and $O^j_{l+k}P^k_l$, will be the O^{j+k}_l themselves, if they exist (i.e., if $0\leqslant|j+k|\leqslant 3$). By straightforward calculation we have arrived at the following final results for the cases where $s\leqslant 0$:

$$
\begin{aligned}
&-(l+1)(l+2)^2(l+3)^2(2l+3)(2l+5)^2 l(2l-1)O^{+3}_{l-3}P^{-3}_l/l^2(l-1)^2(l-2)^2\\
&+12(l+3)^2(2l+5)^2 l(2l-3)(l+1)(l+2)(2l+3)O^{+2}_{l-2}P^{-2}_l/l^2(l-1)^2\\
&-7\times 12(2l-1)(l+1)(l+2)(l+3)^2(2l+3)(2l+5)(l-2)O^{+1}_{l-1}P^{-1}_l/l^2\\
&-7\times 12(l+1)(l+2)(l+3)(2l+3)(2l+5)(l-2)(2l-3)O^0_lP^0_l\\
&-7\times 12l(l-1)(l-2)(2l-1)(2l-3)(l+3)(2l+5)P^0_lO^0_l\\
&+7\times 12l(l-1)(l-2)^2(2l-1)(2l-3)(l+3)(2l+3)P^{-1}_{l+1}O^{+1}_l/(l+1)^2\\
&+12l(l-1)(l-2)^2(2l-1)(2l-3)^2(l+1)(2l+5)P^{-2}_{l+2}O^{+2}_l/(l+1)^2(l+2)^2
\end{aligned}
$$

[a] Research Assistant N.F.W.O. (Belgium).
[b] Research Associate N.F.W.O. (Belgium).

 0022-2488/83/051339-06$02.50

$$+ l(l-1)^2(l-2)^2(2l-1)(2l-3)^2(l+1)(2l+3)P^{-3}_{l+3}O^{+3}_{l}/(l+1)^2(l+2)^2(l+3)^2$$
$$- (8/3\sqrt{3})l(l-1)(l-2)^2(2l-1)(2l-3)^2(l+1)(l+2)(l+3)^2(2l+1)(2l+3)(2l+5)^2O^{0}_{l} = 0, \qquad (2.1)$$
$$- l(l+1)(l+2)^2(l+3)^2(2l+1)(2l+5)^2(2l+7)O^{+3}_{l-3}P^{-3}_{l}/l^2(l-1)^2(l-2)^2$$
$$+ 8(2l-3)(l+1)(l+2)(l+3)^2(2l+1)(2l+5)^2(2l+7)O^{+2}_{l-2}P^{-2}_{l}/l^2(l-1)^2$$
$$- 7\times 24(l-2)(2l-1)(l+1)(l+2)(l+3)^2(2l+5)(2l+7)O^{+1}_{l-1}P^{-1}_{l}/l^2$$
$$- 7\times 24(l-2)(2l-3)(l+2)(l+3)(2l+1)(2l+5)(2l+7)O^{0}_{l}P^{0}_{l}$$
$$- 15\times 28(l-1)(l-2)(2l-3)(l+2)(l+3)(2l+5)O^{-1}_{l+1}P^{+1}_{l}/(l+1)^2$$
$$+ 28l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+3)(2l+1)P^{-1}_{l+1}O^{+1}_{l}/(l+1)^2$$
$$+ 4l(l-1)(l-2)(2l-1)(2l-3)^2(l+1)(2l+1)(2l+5)P^{-2}_{l+2}O^{+2}_{l}/(l+1)^2(l+2)^2$$
$$+ l(l-1)^2(l-2)(2l-1)(2l-3)^2(l+1)(l+2)(2l+1)P^{-3}_{l+3}O^{+3}_{l}/(l+1)^2(l+2)^2-(l+3)^2$$
$$- (8/3\sqrt{3})l^2(l-1)(l-2)(2l-1)(2l-3)^2(l+1)(l+2)(l+3)^2(2l+1)(2l+5)^2(2l+7)O^{0}_{l} = 0, \qquad (2.2)$$
$$- (2l-1)(l+1)(l+2)^2(l+3)^2(l+4)(2l+1)(2l+3)(2l+5)^2(2l+7)O^{+3}_{l-3}P^{-3}_{l}/l^2(l-1)^2(l-2)^2$$
$$+ 30l(2l-3)(l+1)(l+2)(l+3)^2(l+4)(2l+3)(2l+5)^2(2l+7)O^{+2}_{l-2}P^{-2}_{l}/l^2(l-1)^2$$
$$- 140(l-2)(2l-1)(l+2)(l+3)^2(l+4)(2l+1)(2l+3)(2l+5)(2l+7)O^{+1}_{l-1}P^{-1}_{l}/l^2$$
$$- 420(l-2)(2l-3)(l+1)(l+2)(l+3)(l+4)(2l+1)(2l+5)(2l+7)O^{0}_{l}P^{0}_{l}$$
$$- 420(l-1)(l-2)(2l-1)(2l-3)(l+2)(l+3)(2l+5)O^{-2}_{l+2}P^{+2}_{l}/(l+1)^2(l+2)^2$$
$$+ 70l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+3)(l+4)(2l+1)(2l+3)P^{-1}_{l+1}O^{+1}_{l}/(l+1)^2$$
$$+ 2l(l-1)(l-2)(2l-1)(2l-3)(l+1)(2l+1)(2l+3)(2l+5)(9l^2+21l-64)P^{-2}_{l+2}O^{+2}_{l}/(l+1)^2(l+2)^2$$
$$+ l(l-1)^2(l-2)(2l-1)(2l-3)(l+1)(l+2)(2l+1)(2l+3)(4l^2+8l-35)P^{-3}_{l+3}O^{+3}_{l}/(l+1)^2(l+2)^2(l+3)^2$$
$$- (8/3\sqrt{3})l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+2)(l+3)^2(l+4)(2l+1)(2l+3)(2l+5)^2(2l+7)(4l^2-9l+1)O^{0}_{l} = 0, \qquad (2.3)$$
$$- 5l(l+1)(l+2)^2(l+3)^2(l+4)(2l+3)(2l+5)^2(2l+7)(2l+9)(2l-1)O^{+3}_{l-3}P^{-3}_{l}/l^2(l-1)^2(l-2)^2$$
$$+ 36l(2l-3)(l+2)(l+3)^2(l+4)(2l+1)(2l+3)(2l+5)^2(2l+7)(2l+9)O^{+2}_{l-2}P^{-2}_{l}/l^2(l-1)^2$$
$$- 630(2l-1)(l+1)(l+2)(l+3)^2(l+4)(2l+1)(2l+5)(2l+7)(2l+9)(l-2)O^{+1}_{l-1}P^{-1}_{l}/l^2$$
$$- 420(l-2)(2l-3)(l+1)(l+3)(l+4)(2l+1)(2l+3)(2l+5)(2l+7)(2l+9)O^{0}_{l}P^{0}_{l}$$
$$- 35\times 36l(l-1)(l-2)(2l-1)(2l-3)(l+3)(2l+5)O^{-3}_{l+3}P^{+3}_{l}/(l+1)^2(l+2)^2(l+3)^2$$
$$+ 6\times 35l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+3)(l+4)(2l+1)(2l+3)(2l+9)P^{-1}_{l+1}O^{+1}_{l}/(l+1)^2$$
$$+ 12l(l-1)(l-2)(2l-1)(2l-3)(l+1)(2l+1)(2l+3)(2l+5)(2l+9)(4l^2+8l-35)P^{-2}_{l+2}O^{+2}_{l}/(l+1)^2(l+2)^2$$
$$+ l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+2)(2l+1)(2l+3)(20l^4+100l^3-193l^2$$
$$- 795l+1260)P^{-3}_{l+3}O^{+3}_{l}/(l+1)^2(l+2)^2(l+3)^2$$
$$- (8/3\sqrt{3})l(l-1)(l-2)(2l-1)(2l-3)(l+1)(l+2)(l+3)^2(l+4)(2l+1)(2l+3)$$
$$\times(2l+5)^2(2l+7)(2l+9)(10l^2-29l+15)O^{0}_{l} = 0, \qquad (2.4)$$
$$- (l-1)(l+2)^2(l+1)(l+3)(2l+5)^2(2l+3)^2(2l-3)O^{+3}_{l-4}P^{-4}_{l}/(l-1)^2(l-2)^2(l-3)^2$$
$$+ 27(l+1)(l-1)(l+2)^2(2l-5)(2l+5)^2(2l+3)(l+3)O^{+2}_{l-3}P^{-3}_{l}/(l-1)^2(l-2)^2$$
$$- 4\times 27(2l-3)(l+1)(l+2)(l+3)(2l+3)(2l+5)^2(l-3)O^{+1}_{l-2}P^{-2}_{l}/(l-1)^2$$
$$+ 7\times 36(l+3)(l+2)(l+1)(2l+3)(2l+5)(2l-5)(l-3)O^{0}_{l-1}P^{-1}_{l}$$
$$+ 7\times 36(2l+5)(l+3)(l-1)(l-2)(l-3)(2l-3)(2l-5)P^{-1}_{l}O^{0}_{l}$$
$$+ 4\times 27(2l-5)^2(2l+3)(l+3)(l-1)(l-2)(l-3)(2l-3)P^{-2}_{l+1}O^{+1}_{l}/(l+1)^2$$
$$+ 27(l+1)(l-2)^2(2l+5)(2l-5)^2(l-1)(l-3)(2l-3)P^{-3}_{l+2}O^{+2}_{l}/(l+1)^2(l+2)^2$$
$$+ (l+1)(l-1)(l-2)^2(l-3)(2l-3)^2(2l-5)^2(2l+3)P^{-4}_{l+3}O^{+3}_{l}/(l+1)^2(l+2)^2(l+3)^2$$
$$+ 8\sqrt{3}(l+3)(l+2)^2(l+1)l(l-1)(l-2)^2(l-3)(2l-3)(2l+3)(2l-5)^2(2l+5)^2O^{-1}_{l} = 0, \qquad (2.5)$$
$$l(l-1)(2l-1)(2l-3)(l+2)^2(l+3)^2(2l+3)(2l+5)^2O^{+3}_{l-4}P^{-4}_{l}/(l-1)^2(l-2)^2(l-3)^2$$
$$- 18l(2l-1)(2l-3)(l+2)^2(l+3)^2(2l+5)^2(2l-5)O^{+2}_{l-3}P^{-3}_{l}/(l-1)^2(l-2)^2$$
$$+ 6^3l(l-3)(2l-3)^2(l+2)(l+3)^2(2l+5)^2O^{+1}_{l-2}P^{-2}_{l}/(l-1)^2$$

$$
\begin{aligned}
&-7\times 8\times 9(l-3)(2l-1)(2l-3)(2l-5)(l+2)(l+3)^2(2l+5)O^{0}_{l-1}P_l^{-1}\\
&+21\times 36(l-2)(l-3)(2l-3)(2l-5)(l+2)(l+3)(2l+5)O_l^{-1}P_l^{0}\\
&-36l(l-1)(l-2)(l-3)(2l-1)(2l-3)^2(2l-5)(l+3)P^{-2}_{l+1}O_l^{+1}/(l+1)^2\\
&-9l(l-1)(l-2)^2(l-3)(2l-1)(2l-3)^2(2l-5)(2l+5)P^{-3}_{l+2}O_l^{+2}/(l+1)^2(l+2)^2\\
&-l(l-1)(l-2)^2(l-3)(2l-1)(2l-3)^3(2l-5)(l+2)P^{-4}_{l+3}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&-4\sqrt{3}l(l-1)(l-2)^2(l-3)(2l-1)^2(2l-3)^2(2l-5)(l+2)^2(l+3)^2(2l+5)^2O_l^{-1}=0,
\end{aligned}
\tag{2.6}
$$

$$
\begin{aligned}
&-2(2l-1)(2l-3)(l)(l+2)^2(l+3)^2(2l+1)(2l+3)(2l+5)^2(2l+7)O^{+3}_{l-4}P_l^{-4}/(l-1)^2(l-2)^2(l-3)^2\\
&+135l(l-1)(2l-5)(l+2)^2(l+3)^2(2l+1)(2l+5)^2(2l+7)O^{+2}_{l-3}P_l^{-3}/(l-1)^2(l-2)^2\\
&-360(l-3)(2l-1)(2l-3)(l+2)(l+3)^2(2l+1)(2l+5)^2(2l+7)O^{+1}_{l-2}P_l^{-2}/(l-1)^2\\
&+2520l(l-3)(2l-1)(2l-5)(l+2)(l+3)^2(2l+5)(2l+7)O^{0}_{l-1}P_l^{-1}\\
&+21\times 180(l-2)(l-3)(2l-3)(2l-5)(l+2)(l+3)(2l+5)O^{-2}_{l+1}P_l^{+1}/(l+1)^2\\
&+180l(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)(l+3)(2l+1)(2l+7)P^{-2}_{l+1}O_l^{+1}/(l+1)^2\\
&+27l(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)(2l+1)(2l+5)(3l^2+4l-24)P^{-3}_{l+2}O_l^{+2}/(l+1)^2(l+2)^2\\
&+l(l-1)(l-2)(l-3)(2l-1)(2l-3)^2(2l-5)(l+2)(2l+1)(8l^2+8l-75)P^{-4}_{l+3}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&+4\sqrt{3}l(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)(l+2)^2(l+3)^2(2l+1)(2l+5)^2(2l+7)(8l^2-25l+12)O_l^{-1}=0,
\end{aligned}
\tag{2.7}
$$

$$
\begin{aligned}
&-5l(l+1)(l+2)^2(l+3)^2(l+4)(2l+1)(2l+3)(2l+5)^2(2l+7)(l-1)(2l-3)O^{+3}_{l-4}P_l^{-4}/(l-1)^2(l-2)^2(l-3)^2\\
&+81(l+1)(l+2)^2(l+3)^2(l+4)(2l+1)(2l+5)^2(2l+7)(l-1)(2l-1)(2l-5)O^{+2}_{l-3}P_l^{-3}/(l-1)^2(l-2)^2\\
&-810l(l+1)(l+2)(l+3)^2(l+4)(2l+5)^2(2l+7)(l-3)(2l-1)(2l-3)O^{+1}_{l-2}P_l^{-2}/(l-1)^2\\
&+90\times 14l(l+2)(l+3)^2(l+4)(2l+1)(2l+5)(2l+7)(l-3)(2l-1)(2l-5)O^{0}_{l-1}P_l^{-1}\\
&+70\times 54(l+2)(l+3)(2l+5)(l-1)(l-2)(l-3)(2l-3)(2l-5)O^{-3}_{l+2}P_l^{+2}/(l+1)^2(l+2)^2\\
&+270l(l+1)(l+3)(l+4)((2l+1)(2l+7)(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)P^{-2}_{l+1}O_l^{+1}/(l+1)^2\\
&+54l(l+1)(l+4)(2l+1)(2l+5)(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)(2l^2+2l-19)P^{-3}_{l+2}O_l^{+2}/(l+1)^2(l+2)^2\\
&+l(l+1)(l+2)(2l+1)(l-1)(l-2)(l-3)(2l-1)(2l-3)(2l-5)\\
&\times(20l^4+60l^3-311l^2-534l+1575)P^{-4}_{l+3}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&+4\sqrt{3}l(l+1)(l+2)^2(l+3)^2(l+4)(2l+1)(2l+5)^2(2l+7)(l-1)^2(l-2)(l-3)(2l-1)(2l-3)(2l-5)(10l-27)O_l^{-1}=0,
\end{aligned}
\tag{2.8}
$$

$$
\begin{aligned}
&-(l+1)^2(l+2)^2(l+3)(l-2)(2l+3)^2(2l+5)(2l-5)O^{+3}_{l-5}P_l^{-5}/(l-2)^2(l-3)^2(l-4)^2\\
&+15(l+1)(l+2)^2(l+3)(2l+3)^2(2l+5)(l-2)(2l-7)O^{+2}_{l-4}P_l^{-4}/(l-2)^2(l-3)^2\\
&-9\times 15(l+1)(l+2)^2(l+3)(l-4)(2l+3)(2l+5)(2l-5)O^{+1}_{l-3}P_l^{-3}/(l-2)^2\\
&+180(l+1)(l+2)(l+3)(l-4)(2l+3)(2l+5)(2l-7)O^{0}_{l-2}P_l^{-2}\\
&+180(l+3)(l-2)(l-3)(l-4)(2l-5)(2l-7)(2l+5)P_l^{-2}O_l^{0}\\
&+9\times 15(l+3)(l-2)(l-3)^2(l-4)(2l-5)(2l-7)(2l+3)P^{-3}_{l+1}O_l^{+1}/(l+1)^2\\
&+15(l+1)(l-2)(l-3)^2(l-4)(2l-5)^2(2l-7)(2l+5)P^{-4}_{l+2}O_l^{+2}/(l+1)^2(l+2)^2\\
&+(l+1)(l-2)^2(l-3)^2(l-4)(2l-5)^2(2l-7)(2l+3)P^{-5}_{l+3}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&-4\sqrt{3}(l+3)(l+2)^2(l+1)(l-2)(l-3)^2(l-4)(2l-1)(2l-5)^2(2l-7)(2l+3)^2(2l+5)O_l^{-2}=0,
\end{aligned}
\tag{2.9}
$$

