Ternary composition algebras and Hurwitz's theorem

Ronald Shaw
School of Mathematics, University of Hull, Hull, HU6 7RX, United Kingdom

(Received 2 February 1988; accepted for publication 20 July 1988)

Axioms are proposed for a certain "alternative" kind of ternary composition algebra, termed a 3Cn algebra. The axioms are shown to be (for $n > 2$) in a simple correspondence with the axioms for a ternary vector cross product algebra. The axioms imply that $n = 1, 2, 4$, or 8 (from which the usual Hurwitz theorem is deduced). The existence of 3C8 algebras is demonstrated by an explicit construction in four-dimensional Hilbert space, without appeal to the properties of the algebra of octonions.

I. INTRODUCTION: 3Xn ALGEBRAS AND 3Cn ALGEBRAS

Throughout this paper let E denote a real n-dimensional vector space equipped with a positive definite inner product $\langle , \rangle$. For $n \geqslant 3$ a *ternary vector cross product* for $(E, \langle , \rangle)$ is a map X: $E^3 \to E$ that satisfies the axioms[1,2]

$$\begin{aligned} &\text{X1} \quad X \text{ is trilinear,} \\ &\text{X2} \quad X(a,b,c) \text{ is orthogonal to each of } a,b,c, \\ &\text{X3} \quad \|X(a,b,c)\| = \|a \wedge b \wedge c\|, \end{aligned} \qquad (1.1)$$

for all $a,b,c \in E$. Here $\|a\|^2 = \langle a,a \rangle$ and $\|a_1 \wedge a_2 \wedge a_3\|^2 = \det(\langle a_i,a_j \rangle)$. In such circumstances we refer to the triple $\mathfrak{A} = (E, \langle , \rangle, X)$ as a *3Xn algebra.*

With E and $\langle , \rangle$ as above, but allowing now any dimension $n \geqslant 1$, suppose that there is a map $\{ \ \}$: $E^3 \to E$ that satisfies

$$\begin{aligned} &\text{C1} \quad \{ \} \text{ is trilinear,} \\ &\text{C2} \quad \{aac\} = \langle a,a \rangle c = \{caa\}, \\ &\text{C3} \quad \|\{abc\}\| = \|a\| \, \|b\| \, \|c\|, \end{aligned} \qquad (1.2)$$

for all $a,b,c \in E$. In such circumstances we refer to the triple $\mathfrak{C} = (E, \langle , \rangle, \{ \})$ as a *3Cn algebra.*

It is known[1,2] (see also Theorems 2.3 and 3.2 below) that 3Xn algebras exist only in dimensions $n = 4$ and $n = 8$. Since the properties of 3X4 algebras are readily obtained, as indicated[3] in Sec. II, the chief interest lies in the "exceptional" 3X8 algebras. Since the latter can be defined in terms of the (not-associative) algebra $\mathbb{O}$ of the real octonions, previous authors[2,4,5] have studied the properties of 3X8 algebras by appeal to those of the algebra $\mathbb{O}$. However, a 3X8 algebra has more symmetry than $\mathbb{O}$, the respective automorphism groups being[5,6] Spin(7) and G_2, of dimensions 21 and 14. Consequently one would expect that a study of 3X8 algebras in their own right, without appeal to the properties of $\mathbb{O}$, would increase one's understanding of this area.

These expectations are, in fact, borne out. For example, certain results for binary multiplication (the Hurwitz theorem, the existence of quaternionic subalgebras of $\mathbb{O}$) receive a cleaner formulation in terms of ternary multiplication (see Theorems 3.2 and 4.1 below). In particular, the proof we give of Theorem 3.2, the "3Cn Hurwitz theorem," does *not* involve us in having to choose a preferred unit element $e \in E$, nor in computations with the associated *conjugation* $a \mapsto Ka = \bar{a}$:

$$\bar{a} = 2\langle a,e \rangle e - a. \qquad (1.3)$$

The present paper represents an advance on a previous study[7] of 3Xn algebras in two chief respects. First, as a result of the inclusion of the "alternative axiom" C2 in our present definiton of a 3Cn algebra, we show—see Theorem 2.3—that 3Cn algebras and 3Xn algebras are, for $n \geqslant 3$, in a simple one to one correspondence (under which the automorphism groups of the two algebras are identical). Second, we demonstrate the existence of 3X8 and 3C8 algebras, without appeal to the existence and properties of the algebra $\mathbb{O}$, by displaying, in Sec. IV, surprisingly simple explicit expressions for $X(a,b,c)$ and $\{abc\}$. (It is true that previously[7] we obtained a canonical form for a 3X8 algebra. However, owing to the nonlinearity of property X3, it would have been exceedingly tedious to check directly that the purported ternary vector cross product X really did satisfy Axioms X1–X3.) In this paper we also take the opportunity to fill in many details of proofs that were omitted in Ref. 7.

Remark: We ought to mention the ingenious and intricate investigations of McCrimmon[8] into general ternary composition algebras. The disappointing conclusion of McCrimmon's work is that the omission of Axiom C2 does not, up to isotopy (and up to permutation of the variables a,b,c in $\{abc\}$), lead to anything new.

II. CONSEQUENCES OF THE AXIOMS

In this section we consider some immediate consequences, first, of Axioms X1–X3 and, second, of Axioms C1–C3. We are then able to demonstrate that the two axiom systems deal in effect with the same mathematical object.

Associated with a 3Xn algebra $\mathfrak{A} = (E, \langle , \rangle, X)$ is the *scalar quadruple product* Φ defined by

$$\Phi(a,b,c,d) = \langle a, X(b,c,d) \rangle \qquad (2.1)$$

and also a family $\{T_{a,b};\ a,b \in E\}$ of linear operators on E, where

$$T_{a,b}c = X(a,b,c). \qquad (2.2)$$

Observe that Φ and T are, respectively, quadrilinear and bilinear functions of their arguments.

Lemma 2.1: Functions Φ, X, and T are alternating functions of their arguments. Moreover $T_{a,b} \in \mathrm{Sk}(E,E)$, for each $a,b \in E$.

Proof of Lemma 2.1: By X2, $\Phi(a,b,c,d)$ is zero whenever $a = b$ or $a = c$ or $a = d$. Hence Φ is alternating, and hence so are X and T. Knowing now that $\Phi(a,b,c,d) = -\Phi(d,b,c,a)$ we read off the skew-adjoint property $\widetilde{T}_{b,c} = -T_{b,c}$. (We denote by $\widetilde{A}$ the adjoint of a linear operator A on E: $\langle a,Ad\rangle = \langle\tilde{A}a,d\rangle$.)

Remark: Axioms X2 and X3 clearly rule out the possibility $n = 3$. On the other hand if $n = 4$ we can make $(E, \langle\ ,\ \rangle)$ into a 3X4 algebra in precisely two ways, by taking Φ in (2.1) to be Δ or $-\Delta$, where Δ is a normalized determinant function for E. If, for $n = 4$, $\{a,b,c,d\}$ is any ordered orthonormal basis positive in the sense that $\Phi(a,b,c,d) = +1$, then the nonzero values of X on this basis are obtained by permutation from the values

$$\begin{aligned} &X(b,c,d) = a, \quad X(a,c,d) = -b, \\ &X(a,b,d) = c, \quad X(a,b,c) = -d. \end{aligned} \tag{2.3}$$

Turning now to a 3Cn algebra $\mathfrak{C} = (E, \langle\ ,\ \rangle, \{\ \})$, we define, for each $a,b\in E$, linear operator $\gamma_{a,b}$ and $\sigma_{a,b}$ on E by

$$\gamma_{a,b}c = \{abc\} \quad\text{and}\quad \sigma_{a,b}c = \{cba\}. \tag{2.4}$$

From C1, $\gamma_{a,b}$ and $\sigma_{a,b}$ are bilinear functions of their arguments a and b, and from C2 they satisfy

$$\gamma_{a,a} = \langle a,a\rangle I = \sigma_{a,a} \tag{2.5}$$

and hence satisfy also the linearized form of this last equation:

$$\gamma_{a,b} + \gamma_{b,a} = 2\langle a,b\rangle I = \sigma_{a,b} + \sigma_{b,a}. \tag{2.6}$$

From C3 we have

$$\|\gamma_{a,b}c\| = \|a\|\ \|b\|\ \|c\| = \|\sigma_{a,b}c\|,$$

whence

$$\tilde{\gamma}_{a,b}\gamma_{a,b} = \langle a,a\rangle\langle b,b\rangle I = \tilde{\sigma}_{a,b}\sigma_{a,b} \tag{2.7}$$

and hence also, by linearization, we have

$$\tilde{\gamma}_{a,b}\gamma_{a,c} + \tilde{\gamma}_{a,c}\gamma_{a,b} = 2\langle a,a\rangle\langle b,c\rangle I = \tilde{\sigma}_{a,b}\sigma_{a,c} + \tilde{\sigma}_{a,c}\sigma_{a,b}. \tag{2.8}$$

Setting $c = a$ in (2.8) we obtain

$$\gamma_{a,b} + \tilde{\gamma}_{a,b} = 2\langle a,b\rangle I = \sigma_{a,b} + \tilde{\sigma}_{a,b} \tag{2.9}$$

and hence, from (2.6) and (2.9), we see that

$$\tilde{\gamma}_{a,b} = \gamma_{b,a} \quad\text{and}\quad \tilde{\sigma}_{a,b} = \sigma_{b,a}. \tag{2.10}$$

From (2.7) and (2.9) we see that if $A = \gamma_{a,b}$ or if $A = \sigma_{a,b}$, then A satisfies the quadratic equation

$$A^2 - 2\langle a,b\rangle A + \langle a,a\rangle\langle b,b\rangle I = 0. \tag{2.11}$$

If $\langle e,e\rangle = 1$ then the special cases $b = e$ of (2.6) and (2.10) yield

$$\tilde{\gamma}_{a,e} = \gamma_{\bar{a},e} \quad\text{and}\quad \tilde{\sigma}_{a,e} = \sigma_{\bar{a},e}, \tag{2.12}$$

where $\bar{a}$ is as in (1.3). Incidentally, in ternary notation, we have $\bar{a} = \{eae\}$, as is seen by letting (2.6) act upon b and setting $b = e$. Other special cases of the above results yield the following lemma. The only further point that arises is that, in part (a) of Lemma 2.2, $\gamma_{a,b}$ and $\sigma_{a,b}$ are *proper* isometries because they connect up continuously with I $(= \gamma_{a,a} = \sigma_{a,a})$.

Lemma 2.2: (a) If $\langle a,a\rangle = \langle b,b\rangle = 1$, then $\gamma_{a,b}$ and $\sigma_{a,b}$ belong to the group SO(E) of proper isometries of E.

(b) If $\langle a,b\rangle = 0$, then $\gamma_{a,b} = -\gamma_{b,a}\in\text{Sk}(E,E)$ and $\sigma_{a,b} = -\sigma_{b,a}\in\text{Sk}(E,E)$.

(c) If a,b,c are mutually orthogonal then $\gamma_{a,b}\gamma_{a,c} = -\gamma_{a,c}\gamma_{a,b}$ and $\sigma_{a,b}\sigma_{a,c} = -\sigma_{a,c}\sigma_{a,b}$.

As a lead-in to the next theorem observe that if $Z: E^3\to E$ is defined by

$$Z(a,b,c) = \langle a,b\rangle c + \langle b,c\rangle a - \langle a,c\rangle b, \tag{2.13}$$

then Z satisfies Axioms C1 and C2. Consequently if $\{\ \}$ satisfies C1 and C2 and if $X: E^3\to E$ is defined through the equation

$$\{abc\} = X(a,b,c) + Z(a,b,c), \tag{2.14}$$

then X satisfies $X(a,a,c) = 0 = X(c,a,a)$, whence X, clearly trilinear, is alternating.

Theorem 2.3: If (2.14) is used to define a one to one correspondence between maps $\{\ \}$: $E^3\to E$ and X: $E^3\to E$, then, for $n\geqslant 3$, $\mathfrak{C} = (E,\langle\ ,\ \rangle,\{\ \})$ is a 3Cn algebra if and only if $\mathfrak{A} = (E, \langle\ ,\ \rangle, X)$ is a 3Xn algebra.

Proof of Theorem 2.3: Clearly X1 holds if and only if C1 holds. Given X1 and X2, then (Lemma 2.1) X is alternating, whence C2 follows.

On the other hand given C1, C2, and C3 we can derive X2. By our lead-in to the theorem, we know that X is alternating and thus it suffices to show that $X(a,b,c)$ is orthogonal to c,

$$\begin{aligned} \langle c,X(a,b,c)\rangle &= \langle c, \{abc\}\rangle - \langle c,c\rangle\langle a,b\rangle \\ &= \langle c,(\gamma_{a,b} - \langle a,b\rangle I)c\rangle \\ &= 0, \end{aligned}$$

since, by (2.9), $\gamma_{a,b} - \langle a,b\rangle I$ is skew-adjoint.

To complete the proof we need to show, in the presence of properties X1, X2, C1, and C2, that C3 holds if and only if X3 holds. Applying Pythagoras's theorem to (2.14) yields

$$\|\{abc\}\|^2 = \|X(a,b,c)\|^2 + \|Z(a,b,c)\|^2 \tag{2.15}$$

and so the desired result holds by virtue of the (easily checked) identity

$$\langle a,a\rangle\langle b,b\rangle\langle c,c\rangle = \begin{vmatrix} \langle a,a\rangle & \langle a,b\rangle & \langle a,c\rangle \\ \langle b,a\rangle & \langle b,b\rangle & \langle b,c\rangle \\ \langle c,a\rangle & \langle c,b\rangle & \langle c,c\rangle \end{vmatrix} + \|Z(a,b,c)\|^2. \tag{2.16}$$

Remark: It follows from the correspondence (2.14) that the multiplication operators $\gamma_{a,b}$ and $\sigma_{a,b}$ for the 3Cn algebra $\mathfrak{C}$ are related to the multiplication operators $T_{a,b}$ for the 3Xn algebra $\mathfrak{A}$ by

$$\begin{aligned} \gamma_{a,b} &= \langle a,b\rangle I + T_{a,b} - J_{a,b}, \\ \sigma_{a,b} &= \langle a,b\rangle I - T_{a,b} - J_{a,b}. \end{aligned} \tag{2.17}$$

Here $J_{a,b}\in\text{Sk}(E,E)$ is defined by $J_{a,b}c = \langle a,c\rangle b - \langle b,c\rangle a$. (If a and b are linearly independent, then $J_{a,b}$ is nonzero and generates, by exponentiation, rotations in the plane $<a,b>$ spanned by a and b.)

III. THE 3C*n* HURWITZ THEOREM

In this section we shall determine, without appeal to the usual Hurwitz theorem for (binary) composition algebras, those values of n for which 3Cn algebras exist. Bearing in mind Theorem 2.3, and the remark after Lemma 2.1, we see that $n = 3$ is ruled out, but that 3C4 algebras certainly exist.

The values of $\{\,\}$ upon a positive orthonormal basis for a 3C4 algebra are given by (2.14) in conjuction with (2.3). Any facts needed concerning 3C4 algebras are accordingly easily unearthed. In particular we can easily check that the properties

$$\gamma_{b,a}\gamma_{c,a}\gamma_{d,a} = -I = -\sigma_{b,a}\sigma_{c,a}\sigma_{d,a} \tag{3.1}$$

hold for any positive orthonormal basis $\{a,b,c,d\}$.

Remark: In terms of the associated (see Sec. V) binary composition algebra, isomorphic to the algebra $\mathbb{H}$ of the quaternions, with a taken to be the identity element e in (5.7), we see that (3.1) is, in effect, granted that $\mathbb{H}$ is associative, Hamilton's relation $ijk = -e$.

For $n>4$, 3C4 subalgebras of a 3Cn algebra $\mathfrak{C} = (E,\langle\ \rangle, \{\,\})$ are readily constructed, as in part (a) of the next lemma.

Lemma 3.1: Let $\{b,c,d\}$ be any ordered orthonormal triad of vectors of E; set $a = \{bcd\}$ $[=X(b,c,d)]$ and $H = \langle a,b,c,d\rangle$ (= the subspace spanned by a,b,c,d). Then (a) H is a 3C4 subalgebra of E having $\{a,b,c,d\}$ as positive [i.e., $\Phi(a,b,c,d) = +1$] orthonormal basis; (b) for nonzero $h,k \in H$, $\gamma_{h,k}$ and $\sigma_{h,k}$ map H onto H and $H^\perp$ onto $H^\perp$; (c) for nonzero $h \in H$ and nonzero $p \in H^\perp$, $\gamma_{p,h}$ and $\sigma_{p,h}$ inject H into $H^\perp$; and (d) if $\Pi^H \in O(E)$ denotes the involution which is $+1$ upon H and -1 upon $H^\perp$,

$$\gamma_{b,a}\gamma_{c,a}\gamma_{d,a} = -\Pi^H = -\sigma_{b,a}\sigma_{c,a}\sigma_{d,a}. \tag{3.2}$$

Proof of Lemma 3.1: (a) By C3 and X2, $\{a,b,c,d\}$ is an orthonormal set and so $\dim H = 4$. Moreover H is invariant under $\gamma_{u,v}$ for $u,v \in \{a,b,c,d\}$. For example, by the definition of a, we have $\gamma_{b,c}d = a$, whence $\gamma_{b,c}a = -d$, since, by Lemma 2.2, $(\gamma_{b,c})^2 = -I$. Also, by C2, $\gamma_{b,c}c = b$, and hence $\gamma_{b,c}b = -c$. In this way we see that $\{uvw\}$ $(=\gamma_{u,v}w)$ lies in H for all $u,v,w \in \{a,b,c,d\}$, and so H is a 3C4 subalgebra.

(b) Since H is a subalgebra, $\gamma_{h,k}$ and $\sigma_{h,k}$ map H into H. But, by (2.7), $\gamma_{h,k}$ and $\sigma_{h,k}$ are similitudes, and so map H onto H and $H^\perp$ onto $H^\perp$.

(c) For all $k \in H$ we see that $\gamma_{p,h}k = \{phk\} = \sigma_{k,h}p$ lies in $H^\perp$ on account of part (b). Similarly $\sigma_{p,h}k = \gamma_{k,h}p$ lies in $H^\perp$. Since $\gamma_{p,h}$ and $\sigma_{p,h}$ are, by (2.7), invertible, they must inject H into $H^\perp$.

(d) From (3.1) we see that (3.2) certainly holds when acting upon H. Since $E = H \oplus H^\perp$ the first equality in (3.2) will therefore be proved if we can show that $\gamma_{b,a}\gamma_{c,a}\gamma_{d,a}p = p$ for all $p \in H^\perp$. Now

$$\begin{aligned}\gamma_{c,a}\gamma_{d,a}p &= -\gamma_{c,a}\gamma_{p,a}d \quad (\text{since } X \text{ is alternating})\\ &= \gamma_{p,a}\gamma_{c,a}d \quad [\text{by Lemma 2.2(c)}]\\ &= \gamma_{p,a}b \quad [\text{by (2.3)}]\\ &= -\gamma_{b,a}p.\end{aligned}$$

Hence the desired result, since $(\gamma_{b,a})^2 = -I$. The second equality in (3.2) is proved similarly.

Granted these preliminaries, we offer the following very brief proof of the "3Cn Hurwitz theorem."

Theorem 3.2: For a 3Cn algebra, n must equal 1, 2, 4, or 8.

Proof of Theorem 3.2: We need to show that $n>4$ implies $n=8$. Given $n>4$, choose a 3C4 subalgebra $H = \langle a,b,c,d\rangle$ as in Lemma 3.1, and let p be any nonzero element of $H^\perp$. Then $\gamma_{a,p}$ is an invertible operator, which, by Lemma 2.2 (c) and (3.2), anticommutes with Π^H. Hence Π^H has zero trace, whence $\dim H^\perp = \dim H = 4$, and so $n = 4+4 = 8$.

Remark: For a 3Cn algebra with $n=1$ or $n=2$ we have $X=0$ and hence $\{abc\} = Z(a,b,c)$. Certainly 3C1 and 3C2 algebras exist, since, if we take $E = \mathbb{R}$ and $\|a\| = |a|$, then a 3C1 algebra results from the definition $\{abc\} = abc$, while if we take $E = \mathbb{C}$ and $\|a\| = (a\bar{a})^{1/2}$ then a 3C2 algebra results from the definition $\{abc\} = a\bar{b}c$. Moreover, as discussed at the start of this section, we know that 3C4 algebras exist. Finally, 3C8 algebras exist: see Theorem 4.2 in Sec. IV.

IV. 3C8 ALGEBRAS

These possess the pleasing property that they readily split into an orthogonal direct sum of two 3C4 algebras.

Theorem 4.1: If H is any 3C4 subalgebra of a 3C8 algebra then $H^\perp$ is also a 3C4 subalgebra.

Proof of Theorem 4.1: Since $\tilde{\gamma}_{p,q} = \gamma_{q,p}$, we have the identity

$$\langle h, \{pqr\}\rangle = \langle \sigma_{h,p}q,r\rangle.$$

Now for $h \in H$ and $p,q,r \in H^\perp$ the rhs is zero, on account of Lemma 3.1(c). This is because, when $\dim H^\perp = \dim H$, $\sigma_{h,p}$ will, for nonzero h,p map H *onto* $H^\perp$ and, being proportional to an isometry, will therefore map $H^\perp$ onto H. The identity thus entails that $\{pqr\} \in H^\perp$ for all $p,q,r \in H^\perp$.

We now settle the question of existence of 3C8 algebras by means of an explicit construction. To this end let C^4 denote a four-dimensional complex Hilbert space, with inner product (a,b) linear in a and antilinear in b. Let Δ be a determinant function for C^4, normalized to be $+1$ upon some ordered orthonormal basis $\{e_0,e_1,e_2,e_3\}$. Let $b\times c\times d$ denote the "complex ternary vector cross product" on C^4, which is defined via

$$\Delta(a,b,c,d) = (a, b\times c\times d). \tag{4.1}$$

Observe that $a\times b\times c$ is orthogonal to each of a,b,c, that its length is given by

$$(a_1\times a_2\times a_3, a_1\times a_2\times a_3) = \det(a_i,a_j), \tag{4.2}$$

but that it is *antilinear* in each of its three arguments. [We could equally well define $a\times b\times c$ to be $*(a\wedge b\wedge c)$, using that star operator determined by the choice $e_0\wedge e_1\wedge e_2\wedge e_3$ of unit element in $\wedge^4 C^4$. Now recall that, for complex Hilbert space, the star operator is an antiunitary operator.]

Theorem 4.2: Let E be the realification $(C^4)^{\mathbb{R}}$ of C^4, let $\langle a,b\rangle$ be the real part of (a,b), and define $\{\,\}: E^3 \to E$ by

$$\{abc\} = a\times b\times c + (a,b)c + (b,c)a - (a,c)b. \tag{4.3}$$

Then $(E, \langle\ ,\ \rangle, \{\,\})$ is a 3C8 algebra.

Proof of Theorem 4.2: Clearly $\{\,\}$ satisfies Axioms C1 and C2. It also satisfies Axiom C3—because the identity (2.16) still holds even when $\langle\ ,\ \rangle$ is replaced throughout by $(\ ,\ )$.

Remark: Let $[a,b]$ denote the imaginary part of (a,b):

$$(a,b) = \langle a,b \rangle + i[a,b]. \tag{4.4}$$

Of course we have $[a,b] = \langle a,ib \rangle = -\langle ia,b \rangle = -[b,a]$. Thus E is equipped both with O(8) geometry, via $\langle\ ,\ \rangle$, and Sp(8;$\mathbb{R}$) geometry, via $[\ ,\ ]$. The ternary vector cross product X: $E^3 \rightarrow E$ associated with { } is, from (2.14) and (4.3), seen to be

$$X(a,b,c) = a \times b \times c + i([a,b]c + [b,c]a + [c,a]b) \tag{4.5}$$

and the associated scalar quadruple product is

$$\Phi(a,b,c,d) = \mathrm{Re}(\Delta(a,b,c,d)) + \Lambda(a,b,c,d), \tag{4.6}$$

where

$$\Lambda(a,b,c,d) = [a,b][c,d] + [b,c][a,d] + [c,a][b,d]. \tag{4.7}$$

Remark: According to (4.3) we have

$$\gamma_{iu,u} = J, \quad \text{for all unit vectors } u \in E, \tag{4.8}$$

where $J \in \mathrm{SO}(E) \cap \mathrm{Sk}(E,E)$ denotes the linear operator $a \mapsto ia$ on E. Consequently[9] the ternary vector cross product defined by (4.5) is[10] of type I. Indeed if we adopt

$$\{e_0,e_1,e_2,e_3,e_{0'},e_{1'},e_{2'},e_{3'}\},$$

where $e_{k'} = ie_k$, $k = 0,1,2,3$, as the orthonormal basis for $E = (C^4)^{\mathrm{R}}$, then the values of Φ on this basis are precisely those of the canonical form for $\Phi \in$type I which was described in Ref. 7.

Remark: The explicit expression (4.3) makes manifest the fact that the automorphisms of a 3C8 algebra certainly include at least $\mathrm{SU}(C^4) \subset \mathrm{O}(E) \cap \mathrm{Sp}(E)$, and so may prove useful in the context of the maximal subgroup chain:

$$\mathrm{SU}(4) \subset \mathrm{Spin}(7) \subset \mathrm{SO}(8).$$

Remark: It could be of some interest to determine the invariance group G of Φ for a 3X8 algebra, since G could conceivably be larger than Spin(7).[11] After all, the automorphism group of a 2X3 algebra (i.e., that of the usual $a \times b$ in three-dimensional Euclidean space) is SO(3) while the invariance group of the associated scalar triple product $\phi(a,b,c) = \langle a,b \times c\rangle$ is SL(3;$\mathbb{R}$); similarly the automorphism group of a 3X4 algebra is SO(4) while the invariance group of Φ (a determinant function) is SL(4;$\mathbb{R}$). However, the explicit expression (4.6) is probably *not* a good starting point for the determination of G, since we would still not know G even *if* we knew the separate invariance groups, say, $G_{(1)}$ and $G_{(2)}$, of Re(Δ) and Λ. For G is larger than $G_{(1)} \cap G_{(2)}$, if only because it contains certain Spin(7) transformations [those not in SU(C^4)] that mix Re(Δ) and Λ. The author confesses that he knows neither $G_{(1)}$ nor $G_{(2)}$. As far as $G_{(2)}$ is concerned it obviously contains $\mathrm{Sp}(E) \cong \mathrm{Sp}(8;\mathbb{R})$, but is it perhaps larger? Certainly in four-dimensional real space the invariance group of Λ is not Sp(4;$\mathbb{R}$) but SL(4;$\mathbb{R}$), because Λ is alternating and hence, in dimension 4, proportional to a determinant function.

Remark: Starting from the explicit expression (4.5) and making use of the property

$$(a_1 \times a_2 \times a_3, b_1 \times b_2 \times b_3) = \det(b_i,a_j), \tag{4.9}$$

it is possible[12] to derive the following identity for the (eight-dimensional, type I) ternary vector cross product X:

$$\langle X(a,b,c),X(u,v,w)\rangle = \langle a \wedge b \wedge c | u \wedge v \wedge w\rangle + F(a,b,c,u,v,w), \tag{4.10}$$

where, writing $f(a,b,c) + f(b,c,a) + f(c,a,b)$ as $\mathrm{Cyc}_{a,b,c}f(a,b,c)$,

$$F(a,b,c,u,v,w) = \mathrm{Cyc}_{a,b,c}\mathrm{Cyc}_{u,v,w}\langle a,u\rangle\Phi(b,c,v,w). \tag{4.11}$$

In its coordinate form the identity (4.10) has previously appeared[4,13] in the physics literature on $d = 11$ supergravity theories.

V. THE CONNECTION WITH 2C*n* ALGEBRAS

The pair $(E,\langle\ ,\ \rangle)$ becomes a *composition algebra*, in the usual (binary) sense, if E is equipped with a bilinear multiplication ab (not necessarily associative), which possesses an identity element e: $ea = a = ae$, and which satisfies the composition law

$$\|ab\| = \|a\|\,\|b\|. \tag{5.1}$$

If dim $E = n$ we will refer to such an algebra as a 2Cn algebra.

If we start out from a 2Cn algebra we can construct a 3Cn algebra by defining

$$\{abc\} = (a\bar{b})c, \tag{5.2}$$

with $\bar{b}$ as in (1.3). In order to check that { } satisfies Axiom C2 we need to know the following two facts[14] concerning 2Cn algebras. First, they satisfy the left and right *alternative laws*

$$a(ab) = a^2b \quad \text{and} \quad (ba)a = ba^2. \tag{5.3}$$

Second, the conjugation $a \rightarrow \bar{a}$ satisfies

$$\bar{a}a = \langle a,a\rangle e = a\bar{a}. \tag{5.4}$$

Granted these facts, { } as defined in (5.2) satisfies the axioms for a 3Cn algebra. So Theorem 3.2 entails that 2Cn algebras can exist only in dimensions $n = 1, 2, 4$, or 8—that is, Theorem 3.2 entails the well-known Hurwitz theorem for binary composition algebras.

In fact, we can construct a 3Cn algebra out of a 2Cn algebra in four ways, by defining $\{abc\}$ to be

$$\begin{matrix} (1)\ (a\bar{b})c, & (1')\ \ c(\bar{b}a), \\ (2)\ (c\bar{b})a, & (2')\ \ a(\bar{b}c). \end{matrix} \tag{5.5}$$

In the case $n = 8$ all four kinds of $\{abc\}$ are distinct, in the sense that they lie on different SO(8) orbits. Indeed we can see[10] that (1) and (1′), in some order, provide examples of types I^R and I^L, while (2) and (2′), in the same order, provide examples of types II^R and II^L.

Incidentally in order to see that (1) is mapped onto (1′) under the natural action $\{abc\} \mapsto K\{\bar{a}\,\bar{b}\,\bar{c}\}$ of the conjugation $K \in \mathrm{O}_-(E)$ we need a third property of 2Cn algebras, namely,

$$\overline{ab} = \bar{b}\bar{a}. \tag{5.6}$$

Going in the other direction (the viewpoint of the present work) if we start out from a 3Cn algebra we can construct a 2Cn algebra, *moreover with any chosen unit vector* $e \in E$ *as identity element*, simply by defining

$$ac = \{aec\}. \tag{5.7}$$

In ternary terms the alternative laws (5.3) state that

$$(\gamma_{a,e})^2 = \gamma_{\{aea\},e} \quad \text{and} \quad (\sigma_{a,e})^2 = \sigma_{\{aea\},e}, \tag{5.8}$$

which is seen to be the special case $b = e$ of (2.11) [since $\{aea\} = Z(a,e,a)$]. Similarly the special case $b = e$ of (2.7) entails (5.4). Moreover the conjugation K: $a \mapsto \bar{a}$ commutes with $T_{a,e}$ and anticommutes with $J_{a,e}$, whence we derive from (2.17) the result

$$K\gamma_{a,e} = \tilde{\sigma}_{a,e}K. \tag{5.9}$$

On acting upon b with the two sides of (5.9) we obtain (5.6). Alternatively (5.6) can be obtained from (2.8) upon setting $a = e$, $b = \bar{a}$, $c = b$.

Remark: The relation (5.9) occurred also in some recent work[15] of the present author where certain SO(7)–Clifford algebra aspects of ternary vector cross products (related to the involutions Π^H) were investigated.

[1] B. Eckmann, Comment. Math. Helv. **15**, 318 (1942-3).

[2] R. B. Brown and A. Gray, Comment. Math. Helv. **42**, 222 (1967).

[3] See also R. Shaw, Int. J. Math. Educ. Sci. Technol. **18**, 803 (1987).

[4] R. Dündarer, F. Gürsey, and C-H. Tze, J. Math. Phys. **25**, 1496 (1984).

[5] R. Harvey and H. B. Lawson, Acta. Math. **148**, 47 (1982).

[6] A. Gray, Trans. Am. Math. Soc. **141**, 465 (1965).

[7] R. Shaw, J. Phys. A: Math. Gen. **20**, L689 (1987).

[8] K. McCrimmon, Trans. Am. Math. Soc. **275**, 107 (1983).

[9] See Ref. 7, Eq. (12) and subsequent parenthetic remark.

[10] See Ref. 7, Theorem E.

[11] In the case when E has neutral signature (4,4) one certainly finds that G is at least slightly larger than the automorphism group, as a result of the presence of "counter-automorphisms" which send both X and $\langle\ ,\ \rangle$ to their negatives, but which, by (2.1), still preserve Φ. See R. Shaw, "Ternary composition, neutral signature and Spin(3,4)," in *Proceedings of XVIIth International Colloquium on Group Theoretical Methods in Physics, Montreal, 1988* (World Scientific, Singapore, to be published).

[12] R. Shaw, J. Phys. A: Math. Gen. **21**, 593 (1988).

[13] B. DeWit and H. Nicolai, Nucl. Phys. B **231**, 506 (1984).

[14] N. Jacobson, *Basic Algebra I* (Freeman, San Francisco, 1974).

[15] R. Shaw, J. Phys. A: Math. Gen. **21**, 7 (1988).

Notes on a cross product of vectors

J. F. Plebański[a]
Departamento de Física, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México 14, D.F., Mexico
M. Przanowski
Institute of Physics, Technical University of Łódź, Wólczańska 219, 93-005 Łódź, Poland

(Received 25 September 1987; accepted for publication 20 July 1988)

The relation between quaternionlike algebras and cross products of vectors is demonstrated. A classification of all cross products of vectors is given.

I. INTRODUCTION

During the preparation of our paper on quaternionlike algebras,[1] Professor José Adem suggested that those considerations seemed to be closely related to the problem of a definition of a cross product of vectors in a vector space of an arbitrary (finite) dimension. The purpose of the present paper is to find this relation.

There are several possibilities to generalize the usual cross product of vectors in three-dimensional real Euclidean vector space R^3 to arbitrary finite-dimensional vector spaces. Thus according to Eckmann,[2] Whitehead,[3] and Zvengrowski[4,5] (see also Refs. 6 and 7), one can define a vector cross product in n-dimensional real Euclidean vector space R^n to be a mapping

$$P_r : R^{nr} = \underbrace{R^n \times \cdots \times R^n}_{r\ \text{times}} \to R^n, \quad 1 \leqslant r \leqslant n,$$

satisfying the following axioms.

(a_1) P_r is a continuous mapping of R^{nr} $(1 \leqslant r \leqslant n)$ into R^n.

(b_1) $P_r(v_1,...,v_r)\cdot v_i = 0$, for every set of vectors $(v_1,...,v_r) \in R^{nr}$ and $i = 1,...,r$.

(c_1) $P_r(v_1,...,v_r)\cdot P_r(v_1,...,v_r) = \det(v_i \cdot v_j), i,j = 1,...,r$, for every set of vectors $(v_1,...,v_r) \in R^{nr}$.
Here a dot stands for the usual Euclidean scalar product of vectors in R^n.

Another, more "algebraic" definition of the cross product of vectors can be extracted from the works given by Brown and Gray,[8] Gray,[9] and Dündarer, Gürsey, and Tze[10] (see also Ref. 6). According to these works we have the following definition: Let V be an n-dimensional vector space over a field F of the characteristic $\neq 2$ and let $\langle \cdot,\cdot \rangle$ be a nondegenerate, bilinear, symmetric form on V. A vector cross product in V is a mapping

$$P_r : V^r = \underbrace{V \times \cdots \times V}_{r\ \text{times}} \to V, \quad 1 \leqslant r \leqslant n,$$

satisfying the following axioms.

(a_2) P_r is an r-linear mapping of V^r $(1 \leqslant r \leqslant n)$ into V.

(b_2) $\langle P_r(v_1,...,v_r), v_i \rangle = 0$ for every set of vectors $(v_1,...,v_r) \in V^r$ and $i = 1,...,r$.

(c_2) $\langle P_r(v_1,...,v_r), P_r(v_1,...,v_r) \rangle = \det(\langle v_i,v_j \rangle), i,j = 1,...,r$, for every set of vectors $(v_1,...,v_r) \in V^r$.

Then the following important theorem holds (see Refs. 2–10).

Theorem 1: The vector cross product satisfying the axioms (a_1)–(c_1) or (a_2)–(c_2) exists if and only if (1) $r = 1$ and n even; (2) $r = n - 1$ and n arbitrary; (3) $r = 2$ and $n = 3,7$; (4) $r = 3$ and $n = 4,8$. ■

Explicit formulas for the vector cross products are also known.[4,5,8,9]

The cases of $r = 2$ and $n = 3,8$ or $r = 1$ and $n = 8$ are, perhaps, the most interesting ones. Thus the vector cross product in R^3 with $r = 2$ is the usual cross product. The vector cross product in R^8 with $r = 3$ can be defined in an elegant way in terms of octonions.[4,5,8–10] Moreover, Dündarer *et al.*[10] were able to give a compact, unified, covariant, and explicit formulation of various "generalized vector cross products" in R^8 (compare also with Kleinfeld[11]). Then the cross product of vectors in R^7 with $r = 2$ can be defined in terms of pure octonions.[5,8,12–14]

Now it appears that in the case of $r = 2$ the axioms (a_2)–(c_2) for P_2 in R^n are equivalent to the following axioms.[12–14]

(a_3) P_2 is a bilinear and skew-symmetric mapping of $R^n \times R^n$ into R^n.

(b_3) $P_2(v_1,v_2)\cdot v_1 = P_2(v_1,v_2)\cdot v_2 = 0,$

for any $(v_1,v_2) \in R^n \times R^n$.

(c_3) $= P_2(v,P_2(v,w)) = (v\cdot w)v - (v \cdot v)w,$

for any $v,w \in R^n \times R^n$.

In the present paper we consider the vector cross product satisfying the conditions that generalize in a natural manner the axioms (a_3)–(c_3).

II. CROSS PRODUCT OF VECTORS

Let V be an n-dimensional vector space over the number field F ($= R$ or C). Then a mapping $\cdot \times \cdot : V \times V \to V$ is said to be a cross product in V if there exists a bilinear form on $V, (\cdot|\cdot): V \times V \to F$ such that the following conditions hold.

(a) The mapping $\cdot \times \cdot : V \times V \to V$ is bilinear and skew symmetric.

(b) $(v \times w | v) = 0$, for any $v,w \in V$.

(c) $u \times (v \times w) = (u|w)v - (u|v)w$, for any $u,v,w \in V$.

A bilinear form on $V, (\cdot|\cdot): V \times V \to F$, for which (b) and (c) hold true, is called a bilinear form associated with

---

[a] On leave of absence from University of Warsaw, Warsaw, Poland.

the cross product $\cdot\times\cdot$. If $\cdot\times\cdot$ is the usual cross product in R^3 then the properties (a), (b), and (c) are satisfied with $(\cdot|\cdot)$ as the usual scalar product in R^3.

First we prove a simple Proposition.

Proposition 1: The conditions (a) and (b) are equivalent to (a) and (b′), where (b′) reads

(b′) $(v\times w|u) = (u\times v|w)$, for any $u,v,w\in V$.

Proof: Utilizing (a) and (b) to $((u+w)\times v|u+w)$ one finds (b′). Thus (a), (b) $\Rightarrow$ (a), (b′). The implication (a),(b′) $\Rightarrow$ (a), (b) is evident. The proof is completed. ■

Employing Proposition 1 we get easily that (a), (b), and (c) yield the following analog of the conditions (c_1) or (c_2):

$$(v\times w|v\times w) = (v|v)(w|w) - (v|w)^2, \quad \text{for } v,w\in V. \qquad (1)$$

Now we prove the following proposition.

Proposition 2: Let V be an n-dimensional vector space over F and let f be a bilinear skew-symmetric mapping of $V\times V$ into V. If there exists a bilinear form g: $V\times V\rightarrow F$ such that

$$f(u,f(v,w)) = g(u,w)v - g(u,v)w, \quad \text{for any } u,v,w\in V, \qquad (2)$$

then g is a symmetric form and, for $n>1$, g is uniquely defined by the mapping f.

Proof: If $n=1$, then every bilinear form on v is symmetric. Consider $n>1$. Let $e_1,...,e_n$ be a basis of V. We set

$$g(e_i,e_j) = g_{ij}\in F \qquad (3)$$

and

$$f(e_i,e_j) = \sum_k C^k{}_{ij}\, e_k, \quad F\ni C^k{}_{ij} = -C^k{}_{ji}, \qquad (4)$$

where small Latin indices are assumed to run through $1,...,n$. By applying Eq. (2) to the basic vectors e_i with the use of (3) and (4) one gets

$$\sum_m C^l{}_{im}\, C^m{}_{jk} = g_{ik}\,\delta^l_j - g_{ij}\,\delta^l_k, \qquad (5)$$

where δ^l_j is the Kronecker delta. Contracting both sides of the formula (5) with respect to the indices l, k, we obtain

$$\sum_{l,m} C^l{}_{im}\, C^m{}_{jl} = -(n-1)g_{ij}. \qquad (6)$$

Since $n>1$,

$$g_{ij} = g_{ji} = -\frac{1}{n-1}\sum_{l,m} C^l{}_{im}\, C^m{}_{jl}. \qquad (7)$$

Thus the proof is completed. ■

As a consequence of Proposition 2 one has immediately the following Corollary.

Corollary 1: If a mapping $\cdot\times\cdot$: $V\times V\rightarrow V$ is a cross product in V, then a bilinear form on V, $(\cdot|\cdot)$: $V\times V\rightarrow F$, fulfilling the condition (c), is symmetric and if $\dim V>1$, then $(\cdot|\cdot)$ is uniquely defined by $\cdot\times\cdot$. ■

If $\dim V=1$, then (evidently) any bilinear form on V is symmetric and it is associated with the cross product in V which is now uniquely defined, i.e., $v\times w=0$ for any $v,w\in V$. If $\dim V>1$, then a bilinear form associated with a cross product in V is symmetric and it is uniquely defined by the given cross product.

Let us recall the definition of a quaternionlike algebra[1] (ql algebra). An $(n+1)$-dimensional algebra $(n\geqslant 1)$ Q with unity e_0 over F is said to be a quaternionlike algebra (ql algebra) if it is associated and if there exists a decomposition

$$Q = Fe_0\oplus V, \qquad (8)$$

where V is an n-dimensional vector subspace of Q such that for every vector $v\in V$, $vv\in Fe_0$.

As it has been shown in Ref. 1, if some algebra, associative or not, admits a decomposition of the form (8), then this decomposition is unique. We need the notion of equivalent cross products. Let v_1, v_2 be n-dimensional vector spaces over F and let f_1: $V_1\times V_1\rightarrow V_1$, f_2: $V_2\times V_2\rightarrow V_2$ be cross products in V_1, V_2, respectively. Then the cross product f_1 is said to be equivalent to the cross product f_2 if there exists an isomorphism i: $V_1\rightarrow V_2$ such that

$$f_2(iv,iw) = i(f_1(v,w)), \quad \text{for any } v,w\in V_1. \qquad (9)$$

Now we can prove the main theorem of this paper.

Theorem 2: Given an $(n+1)$-dimensional ql algebra Q over F that decomposes according to (8), we define two mappings g: $V\times V\rightarrow F$ and f: $V\times V\rightarrow V$ as follows:

$$vw = -g(v,w)e_0 + f(v,w), \quad \text{for any } v,w\in V. \qquad (10)$$

Then f is a cross product in V and g is a bilinear form associated with f.

Conversely, given an n-dimensional $(n\geqslant 1)$ vector space $\widetilde V$ over F, a cross product in $\widetilde V$, $\cdot\times\cdot$: $\widetilde V\times\widetilde V\rightarrow\widetilde V$, and a bilinear form $(\cdot|\cdot)$: $\widetilde V\times\widetilde V\rightarrow F$ associated with $\cdot\times\cdot$, there exists a unique (with the precision to an isomorphism) $(n+1)$-dimensional ql algebra Q over F that decomposes according to (8) admitting an isomorphism i: $\widetilde V\rightarrow V$ such that

$$f(i\tilde v, i\tilde w) = i(\tilde v\times\tilde w), \quad \text{for any } \tilde v,\tilde w\in\widetilde V, \qquad (11)$$

i.e., $\cdot\times\cdot$ is equivalent to f, and

$$g(i\tilde v, i\tilde w) = (\tilde v|\tilde w), \quad \text{for any } \tilde v,\tilde w\in\widetilde V, \qquad (12)$$

where the mappings f: $V\times V\rightarrow V$ and g: $V\times V\rightarrow F$ are defined by (10); moreover if $\dim\widetilde V = n>1$, then Q is uniquely (i.e., with the precision to an isomorphism) defined by the pair $(\widetilde V, \cdot\times\cdot)$.

Proof: Let Q be an $(n+1)$-dimensional ql algebra over F that decomposes according to (8), and let g: $V\times V\rightarrow F$ and f: $V\times V\rightarrow V$ be the mappings defined by (10). From the fact that Q is an algebra it follows that g is a bilinear form on V, and f is a bilinear mapping of $V\times V$ into V. Since Q is a ql algebra, $vv\in Fe_0$, for every $v\in V$. Hence, by (10), $f(v,v)=0$ for every $v\in V$. Consequently, as f: $V\times V\rightarrow V$ is a bilinear mapping, it is skew symmetric.

Since Q is an associative algebra,

$$(u,v)w = u(vw), \quad \text{for any } u,v,w\in V. \qquad (13)$$

From (13) and (10) it follows that

$$g(f(u,v),w) = g(u,f(v,w)), \quad \text{for any } u,v,w\in V, \qquad (14)$$

$$f(f(u,v),w) - f(u,f(v,w)) = g(u,v)w - g(v,w)u,$$
$$\text{for any } u,v,w\in V. \qquad (15)$$

Let $e_1,...e_n$ be a basis of V and let $g_{ij}\in F$, $F\ni C^k{}_{ij} = -C^k{}_{ji}$ be the numbers defined by (3) and (4); small Latin indices are assumed to run through $1,...,n$. The formula (15), when applied to the basic vectors e_i, yields

$$-\sum_m (C^l{}_{im}\, C^m{}_{jk} + C^l{}_{km}\, C^m{}_{ij}) = g_{ij}\,\delta^l{}_k - g_{jk}\,\delta^l{}_i\,. \quad (16)$$

Contracting both sides of (16) with respect to the indices l, k, and then with respect to l, i, one finds

$$-\sum_{l,m} (C^l{}_{im}\, C^m{}_{jl} + C^l{}_{lm}\, C^m{}_{ij}) = n g_{ij} - g_{ji}\,, \quad (17)$$

$$-\sum_{l,m} (C^l{}_{lm}\, C^m{}_{jk} + C^l{}_{km}\, C^m{}_{lj}) = g_{kj} - n g_{jk}\,. \quad (18)$$

Writing (18) for $k = i$ and adding the results to (17), we get $(n+1)(g_{ij} - g_{ji}) = 0$. Hence $g_{ij} = g_{ji}$. Thus we arrive at the conclusion that the bilinear form g is symmetric. Executing the cyclic sum with respect to u,v,w for both sides of (15) one obtains the "Jacobi identity,"

$$f(f(u,v),w) + f(f(w,u),v) + f(f(v,w),u) = 0,$$
$$\text{for any } u,v,w\in V. \quad (19)$$

Finally, from (15) and (19) one finds

$$f(v,f(w,u)) = g(u,v)w - g(v,w)u, \quad \text{for any } u,v,w\in V. \quad (20)$$

The condition (14) for $u = v$ yields

$$g(v,f(u,w)) = 0, \quad \text{for any } v,w,\in V. \quad (21)$$

Comparing (20) and (21) with (c) and (b), and also employing the symmetry condition of g, we conclude that the mapping f: $V\times V \to V$ is a cross product in V and the mapping g: $V\times V\to F$ appears to be a bilinear form associated with f. Thus the first part of our theorem has been proved.