$$
\begin{aligned}
&-(l+3)(l+2)^2(l+1)(l-1)(l-2)(2l-3)(2l+3)(2l+5)^2O^{+3}_{l-5}P_l^{-5}/(l-2)^2(l-3)^2(l-4)^2\\
&+10(l+3)(l+2)^2(l-1)(2l-3)(2l-7)(2l+3)(2l+5)^2O^{+2}_{l-4}P_l^{-4}/(l-2)^2(l-3)^2\\
&-270(l+3)(l+2)^2(l-1)(l-4)(2l-5)(2l+5)^2O^{+1}_{l-3}P_l^{-3}/(l-2)^2\\
&+360(l+3)(l+2)(l-4)(2l-3)(2l-7)(2l+5)^2O^{0}_{l-2}P_l^{-2}+28\times 45(l+3)(l+2)(l-3)(l-4)(2l-7)(2l+5)O^{-1}_{l-1}P_l^{-1}\\
&+45(l+3)(l-1)(l-2)(l-3)(l-4)(2l-3)(2l-5)(2l-7)P^{-3}_{l+1}O_l^{+1}/(l+1)^2\\
&+5(l-1)(l-2)(l-3)(l-4)(2l-3)(2l-5)^2(2l-7)(2l+5)P^{-4}_{l+2}O_l^{+2}/(l+1)^2(l+2)^2\\
&+(l+2)(l-1)(l-2)^2(l-3)(l-4)(2l-3)(2l-5)^2(2l-7)P^{-5}_{l+3}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&-4\sqrt{3}(l+3)(l+2)^2(l-1)^2(l-2)(l-3)(l-4)(2l-3)(2l-5)^2(2l-7)(2l+3)(2l+5)^2O_l^{-2}=0,
\end{aligned}
\tag{2.10}
$$

$$
-2(l+3)^2(l+2)^2(l+1)(l-1)(2l-1)(2l-3)(2l-5)(2l+3)(2l+5)^2O^{+3}_{l-5}P_l^{-5}/(l-2)^2(l-3)^2(l-4)^2
$$

$$
\begin{aligned}
&+75(l+3)^2(l+2)^2(l-1)(l-2)(2l+3)(2l+5)^2(2l-1)(2l-7)O_{l-4}^{+2}P_l^{-4}/(l-2)^2(l-3)^2\\
&-450(l+3)^2(l+2)^2(2l+5)^2(2l-1)(2l-3)(2l-5)(l-4)O_{l-3}^{+1}P_l^{-3}/(l-2)^2\\
&+1800(l+3)^2(l+2)(l-1)(2l-3)(2l+5)^2(2l-7)(l-4)O_{l-2}^{0}P_l^{-2}\\
&+12\times 21\times 15(l+2)(2l+5)(l+3)(l-3)(2l-5)(2l-7)(l-4)O_l^{-2}P_l^{0}\\
&+15\times 15(l+3)^2(l-1)(l-2)(2l-1)(2l-3)(l-3)(2l-5)(2l-7)(l-4)P_{l+1}^{-3}O_l^{+1}/(l+1)^2\\
&+15(l-1)(l-2)(2l-1)(2l-3)(2l+5)(l-3)(l-4)(2l-5)(2l-7)(3l^2+l-25)P_{l+2}^{-4}O_l^{+2}/(l+1)^2(l+2)^2\\
&+(l+2)(l-1)(l-2)^2(2l-1)(2l-3)(l-3)(l-4)(2l-5)(2l-7)(8l^2-75)P_{l+3}^{-5}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&-4\sqrt{3}(l+3)^2(l+2)^2(l-1)(l-2)(l-3)(l-4)(2l-1)(2l-3)(2l-5)(2l-7)(2l+3)(2l+5)^2(8l^2-32l+25)O_l^{-2}=0,
\end{aligned}
\tag{2.11}
$$

$$
\begin{aligned}
&2l(l+2)^2(l+3)^2(l+1)(2l+3)(2l+5)^2(2l+7)(l-2)(2l-1)(2l-5)(l-1)O_{l-5}^{+3}P_l^{-5}/(l-2)^2(l-3)^2(l-4)^2\\
&-18l(2l-3)(l+2)^2(l+3)^2(2l-1)(2l+5)^2(2l+7)(2l-7)(l-2)(2l+3)O_{l-4}^{+2}P_l^{-4}/(l-2)^2(l-3)^2\\
&+405l(l-1)(2l-3)(2l-5)(l+2)^2(l+3)^2(2l+5)^2(2l+7)(l-4)P_{l-3}^{+4}P_l^{-3}/(l-2)^2\\
&-360(l+3)^2(l+2)(2l+5)^2(2l+7)(l-1)(2l-1)(2l-3)(2l-7)(l-4)O_{l-2}^{0}P_l^{-2}\\
&+12\times 7\times 45(l+2)(2l+5)(l+3)(l-2)(l-3)(l-4)(2l-5)(2l-7)O_{l+1}^{-3}P_l^{+1}/(l+1)^2\\
&-135l(l-1)(l-2)(l-3)(l-4)(2l-1)(2l-3)(2l-5)(2l-7)(l+3)^2(2l+7)P_{l+1}^{-3}O_l^{+1}/(l+1)^2\\
&-3(2l+5)l(l-1)(l-2)(l-3)(l-4)(2l-1)(2l-3)(2l-5)(2l-7)(2l+7)(8l^2-77)P_{l+2}^{-4}O_l^{+2}/(l+1)^2(l+2)^2\\
&-(l+2)l(l-1)(l-2)(l-3)(l-4)(2l-1)(2l-3)(2l-5)(2l-7)(8l^4+8l^3-146l^2\\
&-74l+675)P_{l+3}^{-5}O_l^{+3}/(l+1)^2(l+2)^2(l+3)^2\\
&+4\sqrt{3}(l+3)^2(l+2)^2l(l-1)(l-2)(l-3)(l-4)(2l+3)(2l+5)^2(2l+7)(2l-1)\\
&\times(2l-3)(2l-5)(2l-7)(4l^2-18l+17)O_l^{-2}=0,
\end{aligned}
\tag{2.12}
$$

$$
\begin{aligned}
&-(l+1)^2(l+2)^2(2l+3)^2(l-2)(2l-5)O_{l-5}^{+2}P_l^{-5}/(l-3)^2(l-4)^2\\
&+15(l+1)(l+2)^2(2l+3)^2(l-2)(2l-7)O_{l-4}^{+1}P_l^{-4}/(l-3)^2\\
&-135(l+1)(l+2)^2(2l+3)(2l-5)(l-4)O_{l-3}^{0}P_l^{-3}-180(l+1)(l+2)(2l+3)\\
&\times(l-4)(2l-7)O_{l-2}^{-1}P_l^{-2}+45(l+2)(l-2)(l-3)(l-4)(2l-5)(2l-7)P_l^{-3}O_l^{0}\\
&+5(2l+3)(l-2)(l-3)(l-4)(2l-5)(2l-7)^2P_{l+1}^{-4}O_l^{+1}/(l+1)^2+(l-2)(l-3)^2\\
&\times(l-4)(2l-5)(2l-7)^2P_{l+2}^{-5}O_l^{+2}/(l+1)(l+2)^2+(4/\sqrt{3})(l+2)^2(l+1)^2(l-2)\\
&\times(l-3)^2(l-4)(2l-5)(2l-7)^2(2l+3)^2O_l^{-3}=0,
\end{aligned}
\tag{2.13}
$$

$$
\begin{aligned}
&-3(l+1)^2(l+2)^2(2l+3)^2(2l+5)(l-2)(l-3)O_{l-5}^{+2}P_l^{-5}/(l-3)^2(l-4)^2\\
&+10(l+1)(l+2)^2(2l+3)^2(2l+5)(2l-3)(2l-5)(2l-7)O_{l-4}^{+1}P_l^{-4}/(l-3)^2\\
&-270(l+1)(l+2)^2(2l+3)(2l+5)(l-2)(l-4)(2l-5)O_{l-3}^{0}P_l^{-3}\\
&+5\times 14\times 18(l+1)(l+2)(2l+3)(l-3)(l-4)(2l-7)O_{l-1}^{-2}P_l^{-1}\\
&+90(l+2)(2l+5)(l-2)(l-3)(l-4)(2l-3)(2l-5)(2l-7)P_l^{-3}O_l^{0}\\
&+5(2l+3)(l-2)(l-3)(l-4)(2l-3)(2l-5)(2l-7)(7l^2-8l-69)P_{l+1}^{-4}O_l^{+1}/(l+1)^2\\
&+(l-2)(l-3)^2(l-4)(2l-3)(2l-5)(2l-7)(6l^3-3l^2-76l-67)P_{l+2}^{-5}O_l^{+2}/(l+1)^2(l+2)^2\\
&+(4/\sqrt{3})(l+2)^2(l+1)^2(l-2)(l-3)^2(l-4)(2l-3)(2l-5)(2l-7)(2l+3)^2(2l+5)(3l-13)O_l^{-3}=0,
\end{aligned}
\tag{2.14}
$$

$$
\begin{aligned}
&-2(l+3)(l+2)^2(l+1)^2(l-1)(2l-3)(2l+3)^2(2l+5)O_{l-5}^{+2}P_l^{-5}/(l-3)(l-4)^2\\
&+25(l+3)(l+2)^2(l+1)(l-1)(l-2)(2l-7)(2l+3)^2(2l+5)O_{l-4}^{+1}P_l^{-4}/(l-3)^2\\
&-150(l+3)(l+2)^2(l+1)(l-2)(2l-3)(2l+3)(2l+5)(l-4)O_{l-3}^{0}P_l^{-3}\\
&+30\times 42(l+2)(l+1)(2l+3)(l-3)(l-4)(2l-7)O_l^{-3}P_l^{0}\\
&+150(l+3)(l+2)(2l+5)(l-1)(l-2)(l-3)(l-4)(2l-3)(2l-7)P_l^{-3}O_l^{0}\\
&+25(l+3)(2l+3)(l-1)(l-2)(l-3)(l-4)(2l-3)(2l-7)(2l^2-3l-23)P_{l+1}^{-4}O_l^{+1}/(l+1)^2\\
&+(l-1)(1-2)(l-3)(l-4)(2l-3)(2l-7)(8l^4-16l^3-170l^2+178l+1005)P_{l+2}^{-5}O_l^{+2}/(l+1)(l+2)^2\\
&+(4/\sqrt{3})(l+3)(l+2)^2(l+1)^2(l-1)(l-2)(l-3)(l-4)(2l+3)^2(2l+5)(2l-3)(2l-7)(4l^2-32l+65)O_l^{-3}=0,
\end{aligned}
\tag{2.15}
$$

$$\begin{aligned}&l(l+1)^2(2l+1)(2l+3)(l-3)O_{l-5}^{+1}P_l^{-5}/(l-4)^2-5l(l+1)(2l+1)(2l+3)\\&\times(2l-7)O_{l-4}^{0}P_l^{-4}-45l(l+1)(2l+1)(l-4)O_{l-3}^{-1}P_l^{-3}+45(l+1)(l-3)\\&\times(l-4)(2l-7)P_{l-1}^{-3}O_l^{-1}-5(2l+1)(l-3)(l-4)(2l-7)(2l-9)P_l^{-4}O_l^{0}\\&-l(l-4)^2(l-3)(2l-7)(2l-9)P_{l+1}^{-5}O_l^{+1}/(l+1)^2=0,\end{aligned}\tag{2.16}$$

$$\begin{aligned}&-(l+1)^2(l+2)(2l+3)(l-2)(2l-5)O_{l-5}^{+1}P_l^{-5}/(l-4)^2+15(l+1)(l+2)(2l+3)\\&\times(l-2)(2l-7)O_{l-4}^{0}P_l^{-4}+135(l+1)(l+2)(2l-5)(l-4)O_{l-3}^{-1}P_l^{-3}+180(l+1)\\&\times(l-4)(2l-7)O_{l-2}^{-2}P_l^{-2}+5(l-2)(l-3)(l-4)(2l-5)(2l-7)P_l^{-4}O_l^{0}+(l-2)\\&\times(l-3)(l-4)^2(2l-5)(2l-7)P_{l+1}^{-5}O_l^{+1}/(l+1)^2=0,\end{aligned}\tag{2.17}$$

$$\begin{aligned}&-3(l+1)^2(l+2)(2l+3)(2l+5)(l-2)(l-3)(2l-3)O_{l-5}^{+1}P_l^{-5}/(l-4)^2+10(l+1)\\&\times(l+2)(2l+3)(2l+5)(2l-3)(2l-5)(2l-7)O_{l-4}^{0}P_l^{-4}+270(l+1)(l+2)(2l+5)\\&\times(l-2)(l-4)(2l-5)O_{l-3}^{-1}P_l^{-3}-1260(l+1)(l-3)(l-4)(2l-7)O_{l-1}^{-3}P_l^{-1}\\&+10(2l+5)(l-2)(l-3)(l-4)(2l-3)(2l-5)(2l-7)P_l^{-4}O_l^{0}+3(l-2)(l-3)\\&\times(l-4)(2l-3)(2l-5)(2l-7)(l^2-2l-13)P_{l+1}^{-5}O_l^{+1}/(l+1)^2=0,\end{aligned}\tag{2.18}$$

$$\begin{aligned}&(l-3)l(l+1)(2l+3)O_{l-5}^{0}P_l^{-5}+5(2l-7)(2l+3)lO_{l-4}^{-1}P_l^{-4}+45l(l-4)\\&\times O_{l-3}^{-2}P_l^{-3}+5(l-3)(l-4)(2l-7)P_{l-1}^{-4}O_l^{-1}-(l-3)(l-4)(l-5)(2l-7)P_l^{-5}O_l^{0}=0,\end{aligned}\tag{2.19}$$

$$\begin{aligned}&(2l-5)l(l+1)(l+2)(2l+3)O_{l-5}^{0}P_l^{-5}+15(l-3)l(l+2)(2l+3)O_{l-4}^{-1}P_l^{-4}\\&-180l(l-4)O_{l-2}^{-3}P_l^{-2}+15(l-3)(l-4)(2l-5)(l+2)P_{l-1}^{-4}O_l^{-1}-(l-3)(l-4)(l-6)(2l-5)(2l+5)P_l^{-5}O_l^{0}=0,\end{aligned}\tag{2.20}$$

$$(2l+3)(l+1)(l-3)O_{l-5}^{-1}P_l^{-5}-(l-3)(l-4)(2l-7)P_{l-1}^{-5}O_l^{-1}+5(2l+3)(2l-7)O_{l-4}^{-2}P_l^{-4}+45(l-4)O_{l-3}^{-3}P_l^{-3}=0,\tag{2.21}$$

$$(l-1)(l+1)O_{l-5}^{-1}P_l^{-5}-(l-4)(l-6)P_{l-1}^{-5}O_l^{-1}+5(l-1)O_{l-4}^{-2}P_l^{-4}+5(l-4)P_{l-2}^{-4}O_l^{-2}=0,\tag{2.22}$$

$$(l+1)O_{l-5}^{-2}P_l^{-5}-(l-4)P_{l-2}^{-5}O_l^{-2}+5O_{l-4}^{-3}P_l^{-4}=0,\tag{2.23}$$

$$P_{l-3}^{-5}O_l^{-3}-O_{l-5}^{-3}P_l^{-5}=0.\tag{2.24}$$

3. DISCUSSION

It has been originally pointed out by Hughes and Yadegar[11] that it is always possible to turn an l-lowering shift operator into an l-raising one and vice-versa by a formal change of the parameter l in the definition of that shift operator. This property can be expressed by the operator equalities $P_l^{-k}=P_{-l-1}^{k}$ and $O_l^{-k}=O_{-l-1}^{k}$. On account of these equalities the relations established in the present paper can be simply transformed into relations among product operators that raise l by 1, 2,... or 8 units.

Moreover, every relation among product operators which shift l to $l+s$ can be turned immediately into another relation (or exceptionally the same) of similar kind by carrying out the following operations:

(i) the parameter l in the coefficients and in these only is formally replaced by $-l-s-1$,

(ii) each operator product of the form $P_{l+j}^{k}O_l^{j}$ $(j+k=s)$, respectively, $O_{l+k}^{j}P_l^{k}$, is replaced by the product $O_{l+k}^{j}P_l^{k}$, respectively, $P_{l+j}^{k}O_l^{j}$,

(iii) shift operators occurring linearly in the relation are kept unchanged.

The proof of the validity of this transformation, which we shall denote by "$l\to -l-s-1$", has been given elsewhere.[12]

In the present results, we discover a very simple rule for the total number of independent relations among product operators of the type $P_{l+j}^{k}O_l^{j}$ or $O_{l+k}^{j}P_l^{k}$. Indeed, if $2n(s)$ is the total number of "mixed product operators" [this means: $n(s)$ $P_{l+j}^{k}O_l^{j}$ operators and $n(s)$ $O_{l+k}^{j}P_l^{k}$ operators; $j+k=s$], the number of independent relations among those mixed product operators is given by $n(s)$, and the number of product operators appearing in one relation is $n(s)+1$. However, because of our transformation rule we did not have to construct all $n(s)$ relations. If $n(s)$ is even, we explicitly constructed $n(s)/2$ independent relations; the other $n(s)/2$ independent ones can then be deduced by the transformation "$l\to -l-s-1$". If $n(s)$ is odd, we explicitly constructed $[n(s)+1]/2$ independent relations. In the latter case also one of the expressions turns into itself by the transformation rule "$l\to -l-s-1$" [hence, again $n(s)$ independent ones exist].

The fact that the number of product operators appearing in one relation is $n(s)+1$ also implies that it is impossible to construct a relation consisting of $O_{l+k}^{j}P_l^{k}$ operators (respectively, $P_{l+j}^{k}O_l^{j}$ operators) exclusively. This could point out that there is always a commutator of an O_l^{k}-operator and a P_l^{k}-operator hidden in one relation. Further investigations showed that certain linear combinations of our expressions could result in relations that are built up with a "sort of commutators." For the case $s=0$, for instance, we have set up the following expression:

$$(l+3)(l+4)(2l+5)(2l+7)(2l+9)[O_l^0,P_l^0]$$
$$+15\frac{(l+3)(l+4)(2l+3)(2l+9)}{(l+1)^2}$$
$$\times(O_{l+1}^{-1}P_l^{+1}-P_{l+1}^{-1}O_l^{+1})$$
$$+30\frac{(l+1)(2l+5)(2l+9)}{(l+1)^2(l+2)^2}$$
$$\times(O_{l+2}^{-2}P_l^{+2}-P_{l+2}^{-2}O_l^{+2})+30\frac{(l+1)(2l+3)}{(l+1)^2(l+2)^2(l+3)^2}$$
$$\times(O_{l+3}^{-3}P_l^{+3}-P_{l+3}^{-3}O_l^{+3})=0. \tag{3.1}$$

However, since those kind of relations were of no direct use for further application, we did not bring them all under the commutatorlike form.

In a following paper, we shall demonstrate how the appropriate use of the relations between mixed product operators leads to very general formulas for the O_l^0-eigenvalues and eigenstates, even for l-degenerate states.

ACKNOWLEDGMENT

We would like to express our gratitude to Professor C. C. Grosjean for his continuous interest.

APPENDIX: THE COMMUTATION RELATIONS $[q_\mu, p_\mu]$

$$[q_0,p_{\pm3}]=\mp\frac{1}{\sqrt{3}}q_{\pm3},\quad [q_{\pm2},p_{\mp3}]=\mp\frac{1}{\sqrt{2}\sqrt{3}}q_{\mp1},$$
$$[q_0,p_{\pm2}]=\mp\frac{1}{\sqrt{3}}q_{\pm2},\quad [q_{\pm2},p_{\mp2}]=\mp\frac{1}{\sqrt{3}}q_0,$$
$$[q_0,p_{\pm1}]=\mp\frac{\sqrt{5}}{\sqrt{2}\sqrt{3}\sqrt{7}}q_{\pm1},\quad [q_{\pm2},p_{\mp1}]=\mp\frac{3}{2\sqrt{7}}q_{\pm1},$$
$$[q_{\pm1},p_{\pm2}]=\pm\frac{1}{\sqrt{2}\sqrt{3}}q_{\pm3},\quad [q_{\pm3},p_{\mp5}]=\pm\frac{1}{\sqrt{2}}q_{\mp2},$$
$$[q_{\pm1},p_{\pm1}]=\pm\frac{3}{2\sqrt{7}}q_{\pm2},\quad [q_{\pm3},p_{\mp4}]=\pm\frac{1}{\sqrt{2}}q_{\mp1},$$
$$[q_{\pm1},p_{\mp4}]=\mp\frac{1}{\sqrt{2}}q_{\mp3},\quad [q_{\pm3},p_{\mp3}]=\pm\frac{1}{\sqrt{3}}q_0,$$
$$[q_{\pm1},p_{\mp3}]=\mp\frac{1}{\sqrt{2}\sqrt{3}}q_{\mp2},\quad [q_{\pm3},p_{\mp2}]=\pm\frac{1}{\sqrt{2}\sqrt{3}}q_{\pm1},$$
$$[q_{\pm1},p_{\mp2}]=0,\quad [q_{\pm3},p_{\mp1}]=\pm\frac{5}{2\sqrt{3}\sqrt{7}}q_{\pm2},$$
$$[q_{\pm1},p_{\mp1}]=\pm\frac{\sqrt{5}}{\sqrt{2}\sqrt{3}\sqrt{7}}q_0,\quad [p_0,q_{\pm3}]=\mp\frac{1}{2\sqrt{3}\sqrt{7}}q_{\pm3},$$
$$[q_{\pm2},p_{\pm1}]=\mp\frac{\sqrt{5}}{2\sqrt{3}\sqrt{7}}q_{\pm3},\quad [p_0,q_{\pm2}]=\pm\frac{2}{\sqrt{3}\sqrt{7}}q_{\pm2},$$
$$[q_{\pm2},p_{\mp5}]=\pm\frac{1}{\sqrt{2}}q_{\mp3},\quad [p_0,q_{\pm1}]=\mp\frac{5}{2\sqrt{3}\sqrt{7}}q_{\pm1},$$
$$[q_{\pm2},p_{\mp4}]=0,\quad [p_0,q_0]=0.$$

The not-mentioned commutators are all trivially zero.

[1]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 1902 (1980).
[2]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1906 (1980).
[3]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1913 (1980).
[4]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 1967 (1980).
[5]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1973 (1980).
[6]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 2632 (1980).
[7]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 2635 (1980).
[8]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **22**, 2326 (1981).
[9]G. Vanden Berghe, H. E. De Meyer, and J. Van der Jeugt, J. Math. Phys. **22**, 2332 (1981).
[10]H. E. De Meyer, G. Vanden Berghe, and J. Van der Jeugt, J. Math. Phys. **23**, 1223 (1982).
[11]J. W. B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978).
[12]G. Vanden Berghe, H. E. De Meyer, and J. Van der Jeugt, Lett. Nuovo Cimento **33**, 120 (1982).

Shift operator techniques for the classification of multipole-phonon states. XII. O_l^0 eigenstate and eigenvalue determination in R(7)

J. Van der Jeugt,[a)] G. Vanden Berghe, and H. E. De Meyer[b)]
Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit-Gent Krijgslaan 281-S9, B-9000 Gent, Belgium

(Received 10 February 1982; accepted for publication 26 March 1982)

On account of previously derived relations between quadratic shift operator products of the type $P^j_{l+k}O^k_l$, $O^k_{l+j}P^j_l$, and $O^k_{l+j}O^j_l$ in the group R(7), part of the eigenvalue spectrum of the scalar shift operator O^0_l is derived in closed form. The corresponding eigenstates which are closely related to the octupole-phonon state vectors are defined in terms of angular momentum lowering shift operator actions upon the maximum angular momentum state. In the case of l-degenerate states the relation between the previously constructed P^0_l eigenstates and the derived O^0_l eigenstates is discussed. A short comment on a numerical method for O^0_l eigenvalue determination is included.

PACS numbers: 63.20.Dj, 02.20.Nq, 02.30.Tb

1. INTRODUCTION

In two previous papers[1,2] (to be referred to as V and VI) we have obtained relations which connect quadratic products of the shift operators O^k_l ($|k|\leqslant 3$), built up with R(7) group generators. In V four independent relations have been constructed among the R(3) scalar operators $O^{-k}_{l+k}O^k_l$ ($|k|\leqslant 3$), V^*, and the scalar shift operator P^0_l built up out of the G_2 generators. The R(3) nonscalar extension (VI) leads to eight independent relations among products of the type $O^j_{l+k}O^k_l$ ($|j|, |k|\leqslant 3$; $-5\leqslant j+k<0$). In a preceding paper[3] (to be referred to as XI) mixed product operators of the type $P^j_{l+k}O^k_l$ and $O^k_{l+j}P^j_l$ ($|k|\leqslant 3$; $|j|\leqslant 5$) were introduced with P^j_l shift operators within the G_2 group.

In the present paper we want to derive formulas expressing the O^0_l eigenvalues and eigenstates. The obvious method to do this consists in using the scalar and nonscalar relations (V.2.5–V.2.8 and VI.3.1–VI.3.8), as was the case for R(5) (see Ref. 4) and G_2.[5] However, in the present investigation there are additional problems. Indeed, the scalar and nonscalar relations between the product operators $O^j_{l+k}O^k_l$ also contain an operator of the type P^{j+k}_l. This means that the P^0_l eigenvalues and the action of the P^k_l upon states have to be known before one can start with deriving the O^0_l eigenvalues. On the other hand, the commutator $[P^0_l,O^0_l]$ is not identically zero. As a consequence, the constructed eigenstates of P^0_l will not be eigenstates of O^0_l in general. Considering these problems, it was practically impossible to calculate the O^0_l eigenvalues for degenerate states exclusively out of the relations (V.2.5–2.8) and (VI.3.1–3.8). Therefore, the relations between mixed product operators were introduced in the preceding paper.[3] Since they also contain the P^k_l operators, and since the P^0_l eigenvalues are known in general for $3v-5\leqslant l\leqslant 3v$ (see Ref. 5, to be referred to as X), this will enable us to calculate the O^0_l eigenvalues for the same l-interval. Furthermore, we may assume for all l that the precise number of degenerate states is known in advance.

The eigenvalues of P^0_l are denoted by $\alpha^{(i)}_{v,l}$, where i takes on integer values between one and the number indicating the l-multiplicity of states with seniority v and angular momentum l. In a somewhat modified form we can write formula (X.1.2) as

$$P^0_l|v,l,\alpha^{(i)}_{v,l}\rangle=\alpha^{(i)}_{v,l}|v,l,\alpha^{(i)}_{v,l}\rangle. \tag{1.1}$$

In X closed expressions for all eigenvalues $\alpha^{(i)}_{v,l}$ with $3v-5\leqslant l\leqslant 3v$ have been constructed. In the following sections we shall derive closed expressions for the eigenvalues of O^0_l, denoted by $\beta^{(i)}_{v,l}$. Hence, we write

$$O^0_l|v,l,\beta^{(i)}_{v,l}\rangle=\beta^{(i)}_{v,l}|v,l,\beta^{(i)}_{v,l}\rangle, \tag{1.2}$$

where the states $|v,l,\beta^{(i)}_{v,l}\rangle$ form a set of orthonormalized eigenstates of O^0_l. Proceeding as in X, we introduce the coefficients $b^{(j)}_{v,l-k}\,[l,(i)]$ by means of the following formula:

$$O^{-k}_l|v,l,\beta^{(i)}_{v,l}\rangle=\sum_j b^{(j)}_{v,l-k}[l,(i)]\,|v,l-k,\beta^{(j)}_{v,l-k}\rangle\quad(0<k\leqslant 3). \tag{1.3}$$

The additional labels $\alpha^{(i)}_{v,l}$ or $\beta^{(i)}_{v,l}$ will be omitted whenever $|v,l\rangle$ is not degenerate.

2. THE HIGH ANGULAR MOMENTUM STATES AND THEIR EIGENVALUES

The determination of the eigenvalues $\beta_{v,3v}$ of the nondegenerate maximum angular momentum states $|v,3v\rangle$ ($v=0,1,2,\ldots$) must be obtained out of Eqs. (V.2.5) and (V.2.6). Indeed, we know that the states $|v,3v-1\rangle$ and $|v,3v+i\rangle$ ($i=1,2,3$) do not exist. So, if we let Eqs. (V.2.5) and (V.2.6) act upon $|v,3v\rangle$, and multiply on the left by $\langle v,3v|$, we immediately obtain on account of the normalization of states a system of two quadratic equations in the unknown $\beta_{v,3v}$. This yields as a unique solution the expression

$$\beta_{v,3v}=-\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}v(2v+1)(3v+1)(3v+2). \tag{2.1}$$

In the application of the mixed relations presented in XI, we shall explicitly use the property that the nondegenerate states $|v,3v\rangle$, $|v,3v-2\rangle$, and $|v,3v-3\rangle$ are eigenstates of both P^0_l and O^0_l. The derivation of $\beta_{v,3v-2}$ eigenvalues is a

---

[a)] Research assistant N.F.W.O. (Belgium).
[b)] Research associate N.F.W.O. (Belgium).

 0022-2488/83/051345-05$02.50

first example where we can make use of relations between these product operators of the mixed type. If we let Eqs. (XI.2.2) and (XI.2.3) act on $|v,3v\rangle$, and eliminate the term $O^{+3}_{3v-3}P^{-3}_{3v}|v,3v\rangle$, we obtain

$$\Big\{\frac{(3v+1)(v+1)(2v+1)(6v+5)}{v^2(3v-1)^2}O^{+2}_{3v-2}P^{-2}_{3v} - 126(3v-1)(6v+1)O^0_{3v}P^0_{3v} - 36\sqrt{3}v^2(3v-1)(6v-1)(3v+1)(v+1) \times(6v+1)(2v+1)(6v+5)O^0_{3v}\Big\}|v,3v\rangle = 0. \tag{2.2}$$

Since the $\alpha_{v,3v}$, $\beta_{v,3v}$, and $a_{v,3v-2}[3v]$ values [defined as in (X.1.1)] are already known, it follows from Eq. (2.2) that

$$\langle v,3v|O^{+2}_{3v-2}|v,3v-2\rangle = -12\sqrt{3}v^2(3v-1)(3v+1)[(v-1)(2v-1)(3v-1)]^{1/2}. \tag{2.3}$$

Making use of the Hermiticity properties of shift operators (VI.2.1), one obtains out of (2.3)

$$b_{v,3v-2}[3v] = \langle v,3v-2|O^{-2}_{3v}|v,3v\rangle = -4\sqrt{3}v^2(3v-1)(3v+1)(6v+1) \times[(v-1)(3v-1)/(2v-1)]^{1/2}. \tag{2.4}$$

It should be noticed that the sign of $b_{v,3v-2}[3v]$ is unambiguously determined once a phase convention for $a_{v,3v-2}[3v]$ is adopted [see (X.2.7)].