Now let $\widetilde{V}$ be an n-dimensional $(n\geqslant 1)$ vector space over F and let the mapping $\cdot\times\cdot$: $\widetilde{V}\times\widetilde{V}\to\widetilde{V}$ be a cross product in $\widetilde{V}$; moreover, let the mapping $(\cdot|\cdot)$:$\widetilde{V}\times\widetilde{V}\to F$ be a bilinear form associated with $\cdot\times\cdot$. Let us define $Q{:}= F\oplus\widetilde{V}$. We have $Q = Fe_0\oplus V$, where e_0: $(1,0)\in F\oplus\widetilde{V}$ and V is the vector subspace of $F\oplus\widetilde{V}$ consisting of the vectors of the form $(0,\tilde{v})$, where $\tilde{v}\in\widetilde{V}$. Define a multiplication on Q as follows:

$$Q\times Q\ni((a,\tilde{v}),(b,\tilde{w}))\mapsto(a,\tilde{v})(b,\tilde{w})\in Q,$$

where

$$(a,\tilde{v})(b,\tilde{w}){:}= (ab - (\tilde{v}|\tilde{w}), a\tilde{w} + b\tilde{v} + \tilde{v}\times\tilde{w}) \quad (22)$$

[compare (22) with Refs. 2 and 6]. It is a straightforward matter to show that Q with the above defined multiplication (22) constitutes an $(n+1)$-dimensional ql algebra over F. Let i: $\widetilde{V}\to V$ be the natural isomorphism of $\widetilde{V}$ onto V defined as follows:

$$i\colon \widetilde{V}\ni\tilde{v}\mapsto(0,\tilde{v})\in V. \quad (23)$$

Then, from (22) and (23) we obtain

$$(0,\tilde{v})(0,\tilde{w})$$
$$= (-(\tilde{v}|\tilde{w}),\tilde{v}\times\tilde{w}) = -(\tilde{v}|\tilde{w})(1,0) + (0,\tilde{v}\times\tilde{w})$$
$$= -(\tilde{v}|\tilde{w})e_0 + i(\tilde{v}\times\tilde{w}), \quad \text{for any } \tilde{v},\tilde{w}\in\widetilde{V}. \quad (24)$$

Comparing with (10) one gets (11) and (12).

Now let $Q_1 = Fe_0^{(1)}\oplus V_1$ be a ql algebra such that there exists an isomorphism i_1: $\widetilde{V}\to V_1$ for which the analogs of (11) and (12) hold. Define an isomorphism i_0: $Fe_0^{(1)}\to Fe_0$, $i_0(ae_0^{(1)}) = ae_0$, for every $a\in F$. Then it easy to check that the mapping $i_0\oplus i\circ i_1^{-1}$ is an isomorphism of the ql algebra Q_1 onto the ql algebra Q.

Finally, utilizing Corollary 1 we complete the proof. ■

The main consequence of Theorem 2 is that there exists a 1:1 correspondence between the class of all nonequivalent vector cross products in vector spaces of dimension $n > 1$ and the class of all nonisomorphic ql algebras of dimension $n+1 > 2$. Therefore, employing the results of our previous paper[1] concerning the classification of ql algebras, we arrive at the following conclusion.

A cross product in an n-dimensional real vector space V, $\cdot\times\cdot$: $V\times V\to V$, belongs to one of the following types: (I) a trival cross product, i.e.,

$$v\times w = 0, \quad \text{for any } v,w\in V; \quad (25)$$

(II) a nilpotent cross product of the nilpotency class 2, i.e., a nontrivial cross product such that

$$u\times(v\times w) = 0, \quad \text{for any } u,v,w\in V; \quad (26)$$

(III) there exists a basis $e_1,\ldots,e_n$ of V such that

$$e_\alpha\times e_\beta = 0, \quad e_n\times e_\alpha = \epsilon_\alpha e_\alpha,$$
$$\epsilon_\alpha = \pm 1; \quad \alpha,\beta = 1,\ldots,n-1; \quad (27)$$

(III′) n is odd and there exists a basis $e_1,\ldots,e_n$ of V such that

$$e_\alpha\times e_\beta = 0, \quad e_n\times e_{(n-1)/2+A} = e_A,$$
$$e_n\times e_A = -e_{(n-1)/2+A}, \quad (28)$$
$$\alpha,\beta = 1,\ldots,n-1, \quad A = 1,\ldots(n-1)/2,$$

(IV) $n = 3$, the usual cross product,

$$e_1\times e_2 = e_3, \quad e_3\times e_1 = e_2, \quad e_2\times e_3 = e_1, \quad (29)$$

for some basis e_1, e_2, e_3; (IV′) $n = 3$ and there exists a basis e_1, e_2, e_3 such that

$$e_1\times e_2 = -e_3, \quad e_3\times e_1 = e_2, \quad e_2\times e_3 = e_1. \quad (30)$$

Employing Proposition 2 one can easily find the bilinear forms associated with the above listed cross products.

(I) For $n = 1$ there exists a nonzero vector e_1 such that

$$(e_1|e_1) = \pm 1,$$

or one has

$$(v|w) = 0, \quad \text{for any } v,w\in V. \quad (31)$$

For $n > 1$, Eq. (31) holds.

(II) The formula (31) holds true.

(III) $(e_\alpha|e_\beta) = (e_\alpha|e_n) = 0$, $(e_n|e_n) = -1$, $\alpha,\beta = 1,\ldots,n-1$.

(III′) $(e_\alpha|e_\beta) = (e_\alpha|e_n) = 0$, $(e_n|e_n) = 1$, $\alpha,\beta = 1,\ldots,n-1$.

(IV) $(e_i|e_j) = \delta_{ij}$, $i,j = 1,2,3$.

(IV′) $(e_i|e_j) = \epsilon_i\delta_{ij}$, $i, = 1,2,3$, $\epsilon_1 = \epsilon_2 = -1$, $\epsilon_3 = 1$.

In the case of complex V we have the types (I), (II), (III), and (IV) [for $n = 1$ there exists a nonzero vector e_1 such that $(e_1|e_1) = 1$, or one has (31)].

Remark: Given a cross product $\cdot\times\cdot$: $\widetilde{V}\times\widetilde{V}\to\widetilde{V}$, the algebra $(\widetilde{V},\cdot\times\cdot)$ appears to be a Lie algebra isomorphic to the Lie algebra $(V,[\cdot,\cdot])$ induced by $Q = F\oplus\widetilde{V}$ with multiplication defined by (22). An isomorphism is given as follows: $\widetilde{V}\ni\tilde{v}\mapsto\frac{1}{2}(0,\tilde{v})\in V$. (For the notation see Theorem 2. Regarding induced Lie algebras, see Ref. 1.)

Concluding the present paper we would like to deal with two problems.

The first one concerns the possibility of generalization of the cross product of vectors on the vector spaces over an arbitrary field F. Employing the results of our recent work[15] one can easily realize that such a generalization does exist for an arbitrary field F of the characteristic "not 2" and, in fact, it is almost "automatic." In particular, the main theorem of the present paper, i.e., Theorem 2, holds true in that general case. However, the proof of this theorem given here must be slightly changed to be valid generally. Namely, the formula (7) should be replaced by the formula

$$g_{ij} = g_{ji} = -\sum_m C^k{}_{im}\, C^m{}_{jk}, \quad k \neq j, \tag{7'}$$

which follows from (5) or (20). [Thus the formula (20) yields the symmetry of the bilinear form g and we do not need the formulas (16)–(18).]

The canonical forms of all possible vector cross products can be written down in the case when a field F is of the characteristic "not 2" and, if $n \neq 3$, also "not a divisor of $n-3$"(for details, see Ref. 15).

The second problem we would like to deal with concerns a relation of our cross product of vectors to Clifford algebras. This problem in all its details has been considered in Ref. 15. Now we cite the main results obtained. Given an $(n+1)$-dimensional $(n \geqslant 1)$ quaternionlike algebra Q over a field F of the characteristic "not 2" which decomposes according to (8), one defines the quadratic space (V, q), where q: $V \to F$ is the quadratic form on v defined by

$$vv = q(v)e_0, \quad \text{for any } v \in V. \tag{32}$$

Then we construct the Clifford algebra $C(V,q)$ for (V,q) (see also Ref. 16). From the universal property of Clifford algebras it follows that there exists a unique homomorphism φ: $C(V,q) \to Q$ such that $\varphi(v) = v$ for any $v \in V$. In particular, for any $v,w \in V$,

$$\varphi(v \# w) = vw = -g(v,w)e_0 + f(v,w) \in Q, \tag{33}$$

where the symbol # stands for the multiplication in Clifford algebra $C(V,q)$, and the mappings g: $V \times V \to F$ and f: $V \times V \to V$ are the same as in the formula (10). The mapping is a cross product of vectors in V and g is a bilinear form associated with f.

Another interesting question arises: is the Clifford algebra $C(V,q)$ a ql algebra? The answer to this question follows from the general proposition (see Ref. 15).

Proposition 3: Let V' be a vector space of dimension n' over a field F of the characteristic "not 2" and let q': $V' \to F$ be a quadratic form on V'. Then, the Clifford algebra $C(V',q')$ for the quadratic space (V',q') is a ql algebra if and only if $n' \leqslant 2$. Moreover, for $n' = 1$ every ql algebra over F is a Clifford algebra over F; for $n' = 2$ there exists exactly two nonisomorphic ql algebras of dimension four over F which appear to be nonisomorphic to a four-dimensional Clifford algebra over F. The structures of these ql algebras are defined as follows:

(1) $Q = Fe_0 \oplus V$, dim $V = 3$
$e_i e_j = 0$, $i,j = 1,2,3$;
(2) $Q = Fe_0 \oplus V$, dim $V = 3$,
$e_1e_1 = e_2e_2 = 0$, $e_3e_3 = e_0$,
$e_1e_2 = 0$, $e_3e_1 = e_1$, $e_2e_3 = -e_2$;

where the set of vectors (e_1, e_2, e_3) constitutes a basis for V. ■

Consequently we conclude that Clifford algebras over a field F of characteristic $\neq 2$ define vector cross products according to the "natural" scheme given in Theorem 2 if and only if their dimension $= 2$ or 4.

ACKNOWLEDGMENTS

One of us (M. P.) is grateful to the staff of the Department of Physics at the Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, and especially to Dr. Arnulfo Zepeda and Dr. Alberto García for warm hospitality during his stay at this Department.

His work was supported in part by the CONACYT, México D.F., and by the Centro de Investigación y de Estudios Avanzados del Instituto Politénico Nacional México D.F., Mexico.

[1] J. F. Plebański and M. Przanowski, J. Math. Phys. **29**, 529 (1988).
[2] B. Eckmann, Comment, Math. Helv. **15**, 318 (1942/43).
[3] G. W. Whitehead, Comment. Math. Helv. **37**, 239 (1963).
[4] P. Zvengrowski, thesis, Univ. of Chicago, 1965, Chap. 4.
[5] P. Zvengrowski, Comment. Math. Helv. **40**, 149 (1966).
[6] W. S. Massey, Am. Math. Monthly **90**, 697 (1983).
[7] P. Hilton, *General Cohomology Theory and K-Theory* (Cambridge U.P., Cambridge, 1971).
[8] R. Brown and A. Gray, Comment. Math. Helv. **42**, 222 (1967).
[9] A. Gray, Trans. Amer. Math. Soc. **141**, 465 (1969).
[10] R. Dündarer, F. Gürsey, and C. H. Tze, J. Math. Phys. **25**, 1496 (1984).
[11] E. Kleinfeld, "A characterization of the Cayley numbers," in *Studies in Modern Algebra*, edited by A. A. Albert (Prentice–Hall, Englewood Cliffs, NJ, 1963), Vol. 2.
[12] M. Zorn, Proc. Natl. Acad. Sci. USA **21**, 355 (1935).
[13] R. Hermann, *Spinors, Clifford, and Cayley Algebras* (Math Sci., Brookline, MA, 1974).
[14] E. Calabi, Trans. Am. Math. Soc. **87**, 407 (1958).
[15] J. F. Plebański and M. Przanowski, to appear in Bol. Soc. Mat. Mexicana.
[16] T. Y. Lam, *The Algebraic Theory of Quadratic Forms* (Benjamin, Reading, MA, 1973).

The structure of the Clifford algebra

G. P. Wilmot
Department of Mathematical Physics, University of Adelaide, GPO Box 498, Adelaide, 5001, Australia

(Received 29 December 1987; accepted for publication 20 July 1988)

The properties of the Clifford algebra are concisely summarized in four identities that extend the usual Lie algebra to the entire algebra. This formulation may be used to analyze the general structure of the Clifford algebra, but it is found that the symmetry properties are not consistent with the geometrical content. The extension of the graded Lie algebra to the entire algebra generates a geometrical algebra which simplifies the analysis of the structure of the Clifford algebra. This produces a direct proof of the fundamental theorem that relates the Clifford and Grassmann algebras via the Pfaffian.

I. INTRODUCTION

The abundance of terms contained in an arbitrary Clifford element conceals a wealth of geometrical information describing all relative knowledge of a set of vectors. It is easy to study its structure in simple cases, but for more general elements it is necessary to resort to implicit statements of the geometrical content. In this paper a nonrecursive technique to analyze a Clifford element is developed, and this will naturally lead us to the fundamental theorem of Clifford algebra.

The usual approach to analyzing the structure is via the Lie algebra of the Clifford algebra, which corresponds to the commutator algebra, since we are dealing with an associative algebra. It is well known that this is the Lie algebra of the corresponding orthogonal group. The remaining structure of the Clifford algebra is given by the anticommutator bracket, which, together with the commutator bracket, presents the entire structure in four identities. The restriction of these identities to simple Clifford elements presents, in Sec. II, a concise summary of the properties of the orthogonal group.

The structure of arbitrary Clifford elements can be analyzed using what has been called the multivector formalism. Basically, this approach builds up any member of the algebra from the sum of products of Clifford vectors, as is done using the Feynman slash notation for the Dirac algebra, a well known Clifford algebra. The structure of a multivector can be analyzed by repeated application of the commutator and anticommutator brackets. Such analysis is not performed on arbitrary Clifford elements simply because the bracket products do not express the geometrical content of the multivector in a consistent fashion. To progress, we must abandon the ordinary brackets and look at the generalized brackets. The generalized or graded commutator for an associative algebra defines the graded Lie algebra or, in the case of the Clifford algebra, the semigraded Lie algebra. Introducing the graded anticommutator again gives a complete decomposition of the Clifford product and we will see in Sec. III that these products are consistent in the geometrical operations that the products represent.

This geometrical approach does not alter the inductive analysis referred to above, it only clarifies our perception of the operations. The study of multivectors without using induction techniques becomes feasible, up to a point, because of this geometrical insight. It also becomes obvious that the Pfaffian is the required object to represent the structure of the multivector algebra. This theorem is proved in Sec. IV using the generalized brackets.

The Pfaffian plays, in the Clifford algebra, a role analogous to the determinant in the Grassmann algebra. The properties of both algebras correspond to related properties of the mathematical objects used to define them except that, in the case of the Clifford algebra, this object is more complicated than in the analogous Grassmann case. This extra complication is caused by the semigradation of the algebra and is studied in a separate paper[1] which deals mainly with the associativity property. In the Grassmann algebra this property leads to the Laplace expansion of a determinant, and in the Clifford case it leads to an expansion of the Pfaffian analogous to the Laplace expansion.

II. PROPERTIES OF THE CLIFFORD ALGEBRA

The presentation of the Clifford algebra given here is taken from Raševskiĭ,[2] since this approach reveals the exterior subspace and introduces the concept of a versor most directly. The Clifford algebra will be assumed to be finite and related to a positive definite quadratic form, so that we use a vector space of dimension n, denoted by $\mathbf{R}^n$, along with the standard metric. Following the notation of Porteous,[3] the universal Clifford algebra over $\mathbf{R}^n$ is denoted by a lowered index $\mathbf{R}_n$. The basis of $\mathbf{R}^n$, $\mathbf{e}_1,...,\mathbf{e}_n$, is identified with the basis of the vector subspace of $\mathbf{R}_n$ denoted by $\mathbf{R}_n^{(1)}$, and the quadratic form of $\mathbf{R}_n$ is defined by a symmetric, bilinear mapping on $\mathbf{R}^n$, (,), called the interior product. The scalar subspace is taken to be the real numbers, $\mathbf{R}_n^{(0)} = \mathbf{R}$. In the following we denote the Clifford product by the absence of a product symbol and an arbitrary element of $\mathbf{R}_n$ by capitals. Lowercase letters denote scalars and boldface letters are the vectors of $\mathbf{R}_n^{(1)}$.

The following are properties of the Clifford algebra, given by Raševskiĭ:

$$(A+B)C = AC + BC, \tag{2.1a}$$

$$C(A+B) = CA + CB, \tag{2.1b}$$

$$(AB)C = A(BC), \tag{2.2}$$

$$aB = Ba, \tag{2.3}$$

$$\mathbf{aa} = (\mathbf{a},\mathbf{a}). \tag{2.4a}$$

Property (2.4a) will be called the contraction property, and it is this property that characterizes the Clifford algebra.

0022-2488/88/112338-08$02.50

Introducing the metric tensor, $(\mathbf{e}_i, \mathbf{e}_j) = g_{ij}$, then $(\mathbf{e}_i + \mathbf{e}_j)^2$ in (2.4a) gives the Jordan relation[4]

$$\mathbf{e}_i \mathbf{e}_j + \mathbf{e}_j \mathbf{e}_i = 2g_{ij}. \tag{2.4b}$$

This relation contains the fundamental, geometric property that orthogonal vectors anticommute, since $g_{ij} = 0$ for $i \neq j$ with an orthogonal basis. This property may be used inductively to generate the basis of the entire algebra, called the polyvector basis.

A simple polyvector is defined by Raševskiĭ as complete antisymmetrization of vectors under the Clifford product. Using the alteration operation Alt, a basis polyvector is denoted by

$$\begin{aligned}\mathbf{e}_{[i_1,i_2,\ldots,i_r]} &= \mathrm{Alt}\,\frac{1}{r!}\,\mathbf{e}_{i_1}\mathbf{e}_{i_2}\ldots\mathbf{e}_{i_r}\\ &= \frac{1}{r!}\sum_{\pi\in S^r}\sigma(\pi)\mathbf{e}_{i_{\pi_1}}\mathbf{e}_{i_{\pi_2}}\cdots\mathbf{e}_{i_{\pi_r}},\end{aligned} \tag{2.5a}$$

where S^r is the set of permutations of $1,2,\ldots,r$ and $\sigma(\pi)$ is the sign of the permutation π.

There are $C_r^n = \binom{n}{r}$ basis polyvectors of tensor degree r and these form a basis of the subspace $\mathbf{R}_n^{(r)}$. Members of this space are called polyvectors of valence r. Since there are no polyvectors having valence greater than n, $\mathbf{R}_n$ contains 2^n basis polyvectors. Thus we may view the Clifford algebra as an aggregate of vector spaces, $\mathbf{R}_n \cong \oplus_{r=0}^{n} \mathbf{R}^{C_r^n}$, and denote elements of the polyvector subspaces by suffixing the bracketed valence of the polyvector.

Hence the grading over the polyvector basis of $\mathbf{R}_n$ is denoted by

$$A = A^{(0)} + A^{(1)} + \cdots + A^{(n)},$$

where

$$A^{(r)} \in \mathbf{R}_n^{(r)},$$

and any polyvector part may be expanded on the polyvector basis as

$$A^{(r)} = \sum_{i_1 < i_2 < \cdots < i_r} a^{i_1 i_2 \cdots i_r} \mathbf{e}_{[i_1,i_2,\ldots,i_r]}, \quad a^{i_1,i_2,\ldots,i_r} \in \mathbf{R}.$$

The scalar part is taken to be simply $A^{(0)} = a \in \mathbf{R}$. Also we denote the semigraded elements by $A^{(\pm)} \in \mathbf{R}_n^{(\pm)}$, where $\mathbf{R}_n^{(+)}$ denotes the subalgebra of $\mathbf{R}_n$ consisting of polyvectors of even valence and $\mathbf{R}_n^{(-)}$ is the subspace of odd valence polyvectors.

The construction of the basis polyvector (2.5a) produces a Grassmann form in the tensor algebra. For Clifford forms we need to consider basis elements that are graded in the tensor representation because of the contraction property. For this purpose we define the versor of degree r or r-versor to consist of the Clifford product of r vectors: $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, $\mathbf{a}_i \in \mathbf{R}_n^{(1)}$. These may have many different valence polyvector parts given, in general, by the $\mathbf{Z}_2$ graded element

$$A = \mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_r = A^{(r)} + A^{(r-2)} + \cdots + A^{(r \bmod 2)}. \tag{2.6}$$

Versors consisting of only single valence polyvectors, $A = A^{(r)}$, are completely antisymmetric and will be called exterior.

The final property of the Clifford algebra, given by Raševskiĭ,[2] is simply the statement that the polyvector basis is isomorphic to the basis of the Grassmann or exterior algebra:

$$\mathrm{Alt}\,\mathbf{a}_1\mathbf{a}_2\cdots\mathbf{a}_r = \sum_{i_1,i_2,\ldots,i_r} |a_1^{i_1} a_2^{i_2} \ldots a_r^{i_r}|\,\mathbf{e}_{[i_1,i_2,\ldots,i_r]}, \tag{2.5b}$$

where $|a_1^{i_1} a_2^{i_2} \cdots a_r^{i_r}|$ is the determinant of the matrix of components $\|a_i^j\|$, $j = 1,\ldots,n$, of the arbitrary rank-1 tensors a_i, $i = 1,2,\ldots,r$. This is the exterior component $A^{(r)}$ of the multivector (2.6). To find the structure of Clifford algebra we must construct the remaining polyvector components of the versor (2.6) in terms of functions of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$. Chevalley[5] and Raševskiĭ[2] both give an algorithm to produce this structure given the components of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ over an orthonormal basis. What we wish to arrive at is an explicit statement of the versor expansion of (2.6), which is independent of a basis.

Having defined the main elements of the algebra, the versor and the polyvector, we now move to the Lie algebra, which is extended to the entire Clifford algebra through the use of the anticommutator product.

The commutator and anticommutator products will be denoted by square brackets and braces, respectively. Since the Clifford algebra is associative, the commutator algebra is a Lie algebra satisfying the distributive property and the following derivation:

$$[A, B_1 B_2 \cdots B_r] = \sum_{j=1}^{r} B_1 B_2 \cdots B_{j-1} [A, B_j] B_{j+1} \cdots B_r.$$

For $r = 2$, this may be polarized into the Jacobi identity and an identical relation for the anticommutator:

$$[A,[B,C]] = [[A,B],C] + [B,[A,C]], \tag{2.7}$$

$$[A,\{B,C\}] = \{[A,B],C\} + \{B,[A,C]\}. \tag{2.8}$$

Alternative expansions for both of these bracket relations may be proved by expansion:

$$[A,[B,C]] = \{\{A,B\},C\} - \{B,\{A,C\}\}, \tag{2.7'}$$

$$[A,\{B,C\}] = [\{A,B\},C] - [B,\{A,C\}]. \tag{2.8'}$$

These are the required four identities that express all relations between the two brackets.

These relations can be used to decompose any Clifford element using the multivector formalism. That is, an arbitrary Clifford element is equivalent to a sum of versors and each versor can be expanded via polarization of each vector product. For example, applying the identities above to vectors reproduces the Gibb's vector identities in a more general form that applies to an arbitrary dimension space:

$$\begin{aligned}[\mathbf{a},[\mathbf{b},\mathbf{c}]] &= [[\mathbf{a},\mathbf{b}],\mathbf{c}] + [\mathbf{b},[\mathbf{a},\mathbf{c}]]\\ &= \{\mathbf{a},\mathbf{b}\}\mathbf{c} - \{\mathbf{a},\mathbf{c}\}\mathbf{b},\end{aligned} \tag{2.9}$$

$$\{\mathbf{a},[\mathbf{b},\mathbf{c}]\} = \{[\mathbf{a},\mathbf{b}],\mathbf{c}\}. \tag{2.10}$$

The usual Gibb's identities are attained by restricting ourselves to $\mathbf{R}_3$ and defining the dot and cross products:

$$\mathbf{a}\cdot\mathbf{b} = \tfrac{1}{2}\{\mathbf{a},\mathbf{b}\}, \tag{2.11}$$

$$\mathbf{a}\times\mathbf{b} = \tfrac{1}{2}\,\mathbf{e}_{[3,2,1]}[\mathbf{a},\mathbf{b}], \tag{2.12}$$

where $\mathbf{e}_{[3,2,1]}$, the basis polyvector of valence 3 or volume form, belongs to the center of $\mathbf{R}_3$. It is interesting to note that the axial vector (2.12) cannot be defined without some knowledge of a basis of $\mathbf{R}_3$. Thus in order to be able to interpret the geometrical content of the Gibb's identities, we have placed ourselves in the double bind of being restricted to three dimensions and having basis-dependent identities.

The Clifford algebraic approach to escape this bind is simply to interpret the polyvectors as geometric objects. The bivector, denoted using the wedge product of two vectors,

$$\mathbf{a}\wedge\mathbf{b}=\tfrac{1}{2}\,[\mathbf{a},\mathbf{b}],$$

represents the planar subspace containing the two vectors, unless they are colinear. At this stage it is not easy to see that the bivectors are generators of rotations, but it is easy to verify, using (2.7), (2.9), and (2.11), that the Lie algebra of the subspace $\mathbf{R}_n^{(2)}$ is the Lie algebra of the orthogonal group:

$$\begin{aligned}[\mathbf{a}\wedge\mathbf{b},\mathbf{c}\wedge\mathbf{d}] &= \tfrac{1}{2}\,[\mathbf{a},[\mathbf{b},\mathbf{c}\wedge\mathbf{d}]]-\tfrac{1}{2}\,[\mathbf{b},[\mathbf{a},\mathbf{c}\wedge\mathbf{d}]]\\ &= 2(\mathbf{b}\cdot\mathbf{c})\mathbf{a}\wedge\mathbf{d}-2(\mathbf{b}\cdot\mathbf{d})\mathbf{a}\wedge\mathbf{c}\\ &\quad -2(\mathbf{a}\cdot\mathbf{c})\mathbf{b}\wedge\mathbf{d}+2(\mathbf{a}\cdot\mathbf{d})\mathbf{b}\wedge\mathbf{c}. \qquad (2.13)\end{aligned}$$

Other elements of possible interest that we could decompose, such as the anticommutator of bivectors, are left to Sec. III. First we analyze the versor of degree 2 to find its "complex" exponential form. Evaluating the square of the sum of two vectors introduces the cosine function, $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\cdot|\mathbf{b}|\cos\theta$, where $|\mathbf{a}|^2=\mathbf{a}^2$ and θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. Squaring a bivector and evaluating using the Clifford identities gives $(\mathbf{a}\wedge\mathbf{b})^2=-|\mathbf{a}|^2|\mathbf{b}|^2\sin^2\theta$, so that we may define the "unit" bivector $B^{(2)}=\hat{\mathbf{a}}\wedge\hat{\mathbf{b}}/\sin\theta$, where $\hat{\mathbf{a}}=\mathbf{a}/|\mathbf{a}|$. This enables us to evaluate a two-versor in $\mathbf{R}_n$ as

$$\mathbf{ab}=\mathbf{a}\cdot\mathbf{b}+\mathbf{a}\wedge\mathbf{b}=|\mathbf{a}||\mathbf{b}|(\cos\theta+B^{(2)}\sin\theta). \qquad (2.14)$$

The element in brackets is the required exponential $\hat{\mathbf{a}}\hat{\mathbf{b}}=\exp(B^{(2)}\theta)$ and in $\mathbf{R}_3$ it is identical, apart from sign, to the quaternion versor since $\mathbf{R}_3^{(+)}\cong\mathbf{H}$, the quaternion algebra.

Using the Clifford identities and the multivector formalism, we have seen how easy it is to expose a powerful geometric algebra from fundamental algebraic concepts. The main importance of this approach is that the results are basis independent, the proofs relying only on the symmetry properties of the Clifford product. For products involving only vectors, the brackets have a simple geometric interpretation. But this breaks down for higher valence polyvectors, making the multivector decomposition impractical for Clifford elements of arbitrary valence. For example, the Gibb's scalar triple product, represented by (2.10), is a geometrical expression of the volume form. Hence the geometrical interpretation of the anticommutator depends upon the valence of the polyvector products. The graded products of Sec. III rectify this situation.

III. GRADATION OF THE CLIFFORD ALGEBRA

Similar to the Lie algebra, a graded algebra for an associative algebra is defined using the graded commutator.[8] This also holds for semigraded algebras, so we define the graded commutator, denoted by a circle, as

$$A^{(r)}\circ B^{(s)}=\tfrac{1}{2}\,(A^{(r)}B^{(s)}-(-1)^{rs}B^{(s)}A^{(r)}). \qquad (3.1)$$

This will be called the superproduct and produces a $\mathbf{Z}_2$-graded Lie structure:

$$\begin{aligned}&A^{(\pm)}\circ B^{(\pm)}\in\mathbf{R}_n^{(+)},\quad A^{(\pm)}\circ B^{(\mp)}\in\mathbf{R}_n^{(-)},\\ &A^{(r)}\circ B^{(s)}=-(-1)^{rs}B^{(s)}\circ A^{(r)}, \qquad (3.2)\\ &A^{(r)}\circ(B_1^{(s_1)}B_2^{(s_2)}\cdots B_k^{(s_k)})\\ &\quad=\sum_{j=1}^{k}(-1)^{r(s_1+s_2+\cdots+s_{j-1})}B_1^{(s_1)}B_2^{(s_2)}\cdots\\ &\qquad\times B_{j-1}^{(s_j-1)}(A^{(r)}\circ B_j^{(s_j)})B_{j+1}^{(s_j+1)}\cdots B_k^{(s_k)}. \qquad (3.3)\end{aligned}$$

The superproduct generates the semigraded Lie algebra of the Clifford algebra. Further subalgebras are again given by $\mathbf{R}_n^{(2)}$ and $\mathbf{R}_n^{(+)}$, but these reduce to the ordinary Lie subalgebras. We extend the graded subalgebra to the entire Clifford algebra by defining the generalized exterior product which is denoted by the wedge–bar symbol:

$$A^{(r)}\overline{\wedge}B^{(s)}=\tfrac{1}{2}\,(A^{(r)}B^{(s)}+(-1)^{rs}B^{(s)}A^{(r)}). \qquad (3.4)$$

This product has the same commutivity as the Grassmann or exterior product:

$$A^{(r)}\overline{\wedge}B^{(s)}=(-1)^{rs}B^{(s)}\overline{\wedge}A^{(r)}. \qquad (3.5)$$

In general, the generalized exterior product has the following graded polyvector representation:

$$\begin{aligned}A^{(r)}\overline{\wedge}B^{(s)}=C&=C^{(r+s)}+C^{(r+s-4)}\\ &\quad+\cdots+C^{[(r+s)\bmod 4]}.\end{aligned}$$

This is a generalization of the exterior product since it contains the exterior polyvector part $C^{(r+s)}$. All terms $C^{(k)}$, where $k<|r-s|$, are zero, since $C^{(|r-s|)}$ is the lowest valence polyvector part of $A^{(r)}B^{(s)}$ and hence the series terminates at the polyvector of valence $r+s-4[\min(r/2,s/2)]$, where $[x]$ is the integer part of x. The remaining polyvector components of the Clifford product of $A^{(r)}$ and $B^{(s)}$ are given by (3.1), thus giving a complete decomposition of $A^{(r)}B^{(s)}$. The lowest valence polyvector part, $C^{(|r-s|)}$, may be contained in either (3.1) or (3.4).

The graded derivation (3.3), for the case $r=2$, leads to the graded Jacobi identity and a similar expression for the generalized exterior product:

$$\begin{aligned}A^{(r)}\circ(B^{(s)}\circ C)&=(A^{(r)}\circ B^{(s)})\circ C\\ &\quad+(-1)^{rs}B^{(s)}\circ(A^{(r)}\circ C), \qquad (3.6)\\ A^{(r)}\circ(B^{(s)}\overline{\wedge}C)&=(A^{(r)}\circ B^{(s)})\overline{\wedge}C\\ &\quad+(-1)^{rs}B^{(s)}\overline{\wedge}(A^{(r)}\circ C). \qquad (3.7)\end{aligned}$$

Again we provide another two identities that may be proved by simple expansion:

$$\begin{aligned}A^{(r)}\overline{\wedge}(B^{(s)}\overline{\wedge}C)&=(A^{(r)}\overline{\wedge}B^{(s)})\overline{\wedge}C\\ &\quad-(-1)^{rs}B^{(s)}\circ(A^{(r)}\circ C), \qquad (3.6')\\ A^{(r)}\circ(B^{(s)}\overline{\wedge}C)&=(A^{(r)}\overline{\wedge}B^{(s)})\circ C\\ &\quad-(-1)^{rs}B^{(s)}\circ(A^{(r)}\overline{\wedge}C). \qquad (3.7')\end{aligned}$$

The first expresses the essential property of nonassociativity of the generalized exterior product. Note that these identities are actually semigraded and so can be applied to versors whose degree replaces the valence of the polyvectors.

We reduce the four identities to relations for vectors and bivectors. Setting $A=\mathbf{a}$ and $C=\mathbf{b}$ in (3.6b) gives $\mathbf{a}\,\overline{\wedge}\,(B\,\overline{\wedge}\,\mathbf{b})=(\mathbf{a}\,\overline{\wedge}\,B)\,\overline{\wedge}\,\mathbf{b}$, which may be written as

$$\mathbf{a}\,\overline{\wedge}\,(\mathbf{b}\,\overline{\wedge}\,C)=-\mathbf{b}\,\overline{\wedge}\,(\mathbf{a}\,\overline{\wedge}\,C). \tag{3.8}$$

Also, setting $r=s=1$ in (3.6) and (3.7) gives

$$\mathbf{a}\circ(\mathbf{b}\circ C)=-\mathbf{b}\circ(\mathbf{a}\circ C), \tag{3.9}$$

$$(\mathbf{a}\cdot\mathbf{b})C=\mathbf{a}\circ(\mathbf{b}\,\overline{\wedge}\,C)+\mathbf{b}\,\overline{\wedge}\,(\mathbf{a}\circ C), \tag{3.10}$$

$$(\mathbf{a}\,\overline{\wedge}\,\mathbf{b})\,\overline{\wedge}\,C=\mathbf{a}\,\overline{\wedge}\,(\mathbf{b}\,\overline{\wedge}\,C)+\mathbf{a}\circ(\mathbf{b}\circ C), \tag{3.11}$$

$$(\mathbf{a}\,\overline{\wedge}\,\mathbf{b})\circ C=\mathbf{a}\circ(\mathbf{b}\,\overline{\wedge}\,C)-\mathbf{b}\circ(\mathbf{a}\,\overline{\wedge}\,C), \tag{3.12a}$$

$$=\mathbf{a}\,\overline{\wedge}\,(\mathbf{b}\circ C)-\mathbf{b}\,\overline{\wedge}\,(\mathbf{a}\circ C), \tag{3.12b}$$

where (3.10) has been used twice in (3.12a) to derive (3.12b). Here (3.8) and (3.9) formalize identities employed by Greider[9]: $\mathbf{a}\circ(\mathbf{a}\circ C)=\mathbf{a}\,\overline{\wedge}\,(\mathbf{a}\,\overline{\wedge}\,C)=0$.

This may be generalized to arbitrary simple elements using equations (2.8); $[A,\{A,B\}]=0$ and $\{A,[A,C]\}=0$. These ordinary bracket identities have graded representations that depend on the semigrading of A in the following way:

$$A\circ(A\circ C)=A\,\overline{\wedge}\,(A\,\overline{\wedge}\,C)=0,\quad \text{if } A=A^{(-)}; \tag{3.13a}$$

$$A\,\overline{\wedge}\,(A\circ C)=A\circ(A\,\overline{\wedge}\,C)=0,\quad \text{if } A=A^{(+)}. \tag{3.13b}$$

The vector and bivector identities provide a complete set of relations that can be used to decompose any multivector from the left. This process is aided by the geometrical interpretation of the graded products that is displayed in the identities. For example, $\mathbf{a}\,\overline{\wedge}\,C^{(r)}$ is the exterior or valence $r+1$ part of $\mathbf{a}C^{(r)}$, while $\mathbf{a}\circ C^{(r)}$ is the contracted or valence $r-1$ polyvector part. This exterior property of the wedge–bar product for vectors has been used by Jacobson[6] and Hestenes[10] to inductively define an exterior multivector of degree r:

$$\mathbf{e}_{[i_1,i_2,\ldots,i_r]}=\mathbf{e}_{i_1}\,\overline{\wedge}\,(\mathbf{e}_{i_2}\,\overline{\wedge}\cdots(\mathbf{e}_{i_{r-1}}\,\overline{\wedge}\,\mathbf{e}_{i_r})). \tag{3.14}$$

This is indeed exterior since, by (3.8), any exchange of vectors brings about a change in sign. Because of this alternating property, we expect (3.14) to be identical to the basis polyvector (2.5) in the Clifford space. In fact, the contraction property cancels any difference between the two definitions of this element, and so there is no need to distinguish between them. Hence it is possible to define an exterior product in Clifford algebra without resorting to complete antisymmetrization. This product is denoted by the usual wedge symbol,

$$\mathbf{A}^{(r)}=\mathbf{a}_1\wedge\mathbf{a}_2\wedge\cdots\wedge\mathbf{a}_r=\mathbf{a}_1\,\overline{\wedge}\,(\mathbf{a}_2\,\overline{\wedge}\cdots\overline{\wedge}\,(\mathbf{a}_{r-1}\,\overline{\wedge}\,\mathbf{a}_r)).$$

The following graded derivation for basis polyvectors may be proved inductively using (3.10):

$$\mathbf{e}_i\circ\mathbf{e}_{[i_1,\ldots,i_r]}=\sum_{j=1}^{r}(-1)^{j-1}g_{ii_j}\mathbf{e}_{[i_1,\ldots,\hat{\imath}_j,\ldots,i_r]}, \tag{3.15}$$

where the index denoted with a circumflex is omitted.

The identities (3.8)–(3.12) carry the same information as similar identities given, using the ordinary bracket products, but are geometrically easier to decipher. For example, considering either (3.8) or (3.11) with $C=\mathbf{c}$ gives

$$\mathbf{a}\,\overline{\wedge}\,(\mathbf{b}\,\overline{\wedge}\,\mathbf{c})=(\mathbf{a}\,\overline{\wedge}\,\mathbf{b})\,\overline{\wedge}\,\mathbf{c}. \tag{3.16}$$

This is immediately recognized as a three-volume element and shows that in $\mathbf{R}_3$ the generalized exterior product is associative and thus defines an exterior product. In $\mathbf{R}_3$ it is equivalent to the scalar triple product with the transformation between the two being carried out by the duality transformation. This is multiplication by $\mathbf{e}_{[3,2,1]}$ which commutes in $\mathbf{R}_3$ but changes the gradation of the products in (3.16).

Using this geometrical identification of the graded products, it is now easier to analyze the anticommutator of two bivectors that, in this case, corresponds to the generalized exterior product. Using (3.11), we find that this contains both the exterior part and the scalar part of the polyvector product:

$$(\mathbf{a}\wedge\mathbf{b})\,\overline{\wedge}\,(\mathbf{c}\wedge\mathbf{d})=\mathbf{a}\wedge\mathbf{b}\wedge\mathbf{c}\wedge\mathbf{d}-(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})+(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}).$$

Thus the evaluation of the square of a bivector is simply

$$(\mathbf{a}\wedge\mathbf{b})^2=-(\mathbf{a}^2\mathbf{b}^2-(\mathbf{a}\cdot\mathbf{b})^2)=-\mathbf{a}^2\mathbf{b}^2\sin^2\theta,$$

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. Apart from the negative sign, this is the norm of the Clifford algebra, so the bivector "length" is $|\mathbf{b}\,\overline{\wedge}\,\mathbf{c}|=|\mathbf{b}|\,|\mathbf{c}|\sin\theta$.

The graded identity that reproduces the triple vector identity, represented by (2.9), is given by (3.10) with $C=\mathbf{c}$:

$$\mathbf{a}\circ(\mathbf{b}\,\overline{\wedge}\,\mathbf{c})=(\mathbf{a}\cdot\mathbf{b})\mathbf{c}-(\mathbf{a}\cdot\mathbf{c})\mathbf{b}. \tag{3.17}$$

This identity exposes the triple vector identity as a simple statement of the graded derivation. The action of the bivector on the vector $\mathbf{a}$ under the superproduct is immediately recognized as a rotation of $\mathbf{a}$ by 90°, for if we try to contract (3.17) with $\mathbf{a}$, then by (3.13a) we find they are orthogonal: $\mathbf{a}\circ(\mathbf{a}\circ(\mathbf{b}\,\overline{\wedge}\,\mathbf{c}))=0$. Of course, the rotation is within the plane of the bivector and, by setting $\mathbf{a}=\mathbf{b}$, we find that the rotation has the sense of $\mathbf{b}$ moving towards $\mathbf{c}$. The geometrical action of the superproduct of a bivector with an arbitrary element may be analyzed using (3.6′), with $A=B=B^{(2)}$,

$$\begin{aligned}B\circ(B\circ C)&=(B\,\overline{\wedge}\,B)\,\overline{\wedge}\,C-B\,\overline{\wedge}\,(B\,\overline{\wedge}\,C)\\&=-|B|^2C-B\,\overline{\wedge}\,(B\,\overline{\wedge}\,C).\end{aligned} \tag{3.18}$$

Using (3.13b), we find that $-BBC$ is a polarization of C into components with even or odd numbers of vectors lying in the plane of the bivector B. If $C=B\circ C'$ then we see that (3.18) is a rotation of 180° since, by (3.13b), $B\circ C'$ is already contracted. Therefore $B\circ C$ is a selection of that part of C with a single vector in the plane of B and this vector is rotated by 90°. This is also the interpretation of the Lie product of bivectors (2.13). The remaining terms in the product BC are the semigraded exterior parts $B\,\overline{\wedge}\,C$, and we can use the polarization of (3.18) to select either part of C. With respect to the unit bivector $B^{(2)}=\mathbf{b}\,\overline{\wedge}\,\mathbf{c}/|\mathbf{b}\,\overline{\wedge}\,\mathbf{c}|$, we define the parallel or contracted component of A,

$$A^{\parallel}=-B^{(2)}\circ(B^{(2)}\circ A)=\tfrac{1}{2}(B^{(2)}AB^{(2)}+A).$$

For example, if $A=\mathbf{a}$ then this is the projection of the vector onto the plane of $\mathbf{b}\wedge\mathbf{c}$. The perpendicular part is $A^{\perp}=A-A^{\parallel}=-B^{(2)}\,\overline{\wedge}\,(B^{(2)}\,\overline{\wedge}\,A)$.

We can now easily interpret the transformation law, $A'=VAV^{-1}$, where V is two-versor, as a rotation. It is decomposed using the commutation relations $A^{\perp}\circ B^{(2)}=A^{\parallel}\,\overline{\wedge}\,B^{(2)}=0$ in (3.2) and (3.5):

$$
\begin{aligned}
A' &= \exp(-B^{(2)}\theta/2)A\exp(B^{(2)}\theta/2)\\
&= A^{\perp} + A^{\parallel}\exp(B^{(2)}\theta)\\
&= A^{\perp} + A^{\parallel}(\hat{\mathbf{b}}\cdot\hat{\mathbf{c}}) + A\circ(\hat{\mathbf{b}}\,\overline{\wedge}\,\hat{\mathbf{c}}),
\end{aligned}
$$

where we have used $\exp(B^{(2)}\theta) = \hat{\mathbf{b}}\hat{\mathbf{c}} = \hat{\mathbf{b}}\circ\hat{\mathbf{c}} + \hat{\mathbf{b}}\,\overline{\wedge}\,\hat{\mathbf{c}}$. This is recognized as a rotation of vector components lying in the bivector plane through angle θ. For vectors it is identical to the usual vector notation apart from the generalized triple vector product (3.17) which applies in a vector space of arbitrary dimension. The geometrical identification of the graded products allows for a concise evaluation of a rotation for any multivector. For more general versors V, Raševskiĭ[2] has shown that the transformation $A' = VAV^{-1}$ is an automorphism of the Clifford algebra defining the orthogonal group, proper or improper, for the versor being, respectively, of even or odd degree.

The differences between the identities for the graded and ordinary products are only cosmetic, but it is the straightforward expression of the geometrical content that makes the graded products easier to manipulate. The geometrical description of the graded products, in general, is simply that the superproduct contains those parts of the Clifford product that consist of an odd number of contractions, while the generalized exterior product selects pairs of contradictions including, formally, the exterior part. This is consistent with the supersymmetry properties of the graded products and the geometrical content of (3.11) and (3.12). The ordinary brackets mix this natural grading and thus do not reflect the geometrical operations in a consistent way, and it is for this reason that the graded products may be considered superior to the ordinary products.

IV. THE FUNDAMENTAL THEOREM

The analysis of the structure of a versor reveals that the $\mathbf{Z}_2$ gradation (2.6) may be described as being completely contracting. By this it is meant that the contraction property (2.4b) is applied between all possible pairs of factors of the versor remembering that noncontracting pairs must anticommute. This structure is conveniently given by the Pfaffian, which is simply a subset of the expansion of the determinant. Before defining the Pfaffian, it is necessary to introduce certain well known sets to be used extensively in what follows.

The sets consist of arrangements of labels which may be taken to be subsets of the natural numbers $\mathbf{N}$. We denote ordered sequences of subsets of $\mathbf{N}$ by

$$\mathbf{N}_p^q = (p,p+1,\ldots,q)$$

where

$$1 \leqslant p \leqslant q \in \mathbf{N}.$$

We adopt the notation that Greek letters denote the reorderings of the sequences and thus represent bijections of the labels $\mathbf{N}_p^q$. The ith component of a particular arrangement ρ_i is the result of ρ acting on $i \in \mathbf{N}_p^q$. Of course, the original order of $\mathbf{N}_p^q$ is important and is always taken to be the natural order.

The first set, which contains the other sets to be defined, is the permutation or symmetric set, $S^n(\mathbf{N}_{p+1}^{p+n})$. This is the set of $n!$ permutations of the n labels $\mathbf{N}_{p+1}^{p+n}$. Any particular permutation $\pi \in S^n(\mathbf{N}_1^n)$ has sgn $\sigma(\pi)$ given by the parity of the rearrangement

$$\begin{pmatrix} 1,2,\ldots,n \\ \pi_1,\pi_2,\ldots,\pi_n \end{pmatrix}.$$

The remaining sets are defined as subsets of S^n by placing ordering restrictions on the components of the bijections. The normal ordered set, $N^{2r}(\mathbf{N}_1^{2r})$, takes all members $\rho \in S^{2r}(\mathbf{N}_1^{2r})$ such that (i) for any $k \in \mathbf{N}_1^r$, $\rho_{2k-1} < \rho_{2k}$ and (ii) for any $h,k \in \mathbf{N}_1^r$, $h < k \Rightarrow \rho_{2h-1} < \rho_{2k-1}$. For each k in condition (i) the number of members in the set is halved, while condition (ii) selects only one of each $r!$ ordering given by S^{2r}. Hence the cardinality of the set is $(2r)!/(2^r r!)$.

The partition set $P_{p,q}^n(\mathbf{N}_1^n)$ has members $\mu \in S^n(\mathbf{N}_1^n)$, such that $\mu_1 < \cdots < \mu_p$, $\mu_{p+1} < \cdots < \mu_{p+q}$, and $\mu_{p+q+1} < \cdots < \mu_n$. For $q = 0$ this reduces to the combination set, $C_p^n(\mathbf{N}_1^n)$, which has $n!/((n-p)!p!)$ members as a result of the restrictions $\mu_1 < \mu_2 < \cdots < \mu_p$ and $\mu_{p+1} < \mu_{p+2} < \cdots < \mu_n$. This, of course, corresponds to the combinatorics problem of partitioning n labels into two parts and satisfies $C_0^n(\mathbf{N}_1^n) = C_n^n(\mathbf{N}_1^n) = \mathbf{N}_1^n$. The reason for formulating this set in this way is simply to expose the parity assigned to each member of these sets.