The elimination of the operators $O^{+3}_{l-5}P^{-5}_l$, $O^{+2}_{l-4}P^{-4}_l$, and $O^{+1}_{l-3}P^{-3}_l$ out of Eqs. (XI.2.9), (XI.2.10), (XI.2.11), (XI.2.12) leads to a new relation which we let act upon the $|v,3v\rangle$ state, i.e.,

$$\{-5(3v+1)(3v+2)(v+1)(2v+1)(6v+5)O^0_{3v-2}P^{-2}_{3v} + 5v(3v-1)(6v-1)(v+1)(6v+5)P^{-2}_{3v}O^0_{3v} + 105(3v-1)(3v+2)(2v+1)O^{-2}_{3v}P^0_{3v} - 6\sqrt{3}v(3v-1)(6v-1)(3v+1)(3v+2)^2(v+1) \times(2v+1)^2(6v+5)O^{-2}_{3v}\}|v,3v\rangle = 0. \tag{2.5}$$

Taking into account (2.4), (X.2.3), (X.2.6) and (X.2.7), (X.2.5), and (2.1) one can immediately derive that

$$\beta_{v,3v-2} = -(\sqrt{2}\sqrt{3}/\sqrt{5})v(3v-1)(2v+1)(3v-8) \quad (v\geqslant 2). \tag{2.6}$$

The calculation of the eigenvalue expression for $\beta_{v,3v-3}$ proceeds in exactly the same manner as for $\beta_{v,3v-2}$. First of all, from Eq. (XI.2.2) follows the explicit expression of the matrix element $\langle v,3v|O^{+3}_{3v-3}|v,3v-3\rangle$, from which we derive, taking into account the phase convention (X.2.12),

$$b_{v,3v-3}[3v] = \langle v,3v-3|O^{-3}_{3v}|v,3v\rangle = 18\sqrt{3}v^2(3v-1)(3v-2)(6v+1) \times[(v-1)(v-2)(2v-1)(6v-1)/ \times(6v-5)]^{1/2}. \tag{2.7}$$

Out of the relations (XI.2.14), (XI.2.15), and the "$l\rightarrow -l-s-1$" transformed relation (XI.2.13) (the transformation rule "$l\rightarrow -l-s-1$" is explained in XI and Ref. 6) we can eliminate the product operators $O^{+2}_{l-5}P^{-5}_l$ and $O^{+1}_{l-4}P^{-4}_l$. Letting the obtained result act upon $|v,3v\rangle$, and using (2.7), (X.2.10), and (X.2.14), one finally obtains

$$\beta_{v,3v-3} = -(\sqrt{2}\sqrt{3}/\sqrt{5})(2v-1)(9v^3-18v^2-43v+6) \quad (v>2). \tag{2.8}$$

To complete the $|v,3v-3\rangle$ analysis, we give the expression for the remaining matrix element, which is straightforwardly deduced by suitable combination of the relations (XI.2.5–2.8):

$$b_{v,3v-3}[3v-2] = \langle v,3v-3|O^{-1}_{3v-2}|v,3v-2\rangle = -2\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}(3v-2)(9v+1)(2v-1)[(v-2)(3v-1)(6v-1)/(6v-5)]^{1/2}. \tag{2.9}$$

It has to be noticed that, as was the case for $a_{v,3v-3}[3v]$ and $a_{v,3v-3}[3v-2]$ [see (X.2.8) and (X.2.9)], $b_{v,3v-3}[3v]$ and $b_{v,3v-3}[3v-2]$ also vanish when $v=2$, showing that the particular state $|2,3\rangle$ does not exist. For this reason $v=2$ has been excluded already in (2.8).

3. ANALYSIS OF THE $l=3v-4$ STATES

It is known that, for $v>3$, a twofold degeneracy exists. In X it was possible to define two orthonormal states as the eigenstates of P^0_{3v-4}, i.e., $|v,3v-4,\alpha^{(1)}_{v,3v-4}\rangle$ and $|v,3v-4,\alpha^{(2)}_{v,3v-4}\rangle$. But since O^0_l and P^0_l do not commute, these states will not be eigenstates of O^0_{3v-4}. However, by the aid of the relations introduced in XI, we can calculate the action of O^0_{3v-4} on such states, which we can formally denote as

$$O^0_{3v-4}|v,3v-4,\alpha^{(1)}_{v,3v-4}\rangle = \mu_{11}|v,3v-4,\alpha^{(1)}_{v,3v-4}\rangle + \mu_{12}|v,3v-4,\alpha^{(2)}_{v,3v-4}\rangle, \tag{3.1}$$

$$O^0_{3v-4}|v,3v-4,\alpha^{(2)}_{v,3v-4}\rangle = \mu_{21}|v,3v-4,\alpha^{(1)}_{v,3v-4}\rangle + \mu_{22}|v,3v-4,\alpha^{(2)}_{v,3v-4}\rangle. \tag{3.2}$$

Once these results are achieved, the calculation of the $\beta^{(i)}_{v,3v-4}$ and of the eigenstates $|v,3v-4,\beta^{(i)}_{v,3v-4}\rangle$ is obvious. A first equation useful in the derivation of the μ_{ij} coefficients in (3.1) and (3.2) is found by eliminating the $O^{+1}_{3v-5}P^{-5}_{3v}|v,3v\rangle$ and $O^{-1}_{3v-3}P^{-3}_{3v}|v,3v\rangle$ terms out of Eqs. (XI.2.16), (XI.2.18), and the "$l\rightarrow -l-s-1$" transformed Equation (XI.2.17), all acting upon the $|v,3v\rangle$ state, i.e.,

$$O^0_{3v-4}|+\rangle = -\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\frac{(3v-2)}{(2v-1)}(12v^4-16v^3-99v^2+102v+5)|+\rangle - \frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\frac{(3v-2)}{(2v-1)}\sqrt{\Gamma}|-\rangle, \tag{3.3}$$

where Γ is defined in (X.3.5), and $|+\rangle$ and $|-\rangle$ are shorthand notations for

$$|\pm\rangle = a^{(1)}_{v,3v-4}[3v]\,|v,3v-4,\alpha^{(1)}_{v,3v-4}\rangle \pm a^{(2)}_{v,3v-4}[3v]\,|v,3v-4,\alpha^{(2)}_{v,3v-4}\rangle. \tag{3.4}$$

It is possible to derive from the relations in XI, which shift the l-value with -4 and -2, a second useful equation, namely

$$O^0_{3v-4}|-\rangle = \frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\frac{(3v-2)}{(2v-1)}\{(720v^6 - 5184v^5 + 12408v^4 - 16320v^3 + 7841v^2 + 760v - 25)|+\rangle/\sqrt{\Gamma} - (12v^4 - 88v^3 + 141v^2 - 120v + 5)|-\rangle\}. \tag{3.5}$$

The μ_{ij} coefficients in (3.1) and (3.2) follow directly from (3.3)–(3.5). The eigenvalue of O^0_{3v-4} can be deduced as well from Eqs. (3.3)–(3.5). One finds

$$\begin{cases} \beta^{(i)}_{v,3v-4} = -\dfrac{\sqrt{2}\sqrt{3}}{\sqrt{5}}(3v-2)\{(6v^3 - 23v^2 - v - 5) + (-1)^{i-1}\sqrt{\gamma}\} \quad (i=1,2),(v>3), \\ \gamma = 144v^4 - 720v^3 + 1740v^2 - 660v + 25. \end{cases} \tag{3.6}$$

The eigenstates of O^0_{3v-4} are expressible in terms of the eigenstates of P^0_{3v-4} as follows:

$$|v,3v-4,\beta^{(i)}_{v,3v-4}\rangle = \sum_j \lambda_{ij}|v,3v-4,\alpha^{(j)}_{v,3v-4}\rangle \quad (i,j=1,2). \tag{3.7}$$

Herein $\Lambda = (\Lambda_{ij})$ is the matrix of an orthogonal transformation:

$$\lambda_{11} = (1/\sqrt{2})[1 + (-144v^5 + 936v^4 - 1992v^3 + 3030v^2 - 655v - 25)/\sqrt{\Gamma}\sqrt{\gamma}]^{1/2}, \tag{3.8}$$

$$\lambda_{12} = (1/\sqrt{2})[1 - (-144v^5 + 936v^4 - 1992v^3 + 3030v^2 - 655v - 25)/\sqrt{\Gamma}\sqrt{\gamma}]^{1/2}, \tag{3.9}$$

$$\lambda_{21} = \lambda_{12} \quad \text{and} \quad \lambda_{22} = -\lambda_{11}.$$

Again it is possible to exploit only the relations of the mixed type to extract the $b^{(i)}_{v,3v-4}$ coefficients which agree with the phase convention fixed in X for the $a^{(i)}_{v,3v-4}$ coefficients. Without going into the details of the straightforward but lengthy calculations, we just mention the following results:

$$b^{(i)}_{v,3v-4}[3v-2] = 6\sqrt{2}\sqrt{3}(v-1)(3v-1)(3v-2)[(3v-2)(6v-1)(2v-1)/(6v-7)]^{1/2} \times\{(2v^3 - 8v^2 + 6v + 5) + (-1)^{i-1}(-24v^5 + 156v^4 - 382v^3 + 640v^2 - 680v + 175)/\sqrt{\gamma}\}^{1/2}, \tag{3.10}$$

$$b^{(i)}_{v,3v-4}[3v-3] = 3\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}(v-1)[(2v-1)(3v-1)(3v-2)(6v-5)/(v-2)(6v-7)]^{1/2} \times\{(108v^4 - 344v^3 - 259v^2 + 1036v - 12) + (-1)^{i-1}(1296v^6 - 11208v^5 + 30192v^4 - 21758v^3 - 3049v^2 - 13460v + 5820)/\sqrt{\gamma}\}^{1/2}, \tag{3.11}$$

whereby γ has the same meaning as in (3.6).

In order to complete the discussion of the $|v,3v-4\rangle$ states, we can draw attention upon the fact that, for $v=3$,

$$\lambda_{11} = 1 \quad \text{and} \quad \lambda_{12} = 0.$$

This shows that the unique nondegenerate $|3,5\rangle$ state, which is an eigenstate of P^0_{3v-4}, is also an eigenstate of O^0_{3v-4}, and we can write

$$P^0_{3v-4}|3,5\rangle = \alpha_{3,5}|3,5\rangle, \quad O^0_{3v-4}|3,5\rangle = \beta_{3,5}|3,5\rangle,$$

where

$$\alpha_{3,5} = -2^5\times 3\times 5\sqrt{3} \tag{3.12}$$

and

$$\beta_{3,5} = -2^3\times 3\times 7\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}. \tag{3.13}$$

4. ANALYSIS OF THE $l = 3v - 5$ STATES

The analysis of the $|v,3v-5\rangle$ states is quite analogous to that of the $|v,3v-4\rangle$ states in Sec. 3. Combining the suitable relations of XI, and using the results of X, we end up again with two equations ($v>4$)

$$O^0_{3v-5}|+\rangle = -\frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\{(54v^5 - 171v^4 + 21v^3 - 551v^2 + 977v - 270)|+\rangle + 5\sqrt{\Psi}|-\rangle\}/(3v+1), \tag{4.1}$$

$$O^0_{3v-5}|-\rangle = \frac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\{(1458v^7 - 24462v^6 + 2376v^5 + 148770v^4 - 304728v^3 + 264352v^2 - 97706v + 12100)|+\rangle/\sqrt{\Psi} - (54v^5 - 333v^4 - 114v^3 + 979v^2 - 988v + 222)|-\rangle\}/(3v+1), \tag{4.2}$$

where Ψ is defined in (X.4.1), and $|+\rangle$ and $|-\rangle$ are short-hand notations for

$$|\pm\rangle = a^{(1)}_{v,3v-5}[3v]\,|v,3v-5,\alpha^{(1)}_{v,3v-5}\rangle \pm a^{(2)}_{v,3v-5}[3v]\,|v,3v-5,\alpha^{(2)}_{v,3v-5}\rangle. \tag{4.3}$$

Out of these equations, the calculation of the $\beta^{(i)}_{v,3v-5}$ is straightforward:

$$\begin{cases} \beta^{(i)}_{v,3v-5} = -\dfrac{1}{2}\dfrac{\sqrt{2}\sqrt{3}}{\sqrt{5}}\left[(36v^4-180v^3+29v^2+133v-48)+(-1)^{i-1}\sqrt{\psi}\right], \\ \psi = 2916v^6-324v^5+1197v^4+18798v^3-41927v^2+20176v+64 \quad (v>4, i=1,2). \end{cases} \tag{4.4}$$

Once again we could also report on the $b^{(i)}_{v,3v-5}$ values. However, since the way of solving this problem is completely analogous to the $l=3v-4$ case, we do not like to insist on these results here. To finish the analysis of all $|v,3v-5\rangle$ states we have still to consider the two particular cases $v=3$ and $v=4$. However, as was the case in the derivation of the $P^{0,5}_{l}$ the formulas (4.4) hold for $v=4$ as well as for $v=3$; since the $b^{(i)}_{v,3v-5}$'s have not been reported, one cannot check for which superscript the $\beta^{(i)}_{v,3v-5}$ produces the exact $v=3$ and $v=4$ eigenvalue. Nevertheless, this problem will be solved by means of an alternative method, presented in the following section.

5. A NUMERICAL DERIVATION OF THE O^0_l EIGENVALUES

In some previous papers[7,8] we have shown that the quadrupole R(3) scalar shift operator could be expressed in a biquadratic form in the generators of R(5). By an analogous reasoning one can verify that O^0_l, introduced for the R(7) group in V, can be expressed as follows:

$$O^0_l = 2^4\sqrt{2}7^2\left[((b_3^+b_3)^3(b_3^+b_3)^1)^2((b_3^+b_3)^1(b_3^+b_3)^1)^2\right]^{(0)}. \tag{5.1}$$

The R(7) generators $(b_3^+b_3)^3$ and $(b_3^+b_3)^1$ are denoted here in terms of the octupole phonon creation and annihilation operators. In order to derive numerical eigenvalues of O^0_l the right-hand side of (5.1) has to be brought into the so-called canonical form introduced in Ref. 9, i.e.,

$$\begin{aligned}
\left(\frac{15}{2}\right)^{1/2} O^0_l = & -2^2.3^2.5O + 2^2.3^2.5O^0_0 + 2^2.3.17O^2_2 + 2.3.5.11O^4_4 - 2^2.3.5.11O^6_6 \\
& - 2.3^2.7.29/11O^{21}_{21} - 2^3.3^3.5^2.7/(11.13)O^{03}_{03} - 2.3^3.7^2.17/(11.13)O^{23}_{23} \\
& + 2.3^2.113/11O^{24}_{24} + 3^3.7O^{25}_{25} + 3^3.5.11/2^3O^{46}_{46} + 2^2.3^2.11.13/17O^{47}_{47} \\
& - 3^4.5O^{69}_{69} + 2^2.3^5.5.7\sqrt{5}/(11.13)(O^{03}_{23} + O^{23}_{23}) + 117.47O^{030}_{030} + 97.798O^{230}_{230} - 294.07O^{032}_{032} + 43.685O^{212}_{212} - 28.788O^{213}_{213} \\
& + 16.295O^{034}_{034} - 176.63O^{214}_{214} - 12.710O^{234}_{234} - 11.081O^{235}_{235} - 232.97O^{036}_{036} \\
& - 237.29O^{236}_{236} + 82.820O^{246}_{246} + 74.116O^{247}_{247} + 27.715O^{258}_{258} + 123.03O^{468}_{468} \\
& - 2^2.7O^{469}_{469} + 43.046O^{4710}_{4710} - 2.3^3O^{6912}_{6912} - 236.43(O^{030}_{230} + O^{230}_{030}) \\
& + 124.14(O^{032}_{212} + O^{212}_{032}) + 128.11(O^{034}_{214} + O^{214}_{034}) - 211.15(O^{034}_{234} + O^{234}_{034}) + 42.965(O^{214}_{234} + O^{234}_{214}) \\
& + 375.22(O^{036}_{236} + O^{236}_{036}) + 62.696(O^{036}_{246} + O^{246}_{036}) - 29.860(O^{236}_{246} + O^{246}_{236}) - 54.765(O^{258}_{468} + O^{468}_{258}).
\end{aligned} \tag{5.2}$$

The coefficients of the fourth-order terms are numerically derived and are by this only accurate up to the fifth digit.

The expression (5.1) can now be applied to previously constructed octupole-phonon states.[10] These calculations involve quite a lot of tedious Racah algebra, which we shall not discuss here. Examples of similar calculations, but for the quadrupole case, have been given elsewhere.[7] As long as the considered states are nondegenerate with respect to the angular momentum value, the numerically constructed states[10] are also eigenstates of O^0_l and the derived eigenvalues are in accordance with the results (2.1), (2.6), (2.8), and (3.13). If, however, an n-fold degeneracy occurs the action of O^0_l given in (5.2) upon one of the numerically derived states[10] provides us with a linear combination of the n existing states. Let us treat as an example and as a test for the expression (5.2) the case of the two $v=4$, $l=8$ $(3v-4)$ states. The state vectors which are simultaneous eigenvectors of the phonon number operator N, the R(7) Casimir operator V^*, the angular momentum L^2, and its projection l_0 have been constructed in Ref. 10 and can be denoted by

$$|4,8\rangle_1 = 0.41153|0,3,2,5,8\rangle, \tag{5.3}$$

$$|4,8\rangle_2 = -0.10365|0,3,2,5,8\rangle + 0.30574|0,3,4,6,8\rangle. \tag{5.4}$$

The kets on the right-hand sides of (5.3) and (5.4) are defined in formula (2.1) of Ref. 10. After lengthy numerical calculations, where computer assistance is unavoidable, the application of (5.2) to (5.3) and (5.4) gives the following results:

$$\left(\frac{15}{2}\right)^{1/2} O^0_l|4,8\rangle_1 = -3861.3|4,8\rangle_1 - 1037.4|4,8\rangle_2, \tag{5.5}$$

$$\left(\frac{15}{2}\right)^{1/2} O^0_l|4,8\rangle_2 = -1037.4|4,8\rangle_1 + 3441.3|4,8\rangle_2, \tag{5.6}$$

Due to the fact that O^0_l is a Hermitian operator, one expects that

$${}_2\langle 4,8|O^0_l|4,8\rangle_1 = {}_1\langle 4,8|O^0_l|4,8\rangle_2$$

a condition, which within the accuracy of our numerical calculations, is fulfilled. It is now a matter of straightforward calculation to determine out of (5.5) and (5.6) the $|v=4,8\ \beta^{(i)}_{4,8}\rangle$ orthonormalized wavefunctions which are simultaneously eigenstates of N, V^*, L^2, l_0, and O_l^0, together with the corresponding eigenvalues. The following results have been obtained:

$$\beta_{4,8}^{(i)} = \left(\frac{2}{15}\right)^{1/2}(-210.0 + (-1)^{i-1}3795.8) \quad (i = 1,2) \tag{5.7}$$

and

$$|v = 4, l = 8, \beta_{4,8}^{(1)}\rangle = 0.13797|4,8\rangle_1 - 0.99044|4,8\rangle_2, \tag{5.8}$$

$$|v = 4, l = 8, \beta_{4,8}^{(2)}\rangle = 0.99044|4,8\rangle_1 + 0.13797|4,8\rangle_2. \tag{5.9}$$

The exact result for the O_l^0 eigenvalues follows from (3.6) and is

$$\beta_{4,8}^{(i)} = \left(\frac{2}{15}\right)^{1/2}(-210 + (-1)^{i-1}30\sqrt{16\,009}) \quad (i = 1,2),$$

showing that the numerical result is very close to the exact value, namely within 10^{-5} of relative difference.

By this method we also found

$$\beta_{3,4} = \left(\frac{2}{15}\right)^{1/2}(-294) \quad \text{and} \quad \beta_{4,7} = \left(\frac{2}{15}\right)^{1/2}(-3276).$$

It is easy to verify numerically that these values correspond to the $\beta_{v,3v-5}^{(1)}$ $(v = 3,4)$ eigenvalues as given in (4.4).

6. CONCLUSION

We have shown that by a suitable combination of relations between scalar and nonscalar shift opertor products in R(7) part of the eigenvalue spectrum of the scalar shift operator O_l^0 and the corresponding eigenvectors can be derived. It has become clear that this could only be achieved after having knowledge of the P_l^0 eigenvalues and P_l^k shift actions. Since these last results are only derived up to the case $l = 3v - 5$, we also have limited our investigation here at this point.

From group-theoretical principles one knows that the reduction of symmetric irreducible R(7) representations to R(3) yields a three missing label problem. In our investigation of the classification of octupole-phonon states we have considered two R(3) scalars P_l^0 and O_l^0. Generally these operators do not mutually commute. This follows quite clearly from (XI.3.1). As a consequence we could not diagonalize O_l^0 and P_l^0 simultaneously, and thus no set of orthogonal phonon states could be generally constructed such that they should be eigenstates of both O_l^0 and P_l^0. Moreover, P_l^0 and O_l^0 applied to any state with total angular momentum 0 or 1 yield zero eigenvalues. This can be verified from their explicit forms.

The shift operator terms which we have considered in the present series of papers have all the special form to be composed of one component of a R(3) tensor representation and a number of generators l_i $(i = 0, \pm)$ of R(3). If we like to construct and to investigate other scalars in the R(7) reduction to R(3) with the present techniques their terms should be certainly composed with the help of two or more components of the available R(3) tensor representations. This investigation falls outside the scope of the present work.

ACKNOWLEDGMENT

We like to express our gratitude to Professor C. C. Grosjean for his continuous interest.

[1]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 1973 (1980).
[2]G. Vanden Berghe and H. E. De Meyer, J. Math. Phys. **21**, 2632 (1980).
[3]J. Van der Jeugt, H. E. De Meyer, and G. Vanden Berghe, J. Math. Phys. **24**, 1339 (1983).
[4]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 2635 (1980).
[5]H. E. De Meyer, G. Vanden Berghe, and J. Van der Jeugt, J. Math. Phys. **23**. 1223 (1982).
[6]G. Vanden Berghe, H. E. De Meyer, and J. Van der Jeugt, Lett. Nuovo Cimento, **33**, 120 (1982).
[7]G. Vanden Berghe and H. E. De Meyer, Lett. Nuovo Cimento **26**, 637 (1979).
[8]G. Vanden Berghe and H. E. De Meyer, J. Phys. A. **13**, 1543 (1980).
[9]H. E. De Meyer and G. Vanden Berghe, J. Math. Phys. **21**, 486 (1980).
[10]G. Vanden Berghe and H. E. De Meyer, Nucl. Phys. A **323**, 302 (1979).

Eigenvalues of the Chandrasekhar–Page angular functions

Kevin G. Suffern[a)]
School of Mathematics and Physics, Macquarie University, North Ryde, Australia 2113
Edward D. Fackerell and Christopher M. Cosgrove[b)]
Department of Applied Mathematics, University of Sydney, Sydney, Australia 2006

(Received 2 June 1981; accepted for publication 23 April 1982)

The Chandrasekhar–Page angular functions for the Dirac equation in the Kerr–Newman background are expanded as series of hypergeometric polynomials, and a three-term recurrence relation is derived for the coefficients in these series. This leads to a transcendental equation for the determination of the separation constant which is obtained initially as a power series and is then iterated by the method of Blanch and Bouwkamp.

PACS numbers: 97.60.Lf, 11.10.Qr, 02.30.Mv, 02.70. + d

1. INTRODUCTION

In 1976 an important advance was made in understanding processes involving black holes when Chandrasekhar[1] separated the Dirac equation in the Kerr geometry and when Page[2] and independently Toop[3] separated the Dirac equation in the Kerr–Newman geometry. This separation has been obtained more directly by Carter and McLenaghan[4] using Carter's symmetric tetrad for the Kerr–Newman geometry. In addition, Carter and McLenaghan have shown the existence of a generalized angular momentum operator which commutes with the Dirac operator for the Kerr–Newman geometry.

However, discussion of the rate at which massive Dirac particles are emitted by charged rotating black holes has been hindered by the absence of a satisfactory method of obtaining the separation constant λ which is to be determined as the eigenvalue of the coupled system of first-order equations

$$\left(\frac{d}{d\theta}+a\sigma\sin\theta+\frac{m}{\sin\theta}+\tfrac{1}{2}\cot\theta\right)S_{1/2}(\theta) = -(\lambda-a\mu\cos\theta)S_{-1/2}(\theta) \tag{1}$$

and

$$\left(\frac{d}{d\theta}-a\sigma\sin\theta-\frac{m}{\sin\theta}+\tfrac{1}{2}\cot\theta\right)S_{-1/2}(\theta) = (\lambda+a\mu\cos\theta)S_{1/2}(\theta) \tag{2}$$

for the angular functions $S_{1/2}(\theta)$ and $S_{-1/2}(\theta)$. These functions and their defining differential equations were originally introduced in a slightly different notation by Chandrasekhar[1] in his analysis of the uncharged case. However, as Page[2] and also Toop[3] discovered, the same equations are also valid for the charged case. Here a is the Kerr parameter and μ is the rest mass of the Dirac particle, expressed, along with all other quantities in this paper, in the Planck units described in Page.[5] We take the time and azimuthal dependence in the form $\exp(i\sigma t+im\phi)$, where σ is the energy as measured at infinity and m (half an odd integer) is the axial angular momentum. The quantities θ, ϕ, and t here are the Boyer and Lindquist coordinates for the Kerr–Newman metric.[6]

The difficulty of determining λ is only accentuated if, following Chandrasekhar and Page, we eliminate $S_{1/2}(\theta)$ to obtain the second-order equation

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{ds}{d\theta}\right)+\frac{a\mu\sin\theta}{\lambda+a\mu\cos\theta}\frac{ds}{d\theta} + \left[(\tfrac{1}{2}-a\sigma\cos\theta)^2-\left(\frac{m-\frac{1}{2}\cos\theta}{\sin\theta}\right)^2 - \tfrac{3}{4}-2a\sigma m-a^2\sigma^2+\frac{a\mu(\frac{1}{2}\cos\theta-a\sigma\sin^2\theta-m)}{\lambda+a\mu\cos\theta} - a^2\mu^2\cos^2\theta+\lambda^2\right]S=0 \tag{3}$$

for $S=S_{-1/2}(\theta)$. Since this equation has an apparent singularity (Ince[7], p. 406) at $\cos\theta=-\lambda/a\mu$, the eigenvalue appears nonlinearly and a straightforward expansion of S in terms of $\cos\theta$ by the Frobenius method leads to a seven-term recurrence relation, and even if the appropriate behavior at the irregular singular point is taken out, a four-term recurrence relation remains.

Now since the angular functions $S_{1/2}(\theta)$ and $S_{-1/2}(\theta)$ reduce to the appropriate Teukolsky spin weighted angular spheroidal functions when $\mu=0$ and since Fackerell and Crossman[8] have shown that three-term recurrence relations may be obtained for the coefficients in suitable expansions of the Teukolsky functions, we are led to ask whether there is some way of expanding the Chandrasekhar–Page angular functions so that the expansion coefficients satisfy a three-term recurrence relation.

The purpose of this paper is to derive such expansions for $S_{1/2}(\theta)$ and $S_{-1/2}(\theta)$ and to use the resulting three-term recurrence relation to develop continued fraction techniques for the determination of λ in terms of m, $a\sigma$, $a\mu$, and the total angular momentum j of the particle. We first became aware of the possibility of carrying out this program when one of us (C.M.C.) showed that Eq. (3) could be manipulated to give rise to a third-order equation for S which led to three-term recurrences. This treatment is described in the Appendix. However, it then became clear that an easier treatment was possible starting directly from the first-order equation (1) and (2).

A brief outline of the remainder of the paper is as follows: Sections 2 to 4 describe the necessary transformations of Eqs. (1) and (2), the solutions when $a=0$, the expansion of the solutions in terms of hypergeometric polynomials, and

[a)] Permanent address: School of Computing Sciences, The New South Wales Institute of Technology, P. O. Box 123, Broadway 2007, Australia.
[b)] Present address: Department of Mathematics, Clarkson College, Potsdam, New York 13676.

 0022-2488/83/051350-09$02.50

the derivation of the three-term recurrence relation. The transcendental equation for λ is then derived in Sec. 5 and the enumeration of the angular momentum modes is discussed in Sec. 6. Some special solutions of the transcendental equation for λ are discussed in Sec. 7 and a general power series for λ is obtained in Sec. 8. Finally, Secs. 9 and 10 cover the use of the iteration scheme of Blanch and Bouwkamp and discuss the results of the numerical calculations.

2. TRANSFORMATION OF THE DIFFERENTIAL EQUATIONS

We have found it convenient to introduce a new independent variable

$$x = \tfrac{1}{2}(1 + \epsilon_1 \cos \theta)$$

where $\epsilon_1 = m/|m|$, and to introduce the new dependent variables

$$U(x) = x^{-p/2}(1-x)^{-(p+1)/2}S_{1/2},$$

and

$$V(x) = (1-x)^{-p/2}x^{-(p+1)/2}S_{-1/2},$$

where

$$p = \epsilon_1 m - \tfrac{1}{2} = |m| - \tfrac{1}{2}.$$

Then $U(x)$ and $V(x)$ satisfy the coupled differential equations

$$\left[(1-x)\frac{d}{dx} - (p+1) - 2\epsilon_1 a\sigma(1-x)\right]U(x) = \epsilon_1[\lambda + \epsilon_1 a\mu(1-2x)]V(x), \tag{4}$$

and

$$\left[x\frac{d}{dx} + p + 1 + 2\epsilon_1 a\sigma x\right]V(x) = -\epsilon_1[\lambda - \epsilon_1 a\mu(1-2x)]U(x). \tag{5}$$

3. SOLUTIONS FOR $a = 0$

The key to our goal of obtaining expansions that give rise to a three-term recurrence relation for the coefficients is to expand $U(x)$ and $V(x)$ in terms of the solutions of Eqs. (4) and (5) when $a = 0$. It is convenient to introduce the functions

$$Q_n^+(x) = {}_2F_1(-n, n+2p+2, p+1, x),$$
$$Q_n^-(x) = {}_2F_1(-n, n+2p+2, p+2, x),$$

where ${}_2F_1(a,b,c,x)$ is the usual Gauss hypergeometric function. Solutions for $a = 0$ are then given by

$$U(x) = Q_N^+(x)$$

and

$$V(x) = -\frac{\lambda_N(0)}{p+1}Q_N^-(x),$$

with eigenvalue

$$\lambda = \lambda_N(0) = \epsilon_\lambda(N + p + 1), \tag{6}$$

where

$$\epsilon_\lambda = \operatorname{sgn} \lambda,$$

and N is a nonnegative integer.

The hypergeometric polynomials Q_n^+ and Q_n^- satisfy a number of useful identities, including

$$Q_n^+(1-x) = (-1)^n\frac{n+p+1}{p+1}Q_n^-(x), \tag{7}$$

$$Q_n^-(1-x) = (-1)^n\frac{p+1}{n+p+1}Q_n^+(x),$$

$$\left[(1-x)\frac{d}{dx} - p - 1\right]Q_n^+(x) = -\frac{(n+p+1)^2}{p+1}Q_n^-(x), \tag{8}$$

$$\left[x\frac{d}{dx} + p + 1\right]Q_n^-(x) = (p+1)Q_n^+(x), \tag{9}$$

and the recurrence relations

$$(1-2x)Q_n^+(x) = \frac{(n+2p+2)}{2n+2p+3}Q_{n+1}^+(x) + \frac{(2p+1)Q_n^+(x)}{(2n+2p+1)(2n+2p+3)} + \frac{nQ_{n-1}^+(x)}{2n+2p+1}, \tag{10}$$

$$(1-2x)Q_n^-(x) = \frac{(n+2p+2)(n+p+2)}{(n+p+1)(2n+2p+3)}Q_{n+1}^-(x) - \frac{(2p+1)Q_n^-(x)}{(2n+2p+1)(2n+2p+3)} + \frac{n(n+p)Q_{n-1}^-(x)}{(n+p+1)(2n+2p+1)}, \tag{11}$$

$$xQ_n^-(x) = \frac{(p+1)}{n+p+1}\left[\frac{-(n+2p+2)}{2(2n+2p+3)}Q_{n+1}^+(x) + \frac{(2p+1)(n+p+1)Q_n^+(x)}{(2n+2p+1)(2n+2p+3)} + \frac{nQ_{n-1}^+(x)}{2(2n+2p+1)}\right] \tag{12}$$

and

$$(1-x)Q_n^+(x) = \frac{1}{p+1}\left[\frac{(n+2p+2)(n+p+2)}{2(2n+2p+3)}Q_{n+1}^-(x) + \frac{(2p+1)(n+p+1)^2Q_n^-(x)}{(2n+2p+1)(2n+2p+3)} - \frac{n(n+p)Q_{n-1}^-(x)}{2(2n+2p+1)}\right]. \tag{13}$$

4. EXPANSIONS FOR $U(x)$ AND $V(x)$ AND DERIVATION OF THE RECURRENCE RELATION

Since $S_{1/2}(\pi - \theta)$ is known to be a multiple of $S_{-1/2}(\theta)$ it follows that $V(x)$ must be a multiple of $U(1-x)$. Thus if we expand $U(x)$ in terms of the $Q_n^+(x)$ as

$$U(x) = \sum_{n=0}^{\infty} C_m Q_n^+(x) \tag{14}$$

it follows from Eq. (7) that

$$V(x) = K\sum_{n=0}^{\infty}(-1)^n\frac{(n+p+1)}{p+1}C_n Q_n^-(x), \tag{15}$$

where K is a constant. In fact, K is determined by the requirement that the recurrence relation arising from substituting Eq. (14) and (15) into Eq. (4) be identical with that arising from Eq. (5). In order to obtain the recurrence relations it is necessary to use the identities (8) and (9) and the recurrence relations (10)–(13). We find that the two recurrences are identical provided $K^2 = 1$, and the choice

$$K = -(-1)^N \epsilon_1 \epsilon_\lambda$$

then ensures that the eigenvalue for $a = 0$ agrees with that of Eq. (6). We consequently obtain the unique three-term recurrence relation

$$-\frac{(n+1)\epsilon_1}{2n+2p+3}[a\sigma - (-1)^{N+n}\epsilon_\lambda a\mu]C_{n+1} + \Big[n+p+1-\epsilon_\lambda(-1)^{N+n}\lambda + (2p+1)\epsilon_1\frac{(2a\sigma(n+p+1)+(-1)^{N+n}\epsilon_\lambda a\mu)}{(2n+2p+1)(2n+2p+3)}\Big]C_n + \frac{(n+2p+1)\epsilon_1}{(2n+2p+1)}[a\sigma + (-1)^{N+n}\epsilon_\lambda a\mu]C_{n-1} = 0. \tag{16}$$

5. TRANSCENDENTAL EQUATION FOR λ

The advantage of the three-term recurrence relation (16) over recurrence relations with more than three terms is that it can be written in terms of continued fractions. This provides a simple and accurate method for calculating the separation constant λ by means of a transcendental equation, and such a technique was used very successfully by Fackerell and Crossman[8] in their study of Teukolsky's spin weighted angular spheroidal functions. In this technique an initial value for λ is obtained by a series expansion which is then improved by an iteration technique.