It may be noticed that the partition set $P_{p,q}^n(\mathbf{N}_1^n)$ can be constructed as the composition of two combination sets, but that this may be done in two ways. Thus we have the following identity to be used later:

$$
\begin{aligned}
&\sum_{\mu \in P_{p,q}^n(\mathbf{N}_{m+1}^{m+n})} \sigma(\mu)\, f_{\mu_1\mu_2\cdots\mu_n}\\
&\quad= \sum_{\mu \in C_{p+q}^n(\mathbf{N}_{m+1}^{m+n})} \sigma(\mu)\\
&\qquad\times \sum_{\nu \in C_p^{p+q}(\mathbf{N}_1^{p+q})} \sigma(\nu) f_{\mu_{\nu_1}\mu_{\nu_2}\cdots\mu_{\nu_{p+q}}\mu_{p+q+1}\cdots\mu_n}\\
&\quad= \sum_{\mu \in C_p^n(\mathbf{N}_{m+1}^{m+n})} \sigma(\mu)\\
&\qquad\times \sum_{\nu \in C_q^{n-p}(\mathbf{N}_{p+1}^n)} \sigma(\nu)\, f_{\mu_1\mu_2\cdots\mu_p\mu_{\nu_1}\cdots\mu_{\nu_{n-p}}}.
\end{aligned}
\tag{4.1}
$$

We can now return to the introduction of the Pfaffian. Because of their close association, we define both the Pfaffian and the determinant using the above sets.

A Pfaffian is characterized as being the square root of the determinant of an even-dimensional, antisymmetric matrix. For matrix A having components a_{ij}, $i,j \in \mathbf{N}_1^n$, its determinant is denoted by $|a_{ij}|$ or, in terms of the diagonal components, as $|a_{11}a_{22}\cdots a_{nn}|$. If this matrix has components $a_{ij} = -a_{ji}$, $i,j \in \mathbf{N}_1^{2r}$, then the Pfaffian of the upper diagonal half of A has notation $\backslash a_{ij}|$, or in terms of its upper components, $\backslash a_{12}a_{23}\cdots a_{n-1n}|$, $n = 2r$.

The determinant of rank n and Pfaffian of order r are defined by Porteous[3] as

$$|a_{11},a_{22},\ldots,a_{nn}| = \sum_{\pi \in S^n(\mathbf{N}_1^n)} \sigma(\pi) \prod_{k \in \mathbf{N}_1^n} a_{k\pi_k}, \tag{4.2}$$

$$\backslash a_{12},a_{23},\ldots,a_{2r-1\,2r}| = \sum_{\rho \in N^{2r}(\mathbf{N}_1^n)} \sigma(\rho) \prod_{k \in \mathbf{N}_1^r} a_{\rho_{2k-1}\rho_{2k}}. \tag{4.3}$$

The Pfaffian has many properties in common with the determinant, the main one of interest here being the Pfaffian cofactor expansion,[11] which has identical form to the determinant cofactor expansion. Before this well known result is presented, an important mathematical object is introducted. This is the Pfaffian containing components of the metric tensor (2.4b). We define

$$T_{i_1 i_2 \cdots i_k} = \begin{cases} \backslash g_{i_1 i_2} \cdots g_{i_{k-1} i_k} |, & k \text{ even;} \\ 0, & k \text{ odd.} \end{cases} \tag{4.4}$$

In terms of this object the Pfaffian cofactor expansion is

$$\begin{aligned} T_{i_1 i_2 \cdots i_k} &= \sum_{\nu \in C^{k-1}_{k-2}(N^k_2)} \sigma(\nu) g_{i_1 i_{\nu_{k-1}}} T_{i_{\nu_2} i_{\nu_3} \cdots i_{\nu_{k-1}}} \\ &= \sum_{\nu \in C^{k-1}_{1}(N^k_2)} \sigma(\nu) g_{i_1 i_{\nu_1}} T_{i_{\nu_2} i_{\nu_3} \cdots i_{\nu_{k-1}}} \\ &= \sum_{j=2}^{k} (-1)^j g_{i_1 j_j} T_{i_2 i_3 \cdots \hat{i}_j \cdots i_k}, \end{aligned} \tag{4.5}$$

where again the circumflex denotes the omission of an index.

If we continue this expansion we find that $T_{i_1 i_2 \cdots i_k}$ contains terms giving all possible ways of pairing the indices of the tensor via the metric tensor. This is the property that makes Pfaffians useful in statistical mechanics[12] and it is inherited from the normal ordered set which pairs off the labels (ρ_{2i-1}, ρ_{2i}) in all possible ways. It is also exactly what is wanted in the structure of a versor, except that the factors not contracted using the metric tensor must be antisymmetric and representable by polyvectors. The tensor $T_{i_1 i_2 \cdots i_k}$ may be referred to as the contraction tensor and must be applied to each even subset of the basis versor indices. This is given in the following fundamental theorem connecting Grassmann and Clifford algebras.[13]

Theorem:

$$\begin{aligned} \mathbf{a}_1 \mathbf{a}_2 \cdots \mathbf{a}_r = \sum_{k=0}^{[r/2]} \sum_{\mu \in C^r_{2k}(N^r_1)} \sigma(\mu) \backslash \mathbf{a}_{\mu_1} \cdot \mathbf{a}_{\mu_2}, \ldots, \mathbf{a}_{\mu_{2k-1}} \cdot \mathbf{a}_{\mu_{2k}} \\ \times | \mathbf{a}_{\mu_{2k+1}} \wedge \cdots \wedge \mathbf{a}_{\mu_r}. \end{aligned} \tag{4.6}$$

Proof: It is only necessary to consider the versor basis so that the proposition becomes $\mathbf{e}_{i_1} \mathbf{e}_{i_2} \cdots \mathbf{e}_{i_r} = E_{i_1 i_2 \cdots i_r}$, where

$$E_{i_1 i_2 \cdots i_r} = \sum_{k=0}^{[r/2]} \sum_{\mu \in C^r_{2k}(\mathbf{N}^r_1)} \sigma(\mu) T_{i_{\mu_1} i_{\mu_2} \cdots i_{\mu_{2k}}} \mathbf{e}_{[i_{\mu_{2k+1}}, \ldots, i_{\mu_r}]}. \tag{4.7}$$

The proof of this statement is by induction with the basis being the $r = 2$ case. Using (2.14) we have

$$E_{i_1 i_2} = \mathbf{e}_{[i_1, i_2]} + g_{i_1 i_2} = \mathbf{e}_{i_1} \mathbf{e}_{i_2}.$$

We proceed to the general case by decomposing $E_{i_1 i_2 \cdots i_r}$ into terms with $\mu_1 = 1$ or $\mu_{2k+1} = 1$ [this task has been simplified by our definition of the combination set $C^r_{2k}(\mathbf{N}^r_1)$]:

$$\begin{aligned} E_{i_1 i_2 \cdots i_r} = &\sum_{k=0}^{[(r-1)/2]} \sum_{\mu \in C^{r-1}_{2k}(\mathbf{N}^r_2)} \sigma(\mu) T_{i_{\mu_1} i_{\mu_2} \cdots i_{\mu_{2k}}} \\ &\times \mathbf{e}_{[i_1, i_{\mu_{2k+1}}, \ldots, i_{\mu_{r-1}}]} \\ &+ \sum_{k=1}^{[r/2]} \sum_{\mu \in C^{r-1}_{2k-1}(\mathbf{N}^r_2)} \sigma(\mu) T_{i_1 i_{\mu_1} \cdots i_{\mu_{2k-1}}} \mathbf{e}_{[i_{\mu_{2k}}, \ldots, i_{\mu_{r-1}}]}. \end{aligned} \tag{4.8}$$

Note that for $k = 0$ there is no Pfaffian part, and for $k = [r/2]$ there is no polyvector part if r is even. Hence the changes to the limits of k. Also, there is no sign change in the first term as given by an even parity rearrangement of (4.7). Now we apply the Pfaffian cofactor expansion (4.5) to the second term

$$T_{i_1 i_{\mu_1} \cdots i_{\mu_{2k-1}}} = \sum_{\mu \in C^{2k-1}_{2k-2}(\mathbf{N}^{2k-1}_1)} \sigma(\mu) g_{i_1 i_{\mu_{\nu_{2k-1}}}} T_{i_{\mu_{\nu_1}} i_{\mu_{\nu_2}} \cdots i_{\mu_{\nu_{2k-2}}}}.$$

Exchanging the order of the combination sets in the second term of (4.8) and using (4.1) with $p = 2k - 2$, $q = 1$, $n = r - 1$, and $m = 1$ gives

$$\begin{aligned} &\sum_{\mu \in C^{r-1}_{2k-1}(\mathbf{N}^r_2)} \sigma(\mu) T_{i_1 i_{\mu_1} \cdots i_{\mu_{2k-1}}} \mathbf{e}_{[i_{\mu_{2k}}, \ldots, i_{\mu_{r-1}}]} \\ &= \sum_{\mu \in C^{r-1}_{2k-2}(\mathbf{N}^r_2)} \sigma(\mu) \sum_{\nu \in C^{r-2k+1}_1(\mathbf{N}^{r-1}_{2k-1})} \sigma(\nu) g_{i_1 i_{\mu_{\nu_1}}} \\ &\quad \times T_{i_{\mu_1} i_{\mu_2} \cdots i_{\mu_{2k-2}}} \mathbf{e}_{[i_{\mu_{\nu_2}}, \ldots, i_{\mu_{\nu_{r-2k+1}}}]} \\ &= \sum_{\mu \in C^{r-1}_{2k-2}(\mathbf{N}^r_2)} \sigma(\mu) \sum_{j=2k-1}^{r-1} (-1)^{j-1} g_{i_1 i_{u_j}} \\ &\quad \times T_{i_{u_1} i_{\mu_2} \cdots i_{\mu_{2k-2}}} \mathbf{e}_{[i_{\mu_{2k-1}} \cdots \hat{i}_{\mu_j} \cdots i_{\mu_{r-1}}]} \\ &= \sum_{\mu \in C^{r-1}_{2k-2}(\mathbf{N}^r_2)} \sigma(\mu) T_{i_{\mu_1} i_{\mu_2} \cdots i_{\mu_{2k-2}}} \mathbf{e}_{i_1} \circ \mathbf{e}_{[i_{\mu_{2k-1}}, \ldots, i_{\mu_{r-1}}]}, \end{aligned}$$

where we have also used (3.15). Notice that for $k = [(r+1)/2]$ there is no polyvector part for r odd, while, for r even, $[(r+1)/2] = [r/2]$. Hence there is no change if we take the upper k limit in the second term of (4.8) to be $[(r+1)/2]$. Changing the step of this k value to bring it in line with the other term in (4.8), and using (3.14) finally gives the statement of the inductive step:

$$\begin{aligned} E_{i_1 i_2 \cdots i_r} &= \sum_{k=0}^{[(r-1)/2]} \sum_{\mu \in C^{r-1}_{2k}(\mathbf{N}^r_2)} \sigma(\mu) T_{i_{\mu_1} i_{\mu_2} \cdots i_{\mu_{2k}}} \\ &\quad \times (\mathbf{e}_{i_1} \circ \mathbf{e}_{[i_{\mu_{2k+1}}, \ldots, i_{\mu_{r-1}}]} + \mathbf{e}_{i_1} \overline{\wedge} \mathbf{e}_{[i_{\mu_{2k+1}}, \ldots, i_{\mu_{r-1}}]}) \\ &= \mathbf{e}_{i_1} \circ E_{i_2 i_3 \cdots i_r} + \mathbf{e}_{i_1} \overline{\wedge} E_{i_2 i_3 \ldots i_r} \\ &= \mathbf{e}_{i_1} E_{i_2 i_3 \cdots i_r}. \end{aligned}$$

By the inductive argument, $E_{i_2 i_3 \cdots i_r} = \mathbf{e}_{i_2} \mathbf{e}_{i_3} \cdots \mathbf{e}_{i_r}$, and thus we have proved the proposition. The statement of the theorem follows from the multilinearity of the versor.

The Pfaffian in the versor expansion (4.6) carries a determinant part which is a fundamental part of the property of associativity. This structure is exposed in the following theorem on the Clifford product of two polyvectors.

Theorem:

$$\begin{aligned} &(\mathbf{a}_1 \wedge \mathbf{a}_2 \wedge \cdots \wedge \mathbf{a}_r)(\mathbf{b}_1 \wedge \mathbf{b}_2 \wedge \cdots \wedge \mathbf{b}_s) \\ &= \sum_{k=0}^{\min(r,s)} (-1)^{k(2r-k-1)/2} \sum_{\mu \in C^r_k(\mathbf{N}^r_1)} \sigma(\mu) \\ &\quad \times \sum_{\nu \in C^s_k(\mathbf{N}^s_1)} \sigma(\nu) (| \mathbf{a}_{\mu_1} \cdot \mathbf{b}_{\nu_1}, \mathbf{a}_{\mu_1} \cdot \mathbf{b}_{\nu_2}, \ldots, \mathbf{a}_{\mu_k} \cdot \mathbf{b}_{\nu_k} | \\ &\quad \cdot \mathbf{a}_{\mu_{k+1}} \wedge \cdots \wedge \mathbf{a}_{\mu_r} \wedge \mathbf{b}_{\nu_{k+1}} \wedge \cdots \wedge \mathbf{b}_{\nu_s}). \end{aligned} \tag{4.9}$$

Proof: Again, by multilinearity, we need only consider the polyvector basis. The basis of (4.9) will be some subset of the $(r+s)$-versor:

$$\mathbf{e}_{i_1}\mathbf{e}_{i_2}\cdots\mathbf{e}_{i_{r+s}}$$

$$= \sum_{k=0}^{[(r+s/2)]} \sum_{\mu\in C_{2k}^{r+s}(\mathbf{N}_1^{r+s})} \sigma(\mu) T_{i_{\mu_1} i_{\mu_2}\cdots i_{\mu_{2k}}} \mathbf{e}_{[i_{\mu_{2k+1}},\ldots,i_{\mu_{r+s}}]}. \tag{4.10}$$

To derive the structure of (4.9) we simply antisymmetrize separately the first r and the last s vector factors of (4.10), thus eliminating any components of $T_{i_{\mu_1} i_{\mu_2}\cdots i_{\mu_{2k}}}$ that represent contractions within the rank r and rank s polyvectors of (4.9). For example, the first term of the cofactor expansion will be annihilated for r greater than unity. Hence it is convenient to rearrange the indices of (4.10) before antisymmetrizing. For this rearrangement we choose

$$\begin{pmatrix} \mu_1,\mu_2,\mu_3,\mu_4,\ldots,\mu_{2k-1},\mu_{2k},\mu_{2k+1},\ldots,\mu_{r+k},\mu_{r+k+1},\ldots,\mu_{r+s} \\ \mu_1,\mu_{r+1},\mu_2,\mu_{r+2},\ldots,\mu_k,\mu_{r+k},\mu_{k+1},\ldots,\mu_r,\mu_{r+k+1},\ldots,\mu_{r+s} \end{pmatrix}.$$

This fixed permutation shuffles the elements of $C_{2k}^{r+s}(\mathbf{N}_1^{r+s})$ and does not violate the normal ordering rules for the indices of the contraction tensor. The parity of this arrangement is given by the number of exchanges involved in moving indices μ_{r+i} through μ_j, $\forall i\in\mathbf{N}_1^k$ and $\forall j\in\mathbf{N}_{i+1}^r$. This is $(r-1)+(r-2)+\cdots+(r-k)=kr-\frac{1}{2}k(k+1)$.

It is now straightforward to form the basis polyvectors, $\mathbf{e}_{[i_1,\ldots,i_r]}$ and $\mathbf{e}_{[i_{r+1},\ldots,i_{r+s}]}$, in (4.10), using the alternating maps $\pi\in S^r(\mathbf{N}_1^r)$ and $\rho\in S^s(\mathbf{N}_{r+1}^{r+s})$, respectively. These maps have no effect on the polyvector in (4.10), and so we concentrate on the normal ordered indices. The modified contraction tensor, $T_{i_{\pi_{\mu_1}} i_{\rho_{\mu_{r+1}}} i_{\pi_{\mu_2}} i_{\rho_{\mu_{r+2}}}\cdots i_{\pi_{\mu_k}} i_{\rho_{\mu_{r+k}}}}$, has labels that satisfy $\pi_{\mu_i}<\rho_{\mu_{r+i}}$ because of the normal ordering condition. The mappings π and ρ extend this inequality to $\mu_i<\mu_{r+j}$, $\forall i\in\mathbf{N}_1^r$ and $j\in\mathbf{N}_1^s$. Thus the Pfaffian (4.3), under these conditions, reduces to the determinant (4.2) which absorbs both alternating maps. The combination set $C_{2k}^{r+s}(\mathbf{N}_1^{r+s})$ now splits into separate combinations for the label subsets $\mathbf{N}_1^r$ and $\mathbf{N}_{r+1}^{r+s}$. Relabeling the indices μ_{r+k} by ν_k and the polyvector labels i_{r+k} by j_k finally gives the Clifford product of the basis polyvectors:

$$\mathbf{e}_{[i_1,\ldots,i_r]}\mathbf{e}_{[j_1,\ldots j_s]}$$

$$= \sum_{k=0}^{\min(r,s)} (-1)^{(1/2)k(2r-k-1)} \sum_{\mu\in C_k^r(\mathbf{N}_1^r)} \sigma(\mu)$$

$$\times \sum_{\nu\in C_k^s(\mathbf{N}_1^s)} \sigma(\nu)(|g_{i_{\mu_1} j_{\nu_1}}\cdots g_{i_{\mu_k} j_{\nu_k}}|$$

$$\times \mathbf{e}_{[i_{\mu_{k+1}},\ldots,i_{\mu_r}]}\wedge\mathbf{e}_{[j_{\nu_{k+1}},\ldots j_{\nu_s}]}).$$

Note that for $k>\min(r,s)$, one of the k metric tensors must have both index labels from one antisymmetrized set and all such terms are excluded. Thus the proof of the theorem is complete.

This theorem on the structure of the polyvector product is an explicit statement of the semigradation given by Hestenes[10]:

$$\mathbf{R}_n^{(r)}\mathbf{R}_n^{(s)}\subset\sum_{k=0}^{\min(r,s)}\mathbf{R}_n^{(r+s-2k)}.$$

The smallest polyvector part of this semigradation has valence $|r-s|$ and has been called the generalized interior product by Hestenes. For $s>r$ this is

$$(\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_r)\cdot(\mathbf{b}_1\wedge\cdots\wedge\mathbf{b}_s)$$

$$=(-1)^{r(r-1)/2}$$

$$\times\sum_{\nu\in C_r^s(\mathbf{N}_1^s)}\sigma(\nu)|\mathbf{a}_1\cdot\mathbf{b}_{\nu_1},\ldots,\mathbf{a}_r\cdot\mathbf{b}_{\nu_r}|\mathbf{b}_{\nu_{r+1}}\wedge\cdots\wedge\mathbf{b}_{\nu_s}.$$

The commutation law for this product is $A^{(r)}\cdot B^{(s)}=(-1)^{rs+\min(r,s)}B^{(s)}\cdot A^{(r)}$.

The largest polyvector or exterior part of the product can be selected from (4.9) in the following way:

$$\frac{r!s!}{(r+s)!}\sum_{\mu\in C_r^{r+s}(\mathbf{N}_1^{r+s})}\sigma(\mu)(\mathbf{a}_{\mu_1}\wedge\cdots\wedge\mathbf{a}_{\mu_r})$$

$$\times(\mathbf{a}_{\mu_{r+1}}\wedge\cdots\wedge\mathbf{a}_{\mu_{r+s}})=\mathbf{a}_1\wedge\cdots\wedge\mathbf{a}_{r+s}.$$

This, of course, is a statement of the Laplace expansion of a determinant, resulting from (2.5′), and emphasizes the fact that we have assumed that the polyvector basis forms an exterior algebra under the alternating mapping. When expanded over an orthogonal basis, this statement becomes almost trivial in the Clifford algebra, but this does not diminish the content of (4.9) for the definition of the basis polyvector (2.5) that did not assume the basis vectors were orthogonal. We also look at a part of (4.6); in particular, the scalar part. Denoting the function that selects this grade by sp( ) we see immediately that $\mathrm{sp}(\mathbf{e}_{i_1}\mathbf{e}_{i_2}\cdots\mathbf{e}_{i_k})=T_{i_1 i_2\cdots i_k}$, where $T_{i_1 i_2\cdots i_k}$, is defined by (4.4). Thus the properties of the scalar part of an even-versor correspond to the properties of the Pfaffian. This generalizes the properties of the trace operation in the Dirac algebra[14] to an arbitrary Clifford algebra.

Salingaros[15] has considered generating the Clifford algebra using (4.6) and (4.9) as the definition of the Clifford product. Unfortunately, his form of (4.9) contained an incorrect sign factor, presumably because it was not derived from the more fundamental (4.6). By the universality of the Clifford algebra it is sufficient to verify properties (2.1)–(2.5), for (4.6), to prove that it defines the Clifford product, and this is the path followed by Salingaros. We can immediately see that, except for associativity, these properties are easily satisfied. The contraction property (2.4) follows from the use of the Pfaffian contraction tensor, while (2.5b) is guaranteed by the construction of the versor over the polyvector basis.

But it is the property of associativity (2.2) that is most revealing. This property is not trivial as suggested by Salingaros, and thus the proof that (4.6) is a definition of the Clifford algebra was not completed. Although it is not necessary, we could go on to verify that (4.6) generates an associative product. Such a proof involves a Pfaffian expansion analogous to the Laplace expansion of a determinant in terms of complementary minors, given by Caianiello.[13] It involves Pfaffians and determinants of different orders and ranks with the largest determinant part being displayed in (4.9). To complicate matters further, the associativity of the product of versors contains many Pfaffian expansions because of the $\mathbf{Z}_2$-graded structure. In the analogous case of associativity for the exterior algebra, a similar expansion relies upon only one determinant and so we are led directly to the Laplace expansion. In the Clifford algebra the complexity of the situation warrants a more detailed analysis via the tensor representation. This is presented in a separate paper[1] which deals mainly with the associativity of the Clifford algebra.

ACKNOWLEDGMENTS

The author wishes to thank Dr. L. R. Dodd for kindling an interest in Clifford algebras and to Professor J. S. R. Chisholm and Dr. M. G. Eastwood for invaluable suggestions and encouragement.

Also the assistance of a Commonwealth Postgraduate Scholarship is gratefully acknowledged.

[1] G. Wilmot, "Clifford associativity and the Pfaffian expansion," J. Math. Phys. **29**, 2346 (1988).

[2] P. Raševskiĭ, Am. Math. Soc. Trans. Ser. 2 **6**, 1 (1957).

[3] I. R. Porteous, *Topological Geometry* (Van Nostrand–Reinhold, London, 1969).

[4] J. G. F. Belinfante and B. Kolman, *A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods* (SIAM, Philadelphia, 1972).

[5] C. Chevalley, *Theory of Lie Groups I* (Princeton U.P., Princeton, NJ, 1946).

[6] N. Jacobson, *Lie Algebras* (Interscience, New York, 1962).

[7] Z. Wan, *Lie Algebras* (Pergamon, Hungary, 1975).

[8] L. Corwin, Y. Ne'eman, and S. Sternberg, "Graded Lie algebras in mathematics and physics (Bose–Fermi symmetry)," Rev. Mod. Phys. **47**, 573 (1975).

[9] K. R. Greider, "A unifying Clifford algebra formalism for relativistic fields," Found. Phys. **14**, 467 (1984).

[10] D. Hestenes, *Space-Time Algebra* (Gordon and Breach, New York, 1966).

[11] T. Muir, *The Theory of Determinants* (Macmillan, London, 1923), Vol. iv, p. 264.

[12] H. S. Green and C. A. Hurst, *Order–Disorder Phenomena* (Interscience, London, 1964).

[13] E. R. Caianiello, *Combinatorics and Renormalisation in Quantum Field Theory* (Benjamin, Reading, MA, 1973).

[14] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Relativistic Quantum Theory*, Course of Theoretical Physics (Pergamon, Oxford, 1971), Vol. 4, Part 1.

[15] N. Salingaros, "Physical algebras in four dimensions. 1. The Clifford algebra in Minkowski spacetime," Adv. Appl. Math. **4**, 1 (1983).

Clifford associativity and the Pfaffian expansion

G. P. Wilmot
Department of Mathematical Physics, University of Adelaide, GPO Box 498, Adelaide, 5001, Australia

(Received 29 December 1987; accepted for publication 20 July 1988)

With the help of the Pfaffian, an operator on the tensor algebra which generates the explicit coset structure of the Clifford quotient of the tensor algebra is defined. The associativity property of this tensor representation of Clifford algebra leads to a Pfaffian expansion analogous to the Laplace expansion of a determinant. Further, the explicit statement of the algebra norm provides a proof for the Hadamard theorem.

I. INTRODUCTION

The connection between the Grassmann algebra and the determinant is a fundamental result of the theory of exterior forms. It is an elementary but powerful theorem since it exposes the properties of the determinant from properties of the Grassmann product. The property of interest here is that of associativity, which leads directly to the Laplace expansion of a determinant.

The Clifford algebra has not been as popular as the exterior algebra in the mathematical literature, and yet it is an ideal formalism for geometric operations. The Pfaffian also has been overshadowed by the determinant and indeed was originally perceived as a special determinant. In fact, this situation can be reversed with the determinant being a special case of the Pfaffian, and so arguments as to which one is more fundamental are futile.

The properties of the Pfaffian are presented in a chronological survey of determinants by Muir.[1] The simplest is that the Pfaffian is characterized as being the square root of an antisymmetric determinant.[2] The connection between the Grassmann and Clifford algebras has been studied by Chevalley[3,4] and Raševskiĭ[5] in a restricted form, using an orthonormal basis. More generally, the fundamental theorem giving the Pfaffian as the functional relation between the two algebras is stated by Caianiello[6] and has been previously proved by the author.[7] This result is necessary for a proof of the uniqueness of the operator that generates the Clifford algebra from the tensor algebra.

After introducing the Grassmann and Clifford algebras on the tensor space, we prove that the property of associativity of the Clifford algebra leads to an expansion of the Pfaffian analogous to the Laplace expansion of a determinant in terms of complementary minors. Because of the gradation of the Clifford algebra over the tensor space, Clifford associativity requires that an arbitrary Clifford element contain combinations of Pfaffian expansions for each grade, apart from the lowest grade.

The following notation for certain subsets of the symmetric set S^r will be adopted from Ref. 7. Members of these sets will be denoted by Greek letters and will be considered to be bijections of the natural number sequence $\mathbf{N}_1^r = (1,2,...,r)$. Note that $S^r(\mathbf{N}_p^{p+r})$ denotes all permutations of the labels $p,p+1,...,p+r \subset \mathbf{N}$. For $\mu \in S^r$, μ has sign $\sigma(\mu)$ determined by the parity of μ with action $\mu(i) = \mu_i$, $i \in \mathbf{N}_1^r$. The combination set $C_p^r \subset S^r$ is defined by the ordering restriction $\nu_1 < \nu_2 < \cdots < \nu_p$ for $\nu \in C_p^r$ and the partition set $P_{p,q}^r$ may be defined with extra condition $\nu_{p+1} < \cdots < \nu_{p+q}$ for $\nu \in P_{p,q}^r$ or as the composition of C_p^r and $C_q^{r-p}(\mathbf{N}_{p+1}^r)$.

The normal ordered set[8] is composed with the combination set to give the complete Pfaffian set with members $\rho \in N_{2m}^r \subset S^r$, which satisfy (i) for any $k \in \mathbf{N}_1^r$, $\rho_{2k-1} < \rho_{2k}$ and (ii) for any $h,k \in \mathbf{N}_1^r$, $h < k \Rightarrow \rho_{2h-1} < \rho_{2k-1}$. The normal ordered set N_{2m}^{2m}, or simply N^{2m}, defines a Pfaffian, since the set describes all possible ways of pairing the labels $\mathbf{N}_1^{2m}$. All sets carry the parity defined in the symmetric set given by the sign of $\sigma(\rho)$.

The determinant and the Pfaffian are now defined using the above sets. For matrices with components a_{ij} and $b_{k,l} = -b_{l,k,}$ $i,j \in \mathbf{N}_1^n$, $k,l \in \mathbf{N}_1^{2m}$, the determinant of rank n and Pfaffian of order r are defined by Porteous[8] as

$$|a_{11},a_{22},...,a_{nn}| = \sum_{\pi \in S^n} \sigma(\pi) \prod_{k \in \mathbf{N}_1^n} a_{k\pi_k},$$

$$\backslash b_{12},b_{23},...,b_{2r-1\,2r}| = \sum_{\rho \in N_r^{2r}} \sigma(\rho) \prod_{k \in \mathbf{N}_1^r} b_{\rho_{2k-1}\rho_{2k}}.$$

II. THE TENSOR AND EXTERIOR ALGEBRAS

The tensor algebra $\mathscr{T}(V) = \otimes V$ is a graded, associative, and infinite-dimensional algebra consisting of the multilinear forms over the vector space V. The grading of the algebra is denoted by $\mathscr{T}^r(V) = \otimes^r V$, so that $\mathscr{T}(V) = \Sigma_{r=0}^{\infty} \mathscr{T}^r(V)$. A multilinear form belonging to the subspace $\mathscr{T}^r(V)$ is said to have degree r and, if simple, is denoted by

$$v_{1,...,r} = v_1 \otimes \cdots \otimes v_r \in \otimes^r V,$$

where

$$v_i \in V, \quad i \in \mathbf{N}_1^r.$$

Note that if $r = 0$, then $v \in \mathscr{T}^0(V) \cong \mathbf{F}$, the field over which V is developed, and $\mathscr{T}^1(V) \cong V$, so that the natural map $V \rightarrow \mathscr{T}(V)$ is an embedding.

If the vector space V is a nondegenerate inner product space with inner product denoted by $\langle\ ,\ \rangle$, then this extends to an inner product for the algebra:

$$\langle v_{1,...,r}, w_{1,...,r} \rangle = \prod_{i=1}^{r} \langle v_i, w_i \rangle,$$

where

$$w_{1,...,r} = w_1 \otimes \cdots \otimes w_r. \tag{1}$$

This can be used to identify the dual algebra $(\otimes V)^* \equiv \otimes V^* = \mathscr{T}(V^*)$, where V^* is the dual vector space of V. The dual

0022-2488/88/112346-05$02.50

vectors are denoted with an asterisk or by raising the suffix, and their action on V is given in terms of the inner product by $v^*(w) = \langle v,w \rangle$, $\forall v^* \in V^*$ and $w \in V$. For V of finite dimension n, and with signature p, q, $n = p + q$, we can construct an orthonormal basis $e_1,...,e_n$ with dual basis $e^1,...,e^n$, such that $e^i(e_j) = \langle e_i,e_j \rangle = \delta_{ij}$ and $e^i(e_k) = \langle e_i,e_k \rangle = -\delta_{ik}$, for $i \in N_1^n, j \in N_1^p$, and $k \in N_{p+1}^n$.

The action of the dual vectors becomes multiplication in the Clifford algebra and is referred to as the contraction property of the algebra. The construction of the Clifford quotient algebra of the tensor algebra relies upon the definition of a graded operator Γ^k called the Pfaffian contraction operator of order k and is defined by its action on tensor forms as

$$\Gamma^m v_{1,...,r} = \begin{cases} \sum\limits_{\mu \in C_{2m}^r} \sigma(\mu) \backslash v^{\mu_1}(v_{\mu_2}), v^{\mu_2}(v_{\mu_3}),...,v^{\mu_{2m-1}}(v_{\mu_{2m}}) | v_{\mu_{2m+1},...,\mu_r}, & 2m \leqslant r; \qquad (2) \\ 0, & 2m > r. \qquad (3) \end{cases}$$

Note that this Pfaffian is the square root of $|b_{ij}|$, where $b_{ij} = -b_{ji} = \langle v_i,v_j \rangle$. For any combination C_{2m}^r, the ordering of the Pfaffian indices $\mu_1 < \mu_2 \cdots < \mu_{2m}$ can always be achieved by the properties of the Pfaffian[1] with symmetric components and, in the Clifford algebra, the ordering of the tensor indices determines the sign of $\sigma(\mu)$. Explicitly, the sign term in (2) is $\sigma(\mu) = (-1)^{t - m(2m-1)}$, where $t = \Sigma_{j=1}^{2m} \mu_j$.

Of course, the contraction operator of zero degree is taken to be the identity, $\Gamma^0 \equiv 1$. In general, $\Gamma^m v_{1,...,r}$ is multilinear, so that it is a tensor of degree $r - 2m$, but it is not simple. For example, the order-1 contraction in (2) gives

$$\Gamma v_{1,...,r} = \sum_{i=1}^{r-1} \sum_{j=i+1}^{r} (-1)^{j-i-1} v^i(v_j) v_{1,...,\hat{i},...,\hat{j},...,r}, \qquad (4)$$

where the indices denoted with a circumflex are omitted. Fundamental to the structure of the tensor representation of Clifford algebra is the grading of the contraction operator presented in the following theorem. With an inductive argument using (4), we find that the Pfaffian operator produces contractions between each pair of vectors, the number of pairs being given by the order. Hence the operator $\mathrm{Com} \equiv \Sigma_{m=0}^{\infty} \Gamma^m$ is called the complete contraction operator.

Theorem 1:

$\Gamma^m \cdot \Gamma^n \equiv [(m+n)!/m!n!] \Gamma^{m+n}$.

Proof:

$$\begin{aligned} \Gamma^n(\Gamma^m v_{1,...,r}) &= \sum_{\mu \in C_{2m}^r} \sigma(\mu) \sum_{\nu \in C_{2n}^{r-2m}} \sigma(\nu) \\ &\quad \times \backslash v^{\mu_1}(v_{\mu_2}),...,v^{\mu_{2m-1}}(v_{\mu_{2m}}) | \\ &\quad \times \backslash v^{\mu_{\nu_1}}(v_{\mu_{\nu_2}}),...,v^{\mu_{\nu_{2n-1}}}(v_{\mu_{\nu_{2n}}}) | \\ &\quad \times v_{\mu_{2(m+n)+1},...,\mu_r} \\ &= \sum_{\mu \in P_{2m,2n}^r} \sigma(\mu) \backslash v^{\mu_1}(v_{\mu_2}),...,v^{\mu_{2m-1}}(v_{\mu_{2m}}) | \\ &\quad \times \backslash v^{\mu_{2m+1}}(v_{\mu_{2m+2}}),...,v^{\mu_{2(m+n)-1}}(v_{\mu_{2(m+n)}}) | \\ &\quad \times v_{\mu_{2(m+n)+1},...,\mu_r} \\ &= \sum_{\mu \in N_{2m,2n}^r} \sigma(\mu) v^{\mu_1}(v_{\mu_2}) \cdots v^{\mu_{2(m+n)-1}} \\ &\quad \times (v_{\mu_{2(m+n)}}) v_{\mu_{2(m+n)+1},...,\mu_r}, \end{aligned}$$

where $\mu \in N_{2m,2n}^r \subset S^r$ satisfies $\mu_{2i-1} < \mu_{2i}$, for $i \in \mathbf{N}_1^m$ or $i \in \mathbf{N}_{m+1}^{m+n}$, $\mu_1 < \mu_3 < \cdots < \mu_{2m-1}$, and $\mu_{2m+1} < \mu_{2m+3} < \cdots < \mu_{2(m+n)-1}$. To obtain $\mu \in N_{2(m+n)}^r$ we require that $\mu_{2i-1} < \mu_{2j-1}$, for $i \in \mathbf{N}_1^m$ and $j \in \mathbf{N}_{m+1}^{m+n}$. But from the properties of the Pfaffian,[1] exchanging μ_{2i-1} and μ_{2j-1} is equivalent to exchanging μ_{2i} and μ_{2j}, for $i \in \mathbf{N}_1^m$ for $j \in \mathbf{N}_{m+1}^{m+n}$. This is already included in $\mathbf{N}_{2(m+n)}^r$. Hence removing the restriction $\mu_{2i-1} < \mu_{2j-1}$ in the normal ordered set $N^{2(m+n)}$, gives multiple copies of $N^{2(m+n)}$ corresponding to even parity reordering of members of $N_{2m,2n}^{2(m+n)}$. The factor of multiplication is given by C_m^{m+n}, since we are partitioning the secondary ordered indices into two parts. This completes the proof of the theorem.

The contraction property is only one-half of the multiplication by a vector in the Clifford algebra. The remaining part generates the exterior algebra and so a brief consideration of the construction of the Grassmann cosets in the tensor algebra will introduce another operator necessary for the construction of the Clifford cosets. Both constructions are analogous so that this approach will summarize the method used in the Clifford algebra case.

The Grassmann or exterior algebra $\mathscr{E}(V)$, consisting of all antisymmetric tensor forms, is obtained from the tensor algebra by taking the quotient of $\mathscr{T}(V)$ with the ideal generated by the symmetric dyads, $\mathscr{E}(V) \cong \mathscr{T}(V)/I$, where $I = \{x \otimes x | x \in V\}$. This is a graded and associative algebra which is finite dimensional for V finite and, in such a case, the tensor algebra must be filtered. The ideal may be removed from the cosets $x + I$, $\forall x \in \mathscr{E}(V)$, by antisymmetrizing the graded tensor space using the alternation operator, Alt. This is defined in terms of the antisymmetrization operator Λ^k where $\Lambda^0 1 \equiv 1$, and whose action more generally for a tensor form of degree l is $\Lambda^l v_{1,...,l} = (1/l!) \Sigma_{\mu \in S^l} \sigma(\mu) v_{\mu_1,...,\mu_l}$. Hence by definition,

$$v_{[1,...,l]} = \Lambda^l v_{1,...,l} = \sum_{i_1,...,i_l} |v_1^{i_1},...,v_l^{i_l}| e_{[i_1,...,i_l]}, \qquad (5)$$

where e_i, $i \in \mathbf{N}_1^n$, is a basis of V. This is called an l-form and belongs to $\mathscr{E}^l(V) \cong \mathscr{T}^l(V)/I$. If the action of Λ^l on tensors of degree not equal to l is taken to be annihilation, then we can define the alternation operator to be $\mathrm{Alt} \equiv \Sigma_{l=0}^{\infty} \Lambda^l$. This operator is a mapping from the tensor algebra to the tensor representation of the exterior algebra. We now derive the relation of the Pfaffian to the determinant using Γ.

Lemma 1:

$$\Gamma^l(v_{[1,\ldots,l\,]} \otimes w_{[1,\ldots,l\,]}) = (-1)^{[1/2]}|v^1(w_1),\ldots,v^l(w_l)|,$$

where $[x]$ is the integer part of x.

Proof: Using Theorem 1 and then (4) we have

$$\begin{aligned}
&\Gamma^l(v_{[1,\ldots,l\,]} \otimes w_{[1,\ldots,l\,]})\\
&\quad = \frac{1}{l}\Gamma^{l-1}\Bigg(\sum_{i=1}^{l}\sum_{j=1}^{l}(-1)^{j-i-1+l}v^i(w_j)\\
&\qquad \times v_{[1,\ldots,\hat{i},\ldots,l\,]} \otimes w_{[1,\ldots,\hat{j},\ldots,l\,]}\Bigg)\\
&\quad = \Gamma^{l-1}\Bigg((-1)^{l-1}\frac{1}{l!(l-1)!}\sum_{\mu\in S^l}\sigma(\mu)\\
&\qquad \times \sum_{\nu\in S^l}\sigma(\nu)v^{\mu_1}(w_{\nu_1})v_{\mu_2,\ldots,\mu_l} \otimes w_{\nu_2,\ldots,\nu_l}\Bigg).
\end{aligned}$$

Continuing this reduction using (4) we arrive at the statement of the lemma:

$$\begin{aligned}
&\Gamma^l(v_{[1,\ldots,l\,]} \otimes w_{[1,\ldots,l\,]})\\
&\quad = (-1)^{(l-1)+(l-2)+\cdots+1}\frac{1}{l!}\sum_{\mu\in S_l}\sigma(\mu)\\
&\qquad \times \sum_{\nu\in S^l}\sigma(\nu)v^{\mu_1}(w_{\nu_1})v^{\mu_2}(w_{\nu_2})\cdots v^{\mu_l}(w_{\nu_l})\\
&\quad = (-1)^{l(l-1)/2}\sum_{\nu\in S^l}\sigma(\nu)v^1(w_{\nu_1})v^2(w_{\nu_2})\cdots v^l(w_{\nu_l})\\
&\quad = (-1)^{[l/2]}|v^1(w_1),v^2(w_2),\ldots,v^l(w_l)|.
\end{aligned}$$

It is a simple matter to prove associativity for the quotient algebra $\mathscr{T}(V)/I$ with the statement being $\mathrm{Alt}(v_{1,\ldots,k+l}) = \mathrm{Alt}(\mathrm{Alt}(v_{1,\ldots,k}) \otimes \mathrm{Alt}(v_{k+1,\ldots,k+l}))$. More explicitly, $v_{[1,\ldots,k+l\,]} = \Lambda^{k+l}(v_{[1,\ldots,k\,]} \otimes v_{[k+1,\ldots,k+l\,]})$ and this is a direct statement of the Laplace expansion of a determinant in terms of complementary minors, as a result of (5). The analogous statement for Clifford algebra involves the grading over the exterior basis of many Pfaffian contractions Γ^m, and so it is a statement of many Pfaffian expansions.

III. THE CLIFFORD ALGEBRA

Similar to the exterior algebra, the Clifford algebra is defined as the quotient of the tensor algebra with an ideal. In this case, we are dealing with an inner product vector space, and we denote this space and the inner product, which we assume to be nondegenerate, by $\mathbf{V}$. Then the Clifford algebra is defined as $\mathscr{C}(\mathbf{V}) \cong \mathscr{T}(\mathbf{V})/I(\mathbf{V})$, where the two-sided ideal $I(\mathbf{V})$ is generated by elements $x \otimes x - \langle x,x\rangle$, $\forall x \in \mathscr{T}(V)$. For a completely degenerate inner product, the Clifford algebra reduces to the exterior algebra.

The ideal $I(\mathbf{V})$ is homogeneous of even degree in the semigrading of the tensor algebra, and so $\mathscr{C}(\mathbf{V})$ is not $\mathbf{Z}$ graded, but $\mathbf{Z}_2$ graded. However, the exterior subspace of the tensor algebra is isomorphic (as a vector space, not as an algebra) to the Clifford algebra[8,9] under the mapping generated by the alternation operator Alt. In particular, there is a canonical form for tensor representations of Clifford elements, whereby the space is invariant under the action of Alt. The canonical form is gained by applying the contraction operator Com, to any tensor representation of a Clifford element, thus producing tensors $v_{1,\ldots,r}$, such that for $i, j \in \mathbf{N}_1^r$, not equal, $v^i(v_j) = 0$. Hence the ideal $I(\mathbf{V})$ reduces to the exterior ideal I, which may then be annihilated using Alt.

Composing the contraction and alternating operators produces an idempotent operator which gives a mapping from the tensor space to the canonical tensor space isomorphic to the exterior space. We call this operator the Clifford contraction operator, Con $\equiv$ Alt Com, and it only remains to prove that this maps onto the exterior space and is unique to show that it generates the Clifford algebra.

Lemma 2: The Clifford contraction operator, Con $\equiv$ Alt Com, generates the Clifford algebra.

Proof: Denoting the canonical form of a tensor belonging to $\mathscr{C}(\mathbf{V})$ by $v_{(1,\ldots,p)}$, called a versor, we expect

$$v_{(1,\ldots,p)} = \mathrm{Con}(v_{1,\ldots,p}),$$

where

$$\begin{aligned}
\mathrm{Con}(v_{1,\ldots,p}) = \sum_{k=0}^{[p/2]}\sum_{\mu\in C_{2k}^r}\sigma(\mu)\backslash v^{\mu_1}(v_{\mu_2}),v^{\mu_2}(v_{\mu_3}),\ldots,v^{\mu_{2k-1}}\\
\times(v_{\mu_{2k}})|v_{[\mu_{2k+1},\ldots,\mu_p]}\,. \qquad (6)
\end{aligned}$$

This explicit form of the exterior structure of a versor is proved by induction in a separate paper[7] and need not be repeated here. It extends the Jordan relation, $v_{(1,2)} + v_{(2,1)} = 2v^1(v_2)$, to the entire algebra and thus, by the universality theorem[3], the operator Con is the required generator.