We start by writing Eq. (16) as

$$e_n C_{n+1} - p_n C_n + g_n C_{n-1} = 0, \tag{17}$$

where

$$e_n = -\frac{(n+1)\epsilon_1}{2n+2p+3}[a\sigma - (-1)^{N+n}\epsilon_\lambda a\mu], \tag{18}$$

$$p_n = -(n+p+1) + \epsilon_\lambda(-1)^{N+n}\lambda - \frac{(2p+1)\epsilon_1}{(2n+2p+1)(2n+2p+3)} \times [2a\sigma(n+p+1) + (-1)^{N+n}\epsilon_\lambda a\mu], \tag{19}$$

and

$$g_n = \frac{(n+2p+1)\epsilon_1}{2n+2p+1}[a\sigma + (-1)^{N+n}\epsilon_\lambda a\mu].$$

We next define

$$D_n = e_{n-1}C_n/C_{n-1}$$

so that Eq. (17) becomes

$$D_{n+1} = p_n + q_n/D_n, \tag{20}$$

where

$$q_n = -e_{n-1}g_n. \tag{21}$$

The form of Eq. (20) is such that it can be written as the finite continued fraction

$$D_{n+1} = p_n + \frac{q_n}{p_{n-1}+}\frac{q_{n-1}}{p_{n-2}+}\cdots\frac{q_1}{+p_0}, \tag{22}$$

which terminates at the q_1/p_0 term because q_0 is zero as a consequence of Eqs. (18) and (21).

Equation (20) can also be rearranged to give

$$D_n = q_n/(-p_n + D_{n+1}),$$

which then becomes the infinite continued fraction

$$D_n = \frac{q_n}{-p_n+}\frac{q_{n+1}}{-p_{n+1}}\cdots. \tag{23}$$

We now equate the two continued fractions (22) and (23) [with the substitution $n \to n+1$ in (23)] to obtain the transcendental equation for λ. With $n = N$ this equation is

$$p_N + \frac{q_N}{p_{N-1}+}\frac{q_{N-1}}{p_{N-2}+}\cdots\frac{q_1}{+p_0} = \frac{q_{N+1}}{-p_{N+1}+}\frac{q_{N+2}}{-p_{N+2}+}\cdots. \tag{24}$$

Before we evaluate λ it is convenient to introduce the parameters

$$\alpha = a(\sigma - \mu)$$

and

$$\beta = a(\sigma + \mu) \tag{25}$$

because then the rotation parameter a of the hole and the energy σ of the Dirac particle appear in Eq. (24) only through α and β. Provided they are not too large, α and β would thus make very convenient parameters for expanding λ in a power series because the coefficients in such an expansion would depend only on the angular momentum quantum numbers of the Dirac particle in the gravitational field of the hole. The magnitudes of α and β are discussed in Sec. 8.

When $\epsilon_\lambda = +1$, the quantities p_n of Eq. (19) can be expressed in terms of α and β as

$$p_{N+2l} = -G_{N+2l} - E_{N+2l}\alpha - F_{N+2l}\beta + \lambda,$$

for $l = 0, \pm 1, \pm 2, \cdots$,

where

$$E_{N+2l} = (2p+1)\epsilon_1/2(2N+2p+4l+3),$$
$$F_{N+2l} = (2p+1)\epsilon_1/2(2N+2p+4l+1),$$
$$G_{N+2l} = N+2l+p+1,$$

and

$$p_{N+2l+1} = -G_{N+2l+1} - E_{N+2l+1}\alpha - F_{N+2l+1}\beta - \lambda,$$

where

$$E_{N+2l+1} = (2p+1)\epsilon_1/2(2N+2p+4l+3),$$
$$F_{N+2l+1} = (2p+1)\epsilon_1/2(2N+2p+4l+5),$$
$$G_{N+2l+1} = N+2l+p+2.$$

In addition, the quantities q_n from Eq. (21) become

$$q_{N+2l} = R_{N+2l}\beta^2,$$

where

$$R_{N+2l} = (N+2l)(N+2p+2l)/(2N+2p+4l+1)^2$$

and

$$q_{N+2l+1} = R_{N+2l+1}\alpha^2,$$

with

$$R_{N+2l+1} = (N+2l+1)(N+2p+2l+2)/(2N+2p+4l+3)^2.$$

The continued fractions in Eq. (24) thus exhibit the interesting structure of successive terms alternating in α^2 and β^2.

When $\epsilon_\lambda = -1$ the expressions for p_n and q_n are different from those above, but it is not necessary to consider this case separately. This is because α and β interchange in p_n and q_n when the sign of ϵ_λ is changed and in addition the sign of λ changes. We thus have the symmetry

$$\lambda^{(-)}(\alpha,\beta) = -\lambda^{(+)}(\beta,\alpha). \tag{26}$$

6. ENUMERATION OF THE ANGULAR MOMENTUM MODES

In physical applications of the present theory, for example the emission of leptons from rotating black holes, the physical processes must be summed over all the angular momentum modes that make a significant contribution to the total. Here we enumerate the angular momentum modes and connect the parameter N with the angular momentum quantum numbers of a Dirac particle in a spherical electric field.

The angular momentum eigenstates for the Dirac equation are discussed in most standard texts on relativistic quantum mechanics,[9] and in the spherical case when $a = 0$ Eq. (3) reduces to

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dS}{d\theta}\right) + \left[\lambda^2 - \left(\frac{m - \frac{1}{2}\cos\theta}{\cos\theta}\right)^2 - \frac{1}{2}\right]S = 0.$$

This is the angular part of the Dirac equation in Minkowski space-time with a spherical potential for which the eigenvalues are[9]

$$\lambda = j(j+1) - l(l+1) + \tfrac{1}{4}. \tag{27}$$

where j is the total angular momentum of the particle and l is the orbital angular momentum. The total angular momentum is a positive half-odd integer and there are two helicity states in which j and l are related by

$$j = l \pm \tfrac{1}{2}. \tag{28}$$

In addition, the axial angular momentum m satisfies $-j \leqslant m \leqslant j$ in integer steps, and when $j = l + \frac{1}{2}$ there are solutions for $l = 0,1,2,...$ but when $j = l - \frac{1}{2}$ there is no solution for $l = 0$.

As a consequence of (27) and (28),[9]

$$\begin{aligned}\lambda &= (j+\tfrac{1}{2})\operatorname{sgn}(j-l)\\ &= \pm(j+\tfrac{1}{2}),\end{aligned} \tag{29}$$

so that λ is an integer. Equation (29) is consistent with the expression (6) provided

$$N = j - |m|,$$

and the fact that λ is an integer when $a = 0$ was the reason for introducing N in Eq. (6). Note that the sign of λ in the nonrotating case labels the two helicity states in Eq. (29).

7. THE CASE $\alpha = 0$

When $\alpha = 0$ the eigenvalue λ can be obtained analytically from either the recurrence relation (16) or the transcendental equation (24) since then both continued fractions terminate. The case $\alpha = 0$ corresponds physically to a particle that is marginally bound to the hole because the energy σ, which includes the rest mass μ, is then equal to μ.

We consider both cases $\epsilon_\lambda = \pm 1$ explicitly here and note that the structure of the solution is clearer when the original notation of the recurrence relation (16) is used.

When $\epsilon_\lambda = +1$, Eq. (16) reduces to

$$\left[N+p+1-\lambda+\frac{(2p+1)\epsilon_1 a\mu}{2N+2p+1}\right]C_N + \frac{2(N+2p+1)\epsilon_1 a\mu}{2N+2p+1}C_{N-1} = 0, \tag{30}$$

$$\frac{2N\epsilon_1 a\mu}{2N+2p+1}C_N - \left[N+p+\lambda+\frac{(2p+1)\epsilon_1 a\mu}{2N+2p+1}\right]C_{N-1} = 0, \tag{31}$$

since every coefficient other than C_{N-1} and C_N can be set to zero. Equations (30) and (31) lead to a quadratic equation for λ, specifically

$$(\lambda - \tfrac{1}{2})^2 = (N+p+\tfrac{1}{2})^2 + (2p+1)\epsilon_1 a\mu + a^2\mu^2,$$

which gives

$$\lambda = \tfrac{1}{2} + [(N+p+\tfrac{1}{2})^2 + (2p+1)\epsilon_1 a\mu + a^2\mu^2]^{1/2}. \tag{32}$$

When $N = 0$ this simplifies to

$$\lambda = p + 1 + \epsilon_1 a\mu. \tag{33}$$

For $\epsilon_\lambda = -1$ it is necessary to consider the two cases $N = 0$ and $N > 0$ separately. When $N = 0$ Eq. (16) reduces to

$$\left[p+1+\lambda+\frac{2p+1}{2p+3}\epsilon_1 a\mu\right]C_0 = 0,$$

and since $C_0 \neq 0$ this gives

$$\lambda = -p-1-\frac{2p+1}{2p+3}\epsilon_1 a\mu. \tag{34}$$

When $N > 0$ the relevant equations are

$$\left[N+p+2-\lambda+\frac{(2p+1)a\mu}{2N+2p+3}\right]C_{N+1} + \frac{2(N+2p+2)\epsilon_1 a\mu}{2N+2p+3}C_N = 0,$$

$$\frac{2(N+1)\epsilon_1 a\mu}{2N+2p+3}C_{N+1} - \left[N+p+1+\lambda+\frac{(2p+1)\epsilon_1 a\mu}{2N+2p+3}\right]C_N = 0,$$

and consequently

$$\lambda = \tfrac{1}{2} - [(N+p+\tfrac{3}{2})^2 + (2p+1)\epsilon_1 a\mu + a^2\mu^2]^{1/2}. \tag{35}$$

The above expressions (32)–(35) for λ all reduce to Eq. (6) when the hole is nonrotating.

8. GENERAL SERIES EXPANSION

The general series expansion for λ has the form

$$\lambda = \sum_{r=0}^{R}\sum_{s=0}^{S} C_{rs}\alpha^r\beta^s, \tag{36}$$

where the coefficients C_{rs} depend only on ϵ_λ, N, and m, but we only keep terms up to and including $R + S = 5$. Written out explicitly, this expansion is

$$\lambda = C_{00} + C_{01}\beta + C_{02}\beta^2 + C_{03}\beta^3 + C_{04}\beta^4 + C_{05}\beta^5 + C_{10}\alpha + C_{11}\alpha\beta + C_{12}\alpha\beta^2 + C_{13}\alpha\beta^3 + C_{14}\alpha\beta^4 + C_{20}\alpha^2 + C_{21}\alpha^2\beta + C_{22}\alpha^2\beta^2 + C_{23}\alpha^2\beta^3 + C_{30}\alpha^3 + C_{31}\alpha^3\beta + C_{32}\alpha^3\beta^2 + C_{40}\alpha^4 + C_{41}\alpha^4\beta + C_{50}\alpha^5. \tag{37}$$

We now discuss the magnitudes of α and β that will be required for a study of rotating holes. Since the rotation parameter a is $a = a_* M$ where the dimensionless parameter a_* satisfies $0 \leqslant a_* \leqslant 1$ and M is the mass of the hole, it follows that $\alpha = a_*(M\sigma - M\mu)$ and $\beta = a_*(M\sigma + M\mu)$. In the nonrotating case Page[10] finds the maximum power is emitted for leptons near $M\sigma = 0.2$ and that for $M\sigma = 1.0$ the power is reduced from its maximum value by a factor of $\approx 10^6$. Although rotation of the hole is likely to have some effect on these results, it is unlikely that values of $M\sigma$ considerably greater than 1.0 would be required to obtain the significant part of the spectra for rotating holes. Since $\sigma \geqslant \mu$ for massive particles that escape to infinity, it appears that only values of α and β of the order of unity or less will be required to study most physical processes. Consequently, α and β are suitable expansion parameters for λ.

The most practical way to calculate coefficients C_{rs} is to define the higher order coefficients recursively in terms of the lower order ones. The structure of the continued fractions in Eq. (24) is such that a knowledge of all coefficients up to $(r,s) = (R-2,S)$ and $(r,s) = (R,S-2)$ is sufficient to simultaneously determine the four coefficients with $(r,s) = (R-1,S-1), (R-1,S), (R,S-1)$, and (R,S). In addition, the fact that both continued fractions in Eq. (24) terminate when either α or β is zero means that expressions can be written down for an arbitrary number of coefficients in the first row and column of the expansion.

The coefficients in the first row follow directly from Eqs. (32) and (33) and, with the notation introduced in Sec. 5, we have for $N \geqslant 0$

$$C_{00} = G_N,$$
$$C_{01} = F_N,$$

and for $N > 0$

$$C_{02} = N(N+2p)/(2N+2p+1)^3,$$

with

$$C_{0s} = \tfrac{1}{2}(G_N + G_{N-1}) \sum_{m=0}^{[s/2]} a_{sm}$$

for $s > 2$. Here $[s/2]$ is the integer part of $s/2$ and

$$a_{sm} = \binom{1/2}{s-m}\binom{s-m}{m} 2^{s-2m} \times (F_N + F_{N-1})^{s-2m}(G_N + G_{N-1})^{-s} \times [(F_N + F_{N-1})^2 + 4R_N]^m.$$

When $N = 0$, $C_{0s} = 0$ for all $s \geqslant 2$.

The first column expansion is obtained from Eq. (24), which reduces to

$$\lambda - G_N - E_N\alpha = R_{N+1}\alpha^2/(\lambda + G_{N+1} + E_{N+1}\alpha)$$

when $\beta = 0$. This is a quadratic in λ whose positive root is

$$\lambda = \tfrac{1}{2}[(G_N - G_{N+1}) + (E_N - E_{N+1})\alpha] + \tfrac{1}{2}\{[G_N - G_{N+1} + (E_N + E_{N+1})\alpha]^2 + 4R_{N+1}\alpha^2\}^{1/2},$$

and which can be expanded to give, for $N \geqslant 0$,

$$C_{10} = E_N,$$
$$C_{20} = (N+1)(N+2p+2)/(2N+2p+3)^3,$$

and for $r > 2$,

$$C_{r0} = \tfrac{1}{2}(G_N + G_{N+1}) \sum_{m=0}^{[r/2]} b_{rm},$$

where

$$b_{rm} = \binom{1/2}{r-m}\binom{r-m}{m} 2^{r-2m}(E_N + E_{N+1})^{r-2m} \times (G_N + G_{N+1})^{-r}[(E_N + E_{N+1})^2 + 4R_{N+1}]^m.$$

Three other coefficients are relatively simple and are worthwhile writing out explicitly. These are C_{11}, which is always zero because of the quadratic dependence of the continued fractions on α and β, and C_{21} and C_{12}, which are given by

$$C_{21} = -\frac{(N+1)(N+2p+2)(2p+1)\epsilon_1}{(2N+2p+3)^3(2N+2p+1)(2N+2p+5)}$$

and

$$C_{12} = -\frac{N(N+2p)(2p+1)\epsilon_1}{(2N+2p+1)^3(2N+2p+3)(2n+2p-1)}.$$

The remaining coefficients in the expansion (37) can now be expressed recursively in terms of the above coefficients, but before writing them down it is convenient to define the following quantities:

$$X_{01} = (F_N + F_{N+1})/(G_N + G_{N+1}),$$
$$X_{02} = (G_N + G_{N+1})^{-1}\left[C_{02} - \frac{R_{N+2}}{G_N + G_{N+2}}\right],$$
$$X_{03} = (G_N + G_{N+1})^{-1}\left[C_{03} + \frac{R_{N+2}(F_N - F_{N+2})}{(G_N - G_{N+2})^2}\right],$$
$$Y_{10} = (E_N + E_{N-1})/(G_N + G_{N-1}),$$
$$Y_{01} = (F_N + F_{N-1})/(G_N + G_{N-1}),$$
$$Y_{21} = (G_N + G_{N-1})^{-1}\left[C_{21} + \frac{R_{N-1}(F_N - F_{N-2})}{(G_N - G_{N-2})^2}\right],$$
$$Y_{20} = (G_N + G_{N-1})^{-1}\left[C_{20} - \frac{R_{N-1}}{G_N - G_{N-2}}\right],$$
$$Y_{02} = C_{02}/(G_N + G_{N-1}),$$
$$Y_{12} = C_{12}/(G_N + G_{N-1}),$$
$$Z_{10} = (E_N + E_{N+1})/(G_N + G_{N+1}),$$
$$Z_{20} = C_{20}/(G_N + G_{N+1}),$$
$$Z_{21} = C_{21}/(G_N + G_{N+1}),$$
$$Z_{12} = (G_N + G_{N+1})^{-1}\left[C_{12} + \frac{R_{N+2}(E_N - E_{N+2})}{(G_N - G_{N+2})^2}\right]$$

and

$$Z_{30} = (G_N + G_{N+1})^{-1}\left[C_{30} + \frac{R_{N-1}(E_N - E_{N-2})}{(G_N - G_{N-2})^2}\right].$$

In terms of these quantities we have

$$C_{22} = \frac{R_N}{G_N + G_{N-1}}(Y_{10}^2 - Y_{20}) - \frac{R_{N+1}}{G_N + G_{N+1}}(X_{01}^2 - X_{02}),$$

$$C_{31} = \frac{2R_{N+1}X_{01}Z_{10}}{G_N + G_{N+1}},$$

$$C_{41} = -\frac{R_{N+1}}{G_N + G_{N+1}}(Z_{21} - 2X_{01}Z_{20} + 3X_{01}Z_{10}^2),$$

$$C_{32} = -\frac{R_{N+1}}{G_N + G_{N+1}}(Z_{12} - 2X_{02}Z_{10} + 3X_{01}^2 Z_{10}) + \frac{R_N}{G_N + G_{N-1}}(2Y_{10}Y_{20} - Y_{10}^3 - Z_{30}),$$

$$C_{13} = \frac{2R_N Y_{10}Y_{01}}{G_N + G_{N-1}},$$

$$C_{23} = -\frac{R_{N+1}}{G_N + G_{N+1}}(X_{01}^3 - 2X_{01}X_{02} + X_{03}) - \frac{R_N}{G_N + G_{N-1}}(Y_{21} - 2Y_{20}Y_{01} + 3Y_{10}^2 Y_{01}),$$

and

$$C_{14} = -\frac{R_N}{G_N + G_{N-1}}(Y_{12} - 2Y_{10}Y_{02} + 3Y_{10}Y_{01}^2).$$

Table I displays the expansion coefficients in Eq. (37) rounded to six significant figures for the angular momentum modes with $\epsilon_\lambda = +1$ and $(j,m) = (\frac{1}{2},\frac{1}{2})$, $(\frac{5}{2},\frac{5}{2})$, and $(\frac{9}{2},\frac{7}{2})$. Because of the symmetry relation (26) the coefficients with $\epsilon_\lambda = -1$ satisfy $C_{rs}^{(-)} = -C_{sr}^{(+)}$. Note that for the first two sets of coefficients $N = 0$, and that apart from C_{10}, the second row is zero as a consequence of $R_0 = 0$.

As is evident from Table I, the magnitudes of C_{rs} decrease rapidly as $r + s$ increases, although the decrease is not always monotonic. This behavior is of great assistance in the numerical evaluation of λ since it means that very accurate values of λ can be obtained from the expansion (37) [up to six significant figures, see Sec. 10.] when α and β are small compared with unity. Even when α and β are of order unity the expansion (36) converges very rapidly as $R + S$ is increased.

9. THE ITERATION SCHEME OF BLANCH AND BOUWKAMP

The iteration scheme of Blanch[11] and Bouwkamp[12] can be used to improve the value of λ when the series (37) does not give an accurate enough value for use in numerical calculations. This can typically occur when α and β are of order unity or larger and, in particular, when the resulting value of λ is small compared with unity. The scheme is discussed in Flammer[13] and we describe it here specifically for the transcendental equation (24).

We start by writing (24) as

$$U(\lambda) + V(\lambda) = 0, \tag{38}$$

where

$$U(\lambda) = P_N + \frac{q_N}{p_{N-1} +} \frac{q_{N-1}}{p_{N-2} +} \cdots$$

and

$$V(\lambda) = -\frac{q_{N+1}}{-p_{N+1} +} \frac{q_{N+2}}{-p_{N+2} +} \cdots.$$

TABLE I. Typical expansion coefficients C_{rs}.

		j = ½	m = ½		
1.000000	0.500000	0.000000	0.000000	0.000000	0.000000
0.166667	0.000000	0.000000	0.000000	0.000000	
0.740741(-1)	-0.148148(-1)	-0.987655(-3)	0.987654(-4)		
-0.823045(-2)	0.329218(-2)	-0.642759(-3)			
-0.914495(-3)	0.548697(-3)				
0.508053(-3)					
		j = 5/2	m = 5/2		
3.000000	0.500000	0.000000	0.000000	0.000000	0.000000
0.357143	0.000000	0.000000	0.000000	0.000000	
0.174927(-1)	-0.194363(-2)	-0.308514(-4)	0.158542(-4)		
-0.178497(-2)	0.396660(-3)	-0.403529(-4)			
0.138426(-3)	-0.461421(-4)				
-0.520400(-5)					
		j = 9/2	m = 7/2		
5.000000	0.388889	0.960220(-2)	-0.829819(-3)	0.614681(-4)	-0.354137(-5)
0.318182	0.000000	-0.872927(-3)	0.150876(-3)	-0.167640(-4)	
0.135237(-1)	-0.809108(-3)	0.126534(-3)	-0.609217(-5)		
-0.782361(-3	0.936158(-4)	-0.267375(-5)			
0.286342(-4)	-0.513948(-5)				
0.267182(-6)					

If $\lambda^{(1)}$ is an approximate value of λ obtained from the series (37), a correction $\delta\lambda^{(1)}$ can be obtained from (38) because

$$U(\lambda^{(1)} + \delta\lambda^{(1)}) + V(\lambda^{(1)} + \delta\lambda^{(1)}) = 0,$$

and consequently

$$U(\lambda^{(1)}) + V(\lambda^{(1)}) + \delta U(\lambda^{(1)}) + \delta V(\lambda^{(1)}) = 0. \tag{39}$$

To evaluate δU we use Eq. (20) with n replaced by N and note that when a variation $\delta\lambda$ is made to λ the resulting variation in D_{N+1} is

$$\delta D_{N+1} = \epsilon_\lambda \delta\lambda - \frac{q_N}{D_N^2}\delta D_N.$$

This process can be iterated, but only a finite number of times since we are dealing with the finite continued fraction with $q_0 = 0$. For $N \geqslant 1$ we obtain

$$\delta U = \delta D_{N+1} = \epsilon_\lambda \delta\lambda \left[1 + \sum_{l=1}^{N} \prod_{r=0}^{l-1} \frac{q_{N-r}}{D^2_{N-r}}\right], \tag{40}$$

and when $N = 0$, $U(\lambda)$ is equal to p_0 with the result that

$$\delta U = \delta D_1 = \epsilon_\lambda \delta\lambda. \tag{41}$$

If we denote $V(\lambda)$ by H_{N+1} so that H_{N+1} satisfies

$$H_{N+1} = -q_{N+1}/(-p_{N+1} + H_{N+2}),$$

then a variation of λ in this expression results in

$$\delta H_{N+1} = \frac{H^2_{N+1}}{q_{N+1}}(\epsilon_\lambda \delta\lambda^{(1)} + \delta H_{N+1}).$$

This may now be iterated to give

$$\delta H_{N+1} = \epsilon_\lambda \delta\lambda^{(1)} \sum_{l=1}^{\infty} \prod_{r=1}^{l} \frac{H^2_{N+r}}{q_{N+r}} \tag{42}$$

for $N \geqslant 0$, because in this case there is no limit to the number of iterations.

The correction to $\lambda^{(1)}$ is now obtained by combining Eqs. (39)–(42), and there results

$$\delta\lambda^{(1)} = -\epsilon_\lambda \frac{p_0 + V(\lambda^{(1)})}{1 + \sum_{l=1}^{\infty} \prod_{r=1}^{l} \frac{H^2_{N+r}}{q_{N+r}}} \tag{43}$$

for $N = 0$, and for $N \geqslant 1$

$$\delta\lambda^{(1)} = -\epsilon_\lambda \frac{U(\lambda^{(1)}) + V(\lambda^{(1)})}{1 + \sum_{l=1}^{N} \prod_{r=0}^{l-1} \frac{q_{N-r}}{D^2_{N-r}} + \sum_{l=1}^{\infty} \prod_{r=1}^{l} \frac{H^2_{N+r}}{q_{N+r}}}. \tag{44}$$

Once $\lambda^{(1)}$ has been calculated from either of the above expressions the process can be iterated with $\lambda^{(i+1)} = \lambda^{(i)} + \delta\lambda^{(i)}$.

In expressions (40)–(44), the quantities D_{N-r} are, in obvious notation, the finite continued fractions

$$D_{N-r} = p_{N-r-1} + \frac{q_{N-r-1}}{q_{N-r-2}+} \cdots \frac{q_1}{+p_0}$$

for $0 \leqslant r \leqslant N-1$, and the quantities H_{N+r} are the infinite continued fractions

$$H_{N+r} = -\frac{q_{N+r}}{-p_{N+r}+} \frac{q_{N+r+1}}{-p_{N+r+1}+} \cdots$$

for $r \geqslant 1$.

10. RESULTS

We have calculated some values of λ for the five lowest angular momentum modes for which $j = \frac{1}{2}$ to $\frac{9}{2}$. For each helicity state $\epsilon_\lambda = \pm 1$ two pairs of values for α and β were used, specifically $(\alpha,\beta) = (0.01, 0.02)$ and $(0.5, 1.0)$, and the

TABLE II. Eigenvalues λ for the 5 lowest angular momentum modes for selected values of α and β.

$\epsilon_\lambda = +1 \quad \alpha = 0.01 \quad \beta = 0.02$

m \ j	1/2	3/2	5/2	7/2	9/2
-9/2					4.98591
-7/2				3.98611	4.98905
-5/2			2.98643	3.99009	4.99218
-3/2		1.98700	2.99187	3.99406	4.99531
-1/2	0.98834	1.99567	2.99730	3.99803	4.99845
1/2	1.01167	2.00435	3.00273	4.00180	5.00158
3/2		2.01300	3.00815	4.00596	5.00471
5/2			3.01357	4.00993	5.00784
7/2				4.01389	5.01096
9/2					5.01409

$\epsilon_\lambda = +1 \quad \alpha = 0.5 \quad \beta = 1.0$

m \ j	1/2	3/2	5/2	7/2	9/2
-9/2					4.29756
-7/2				3.30870	4.46676
-5/2			2.32657	3.52651	4.63150
-3/2		1.35984	2.63036	3.73453	4.79235
-1/2	0.44058	1.84225	2.90717	3.93475	4.94973
1/2	1.59764	2.22587	3.17408	4.13127	5.10533
3/2		2.65654	3.43391	4.32468	5.25934
5/2			3.68229	4.51300	5.41065
7/2				4.69685	5.55956
9/2					5.70622

$\epsilon_\lambda = -1 \quad \alpha = 0.01 \quad \beta = 0.02$

m \ j	1/2	3/2	5/2	7/2	9/2
-9/2					-4.98682
-7/2				-3.98723	-4.98975
-5/2			-2.98786	-3.99088	-4.99269
-3/2		-1.98901	-2.99273	-3.99454	-4.99562
-1/2	-0.99170	-1.99636	-2.99759	-3.99819	-4.99855
1/2	-1.00836	-2.00369	-3.00245	-4.00184	-5.00148
3/2		-2.01101	-3.00730	-4.00549	-5.00440
5/2			-3.01215	-4.00914	-5.00733
7/2				-4.01278	-5.01026
9/2					-5.01318

$\epsilon_\lambda = -1 \quad \alpha = 0.5 \quad \beta = 1.0$

m \ j	1/2	3/2	5/2	7/2	9/2
-9/2					-4.34936
-7/2				-3.37371	-4.50494
-5/2			-2.41349	-3.56913	-4.65756
-3/2		-1.48903	-2.67521	-3.75867	-4.80756
-1/2	-0.67315	-1.87948	-2.92301	-3.94336	-4.95513
1/2	-1.47645	-2.23549	-3.16265	-4.12446	-5.10083
3/2		-2.57663	-3.39585	-4.30255	-5.24479
5/2			-3.62219	-4.47702	-5.38684
7/2				-4.64856	-5.52717
9/2					-5.66583

results are shown in Table II. Several values of λ for the higher angular momentum mode with $j = \frac{21}{2}$ were also calculated and these are shown in Table III. In each case the series (37) was used to obtain the initial estimate for λ, which was then iterated by the Blanch–Bouwkamp scheme until $|\delta\lambda / \lambda| < 10^{-7}$ or a maximum of five iterations. Ten terms were kept in the infinite sums in the expressions (43) and (44), since increasing the number beyond 10 had no effect on λ to seven significant figures.

For $(\alpha,\beta) = (0.01,0.02)$ the difference between the values of λ calculated with (37) and the iterated value was at most one part in the seventh significant figure. In fact, for these small values of α and β the three-term expansion

$$\lambda = G_N + E_N\alpha + E_N\beta$$

gives six-figure accuracy for the range of j and m in Tables II and III. The series (37) is naturally not as accurate for $(\alpha,\beta) = (0.5,1.0)$ but, even so, the error in Tables II and III was only about one part in the fifth significant figure. This is a result of the rapid decrease in $|C_{rs}|$ as $r + s$ increases.

An interesting situation occurs for larger values of α and β, where λ can be zero for certain angular momentum modes. For example, it is readily seen from Eq. (33) that λ can be zero when $N = 0$ and $\alpha = 0$, since setting $\lambda = 0$ in this equation gives $\beta = -\epsilon_1(2p + 2)$. Consequently,

$$\beta = 2j + 1$$

when $\epsilon_1 = -1$ because then $m = -j$ when $N = 0$ and so $p = j - \frac{1}{2}$. For $\alpha = 0$ and $m = -j$ there is thus one value of β for which λ is zero for every value of j. This behavior is not restricted to $\alpha = 0$, and for $m = -j = -\frac{1}{2}$ and $m = -j = -\frac{3}{2}$, Fig. 1 shows the curves in the (α,β) plane along which $\lambda = 0$. An infinite number of similar curves exist for higher values of $m = -j$.

As the $\lambda = 0$ curves are approached the series (37) becomes progressively less accurate until quite near the curves the series gives no significant figures at all. However, this does not prevent the iteration (43), with an absolute error criterion of $|\delta\lambda| < 10^{-7}$ from calculating λ to four significant figures in three iterations or less, provided $|\lambda| > 10^{-2}$. Although (37) does not give accurate values of λ near $\lambda = 0$, it does predict the positions of the two $\lambda = 0$ curves in Fig. 1 to within $\beta \lesssim 0.02$.