The associativity of the Clifford algebra can now be stated as

$$\mathrm{Con}(v_{1,\ldots,r} \otimes w_{1,\ldots,s}) = \mathrm{Con}(\mathrm{Con}(v_{1,\ldots,r}) \otimes \mathrm{Con}(w_{1,\ldots,s}))\,.$$

This is the basis of the following theorem which presents the Pfaffian analogy to the Laplace expansion of a determinant. This theorem has been proved by combinatorial arguments by Caianiello.[10]

Theorem 2:

$$\begin{aligned}
&\backslash v^1(v_2),\ldots,v^{2k-1}(v_{2k})|\\
&\quad = \sum_{l'=0}^{\min([r/2],[k-r/2])}(-1)^{l'}\sum_{\mu\in C_{2m}^r}\sigma(\mu)\\
&\qquad \times \sum_{\nu\in C_{2n}^{2k-r}(N_{r+1}^{2k})}\sigma(\nu)\cdot\backslash v^{\mu_1}(v_{\mu_2}),\ldots,v^{\mu_{2m-1}}(v_{\mu_{2m}})|\\
&\qquad \times \backslash v^{\nu_1}(v_{\nu_2}),\ldots,v^{\nu_{2n-1}}(v_{\nu_{2n}})|\\
&\qquad \times |v^{\mu_{2m+1}}(v_{\nu_{2n+1}}),\ldots,v^{\mu_r}(v_{\nu_s})|\,,
\end{aligned}$$

where

$$m = ([r/2]-l'), \quad n = ([k-r/2]-l')\,.$$

Proof: More explicitly, associativity may be written as

$$\begin{aligned}
&\sum_{k=0}^{[(r+s)/2]}\Lambda^{r+s-2k}\Gamma^k(v_{1,\ldots,r} \otimes w_{1,\ldots,s})\\
&\quad = \mathrm{Con}\Bigg[\sum_{m=0}^{[r/2]}\Lambda^{r-2m}\Gamma^m(v_{1,\ldots,r})\\
&\qquad \otimes \sum_{n=0}^{[s/2]}\Lambda^{s-2n}\Gamma^n(w_{1,\ldots,s})\Bigg]\,. \qquad (7)
\end{aligned}$$

The term in square brackets is multiplied by $\mathrm{Com} = \Sigma_{l=0}^{\infty}\Gamma^l$, so the total order of the contraction operator on the right-hand side is $l + m + n$. This can be equated to the order k operator on the other side by substituting $k = l + m + n$ and replacing the summation over n by a summation over k. The last factor then becomes $\Sigma_{k=m+l}^{m+l+[s/2]} \times \Lambda^{s-2(k-m-l)}\Gamma^{k-m-l}(w_{1,\dots,s})$. Now the order of the summations over m and k are interchanged:

$$\sum_{m=0}^{[r/2]} \sum_{k=m+l}^{m+l+[s/2]} \equiv \sum_{k=l}^{[r/2]+[s/2]+l} \sum_{m=p}^{q},$$

where

$$p = \begin{cases} 0, & \text{for } k < l + [s/2], \\ k - l - [s/2], & \text{otherwise;} \end{cases}$$
$$q = \begin{cases} k - l, & \text{for } k < l + [r/2], \\ [r/2], & \text{otherwise.} \end{cases} \tag{8}$$

Also, nonzero terms in (7) require $2k < r + s$ so that we may take the upper limits of both the k and l summations to be $[(r+s)/2]$. Exchanging the order of these summations gives a graded statement of the right-hand side of (7). With p and q defined by (8), this is

$$\sum_{k=0}^{[(r+s)/2]} \Lambda^{r+s-2k}\Gamma^k(v_{1,\dots,r} \otimes w_{1,\dots,s})$$
$$= \sum_{k=0}^{[(r+s)/2]} \sum_{l=0}^{k} \sum_{m=p}^{q} \Gamma^l[(\Lambda^{r-2m}\Gamma^m(v_{1,\dots,r}))$$
$$\otimes (\Lambda^{s-2(k-m-l)}\Gamma^{k-m-l}(w_{1,\dots,s}))]\,. \tag{9}$$

For each k, this expresses many Pfaffian expansions as coefficients in the exterior subspace $\mathscr{E}^{r+s-2k}(\mathrm{V}) \cong \mathscr{T}^{r+s-2k}(\mathrm{V})/I$. Choosing the scalar part, with $2k = r + s$ in (9), produces a single Pfaffian on the left-hand side. Equating the scalar terms gives

$$\Gamma^k(v_{1,\dots,r} \otimes w_{1,\dots,s})$$
$$= \sum_{l=0}^{k} \sum_{m=p}^{q} \Gamma^l[\Lambda^{r-2m}\Gamma^m(v_{1,\dots,r})$$
$$\otimes (\Lambda^{s-2(k-m-l)}\Gamma^{k-m-l}(w_{1,\dots,s}))]\,.$$

Now, as a result of (3), $\Gamma^l[\Lambda^{r-2m}\Gamma^m(v_{1,\dots,r}) \otimes w_{[1,\dots,s]}] = 0$ for $l > \min(r,s)$ or $l + 2m > r$ and $\Gamma^l[v_{[1,\dots,r]} \otimes (\Lambda^{2(l+m)-r}\Gamma^{k-l-m}w_{1,\dots,s})] = 0$ for $l + 2(k-l-p) > s$ or $l + 2m < r$. Hence, for nonzero terms in (9), we require $l \leqslant \min(r,s)$ and $2m = r - l$. Denoting equivalence modulo 2 by $\sim$, we have $r \sim s \sim l$, which gives the following relations:

$$\frac{r-l}{2} = \left[\frac{r}{2}\right] - \left[\frac{l}{2}\right], \quad \frac{s-l}{2} = \left[\frac{s}{2}\right] - \left[\frac{l}{2}\right],$$
$$k = \frac{r+s}{2} = \left[\frac{r}{2}\right] + \left[\frac{s}{2}\right] + l - 2\left[\frac{l}{2}\right].$$

Using these expressions in (8), it is easy to check that $p \leqslant (r-l)/2 \leqslant q$, so that it is always possible to choose $m = (r-l)/2$, $\forall l \sim r$. Hence changing the summation over l and substituting $m = (r-l)/2$ gives

$$\Gamma^k(v_{1,\dots,r} \otimes w_{1,\dots,s}) = \sum_{\substack{l=r-2[r/2] \\ \text{step } 2}}^{\min(r,s)} \Gamma^l(\Lambda^l\Gamma^{(r-l)/2}(v_{1,\dots,r})$$
$$\otimes \Lambda^l\Gamma^{k-l-((r-1)/2)}(w_{1,\dots,s}))\,.$$

Changing the summation variable $l = l' + r - 2[r/2]$ to $l' = [l/2]$ by changing the summation step and substituting $m = [r/2] - l'$ and $n = [k - r/2] - l'$ finally gives the statement of the Pfaffian expansion:

$$\Gamma^k(v_{1,\dots,r} \otimes w_{1,\dots,s})$$
$$= \sum_{l'=0}^{\min([r/2],[k-r/2])} \Gamma^l(\Lambda^l\Gamma^m(v_{1,\dots,r}) \otimes \Lambda^l\Gamma^n(w_{1,\dots,s}))$$
$$= \sum_{l'=0}^{\min([r/2],[k-r/2])} \sum_{\mu\in C_{2m}^r} \sigma(\mu) \sum_{\nu\in C_{2n}^s} \sigma(\mu)(\Gamma^m(v_{1,\dots,2m})$$
$$\times \Gamma^n(w_{1,\dots,2n})\Gamma^l(v_{[\mu_{2m+1},\dots,\mu_r]} \otimes w_{[\nu_{2n+1},\dots,\nu_s]})),$$

where we have removed the combinatorial parts of Γ^m and Γ^n contained in definition (2). Applying this definition again and using Lemma 2 produces the final form of the Pfaffian expansion. With $w_{1,\dots,s} \equiv v_{r+1,\dots,r+s}$, this is the statement of the theorem, thus completing the proof.

Theorem 2 is a partitioning of the Pfaffian of order k giving an action on the first r indices and a separate action on the remaining $2k - r$ indices of the Pfaffian. These factors describe pairs of contractions whereby m pairs connect indices of $v_{(1,\dots,r)}$; n pairs connect indices of $v_{(r+1,\dots,2k)}$, and the remaining l pairs connect an index from each side of the partition. This theorem results in an explicit statement of the norm of a versor.

The norm of the Clifford algebra is defined in terms of the reversion involution $\check{v}$ which is an antiautomorphism of the algebra sending $v_{(1,\dots,p)} \mapsto v_{(p,\dots,1)}$. Denoting the projection of v on the polyvector subspace of valence s by $v^{(s)}$, reversion involves $s(s-1)/2$ exchanges of the s vector factors, equivalent modulo 2 to the integer of $s/2$: $\check{v}^{(s)} = (-1)^{[s/2]}v^{(s)}$.

The inner product for the tensor algebra (1) is

$$\langle u_{[1,\dots,r]}, v_{[1,\dots,r]}\rangle = \sum_{\mu\in S^r} \sigma(\mu) \sum_{\nu\in S^r} \sigma(\nu) \prod_{i\in \mathrm{N}_1^r} \langle u_{\mu_i}, v_{\nu_i}\rangle$$
$$= |u^1(v_1),\dots,u^r(v_r)|$$
$$= \langle u_{[1,\dots,r]}\check{v}_{[1,\dots,r]}, 1\rangle$$
$$= \langle 1, \check{u}_{[1,\dots,r]}v_{[1,\dots,r]}\rangle\,, \tag{10}$$

where Lemma 2 has been used.

Defining the Clifford algebra norm $\|v\| = \langle v,v\rangle$ and using linearity, the inner product polarizes to $\langle u,v\rangle = \langle (u\check{v} + v\check{u})/2, 1\rangle = \langle u\check{v}, 1\rangle$. For versors, this reduces to the algebraic expression $\|u\| = u\check{u}$. The norm for a polyvector is given by (10) as the square of the volume of the parallelepiped described by its vector factors. The maximum value for this volume squared corresponds to the norm of the versor formed from these vectors

$$\|v_{(1,\dots,p)}\| = v_{(1,\dots,p)}v_{(p,\dots,1)} = \prod_{i\in \mathrm{N}_1^p} \|v_i\|\,.$$

The versor expansion (6) provides, in the following theorem,[9] an explicit relation between the norms of these two elements.

Hadamard Theorem:

$$\|v_{(1,\dots,p)}\| = \sum_{k=0}^{[p/2]} \|v_{(1,\dots,p)}^{(p-2k)}\|\,,$$

where $v_{(1,\dots,p)}^{(p-2k)} = \Lambda^{p-2k}\Gamma^k v_{1,\dots,p}$ is the projection of $v_{1,\dots,p}$ on the polyvector subspaces of valence $p - 2k$.

Proof:

$$\begin{aligned}\|v_{(1,\ldots,p)}\| &= v_{(1,\ldots,p)}\check{v}_{(1,\ldots,p)}\\ &= \sum_{k=0}^{[p/2]}\sum_{l=0}^{[p/2]}\sum_{\mu\in C^{\wp}_{2k}}\sigma(\mu)\sum_{\nu\in C^{\wp}_{2l}}\sigma(\nu)\backslash v^{\mu_1}(v_{\mu_2}),\ldots,v^{u_{2k-1}}(v_{\mu_{2k}})|\backslash v^{\nu_1}(v_{\nu_2}),\ldots,v^{\nu_{2l-1}}(v_{\nu_{2l}})|v_{[\mu_{2k+1},\ldots,\mu_p]}\check{v}_{[\nu_{2l+1},\ldots,\nu_p]}\,. \end{aligned}\qquad (11)$$

For a scalar result we must choose $k = l$, and hence

$$\begin{aligned}\|v_{(1,\ldots,p)}\| &= \sum_{k=0}^{[p/2]}\Big\|\sum_{\mu\in C^{\wp}_{2k}}\sigma(\mu)\backslash v^{\mu_1}(v_{\mu_2}),\ldots,v^{\mu_{2k-1}}(v_{\mu_{2k}})|\cdot v_{[\mu_{2k+1},\ldots,\mu_p]}\Big\|\\ &= \sum_{k=0}^{[p/2]}\|v^{(p-2k)}_{(1,\ldots,p)}\|\,.\end{aligned}$$

Exchanging the polyvector norm in (11) with $v_{[\mu_m,\ldots,\mu_p]}\check{v}_{[\nu_m,\ldots,\nu_p]} = |v^{\mu_m}(v_{\nu_m}),\ldots,v^{\mu_p}(v_{\nu_p})|$, Theorem 2 identifies the analogous versor norm as

$$\|v_{(1,\ldots,p)}\| = \backslash v^1(v_1),v^1(v_2),v^2(v_2),\ldots,v^p(v_p)|\,,$$

since

$$\backslash v^1(v_2),\ldots,v^p(v_1),\ldots,v^{p-1}(v_p)| = (-1)^{[p/2]}\backslash v^1(v_2),\ldots,\nu^{p-1}(v_p),v^p(v_{p-1}),\ldots,v^2(v_1)|\,,$$

from the work of Caianiello.[10]

Hence the Hadamard theorem is a special case of the Pfaffian partition expansion corresponding to a matrix which is antisymmetric about the usual diagonal and symmetric about the cross diagonal.

Finally, it is worthwhile giving an example of the Pfaffian expansion. Of course, partitioning just one index leads to the Pfaffian cofactor expansion. Partitioning the order-3 Pfaffian into halves gives the following "matrix" expansion:

$$\left|\begin{array}{ccccc} a_{12} & a_{13} & a_{14} & a_{15} & a_{16}\\ & a_{23} & a_{24} & a_{25} & a_{26}\\ & & a_{34} & a_{35} & a_{36}\\ & & & a_{45} & a_{46}\\ & & & & a_{56}\end{array}\right| = \sum_{\mu\in C^3_2(\mathbf{N}^3_1)}\sigma(\mu)\sum_{\nu\in C^3_2(\mathbf{N}^6_4)}\sigma(\nu)\backslash a_{\mu_1\mu_2}|\cdot\backslash a_{\nu_1\nu_2}|\cdot|a_{\mu_3\nu_3}|$$

$$- \sum_{\mu\in C^3_3(\mathbf{N}^3_1)}\sigma(\mu)\sum_{\nu\in C^3_3(\mathbf{N}^6_4)}\sigma(\nu)|a_{\mu_1\nu_1}a_{\mu_2\nu_2}a_{\mu_3\nu_3}|$$

$$= \begin{array}{l} a_{12}a_{45}a_{36} - a_{12}a_{46}a_{35} + a_{12}a_{56}a_{34}\\ -a_{13}a_{45}a_{26} + a_{13}a_{46}a_{25} - a_{13}a_{56}a_{24} -\\ +a_{23}a_{45}a_{16} - a_{23}a_{46}a_{15} + a_{23}a_{56}a_{14}\end{array}\left|\begin{array}{ccc} a_{14} & a_{15} & a_{16}\\ a_{24} & a_{25} & a_{26}\\ a_{34} & a_{35} & a_{36}\end{array}\right|\,.$$

ACKNOWLEDGMENTS

The author wishes to thank Dr. L. R. Dodd for assistance and supervision and Dr. M. G. Eastwood for making available private communications with R. Harvey.

Also, the assistance of a Commonwealth Postgraduate Scholarship is gratefully acknowledged.

---

[1] T. Muir, *The Theory of Determinants* (Macmillan, London, 1923), Vol. IV, p. 264.

[2] H. S. Green and C. A. Hurst, *Order–Disorder Phenomena* (Interscience, London, 1964).

[3] C. Chevalley, "The construction and study of certain important algebras," Math. Soc. Jpn. **1**, 33 (1955).

[4] C. Chevalley, *Theory of Lie Groups I* (Princeton U.P., Princeton, NJ, 1946).

[5] P. Raševskiĭ, "The theory of spinors," Amer. Math. Soc. Transl. Ser. 2 **6**, 1 (1957).

[6] E. R. Caianiello, *Combinatorics and Renormalization in Quantum Field Theory* (Benjamin, Reading, MA, 1973).

[7] G. Wilmot, "The structure of Clifford algebra," J. Math. Phys. **29**, 2338 (1988).

[8] I. R. Porteous, *Topological Geometry* (Van Nostrand–Reinhold, London, 1969).

[9] R. Harvey, *Spinors* (Academic, New York, in press).

[10] E. R. Caianiello, "On the algorithm of Dirac spurs," Nuovo Cimento **9**, 1218 (1952).

On biorthogonal and orthonormal Clebsch–Gordan coefficients of SU(3): Analytical and algebraic approaches

Sigitas Ališauskas
Institute of Physics, Academy of Sciences of the Lithuanian SSR, Vilnius, 232600, Union of Soviet Socialist Republics

(Received 2 January 1988; accepted for publication 25 May 1988)

Minimal biorthogonal systems of the Clebsch–Gordan (Wigner) coefficients of SU(3) ⊃U(2) are discussed as well as the dual coupled bases. The closed system of analytical expressions for the dual isofactors (reduced Wigner coefficients) and the overlaps of coupled states is obtained with the help of analytical inversion symmetry. The Regge-type symmetry of the overlaps and the boundary orthonormal isofactors (orthogonalization coefficients) is discovered. The polynomial structure of the alternative complete algebraic systems of the orthonormal SU(3) isofactors (characterized by the null spaces, symmetries, and additional selection rules and obtained by means of the Hecht or Gram–Schmidt process) is considered. The realizations of the external "missing label" operators of the third and the fourth orders in the minimal coupled bases, which lead to preferable algorithms to evaluate the orthonormal SU(3) coupling coefficients satisfying different symmetry properties, are presented. With the help of the $6j$ coefficients of SU(2) or inverted truncated SU(2) recoupling matrices, the biorthogonal systems associated with the SU(3) canonical tensor operators are expanded in terms of minimal ones.

I. INTRODUCTION

The missing label problem for SU(3) ×SU(3) ⊃SU(3) and the theory of the non-multiplicity-free Clebsch–Gordan (Wigner) coefficients of SU(3) ⊃U(2) have a long history beginning immediately after the first applications of those coefficients in the nuclear and particle physics. Three main directions of this theory are known that are grounded (i) on the construction of complete nonorthonormal systems, (ii) on the use of the SU(3) invariant classifying operators, or (iii) on the canonical unit tensor operators, respectively.

The purpose of this paper is to discuss from a uniform viewpoint different solutions of the SU(3) outer multiplicity problem, which allow one finally to obtain the numerical algorithms and analytical or algebraic expressions for the coupling coefficients of SU(3) in the most convenient forms.

General analytical expressions for nonorthogonal Clebsch–Gordan (CG) coefficients of SU(3) have been constructed by means of different methods, including the recursive-recoupling techniques,[1–6] the use of different generating invariants,[7–9] the projection operators,[10–13] or the integration over the group.[14]

The most convenient expressions for corresponding nonorthogonal isofactors (reduced Clebsch–Gordan–Wigner coefficients) of SU(3) ⊃U(2) [with the minimal number of sums (six), some of them being included in the standard multiplicity-free functions of the SU(2)- or SU(3)-Wigner–Racah calculus] form the analytical biorthogonal systems.[5,15] The use of the biorthogonal systems allows one to simplify considerably many operations of the Wigner–Racah algebra. The usage of the biorthogonal systems is particularly preferable from the computational point of view in all the cases when the summation over the multiplicity labels of the irreducible representations (irreps) takes place.

General properties of the biorthogonal systems of noncanonical bases and coupling coefficients are discussed in Refs. 16 and 17. As a rule, the biorthogonal system is formed by dual isofactors represented by specific bilinear combinations[6,16] of the orthonormal isofactors (i.e., by special matrix elements of the projection operators[10–13]) and by analytical solutions of some discrete boundary value problems[6,16,18] (i.e., by the explicit integrals of the recursion formulas[19]). Due to the one-to-one correspondence between the parameters of the biorthogonal systems of the isofactors (or, respectively, of the coupled states) and the Weyl (and Littlewood–Richardson) direct product decomposition rules, the labeling of the minimal analytical biorthogonal systems of the SU(3) isofactors[6] presents alternatives to the canonical labeling scheme introduced by Biedenharn, Louck, and collaborators.[20,21] Particularly, the Gram–Schmidt producers applied to the minimal biorthogonal systems lead in a simple way to the orthogonal system of isofactors introduced long ago by Hecht[19] (see also Refs. 5 and 14), some properties of which have been discussed recently by Le Blanc and Rowe.[22] The results of Refs. 5, 6, and 15 necessary for our investigation (the review of which is given in Sec. II of our paper) allowed us to reveal the additional symmetry properties of the introduced paracanonical orthonormal isofactors that reduce the supposed arbitrariness of the chosen labeling scheme considerably. Thus in Sec. II, some important universal constructive elements for the SU(3) Wigner–Racah calculus are presented.

The closed final form of the minimal biorthogonal systems is achieved in Sec. III, where the analytical inversion symmetry[17,23] of the biorthogonal systems allowed us to obtain new expressions for isofactors and overlaps of the nonorthogonal coupled states. In Sec. IV, in a rather simple way, the polynomial and other properties of the paracanonical isofactors are established that have some analogy with and

0022-2488/88/112351-16$02.50

differ from the properties of the canonical SU(3) tensor operators.[21] In general, the paracanonical SU(3) operators as the basis for the Wigner–Racah algebra may be found with less effort than the canonical ones. The null-space properties,[22] symmetries, and the additional selection rules of the isofactors are different as well as the reduction formulas (more exactly, the Regge-type symmetries of the boundary isofactors) and the structures of the denominator functions (cf. Ref. 21). The alternative pseudocanonical systems of orthonormal isofactors that imitate the null-space structure of the canonical tensor operators are also discussed in Sec. IV.

The minimal biorthogonal systems are also very important for calculating the uniquely labeled orthonormal isofactors of SU(3) in the case of other splitting schemes. The usage of classifying operators may be preferable from the physical point of view. The external labeling operator of the third order for SU(3) has been proposed by Hecht.[19] [The proof of equivalence of this operator with that proposed by Moshinsky[24] acting in the complementary group $U(4) \supset U(2)\times U(2)$ space is not trivial.] The realizations of this operator in different nonminimal coupled bases has been proposed in Refs. 4, 8, 12, 13, 18, 25, and 26. Particularly in Refs. 12, 13, 18, and 26, some auxiliary coupled bases of SU(3), in fact, have been expanded in terms of the minimal ones.[27]

Ališauskas and Kulish[28] have demonstrated that the external labeling operator of the fourth order, suggested by Sharp,[29] is also indispensable for the spectral resolution of the SU(3)-invariant solutions of the Yang–Baxter equation. They found the matrix elements of both external labeling operators in the minimal coupled bases.

The symmetry properties of the isofactors labeled by the proper values of the classifying operator of the fourth order better satisfy the pattern of Derome[30] than the case of the operator of the third order.[12,13,18] The minimal algorithms of evaluation of the boundary values of isofactors for both classification schemes are given in Sec. V of this paper, as well as the representation of the $SU(3)\times SU(3)\supset SU(3)$ generators and some tensor operators. Particularly for low multiplicities, the expressions in the algebraic-polynomial form are possible.

In spite of the considerable aesthetic fascination, the problems remain in the general explicit algebraic construction of the canonical SU(3) tensor operators.[20,21,31–33] The recursive-numerical algorithm for the SU(3) isofactors corresponding to this external classification scheme has been given in Ref. 34, while in Ref. 21 the generating function technique is used for the algebraic construction. There was a rather strange situation concerning the absence of connection between the SU(3) canonical tensors and the analytical systems of SU(3) isofactors.

In Sec. VI of this paper, a biorthogonal system is introduced associated with canonical SU(3) tensor operators, which is expanded in terms of the minimal biorthogonal system. This expansion corresponds to the Weyl transformation between different SU(2) subgroups in SU(3). The canonical splitting of the multiplicity similarly to the paracanonical one is specified precisely by the proper sequence of the Schmidt process, which may be performed explicitly because for any fixed SU(3) tensor operator the overlaps of the initially nonorthogonal coupled states take algebraic-polynomial form. The algebraic-polynomial structure of the explicit expressions for orthonormal isofactors [i.e., the Biedenharn–Louck canonical U(3): U(2) projective operators] is also ensured.

In this paper, the problem of polynomial representation of analytical expressions for isofactors and overlaps (norms) is emphasized, taking into account both the possibility of the algebraic computer based calculations[35] and the problems that appear in the case of large values of some parameters. Analytical expressions usually are represented as factorial sums. For fixed values of all summation intervals [the Regge–Bargmann–Shelepin (RBS) parameters of the type (a)[17]] they turn into elementary algebraic-polynomial functions of the remaining (free) parameters or, in particular, are summed up. Contrary to the SU(2) case for which all the Regge or Bargmann–Shelepin parameters of the Clebsch–Gordan or Racah coefficients may belong to the type (a) (see Ref. 36, Secs. 13 and 29), there are no universal, in this respect, expansions for general isofactors of SU(3). Otherwise, the alternative approaches allow us to find solutions most convenient for concrete aims. For example, different expressions of Ref. 37 (see also Ref. 16) exhaust all the possibilities to choose the sets of the RBS parameters of the type (a) in the case of the multiplicity-free SU(n) isofactors for coupling $\lambda \times p$ (where λ is an arbitrary irrep and p is a symmetric one).[38]

The unitary irreps of SU(3) will be denoted below as mixed tensor irreps (ab), where $a=m_{13}-m_{23}$, $b=m_{23}-m_{33}$, and $[m_{13}m_{23}m_{33}]$ is the Young scheme. The group generators $E_{\rho\sigma}$ $(\rho,\sigma=1,2,3)$ satisfy the usual commutation relations,

$$[E_{\rho\sigma},E_{\rho'\sigma'}]=\delta_{\rho'\sigma}E_{\rho\sigma'}-\delta_{\rho\sigma'}E_{\rho'\sigma}. \tag{1.1}$$

The basis states are labeled by the hypercharge $y=m_{12}+m_{22}-\frac{2}{3}(m_{13}+m_{23}+m_{33})$, the isospin $i=\frac{1}{2}(m_{12}-m_{22})$, and its projection $i_z=m_{11}-\frac{1}{2}(m_{12}+m_{22})$, where the integers m_{ij} form the Gelfand–Tzetlin pattern. Frequently the parameter

$$z=\tfrac{1}{3}(b-a)-\tfrac{1}{2}y=m_{23}-\tfrac{1}{2}(m_{12}+m_{22}) \tag{1.2}$$

is more convenient than y because the linear combinations

$$i\pm z,\quad a+z-i,\quad b-z-i, \tag{1.3a}$$

are non-negative integers. In the case of the coupling $(a'b')\times(a''b'')$ to (ab)

$$z=z'+z''+v, \tag{1.3b}$$

where

$$v=\tfrac{1}{3}(a'-b'+a''-b''-a+b) \tag{1.3c}$$

is an integer. The parameters of the highest weight state (HWS) take the values

$$y_0=\tfrac{1}{3}(a+2b),\quad i_0=\tfrac{1}{2}a=-z_0, \tag{1.4a}$$

while for the lowest weight state (LWS)

$$\bar{y}_0=-\tfrac{1}{3}(2a+b),\quad \bar{i}_0=\tfrac{1}{2}b=\bar{z}_0, \tag{1.4b}$$

and for the maximal isospin state (MIS)

$$y_m = \tfrac{1}{3}(a-b), \quad i_m = \tfrac{1}{2}(a+b), \quad z_m = \tfrac{1}{2}(b-a). \qquad (1.4c)$$

The conditions of the biorthogonality of isofactors

$$\sum_{\mu',\mu''} (\lambda'\mu'\lambda''\mu''\|\lambda\mu)^{\rho}(\lambda'\mu'\lambda''\mu''\|\underline{\lambda}\mu)_{\varrho} = \delta_{\rho\varrho}\delta_{\lambda\underline{\lambda}}, \qquad (1.5a)$$

$$\sum_{\lambda,\rho} (\lambda'\mu'\lambda''\mu''\|\lambda\mu)^{\rho}(\lambda'\underline{\mu}'\lambda''\underline{\mu}''\|\lambda\mu)_{\rho} = \delta_{\mu'\underline{\mu}'}\delta_{\mu''\underline{\mu}''}, \qquad (1.5b)$$

where $\lambda = (ab)$ and $\mu = yi$ generalizes usual conditions of the orthogonality and completeness of the orthonormal isofactors.

II. MINIMAL BIORTHOGONAL SYSTEMS AND THE PARACANONICAL SPLITTING

Complete sets of the SU(3) coupled basis states may be chosen among the vectors

$$P^{(ab)}_{yii_z;\mathbf{yii}_z}|a'b'\mathbf{y}'\mathbf{i}'\mathbf{i}'_z\rangle|a''b''\mathbf{y}''\mathbf{i}''\mathbf{i}''_z\rangle$$

$$= (\mathbf{i}'\mathbf{i}'_z\mathbf{i}''\mathbf{i}''_z|\mathbf{ii}_z)\sum_{\rho}(a'b'\mathbf{y}'\mathbf{i}'a''b''\mathbf{y}''\mathbf{i}''\|ab\mathbf{yi};\rho)$$

$$\times \sum_{y',y'',i',i'',i'_z,i''_z}(a'b'y'i'a''b''y''i''\|abyi;\rho)(i'i'_zi''i''_z|ii_z)$$

$$\times |a'b'y'i'i'_z\rangle|a''b''y''i''i''_z\rangle, \qquad (2.1)$$

where the left-hand side general projection operator of SU(3) acts on the direct product state, and the right-hand side CG coefficients of SU(2) and the bilinear combination of the SU(3) isofactors are used as coupling coefficients. (Here ρ is a multiplicity label of arbitrary orthogonal coupled states.) It is convenient to denote the states, coupled with the help of bilinear combinations, as projected ones.

A complete and most convenient basis for evaluation of matrix elements of the labeling operators is formed by the vectors

$$|\eta_{-,+,\tilde I}\rangle \equiv |(a'b'')(a''b'')abyii_z\rangle_{-,+,\tilde I}$$

$$= P^{(ab)}_{yii_z;\tilde y\tilde I\tilde i_z}|a'b'\rangle_{\mathrm{HWS}}|a''b''\rangle_{\mathrm{LWS}}, \qquad (2.2)$$

with the extremal parameters $\mathbf{i}' = \mathbf{i}'_z \to i'_0$, $\mathbf{y}' \to y'_0$, $\mathbf{i}'' = -\mathbf{i}''_z \to \bar{i}''_0$, $\mathbf{y}'' \to \bar{y}''_0$, and subscript $\tilde I \geqslant B$, where

$$B = \tfrac{1}{2}(a+b-b'-a''+|v|), \qquad (2.3)$$

for the linearly independent states. Similarly the coupled bases $|\eta_{+,\tilde I''_{-},+}\rangle$ or $|\eta_{-,\mathbf{I}'',-}\rangle$ with the extremal parameters $\mathbf{i}' = -\mathbf{i}'_z \to \bar{i}'_0$, $\mathbf{y}' \to \bar{y}'_0$, $\mathbf{i} = -\mathbf{i}_z \to \bar{i}_0$, $\mathbf{y} \to \bar{y}_0$, or $\mathbf{i}' = \mathbf{i}'_z \to i'_0$, $\mathbf{y}' \to y'_0$, $\mathbf{i} = \mathbf{i}_z \to i_0$, $\mathbf{y} \to y_0$ may be introduced as well as the bases $|\eta_{\mathbf{I}',+,+}\rangle$, $|\eta_{\tilde I',-,-}\rangle$, and $|\eta_{+,-,\mathbf{I}}\rangle$ (the signs $+,-$ in the subscripts are correlated with the signs of the chosen extremal values of the parameters $\mathbf{z}',\mathbf{z}'',\mathbf{z}$ in the corresponding position). The multiplicity labels ("intrinsic isospins") are in one-to-one correspondence with the Weyl rules for the decomposition of the direct product of irreps, similarly to the Gelfand–Weyl–Biedenharn pattern.[20,21] The bilinear combinations of isofactors for coupling to the states $|\eta_{-,+,\tilde I}\rangle$ will be discussed later.

The basis dual with respect to $|\eta_{-,+,\tilde I}\rangle$ (for $\tilde I \geqslant B$) may be constructed as follows:

$$|\eta^{-,+,\tilde I}\rangle \equiv |(a'b')(a''b'')abyii_z\rangle^{-,+,\tilde I}$$

$$= (i'_0 i'_0 \bar{i}''_0 - \bar{i}''_0|\tilde I\tilde i_z)^{-1}$$

$$\times \sum_{y',y'',i',i''}(a'b'y'i'a''b''y''i''\|abyi)^{-,+,\tilde I}$$

$$\times \sum_{i'_z,i''_z}(i'i'_zi''i''_z|ii_z)|a'b'y'i'i'_z\rangle|a''b''y''i''i''_z\rangle. \qquad (2.4)$$

The nonorthonormal isofactors in Eq. (2.4) satisfy the standard boundary condition[6]

$$(a'b'y'_0i'_0a''b''\bar{y}''_0\bar{i}''_0\|ab\tilde y i)^{-,+,\tilde I} = \delta_{i\tilde I}, \qquad (2.5)$$

for $i \geqslant B$. Similarly the bases $|\eta^{+,\tilde I'',+}\rangle$, $|\eta^{-,\mathbf{I}'',-}\rangle$, etc., may be defined. The constructive elements of all those bases (the isofactors and overlaps) are mutually related by the symmetry relations of the SU(3) isofactors and CG coefficients of SU(2). It may be shown that the nonorthonormal coupled states introduced by Quesne[39] are equivalent (up to normalization) to our states $|\eta^{-,\mathbf{I}'',-}\rangle$ or $|\eta^{\mathbf{I}',+,+}\rangle$.

The corresponding solutions of the discrete boundary value problem may be used in the following modifications of the Wigner–Eckart theorem:

$$\langle abyii_z|T^{(a''b'')}_{y''i''i''_z}|a'b'y'i'i'_z\rangle$$

$$= \sum_{\mathbf{I}''}\langle aby_0i_0\|T^{(a''b'')}_{\mathbf{y}''\mathbf{I}''}\|a'b'y'_0i'_0\rangle$$

$$\times (a'b'y'i'a''b''y''i''\|abyi)^{-,\mathbf{I}'',-}(i'i'_zi''i''_z|ii_z) \qquad (2.6a)$$

$$= \sum_{\tilde I}\langle ab\tilde y\tilde I\|T^{(a''b'')}_{\bar y''_0\bar i''_0}\|a'b'y'_0i'_0\rangle$$

$$\times (a'b'y'i'a''b''y''i''\|abyi)^{-,+,\tilde I}(i'i'_zi''i''_z|ii_z), \qquad (2.6b)$$

where in the right-hand sides special SU(2)-reduced matrix elements of the SU(3) tensor operator $T(a''b'')$ appear.

The dual external multiplicity labels appear quite naturally in the usual relation[19]

$$\sum_{\rho_{1,23}} U(\lambda_1\lambda_2\lambda\lambda_3;\lambda_{12},\lambda_{23})_{\rho_{1,23}}(\lambda_1\mu_1\lambda_{23}\mu_{23}\|\lambda\mu)^{\rho_{1,23}}$$

$$= \sum_{\mu_2,\mu_3,\mu_{12}}(\lambda_1\mu_1\lambda_2\mu_2\|\lambda_{12}\mu_{12})(\lambda_{12}\mu_{12}\lambda_3\mu_3\|\lambda\mu)$$

$$\times (\lambda_2\mu_2\lambda_3\mu_3\|\lambda_{23}\mu_{23})U(\mu_1\mu_2\mu\mu_3;\mu_{12}\mu_{23}) \qquad (2.7)$$

between isofactors and recoupling coefficients U of the group and its subgroup. For simplicity only a single coupling chosen here is non-multiplicity-free. After choosing extremal values of μ_1,μ_{23},μ according to one of the above enumerated patterns, only a single (nonorthogonal) recoupling coefficient with the fixed subscript $\rho_{1,23}$ remains in the left-hand side when the general case of the right-hand side of Eq. (2.7) (with arbitrary μ_1,μ_{23},μ) may be expanded in terms of its corresponding boundary values.

Otherwise, Eq. (2.7) allows us to fix different nonorthonormal systems of SU(3) isofactors. For

$$\lambda_1 = (a'b'), \quad \lambda_{23} = (a''b''), \quad \lambda = (ab), \quad \lambda_2 = (b''0),$$

$$\lambda_3 = (a''+b'',0), \quad \lambda_{12} = (2\tilde I, \tfrac{1}{2}(a'+b'')+b'-\tilde I), \qquad (2.8)$$

the recoupling coefficients of SU(3) in the left-hand side of Eq. (2.7) are multiples of special isofactors of SU(3) ⊃ U(2).[16,40,41] There is one-to-one correspondence of λ_{12} with the Littlewood–Richardson rules for the decomposition $(ba)\times(a''b'')$ to $(b'a')$. In this way Eq. (2.7) allows us to express bilinear combinations of isofactors necessary for constructing states (2.2) in terms of the multiplicity-free isofactors.[5,6]

For

$$\lambda_1=(a'b'),\quad \lambda_{23}=(a''b''),\quad \lambda=(ab),\quad \lambda_2=(0b''),$$

$$\lambda_3=(a''0),\quad \lambda_{12}=(2\tilde{I}',\tfrac{1}{2}(a-a'')+b-\tilde{I}'), \tag{2.9}$$

and

$$a+2b-a'-2b'-a''+b''\geqslant 0, \tag{2.10a}$$

the relation (2.7) gives the solution[6] of the boundary value problem, which is needed for constructing the coupled states $|\eta^{\tilde{I}',-,-}\rangle$ [in this case the recoupling coefficient in the left-hand side of Eq. (2.7) is reduced to a simple factor that does not vanish for only the single value of λ_{12} correlated with the multiplicity label]. Similarly, the set (2.9) with λ_{12} replaced by $(2\tilde{I},\tfrac{1}{2}(a'-b'')+b'-\tilde{I})$ leads to the isofactors associated with the states $|\eta^{-,+,\tilde{I}}\rangle$ in the region

$$a'+2b'+a''-b''-a-2b\geqslant 0. \tag{2.10b}$$

Thus general nonorthonormal SU(3) isofactors of the dual types may be expressed in terms of the SU(2) Racah coefficients and the multiplicity-free SU(3) isofactors. The minimum of the sum (6) has been achieved as a result of the proper choice of the expressions for auxiliary isofactors[16,37] and is caused by the construction asymmetry. The different choice of $\lambda_2,\lambda_3,\lambda_{12}$ in Eq. (2.7) leads to alternative (more complicated) expressions for the nonorthonormal isofactors of SU(3). For example, relation (2.7) with

$$\begin{aligned}&\lambda_2=(\tfrac{1}{2}(a'-a)+a''-v-\mathbf{I}'',2\mathbf{I}''),\\ &\lambda_3=(0,a''-b'+b-v),\\ &\lambda_{12}=(a,b'-a''+v),\end{aligned} \tag{2.9'}$$

solves the boundary value problem associated with the coupled states $|\eta^{-,\mathbf{I}'',-}\rangle$ and equivalent to construction of the SU(3) coupled states proposed by Quesne.[39]

The expansion coefficients represented as boundary values of isofactors

$$\begin{aligned}&(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{i})^{\tilde{I}',-,-}\\ &=(-1)^{\tilde{i}+(a'-b'')/2+2\tilde{I}'}\frac{\nabla(\tfrac{1}{2}b'',\tfrac{1}{2}a',\tilde{i})H(ab\tilde{i}\tilde{z})}{\nabla(\tfrac{1}{2}a'',\tfrac{1}{2}a,\tilde{I}')H(a'b'\tilde{I}'\tilde{z}')}\left(\frac{(2\tilde{I}'+1)(a+1)b'!(a'+b'+1)!a''!(\tilde{I}'-\tilde{z}')!(\tilde{i}+\tilde{z})!}{b''!b!(a+b+1)!(\tilde{I}'+\tilde{z}')!(\tilde{i}-\tilde{z})!}\right)^{1/2}\\ &\quad\times\frac{[\tilde{i}+\tilde{I}'+\tfrac{1}{2}(b'-b+v)]!(b-b'-v)^{(\tilde{I}'+(b-b'-v)/2-\tilde{i})}}{[\tfrac{1}{2}(b-b'-v)-\tilde{i}+\tilde{I}']![\tfrac{1}{2}(b-b'-v)+\tilde{i}+\tilde{I}'+1]!}\quad[\tilde{z}=\tfrac{1}{2}(b''-a')+v,\quad \tilde{z}'=\tfrac{1}{2}(a''-a)-v]\end{aligned} \tag{2.11}$$

allow us to join both regions (2.10a) and (2.10b). They also allow us to expand the isofactors with superscript $\tilde{I}',-,-$ in terms of the isofactors with superscript $-,+,\tilde{i}$, as well as the isofactors with subscript $-,+,\tilde{i}$ in terms of the isofactors with subscript $\tilde{I}',-,-$. Here and below the quasipowers

$$A^{(x)}=A(A-1)(A-2)\cdots(A-x+1)=(A-x)^{(-1)(x)} \tag{2.12}$$

and other notations

$$H(abiz)=[(a+z-i)!(a+z+i+1)!(b-z-i)!(b-z+i+1)!]^{1/2}, \tag{2.13}$$

$$\nabla(abc)=[(a+b-c)!(a-b+c)!(a+b+c+1)!/(b+c-a)!]^{1/2}, \tag{2.14}$$

$$\Gamma(abiz)=\left[\frac{(a+z-i)!(a+z+i+1)!(i+z)!}{(b-z-i)!(b-z+i+1)!(i-z)!}\right]^{1/2} \tag{2.15}$$

are used.

It is remarkable that Eq. (2.11) [obtained in region (2.10a) immediately from Eqs. (2.7) and (2.9) and in the region (2.10b) after inverting the corresponding triangular matrix[6]] accepts unified analytical form. In the region (2.10a) the allowed values of the multiplicity label $\tilde{I}'$ are completely determined by the properties of discrete functions (2.13) and (2.14) in Eq. (2.11), as well as the values of $\tilde{I}$ in the region (2.10b). So the bases $|\eta_{\tilde{I}',-,-}\rangle$ and $|\eta_{-,+,\tilde{I}}\rangle$ are never overcomplete at the same time. The extremal values of $\tilde{I}$ (or $\tilde{I}$) may be found from the triangular and betweenness conditions (including three inequalities for the maximum and six for the minimum).

The lengths of intervals for $\tilde{I}$ and $\tilde{I}'$ form the pattern

$$|q_{ij}|=\begin{vmatrix} b'-a''+a+v & a'-b''+b-v & b-v & b & b' & b'+v\\ a & a+v & a-a'+b''+v & a''-b'+b-v & a''-v & a''\\ a'-v & a' & b'' & b''+v & b'+b''-b+v & a'+a''-a-v\end{vmatrix}, \tag{2.16}$$

where $\min q_{ij}+1$ gives the external multiplicity[6] of (ab) in $(a'b')\times(a''b'')$. As demonstrated below, symmetries of the overlaps and boundary paracanonical isofactors (orthogonalization coefficients) may be described by some of the $72=6\times6\times2$ transformations of pattern (2.16) which include the row permutations, the permutations of the couples of columns (12,34,56), and the permutation of the even and odd columns (with the change of the sign of v in the last case).

The triangular nature of the transformations with the help of Eq. (2.11) or with the related matrices allows us to fix the following correspondence:

$$\tilde{i}\rightarrow\tfrac{1}{2}(b-b'-v)+\tilde{I}'\rightarrow\tfrac{1}{2}(a-a''+v)+\tilde{I}''\rightarrow\tilde{i} \tag{2.17}$$

between the states of the biorthogonal systems with the labels $-,+,\tilde{i};\tilde{I}',-,-$; and $+,\tilde{I}'',+$. The proper sequence of the Schmidt process beginning with the highest values of the $\tilde{i}, \tilde{I}'$, or $\tilde{I}''$ for the coupled states with subscript and with the lowest values for the states with superscript reduces completely (up to the phase convention) the arbitrarity of the orthonormal coupled states obtained. The new splitting, which may be called paracanonical, leaves only two versions of the orthogonal coupled bases [two systems of the paracanonical SU(3) unit tensor operators or isofactors, respectively] instead of 12 original types of nonorthogonal bases (or systems of isofactors).

For the first version of the paracanonical splitting the orthogonal isofactors vanish unless

$$\begin{aligned} &i-\tilde{i}\leqslant i_0'+z'+\bar{i}_0''-z'',\\ &i'+\tfrac{1}{2}(b-b'-v)-\tilde{i}\leqslant i_0''+z''+i_0+z, \\ &i''+\tfrac{1}{2}(a-a''+v)-\tilde{i}\leqslant\bar{i}_0-z+\bar{i}_0'-z'.\end{aligned} \tag{2.18}$$

This property may be proved with the help of Eq. (2.22). The restriction of the boundary isofactors proposed by Hecht[19] (see also Refs. 5, 14, and 22) is generalized here for arbitrary values of parameters.

Thus the paracanonical classification is invariant with respect to the cyclic permutations of the parameters

$$ba-yi\rightarrow a'b'y'i'\rightarrow a''b''y''i''\rightarrow ba-yi \tag{2.17'}$$

in isofactors along with the relabeling (2.17).

Let us postpone the analysis of the paracanonical splitting and return to construction of the nonorthogonal coupled states. In the complementary region

$$\begin{aligned} &2a''+b''-a'+b'-2a-b<0,\\ &a'+2b'+a''-b''-a-2b<0,\end{aligned} \tag{2.10c'}$$

the nonvanishing values of the isofactors of the type (2.5) also appear for $i<B$. They may be found with the help of Eq. (2.11) and the inverse transformation. In this case the states $|\eta_{-,+,\tilde{i}}\rangle$ with $\tilde{i}<B$ are linearly dependent and may be expanded as follows[15,17,28]:

$$|\eta_{-,+,\tilde{i}}\rangle=\sum_{\tilde{I}}R_{\tilde{i}\tilde{I}}|\eta_{-,+,\tilde{I}}\rangle \tag{2.19a}$$

$$=\sum_{\tilde{I}}(i_0'i_0'\bar{i}_0''-\bar{i}_0''|\tilde{i}\,\tilde{i}_z)(i_0'i_0'\bar{i}_0''-\bar{i}_0''|\tilde{I}\tilde{i}_z)^{-1}(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{i})^{-,+,\tilde{I}}|\eta_{-,+,\tilde{I}}\rangle \tag{2.19b}$$

$$\begin{aligned}&=\sum\left\{\delta_{\tilde{i}\tilde{I}}+\left(\frac{(\tilde{I}-\tilde{i}_z)!(\tilde{i}+\tilde{i}_z)!(\tilde{I}-\tilde{z})!(\tilde{i}+\tilde{z})!}{(\tilde{i}-\tilde{i}_z)!(\tilde{I}+\tilde{i}_z)!(\tilde{i}-\tilde{z})!(\tilde{I}+\tilde{z})!}\right)^{1/2\,\mathrm{sign}\,v}[(2\tilde{I}+1)(2\tilde{i}+1)]^{1/2}H(ab\tilde{i}\,\tilde{z})H^{-1}(ab\tilde{I}\,\tilde{z})\right.\\ &\quad\left.\times\frac{(-1)^{\tilde{i}-B}(\tilde{I}+B)!}{(\tilde{I}-\tilde{i})(\tilde{I}+\tilde{i}+1)(\tilde{I}-B)!(\tilde{i}+B)!(B-\tilde{i}-1)!}\right\}|\eta_{-,+,\tilde{I}}\rangle.\end{aligned} \tag{2.19c}$$

Recursive constructions (2.7)–(2.9) allowed us also to express the normalization coefficients and overlaps of the nonorthogonal states in terms of certain bilinear combinations of the recoupling coefficients.

The corresponding recoupling coefficients for the states $|\eta_{-,+\tilde{I}}\rangle$ coincide with the complementary resubducing coefficients of the chains $\mathrm{U}(n)\supset\mathrm{U}(n-2)+\mathrm{U}(2)$ and $\mathrm{U}(n)\supset\mathrm{U}(n-1)\supset\mathrm{U}(n-2)$ and may be found[5,42] with the help of the Löwdin–Shapiro projection operators of the subgroup SU(2)[43,44] (which transform the last two components of weight). In this way one gets the following expression for the overlaps:

$$\begin{aligned}\langle\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}}\rangle&=(i_0'i_0'\bar{i}_0''-\bar{i}_0''|\tilde{I}\tilde{i}_z)(i_0'i_0'\bar{i}_0''-\bar{i}_0''|\tilde{J}\tilde{i}_z)(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})\\ &=\left(\frac{(2\tilde{I}+1)(2\tilde{J}+1)(\tilde{I}-\tilde{i}_z)!(\tilde{J}-\tilde{i}_z)!}{(\tilde{I}+\tilde{i}_z)!(\tilde{J}+\tilde{i}_z)!}\right)^{1/2}\frac{1}{\Gamma(ab\tilde{I}\,\tilde{z})\Gamma(ab\tilde{J}\,\tilde{z})}\\ &\quad\times\frac{(a+1)(b+1)(a+b+2)a'!a''!b''!(a''+b''+1)!}{(b'+b''-b+v)!(b'+b''+v+1)!(a+b'+b''+v+2)!}\\ &\quad\times\sum_{j,z}\frac{(-1)^{a'+b''+2j}(2j+1)(2z-2\tilde{z})!\Gamma^2(abjz)}{(a''-2\tilde{z}+2z+1)!\nabla^2(z-\tilde{z},\tilde{I},j)\nabla^2(z-\tilde{z},\tilde{J},j)}\\ &\quad\times\frac{(b'+b''+v-z-j)!(b'+b''+v-z+j+1)!}{\nabla^2(b''+v-i_0'-z,i_0',j)}.\end{aligned} \tag{2.20}$$

Here

$$(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})=\sum_{\rho}(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{I};\rho)(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{J};\rho). \tag{2.20'}$$

The substitution group technique applied to special recoupling coefficients[6] allow us to find the overlaps of the dual states.[15,16,45] Both classes of overlaps were generalized for SU(n).[5,16]

Equations (2) and (6) of Ref. 15 or Eqs. (4.4) and (4.7) of Ref. 18 (Ref. 46) along with the symmetry relations

$$\left(\frac{(2i+1)(b''+1)}{(2i''+1)(2\tilde{I}+1)}\right)^{1/2}(a'b'y'i'a''b''y''i''\|abyi)^{-,+,\tilde{I}}=(-1)^{i'-i_0'+(y_0'-y')/2}(a'b'y'i'ba-yi\|b''a''-y''i'')^{-,\tilde{I},-} \tag{2.21a}$$

$$=(-1)^{i''-i-\bar{i}_0''+\tilde{I}+(y_0'-y')/2}(b'a'-y'i'abyi\|a''b''y''i'')^{+,\tilde{I},+} \tag{2.21b}$$

give another solution of the boundary value problem. Its explicit form

$$\begin{aligned}
&((a'b'y'i'a''b''y''i''\|abyi))^{-,+,\tilde{i}}\\
&\quad=(-1)^{a'+b''+z'+i''-i}N_{\tilde{i}}^{-1}[(2i'+1)(2i''+1)(b''-2z''+1)]^{1/2}(2\tilde{i}+1)\\
&\qquad\times\sum_{r,r',s}\frac{(-1)^{2r'}(2r+1)(2r'+1)(b'-z'+s-r')!(b'-z'+s+r'+1)!\Gamma^2(a,b,r,\bar{i}_0''+z'+v-s)}{[(2s)!(b''-2z''-2s)!]^{1/2}\nabla(si'r')\nabla(i_0'+z'-s,i_0',r')\nabla(\bar{i}_0''-z''-s,ir)\nabla(i_0'+z'-s,\tilde{i},r)}\\
&\qquad\times\begin{Bmatrix} r' & r & \bar{i}_0'' \\ \tilde{i} & i_0' & i_0'+z'-s \end{Bmatrix}\begin{Bmatrix} r & i & \bar{i}_0''-z''-s \\ r' & i' & s \\ \bar{i}_0'' & i'' & \bar{i}_0''-z'' \end{Bmatrix},
\end{aligned} \tag{2.22}$$

with the 6*j* and stretched 9*j* coefficients of SU(2) [including single and double sums, see Eqs. (29.1) and (32.13) of Ref. 36] on the right-hand side gives an expression for the nonorthonormal SU(3) isofactors with the exception of region (2.10c). Here and in Sec. III,

$$N_{\tilde{i}}=\Gamma(ab\tilde{i}\,\tilde{z})\Gamma(abiz)\Gamma^{-1}(a'b'i'z')\left(\frac{b'!(a'+b'+1)!a''!(a''+b''+1)!}{(b''-2z'')!(a''+z''-i')!(a''+z''+i''+1)!}\right)^{1/2}. \tag{2.23}$$

In region (2.10c), Eq. (2.22) gives only the expansion coefficients of the arbitrary isofactors

$$(a'b'y'i'a''b''y''i''\|abyi;\rho)=\sum_{\tilde{i}}((a'b'y'i'a''b''y''i''\|abyi))^{-,+,\tilde{i}}(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{i};\rho) \tag{2.24}$$

in terms of their boundary values in the region wider than the multiplicity of the irreps. In this last case the additional expansion

$$(a'b'y'i'a''b''y''i''\|abyi)^{-,+,\tilde{I}}=\sum_{\tilde{i}}(a'b'y_0'i_0'a''b''y_0''i_0''\|ab\tilde{y}\tilde{i})^{-,+,\tilde{I}}((a'b'y'i'a''b''y''i''\|abyi))^{-,+,\tilde{i}} \tag{2.25}$$

with the coefficient used in Eq. (2.19) is necessary. Otherwise, Eq. (2.22) may be used for the expansion of the SU(3) direct product states [coupled with the help of the SU(2) CG coefficients] in terms of the SU(3) coupled basis (2.2) which may be overcomplete.

It is remarkable that in the case of the overcomplete basis two alternatives may be chosen for summation intervals in formulas of the type (2.6) or (2.7): one can limit oneself by the linearly independent states or the sum may be taken in a wider region omitting the additional expansion used in Eq. (2.19) and substituting the isofactors in the right-hand side by the pseudoisofactors, i.e., by the expansion coefficients of the type (2.22). Such "painless" [47] expansion is more simple analytically, but demands more time for computation.

III. ANALYTICAL INVERSION AND NEW EXPRESSIONS FOR ISOFACTORS AND OVERLAPS

The analytical inversion[17,23] is a discrete operation of the analytical continuation, the nonorthonormal isofactors, or other resubducing coefficients to the dual ones. It may be associated with group automorphism, which corresponds to the transition to the inverse elements of this group. The analytical inversion symmetry should not be mixed up with the hook permutation[48] or the substitution[49] group symmetry. The analytical inversion is allowed only for the nonorthogonal isofactors represented in the analytical form (in terms of the factorial sums) not being allowed for the usual algebraic-polynomial expressions.

The relation or the analytical inversion between the dual isofactors of SU(3)

$$(a'b'y'i'a''b''y''i''\|abyi)_{-,+,\tilde{i}}$$

$$\equiv\sum_{\rho}(a'b'y'i'a''b''y''i''\|abyi;\rho)(a'b'y_0'i_0'a''b''\bar{y}_0''\bar{i}_0''\|ab\tilde{y}\tilde{i};\rho) \tag{3.1a}$$

$$\begin{aligned}
&=\frac{(a+1)(b+1)(a+b+2)}{(2\tilde{i}+1)[(b'+1)(a'+b'+2)(a''+1)(a''+b''+2)]^{1/2}}\\
&\quad\times((-a'-2,-b'-2,-y',-i'-1,-a''-2,-b''-2,-y'',-i''-1\|-a-2,b-2,-y,-i-1)\ )^{-,+,\tilde{i}-1}
\end{aligned} \tag{3.1b}$$

may be based after examination of the behavior of matrix elements of the SU(3)×SU(3) generators (see Sec. V).

This relation applied to Eq. (2.22) [and used together with Eqs. (29.19) and (31.15) of Ref. 36 for 6*j* and 9*j* coefficients of SU(2) with some negative parameters] allows us to obtain the two following expressions for nonorthonormal isofactors (bilinear combinations of orthonormal isofactors) labeled by the subscript:

$$
\begin{aligned}
&(a'b'y'i'a''b''y''i''\|abyi)_{-,+,\tilde{I}}\\
&\quad=(a+1)(b+1)(a+b+2)[(2i'+1)(2i''+1)]^{1/2}N_{\tilde{I}}\\
&\qquad\times\sum_{r,r',s}\frac{(-1)^{b''+\tilde{I}-\tilde{z}-i''+z-s+r'}(2r+1)(2r'+1)}{(b'-z'+s-r'+1)!(b'-z'+s+r'+2)!\nabla(s-i'_0-z',i'_0,r')}\\
&\qquad\times\left(\frac{(2s+1)!}{(2z''-b''+2s)!}\right)^{1/2}\frac{\nabla(si'r')\Gamma^2(b,a,r,s-\bar{i}''_0-v-z')}{\nabla(z''-\bar{i}''_0+s,i,r)\nabla(s-i'_0-z',\tilde{I},r)}\begin{Bmatrix} r' & r & \bar{i}''_0\\ \tilde{I} & i'_0 & s-i'_0-z'\end{Bmatrix}\begin{Bmatrix} \bar{i}''_0 & i'' & \bar{i}''_0-z''\\ z & i & s+z''-\bar{i}''_0\\ r' & i' & s\end{Bmatrix}
\end{aligned}
\tag{3.2a}
$$

$$
\begin{aligned}
&=(a+1)(b+1)(a+b+2)[(2i'+1)(2i''+1)]^{1/2}N_{\tilde{I}}\sum_{r,r',s}\frac{(-1)^{a'+b+z'+\bar{i}''_0-r+s}(2r+1)(2r'+1)}{(b'-z'-s-r')!(b'-z'-s+r'+1)!\nabla(si'r')}\\
&\quad\times\left(\frac{(b''-2z''+2s+1)!}{(2s)!}\right)^{1/2}\frac{\nabla(i'_0+z'+s,i'_0,r')\nabla(\bar{i}''_0-z''+s,i,r)}{(\bar{i}''_0+z'+v+s-r)!(\bar{i}''_0+z'+v+s+r+1)!}\\
&\quad\times\frac{\nabla(i'_0+z'+s,\tilde{I},r)(b-\bar{i}''_0-z'-v-s+r)!}{(\bar{i}''_0+z'+v-b+s+r)!(a+\bar{i}''_0+z'+v+s-r+1)!(a+\bar{i}''_0+z'+v+s+r+2)!}\\
&\quad\times\begin{Bmatrix} r' & r & \bar{i}''_0\\ \tilde{I} & i'_0 & i'_0+z'+s\end{Bmatrix}\begin{Bmatrix} \bar{i}''_0 & i'' & \bar{i}''_0-z''\\ r' & i' & s\\ r & i & \bar{i}''_0-z''+s\end{Bmatrix},
\end{aligned}
\tag{3.2b}
$$

where $N_{\tilde{I}}$ is defined by Eq. (2.23). Different versions of sums are obtained here depending on whether the summation parameter s has been reflected $(s\rightarrow -s-1)$ or not. Equation (3.2a) is equivalent to the expression[11,50] for the matrix elements of the projection operators of SU(3). In all three expressions (2.22), (3.2a), and (3.2b) the summation intervals differ similarly as in Eqs. (9) and (12)–(15) of Ref. 37 [see also Eq. (22) of Ref. 16] for the multiplicity-free isofactors of SU(n).

The expressions constructed by means of Eqs. (2.7)–(2.9) take algebraic-polynomial forms for the fixed irrep$(a''b'')$, parameters y'', i'', and shifts $a-a'$, $b-b'$, $i-i'$ [the multiplicity label $\tilde{I}$ being correlated according to the correspondence (2.17)]. The same property is satisfied by Eqs. (2.22), but Eq. (3.2a) takes algebraic-polynomial form for the fixed (ab), etc., and Eq. (3.2b) takes such form for fixed $(a'b')$, etc.

As a rule the expression (2.22) demands less time for computation because it vanishes for $|i-\tilde{I}|>i'_0+z'+\bar{i}''_0-z''$. The application of Eq. (3.2b) will be discussed in Sec. VI.

The analytical inversion applied to Eq. (2.20) allows us to obtain the following expression for the overlaps of the dual states (2.4) [compare with the SU(3)⊃SO(3) case, Eq. (5.13) of Ref. 17]:

$$
\begin{aligned}
&(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})(i'_0i'_0\bar{i}''_0-\bar{i}''_0|\tilde{I}\tilde{i}_z)^{-1}(i'_0i'_0\bar{i}''_0-\tilde{i}''_0|\tilde{J}\tilde{i}_z)^{-1}\\
&\quad=\langle\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}}\rangle=\sum_{i,j}R_{\tilde{i}\tilde{I}}R_{\tilde{j}\tilde{J}}\Gamma(ab\tilde{i}\,\tilde{z})\Gamma(ab\tilde{j}\,\tilde{z})\\
&\qquad\times\frac{(b'+1)(a'+b'+2)(b'+b''-b+v+1)!(b'+b''+v+2)!(a+b'+b''+v+3)!}{(a+1)(b+1)(a+b+2)a'!a''!b''!(a''+b''+1)!}\\
&\qquad\times\left(\frac{(2\tilde{i}+1)(2\tilde{j}+1)(\tilde{i}+\tilde{i}_z)!(\tilde{j}+\tilde{i}_z)!}{(\tilde{i}-\tilde{i}_z)!(\tilde{j}-\tilde{i}_z)!}\right)^{1/2}\sum_{j,z}\frac{(2j+1)(2\tilde{z}-2z)!}{(b'+b''+v-z+j+3)!}\\
&\qquad\times\frac{(a''-2\tilde{z}+2z)!\nabla^2(b''+v-i'_0-z,i'_0,j)}{(b'+b''+v-z-j+2)!\nabla^2(\tilde{z}-z,\tilde{i},j)\nabla^2(\tilde{z}-z,\tilde{j},j)\Gamma^2(abjz)}.
\end{aligned}
\tag{3.3}
$$

In region (2.10c) infinite terms with $z<\tilde{z}-a''/2$ appear that vanish only after the expansion coefficients $R_{\tilde{i}\tilde{I}}$ from Eq. (2.19) are used. Therefore, the additional condition $z\geqslant\tilde{z}-a''/2$ is expedient.

The substitution

$$
a'\rightarrow -a'-2,\quad b'\rightarrow a'+b'+1,\quad a''\rightarrow a''+b''+1,\quad b''\rightarrow -b''-2
\tag{3.4}
$$

(and $v\rightarrow b''-a'+v$, $i'_0\rightarrow -i_0-1$, $i''_0\rightarrow -\bar{i}''_0-1$) leaves the dual isofactors represented by Eqs. (2.22) and (3.2a), (3.2b) invariant (up to sign). The same substitution [along with the phase factor $(-1)^{\tilde{I}-\tilde{J}}$] applied to overlaps $(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})$ and $(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})$ [see Eqs. (2.20) and (3.3)] allows us to obtain new expressions for those overlaps:

$$
(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})=\frac{(a+1)(b+1)(a+b+2)(-1)^{\tilde{I}-\tilde{J}}a''!(a''+b''+1)!}{(b'+v)!(a+b'+v+1)!\nabla(\bar{i}''_0,i'_0,\tilde{I})\nabla(\bar{i}''_0,i'_0,\tilde{J})\Gamma(ab\tilde{I}\tilde{z})\Gamma(ab\tilde{J}\tilde{z})}
$$

$$\times\sum_{j,z}\frac{(2j+1)(b'+v-j-z-1)^{(b-j-z)}(b'+v-z+j)!}{(a''+b''-2\tilde{z}+2z+2)!\nabla^{2}(z-\tilde{z},\tilde{I},j)\nabla^{2}(z-\tilde{z},\tilde{J},j)}(2z-2\tilde{z})!\Gamma^{2}(abjz)\nabla^{2}(i_0'-v+z,i_0,j), \tag{3.5}$$

$$(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})=(2\tilde{I}+1)(2\tilde{J}+1)\frac{(b'+1)(a'+b'+2)(b'-b+v)!(b'+v+1)!(a+b''+v+2)!}{(a+1)(b+1)(a+b+2)a''!(a''+b''+1)!}$$
$$\times\nabla(\bar{i}_0'',i_0',\tilde{I})\nabla(\bar{i}_0'',i_0',\tilde{J})\Gamma(ab\tilde{I}\tilde{z})\Gamma(ab\tilde{J}\tilde{z})$$
$$\times\sum_{j,z}\frac{(-1)^{a'+b''+\tilde{I}-\tilde{J}+2j}(2j+1)(2\tilde{z}-2z)!(a''+b''-2\tilde{z}+2z+1)!}{(b'+v-z+j+2)!(b'+v-z-j+1)!\nabla^{2}(i_0'-v+z,i_0',j)}$$
$$\times\nabla^{-2}(\tilde{z}-z,\tilde{I},j)\nabla^{-2}(\tilde{z}-z,\tilde{J},j)\Gamma^{-2}(abjz). \tag{3.6}$$

Expression (3.6) is equivalent (after some symmetry transformations) to the corrected Eq. (11) of Ref. 15. The expression (3.3) outside the region (2.10c′) includes terms of the same sign as well as Eq. (3.5) for $b'-b+v\geqslant 0$ (in both cases the states $|\eta_{-,+,\tilde{I}}\rangle$ are not overcomplete). For $b'-b+v<0$, Eq. (3.6) is indefinite.

The consideration of Eqs. (2.20), (3.3), (3.5), and (3.6) allows us to prove that the functions

$$(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})[(a+1)(b+1)(a+b+2)b'!(a'+b'+1)!a''!(a''+b''+1)!]^{-1}, \tag{3.7a}$$

$$(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})(a+1)(b+1)(a+b+2)b'!(a'+b'+1)!a''!(a''+b''+1)!, \tag{3.7b}$$

are invariant with respect to the $24=6\times2\times2$ transformations of pattern (2.16) (from those of 72 mentioned above) that do not interchange the parameters of the two last columns with the remaining ones.

Separate sums (with respect to $j\pm z$) in Eqs. (2.20), (3.3), (3.5), and (3.6) may be represented as the Saalschutzian[51] ${}_{k+1}F_k(1)$ series. Single sums remaining for extremal values of $\tilde{I}$ or $\tilde{J}$ are equivalent to the Saalschutzian ${}_4F_3(1)$ series; however, they are not of the type that appears in $6j$ coefficients of SU(2).

It is remarkable that both sums in Eq. (2.20) are finite for the fixed values of a single parameter (q_{11}, q_{13}, q_{31}, or q_{33}) from the pattern (2.16), as well as in Eq. (3.3) for fixed q_{22} or q_{24}, in Eq. (3.5)—for fixed q_{11} or q_{12} and in Eq. (3.6)—for fixed q_{23}, q_{24}, q_{33}, or q_{34}. The substitution (permutation)

$$a\leftrightarrow b,\quad a'\leftrightarrow b'',\quad a''\leftrightarrow b'\quad (v\rightarrow -v) \tag{3.8}$$

allow us to express corresponding overlaps in a finite form also for the fixed single parameters q_{22}, q_{24}, q_{32}, q_{34}; q_{11}, q_{13}; q_{23}, q_{24}; q_{11}, q_{12}, q_{31}, or q_{32}, respectively.

For $b-2\tilde{z}=b'-a''+a+v$ fixed (along with fixed $b-\tilde{z}-\tilde{I}$ and $\tilde{i}-\tilde{z}$, where $\tilde{I}\geqslant\tilde{J}$) it is expedient to introduce the renormalized overlaps

$$E_{\tilde{I},\tilde{J}}=(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})[(\tilde{I}+\tilde{z})^{(-1)(b-2\tilde{z}+1)}(\tilde{J}+\tilde{z})^{(-1)(b-2\tilde{z}+1)}]^{1/2}$$
$$\times\frac{a''^{(-1)(b-2\tilde{z}+1)}(a''+b''+1)^{(-1)(b-2\tilde{z}+1)}}{(a+1)(b+1)(a+b+2)}, \tag{3.9a}$$
$$F^{\tilde{I},\tilde{J}}=(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})[(b-\tilde{z}+\tilde{I}+1)^{(b-2\tilde{z}+1)}(b-\tilde{z}+\tilde{J}+1)^{(b-2\tilde{z}+1)}]^{1/2}$$
$$\times\frac{(a+1)(b+1)(a+b+2)b'^{(b-2\tilde{z})}(a'+b'+1)^{(b-2\tilde{z})}}{(2\tilde{I}+1)(2\tilde{J}+1)(a''+1)(a''+b''+2)} \tag{3.9b}$$

equivalent to polynomials in five free parameters (e.g., a', b', a'', b'', v) of the total degree

$$3(b-2\tilde{z}-|\tilde{I}-\tilde{J}|) \tag{3.10}$$

with integer coefficients (and common integer factors under the square roots for $\tilde{I}\neq\tilde{J}$) when Eq. (2.20) or (3.5) is used for $(\eta_{-,+,\tilde{I}}|\eta_{-,+,\tilde{J}})$ and Eq. (3.3) or (3.6) along with substitution (3.8) used for $(\eta^{-,+,\tilde{I}}|\eta^{-,+,\tilde{J}})$. Some concrete expressions for $E_{\tilde{I},\tilde{J}}$ and $F^{\tilde{I},\tilde{J}}$ are given and discussed in Appendix A.

The above mentioned invariance properties of functions (3.7a) and (3.7b) allow us to express the overlaps for other fixed single parameters of q_{ij} ($j\leqslant 4$) type [from pattern (2.16)] in polynomial form.

IV. STRUCTURE OF PARACANONICAL AND OTHER ALGEBRAIC SYSTEMS OF ISOFACTORS

The usual and dual Gram–Schmidt processes give the following expansions of the orthonormal states ξ_α ($1\leqslant\alpha\leqslant n$):

$$\xi_\alpha=[\Gamma^\alpha\Gamma^{\alpha+1}]^{-1/2}\det\begin{bmatrix}\langle\eta^n\eta^n\rangle\cdots\langle\eta^n\eta^{\alpha+1}\rangle\eta^n\\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\\ \langle\eta^\alpha\eta^n\rangle\cdots\langle\eta^\alpha\eta^{\alpha+1}\rangle\eta^\alpha\end{bmatrix} \tag{4.1a}$$

$$=[\Gamma_\alpha\Gamma_{\alpha-1}]^{-1/2}\sum_{\beta=\alpha}^{n}\det\begin{bmatrix}\langle\eta_1\eta_1\rangle\cdots\langle\eta_1\eta_{\alpha-1}\rangle\langle\eta_1\eta_\beta\rangle\\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\\ \langle\eta_\alpha\eta_1\rangle\cdots\langle\eta_\alpha\eta_{\alpha-1}\rangle\langle\eta_\alpha\eta_\beta\rangle\end{bmatrix}\eta^\beta, \tag{4.1b}$$

where $\langle \eta_\alpha \eta^\beta \rangle = \delta_{\alpha\beta}$ and the minors

$$\Gamma^\alpha = \det \begin{bmatrix} \langle \eta^n \eta^n \rangle \cdots \langle \eta^n \eta^\alpha \rangle \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \langle \eta^\alpha \eta^n \rangle \cdots \langle \eta^\alpha \eta^\alpha \rangle \end{bmatrix}, \quad \Gamma_\alpha = \det \begin{bmatrix} \langle \eta_1 \eta_1 \rangle \cdots \langle \eta_1 \eta_\alpha \rangle \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \langle \eta_\alpha \eta_1 \rangle \cdots \langle \eta_\alpha \eta_\alpha \rangle \end{bmatrix}$$

appear.

Now the boundary paracanonical isofactors [i.e., the orthogonalization coefficients of the general paracanonical isofactors according to Eq. (2.24)] for $q_{11} = b' - a'' + a + v$ fixed may be written as follows:

$$(a'b'y_0' i_0' a'' b'' \bar{y}_0'' \bar{i}_0'' \| ab\tilde{y}i;\tilde{I})$$
$$= \left[\frac{(a+1)(b+1)(a+b+2) b'^{(b-\tilde{z}-\tilde{I})} (a'+b'+1)^{(b-\tilde{z}-\tilde{I})} (b-\tilde{z}+i+1)^{(b-\tilde{z}-\tilde{I})}}{a''^{(-1)(\tilde{I}-\tilde{z}+1)} (a''+b''+1)^{(-1)(\tilde{I}-\tilde{z}+1)} (i+\tilde{z})^{(-1)(\tilde{I}-\tilde{z}+1)}} \right]^{1/2} \frac{K_{\tilde{I},i}\, g_{\tilde{I},i}}{[g_{\tilde{I},\tilde{I}}\, g_{\tilde{I}+1,\tilde{I}+1}]^{1/2}} \tag{4.2}$$

for $\min(b-\tilde{z}, a+\tilde{z}, i_0' + \bar{i}_0' + \bar{i}_0'') \geqslant \tilde{I} \geqslant i \geqslant \max(|\bar{i}_z|, |\tilde{z}|)$ and vanish otherwise.

Here

$$K_{\tilde{I},i} = [(\tilde{I} + \bar{i}_z)^{(\tilde{I}-i)} (\tilde{I} - \bar{i}_z)^{(\tilde{I}-i)}$$
$$\times (i_0' + \bar{i}_0'' - i)^{(\tilde{I}-i)} (i_0' + \bar{i}_0'' + \tilde{I} + 1)^{(\tilde{I}-i)}$$
$$\times (a + \tilde{z} - i)^{(\tilde{I}-i)} (a + \tilde{z} + \tilde{I} + 1)^{(\tilde{I}-i)}]^{1/2} k_{\tilde{I},i}, \tag{4.3}$$

where $k_{\tilde{I}i}$ is an irrational factor (the common measure) of

$$\left[\binom{\tilde{I}-\tilde{z}}{\tilde{I}-i} \binom{b-\tilde{z}-i}{\tilde{I}-i} \right]^{1/2}. \tag{4.4}$$

The factors $g_{\tilde{I}i}$ $(b-\tilde{z} \geqslant \tilde{I} \geqslant i \geqslant \tilde{z})$ are polynomials with some integer coefficients in free parameters of the total degree

$$3(b-\tilde{z}-\tilde{I}+1)(\tilde{I}-\tilde{z}) - 3(\tilde{I}-i). \tag{4.5}$$

There are two possibilities to express these polynomials:

$$g_{\tilde{I},i} = \det \begin{bmatrix} E_{b-\tilde{z},b-\tilde{z}} \cdots E_{b-\tilde{z},\tilde{I}+1} & E_{b-\tilde{z},i} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ E_{\tilde{I},b-\tilde{z}} \cdots E_{\tilde{I},\tilde{I}+1} & E_{\tilde{I},i} \end{bmatrix}$$
$$\times (b-\tilde{z}+\tilde{I}+1)^{(b-\tilde{z}-\tilde{I})}$$
$$\times \left[(b-\tilde{z}+i+1)^{(b-\tilde{z}-\tilde{I})} K_{\tilde{I},i} \prod_{j=\tilde{I}}^{b-\tilde{z}-1} e_j \right]^{-1}$$

$$= \det \begin{bmatrix} F^{\tilde{z},\tilde{z}} \cdots F^{\tilde{z},\tilde{I}-1} & 0 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \\ F^{i,\tilde{z}} \cdots F^{i,\tilde{I}-1} & 1 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \\ F^{\tilde{I},\tilde{z}} \cdots F^{\tilde{I},\tilde{I}-1} & 0 \end{bmatrix} \tag{4.6a}$$
$$\times (i+\tilde{z})^{(-1)(\tilde{I}-z+1)}$$
$$\times \left[(2i+1)(\tilde{I}+\tilde{z})^{(-1)(\tilde{I}-\tilde{z})} K_{\tilde{I},i} \prod_{j=\tilde{z}+1}^{\tilde{J}-1} f_j \right]^{-1}, \tag{4.6b}$$

where

$$e_j = b'^{(b-\tilde{z}-j)} (a'+b'+1)^{(b-\tilde{z}-j)}$$
$$\times (a+b'+v+1)^{(b-\tilde{z}-j)}$$
$$\times (a+b'+b''+v+2)^{(b-\tilde{z}-j)}$$
$$\times [(b-\tilde{z}+j+1)^{(b-\tilde{z}-j)}]^2, \tag{4.7a}$$

$$f_j = (a''-a-v)^{(-1)(j-\tilde{z})}$$
$$\times (a'+a''-a-v+1)^{(-1)(j-\tilde{z})}$$
$$\times (a''+1)^{(-1)(j-\tilde{z})}$$
$$\times (a''+b''+2)^{(-1)(j-\tilde{z})} [(j+\tilde{z})^{(-1)(j-\tilde{z})}]^2, \tag{4.7b}$$

and the determinant of the order $b-\tilde{z}-\tilde{I}+1$ appeared in Eq. (4.6a), as well as the minor of the order $\tilde{I}-\tilde{z}$ in Eq. (4.6b). Particularly (see also Table I)

$$g_{b-\tilde{z}+1,b-\tilde{z}+1} = g_{\tilde{z},\tilde{z}} = 1, \quad g_{b-\tilde{z},\tilde{z}} = (-1)^{b-2\tilde{z}},$$
$$g_{b-\tilde{z},i} = E_{b-\tilde{z},i}/K_{b-\tilde{z},i}, \quad g_{\tilde{z}+1,\tilde{z}+1} = F^{\tilde{z},\tilde{z}},$$
$$g_{\tilde{z}+1,\tilde{z}} = -F^{\tilde{z}+1,\tilde{z}}/K_{\tilde{z}+1,\tilde{z}}.$$

The polynomial form of $g_{\tilde{I},i}$ may be proved at first for the denominator factors $g_{\tilde{I},\tilde{I}}$. In this case $K_{\tilde{I},\tilde{I}} = 1$ and the products $e_j f_j$ are completely determined by dual constructions of the type (4.1a), (4.1b) from Eqs. (3.10a), (3.10b). The appearance of the factor $b'^{(b-\tilde{z}-j)}$ in e_j caused by the null space property, i.e., by the appearance of linearly dependent columns of the determinant in Eq. (4.6a). The complementary factor $(a''-a-v)^{(-1)(j-\tilde{z})}$ in f_j allows us to counsel vanishing in region (2.10c) factors of the determinant in Eq. (4.6b). Four first factors of e_J (or f_j) are interrelated by substitution (3.4) and the permutation of the second and third rows of pattern (2.1b). The choice of the last factors of e_j and f_j is completely determined after expansion of the type (2.11) and some cyclic permutation of the parameters of the boundary isofactors.

Elementary special cases of $g_{\tilde{I},i}$ $(g_{b-\tilde{z},i}, g_{\tilde{z},\tilde{z}}$, and $g_{\tilde{z}+1,\tilde{z}})$ along with $g_{\tilde{I},\tilde{I}}$ leave no ambiguity for determination of the general case of Eq. (4.2).

The invariance of functions (3.7a), (3.7b) allows us to deduce the invariancy of the function

$$(a'b'y_0' i_0' a'' b'' \bar{y}_0'' \bar{i}_0'' \| ab\tilde{y}i;\tilde{I})$$
$$\times [(a+b+2)(a+1)(b+1)$$
$$\times b'!(a'+b'+1)!a''!(a''+b''+1)!]^{-1/2} \tag{4.8}$$

under the same 24 transformations of pattern (2.16). This invariancy leads, in fact, to the reduction formula of the boundary paracanonical isofactors as a function in eight independent parameters $(a', b', a'', b'', a, b, \tilde{I}, i)$ with three linear combinations fixed. For example, a polynomial expression with param-

TABLE I. Polynomials $g_{\tilde{I},i}$ for small values of $b-2\tilde{z}$.

$b-2\tilde{z}=1,\quad \tilde{z}=\frac{1}{2}(b-1);\quad \tilde{I},i=\frac{1}{2}(b\pm 1)$
$g_{(b+1)/2,(b+1)/2}=(a''+2)(b+1)(a''+b''-a+3)$
$-a(b-v)(b''+v+2)$
$b-2\tilde{z}=2,\quad \tilde{z}=\frac{1}{2}b-1;\quad \tilde{I},i=\frac{1}{2}b-1,\frac{1}{2}b,\frac{1}{2}b+1$
$k_{b/2+1,b/2}=k_{b/2,b/2-1}=\sqrt{2}$
$g_{b/2,b/2}=b'(b'-1)b(b+1)(a''+b''+v+5)^{(2)}$
$+2(b'-1)(b+1)(a''+b''+v+5)$
$\times(v+1)(a+b)(b''+v+2)$
$+(v+2)^{(2)}(a+b+1)^{(2)}(b''+v+3)^{(2)}$
$g_{b/2+1,b/2+1}=(b'+b''+v+3)^{(2)}(b+2)^{(2)}(a''+3)^{(2)}$
$-2(b'+b''+v+2)(b+1)$
$\times(a''+3)(a-1)(b-v)(b''+v+3)$
$+a^{(2)}(b''+v+3)^{(2)}(b-v)^{(2)}$
$g_{b/2+1,b/2}=a(b-v-1)(b''+v+2)$
$-b(b'+b''+v+2)(a''+3)$
$g_{b/2,b/2-1}=-(b'-1)(b+2)(a''+b''+v+5)$
$-(v+2)(b''+v+3)(a+b+1)$

eters $a,a+\tilde{z}-\tilde{I},i-\tilde{z}$ fixed may be obtained from Eqs. (4.2)–(4.8) after substitution

$$\begin{aligned} &a\to a+b'-a''+v,\quad b\to b-b'+a''-v,\\ &a'\to a',\quad b'\to a''-v,\quad a''\to b'+v,\quad b''\to b''. \end{aligned} \tag{4.9}$$

The polynomials $g_{\tilde{I},i}$ are invariant with respect to the substitution (3.4) and permutation of the second and third rows of pattern (2.16), when the denominator polynomials $g_{\tilde{I},\tilde{I}}$ remain unchanged after the permutation of the couples of the columns (3,4) and (5,6).

Those permutations of pattern (2.16) that transpose the parameters of two right columns with the remaining ones lead to other versions of the boundary region of paracanonical isofactors of the same classification type.

The structure of the null space (the ordered vanishing properties with decreasing multiplicity along with conserving orthogonality) of general paracanonical isofactors is caused particularly by the main (first) factor in expansion (2.24), by the factors $b'^{(b-\tilde{z}-\tilde{I})}$, or by zeros of the polynomial $g_{\tilde{I},i}$ [when the parameters q_{i5} or q_{i6} of pattern (2.16) are minimal, respectively].

Different choices of the boundary region lead to different expressions of isofactors and give different interpretations of the same null spaces. Contrary to the SU(3) canonical tensor operators,[20,21] the null spaces of the paracanonical tensor operators are not lexically ordered. The linearly dependent states appear from below, though some natural inequalities cut off the multiplicity labels from above. After cyclic permutations (2.17′), the null spaces of different nature exchange. Such symmetry is absent in the case of the SU(3) canonical tensor operators; in our case the conjugation symmetry is spoiled: the conjugation connects two alternative versions of the paracanonical classification. Otherwise the conjunction along with transposition $(a'b')\leftrightarrow(a''b'')$ [i.e., permutation (3.9)] does not change the type of paracanonical splitting. Therefore the symmetry group of general paracanonical isofactors includes six elements of 12 considered in Ref. 30.

The Gram–Schmidt process, begun from the opposite end (i.e., when the linearly dependent states appear from above), leads to the pseudocanonical system of orthonormal isofactors, defined by vanishing of isofactors with parameters

$$\tilde{J}-i>i'_0+z'+\bar{i}''_0-z''. \tag{4.10}$$

A conjecture may also be made about the structure of the boundary pseudocanonical isofactors

$$\begin{aligned} &(a'b'y'_0i'_0a''b''\bar{y}''_0\bar{i}''_0\|ab\tilde{y}i;\tilde{J})\\ &=\left[\frac{(a+1)(b+1)(a+b+2)b'^{(\tilde{J}-\tilde{z})}(a'+b'+1)^{(\tilde{J}-\tilde{z})}(i+\tilde{z})^{(-1)(\tilde{J}-\tilde{z})}}{a''^{(-1)(b-\tilde{J}-\tilde{z}+1)}(a''+b''+1)^{(-1)(b-\tilde{J}-\tilde{z}+1)}(b-\tilde{z}+i+1)^{(b-\tilde{J}-\tilde{z}+1)}}\right]^{1/2}\frac{K_{i,\tilde{J}}\,\bar{g}_{i,\tilde{J}}}{[\bar{g}_{\tilde{J}-1,\tilde{J}-1}\,\bar{g}_{\tilde{J},\tilde{J}}]^{1/2}}, \end{aligned} \tag{4.11}$$

for max $(|\tilde{z}|,|\bar{i}_z|)\leqslant\tilde{J}\leqslant i\leqslant\min(b-\tilde{z},a+\tilde{z},i'_0+\bar{i}''_0)$. Here

$$g_{i,\tilde{J}}=\det\begin{bmatrix} E_{\tilde{z},\tilde{z}}\cdots E_{\tilde{z},\tilde{J}-1} & E_{\tilde{z},i}\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\\ E_{\tilde{J},\tilde{z}}\cdots E_{\tilde{J},\tilde{J}-1} & E_{\tilde{J},i} \end{bmatrix}(\tilde{J}+\tilde{z})^{(-1)(\tilde{J}-\tilde{z})}\left[(i+\tilde{z})^{(-1)(\tilde{J}-\tilde{z})}K_{i,\tilde{J}}\prod_{j=\tilde{z}+1}^{\tilde{J}}\bar{e}_j\right]^{-1} \tag{4.12a}$$

$$\begin{aligned} &=\det\begin{bmatrix} F^{b-\tilde{z},b-\tilde{z}}\cdots F^{b-\tilde{z},\tilde{J}+1} & 0\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\\ F^{i,b-\tilde{z}}\cdots F^{i,\tilde{J}+1} & 1\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\\ F^{\tilde{J},b-\tilde{z}}\cdots F^{\tilde{J},\tilde{J}+1} & 0 \end{bmatrix}\\ &\times(b-\tilde{z}+1+i)^{(b-\tilde{z}-\tilde{J}+1)}\left[(2i+1)(b-\tilde{z}+\tilde{J}+1)^{(b-\tilde{z}-\tilde{J})}K_{i,\tilde{J}}\prod_{j=\tilde{J}+1}^{b-\tilde{z}-1}\bar{f}_j\right]^{-1} \end{aligned} \tag{4.12b}$$

are the polynomials with integer coefficients:

$$\bar{e}_j = b'^{(j-\tilde{z})}(a'+b'+1)^{(j-\tilde{z})}(a+b'+v+1)^{(j-\tilde{z})} \times(a+b'+b''+v+2)^{(j-\tilde{z})}[(j+\tilde{z})^{(-1)(j-\tilde{z})}]^2, \tag{4.13a}$$

$$\bar{f}_j = (a''-a-v)^{(-1)(b-\tilde{z}-j)} \times(a'+a''-a-v+1)^{(-1)(b-\tilde{z}-j)} \times(a''+1)^{(-1)(b-\tilde{z}-j)}(a''+b''+2)^{(-1)(b-\tilde{z}-j)} \times[(b-\tilde{z}+j+1)^{(b-\tilde{z}-j)}]^{1/2}. \tag{4.13b}$$

In the pseudocanonical case, the quantity of the type (4.8) is also invariant with respect to 24 permutations of pattern (2.16), though different versions of the boundary region lead to different schemes of the pseudocanonical classification. For example, the isofactors represented by Eq. (4.11) are invariant only with respect to substitutions (3.4) and (3.8). Sometimes [e.g., when the minimal values in pattern (2.16) are accepted by parameters q_{52}, q_{53}, q_{62}, or q_{63}] the indefinites 0/0 may appear in the right-hand side of Eq. (4.11).

It seems that different pseudocanonical and paracanonicals systems of isofactor are related by the same hook permutations[48] or substitutions.[49] Thus both paracanonical and pseudocanonical algebraic systems of isofactors are less symmetric than canonical ones with respect to the substitution group.[48,49]

V. LABELING OPERATORS FOR ORTHONORMAL COUPLING COEFFICIENTS OF SU(3)

The multiplicity problem for decomposition $(a'b')\times(a''b'')$ to (ab) may be solved with the help of the SU(3) invariant operator[28,29]

$$Q(\alpha,\beta)=\alpha\sum_{\rho\sigma}D_{\rho\sigma}E^{\{2\}}_{\sigma\rho}+\beta\sum_{\rho\sigma}D^{\{2\}}_{\rho\sigma}E^{\{2\}}_{\sigma\rho}, \tag{5.1}$$

where

$$D_{\rho\sigma}=E'_{\rho\sigma}-E''_{\rho\sigma},D^{\{2\}}_{\rho\sigma}=\frac{1}{2}\sum_{\tau}(D_{\rho\tau}D_{\tau\sigma}+D_{\tau\sigma}D_{\rho\tau}), \tag{5.2a}$$

$$E_{\rho\sigma}=E'_{\rho\sigma}+E''_{\rho\sigma},\quad E^{\{2\}}_{\sigma\rho}=\frac{1}{2}\sum_{\tau}(E_{\sigma\tau}E_{\tau\rho}+E_{\tau\rho}E_{\sigma\tau}) \tag{5.2b}$$

form the SU(3) irreducible tensors of rank (11) in the enveloping algebra of SU(3)′×SU(3)″. Particularly $D_{\rho\sigma}$ and $E_{\rho\sigma}$ (with $\Sigma_\rho D_{\rho\rho}=0$ and $\Sigma_\rho E_{\rho\rho}=0$) are the generators of SU(3)×SU(3).

The traceless part of $Q(\alpha,\beta)$ [with the eigenvalues denoted by $q(\alpha,\beta)$] is invariant for the even permutations belonging to the symmetry group $S_3\times S_2$ of the Clebsch–Gordan coefficients of SU(3).[30] For the odd permutations it conserves or changes the sign, if $\alpha=0$ or $\beta=0$, respectively.

The SU(3)-reduced matrix elements of the operators $D_{\rho\sigma}$ and $D^{\{2\}}_{\rho\sigma}$ may be found in the basis (2.2) with the help of the transposition formula[12,52] of the tensor and projection operators

$$T^{\lambda_1}_{\mu_1}P^{\lambda}_{\mu,\mathbf{\mu}}=\sum_{\underline{\lambda},\gamma,\underline{\mu}}\dim(\lambda)\dim^{-1}(\underline{\lambda})(\lambda\mu\lambda_1\mu_1|\underline{\lambda}\underline{\mu})^{\gamma} \times\sum_{\underline{\mu},\tilde{\underline{\mu}}}(\lambda\mathbf{\mu}\lambda_1\underline{\mu}_1|\underline{\lambda}\tilde{\underline{\mu}})_{\gamma}P^{\underline{\lambda}}_{\underline{\mu},\tilde{\underline{\mu}}}T^{\lambda_1}_{\underline{\mu}_1}, \tag{5.3}$$

where $\lambda=(ab)$, $\mu=yii_z$.

The explicit form of the projection operator $P^{\lambda}_{\mu,\mathbf{\mu}}$ is unnecessary. The operators $D_{\rho\sigma}$ $(\rho\neq\sigma)$ acting on $|a'b'\rangle_{\rm HWS}|a''b''\rangle_{\rm LWS}$ may be replaced by $\pm E_{\rho\sigma}$ (with + for $\rho>\sigma$ and − for $\rho<\sigma$) and thus they may be included into the projection operator. With considerable effort, many terms obtained on the first stage were united and the following representation of the operator $\alpha D_{\rho\sigma}+\beta D^{\{2\}}_{\rho\sigma}$ in the irreducible form was obtained[28]:

$$\begin{aligned}&[\alpha D^{(11)}_{y_1i_1m_1}+\beta(D^{\{2\}})^{(11)}_{y_2i_1m_1}]|(a'b')(a''b'')abyii_z\rangle_{-,+,\tilde{I}}\\
&=\sum_{\substack{\gamma,\underline{a},\underline{b}\\ y,\underline{i},\underline{i}_z}}\beta\dim(ab)\dim^{-1}(\underline{ab})(abyi11y_1i_1\|\underline{abyi};\gamma)(ii_zi_1m_1|\underline{ii_z})\sum_{\tilde{J}}(ab\tilde{y}\tilde{I}1101\|\underline{ab}\tilde{y}\tilde{J};\gamma)(\tilde{I}i_z10|\tilde{J}i_z)2^{-1/2}\\
&\times[\alpha/\beta+y'_0-\bar{y}''_0+2+\tfrac{1}{3}f_2(\underline{ab})-\tfrac{1}{3}f_2(ab)+(\tilde{I}-\tilde{J})(\tilde{I}+\tilde{J}+1)][a'+b''+2+(\tilde{J}-\tilde{I})(\tilde{I}+\tilde{J}+1)]\\
&+(ab\tilde{y}\tilde{I}1100\|\underline{ab}\tilde{y}\tilde{j};\gamma)6^{-1/2}\delta_{\tilde{I}\tilde{J}}\{[3(\bar{y}''_0-y'_0-2)+\tfrac{2}{3}f_2(ab)-\tfrac{2}{3}f_2(\underline{ab})][\alpha/\beta+\tfrac{1}{2}(\bar{y}''_0-y'_0)-1]\\
&+4\tilde{I}(\tilde{I}+1)-a'(a'+2)-b''(b''+2)+\delta_{a\underline{a}}\delta_{b\underline{b}}[\tfrac{1}{2}\tilde{y}^2-2-4\tilde{I}(\tilde{I}+1)]-2\tilde{z}^2(\delta_{\underline{a},a+1}\delta_{\underline{b},b+1}+\delta_{\underline{a},a-1}\delta_{\underline{b},b-1})\\
&-2(a+\tilde{z}+1)^2\delta_{\underline{b},b\pm2}-2(b-\tilde{z}+1)^2\delta_{\underline{a},a\pm2}\}+\delta_{a\underline{a}}\delta_{b\underline{b}}\delta_{\tilde{I}\tilde{J}}\{\tfrac{1}{9}6^{-1/2}f_3(ab)f_2^{-1/2}(ab)\delta_{\gamma1}+\delta_{\gamma2}f_2^{1/2}(ab)(-1)^{\varphi}\\
&\times[2ab(a+2)(b+2)(a+b+1)(a+b+3)]^{-1/2}[\tfrac{1}{81}f_3(ab)f_2^{-1}(ab)-\tfrac{4}{27}f_2^2(ab)-\tfrac{1}{9}f_2(ab)-\tfrac{2}{27}f_3(ab)\tilde{y}+\tilde{y}^2\\
&+\tfrac{1}{9}f_2(ab)\tilde{y}^2-\tfrac{1}{4}\tilde{y}^4+2\tilde{y}^2\tilde{I}(\tilde{I}+1)+\tfrac{4}{3}f_2(ab)\tilde{I}(\tilde{I}+1)-4\tilde{I}^2(\tilde{I}+1)^2]\})|(a'b')(a''b'')\underline{abyii_z}\rangle_{-,+,\tilde{J}},\end{aligned} \tag{5.4}$$

where $\tilde{y}=y'_0+\bar{y}''_0$, $\tilde{i}_z=i'_0-\bar{i}''_0$, $f_2(ab)$ and $f_3(ab)$ being the eigenvalues of the Casimir operators defined as

$$f_2(ab)=a^2+ab+b^2+3a+3b,\quad f_3(ab)=(a-b)(2a+b+3)(a+2b+3). \tag{5.5}$$

[For $\beta=0$ an elementary limit transition is necessary in Eq. (5.4).]

The use of the coupled basis $|\eta_{-,+,\tilde{I}}\rangle$ (or $|\eta_{+,-,\tilde{I}}\rangle$) is preferable because the reduced matrix elements in the right-hand side of Eq. (5.4) are expressed in terms of the SU(3) Clebsch–Gordan coefficients and the eigenvalues of the Casimir operators of the intrinsic subgroups and depend only on the intrinsic parameters of the bra and ket states [compare Eq. (45) of Ref. 53 or Eqs. (2.3) and (3.3) of Ref. 54 in the SU(3)⊃SO(3) case]. The corresponding isofactors of SU(3)⊃U(2) in the right-hand side of Eq. (5.4) [in the canonical labeling scheme, with $\gamma=1$ corresponding to the SU(3) generator matrix elements] are tabulated in Refs. 19, 55, and 56 (see also Ref. 57). [The factor $(-1)^{\varphi}$ is negative only in the phase system of Ref. 57.] Rather simple matrix elements of the generators $D_{\sigma\rho}$ are useful for motivation of the analytical inversion symmetry, because the corresponding matrix turns into the transposed one (with some factor).

Now the matrix elements of the SU(3) scalar operator (5.1) may be presented in the following tridiagonal form:

$$\begin{aligned}
&\langle \eta^{-,+,\tilde{J}} | Q(\alpha,\beta) | \eta_{-,+,\tilde{I}} \rangle \\
&\quad = \delta_{\tilde{I}\tilde{J}} \{ [-\tfrac{2}{27} f_2^2(ab) - \tfrac{1}{18} f_2(ab) - \tfrac{1}{27} f_3(ab)\tilde{y} + \tfrac{1}{2}\tilde{y}^2 + \tfrac{1}{18} f_2(ab)\tilde{y}^2 - \tfrac{1}{8}\tilde{y}^4]\beta + \tfrac{1}{8}\tilde{y}(a' - b'')(a' + b'' + 2) \\
&\qquad \times [\alpha + \beta(y_0' - \bar{y}_0'' + 2)] - \tfrac{1}{4}[\tfrac{1}{2}\tilde{y}^2 + \tfrac{2}{9} f_2(ab)] [3(y_0' - \bar{y}_0'' + 2)\alpha + \beta(a'(a'+2) + b''(b''+2) \\
&\qquad - \tfrac{3}{2}(y_0' - \bar{y}_0'' + 2)^2 - \tfrac{1}{2}\tilde{y}^2 + 2)] + \tfrac{1}{4}(a' - b'')(a' + b'' + 2)\tilde{I}^{-1}(\tilde{I}+1)^{-1}[\alpha + \beta(y_0' - \bar{y}_0'' + 2)][\tfrac{1}{2} f_3(ab) \\
&\qquad - \tfrac{1}{8}\tilde{y}^3 + \tfrac{1}{6}\tilde{y}(f_2(ab) + 3)] + \tfrac{3}{2}\tilde{I}(\tilde{I}+1)(y' - \bar{y}_0'' + 2)\alpha - \tilde{I}(\tilde{I}+1)\beta[\tfrac{3}{4}(y_0' - \bar{y}_0'' + 2)^2 - \tfrac{3}{4}\tilde{y}^2 \\
&\qquad - \tfrac{1}{2}a'(a'+2) - \tfrac{1}{2}b''(b''+2) - \tfrac{2}{3} f_2(ab) + 2\tilde{I}(\tilde{I}+1) - 1]\} - \delta_{\tilde{I}+1,\tilde{J}}[(\tilde{I} + \tilde{i}_z + 1)(\tilde{I} - \tilde{i}_z + 1) \\
&\qquad \times (\tilde{I} - \tilde{z} + 1)(\tilde{I} + \tilde{z} + 1)(a + \tilde{z} - \tilde{I})(a + \tilde{z} + \tilde{I} + 2)(b - \tilde{z} - \tilde{I})(b - \tilde{z} + \tilde{I} + 2)]^{1/2} \\
&\qquad \times \frac{(a' + b'' + 2\tilde{I} + 4)}{[(2\tilde{I}+1)(2\tilde{I}+3)]^{1/2}(\tilde{I}+1)} \left\{ \frac{1}{2}\alpha + \left[\frac{1}{2}(y_0' - \bar{y}_0'') - \tilde{I} \right]\beta \right\} \\
&\qquad - \delta_{\tilde{I}-1,\tilde{J}}[(\tilde{I} + \tilde{i}_z)(\tilde{I} - \tilde{i}_z)(\tilde{I} + \tilde{z})(\tilde{I} - \tilde{z})(a + \tilde{z} - \tilde{I} + 1)(a + \tilde{z} + \tilde{I} + 1)(b - \tilde{z} - \tilde{I} + 1)(b - \tilde{z} + \tilde{I} + 1)]^{1/2} \\
&\qquad \times \frac{(a' + b'' - 2\tilde{I} + 2)}{[(2\tilde{I}+1)(2\tilde{I}-1)]^{1/2}\tilde{I}} \left\{ \frac{1}{2}\alpha + \left[\frac{1}{2}(y_0' - \bar{y}_0'') + \tilde{I} + 1) \right]\beta \right\}. \qquad (5.6)
\end{aligned}$$

The $\tilde{I}$-independent terms of the diagonal matrix elements may be omitted. The parameter $\tilde{I}$ may be zero only when $a' = b''$. In region (2.10c) the nonvanishing matrix elements

$$\begin{aligned}
&\langle \eta^{-,+,\tilde{J}} | Q'(\alpha,\beta) | \eta_{-,+,B} \rangle \\
&\quad = R_{B-1,\tilde{J}} \langle \eta^{-,+,B-1} | Q(\alpha,\beta) | \eta_{-,+,B} \rangle \qquad (5.7)
\end{aligned}$$

also appear, and the action of operator (5.1) in a complete basis (2.2) may be represented as a sum of Eqs. (5.6) and (5.7). [We failed to eliminate explicitly the trace of the operator $Q(0,1)$.]