TABLE III. Selected values of λ for the angular momentum mode with $j = 21/2$

	$\epsilon_\lambda = +1$		$\epsilon_\lambda = -1$	
m	$\alpha = 0.01$ $\beta = 0.02$	$\alpha = 0.5$ $\beta = 1.0$	$\alpha = 0.01$ $\beta = 0.02$	$\alpha = 0.5$ $\beta = 1.0$
-21/2	10.9854	10.2722	-10.9859	-10.2954
-17/2	10.9882	10.4159	-10.9886	-10.4343
-13/2	10.9910	10.5583	-10.9913	-10.5721
-9/2	10.9938	10.6996	-10.9940	-10.7090
-5/2	10.9965	10.8400	-10.9966	-10.8849
-1/2	10.9993	10.9789	-10.9993	-10.9799
1/2	11.0007	11.0481	-11.0007	-11.0472
5/2	11.0035	11.1861	-11.0034	-11.1900
9/2	11.0063	11.3231	-11.0061	-11.3140
13/2	11.0090	11.4592	-11.0087	-11.4461
17/2	11.0118	11.5944	-11.0114	-11.5776
21/2	11.0146	11.7287	-11.0141	-11.7082

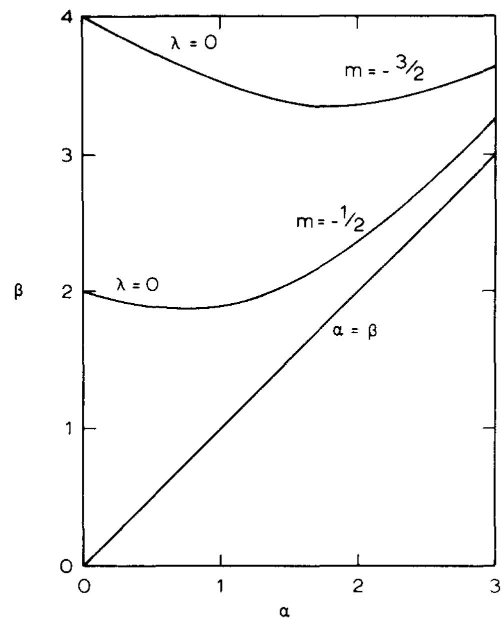

FIG. 1. Curves in the (α,β) plane along which λ is zero for the angular momentum modes with $\epsilon_\lambda = 1$, and $(j,m) = (1/2, -1/2)$ and $(3/2, -3/2)$. For each mode, λ is positive below the curve and negative above it. It is only values of α and β such that $\beta \gg \alpha$ that are physically significant.

For $N > 0$ and $\lambda = 0$ Eq. (32) gives

$$\beta = -(2p + 1)\epsilon_1 + [1 - 4N(N + 2p + 1)]^{1/2},$$

from which it is obvious that no real positive values of β exist when $\alpha = 0$. There is no numerical indication that the situation is different for $\alpha > 0$, with the most likely result that for $\epsilon_\lambda = +1$ the only modes to exhibit $\lambda = 0$ are those with $m = -j$. A similar situation exists for $\epsilon_\lambda = -1$, where Eq. (34) gives

$$\beta = (2j + 1)(j + 1)/j$$

when $N = 0$, $\alpha = 0$, and $\lambda = 0$. An inspection of Eq. (35) readily shows that there are no modes with $N > 0$ and $\lambda = 0$ when $\alpha = 0$.

As a check on the accuracy of the series (37) and the subsequent Blanch–Bouwkamp iteration the eigenvalue was calculated with an effective expansion order of $R + S = 8$ by working directly with Eq. (24). The denominators of the continued fractions in this equation contain series of the form

$$D = \sum_{r=0}^{R} \sum_{s=0}^{S} A_{rs}\alpha^r\beta^s,$$

and the calculation of λ, which was performed by computer, used an algorithm[14] for calculating the coefficients B_{rs} in the corresponding expansion of

$$D^{-1} = \sum_{r=0}^{R} \sum_{s=0}^{S} B_{rs}\alpha^r\beta^s.$$

The resulting expressions for λ were then recursively substituted into Eq. (24) in the same manner as used in Sec. 8 for the evaluation of the coefficients C_{rs}. (This program simply automated the procedure used in Sec. 8 and allowed more terms to be included than could be calculated by hand.)

Computationally this was very inefficient, but after application of the Blanch–Bouwkamp iteration there were no differences in the values of λ in Tables II and III to six significant figures.

11. CONCLUSIONS

Along with the separation of the Dirac equation in the Kerr–Newman background by Chandrasekhar, Page, and Toop, a necessary step in the study of the quantum mechanical processes of massive Dirac particles near rotating black holes is the calculation of the separation constant λ. The double series expansion (37) derived in this paper for λ in terms of the parameters α and β [defined by Eq. (25)] is very accurate considering the small number of terms involved, giving typically four to six significant figures for the values of α and β that are physically important. A reason for this accuracy is the rapid decrease in magnitude of the coefficients in the expansion as the order of the terms increases. When greater accuracy in λ is required than given by the series (37) the iteration method of Blanch and Bouwkamp is very effective in improving the values of λ to at least six-figure accuracy.

With an algorithm for calculating λ now available it is possible to study the emission of massive Dirac particles from rotating black holes in a manner similar to the way Page[10] studied their emission in the nonrotating case. In addition, the expansions (14) and (15) for the angular functions as series of hypergeometric polynomials, and the three-term recurrence relation (16) for the expansion coefficients will be useful in any investigation of the analytic properties of the angular functions.

APPENDIX

In the introduction it was stated that a three-term recurrence relation could be derived from a third-order differential equation related to Eq. (3). In fact it is possible to derive such a third-order equation for functions related to either $S_{1/2}(\theta)$ or $S_{-1/2}(\theta)$, so we introduce the notation $S^{(\epsilon)}$, where $\epsilon = \pm 1$ and $S^{(\epsilon)} = S_{1/2}(\theta)$ if $\epsilon = +1$ and $S^{(\epsilon)} = S_{-1/2}(\theta)$ if $\epsilon = -1$. We also introduce the new independent variable

$$z = \tfrac{1}{2}(1 + \epsilon\epsilon_1 \cos\theta), \tag{A1}$$

and a new dependent variable $w(z)$ defined by

$$w(z) = (1+\cos\theta)^{-(1/2)\epsilon_1(m-(1/2)\epsilon)} \times (1-\cos\theta)^{-(1/2)\epsilon_1(m+(1/2)\epsilon)} e^{-\epsilon\epsilon_1\gamma\cos\theta} S^{(\epsilon)}, \tag{A2}$$

where as before $\epsilon_1 = \operatorname{sgn} m$ and where

$$\gamma = \epsilon_2(\alpha\beta)^{1/2}, \quad \epsilon_2^2 = +1. \tag{A3}$$

Then $w(z)$ satisfies the second-order ordinary differential equation

$$\begin{aligned} z(1-z)w'' &+ \left[4\gamma z(1-z) - \tfrac{1}{2} - (1+\epsilon_1 m)(2z-1) - \frac{2a\mu z(1-z)}{a\mu(2z-1)-\epsilon_1\lambda}\right]w' + \{\lambda^2 - a^2\mu^2 - \gamma \\ &\times [1 + 2(1+\epsilon_1 m)(2z-1)] - 2a\sigma m + \epsilon_1 a\sigma(2z-1) - (\epsilon_1 m + \tfrac{1}{2})^2 - \gamma \\ &\times \frac{4a\mu z(1-z)}{a\mu(2z-1)-\epsilon_1\lambda} + \frac{a\mu}{a\mu(2z-1)-\epsilon_1\lambda} \\ &\times [(2\epsilon_1 m + 1)z + 4\epsilon_1 a\sigma z(1-z)]\}w = 0. \end{aligned} \tag{A4}$$

If we take the combination

$$\frac{d}{dz}(\text{A4}) + \frac{1}{z - \frac{1}{2} - \dfrac{\epsilon_1\lambda}{2a\mu}}(\text{A4}) + \frac{(\gamma + \epsilon_1 a\sigma)\left(1 + \dfrac{\epsilon_1\lambda}{a\mu}\right) + \epsilon_1 m - \frac{1}{2}}{z}(\text{A4})$$

we obtain the required third-order ordinary differential equation, namely,

$$\begin{aligned} &z(1-z)w''' \\ &+ [2p + 2 + A + (-3p - 5 + 4\gamma - A)z - 4\gamma z^2]w'' \\ &\quad + \left[\frac{p(p+1) + (p+1)A}{z} + \lambda^2 - a^2\mu^2 + \gamma(6p+8)\right. \\ &\quad - 2\epsilon_1 a\sigma p - (p+1)(3p+4) \\ &\quad \left. - (4\gamma - 2p - 3)A + (-8\gamma(p+2) - 4\gamma A)z\right]w' \\ &+ \left[\frac{(p+A)(\lambda^2 - a^2\mu^2 + 2)(\gamma - \epsilon_1 a\sigma)(p+1) - (p+1)^2}{z}\right. \\ &\quad \left. - 4\gamma(p+1)(p+2+A)\right]w = 0, \end{aligned} \tag{A5}$$

where, as before, $p = \epsilon_1 m - \frac{1}{2}$ and

$$A = (\gamma + \epsilon_1 a\sigma)(1 + \epsilon_1\lambda/a\mu).$$

If we attempt to solve this equation with the expansion

$$w = k_0 + k_1 z + k_2 z^2 + \cdots$$

we obtain the three-term recurrence relation

$$\begin{aligned} &(n+1)(n+p+1)(n+p+A)k_{n+1} + [-n(n+p+A) \\ &\quad \times (n + 2p + 2 - 4\gamma) \\ &\quad + n(\lambda^2 - a^2\mu^2 + \gamma(2p+4) - 2\epsilon_1 a\sigma p - (p+1)^2) \\ &\quad + (p+A)(\lambda^2 - a^2\mu^2 + 2(\gamma - \epsilon_1 a\sigma)(p+1) \\ &\quad - (p+1)^2]k_n - 4\gamma(n+p)(n+p+1+A)k_{n-1} = 0. \end{aligned} \tag{A6}$$

[1] S. Chandrasekhar, Proc. R. Soc. London A **349**, 571 (1976).
[2] D. N. Page, Phys. Rev. D **14**, 1509 (1976).
[3] N. Toop, unpublished.
[4] B. Carter and R. G. McLenaghan, Phys. Rev. D **19**, 1093 (1979).
[5] D. N. Page, Phys. Rev. D **13**, 198 (1976).
[6] R. H. Boyer and R. W. Lindquist, J. Math. Phys. **8**, 265 (1967).
[7] E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[8] E. D. Fackerell and R. G. Crossman, J. Math. Phys. **18**, 1849 (1977).
[9] Cf. J. S. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).
[10] D. N. Page, Phys. Rev. D **16**, 2402 (1977).
[11] G. Blanch, J. Math. and Phys. **25**, 1 (1946).
[12] C. J. Bouwkamp, J. Math. and Phys. **26**, 79 (1947).
[13] C. Flammer, *Spheroidal Wave Functions* (Stanford U.P., Stanford, Calif., 1957).
[14] K. G. Suffern, "Expand: A Fortran subprogram for performing operations on a finite power series in two variables," preprint, Macquarie University (1981).

A non-Machian solution in Brans–Dicke theory

S. K. Chakraborty
Physics Department, Presidency College, Calcutta, India

(Received 2 April 1982; accepted for publication 25 June 1982)

A solution of the vacuum field equations of the Brans–Dicke theory of gravitation admitting a particular group of motions is found out. The solution is of interest because it is supposed to be at variance with Machian ideas of inertia.

PACS numbers: 98.80.Dr

I. INTRODUCTION

There have been different formulations of Mach's principle originally inspired by the coincidence of the inertial frame and the rest frame of distant stars. It led to the idea that inertia arises in some way from the interaction between different bodies. As an extension of this idea it was sometimes demanded that there should be no regular space–time devoid of all matter. A somewhat weaker formulation was that there should not be any non-Euclidean space–time if there is no matter.

An apparently contrary example in the general theory of relativity was given by Taub[1]—his metric was spatially homogeneous, i.e., admittted a group of motions with three spacelike Killing vectors. He obtained a solution of the equations $R_{\mu\nu}=0$ and though there were singularities, he argued that they could not represent matter because the singular points were equivalent to regular points under a transformation belonging to group of motions. Such solutions were looked upon as non-Machian.

In the present paper, we have addressed ourselves to the corresponding problem in Brans–Dicke theory. True, recent observational studies on the change of periods of pulsars have apparently ruled out this theory, yet from the pedagogic point of view, it is interesting to investigate how this theory, commonly believed to be more consistent with Machian ideas than general relativity, stands so far as the Taub test is concerned. We present here a solution of vacuum Brans–Dicke equations which admits four linearly independent Killing vectors and thus the space–time is both stationary and spatially homogenous. However, the scalar ϕ in Brans–Dicke theory is not constant.

II. THE METRIC AND THE SCALAR FIELD

The line element:

$$ds^2=(dx^0)^2-(dx^1)^2-(dx^2)^2-(1+M^2+N^2-L^2)(dx^3)^2 + 2M\,dx^1dx^3+2N\,dx^2dx^3-2L\,dx^0dx^3, \quad (1)$$

with the scalar field

$$\phi=\left[(A/\alpha^2)(qx^1+rx^2-px^0)+e^{\alpha x^3}(Gx^1+Ex^2+Bx^0) + e^{-\alpha x^3}(Hx^1+Fx^2+Cx^0)^{(1/\omega+1)}\right], \quad (2)$$

where $L=-rx^1+qx^2$, $M=px^2-rx^0$, $N=-px^1+qx^0$, $\alpha^2=q^2+r^2-p^2$, and A, an arbitrary constant, is a solution of the vacuum field equations of the Brans–Dicke theory of gravitation. ω is the coupling constant.

For $\alpha\neq 0$, the above solution is a general one representing a homogeneous space–time allowing the following group of motions;

$$(X_a,X_b)=C^c_{ab}X_c, \quad (3)$$

where C^c_{ab}'s are the structure constants of the group with the following nonzero values:

$$C^1_{30}=r,\quad C^2_{03}=q,\quad C^0_{31}=r,\quad C^2_{31}=p,\quad C^0_{23}=q,\quad C^1_{23}=p. \quad (4)$$

[In defining the operators X_a's the convention is as followed by Landau and Lifshitz.[2] The g_{ab}'s occuring, are chosen as $g_{ab}=(1,-1,-1,-1)$ following the classification of structure constant proposed by Hiromoto and Ozsvath.[3]]

The constants in Eq. (2) satisfy the following conditions:

$$\begin{aligned} &qG+rE+pB=0,\\ &\alpha G-pE+rB=0,\\ &qH+rF+pC=0,\\ &\alpha H+pF+rC=0. \end{aligned} \quad (5)$$

For $\alpha=0$, the scalar field ϕ is given by

$$\phi=\{A\,[(x^3)^2/2](px^0-qx^1-rx^2)+x^3(Fx^0+Bx^1+Dx^2) + Gx^0+Cx^1+Ex^2\}^{1/\omega+1} \quad (6)$$

and the constants satisfy the following relations:

$$\begin{aligned} &qB+rD+pF=0,\\ &pB-rA+qF=0,\\ &pC+D+qG=0,\\ &rC-qE-F=0,\\ &C^2+E^2-G^2=0. \end{aligned} \quad (7)$$

Details of calculations along with some other results may be published elsewhere.

In the present case, $R_{\mu\nu}=0$, so that the metric (1) is a vacuum solution for the general relativity equations also. This has been already noted Hiromoto and Ozavath.[3] Considered from the point of view of Tabensky and Taub,[4] who have shown that any Brans–Dicke solution in vacuum has an alternative interpretation in terms of a massless scalar field or a perfect fluid in irrotational motions, with an equation of state as pressure equals to density, the present case corresponds to a trivial one, i.e., a massless scalar field or a perfect fluid with vanishing energy-stress tensor.

The role of the scalar ϕ is interesting. It is not trivial, i.e., ϕ does neither vanish nor is equal to a constant. This means that the Lie derivative of ϕ with respect to the four Killing vectors does not vanish. However, all the compon-

 0022-2488/83/051359-02$02.50

ents of the tensor $\phi_{\mu;\nu}$ vanish and this makes possible a solution of the Brans–Dicke equations.

ACKNOWLEDGMENTS

The author feels deeply indebted to Professor A. K. Raychaudhuri and Sri N. Bandyopadhyay, Physics Department, Presidency College, Calcutta, for suggesting the problem and extending helpful comments. Thanks are also due to the referee of the *Journal of Mathematical Physics* for his suggestions.

[1]A. Taub, Ann. Math. **53**, 474 (1951).
[2]L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, London, 1971), 3rd. rev. ed., Sec. 112.
[3]E. Hiromoto and I. Ozsvath, Gen. Relativ. Gravit. **9** (4), 299 (1978).
[4]R. Tabensky and A. Taub, Commun. Math. Phys. **29**, 61 (1973).