The eigenvalue problems of the operator (5.1) split into two more elementary problems for special integer values of $3\alpha/\beta$ for which some nondiagonal matrix element in the right-hand side of Eq. (5.6) disappears. The discriminant

$$q^2(\alpha,\beta) = [4f_2(ab) + 9]\alpha^2 - \tfrac{8}{3} f_3(ab)\alpha\beta + \tfrac{16}{9} f_2^2(ab)\beta^2 \qquad (5.8)$$

of the eigenvalue problem for the coupling $(ab)\times(11)$ to (ab) (with a double multiplicity of irreps) gives rational proper values of the traceless operator $Q(0,1)$. In this case the eigenstates of $Q(0,1)$ correspond to the canonical splitting scheme.[20,56] The positive eigenvalue $q(0,1) = \tfrac{4}{3} f_2(ab)$ corresponds to the state coupled with the help of the SU(3) generator matrix elements.

In general, the eigenstates of the operator $Q(0,1)$ belong to the symmetric or antisymmetric subspace of the direct product space $(a'b')\times(a''b'')$ (where $a' = a''$, $b' = b''$) in accordance with the Derome[30] pattern, i.e., it corresponds to a definite one-dimensional irrep of the above-mentioned symmetry group $S_3 \times S_2$ of the Clebsch–Gordan coefficients. The question whether concrete eigenstates belong to a symmetric or antisymmetric subspace remains, in general, open, as well as the problem of the dependence of the phases on this multiplicity label.

However, the coupled eigenstate of the Pluhař operator[18] [$Q(1,0)$ in our notations] is, in general, neither symmetric nor antisymmetric because it belongs to a reducible representation of the group $S_3 \times S_2$ (which is a sum of symmetric and antisymmetric irreps). In this way the independence of the phase factors in Eqs. (5.3) and (5.4) of Ref. 18 from the multiplicity label may be explained. Undoubtedly these phase factors are completely motivated in the multiplicity-free cases.[6]

It is evident that the elements of the transposed matrix (5.6) [with additional terms (5.7), if necessary] give the action of the Hermitian operator (5.1) in the dual basis (2.4). Then the solution of the eigenvalue problem [along with the expression (3.3) or (= 0.6) for overlaps] allows us to expand the $q(\alpha\beta)$-classified SU(3) coupled states in terms of the states (2.4), and thus to find boundary values of the $q(\alpha\beta)$-classified isofactors. Later, with the help of Eqs. (2.22) and (2.24) most general $q(\alpha,\beta)$-classified SU(3)⊃U(2) isofactors may be evaluated.

The use of the symmetry relations of the Clebsch–Gordan coefficients of SU(3) always allows us to omit additional expansions that usually appear simultaneously in Eqs. (2.24), (3.3), and (5.7). In different minimal coupled bases matrix elements of the operator (5.1) are related (up to trace and phases) by some substitutions of parameters.

Only a special case of Eq. (2.22) proportional to the $6j$ coefficient of SU(2)[6,15,17] [i.e., Eq. (4.4) of Ref. 18] is needed for special SU(3)⊃U(2) isofactors necessary for obtaining the SU(3)⊃SO(3) isofactors by means of the slightly modified Engeland[52] method. The realization of both SU(3)⊃SO(3) internal labeling operators in the Elliott[53] basis is given in Refs. 54 and 58.

Thus general SU(3)⊃SO(3) isofactors with completely solved inner and outer labeling problems of the repeating irreps may be evaluated.

VI. ON REALIZATION OF THE CANONICAL SPLITTING

The properties of the canonical SU(3) unit tensor operators $T(a''b'')$ are discussed in detail in Refs. 21, 31, and 32. The multiplicity label $\overset{*}{J}''$, which accepts the same values as the label $\mathbf{I}''$ in $|\eta_-,\mathbf{I}'',-\rangle$ if the linearly dependent states of the latter are chosen from above, may be used instead of the operator Gelfand–Weyl–Biedenharn pattern. The shifts $a-a'$ and $b-b'$ are completely determined by $\mathbf{i}''_z = \frac{1}{2}(a-a')$ and $\mathbf{z}' = \frac{1}{2}(a'-a) - v$ or $\mathbf{y}'' = y_0 - y'_0$.

It should be noted that the pseudocanonical SU(3) tensor operators obtained from $|\eta_-,\mathbf{I}'',-\rangle$ by means of the Schmidt process, beginning from below, have the same null-space structure as the canonical ones. However, the pseudocanonical tensor operators lack the majority of the above mentioned symmetries of the canonical or the paracanonical tensor operators.

The orthonormal canonical isofactors with the parameters

$$i'' + |i - i'| > a'' + b'' - \overset{*}{J}{}''_{\max} + \overset{*}{J}{}'', \tag{6.1a}$$

where

$$\overset{*}{J}{}''_{\max} = \min(a'' + \mathbf{z}'', b'' - \mathbf{z}'') \geqslant \overset{*}{J}{}'' \geqslant \max(|\mathbf{i}''_z|, |\mathbf{z}''|), \tag{6.1b}$$

vanish [cf. Eq. (1.32) of Ref. 21]. As a matter of fact, the canonical splitting is completely determined by (6.1) when the maximal values of $i'' = i''_m$ and $i' = i'$, where $i'_m - i''_m \geqslant |\hat{z}|$, $\hat{z} = \frac{1}{2}(b - a - v)$ [or, respectively, $i'' = i''_m$ and $i = i_m$, where $i_m - i''_m \geqslant |\hat{z}'|, \hat{z}' = \frac{1}{2}(b' - a' + v)$] are chosen.[21,49] This property of the canonical tensor operators is caused by the group generators included in their structure.

Vectors (2.1) with the parameters $\mathbf{i}' = \mathbf{i}'_z \to i'_m$, $\mathbf{y}' \to y'_m$, $\mathbf{i}'' = -\mathbf{i}''_z \to i''_m$, $\mathbf{y}'' \to y''_m$, and $\mathbf{I} \to \hat{I}$ $(i'_m \geqslant i''_m)$ form a complete nonorthogonal coupled basis

$$\begin{aligned}|\eta_{\uparrow,\downarrow,\hat{I}}\rangle &\equiv |(a'b')(a''b'')abyii_z\rangle_{\uparrow,\downarrow,\hat{I}}\\ &= P^{(ab)}_{yii_z;\hat{y}\hat{I}\hat{i}_z}|a'b'y'_m i'_m i'_m\rangle |a''b''y''_m i''_m - i''_m\rangle.\end{aligned} \tag{6.2}$$

The corresponding orthonormal basis obtained by means of the Schmidt process beginning from the lowest value of $\hat{I}$ is equivalent to the coupled basis constructed with the help of the canonical isofactors and labeled by

$$\overset{*}{J}{}'' = i'_m - i''_m + \overset{*}{J}{}''_{\max} - \hat{I} \tag{6.3}$$

at least when the external multiplicity of irreps coincides with the asymptotical multiplicity

$$\mathscr{M} = \min(q_{ij}) + 1 \quad (i = 2,3,\ j = 3,4,5,6). \tag{6.4}$$

[$\mathscr{M}$ is expressed in terms of the parameters of pattern (2.16) depending only on a'', b'', $a - a'$, $b - b'$—cf. Eq. (1.11) of Ref. 21.] The asymptotical and usual multiplicities coincide for $i'_m - i''_m \geqslant |\hat{z}|$ or $i_m - i''_m \geqslant |\hat{z}'|$.

As seen from Eq. (2.1), special bilinear combinations of SU(3) isofactors are needed to construct the states (6.2). However, the known analytical expressions[12,14] for the bilinear combinations of the isofactors in the right-hand side of Eq. (2.1) remain nonpolynomial for fixed $(a''b'')$ and corresponding shifts.

It is remarkable that the coupled basis (6.2) may be expanded in terms of the states (2.2) as follows:

$$|\eta_{\uparrow,\downarrow,\hat{I}}\rangle = \sum_{\tilde{I}} (-1)^{(a''+b')/2+\tilde{I}+\hat{I}}[(2\hat{I}+1)(2\tilde{I}+1)]^{1/2} \times \begin{Bmatrix} \alpha & \beta & \hat{I} \\ \gamma & \frac{1}{2}a & \tilde{I} \end{Bmatrix} |\eta_{-,+,\tilde{I}}\rangle. \tag{6.5}$$

In the right-hand side there appeared the $6j$ coefficient of SU(2) with the parameters

$$\alpha = \tfrac{1}{2}(b - b' + a'' - v), \quad \beta = \tfrac{1}{2}(a' - b'' + b - v),$$
$$\gamma = \tfrac{1}{2}(b - v).$$

Here the summation is taken over the states of the complete or overcomplete coupled basis. This expansion is equivalent to Weyl transformation[59] between different SU(2) subgroups in SU(3). Thus special isofactors

$$\begin{aligned}&((a'b'y'_m i'_m a''b''y''_m i''_m \| ab\hat{y}\hat{i}))^{-,+,\tilde{I}}\\ &\quad= (i'_m i'_m i''_m - i''_m | \hat{i} i'_m - i''_m)^{-1} (i'_0 i'_0 \bar{i}''_0 - \bar{i}''_0 | \tilde{I} i'_0 - \bar{i}''_0)\\ &\quad\quad\times(-1)^{(a''+b')/2+\tilde{I}+\hat{i}}[(2\hat{i}+1)(2\tilde{I}+1)]^{1/2}\\ &\quad\quad\times\begin{Bmatrix} \alpha & \beta & \hat{i} \\ \gamma & \frac{1}{2}a & \tilde{I} \end{Bmatrix}\end{aligned} \tag{6.6}$$

are found, which may also be expressed by means of Eq. (2.22). The double sum obtained has been taken for the selected extremal values of $\hat{i}$ or $\tilde{I}$ with the help of Eq. (14) of Ref. 37.

Thus Eqs. (6.5), (3.2b), and (2.20) or (3.5) (after some permutations of parameters) allow us to express the SU(3) canonical isofactors in the algebraic polynomial form by means of the Schmidt process beginning with $\hat{I}_{\min}$ (i.e., from the maximal null-space case, similarly to Ref. 21).

Otherwise, the dual coupled basis $|\eta^{\uparrow,\downarrow,\hat{I}}\rangle$ allows us to express the SU(3) canonical isofactors by means of the Schmidt process beginning with $\hat{I}_{\max}$, which corresponds to the minimal null space. For

$$\begin{aligned}\mathscr{M} = \min(&a'' - b' + b - v,\\ &b'' - a' + a + v, b'', b'' + v) + 1,\end{aligned} \tag{6.4'}$$

the dual coupled basis may be expanded as follows:

$$\begin{aligned}|\eta^{\uparrow,\downarrow,\hat{I}}\rangle = \sum_{i,I} &(-1)^{(a''+b')/2+\tilde{I}+\hat{I}}[(2\tilde{I}+1)(2\hat{I}+1)]^{1/2}\Bigg\{\delta_{\hat{i}\hat{I}} + \left(\frac{(\hat{I}+\hat{z})!(\hat{i}-\hat{z})!(\hat{i}-\hat{I}_{\min})!(\hat{I}+\hat{I}_{\min})!}{(\hat{i}+\hat{z})!(\hat{I}-\hat{z})!(\hat{I}-\hat{I}_{\min})!(\hat{i}+\hat{I}_{\min})!}\right)^{1/2}\\ &\times\left(\frac{(a+\hat{z}-\hat{I})!(a+\hat{z}+\hat{I}+1)!(b-\hat{z}-\hat{i})!(b-\hat{z}+\hat{i}+1)!}{(a+\hat{z}-\hat{i})!(a+\hat{z}+\hat{i}+1)!(b-\hat{z}-\hat{I})!(b-\hat{z}+\hat{I}+1)!}\right)^{1/2\,\mathrm{sign}\,v}\\ &\times\frac{(-1)^{\hat{i}-\hat{B}}(2\hat{i}+1)(\hat{B}+\hat{i}+1)!}{(\hat{i}-\hat{I})(\hat{i}+\hat{I}+1)(\hat{B}-\hat{I})!(\hat{B}+\hat{I}+1)!(\hat{i}-\hat{B}-1)!}\Bigg\}\begin{Bmatrix} \alpha & \beta & \hat{i} \\ \gamma & \frac{1}{2}a & \tilde{I} \end{Bmatrix}|\eta^{-,+,\tilde{I}}\rangle,\end{aligned} \tag{6.7}$$

where

$$\hat{z}=\tfrac{1}{2}(b-a-v), \quad |\hat{z}|\leqslant\hat{I}_{\min}=\tfrac{1}{2}(a'+b'-a''-b''),$$
$$\hat{I}\leqslant\widehat{B}=\tfrac{1}{2}(a'+b'-a''+b''+v-|v|), \quad \tilde{I}\leqslant\tfrac{1}{2}(a'+b''),$$
$$\hat{I}_{\max}=\min[\tfrac{1}{2}(a+b-|v|),\widehat{B}\,].$$

The second term in braces on the right-hand side of Eq. (6.7) does not vanish only when the basis $|\eta_{\uparrow,\downarrow,\hat{I}}\rangle$ is overcomplete, i.e., $\mathscr{M}=\min(b'',b''+v)+1$ and the recoupling matrix of SU(2) in the right-hand side of Eq. (6.5) is truncated. In this case the expansion coefficients of $|\eta^{\uparrow,\downarrow,\hat{I}}\rangle$ in terms of $|\eta^{-,+,\bar{I}}\rangle$ form a matrix inverse of the truncated one. Their derivation is shown in Appendix B.

In fact, Eq. (6.7) allows us to construct nonorthonormal isofactors that satisfy the boundary condition

$$(a'b'y'_m i'_m a''b''y''_m i''_m\|ab\hat{y}i)^{\uparrow,\downarrow,\hat{I}}=\delta_{\hat{I}i} \qquad (6.8)$$

for $i\leqslant\widehat{B}$ ($\hat{I}\leqslant\widehat{B}$).

When $\mathscr{M}$ is different from the one defined according to Eq. (6.4a), basis $|\eta_{-,+,\bar{I}}\rangle$ is overcomplete and therefore it is more convenient to use the basis $|\eta^{\hat{I}',\uparrow,\uparrow}\rangle$ expanded in terms of $|\eta^{\bar{I}',-,-}\rangle$ instead of $|\eta^{\uparrow,\downarrow,\hat{I}}\rangle$.

If the external multiplicity does not coincide with $\mathscr{M}$, the numerical orthogonalization of bases $|\eta_{\uparrow,\downarrow,\hat{I}}\rangle$ and $|\eta^{\uparrow,\downarrow,\hat{I}}\rangle$ does not always lead to the canonical isofactors. Sometimes the alternative version of the canonical classification appear, which may be associated with tensors of rank $(a'b')$.

Algebraic expressions for the matrix elements of $T(a''b'')$ obtained may be used in this region if the indefinite (vanishing) factors in the numerator and the denominator are eliminated. The knowledge of the properties of the denominator function[21] is undoubtedly very useful for the maximal simplification of algebraic expressions for the orthonormal isofactors of the canonical type.

VII. CONCLUSIONS

In this paper the pluralism of the external multiplicity problem for SU(3) is demonstrated. Depending on the situation, one may choose either the analytical biorthogonal systems or the algebraic or numerical orthonormal isofactors. The later may be labeled by the irrational (in general case) eigenvalues of the classifying operator or by the intrinsic isospins (by the Gelfand–Weyl–Biedenharn operator patterns, respectively). The required symmetry of isofactors and their additional selection rules may serve as arguments for the choice between the canonical and paracanonical splittings.

Completely analytical expression (in all 12 parameters) for the orthonormal isofactors of SU(3) seems impossible. By means of the proper Gram–Schmidt process, analytical expressions for the matrix elements of the canonical and paracanonical tensor operators may be obtained when at least the difference between the multiplicity label and its extremal (minimal or maximal) value (i.e., the number of steps of the Schmidt process) is fixed. For example, the relation between the denominator factors of the minimal null-space case and the maximal null-space case[32,33] corresponds in some aspects to the analytical inversion symmetry of the overlaps.

However, analytical expressions in the canonical case are much more complicated and sometimes the limit transitions are indispensable. When a sufficient number of parameters is fixed, the corresponding expressions accept algebraic-polynomial form. In all such cases the minimal biorthogonal systems remain the universal element of the optimal construction. Although the transformation between the minimal biorthogonal systems and the systems associated with canonical splitting has simple interpretation, the corresponding unitary transformation between the canonical and paracanonical tensor operators is not simple and not related elementary with the SU(2) recoupling matrix or with the Weyl transformation of the operator pattern.

Thus three versions of the canonical splitting [corresponding to fixed $(a''b'')$, (a',b'), or (ab), respectively] and two versions of the paracanonical splitting along with six versions of the pseudocanonical splitting give different algebraic systems of the SU(3) orthonormal isofactors which are determined by the additional selection rules, their null-space structure, and symmetries of isofactors. However, situations exist when different solutions are more convenient, e.g., the bilinear combinations of SU(3) isofactors in the right-hand side of Eq. (2.1) with $v=\mathrm{i}'=\mathrm{i}''=\mathrm{i}=0$ may be expressed as double sums only by the methods of Ref. 60. Such bilinear combinations appear as the expansion coefficients of the SU(3) ⊃ U(2) spherical functions. The specific external multiplicity label in this case is not simply correlated with those discussed above.

APPENDIX A: ON THE POLYNOMIAL FACTORS IN OVERLAPS

The renormalized overlaps [see Eqs. (3.9a) and (3.9b)] for the fixed $b-2\tilde{z}$ accept the following forms:

$$\begin{aligned}E_{\tilde{I},\tilde{J}}=\,&[(\tilde{I}+\hat{i}_z)^{(\tilde{I}-\tilde{J})}(\tilde{I}-\bar{i}_z)^{(\tilde{I}-\tilde{J})}(i'_0+\bar{i}''_0-\tilde{J})^{(\tilde{I}-\tilde{J})}(i'_0+\bar{i}''_0+\tilde{I}+1)^{(\tilde{I}-\tilde{J})}\\&\times(a+\hat{z}-\tilde{J})^{(\tilde{I}-\tilde{J})}(a+\tilde{z}+\tilde{I}+1)^{(\tilde{I}-\tilde{J})}(b-\tilde{z}-\tilde{J})^{(\tilde{I}-\tilde{J})}/(\tilde{I}-\tilde{z})!(\tilde{J}-\tilde{z})!]^{1/2}\\&\times\sum_{x_1,x_2}(-1)^{a'+b''+b+x_1+x_2}\binom{\tilde{J}-\tilde{z}}{x_1}\binom{b-\tilde{z}-\tilde{I}}{x_2}\frac{(b+x_1-x_2+1)}{(b+x_1+1)^{(b-\tilde{z}-\tilde{I}+1)}}\\&\times(\tilde{I}-\tilde{z})^{(x_1)}(b-\tilde{z}-\tilde{J}-x_2)^{(-1)(\tilde{J}-\tilde{z}-x_1)}(a+b'+v+1)^{(x_1+x_2)}(b'+b''-b+v)^{(-1)(x_2)}\\&\times(b'+b''+v+1)^{(-1)(x_1)}(b-\tilde{z}+\tilde{I}+1)^{(x_2)}(b-\tilde{z}+\tilde{J}+1)^{(x_2)}(\tilde{J}+\tilde{z})^{(-1)(x_1)}(\tilde{J}-\bar{i}_z)^{(\tilde{J}-\tilde{z}-x_1)}\\&\times(\tilde{I}+\bar{i}_z)^{(-1)(b-\tilde{z}-\tilde{I}-x_2)}(a+\tilde{z}-\tilde{J})^{(-1)(\tilde{J}-\tilde{z}-x_1)}(a+\tilde{z}+\tilde{I}+1)^{(-1)(b-\tilde{z}-\tilde{I}-x_2)}(\tilde{I}+\tilde{z})^{(-1)(b-\tilde{z}-\tilde{I}-x_2)}\\&\times(i'_0+\bar{\imath}''_0-\tilde{I})^{(b-\tilde{z}-\tilde{I}-x_2)}(i'_0+\bar{i}''_0+\tilde{J}+1)^{(\tilde{J}-\tilde{z}-x_1)},\end{aligned} \qquad (A1)$$

$$
\begin{aligned}
F^{\tilde I,\tilde J} = {} & [(\tilde I+\bar i_z)^{(\tilde I-\tilde J)}(\tilde I-\bar i_z)^{(\tilde I-\tilde J)}(i_0'+\bar i_0''-\tilde J)^{(\tilde I-\tilde J)}(i_0'+\bar i_0''+\tilde I+1)^{(\tilde I-\tilde J)} \\
& \times (a+\tilde z-\tilde J)^{(\tilde I-\tilde J)}(a+\tilde z+\tilde I+1)^{(\tilde I-\tilde J)}(b-\tilde z-\tilde J)^{(\tilde I-\tilde J)}/(\tilde I-\tilde z)!(\tilde J-\tilde z)!]^{1/2} \\
& \times \sum_{x_1,x_2} \binom{\tilde J-\tilde z}{x_1}\binom{b-\tilde z-\tilde I}{x_2} \frac{(b+x_1-x_2+1)}{(b-x_2)^{(-1)(\tilde J-\tilde z+1)}} \\
& \times (\tilde I-\tilde z)^{(x_1)}(b-\tilde z-\tilde J-x_2)^{(-1)(\tilde J-\tilde z-x_1)}(a''-a-v)^{(-1)(x_1+x_2)}(\tilde I+\tilde z)^{(-1)(x_1)} \\
& \times (\tilde J+\tilde z)^{(-1)(x_1)}(\tilde I+\tilde i_z)^{(-1)(b-\tilde z-\tilde I-x_2)}(i_0'+\bar i_0''+\tilde I+1)^{(-1)(b-\tilde z-\tilde I-x_2)} \\
& \times (a+\tilde z-\tilde J)^{(-1)(\tilde J-\tilde z-x_1)}(a+\tilde z+\tilde J+1)^{(-1)(b-\tilde z-\tilde I-x_2)}(i_0'+\bar i_0''-\tilde J)^{(-1)(\tilde J-\tilde z-x_1)} \\
& \times (b-\tilde z+\tilde I+1)^{(x_2)}(\tilde J-\bar i_z)^{(\tilde J-\tilde z-x_1)}(b-\tilde z+\tilde J+1)^{(\tilde J-\tilde z-x_1)}(a'+a''+b-v+3)^{(x_2)}(a'+a''-v+2)^{(x_1)},
\end{aligned}
\tag{A2}
$$

where $\tilde I \geqslant \tilde J$,

$$
\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}, \quad A^{(-1)(x)} = (A+x)^{(x)} = (A+1)\cdots(A+x).
$$

Some efforts are necessary in order to demonstrate that $E_{\tilde I,\tilde J}$ and $F^{\tilde I,\tilde J}$ are equivalent to polynomials in free parameters. The factors $(b+x_1+1)^{(b-\tilde z-\tilde I+1)}$ and $(b-x_2)^{(-1)(\tilde J-\tilde z+1)}$ in denominators may be canceled in an elementary manner for $\tilde I = b-\tilde z$ or $\tilde J=\tilde z$ and less easily for $\tilde I = b - \tilde z - 1$ or $\tilde J = \tilde z + 1$. In order to prove the general case of Eq. (A1) being polynomial, it is sufficient to show the absence of the poles caused by the denominator factors of the type $\mathscr{L}(\tilde I + \tilde z + 1 \leqslant \mathscr{L} \leqslant b - \tilde z + \tilde J + 1)$. Really, both the invariance transformations of functions (3.7a), (3.7b) and the compensating transformations with coefficients of the type (2.11) (along with usual symmetries of isofactors and relabeling of parameters) lead to the mutually excluding sets of possible poles.

APPENDIX B: TWO RELATIONS BETWEEN THE 6*j* COEFFICIENTS OF SU(2)

Similarly to Eq. (6.5) the following expansion of the minimal coupled bases may be found:

$$
|\eta_{+,-,\mathbf{I}}\rangle = \sum_{\tilde I} (-1)^{a'+b'+a''+b''}[(2\tilde I+1)(2\mathbf{I}+1)]^{1/2} \begin{Bmatrix} \frac12(b+a'-b''-v) & \frac12(b-v) & \tilde I \\ \frac12(b+a''-b'-v) & \frac12 a & \mathbf{I} \end{Bmatrix} |\eta_{-,+,\tilde I}\rangle. \tag{B1}
$$

When the basis $|\eta_{-,+,\tilde I}\rangle$ (but not $|\eta_{+,-,\mathbf{I}}\rangle$) is overcomplete, the SU(2) recoupling matrix in the right-hand side is truncated by the additional condition $\mathbf{I} \leqslant \frac12(a''+b')$. It is clear that the recoupling matrix truncated by the conditions $\tilde I \geqslant B$ and $\mathbf{I}\leqslant\frac12(a''+b')$ gives the inverse expansion of the linearly independent states $|\eta_{-,+,\tilde I}\rangle$ in terms of $|\eta_{+,-,\mathbf{I}}\rangle$. Equation (B1) along with (2.19c) gives the expansion in terms of linearly independent states. In such a way the inverse of the truncated SU(2) recoupling matrix is found.

Equations (B1) and (2.19c) lead to the following relation between the 6*j* coefficient:

$$
\begin{Bmatrix} a & b & e \\ d & c & f \end{Bmatrix} = \sum_g \frac{(-1)^{e-B}(2g+1)(B+g)!}{(g-e)(g+e+1)(g-B)!(B-e-1)!(B+e)!} \frac{\nabla(abe)\nabla(dce)}{\nabla(abg)\nabla(dcg)} \begin{Bmatrix} a & b & g \\ d & c & f \end{Bmatrix}, \tag{B2}
$$

where

$$
a+d \geqslant b+c, \quad f \leqslant \min(a+c,b+d) + \max(|a-b|,|c-d|) - B.
$$

The substitutions $a \to a-1$, $e \to -e-1$ [see Eq. (29.21) of Ref. 36] allow to obtain the following equation:

$$
\begin{Bmatrix} a & b & e \\ d & c & f \end{Bmatrix} = \sum_g \frac{(-1)^{e-B'}(2g+1)(B'+e+1)!}{(e-g)(e+g+1)(B'-g)!(B'+g+1)!(e-B'-1)!} \frac{\nabla(abg)\nabla(dce)}{\nabla(abe)\nabla(dcg)} \begin{Bmatrix} a & b & g \\ d & c & f \end{Bmatrix}, \tag{B3}
$$

where $f \leqslant f_{\min} - g_{\min} + B'$. Equation (B3) is valid if the 6*j* coefficients with the parameters $f_{\min} = d-b$, $g_{\min} = a-b$, or $f_{\min} = c-a$, $g_{\min} = c-d$ are not vanishing. The symmetries of the 6*j* coefficients allow us to cover the remaining cases of the recoupling matrices truncated from above. Of course, the solutions in different regions may be joined. Identity (B3) allows us to expand the linearly dependent states of $|\eta_{-,+,\tilde I}\rangle$ in terms of the linearly independent states restricted from above.Thus the inverse SU(2) recoupling matrix, truncated in different ways, may be found explicitly.

[1] T. A. Brody, M. Moshinsky, and I. Renero, J. Math. Phys. **6**, 1540 (1965).

[2] R. T. Sharp and H. C. von Baeyer, J. Math. Phys. **7**, 1105 (1966).

[3] G. Ponzano, Nuovo Cimento A **41**, 142 (1966).

[4] M. Resnikoff, J. Math. Phys. **8**, 63 (1967).

[5] S. J. Ališauskas, Liet. Fiz. Rinkinys. **12**, 721 (1972).

[6] S. J. Ališauskas, Liet. Fiz. Rinkinys. **18**, 701 (1978) [Engl. transl. Sov. Phys.–Coll. Litovsk. Fiz. Sb. **18** (6), 6 (1978)].

[7] Hou Pei-yu, Sci. Sinica **14**, 367 (1965).

[8] C. K. Chew and R. T. Sharp, Nucl. Phys. B **2**, 697 (1967).

[9] L. A. Shelepin and V. P. Karasev, Yad. Fiz. **5**, 221 (1967); V. P. Karasev and L. A. Shelepin, *ibid*, **7**, 1131 (1968); see also V. P. Karasev and N. P. Shchelok, Trudy Fiz. Inst. Lebedev **173**, 115 (1986).

[10] R. M. Asherova and Yu. F. Smirnov, Nucl. Phys. B **134**, 399 (1986).

[11] I. S. Guseva, Yu. F. Smirnov, V. N. Tolstoy, and Yu. I. Kharitonov, Leningrad Inst. Nucl. Phys. preprint, No. 678, 1981 (in Russian).

[12] Z. Pluhař, Yu. F. Smirnov, and V. N. Tolstoy, Charles University Preprint, Praha, 1981 (in Russian).

[13] V. N. Tolstoy, Yu. F. Smirnov, and Z. Pluhař, in *Group Theoretical Methods in Physics*, Proceedings of the 1982 Zvenigorod International Seminar (Nauka, Moscow, 1983), Vol. 2, pp. 523–530 [Engl. transl. (Harwood, Chur, 1985), Vol. 1, pp. 77–86].

[14] A. U. Klimyk and A. M. Gavrilik, J. Math. Phys. **20**, 1624 (1979); A. U. Klimyk, *Matrix Elements and Clebsch–Gordan Coefficients of the Group Representations* (Naukova Dumka, Kiev, 1979) (in Russian).

[15] S. J. Ališauskas, Liet. Fiz. Rinkinys. **22** (2), 13 (1988); Sov. Phys.–Collection **22** (2), 9 (1988).

[16] S. J. Ališauskas, Sov. J. Part. Nucl. **14**, 563, 1336 (1983).

[17] S. Ališauskas, J. Phys. A: Math. Gen. **20**, 1045 (1987).

[18] Z. Pluhař, Yu. F. Smirnov, and V. N. Tolstoy, J. Phys. A: Math. Gen. **19**, 19 (1986).

[19] K. T. Hecht, Nucl. Phys. **62**, 1 (1965).

[20] L. C. Biedenharn, A. Giovannini, and J. D. Louck, J. Math. Phys. **8**, 691 (1967); J. D. Louck, Am. J. Phys. **38**, 3 (1970).

[21] L. C. Biedenharn, M. A. Lohe, and J. D. Louck, J. Math. Phys. **26**, 1458 (1985).

[22] R. Le Blanc and D. J. Rowe, J. Phys. A: Math. Gen. **19**, 2913 (1986).

[23] S. J. Ališauskas, in *Group Theoretical Methods in Physics*, Proceedings of the 1985 Jūrmala International Seminar (Nauka, Moscow, 1986), Vol. 2, pp. 68–76.

[24] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).

[25] C. K. Chew and R. T. Sharp, Can. J. Phys. **44**, 2789 (1966).

[26] I. S. Guseva, Yu. F. Smirnov, V. N. Tolstoy, and Yu. I. Kharitonov, Leningrad Inst. Nucl. Phys. preprint, No. 837, 1983.

[27] In Ref. 18 the former solution of the boundary value problem (Ref. 15) is rediscovered, but the corresponding method to evaluate the boundary values of the orthonormal isofactors is not optimal. Some peculiarities of the complete (but not overcomplete) coupled basis case have been also overlooked in Ref. 18. The two additional summations used in Ref. 18 are unnecessary if the classifying operator is represented in the minimal coupled basis.

[28] S. J. Ališauskas and P. P. Kulish, Zap. Nauch. Sem. LOMI **145**, 3 (1985); J. Sov. Math. **35**, 2563 (1986).

[29] R. T. Sharp, J. Math. Phys. **16**, 2050 (1975).

[30] J. R. Derome, J. Math. Phys. **7**, 612 (1966); **8**, 714 (1967).

[31] J. D. Louck and L. C. Biedenharn, J. Math. Phys. **11**, 2368 (1970).

[32] L. C. Biedenharn, J. D. Louck, E. Chacón, and M. Ciftan, J. Math. Phys. **13**, 1957 (1972); L. C. Biedenharn and J. D. Louck, *ibid*. **13**, 1985 (1972).

[33] J. D. Louck, M. A. Lohe, and L. C. Biedenharn, J. Math. Phys. **16**, 2408 (1975); M. A. Lohe, L. C. Biedenharn, and J. D. Louck, *ibid*. **18**, 1883 (1977).

[34] J. P. Draayer and Y. Akiyama, J. Math. Phys. **14**, 1904 (1973).

[35] J. A. Castilho Alcarás and V. Vanagas, J. Math. Phys. **28**, 1995 (1987).

[36] A. P. Jucys and A. A. Bandzaitis, *Theory of Angular Momentum in Quantum Mechanics* (Mokslas,Vilnius, 1977), 2nd ed. (in Russian).

[37] S. J. Ališauskas, A.-A. A. Jucys, and A. P. Jucys, J. Math. Phys. **13**, 1329 (1972).

[38] This result probably remained unnoticed by R. Le Blanc and D. J. Rowe, J. Phys. A: Math. Gen. **19**, 1093 (1986), who derived with considerable efforts special boundary cases of these isofactors. Namely, the possibilities of the maximal simplification of the boundary cases have been taken into account in Ref. 37 when the alternative expressions for the SU(n) isofactors [different from those of E. Chacón, M. Ciftan, and L. C. Biedenharn, J. Math. Phys. **13**, 577 (1972)] were derived. It should be noted that the result M. Hage Hassan, J. Phys. A: Math. Gen. **16**, 2891 (1983), concerning the multiplicity-free isofactors is incompatible with the results of Ref. 37 and Chacón *et al*. The orthogonality and completeness of the isofactors cannot be reached with a lesser number of sums as it is presented in Ref. 37 and Chacón *et al.*

[39] C. Quesne, J. Phys. A: Math. Gen. **20**, 769, 785, 2259 (1987).

[40] S. J. Ališauskas, V. V. Vanagas, and A. P. Jucys, Dokl. Akad. Nauk SSSR **197**, 804 (1971).

[41] J. J. Sullivan, J. Math. Phys. **14**, 387 (1973).

[42] The corresponding Eq. (2.5) of Ref. 6 should be corrected.

[43] P. O. Löwdin, Rev. Mod. Phys. **36**, 966 (1964)

[44] J. Shapiro, J. Math. Phys. **6**, 1680 (1965).

[45] The corresponding Eq. (11) of Ref. 15 should be corrected.

[46] Up to the choice of the boundary regions, the expansion coefficient in the right-hand side of Eq. (4.4) of Ref. 18 is equivalent to (4.1) of Ref. 6 or (6) of Ref. 15. Expansion (4.7) of Ref. 18 is equivalent to (2) of Ref. 15. [Probably, the purely infinitesimal derivation of Eq. (4.4) of Ref. 18, perhaps with the help of their remarkable Eq. (2.10), is original.] Both equations were generalized[16] [respectively, Eqs. (47) and (49)] for the expansion of arbitrary SU(n) $\supset$U($n-1$) isofactors in terms of their boundary values in a certain hyperplane. The corresponding equations give the explicit expressions for the proper nonorthogonal isofactors of SU(n) in terms of special recoupling coefficients of the subgroup SU($n-1$) when the cardinality of the boundary region coincides with the external multiplicity of irreps.

[47] I. Daubechies, A. Grossmann, and Y. Meyer, J. Math. Phys. **27**, 1271 (1986).

[48] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **5**, 1730 (1964).

[49] S. J. Ališauskas and A. P. Jucys, J. Math. Phys. **8**, 2250 (1967).

[50] The choice of phases in the corresponding expressions of Refs. 12 and 13 is questionable.

[51] L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U.P., Cambridge, 1966).

[52] T. Engeland, Nucl. Phys. **72**, 68 (1965).

[53] J. P. Elliott, Proc. R. Soc. Ser. A **245**, 562 (1958).

[54] H. De Meyer, G. Van den Berghe, and J. Van der Jeught, J. Math. Phys. **26**, 3109 (1985).

[55] J. G. Kuriyan, D. Lurie, and A. J. Macfarlane, J. Math. Phys. **6**, 722 (1965).

[56] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **6**, 1847 (1965).

[57] In the corresponding Tables of Refs. 19 and 55 some corrections are necessary.

[58] B. R. Judd, W. Miller, I. Patera, and P. Winternitz, J. Math. Phys. **15**, 1787 (1974); M. Moshinsky, I. Patera, R. T. Sharp, and P. Winternitz, Ann. Phys. (NY) **95**, 139 (1975); A. Partensky and C. Quesne, J. Math. Phys. **20**, 2014 (1979).

[59] E. Chacón and M. Moshinsky, Phys. Lett. **23**, 567 (1966).

[60] E. Z. Norvaišas and S. J. Ališauskas, Liet. Fiz. Rinkinys. **15**, 501 (1975); Sov. Phys.-Collection **15** (4), 1 (1975).

Unitary lowest weight representations of the noncompact supergroup OSp(2*n*/2*m*,*R*)

M. Günaydin and S. J. Hyun
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 15 March 1988; accepted for publication 29 June 1988)

The oscillator construction of the unitary irreducible lowest (highest) weight representations of the noncompact supergroup $\mathrm{OSp}(2n/2m,R)$ with even subgroup $\mathrm{SO}(2n)\times\mathrm{Sp}(2m,R)$ is given. In particular, simple rules for determining the $\mathrm{SO}(2n)\times\mathrm{Sp}(2m,R)$ decomposition of the unitary lowest weight representations of $\mathrm{OSp}(2n/2m,R)$ are derived.

I. INTRODUCTION

Lie groups play a fundamental role in the formulation of physical theories. Over the last decade Lie supergroups and Lie superalgebras have come to play equally important roles in theoretical physics. They underlie all supersymmetric theories such as superstring and supergravity theories. In practically all the theories in which supersymmetry enters at a fundamental level the related supergroups turn out to be noncompact. For example, the Lie supergroups that contain the space-time symmetry groups such as Poincaré, anti-de Sitter, and conformal groups are all noncompact. Therefore the relevant unitary representations of such noncompact supergroups are all infinite dimensional. Furthermore, one is in general interested in positive energy unitary representations of space-time supergroups. These positive energy unitary representations are said to be of the lowest weight type, since the energy generator belonging to the noncompact superalgebra has a spectrum bounded from below. The lowest weight unitary representations are related to the highest weight representations by a simple involution.

A general method for constructing oscillatorlike unitary representations of noncompact groups and noncompact supergroups was given in Refs. 1 and 2, respectively. These representations are all of the lowest weight type, and for space-time groups and supergroups they correspond to the positive energy unitary representations. The methods of Refs. 1 and 2 were further developed and applied to supergravity and superstring theories.[3–6] In Ref. 3, the oscillator construction of the positive energy unitary representations of the seven-dimensional, $N=4$ anti-de Sitter supergroup $\mathrm{OSp}(8^*/4)\simeq\mathrm{OSp}(6,2/4)$ was given. Furthermore, the spectrum at the S^4 compactification[7] of the 11-dimensional supergravity[8] was shown to fit into an infinite set of short supermultiplets of $\mathrm{OSp}(8^*/4)$ (see Ref. 3). In Ref. 4, the unitary lowest weight representations of the four-dimensional $N=8$ anti-de Sitter supergroup $\mathrm{OSp}(8/4,R)$ were constructed, and the spectrum of the S^7 compactification[9] of the 11-dimensional supergravity was fitted into an infinite set of short supermultiplets of $\mathrm{OSp}(8/4,R)$. The spectrum of the S^5 compactification of the chiral $N=2$ supergravity was first obtained and fitted into unitary supermultiplets of the five-dimensional, $N=8$ anti-de Sitter supergroup $\mathrm{SU}(2,2/4)$ in Ref. 5. Somewhat later the results of Ref. 5 were confirmed by the results of Ref. 10 using differential geometrical methods. In Ref. 6, a complete classification of finite-dimensional conformal superalgebras in two dimensions was given and their lowest weight unitary representations studied. Using the results of Ref. 6, it was shown that the light-cone actions of superstring theories in $d=10$ can be interpreted as singleton field theories of certain three-dimensional anti-de Sitter superalgebras.[11]

In the references cited above the unitary lowest weight representations of noncompact superalgebras L are constructed over the super Fock space of bosonic and fermionic oscillators, which transform in the fundamental representation of a maximal compact subsuperalgebra L_0. The superalgebra L has a Jordan decomposition (three-grading) with respect to L_0. The method readily gives the full content of the unitary lowest weight representations of the noncompact superalgebra L in terms of the finite-dimensional representations of its maximal compact subsuperalgebra L_0. All noncompact Lie groups that admit lowest weight unitary representations have a Jordan structure with respect to their maximal compact subgroups.[12] However, this is not the case for noncompact Lie supergroups.[13] Recently the oscillator method has been generalized to noncompact supergroups that have a Kantor structure (five-grading) with respect to a compact subsupergroup of maximal rank.[13] This generalization allows one to construct unitary lowest weight representations of all simple noncompact supergroups of the classical type whose even subgroups are in the form of a direct product of a simple noncompact group with a compact group. The simple superalgebras of classical type, which include the exceptional and strange superalgebras, have been classified by Kac.[14] In addition to the references cited above, there have appeared in the literature papers studying representations of $\mathrm{OSp}(2n/2m,R)$ for special values of n and m, using different methods.[15]

In this paper we shall generalize the results of Refs. 2 and 4, and give a detailed study of the unitary lowest weight representations of $\mathrm{OSp}(2n/2m,R)$, which has a Jordan structure with respect to its maximal compact subsupergroup $\mathrm{U}(m/n)$. In particular, we shall give simple rules for determining the decomposition of a unitary representation of $\mathrm{OSp}(2n/2m,R)$ with respect to its even subgroup $\mathrm{Sp}(2m,R)\times\mathrm{SO}(2n)$. We illustrate these rules with several examples.

The plan of our paper is as follows: In Sec. II we give the oscillator construction of the representations of the compact

0022-2488/88/112367-10$02.50

group SO($2n$) and the rules for determining the Gelfand–Zetlin and Dynkin labels of an irreducible representation (irrep) of SO($2n$) from its lowest weight state. Section III summarizes the unitary lowest weight representations of Sp($2m$,R) following Ref. 4. Then in Sec. IV we give the oscillator construction and a detailed study of the unitary lowest weight representations of OSp($2n/2m$,R).

II. OSCILLATOR CONSTRUCTION OF THE REPRESENTATIONS OF SO(2*n*)

The oscillator construction of the representations of SO(8) was given in Ref. 4. The results of Ref. 4 can be extended in a rather straightforward manner to the group SO($2n$) for arbitrary n. The only subtlety that arises in this extension is the qualitative difference between even $n = 2k$ and odd $n = 2k + 1$. The orthogonal groups SO($2n$) with odd n have complex representations, while for even n the representations are all real. In this section we shall give the general oscillator construction of the representations of SO($2n$) for arbitrary n. We shall see below that the difference between even and odd n cases is elegantly reflected in the oscillator construction.

The Lie algebra of SO($2n$) has a Jordan decomposition (three-grading) with respect to its maximal subalgebra L_0 of U(n) (see Refs. 4 and 16):

$$L = L_{-1} \oplus L_0 \oplus L_{+1} . \tag{2.1}$$

The generators of SO($2n$) can be realized as bilinears of an arbitrary number $f = 2p + \varepsilon$ ($\varepsilon = 0,1$) of fermionic oscillators transforming in the fundamental representation of U(n):

$$\begin{aligned} A_{\mu\nu} &= \boldsymbol{\alpha}_\mu \cdot \boldsymbol{\beta}_\nu - \boldsymbol{\alpha}_\nu \cdot \boldsymbol{\beta}_\mu + \varepsilon \gamma_\mu \gamma_\nu , \\ A^{\mu\nu} &= \boldsymbol{\alpha}^\mu \cdot \boldsymbol{\beta}^\nu - \boldsymbol{\alpha}^\nu \cdot \boldsymbol{\beta}^\mu + \varepsilon \gamma^\mu \gamma^\nu = -A^\dagger_{\mu\nu} , \\ M^\mu_\nu &= \boldsymbol{\alpha}^\mu \cdot \boldsymbol{\alpha}_\nu - \boldsymbol{\beta}_\nu \cdot \boldsymbol{\beta}^\mu + (\varepsilon/2)(\gamma^\mu \gamma_\nu - \gamma_\nu \gamma^\mu) . \end{aligned} \tag{2.2}$$

The parameter ε takes on the value 0 or 1 depending on whether we have an even or odd number of fermionic oscillators, respectively. The expressions of the type $\boldsymbol{\alpha}^\mu \cdot \boldsymbol{\beta}_\nu$ or $\boldsymbol{\alpha}^\mu \cdot \boldsymbol{\alpha}_\nu$ are a short form for a summation over a family of p oscillators $\alpha(r)$ and $\beta(r)$ ($r = 1,2,...,p$):

$$\begin{aligned} \boldsymbol{\alpha}_\mu \cdot \boldsymbol{\beta}_\nu &\equiv \sum_{r=1}^{p} \alpha_\mu(r) \beta_\nu(r) , \\ \boldsymbol{\alpha}^\mu \cdot \boldsymbol{\alpha}_\nu &\equiv \sum_{r=1}^{p} \alpha^\mu(r) \alpha_\nu(r) . \end{aligned} \tag{2.3}$$

The oscillators $\alpha_\mu(r)$, $\beta_\mu(r)$, and γ_μ satisfy the canonical anticommutation rules

$$\begin{aligned} &\{\alpha_\mu(r), \alpha^\nu(s)\} = \delta^\nu_\mu \delta_{rs} , \\ &\{\beta_\mu(r), \beta^\nu(s)\} = \delta^\nu_\mu \delta_{rs} , \\ &\{\gamma_\mu, \gamma^\nu\} = \delta^\nu_\mu , \\ &\{\alpha_\mu(r), \alpha_\nu(s)\} = \{\beta_\mu(r), \beta_\nu(s)\} = \{\gamma_\mu, \gamma_\nu\} = 0 , \\ &\{\alpha_\mu(r), \beta^\nu(s)\} = \{\alpha_\mu(r), \gamma^\nu\} = \{\beta_\mu(r), \gamma^\nu\} = 0 , \\ &\{\alpha_\mu(r), \beta_\nu(s)\} = \{\alpha_\mu(r), \gamma_\nu\} = \{\beta_\mu(r), \gamma_\nu\} = 0 , \end{aligned} \tag{2.4}$$

where $\mu, \nu, ... = 1,...,n$ and $r,s = 1,...,p$.

The nonvanishing commutators of SO($2n$) are

$$\begin{aligned} [M^\mu_\nu, M^\rho_\sigma] &= \delta^\rho_\nu M^\mu_\sigma - \delta^\mu_\sigma M^\rho_\nu , \\ [A_{\mu\nu}, M^\rho_\sigma] &= \delta^\rho_\nu A_{\mu\sigma} - \delta^\rho_\mu A_{\nu\sigma} , \\ [A^{\mu\nu}, M^\rho_\sigma] &= \delta^\mu_\sigma A^{\nu\rho} - \delta^\nu_\sigma A^{\mu\rho} , \\ [A_{\mu\nu}, A^{\rho\sigma}] &= -\delta^\sigma_\mu M^\rho_\nu + \delta^\rho_\mu M^\sigma_\nu - \delta^\rho_\nu M^\sigma_\mu + \delta^\sigma_\nu M^\rho_\mu . \end{aligned} \tag{2.5}$$

To construct the irreducible representations (irreps) of SO($2n$) one first chooses a set of states $|\Omega\rangle$ in the Fock space of the fermionic oscillators, which transforms irreducibly under the U(n) subgroup and is annihilated by the operators $A_{\mu\nu}$ belonging to the L_{-1} space. Then acting on $|\Omega\rangle$ repeatedly with the operators $A^{\mu\nu}$ belonging to the L_{+1} space one generates a set of states that form the basis R of an irrep of SO($2n$):

$$R = \{|\Omega\rangle, A^{\mu\nu}|\Omega\rangle, A^{\mu\nu}A^{\rho\lambda}|\Omega\rangle, ...\} . \tag{2.6}$$

Since the fermionic oscillators anticommute, we have $(L_{+1})^k = 0$ for $k > nf/2$, and hence the representation space R is finite dimensional. We shall refer to the set of states $|\Omega\rangle$ as the lowest weight state of the corresponding irrep of SO($2n$).

The Fock vacuum $|0\rangle$ is defined to be the state annihilated by all the annihilation operators $\alpha_\mu(r)$, $\beta_\mu(r)$, and γ_μ. If we have only one set of oscillators ($p = 0$ and $\varepsilon = 1$), then the only possible lowest weight states that transform irreducibly under the U(n) subgroup generated by M^μ_ν are the Fock vacuum

$$|0\rangle \tag{2.7a}$$

and the "one-particle" state

$$\gamma^\mu|0\rangle . \tag{2.7b}$$

On the other hand, for two sets of oscillators, i.e., $p = 1$ and $\varepsilon = 0$, we have ($n + 2$) nonequivalent irreducible lowest weight states. They are

$$\begin{aligned} &|0\rangle , \\ &\alpha^\mu|0\rangle , \\ &\alpha^\mu\alpha^\nu|0\rangle , \\ &\quad\vdots \\ &\underbrace{\alpha^\mu\alpha^\nu\cdots\alpha^\rho}_{n \text{ copies}}|0\rangle , \end{aligned} \tag{2.8a}$$

and the symmetric tensor state

$$(\alpha^\mu\beta^\nu + \alpha^\nu\beta^\mu)|0\rangle . \tag{2.8b}$$

One can, of course, construct lowest weight states by replacing the oscillators α^μ in (2.8a) by β^μ. However, the resulting lowest weight states are equivalent to those of (2.8a). If we have $2p$ sets of oscillators, then the possible lowest weight states are those that can be obtained by tensoring p copies of the states of the type (2.8a) and (2.8b). For an odd number ($2p + 1$) of sets of oscillators the possible lowest weight states in the Fock space are those obtained by tensoring the states (2.7a) and (2.7b) with p copies of the states of type (2.8a) and (2.8b).

The U(n) = SU(n)×U(1) transformation properties of the lowest weight states $|\Omega\rangle$ can be conveniently labeled by using Young tableaux of U(n) and the U(1) quantum number of $|\Omega\rangle$. The U(1) quantum number of $|\Omega\rangle$ generated by $Q_F = M^{\mu}_{\mu}$ is uniquely determined by the U(n) Young tableau of $|\Omega\rangle$ and the number f of fermionic oscillators, since

$$Q_F = M^{\mu}_{\mu} = N_F - \tfrac{1}{2}nf, \qquad (2.9)$$

where N_F is the number operator of all the oscillators. For example, the vacuum $|0\rangle$ is an SU(n) singlet with

$$Q_F|0\rangle = -\tfrac{1}{2}nf|0\rangle\,. \qquad (2.10)$$

The U(n) Young tableaux of the possible lowest weight states $|\Omega\rangle$ have at most f columns. By taking a larger number f of fermionic oscillators one can construct a lowest weight state $|\Omega\rangle$ corresponding to an arbitrary Young tableau of U(n) with some definite U(1) charge.

The full U(n) content of an irrep of SO($2n$) with a given lowest weight state $|\Omega\rangle$ can be obtained by tensoring the Young tableau of $|\Omega\rangle$ with the Young tableaux of the symmetric tensor products of the operators $A^{\mu\nu}$. The Young tableau of $A^{\mu\nu}$ is simply $\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline\end{array}$, and since $[A^{\mu\nu}, A^{\lambda\rho}] = 0$, one needs to consider only the symmetric tensor products of

$$A^{\mu\nu} \approx \begin{array}{|c|}\hline \; \\ \hline \; \\ \hline\end{array}\,, \quad A^{\mu\nu}A^{\lambda\rho} \approx \begin{array}{|c|c|}\hline \; & \; \\ \hline \; & \; \\ \hline\end{array} \oplus \begin{array}{|c|}\hline \; \\ \hline \; \\ \hline \; \\ \hline \; \\ \hline\end{array}\,, \ldots\,. \qquad (2.11)$$

In tensoring the Young tableau of $|\Omega\rangle$ with the symmetric powers of $\begin{array}{|c|}\hline \; \\ \hline \; \\ \hline\end{array}$ one has to keep in mind that the resulting Young tableaux have at most f columns. The states corresponding to Young tableau with more than f columns do not exist in the Fock space of f fermionic oscillators. For details on this point we refer to Ref. 4.

The most common labelings for the irreps of SO($2n$) are the Dynkin and Gelfand–Zetlin labelings. It is a rather straightforward exercise to determine the labels of an irrep of SO($2n$) with lowest weight state $|\Omega\rangle$ by studying its U(n) content. Here we shall simply give the Dynkin and Gelfand–Zetlin labels of an irrep of SO($2n$) in terms of the U(n) Young tableaux labels of the corresponding lowest weight state $|\Omega\rangle$. We denote the Young tableaux (YT) for the irreducible representations of U(n) as $[l_1, l_2, \ldots, l_n]_{\rm YT}$, where l_i denotes the number of boxes in the ith row of the corresponding tableaux. The Gelfand–Zetlin (GZ) and Dynkin (D) labels of the irreps of SO($2n$) will be denoted as $(r_1, r_2, \ldots, r_n)_{\rm GZ}$ and $(n_1, n_2, \ldots, n_n)_{\rm D}$, respectively. In Table I we give the Gelfand–Zetlin and Dynkin labeling of the irreps of SO($2n$) for even $n = 2k$ corresponding to given lowest weight states with U(n) Young tableaux $[l_1, \ldots, l_n]_{\rm YT}$. Table II gives the corresponding labeling in the case of SO($2n$) for odd $n = 2k + 1$.

TABLE II. The U(n) Young tableau label of a lowest weight state $|\Omega\rangle$ and the Gelfand–Zetlin and Dynkin labels of the corresponding irreducible representation of SO($2n$) for odd $n = 2k + 1$. Note that $f = 2p + \varepsilon$.

U(n) Young tableau of the lowest weight state $\lvert\Omega\rangle$	$[l_1, l_2, l_3, \ldots, l_n]_{\rm YT}$
Gelfand–Zetlin labeling of the irrep of SO($2n$)	$\left(\frac{f}{2} - l_n, \frac{f}{2} - l_{n-1}, \ldots, \frac{f}{2} - l_2, l_1 - \frac{f}{2}\right)_{\rm GZ}$
Dynkin labeling of the irrep of SO($2n$)	$(l_{n-1} - l_n, l_{n-2} - l_{n-1}, \ldots,$ $l_2 - l_3, f - l_1 - l_2, l_1 - l_2)_{\rm D}$

To every lowest weight state $|\Omega\rangle$ of an irrep of SO($2n$) there corresponds a highest weight state $|{\rm HWS}\rangle$, which is annihilated by the operators $A^{\mu\nu}$ belonging to the L^{+1} space and transforms irreducibly under U(n). If we denote the U(n) Young tableau of a lowest weight state $|\Omega\rangle$ as $(l_1, l_2, \ldots, l_n)_{\rm YT}$, then the U($n$) Young tableau of the corresponding highest weight state $|{\rm HWS}\rangle$ is given as follows:

$$|{\rm HWS}\rangle \to [f - l_n, f - l_{n-1}, \ldots f - l_1]_{\rm YT}, \quad \text{for even } n = 2k\,, \qquad (2.12a)$$

and

$$|{\rm HWS}\rangle \to [f - l_n, f - l_{n-1}, \ldots f - l_2, l_1]_{\rm YT}, \quad \text{for } n = 2k + 1\,, \qquad (2.12b)$$

where $f = 2p + \varepsilon$ is the total number of fermionic oscillators. In Table III we give the Gelfand–Zetlin and Dynkin labels of an irrep of SO($2n$) corresponding to a given highest weight state $|{\rm HWS}\rangle$.

By the oscillator method outlined above one can construct all the irreps of SO($2n$), spinorial as well as tensorial.

TABLE I. The U(n) Young tableau label of a lowest weight state $|\Omega\rangle$ and the Gelfand–Zetlin and Dynkin labels of the corresponding irreducible representation of SO($2n$) for even $n = 2k$. Note that $f = 2p + \varepsilon$.

U(n) Young tableau of the lowest weight state $\lvert\Omega\rangle$	$[l_1, l_2, l_3, \ldots, l_n]_{\rm YT}$
Gelfand–Zetlin labeling of the irrep of SO($2n$)	$\left(\frac{f}{2} - l_n, \frac{f}{2} - l_{n-1}, \ldots, \frac{f}{2} - l_1\right)_{\rm GZ}$
Dynkin labeling of the irrep of SO($2n$)	$(l_{n-1} - l_n, l_{n-2} - l_{n-1}, \ldots,$ $l_1 - l_2, f - l_1 - l_2)_{\rm D}$

TABLE III. The U(n) label of a highest weight state and the Gelfand–Zetlin and Dynkin labels of the corresponding irrep of SO($2n$) for arbitrary n. Note that $f = 2p + \varepsilon$.

U(n) Young tableaux label of the highest weight state $\lvert{\rm HWS}\rangle$	$[k_1, k_2, \ldots, k_n]_{\rm YT}$
Gelfand–Zetlin label of the irrep of SO($2n$)	$\left(k_1 - \frac{f}{2}, k_2 - \frac{f}{2}, \ldots, k_n - \frac{f}{2}\right)_{\rm GZ}$
Dynkin label of the irrep of SO($2n$)	$(k_1 - k_2, k_2 - k_3, \ldots, k_{n-1} - k_n, k_{n-1} + k_n - f)_{\rm D}$

[Therefore, strictly speaking, we should be talking about Spin($2n$). Since at the Lie algebra level there is no distinction, we shall continue using the notation SO($2n$).] If we have an odd number $f = 2p + 1$ (i.e., $\varepsilon = 1$) of families of fermionic oscillators, we obtain spinorial irreps of SO($2n$) in general. This is most obvious from the Gelfand–Zetlin labeling given above. For example, for $f = 1$ (i.e., $p = 0$ and $\varepsilon = 1$) one obtains the two irreducible spinor representations of SO($2n$) (see Refs. 4 and 6).

To make our conventions for the various labeling clearer we give in Tables IV and V the decompositions of some of the lower-dimensional irreps of SO(8) and SO(10).

III. UNITARY HIGHEST (LOWEST) WEIGHT REPRESENTATIONS OF Sp(2*m*,*R*)

Before giving the general construction of the unitary highest weight representations of OSp($2n/2m,R$) we shall review the oscillator construction of the unitary highest weight representations of Sp($2m,R$) (see Refs. 1, 4, and 13).

TABLE IV. Some of the lower-dimensional irreducible representations of SO(8).

Number of oscillators	Young tableaux of lowest weight vector	SO(8)$_{\rm GZ}$	SO(8)$_{\rm D}$	Dimension
$f = 1$:	$\lvert 0\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$	(0,0,0,1)	8_s
	$\lvert\square\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2})$	(0,0,1,0)	8_c
$f = 2$:	$\lvert 0\rangle$	(1,1,1,1)	(0,0,0,2)	35_s
	$\lvert\square\rangle$	(1,1,1,0)	(0,0,1,1)	56_v
	$\left\lvert\begin{array}{c}\square\\\square\end{array}\right\rangle$	(1,1,0,0)	(0,1,0,0)	28
	$\left\lvert\begin{array}{c}\square\\\square\\\square\end{array}\right\rangle$	(1,0,0,0)	(1,0,0,0)	8_v
	$\left\lvert\begin{array}{c}\square\\\square\\\square\\\square\end{array}\right\rangle$	(0,0,0,0)	(0,0,0,0)	1
	$\lvert\square\square\rangle$	(1,1,1, − 1)	(0,0,2,0)	35_c
$f = 3$:	$\lvert 0\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{3}{2},\frac{3}{2})$	(0,0,0,3)	112_s
	$\lvert\square\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{3}{2},\frac{1}{2})$	(0,0,1,2)	224_{sc}
	$\left\lvert\begin{array}{c}\square\\\square\end{array}\right\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{1}{2},\frac{1}{2})$	(0,1,0,1)	160_s
	$\left\lvert\begin{array}{c}\square\\\square\\\square\end{array}\right\rangle$	$(\frac{3}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$	(1,0,0,1)	56_c
	$\left\lvert\begin{array}{c}\square\\\square\\\square\\\square\end{array}\right\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$	(0,0,0,1)	8_s
	$\lvert\square\square\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{3}{2},-\frac{1}{2})$	(0,0,2,1)	224_{cs}
	$\left\lvert\begin{array}{ll}\square&\square\\\square&\end{array}\right\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{1}{2},-\frac{1}{2})$	(0,1,1,0)	160_c
	$\left\lvert\begin{array}{ll}\square&\square\\\square&\\\square&\end{array}\right\rangle$	$(\frac{3}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2})$	(1,0,1,0)	56_s
	$\left\lvert\begin{array}{ll}\square&\square\\\square&\\\square&\\\square&\end{array}\right\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2})$	(0,0,1,0)	8_c
	$\lvert\square\square\square\rangle$	$(\frac{3}{2},\frac{3}{2},\frac{3}{2},-\frac{3}{2})$	(0,0,3,0)	112_c

TABLE V. Some of the lower-dimensional irreducible representations of SO(10).

Number of oscillators	Young tableaux of lowest weight vector	SO(10)$_{\rm GZ}$	SO(10)$_{\rm D}$	Dimension
$f = 1$:	$\lvert 0\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2})$	(0,0,0,1,0)	$\overline{16}$
	$\lvert\square\rangle$	$(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$	(0,0,0,0,1)	16
$f = 2$:	$\lvert 0\rangle$	(1,1,1,1, − 1)	(0,0,0,2,0)	$\overline{126}$
	$\lvert\square\rangle$	(1,1,1,1,0)	(0,0,0,1,1)	210
	$\left\lvert\begin{array}{c}\square\\\square\end{array}\right\rangle$	(1,1,1,0,0)	(0,0,1,0,0)	120
	$\left\lvert\begin{array}{c}\square\\\square\\\square\end{array}\right\rangle$	(1,1,0,0,0)	(0,1,0,0,0)	45
	$\left\lvert\begin{array}{c}\square\\\square\\\square\\\square\end{array}\right\rangle$	(1,0,0,0,0)	(1,0,0,0,0)	10
	$\left\lvert\begin{array}{c}\square\\\square\\\square\\\square\\\square\end{array}\right\rangle$	(0,0,0,0,0)	(0,0,0,0,0)	1
	$\lvert\square\square\rangle$	(1,1,1,1,1)	(0,0,0,0,2)	126

The highest weight representations of Sp($2m,R$) are related by a simple involution to the lowest weight representations. For the oscillator construction the term "lowest weight representations" is more appropriate, since the generator for Sp($2m,R$) whose spectrum is bounded (from below) is essentially the number operator of the bosonic oscillators.

The noncompact group Sp($2m,R$) has a Jordan decomposition with respect to its maximal compact subgroup U(m). Its Lie algebra L can be decomposed as a vector space direct sum:

$$L = L_{-1} \oplus L_0 \oplus L_{+1} = S_{ij} \oplus I^i_j \oplus S^{ij}, \tag{3.1}$$

where $i,j = 1,...,m$ and I^i_j are the generators of the U(m) subgroup. The nonvanishing commutators of Sp($2m,R$) in the above basis are[1,4,13]

$$\begin{aligned} [S_{ij},S^{kl}] &= \delta^l_j I^k_i + \delta^k_i I^l_j + \delta^k_j I^l_i + \delta^l_i I^k_j , \\ [I^i_j,S^{kl}] &= \delta^k_j S^{il} + \delta^l_j S^{ik} , \\ [I^i_j,S_{kl}] &= -\delta^i_k S_{jl} - \delta^i_l S_{jk} , \\ [I^i_j,I^k_l] &= \delta^k_j I^i_l - \delta^i_l I^k_j . \end{aligned} \tag{3.2}$$

The generators of Sp($2m,R$) can be realized as bilinears of an arbitrary number $f = 2p + \varepsilon$ ($p = 0,1,2,..., \varepsilon = 0,1$) of bosonic oscillators transforming in the fundamental representation of U(m):

$$\begin{aligned} S_{ij} &= \mathbf{a}_i\cdot\mathbf{b}_j + \mathbf{a}_j\cdot\mathbf{b}_i + \varepsilon c_i c_j , \\ S^{ij} &= \mathbf{a}^i\cdot\mathbf{b}^j + \mathbf{a}^j\cdot\mathbf{b}^i + \varepsilon c^i c^j , \\ I^i_j &= \mathbf{a}^i\cdot\mathbf{a}_j + \mathbf{b}_j\cdot\mathbf{b}^i + (\varepsilon/2)(c^i c_j + c_j c^i) , \end{aligned} \tag{3.3}$$

where $\mathbf{a}_i\cdot\mathbf{b}_j$ represents the sum $\Sigma^p_{r=1} a_i(r) b_j(r)$, etc. The bosonic oscillators satisfy the canonical commutation relations

$$\begin{aligned}&[a_i(r),a^j(s)] = [b_i(r),b^j(s)] = \delta^j_i\delta_{rs} ,\\ &[c_i,c^j] = \delta^j_i ,\\ &[a_i(r),a_j(s)] = [a_i(r),b_j(s)] = [a_i(r),c_j] = 0 ,\\ &[a_i(r),b^j(s)] = [a_i(r),c^j] = [b_i(r),c^j] = 0 ,\\ &[c_i,c_j] = [b_i(r),b_j(s)] = [c_i,b_j(r)] = 0 ,\end{aligned} \tag{3.4}$$

where $r = 1,2,...,p$, $i, j = 1,2,...,m$. The bosonic annihilation $(a_i(r),b_i(r),c_i)$ and creation operators $(a^i(r),b^i(r),c^i)$ transform covariantly and contravariantly, respectively, under the $\mathrm{U}(m)$ subgroup generated by I^i_j.

Every unitary lowest weight representation of $\mathrm{Sp}(2m,R)$ can be constructed over the Fock space of these bosonic oscillators as follows. One considers a set of states $|\Omega\rangle$ in Fock space transforming in a definite representation of $\mathrm{U}(m)$ and annihilated by all the operators S_{ij} belonging to the L_{-1} space. Then acting on $|\Omega\rangle$ repeatedly by the operators S^{ij} belonging to the L_{+1} space one generates an infinite set of states that form the basis of a unitary lowest weight representation of $\mathrm{Sp}(2m,R)$. If the lowest weight state $|\Omega\rangle$ transforms irreducibly under $\mathrm{U}(m)$, then the corresponding unitary representation of $\mathrm{Sp}(2m,R)$ is also irreducible. The unitary lowest weight irreducible representations of $\mathrm{Sp}(2m,R)$ can therefore be uniquely labeled by the $\mathrm{U}(m)$ labels of their lowest weight states.[1,4,13] If we have one set of bosonic oscillators, i.e., $p = 0$ and $\varepsilon = 1$, then in the corresponding Fock space there are only two nonequivalent irreducible lowest weight states, namely, the Fock vacuum

$$|0\rangle \tag{3.5a}$$

and the one-particle state

$$c^i|0\rangle . \tag{3.5b}$$

On the other hand, if we have two sets of oscillators ($p = 1$ and $\varepsilon = 0$), then the possible irreducible lowest weight states are states of the form

$$a^i a^j \cdots a^k|0\rangle , \tag{3.6a}$$

which correspond to symmetric tensors of arbitrary rank of $\mathrm{U}(m)$, and the antisymmetric tensor state

$$(a^ib^j - a^jb^i)|0\rangle . \tag{3.6b}$$

One can construct lowest weight states of the form (3.6a) using b-type oscillators only. However, they are all equivalent to the ones given above. If we have p pairs of the bosonic oscillators, then the possible lowest weight states are those that can be obtained by tensoring p copies of the states of the form (3.6a) and (3.6b). For an odd number $f = 2p + 1$ of oscillators the possible lowest weight states will be given by tensoring the states of the form (3.5a) and (3.5b) with p copies of the states of the form (3.6a) and (3.6b). We shall denote the $\mathrm{U}(m) = \mathrm{SU}(m)\times\mathrm{U}(1)$ transformation properties of the lowest weight states $|\Omega\rangle$ by their $\mathrm{U}(m)$ Young tableaux and their $\mathrm{U}(1)$ quantum numbers, whose generator is

$$\begin{aligned}&Q_B = I^i_i = \mathbf{a}^i\cdot\mathbf{a}_i + \mathbf{b}_i\cdot\mathbf{b}^i + (\varepsilon/2)(c^ic_i + c_ic^i) ,\\ &Q_B = \mathbf{a}^i\cdot\mathbf{a}_i + \mathbf{b}^i\cdot\mathbf{b}_i + \varepsilon c^ic_i + \tfrac{1}{2}fm ,\\ &Q_B = N_B + \tfrac{1}{2}fm ,\end{aligned} \tag{3.6c}$$

where N_B is the number operator of all the $f = 2p + \varepsilon$ bosonic oscillators. Since the bosonic creation operators all commute with each other, the Young tableaux of the lowest weight states $|\Omega\rangle$ can have at most $f = 2p + \varepsilon$ rows. By choosing f large enough one can construct a lowest weight state $|\Omega\rangle$ that transforms in any given representation of $\mathrm{SU}(m)$ and has a definite $\mathrm{U}(1)$ quantum number that depends on f (see Refs. 4 and 13). The operators S^{ij} belonging to the L_{+1} space transform in the symmetric tensor representation of $\mathrm{U}(m)$ with Young tableau $\square\square$. Since they commute with each other, the higher powers of S^{ij} correspond to symmetric tensor products of the representation $\square\square$ of $\mathrm{U}(m)$:

$$S^{ij} \approx \square\square , \quad S^{ij}S^{kl} \approx \begin{smallmatrix}\square&\square\\ \square&\square\end{smallmatrix} \oplus \square\square\square\square ,\ldots . \tag{3.7}$$

Thus to calculate the full $\mathrm{U}(m)$ content of a unitary lowest weight representation one needs to tensor the symmetric powers of the representation $\square\square$ of $\mathrm{U}(m)$ with the irrep of $\mathrm{U}(m)$ corresponding to the lowest weight state $|\Omega\rangle$. In decomposing the tensor products into irreducible representations of $\mathrm{U}(m)$ one has to keep in mind that the allowed irreps have at most $f = 2p + \varepsilon$ rows in their Young tableaux. For details on this point we refer to Ref. 4.

As an example, consider the case when $p = 0$ and $\varepsilon = 1$, i.e., when we have one set of bosonic oscillators c_i and c^i. Then as stated above there exist only two states in the Fock space that are annihilated by the operators $S_{ij} = c_ic_j$ and that transform irreducibly under the $\mathrm{U}(m)$ generated by $I^i_j = \frac{1}{2}(c^ic_j + c_jc^i)$. These are the Fock vacuum $|0\rangle$ and the one-particle state $c^k|0\rangle$:

$$\begin{aligned}&S_{ij}|0\rangle = 0, \quad I^i_j|0\rangle = \tfrac{1}{2}\delta^i_j|0\rangle ,\\ &S_{ij}c^k|0\rangle = 0, \quad I^i_jc^k|0\rangle = \delta^k_jc^i|0\rangle + \tfrac{1}{2}\delta^i_jc^k|0\rangle .\end{aligned} \tag{3.8}$$

The unitary representation of $\mathrm{Sp}(2m,R)$ with lowest weight vector $|0\rangle$ and $f = 1$ has the following $\mathrm{U}(m)$ content (indicated by Young tableaux):

$$\begin{aligned}&|0\rangle \approx 1, \quad \text{with } Q_B = m/2 ,\\ &S^{ij}|0\rangle \approx \square\square ,\\ &S^{ij}S^{kl}|0\rangle \approx \square\square\square\square\\ &\quad\vdots \qquad\quad \vdots ,\\ &\underbrace{S^{ij}\cdots S^{kl}}_{n\text{ times}}|0\rangle \approx \underbrace{\square\square\cdots\square\square}_{2n\text{ boxes}} .\end{aligned} \tag{3.9}$$

Similarly the unitary representation with the lowest weight vector $c^k|0\rangle$ has the following $\mathrm{U}(m)$ content:

$$\begin{aligned}&c^k|0\rangle \approx \square ,\\ &S^{ij}c^k|0\rangle \approx \square\square\square ,\\ &\quad\vdots \qquad\quad \vdots\\ &\underbrace{S^{ij}\cdots S^{lq}}_{n\text{ times}}c^k|0\rangle \approx \underbrace{\square\square\square\cdots\square}_{(2n+1)\text{ boxes}} .\end{aligned} \tag{3.10}$$

These two unitary lowest weight representations are the singleton representations of $\mathrm{Sp}(2m,R)$ (see Ref. 4). For $m = 2$, they give us the singleton representations of the covering group Sp(4,R) of the four-dimensional anti-de Sitter group SO(3,2), which were first discovered by Dirac.[17]

IV. UNITARY HIGHEST (LOWEST) WEIGHT REPRESENTATIONS OF THE NONCOMPACT SUPERGROUP OSp(*2n/2m,R*)

The noncompact supergroup OSp($2n/2m,R$) has a Jordan structure with respect to its maximal compact subsupergroup U(m/n). The even subgroup U(m)$\times$U(n) of U(m/n) is simply the subgroup with respect to which the even subgroup O($2n$)$\times$Sp($2m,R$) has a Jordan decomposition. The Lie superalgebra of OSp($2n/2m,R$) can be realized as bilinears of superoscillators transforming in the fundamental representation of U(m/n) (see Refs. 2 and 4). These superoscillators are defined as follows:

$$\xi_A(r) = \begin{pmatrix} a_i(r) \\ \alpha_\mu(r) \end{pmatrix}, \quad \eta_A(r) = \begin{pmatrix} b_i(r) \\ \beta_\mu(r) \end{pmatrix}, \quad \zeta_A = \begin{pmatrix} c_i \\ \psi_\mu \end{pmatrix},$$
$$\xi^A(r) = \begin{pmatrix} a^i(r) \\ \alpha^\mu(r) \end{pmatrix}, \quad \eta^A(r) = \begin{pmatrix} b^i(r) \\ \beta^\mu(r) \end{pmatrix}, \quad \zeta^A = \begin{pmatrix} c^i \\ \psi^\mu \end{pmatrix}, \tag{4.1}$$

where $A,B,... = 1,2,...,m+n$. The first m components of these superoscillators are bosonic and the remaining n components are fermionic. Superannihilation (and supercreation) operators ξ_A,η_A,ζ_A (and ξ^A,η^A,ζ^A) transform in the covariant (and contravariant) fundamental representation of U(m/n), respectively. They satisfy the supercommutation relations

$$\begin{aligned}
&[\xi_A(r),\xi^B(s)\} = [\eta_A(r),\eta^B(s)\} = \delta_A^B\delta_{rs}\,,\\
&[\zeta_A,\zeta^B\} = \delta_A^B\,,\\
&[\xi_A(r),\xi_B(s)\} = [\eta_A(r),\eta_B(s)\} = [\zeta_A,\zeta_B\} = 0\,,\\
&[\xi_A(r),\eta_B(s)\} = [\xi_A(r),\zeta_B\} = [\eta_A(r),\zeta_B\} = 0\,,\\
&[\xi_A(r),\eta^B(s)\} = [\xi_A(r),\zeta^B\} = [\eta_A(r),\zeta^B\} = 0\,,
\end{aligned} \tag{4.2}$$

where the superbracket [, } represents an anticommutator among two fermionic indices and a commutator, otherwise. The operators corresponding to the Jordan decomposition of the Lie superalgebra OSp($2n/2m,R$),

$$\mathrm{OSp}(2n/2m,R) \simeq S_{AB} \oplus M^A_B \oplus S^{AB} \simeq L_{-1} \oplus L_0 \oplus L_{+1}\,,$$

have the following realization as bilinears of an arbitrary number $f = 2p + \varepsilon$ of superoscillators[2,4]:

$$\begin{aligned}
S_{AB} &= \boldsymbol{\xi}_A\cdot\boldsymbol{\eta}_B + \boldsymbol{\eta}_A\cdot\boldsymbol{\xi}_B + \varepsilon\zeta_A\zeta_B\,,\\
S^{BA} &= S^\dagger_{AB} = \boldsymbol{\xi}^B\cdot\boldsymbol{\eta}^A + \boldsymbol{\eta}^B\cdot\boldsymbol{\xi}^A + \varepsilon\zeta^B\zeta^A\,,\\
M^A_B &= \boldsymbol{\xi}^A\cdot\boldsymbol{\xi}_B + (-1)^{\deg A \deg B}\boldsymbol{\eta}_B\cdot\boldsymbol{\eta}^A\\
&\qquad + (\varepsilon/2)[\zeta^A\zeta_B + (-1)^{\deg A \deg B}\zeta_B\zeta^A]\,,
\end{aligned} \tag{4.3}$$

where deg $A = 0$ or 1 depending on whether A is a bosonic or a fermionic index, respectively. By restricting the superindices to the purely bosonic or purely fermionic indices we recover the Lie algebras of Sp($2m,R$) and SO($2n$), respectively:

$$\begin{aligned}
&(M^i_j = I^i_j,\ S_{ij},\ S^{ij}) \leftrightarrow \mathrm{Sp}(2m,R)\,,\\
&(M^\mu_\nu,\ S_{\mu\nu} = A_{\mu\nu},\ S^{\mu\nu} = A^{\mu\nu}) \leftrightarrow \mathrm{SO}(2n)\,.
\end{aligned} \tag{4.4}$$

The odd generators of OSp($2n/2m,R$) are those bilinears that carry one bosonic and one fermionic index, i.e., $M^i_\mu,M^\mu_i,S_{i\mu},S^{i\mu}$. Their anticommutators close into the even generators:

$$\begin{aligned}
&\{S_{i\mu},S^{j\nu}\} = \delta^\nu_\mu M^j_i - \delta^j_i M^\nu_\mu\,,\\
&\{M^i_\mu,M^\nu_j\} = \delta^i_j M^\nu_\mu + \delta^\nu_\mu M^i_j\,,\\
&\{S_{i\mu},M^j_\nu\} = -\delta^j_i S_{\mu\nu}\,,\\
&\{S_{i\mu},M^\nu_j\} = \delta^\nu_\mu S_{ij}\,.
\end{aligned} \tag{4.5}$$

The other nonvanishing anticommutators can be obtained from these by Hermitian conjugation. Furthermore, the commutators of the even generators with odd generators are

$$\begin{aligned}
&[M^i_j,M^k_\mu] = \delta^k_j M^i_\mu\,,\\
&[M^i_j,M^\mu_k] = -\delta^i_k M^\mu_j\,,\\
&[M^i_j,S_{k\mu}] = -\delta^i_k S_{j\mu}\,,\\
&[M^i_j,S^{k\mu}] = \delta^k_j S^{i\mu}\,,\\
&[M^\mu_\nu,M^k_\lambda] = -\delta^\mu_\lambda M^k_\nu\,,\\
&[M^\mu_\nu,M^\lambda_k] = \delta^\lambda_\nu M^\mu_k\,,\\
&[M^\mu_\nu,S_{k\lambda}] = -\delta^\mu_\lambda S_{k\nu}\,,\\
&[M^\mu_\nu,S^{k\lambda}] = \delta^\lambda_\nu S^{k\mu}\,,\\
&[S_{ij},M^k_\mu] = \delta^k_i S_{j\mu} + \delta^k_j S_{i\mu}\,,\\
&[S_{ij},S^{k\mu}] = \delta^k_i M^\mu_j + \delta^k_j M^\mu_i\,,\\
&[S_{\mu\nu},M^\lambda_k] = -\delta^\lambda_\mu S_{\nu k} + \delta^\lambda_\nu S_{\mu k}\,,\\
&[S_{\mu\nu},S^{k\lambda}] = \delta^\lambda_\nu M^k_\mu - \delta^\lambda_\mu M^k_\nu\,.
\end{aligned} \tag{4.6}$$

The other nonvanishing commutators between odd and even generators can be obtained from these again by Hermitian conjugation.

The unitary highest weight representations of noncompact supergroups are related to the unitary lowest weight representations by a simple involution. For reasons explained above we shall use the term "unitary lowest weight" representations below.

To construct unitary lowest weight representations of OSp($2n/2m,R$) in the super Fock space of all the superoscillators one considers a set of states $|\Omega\rangle$ that are annihilated by all the operators S_{AB} belonging to the L_{-1} space and that transform in a definite representation of the maximal compact subsupergroup U(m/n). Then acting on $|\Omega\rangle$ repeatedly by the operators S^{AB} belonging to the L_{+1} space one generates an infinite set of states in the super Fock space that forms the basis of a unitary lowest weight representation of OSp($2n/2m,R$) (see Refs. 2 and 4). This unitary representation is irreducible if the corresponding lowest weight state $|\Omega\rangle$ transforms irreducibly under the maximal compact subsupergroup U(m/n) (see Refs. 2 and 4). The irreducible representations of U(m/n) that occur in this construction can be conveniently labeled by supertableaux.[18] For example, the operators S^{AB} of the L_{+1} space transform as the supersymmetric tensor of rank 2 under U(m/n), whose supertableau is simply ⧄⧄ . Under the even subgroup U(m)$\times$U(n) the representation ⧄⧄ of U(m/n) decomposes as

$$\text{⧄⧄} = (\square\square,1) + (\square,\square) + \left(1, \begin{array}{c}\square\\[-4pt]\square\end{array}\right)\,. \tag{4.7}$$

The states that occur in the super Fock space fall into irreps of $U(m/n)$. The supertableaux of these irreps can have at most $f = 2p + \varepsilon$ rows. To calculate the full $U(m/n)$ content of a unitary lowest weight representation of $OSp(2n/2m,R)$ one needs to tensor the supertableau of $|\Omega\rangle$ with the supersymmetrized powers of the representation $\boxslash\boxslash$ of $U(m/n)$, since

$$
\begin{aligned}
&S^{AB} \approx \boxslash\boxslash\ , \\
&S^{AB}S^{CD} \approx (\boxslash\boxslash \otimes \boxslash\boxslash)_S = \begin{array}{c}\boxslash\boxslash\\ \boxslash\boxslash\end{array} \oplus \boxslash\boxslash\boxslash\boxslash\ , \qquad (4.8)\\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
&\underbrace{S^{AB}\cdots S^{CD}}_{k\text{ times}} \approx (\underbrace{\boxslash\boxslash \otimes \cdots \otimes \boxslash\boxslash}_{k\text{ times}})\ .
\end{aligned}
$$

If we have only one set of superoscillators (i.e., $p = 0$ and $\varepsilon = 1$), then the only possible irreducible lowest weight vectors of $OSp(2n/2m,R)$ are the vacuum

$$|0\rangle \qquad (4.9a)$$

and the one-particle state

$$\zeta^A|0\rangle \approx \boxslash\ . \qquad (4.9b)$$

For two sets of superoscillators ($p = 1$ and $\varepsilon = 0$) the possible irreducible lowest weight vectors are

$$
\begin{aligned}
&|0\rangle\ , \\
&\xi^A|0\rangle \approx \boxslash\ , \\
&\xi^A\xi^B|0\rangle \approx \boxslash\boxslash\ , \qquad (4.10a)\\
&\underbrace{\xi^A\cdots\xi^D}_{k\text{ times}}|0\rangle \approx \underbrace{\boxslash\boxslash\cdots\boxslash}_{k\text{ boxes}}\ ,
\end{aligned}
$$

and the states

$$(\xi^A\eta^B - \eta^A\xi^B)|0\rangle \approx \begin{array}{c}\boxslash\\ \boxslash\end{array}\ . \qquad (4.10b)$$

For $2p$ sets of superoscillators the possible lowest weight states can all be obtained by tensoring p copies of the lowest weight states of the form (4.10a) and (4.10b). To obtain all the possible lowest weight vectors for $f = 2p + 1$ we need to tensor p copies of the states of the form (4.10a) and (4.10b) with the states (4.9a) and (4.9b). For the rules concerning the decomposition of the tensor product of the corresponding supertableaux into irreducible supertableaux of $U(m/n)$ we refer the reader to Refs. 2, 3, 4, and 18. To obtain lowest weight states $|\Omega\rangle$ of irreducible representations of $OSp(2n/2m,R)$ we must project out the irreducible representations of $U(m/n)$ from the above set of states. In most physical applications it turns out to be more useful to decompose the unitary irreducible representations of a noncompact supergroup into irreducible representations of its even subgroup. The infinite set of irreps of $U(m/n)$ that occur in a given irreducible unitary highest weight representation of $OSp(2n/2m,R)$ can be combined into a finite set of irreducible representations of its even subgroup $Sp(2m,R)\times SO(2n)$. For the short supermultiplets of the supergroup $OSp(8/4,R)$ this has been done in Ref. 4. Here we shall extend these results to general n and m.

The unitary irreps of $Sp(2m,R)$ that occur in a unitary lowest weight representation of $OSp(2n/2m,R)$ are all of the lowest weight type. Therefore to decompose a unitary lowest weight irrep of $OSp(2n/2m,R)$ into irreps of $SO(2n)\times Sp(2m,R)$ is equivalent to identifying all the states obtained by the repeated action of S^{AB} on $|\Omega\rangle$ that are lowest weight vectors of both $SO(2n)$ and $Sp(2m,R)$ simultaneously. Then from these simultaneous lowest weight vectors of $SO(2n)$ and $Sp(2m,R)$ we can read off the full $SO(2n)\times Sp(2m,R)$ content of the unitary lowest weight irrep of $OSp(2n/2m,R)$. Clearly the lowest weight vector $|\Omega\rangle$ of $OSp(2n/2m,R)$ is a simultaneous lowest weight vector of both $SO(2n)$ and $Sp(2m,R)$. The other simultaneous lowest weight vectors of $SO(2n)$ and $Sp(2m,R)$ can be obtained from $|\Omega\rangle$ by the action of the odd generators corresponding to supersymmetry transformations. More specifically, the relevant operators are the odd generators $S^{i\mu}$ belonging to the L_{+1} space. The operators $S^{i\mu}$ transform in the (m,n) representation of the $U(m)\times U(n)$ subgroup of $Sp(2m,R)\times SO(2n)$ with Young tableau $(\square,\square)$. Since $\{S^{i\mu},S^{i\nu}\} = 0$, the product of the operators $S^{i\mu}$ transform in antisymmetrized powers of $(\square,\square)$, i.e.,

$$
\begin{aligned}
&S^{i\mu} \approx (\square,\square), \quad \text{under } U(m)\times U(n)\ , \\
&S^{i\mu}S^{i\nu} \approx \left(\square\square\ ,\begin{array}{c}\square\\ \square\end{array}\right) + \left(\begin{array}{c}\square\\ \square\end{array},\square\square\right), \\
&S^{i\mu}S^{j\nu}S^{k\rho} \approx \left(\square\square\square\ ,\begin{array}{c}\square\\ \square\\ \square\end{array}\right) + \left(\begin{array}{l}\square\square\\ \square\end{array},\begin{array}{l}\square\square\\ \square\end{array}\right) + \left(\begin{array}{c}\square\\ \square\\ \square\end{array},\ \square\square\square\right), \\
&\qquad\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad\qquad\qquad (4.10c)\\
&\underbrace{S^{i\mu}\cdots S^{l\rho}}_{k\text{ copies}} \approx (\square,\square)^k_{\text{antisymmetrized}}\ .
\end{aligned}
$$

Clearly all the $U(m)$ and $U(n)$ tableaux with more than m and n rows, respectively, vanish. Furthermore, if the number $f = 2p + \varepsilon$ of the superoscillators is less than m or n, then the states whose $U(m)$ or $U(n)$ tableaux have more than f rows or f columns, respectively, do not occur in the super Fock space. The lowest weight vectors of $Sp(2m,R)\times SO(2n)$ will be a tensor product of the lowest weight vectors of $Sp(2m,R)$ and $SO(2n)$, which were discussed in Secs. III and II, respectively. Therefore any state of this tensor product form created by the repeated action of $S^{i\mu}$ on the lowest weight state $|\Omega\rangle$ will be a lowest weight state of $Sp(2m,R)\times SO(2n)$. To identify these states simply one needs to tensor the Young tableaux of the lowest weight state $|\Omega\rangle$ with the antisymmetrized powers of $(\square,\square)$. Let us now illustrate this with examples. Consider the simplest case of $f = 1$ (i.e., $p = 0$ and $\varepsilon = 1$). Then there exist only two lowest weight states transforming irreducibly under the maximal compact subsupergroup $U(m/n)$, namely, the vacuum state $|0\rangle$ and the one particle state $\zeta^A|0\rangle$. Consider now the unitary representation with lowest weight vector $|0\rangle$. By the rules stated above there are only two states that are lowest weight vectors of $Sp(2m,R)\times SO(2n)$ inside this irrep of $OSp(2n/2m,R)$. They are

$$
\begin{aligned}
&|0\rangle \approx (1,1), \quad \text{under } U(m)\times U(n), \\
&S^{i\mu}|0\rangle = c^i\gamma^\mu|0\rangle \approx (\square,\square), \quad \text{under } U(m)\times U(n)\ .
\end{aligned}
$$

All the other states inside the unitary representation of $OSp(2n/2m,R)$ correspond to "excitations" of these states by the action of the operators belonging to the L_{+1} space of $SO(2n)$ and $Sp(2m,R)$. By the results of Secs. II and III we

TABLE VI. The $Sp(2m,R)\times SO(2n)$ content of the unitary irrep of $OSp(2n/2m,R)$ with lowest weight vector $\vert 0\rangle$ for even $n=2k$ and $f=1$.

Lowest weight vector of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$\vert 0\rangle$	$(0,...,0)_{YT};\ Q_B=m/2$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2})_{GZ}=(0,0,0,...,1)_D$
$S^{i\mu}\vert 0\rangle\approx(\square,\square)$	$(1,0,...,0)_{YT};\ Q_B=m/2+1$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2},-\frac{1}{2})_{GZ}=(0,0,...,1,0)_D$

can read off the $Sp(2m,R)\times SO(2n)$ content of these unitary lowest weight representations of $OSp(2n/2m,R)$. They are given in Tables VI and VII for even and odd n, respectively.

The unitary representation of $OSp(2n/2m,R)$ with the lowest weight vector $\zeta^A\vert 0\rangle$ has lowest weight vectors of $Sp(2m,R)\times SO(2n)$, namely, $c^i\vert 0\rangle\approx(\square,1)$ and $\gamma^\mu\vert 0\rangle\approx(1,\square)$. The corresponding $Sp(2m,R)\times SO(2n)$ labels of these lowest weight vectors are given in Tables VIII and IX. The supermultiplets of $OSp(2n/2m,R)$ for $f=1$ are the singleton supermultiplets and agree with those obtained for $OSp(8/4,R)$ (see Ref. 4) and $OSp(2n/2,R)$ (see Ref. 6). The singleton supermultiplets are the shortest supermultiplets of $OSp(2n/2m,R)$ and involve the two spinor representations of $SO(2n)$. Note that the two spinor representations of $SO(2n)$ get interchanged in going from the singleton supermultiplet with lowest weight vector $\vert 0\rangle$ to that with lowest weight vector $\zeta^A\vert 0\rangle$.

Let us next consider the unitary lowest weight representation of $OSp(2n/2m,R)$ with lowest weight vector $\vert 0\rangle$ and $f=2$ (i.e., $p=1$ and $\varepsilon=0$). In this case, possible lowest weight states of $Sp(2m,R)\times SO(2n)$ have the $U(m)\times U(n)$ Young tableaux

$$\begin{aligned}
&(0,0)\approx\vert 0\rangle ,\\
&(\square,\square) ,\\
&\left(\square\square,\begin{array}{c}\square\\\square\end{array}\right)+\left(\begin{array}{c}\square\\\square\end{array},\square\square\right) ,\\
&\left(\square\square\square,\begin{array}{c}\square\\\square\\\square\end{array}\right) ,\\
&\left(\square\square\square\square,\begin{array}{c}\square\\\square\\\square\\\square\end{array}\right) ,\\
&\quad\vdots\\
&\left(\underbrace{\square\square\cdots\square}_{n\text{ boxes}},\left.\begin{array}{c}\square\\\square\\\vdots\\\square\end{array}\right\}n\text{ boxes}\right) .
\end{aligned}\tag{4.11}$$

In Table X we give the full $Sp(2m,R)\times SO(2n)$ content of this unitary representation of $OSp(2n/2m,R)$.

For $f=2$, any state of the form $\xi^A\xi^B\cdots\xi^C\vert 0\rangle$ is a lowest weight vector of $OSp(2n/2m,R)$. Such states transform irreducibly under the maximal compact subsupergroup $U(m/n)$. They correspond to the supersymmetric tensor representations of $U(m/n)$ with supertableaux $\boxslash\boxslash\cdots\boxslash$. In addition, we have the lowest weight vector

$$[\xi^A\eta^B-\eta^A\xi^B]\vert 0\rangle ,\tag{4.12}$$

which also transforms in an irreducible representation of $U(m/n)$ with supertableau $\begin{array}{c}\boxslash\\\boxslash\end{array}$. These lowest weight vectors of $OSp(2n/2m,R)$ can be decomposed into irreducible lowest weight vectors of $Sp(2m,R)\times SO(2n)$ simply by decomposing the supertableaux of $U(m/n)$ into irreps of its even subgroup $U(m)\times U(n)$. For example, we have

$$\begin{aligned}
&U(m/n)\supset U(m)\times U(n) ,\\
&\boxslash=(\square,1)+(1,\square) ,\\
&\begin{array}{c}\boxslash\\\boxslash\end{array}=\left(\begin{array}{c}\square\\\square\end{array},1\right)+(\square,\square)+(1,\square\square) ,\\
&\boxslash\boxslash\ =(\square\square,1)+(\square,\square)+\left(1,\begin{array}{c}\square\\\square\end{array}\right) ,\\
&\underbrace{\boxslash\boxslash\cdots\boxslash}_{k}=(\underbrace{\square\square\cdots\square}_{k},1)\\
&\qquad+(\underbrace{\square\cdots\square}_{k-1},\square)+\cdots+\left(\square,\left.\begin{array}{c}\square\\\square\\\vdots\\\square\end{array}\right\}k-1\right)+\left(1,\left.\begin{array}{c}\square\\\square\\\vdots\\\square\end{array}\right\}k\right) .
\end{aligned}\tag{4.13}$$

For $k>n$ those Young tableaux of $U(n)$ on the right-hand side containing more than n boxes vanish. For further details on the decomposition of supertableaux we refer to Refs. 4 and 18. In addition to the lowest weight vectors of $Sp(2m,R)\times SO(2n)$ contained in the lowest weight vector of $OSp(2n/2m,R)$, there are others created by the action of the odd supersymmetry generators $S^{i\mu}$ within a unitary representation of $OSp(2n/2m,R)$. In Table XI we give the $Sp(2m,R)\times SO(2n)$ content of the irrep of $OSp(2n/2m,R)$

TABLE VII. $Sp(2m,R)\times SO(2n)$ content of the unitary irrep of $OSp(2n/2m,R)$ with lowest weight vectors $\vert 0\rangle$ for odd $n=2k+1$ and $f=1$.

Lowest weight vector of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$\vert 0\rangle$	$(0,...,0)_{YT};\ Q_B=m/2$	$(\frac{1}{2},\frac{1}{2},...,-\frac{1}{2})_{GZ}=(0,0,...,1,0)_D$
$S^{i\mu}\vert 0\rangle\approx(\square,\square)$	$(1,0,...,0)_{YT};\ Q_B=m/2+1$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2},\frac{1}{2})_{GZ}=(0,0,...,0,1)_D$

TABLE VIII. $Sp(2m,R)\times SO(2n)$ content of the unitary irrep of $OSp(2n/2m,R)$ with lowest weight vector $\zeta^A|0\rangle$ for even $n=2k$ and $f=1$.

Lowest weight vector of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$c^i|0\rangle \approx (\square,1)$	$(1,0,...,0)_{YT};\ Q_B=m/2+1$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2})_{GZ}=(0,0,...,0,1)_D$
$\gamma^\mu|0\rangle \approx (1,\square)$	$(0,0,0,...,0)_{YT};\ Q_B=m/2$	$(\frac{1}{2},\frac{1}{2},...,-\frac{1}{2},)_{GZ}=(0,0,...,1,0)_D$

TABLE IX. $Sp(2m,R)\times SO(2n)$ content of the unitary irrep of $OSp(2n/2m,R)$ with lowest weight vector $\zeta^A|0\rangle \approx$ ⧄ for odd $n=2k+1$ and $f=1$.

Lowest weight vector of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$c^i|0\rangle \approx (\square,1)$	$(1,0,...,0)_{YT};\ Q_B=m/2+1$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2},-\frac{1}{2})_{GZ}=(0,0,...,1,0)_D$
$\gamma^\mu|0\rangle \approx (1,\square)$	$(0,0,...,0)_{YT};\ Q_B=m/2+1$	$(\frac{1}{2},\frac{1}{2},...,\frac{1}{2},\frac{1}{2})_{GZ}=(0,0,...,0,1)_D$

TABLE X. The $Sp(2m,R)\times SO(2n)$ decomposition of the unitary representation of $OSp(2n/2m,R)$ with lowest weight vector $|0\rangle$ and $f=2$.

$U(m)\times U(n)$ Young tableaux of the lowest weight vectors of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$(0,0)_{YT}$	$(0,0,...,0)_{YT};\ Q_B=m$	$(1,1,...,1)_{GZ}=(0,0,...,0,2)_D$, for even n $(1,1,...,1,-1)_{GZ}=(0,0,...,0,2,0)_D$, for odd n
$(\square,\square)$	$(1,0,...,0)_{YT};\ Q_B=m+1$	$(1,1,...,1,0)_{GZ}=(0,0,...,0,1,1)_D$, for all n
$\Big(\underbrace{\square\cdots\square}_{\substack{k\text{ boxes}\\ \{2<k<n\}}},\ \left.\begin{smallmatrix}\square\\ \vdots\\ \square\end{smallmatrix}\right\}k\text{ boxes}\Big),$	$(k,0,...,0)_{YT};\ Q_B=m+k$	$(1,1,...,1,\overbrace{0,...,0}^{k\text{ zeros}})_{GZ}=(0,0,...,0,1,\underbrace{0,...,0}_{k\text{ zeros}})_D$

TABLE XI. The $Sp(2m,R)\times SO(2n)$ content of the unitary irrep of $OSp(2n/2m,R)$ with lowest weight vector $(\xi^A\eta^B-\eta^A\xi^B)|0\rangle$ and $f=2$.

$U(m)\times u(n)$ Young tableaux of the lowest weight vector of $Sp(2m,R)\times SO(2n)$	$U(m)$ labels of the lowest weight vector of the irrep of $Sp(2m,R)$	Gelfand–Zetlin and Dynkin labels of the irrep of $SO(2n)$
$\left(\begin{smallmatrix}\square\\ \square\end{smallmatrix},1\right)$	$(1,1,0,...,0)_{YT};\ Q_B=m+2$	$(1,1,...,1)_{GZ}=(0,0,...,0,2)$, for even n $(1,1,...,,-1)_{GZ}=(0,0,...,0,2,0)_D$, for odd n
$(\square,\square)$	$(1,0,...,0)_{YT};\ Q_B=m+1$	$(1,1,...,1,0)_{GZ}=(0,0,...,0,1,1)_D$, for all n
$(1,\square\square)$	$(0,0,...,0)_{YT};\ Q_B=m$	$(1,1,...,1,-1)_{GZ}=(0,...,0,2,0)_D$, for even n $(1,1,...,1,1)_{GZ}=(0,0,...,0,2)_D$, for odd n
$\left(\square\square,\begin{smallmatrix}\square\\ \square\end{smallmatrix}\right)$	$(2,0,...,0)_{YT};\ Q_B=m+2$	$(1,1,...,1,0,0)_{GZ}=(0,0,...,0,1,0,0)_D$
$\Big(\underbrace{\overbrace{\square\square\cdots\square}^{k}}_{3<k<n},\ \left.\begin{smallmatrix}\square\\ \vdots\\ \square\end{smallmatrix}\right\}k\Big),$	$(k,0,0,...,0)_{YT};\ Q_B=m+k$	$(1,1,...,1,\overbrace{0,...,0}^{k})_{GZ}=(0,...,0,1,\underbrace{0,...,0}_{k})_D$

with lowest weight vector given in (4.12) whose supertableau is  .

It is interesting to compare the unitary irreps of OSp(8/4,R) whose lowest weight vectors are $|0\rangle$ and $(\xi^A\eta^B - \eta^A\xi^B)|0\rangle$. As stated above, Sp(4,$R$) is the covering group of the anti-de Sitter group in four space-time dimensions. Following Ref. 4 we shall label its lowest weight irreps by the spin S (SU(2)) and anti-de Sitter energy $E_0 = Q_B/2$ (U(1)). Then the contents of these two irreps of OSp(8/4,R) are as follows.

$|\Omega\rangle = |0\rangle$:

S	E_0	SO(8)
0	1	$(0,0,0,2)_D$
0	2	$(0,0,2,0)_D$
$\frac{1}{2}$	$\frac{3}{2}$	$(0,0,1,1)_D$
1	2	$(0,1,0,0)_D$
$\frac{3}{2}$	$\frac{5}{2}$	$(1,0,0,0)_D$
2	3	$(0,0,0,0)_D$.

$|\Omega\rangle = (\xi^A\eta^B - \eta^A\xi^B)|0\rangle$:

S	E_0	SO(8)
0	2	$(0,0,0,2)_D$
0	1	$(0,0,2,0)_D$
$\frac{1}{2}$	$\frac{3}{2}$	$(0,0,1,0)_D$
1	2	$(0,1,0,0)_D$
$\frac{3}{2}$	$\frac{5}{2}$	$(1,0,0,0)_D$
2	3	$(0,0,0,0)_D$.

From the spin and anti-de Sitter energy content of these supermultiplets it is clear that they correspond to massless $N = 8$ anti-de Sitter supermultiplets.[4] They only differ in the SO(8) transformation properties of the $S = 0$ states with a given anti-de Sitter energy. Even though they are nonequivalent $N = 8$ anti-de Sitter supermultiplets, they reduce to the same massless $N = 8$ supermultiplet in the Poincaré limit! This suggests that corresponding to a given Poincaré extended supergravity theory there may be different, nonequivalent gauged (anti-de Sitter) supergravity theories.

Before concluding, let us now summarize the general rules for determining the SO($2n$)$\times$Sp($2m$,R) content of a given irreducible unitary lowest weight representation of OSp($2n$/$2m$,R).

(1) Decompose the lowest weight state $|\Omega\rangle$ into irreducible representations of the even subgroup U(m)$\times$U(n) of the maximal compact subsupergroup U(m/n). Each one of these irreps of U(m)$\times$U(n) is a lowest weight state of an irrep of SO($2n$)$\times$Sp($2m$,R).

(2) By acting on these lowest weight states of SO($2n$)$\times$Sp($2m$,R) with the odd supersymmetry generators $S^{i\mu} \approx (\square,\square)$ repeatedly, one generates all the other lowest weight states of the irreps of SO($2n$)$\times$Sp($2m$,R) that occur in the irrep of OSp($2n$/$2m$,R) defined by $|\Omega\rangle$. Since $\{S^{i\mu},S^{j\nu}\} = 0$, the higher powers of $S^{i\mu}$ transform in the antisymmetrized tensor product of the representation $(\square,\square)$ of U(m)$\times$U(n). For a given number $f = 2p + \varepsilon$ of oscillators, the possible lowest weight states of SO($2n$) and Sp($2m$,R) were given in Secs. II and III, respectively. Here it is important to keep in mind that only some of the states created by the repeated action of $S^{i\mu}$ will be lowest weight states of SO($2n$)$\times$Sp($2m$,R). The rest will simply be excitations of these lowest weight states.

(3) The irreducible unitary lowest weight representations of Sp($2m$,R) are uniquely labeled by the U(m) quantum numbers of their lowest weight vectors. Their SO($2n$) transformation properties can be read off from the U(n) transformation properties of the corresponding lowest weight vectors of SO($2n$) according to the rules given in Sec. II.

[1] M. Günaydin and C. Saclioglu, Phys. Lett. B **108**, 169 (1982); Commun. Math. Phys. **87**, 159 (1982).

[2] I. Bars and M. Günaydin, Commun. Math. Phys. **91**, 31 (1983); M. Günaydin, in *Proceedings of the XIth International Colloquium on Group Theoretical Methods, Lecture Notes in Physics*, Vol. 180, edited by E. İnönü and M. Serdaroḡlu (Springer, New York, 1982).

[3] M. Günaydin, P. van Nieuwenhuizen, and N. P. Warner, Nucl. Phys. B **255**, 63 (1985).

[4] M. Günaydin and N. P. Warner, Nucl. Phys. B **272**, 99 (1986).

[5] M. Günaydin and N. Marcus, Class. Quantum Grav. **2**, L11 (1985).

[6] M. Günaydin, G. Sierra, and P. K. Townsend, Nucl. Phys. B **274**, 429 (1986).

[7] K. Pilch, P. K. Townsend, and P. van Nieuwenhuizen, Nucl. Phys. B **242**, 377 (1984); P. van Nieuwenhuizen, Class. Quantum Grav. **2**, 1 (1985).

[8] E. Cremmer, B. Julia, and J. Scherk, Phys. Lett. B **76**, 409 (1978).

[9] See M. J. Duff, B. E. W. Nilsson, and C. N. Pope, Phys. Rep. **130**, 1 (1986), and the references therein.

[10] H. J. Kim, L. J. Romans, and P. van Nieuwenhuizen, Phys. Rev. D **32**, 389 (1985).

[11] M. Günaydin, B. E. W. Nilsson, G. Sierra, and P. K. Townsend, Phys. Lett. B **176**, 45 (1986).

[12] Harish-Chandra, Am. J. Math. 77, 743 (1955); **78**, 1 (1956).

[13] M. Günaydin, J. Math. Phys. **29**, 1275 (1988).

[14] V. G. Kac, Adv. Math. **26**, 8 (1977).

[15] D. Z. Freedman and N. Nicolai, Nucl. Phys. B **237**, 342 (1984); L. Castell, W. Heidenreich, and T. Künemund, Starnberg preprint, 1984; C. Fronsdal and T. Hirai, in *Essays on Supersymmetry*, edited by C. Fronsdal (Reidel, Dordrecht, 1986).

[16] M. Günaydin, Ann. Israel Phys. Soc. **2**, 279 (1980).

[17] P. A. M. Dirac, J. Math. Phys. **4**, 901 (1963); C. Fronsdal, Phys. Rev. D **12**, 3819 (1975).

[18] A. B. Balantekin and I. Bars, J. Math. Phys. **22**, 1149, 1810 (1981); **23**, 1239 (1982); J. Q. Chen, X. G. Chen, and M. G. Gao, J. Phys. A **16**, 1361 (1983).

Infinitesimal operators of group representations in noncanonical bases

A. U. Klimyk and I. I. Kachurik
Institute for Theoretical Physics, Kiev-130, Union of Soviet Socialist Republics

(Received 10 March 1988; accepted for publication 29 June 1988)

Explicit expressions are obtained for the action of the infinitesimal operators of some classes of representations of the groups $\mathrm{U}(p+q)$, $\mathrm{SO}(p+q)$, $\mathrm{SU}(n)$, $\mathrm{Sp}(n)$, $\mathrm{SL}(n,R)$, and $\mathrm{GL}(n,C)$ in the bases that differ from the Gel'fand-Zetlin ones.

I. INTRODUCTION

The first expressions for the action of infinitesimal operators of group representations onto elements of a discrete basis were the Gel'fand–Zetlin formulas.[1,2] The results for $\mathrm{SU}(n)$ were independently proved in Ref. 3. Apparently, they are the most utilizable results of the representation theory in physics, especially in particle physics, nuclear physics, and quantum chemistry. Infinitesimal operators of representations are of great importance for dynamical symmetries.

Considering physical applications of group representations, we need various bases (corresponding to different subgroup reductions) of carrier spaces. In this paper we give explicit expressions for infinitesimal operators of some classes of representations of the classical Lie groups in the bases that differ from the Gel'fand–Zetlin ones.

The Gel'fand–Zetlin formulas for the infinitesimal operators of representations of the groups $\mathrm{U}(n)$ and $\mathrm{SO}(n)$ correspond to the so-called canonical reductions,

$$\mathrm{U}(n)\supset\mathrm{U}(n-1)\supset\cdots\supset\mathrm{U}(1),$$
$$\mathrm{SO}(n)\supset\mathrm{SO}(n-1)\supset\cdots\supset\mathrm{SO}(2).$$

These reductions are characterized by the unit multiplicity for irreducible representations of neighboring subgroups. For noncanonical reductions, multiplicities can exceed 1. It is well known that to derive explicit formulas in these cases is a very complicated task. Noncanonical reductions for certain classes of representations are also characterized by the unit multiplicities. We deal with reductions and representations of this type.

II. INFINITESIMAL OPERATORS FOR THE GROUP U($p+q$) IN A U(p)×U(q) BASIS

We consider the irreducible representations $D(M_1,M_2)$ of $\mathrm{U}(p+q)$ with highest weights $(M_1,\dot{0},M_2)$, $M_1\geqslant 0\geqslant M_2$, where $\dot{0}=(0,...,0)$. The reduction of $D(M_1,M_2)$ onto the subgroup $\mathrm{U}(p)\times\mathrm{U}(q)$ decomposes into a sum of the irreducible representations $D_p(m_1,m_2)\otimes D_q(m_1',m_2')$ of $\mathrm{U}(p)\times\mathrm{U}(q)$ with highest weights $(m_1,\dot{0},m_2)(m_1',\dot{0},m_2')$, $m_1\geqslant 0\geqslant m_2$, $m_1'\geqslant 0\geqslant m_2'$, for which[4]

$$m_1+m_2+m_1'+m_2'=M_1+M_2,$$
$$m_1-m_2+m_1'-m_2'\leqslant M_1-M_2.$$

Each of these representations occurs in the decomposition with the unit multiplicity. The set of elements of the Gel'fand–Zetlin bases of all representations $D_p(m_1,m_2)\otimes D_q(m_1',m_2')$ forms an orthonormal basis for the representation $D(M_1,M_2)$. The basis elements are denoted by $|m_1,m_2,m_1',m_2';\alpha,\beta\rangle$, where α and β are the Gel'fand–Zetlin patterns for the representations $D_p(m_1,m_2)$ and $D_q(m_1',m_2')$, respectively, without the first line.

The infinitesimal operators $E_{p,p+1}$ and $E_{p+1,p}$ of the representation $D(M_1,M_2)$ are defined by Eqs. (9) and (10) of Ref. 5. The coefficients $K_{\cdots}^{\cdots}$ and $K_{\cdots}^{\cdots}(\alpha,\beta)$ of these equations are expressed by means of Clebsch–Gordan coefficients (CGC's) of the tensor products $D_p(m_1,m_2)\oplus D_p(1,0)$, $D_p(m_1,m_2)\otimes D_p(0,-1)$ of the representations of $\mathrm{U}(p)$ and of similar tensor products for the group $\mathrm{U}(q)$ (see Sec. I in Ref. 5).

Using in $K_{\cdots}^{\cdots}$ and $K_{\cdots}^{\cdots}(\alpha,\beta)$ the expressions for CGC's from Ref. 3, after some simplifications we obtain that

$$\begin{aligned}
E_{p,p+1}|m_1,m_2,m_1',m_2';\alpha,\beta\rangle &= [(M_1-m_2-m_1'+p)(M_2-m_2-m_1'-q+1)]^{1/2}\\
&\quad\times K_p(m_1,m_2,n_1,n_2)K_q(m_1'-1,m_2',n_1',n_2')|m_1+1,m_2,m_1'-1,m_2';\alpha,\beta\rangle\\
&\quad+[(M_1-m_1-m_1')(M_2-m_1-m_1'-p-q+1)]^{1/2}K_p(m_1,m_2,n_1,n_2)\\
&\quad\times K_q'(m_1',m_2'-1,n_1',n_2')|m_1+1,m_2,m_1',m_2'-1;\alpha,\beta\rangle\\
&\quad+[(M_1-m_1-m_1'+1)(M_2-m_1-m_1'-p-q+2)]^{1/2}K_p'(m_1,m_2,n_1,n_2)\\
&\quad\times K_q(m_1'-1,m_2',n_1',n_2')|m_1,m_2+1,m_1'-1,m_2';\alpha,\beta\rangle\\
&\quad+[(M_1-m_2-m_1'+p-1)(M_2-m_2-m_1'-q)]^{1/2}K_p'(m_1,m_2,n_1,n_2)\\
&\quad\times K_q'(m_1',m_2'-1,n_1',n_2')|m_1,m_2+1,m_1',m_2'-1;\alpha,\beta\rangle, \qquad (1)
\end{aligned}$$

$$
\begin{aligned}
E_{p+1,p}|m_1,m_2,m_1',m_2';\alpha,\beta\rangle = & -[(M_1-m_2-m_1'+p+1)(M_2-m_2-m_1'-q)]^{1/2} \\
& \times K_p(m_1-1,m_2,n_1,n_2)K_q(m_1',m_2',n_1',n_2')|m_1-1,m_2,m_1'+1,m_2';\alpha,\beta\rangle \\
& -[(M_1-m_1-m_1'+1)(M_2-m_1-m_1'-p-q+2)]^{1/2}K_p(m_1-1,m_2,n_1,n_2) \\
& \times K_q'(m_1',m_2',n_1',n_2')|m_1-1,m_2,m_1',m_2'+1;\alpha,\beta\rangle \\
& -[(M_1-m_1-m_1')(M_2-m_1-m_1'-p-q+1)]^{1/2}K_p'(m_1,m_2-1,n_1,n_2) \\
& \times K_q(m_1',m_2',n_1',n_2')|m_1,m_2-1,m_1'+1,m_2';\alpha,\beta\rangle \\
& -[(M_1-m_2-m_1'+p)(M_2-m_2-m_1'-q+1)]^{1/2}K_p'(m_1,m_2-1,n_1,n_2) \\
& \times K_q'(m_1',m_2',n_1',n_2')|m_1,m_2-1,m_1',m_2'+1;\alpha,\beta\rangle, \qquad (2)
\end{aligned}
$$

where n_1 and n_2 are defined by the first line $(n_1,\dot{0},n_2)$ of the pattern α, a highest weight of a representation of $\mathrm{U}(n-1)$, n_1 and n_2 are defined by the first line of the pattern β, and

$$K_s(r_1,r_2,t_1,t_2)=\left[\frac{(r_1-t_1+1)(r_1-t_2+s-1)}{(r_1-r_2+s-1)(r_1-r_2+s)}\right]^{1/2}, \qquad (3)$$

$$K_s'(r_1,r_2,t_1,t_2)=\left[\frac{(t_1-r_2+s-2)(t_2-r_2)}{(r_1-r_2+s-1)(r_1-r_2+s-2)}\right]^{1/2}. \qquad (4)$$

As in the case of Gel'fand–Zetlin formulas, the expressions for the infinitesimal operators E_{rt}, E_{tr}, $r\leqslant p<t$, are obtained by making commutation of $E_{p,p+1}$ and $E_{p+1,p}$ with the infinitesimal operators for the subgroups $\mathrm{U}(p)$ and $\mathrm{U}(q)$. They also can be obtained[6] with the help of CGC's of Ref. 3.

The expressions (3) and (4) define infinitesimal operators of the representations $\pi(\lambda_1,\lambda_2)$ of the group $\mathrm{U}(p,q)$ in Ref. 4.

III. INFINITESIMAL OPERATORS FOR THE GROUP SO($p+q$) IN AN SO(p)×SO(q) BASIS

We consider the irreducible representations $D(M)$ of $\mathrm{SO}(p+q)$ with highest weights $(M,\dot{0})$, $M\geqslant 0$. The reduction of $D(M)$ onto the subgroup $\mathrm{SO}(p)\times\mathrm{SO}(q)$ decomposes into a sum of the irreducible representations of $\mathrm{SO}(p)\times\mathrm{SO}(q)$ with highest weights $(m,\dot{0})(m',\dot{0})$, $m\geqslant 0$, $m'\geqslant 0$, for which $m+m'$ and M are of the same evenness and

$$m+m'\leqslant M,\quad \text{if } p\geqslant q>2,\quad m+m'\leqslant M,\quad m-m'\leqslant M,\quad \text{if } p>q=2.$$

The decomposition is free of multiplicities. A basis of the carrier space of $D(M)$ consists of the Gelfand–Zetlin bases of the representations $D_p(m)\otimes D_q(m')$. As in the case of $\mathrm{U}(p+q)$, the basis elements are denoted by $|m,m';\alpha,\beta\rangle$.

The infinitesimal operators $J_{p,p+1}=E_{p,p+1}-E_{p+1,p}$ of the representation $D(M)$ are defined by Eq. (8.19) of Ref. 7. As in the case of $\mathrm{U}(p+q)$, the coefficients $K^{\cdots}_{\cdots}$ and $K^{\cdots}_{\cdots}(\alpha,\beta)$ of this equation are expressed by means of CGC's of the tensor products $D_p(m)\otimes D_p(1)$ and $D_q(m')\otimes D_q(1)$. We find these CGC's from Eqs. (4.24)–(4.27) of Ref. 7:

$$\left(\begin{matrix} m & 1 \\ n & 0 \end{matrix}\middle|\begin{matrix} m+1 \\ n \end{matrix}\right)=\left[\frac{(m+n+p-2)(m-n+1)}{(m+1)(2m+p-2)}\right]^{1/2},\quad \left(\begin{matrix} m & 1 \\ n & 0 \end{matrix}\middle|\begin{matrix} m-1 \\ n \end{matrix}\right)=\left[\frac{(m+n+p-3)(m+n)}{(m+n-3)(2m+p-2)}\right]^{1/2}.$$

Using these expressions and formulas for dimensions of representations of $\mathrm{SO}(n)$, we obtain, for $p\geqslant q>2$,

$$
\begin{aligned}
& J_{p,p+1}|m,m';\alpha,\beta\rangle \\
& \quad =[(M-m-m')(M+m+m'+p+q-2)]^{1/2}K_p(m,n)K_q(m',n')|m+1,m'+1;\alpha,\beta\rangle \\
& \qquad +[(M-m+m'+q-2)(M+m-m'+p)]^{1/2}K_p(m,n)K_q(m'-1,n')|m+1,m'-1;\alpha,\beta\rangle \\
& \qquad -[(M-m+m'+q)(M+m-m'+p-2)]^{1/2}K_p(m-1,n)K_q(m',n')|m-1,m'+1;\alpha,\beta\rangle \\
& \qquad -[(M-m-m'+2)(M+m+m'+p+q-4)]^{1/2}K_p(m-1,n)K_q(m'-1,n')|m-1,m'-1;\alpha,\beta\rangle, \qquad (5)
\end{aligned}
$$

where n and n' are defined by the first lines $(n,\dot{0})$ and $(n',\dot{0})$ of the patterns α and β, and

$$K_s(r,t)=\left[\frac{(r+t+s-2)(r-t+1)}{(2r+s-2)(2r+s)}\right]^{1/2}. \qquad (6)$$

If $p > q = 2$, we have $|m,m';\alpha\rangle$, $m \geqslant 0$, $-\infty < m' < \infty$, instead of $|m,m';\alpha,\beta\rangle$ and $J_{p,p+1}$ has the form

$$\begin{aligned} J_{p,p+1}|m,m';\alpha\rangle = {} & [(M-m-m')(M+m+m'+p)]^{1/2}K_p(m,n)|m+1,m'+1;\alpha\rangle \\ & + [(M-m+m')(M+m-m'+p)]^{1/2}K_p(m,n)|m+1,m'-1;\alpha\rangle \\ & - [(M-m+m'+2)(M+m-m'+p-2)]^{1/2}K_p(m-1,n)|m-1,m'+1;\alpha\rangle \\ & - [(M-m-m'+2)(M+m+m'+p-2)]^{1/2}K_p(m-1,n)|m-1,m'-1;\alpha\rangle, \end{aligned} \tag{7}$$

where $K_p(m,n)$ is given by Eq. (6).

The expressions for the infinitesimal operators $J_{st} = E_{st} - E_{ts}$, $s \leqslant p < t$, are obtained by making commutation of $J_{p,p+1}$ with the infinitesimal operators for the subgroups SO(p) and SO(q).

By means of Eq. (6) we can obtain infinitesimal operators of the representations π_λ, $\lambda \in C$, of the most degenerate series of SO(p,q) (see Sec. 5 of Chap. 8 in Ref. 7).

IV. INFINITESIMAL OPERATORS FOR THE GROUP SU(n) IN AN SO(n) BASIS

We consider the irreducible representations $D(M,\dot{0})$ of SU(n) with highest weights $(M,\dot{0})$. The reduction of $D(M,\dot{0})$ onto the subgroup SO(n) decomposes into a sum of the irreducible representations $D(m)$ of SO(n) with highest weights $(m,\dot{0})$, for which m and M are of the same evenness and $0 \leqslant m \leqslant M$. Multiplicities equal 1. The basis of a carrier space of $D(M,0)$ consists of the Gel'fand–Zetlin bases of the representations $D(m)$. The basis elements are denoted by $|m,\alpha\rangle$, where α is a Gel'fand–Zetlin pattern for $D(m)$ without the highest weight $(m,\dot{0})$.

The Lie algebra su(n) of SU(n) decomposes into the direct sum su(n) = so(n) + $\mathfrak{p}$, where so(n) is a Lie algebra of SO(n) and the subspace $\mathfrak{p}$ consists of the linear combinations of the matrices

$$J_{ij} = \sqrt{-1}(E_{ij} + E_{ji}), \quad i < j, \tag{8a}$$

$$J_{kk} = \left(\frac{-2}{n(n-1)}\right)^{1/2}\left(E_{kk} - \frac{1}{n}\sum_{j=1}^{n} E_{jj}\right), \quad k = 1,2,\ldots,n. \tag{8b}$$

The operators ad X: (ad X)$Y = [X,Y]$, $X \in$ so(n), realize on $\mathfrak{p}$ the irreducible representation[8,9] of so(n) with highest weight $(2,\dot{0})$. The matrices (8a) and (8b) form the Gel'fand–Zetlin basis for this representation. We have the following one-to-one correspondence:

$$J_{j-i,j} \to \left[\begin{array}{ccccccccc} 2 & & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & 2 & & 0 & \cdots & \cdots & \cdots & & 0 \\ & & 1 & & 0 & \cdots & & 0 & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & & \\ & & & 0 & \cdots & \cdots & 0 & & \\ & & & & \cdots & & & & \\ & & & & 0 & & & & \end{array}\right], \tag{9}$$

where the first $n-j+1$ rows coincide with $(2,\dot{0})$ and the following i rows coincide with $(1,\dot{0})$.

The infinitesimal operators J_{sj} of the representation $D(M,0)$ of U(n) have the form[8,9]

$$\begin{aligned} J_{sj}|m,\alpha\rangle = {} & \left(\frac{n-1}{2n}\right)^{1/2}\Bigg\{[(M-m+2)(M+m+n-2)]^{1/2}\left[\frac{\dim D(m)}{\dim D(m-2)}\right]^{1/2}\sum_{\alpha'}\left(\begin{array}{cc|c} m & 2 & m-2 \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m-2 \\ \alpha & (J_{sj}) & \alpha' \end{array}\right) \\ & \times|m-2,\alpha'\rangle - [(M-m)(M+m+n)]^{1/2}\left[\frac{\dim D(m)}{\dim D(m+2)}\right]^{1/2}\sum_{\alpha'}\left(\begin{array}{cc|c} m & 2 & m+2 \\ 0 & 0 & 0 \end{array}\right) \\ & \times\left(\begin{array}{cc|c} m & 2 & m+2 \\ \alpha & (J_{sj}) & \alpha' \end{array}\right)|m+2,\alpha'\rangle - \sqrt{-1}\left(M+\frac{n}{2}\right)\sum_{\alpha'}\left(\begin{array}{cc|c} m & 2 & m \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m \\ \alpha & (J_{sj}) & \alpha' \end{array}\right)|m,\alpha'\rangle\Bigg\}, \end{aligned} \tag{10}$$

where (J_{sj}) denote the Gel'fand–Zetlin pattern (9) with $j - i = s$. If $s = j = n$, the summation over α' reduces to one summand $\alpha' = \alpha$ and the CGC $\left(\begin{smallmatrix} m & 2 & m\pm 2 \\ \alpha & (J_{sj}) & \alpha' \end{smallmatrix}\right)$ can be written as $\left(\begin{smallmatrix} m & 2 & m\pm 2 \\ k & 0 & k \end{smallmatrix}\right)$, where k is defined by the first row $(k,\dot{0})$ of the pattern α. A product of CGC's of Eq. (10) can be written down in the form

$$\left(\begin{array}{cc|c} m & 2 & m\pm 2 \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m\pm 2 \\ k & 0 & k \end{array}\right) = \dim D(m\pm 2)\int_{S^{n-1}} t^{nm}_{\alpha 0}(\xi)\, t^{n2}_{00}(\xi)\, t^{n,m\pm 2}_{\alpha 0}(\xi)\, d\xi, \tag{11}$$

where S^{n-1} is the sphere $\mathrm{SO}(n)/\mathrm{SO}(n-1)$ and $t^{nm}_{\alpha 0}(\xi)$ are associated spherical functions of the representation $D(m)$ of the group $\mathrm{SO}(n)$ (see Sec. 9.4 of Ref. 10). We represent $t^{nm}_{\alpha 0}(\xi)$ as a product $t^{nm}_{k0}(g_n(\theta))t^{n-1,k}_{\beta 0}(\eta)$, where $t^{n-1,k}_{\beta 0}(\eta)$, $\eta \in S^{n-2}$, is an associated spherical function of the representation $D_{n-1}(k)$ of $\mathrm{SO}(n-1)$. The invariant measure $d\xi$ in (11) has the form

$$d\xi = \frac{\Gamma(n/2)}{\sqrt{\pi}\,\Gamma((n-1)/2)} \sin^{n-2}\theta \, d\theta \, d\eta$$

(see Sec. 9.1 in Ref. 10). Therefore, making use of the orthogonality relation for $t^{n-1,k}_{\beta 0}(\eta)$ in (11), we have

$$\left(\begin{array}{cc|c} m & 2 & m\pm 2 \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m\pm 2 \\ k & 0 & k \end{array}\right) = \frac{\Gamma(n/2)\dim D(m\pm 2)}{\sqrt{\pi}\,\Gamma((n-1)/2)\dim D_{n-1}(k)} I^{\pm}, \tag{12}$$

where

$$I^{\pm} = \int_0^{\pi} t^{nm}_{k0}(g_n(\theta)) t^{n2}_{00}(g_n(\theta)) t^{n,m\pm 2}_{k0}(g_n(\theta)) \sin^{n-2}\theta \, d\theta. \tag{13}$$

The functions $t^{nm}_{k0}(g_n(\theta))$ are expressed by means of Gegenbauer polynomials $C^{k+(n-2)/2}_{m-k}(\cos\theta)$ (see Sec. 9.4 in Ref. 10), and

$$t^{n2}_{00}(g_n(\theta)) = [n/(n-1)](\cos^2\theta - 1/n).$$

According to this formula one has

$$I^{+} = A\int_{-1}^{1} C^{p+k}_{m-k}(t) C^{p+k}_{m-k+2}(t)(t^2 - 1/(2p+2))(1-t^2)^{p+k-1/2}\, dt,$$

where $p = (n-2)/2$ and

$$A = 2^{2k+1}\left[\frac{\Gamma(p+k)}{\Gamma(p)}\right]^2 \frac{(p+1)(2p-1)!m!(m-k)!(2p+k-2)!(2p+2k-1)}{(2p+1)(2p+m-1)!(2p+m+k-1)!k!} \times\left[\frac{(m+1)(m+2)(m-k+1)(m-k+2)}{(2p+m+k)(2p+m+k+1)(2p+m)(2p+m+1)}\right]^{1/2}.$$

Because of the orthogonality relation for Gegenbauer polynomials we obtain

$$I^{+} = A\int_{-1}^{1} [tC^{p+k}_{m-k}(t)][tC^{p+k}_{m-k+2}(t)](1-t^2)^{p+k-1/2}\, dt.$$

For the product $tC^{p+k}_{m-k}(t)$ we make use of the recurrence relation

$$tC^{p+k}_{m-k}(t) = \frac{m-k+1}{2(p+m)} C^{p+k}_{m-k+1}(t) + \frac{2p+m+k-1}{2(p+m)} C^{p+k}_{m-k-1}(t).$$

The same relation is used for the product $tC^{p+k}_{m-k+2}(t)$. Taking into account again the orthogonality relation for Gegenbauer polynomials we have

$$I^{+} = \frac{\sqrt{\pi}(p+1)\Gamma(p+1/2)m!(2p+k-2)!(2p+2k-1)}{2(2p+1)\Gamma(p)k!(2p+m-1)!(p+m)(p+m+1)(p+m+2)} \times\left[\frac{(m+1)(m+2)(m-k+1)(m-k+2)(2p+m+k)(2p+m+k+1)}{(2p+m)(2p+m+1)}\right]^{1/2}.$$

Thus

$$\left(\begin{array}{cc|c} m & 2 & m+2 \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m+2 \\ k & 0 & k \end{array}\right) = \frac{(p+1)[(2p+m)(2p+m+1)(2p+m+k)(2p+m+k+1)(m-k+1)(m-k+2)]^{1/2}}{2(2p+1)(p+m)(p+m+1)[(m+1)(m+2)]^{1/2}}.$$

In the same way we obtain that

$$\left(\begin{array}{cc|c} m & 2 & m-2 \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m-2 \\ k & 0 & k \end{array}\right) = \frac{(p+1)[m(m-1)(m-k)(m-k-1)(2p+m+k-1)(2p+m+k-2)]^{1/2}}{2(2p+1)(p+m)(p+m-1)[(2p+m-1)(2p+m-2)]^{1/2}},$$

$$\left(\begin{array}{cc|c} m & 2 & m \\ 0 & 0 & 0 \end{array}\right)\left(\begin{array}{cc|c} m & 2 & m \\ k & 0 & k \end{array}\right) = \frac{p(m-k)(2p+m+k) - k(p+k-1)}{2(2p+1)(p+m-1)(p+m+1)}.$$

Now it follows from Eq. (10) that

$$J_{nn}|m;\alpha\rangle = -\sqrt{-1}\left(M+\frac{n}{2}\right)\frac{(n-2)(m-k)(m+k+n-2)-k(2k+n-4)}{[2n(n-1)]^{1/2}(2m+n)(2m+n-4)}|m,\alpha\rangle$$

$$+[(M-m+2)(M+m+n-2)]^{1/2}$$

$$\times\frac{[n(m-k)(m-k-1)(m+k+n-3)(m+k+n-4)]^{1/2}}{[2(n-1)(2m+n-2)(2m+n-6)]^{1/2}(2m+n-4)}|m-2,\alpha\rangle-[(M-m)(M+m+n)]^{1/2}$$

$$\times\frac{[n(m+k+1)(m+k+n-1)(m+k+n-2)(m-k+2)]^{1/2}}{[2(n-1)(2m+n+2)(2m+n-2)]^{1/2}(2m+n)}|m+2,\alpha\rangle. \qquad (14)$$

The rest of the infinitesimal operators (8a) and (8b) are obtained by making commutation of J_{nn} with the infinitesimal operators of representations of SO(n).

V. INFINITESIMAL OPERATORS FOR THE GROUPS SL(n,R) IN AN SO(n) BASIS

The results of Sec. IV allow us to obtain the expressions for infinitesimal operators of the representations π_σ^+ and π_σ^- of the most degenerate series[8,9] of SL(n,R). The reduction of $\pi_\sigma^\pm$ onto the subgroup SO(n) decomposes into a sum of the irreducible representations $D(m)$, for which m are even for π_σ^+ and odd for π_σ^-. For the group SL(n,R) we have the matrices

$$I_{ij}=E_{ij}+E_{ji},\quad i<j,\quad I_{kk}=\left(\frac{2}{n(n-1)}\right)^{1/2}\left(E_{kk}-\frac{1}{n}\sum_{j=1}^{n}E_{jj}\right),\quad k=1,2,\ldots,n,$$

instead of the matrices (8a) and (8b). They form the Gel'fand–Zetlin basis for the irreducible representation $D(2)$ of SO(n).

Utilizing the results of Refs. 8 and 9, and of Sec. IV, we obtain

$$\pi_\sigma^\pm(I_{nn})|m,\alpha\rangle=\left(\sigma-\frac{n}{2}\right)\frac{(n-2)(m-k)(m+k+n-2)-k(2k+n-4)}{[2n(n-1)]^{1/2}(2m+n)(2m+n-4)}|m,\alpha\rangle$$

$$+(\sigma+m)\frac{[n(m-k+1)(m-k+2)(m+k+n-1)(m+k+n-2)]^{1/2}}{[2(n-1)(2m+n+2)(2m+n-2)]^{1/2}(2m+n)}|m+2,\alpha\rangle$$

$$+(\sigma-m-n+2)\frac{[n(m-k)(m-k-1)(m+k+n-3)(m+k+n-4)]^{1/2}}{[2(n-1)(2m+n-2)(2m+n-6)]^{1/2}(2m+n-4)}|m-2,\alpha\rangle.$$

By making commutation of $\pi_\sigma^\pm(I_{nn})$ with infinitesimal operators for the subgroup SO(n), we derive the other infinitesimal operators of the representations $\pi_\sigma^\pm$.

VI. INFINITESIMAL OPERATORS FOR THE GROUP Sp(n) IN A U(n) BASIS

We consider the irreducible representation $d(M)$ of Sp(n) with highest weights $(M,\dot{0})$, $M\geqslant 0$. The reduction of $d(M)$ onto the subgroup U(n) decomposes into a sum of the irreducible representations $D(m_1,m_2)$ of U(n) with highest weights $(m_1,\dot{0},m_2)$, for which m_1-m_2 and M are of the same evenness and $m_1-m_2\leqslant M$. Multiplicities equal one. The Gel'fand–Zetlin bases of the representations $D(m_1,m_2)$ constitute the orthonormal basis for $d(M)$. The basis elements are denoted by $|m_1,m_2;\alpha\rangle$, where α are the Gel'fand–Zetlin patterns for the representation $D(m_1,m_2)$ without the first row.

The Lie algebra sp(n) of Sp(n) is represented as the sum u(n) + $\mathfrak{p}$, where u(n) is the Lie algebra of U(n). The complexification $\mathfrak{p}_c$ of $\mathfrak{p}$ forms a carrier space for a sum of two irreducible representations[11] $D(2,0)$ and $D(0,-2)$ of u(n) with respect to the action ad $X, X\in$u(n). We take the matrices

$$J_+=i(E_{nn}-E_{2n,2n})-(E_{n,2n}+E_{2n,n}),\quad J_-=i(E_{nn}-E_{2n,2n})+(E_{n,2n}+E_{2n,n})$$

from $\mathfrak{p}_c$. They are invariant with respect to the operators ad X, $X\in$u($n-1$). The matrices J_+ and J_- belong to the carrier spaces of $D(2,0)$ and $D(0,-2)$, respectively.

The infinitesimal operators J_+ and J_- of the representation $d(M)$ are defined by Eqs. (29) and (30) of Ref. 11. The coefficients of these equations are expressed by means of CGC's of the tensor products $D(m_1,m_2)\otimes D(2,0)$ and $D(m_1,m_2)\otimes D(0,-2)$. Utilizing the expressions for these CGC's from Ref. 12, after some simplifications, we obtain

$$J_+|m_1,m_2;\alpha\rangle = \frac{[(m_1-n_1+1)(m_1-n_1+2)(m_1-n_2+n-1)(m_1-n_2+n)]^{1/2}}{(m_1-m_2+n)[(m_1-m_2+n+1)(m_1-m_2+n-1)]^{1/2}}$$
$$\times[(M-m_1+m_2)(M+m_1-m_2+2n)]^{1/2}|m_1+2,m_2;\alpha\rangle$$
$$-\frac{[(n_1-m_2+n-3)(n_1-m_2+n-2)(n_2-m_2-1)(n_2-m_2)]^{1/2}}{(m_1-m_2+n-2)[(m_1-m_2+n-3)(m_1-m_2+n-1)]^{1/2}}$$
$$\times[(M+m_1-m_2+2n-2)(M-m_1+m_2+2)]^{1/2}|m_1,m_2+2;\alpha\rangle$$
$$+\frac{2[(m_1-n_1+1)(m_1-n_2+n-1)(n_1-m_2+n-2)(n_2-m_2)]^{1/2}}{(m_1-m_2+n)(m_1-m_2+n-2)}(M+n)|m_1+1,m_2+1;\alpha\rangle, \tag{15}$$

$$J_-|m_1,m_2;\alpha\rangle = -\frac{[(m_1-n_2+n-3)(m_1-n_2+n-2)(m_1-n_1-1)(m_1-n_1)]^{1/2}}{(m_1-m_2+n-2)[(m_1-m_2+n-1)(m_1-m_2+n-3)]^{1/2}}$$
$$\times[(M+m_1-m_2+2n-2)(M-m_1+m_2+2)]^{1/2}|m_1-2,m_2;\alpha\rangle$$
$$-\frac{[(n_2-m_2+1)(n_2-m_2+2)(n_1-m_2+n-1)(n_1-m_2+n)]^{1/2}}{(m_1-m_2+n)[(m_1-m_2+n+1)(m_1-m_2+n-1)]^{1/2}}$$
$$\times[(M-m_1+m_2)(M+m_1-m_2+2n)]^{1/2}|m_1,m_2-2;\alpha\rangle$$
$$-\frac{2[(n_2-m_2+1)(n_1-m_2+n-1)(m_1-n_2+n-2)(m_1-n_1)]^{1/2}}{(m_1-m_2+n)(m_1-m_2+n-2)}(M+n)|m_1-1,m_2-1;\alpha\rangle. \tag{16}$$

The other infinitesimal operators of the representation $d(M)$ can be obtained by making commutation of J_+ and J_- with infinitesimal operators for the subgroup $\mathrm{U}(n)$.

Equations (15) and (16) define the infinitesimal operators of the most degenerate series representations of the group $\mathrm{Sp}(n,R)$ in a $\mathrm{U}(n)$ basis.[11]

VII. INFINITESIMAL OPERATORS FOR THE GROUP GL(n,C) IN A U(n) BASIS

We consider the representations $\pi^{\sigma q}$, $\sigma\in C$, $q\in\{0, \pm 1, \pm 2,...\}$, of $\mathrm{GL}(n,C)$ introduced in Ref. 13. The reduction of $\pi^{\sigma q}$ onto the subgroup $\mathrm{U}(n)$ decomposes into a sum of the irreducible representations $D(m_1,m_2)$ with highest weights $(m_1,\dot{0},m_2)$, for which $m_1+m_2=q$. The elements of the Gel'fand–Zetlin bases for the representations $D(m_1,m_2)$ are denoted by $|m_1,m_2;\alpha\rangle$.

The infinitesimal operators of the representation $\pi^{\sigma q}$ are defined by Eq. (25) of Ref. 13. The coefficients of this equation are expressed by means of CGC's

$$K=\left(\begin{matrix} D(m_1,m_2) & D(1,-1) \\ 0 & 0 \end{matrix}\left|\begin{matrix} D(m_1',m_2') \\ 0 \end{matrix}\right.\right)\left(\begin{matrix} D(m_1,m_2) & D(1,-1) \\ D'(k_1,k_2) & 0 \end{matrix}\left|\begin{matrix} D(m_1',m_2') \\ D'(k_1,k_2) \end{matrix}\right.\right), \tag{17}$$

where $D'(k_1,k_2)$ is an irreducible representation of the subgroup $\mathrm{U}(n-1)$ with highest weight $(k_1,\dot{0},k_2)$, $k_1\geqslant 0\geqslant k_2$, and (m_1',m_2') takes the values (m_1+1,m_2-1), (m_1-1,m_2+1), (m_1,m_2). The coefficients (17) are evaluated in the same manner as in the case of the representations of $\mathrm{U}(n)$ in an $\mathrm{SO}(n)$ basis.

Utilizing the expression (54) of Ref. 14 for the coefficient (17), we obtain the formula analogous to Eq. (12),

$$K=2(n-1)\frac{\dim D(m_1',m_2')}{\dim D'(k_1,k_2)}\int_0^{\pi/2} t^{n,m_1,m_2}_{(k_1,k_2)\dot{0}}(g_n(\theta))t^{n,1,-1}_{00}(g_n(\theta))t^{n,m_1',m_2'}_{(k_1,k_2)\dot{0}}(g_n(\theta))\sin^{2n-3}\theta\cos\theta\, d\theta. \tag{18}$$

The matrix elements $t^{n,m_1,m_2}_{(k_1,k_2)\dot{0}}(g_n(\theta))$ are defined by Eq. (48) of Ref. 14 under appropriate meaning of highest weights. After laborious evaluations we find that this matrix element is expressed by means of the hypergeometric function ${}_2F_1$, which, in turn, is expressed through the Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$. One has

$$t^{n,m_1,m_2}_{(k_1,k_2)\dot{0}}(g_n(\theta))=(-1)^{k_1}\Bigg[(n-2)(k_1-k_2+n-2)\frac{(m_1-k_1)!m_1!(k_1+n-3)!}{(k_2-m_2)!(m_1-k_2+n-2)!}$$
$$\times\frac{(-k_2+n-3)!(k_1-m_2+n-2)!(-m_2)!}{(-m_2+n-2)!(m_1+n-2)!k_1!(-k_2)!}\Bigg]^{1/2}(\tan\theta)^{-k_1-k_2}(\cos\theta)^{-m_1-m_2}$$
$$\times(\sin\theta)^{2k_1}P^{(k_1-k_2+n-2,k_1+k_2-m_1-m_2)}_{m_1-k_1}(\cos 2\theta). \tag{19}$$

It follows from here that

$$t_{00}^{n,1,-1}(g_n(\theta)) = 1 - [n/(n-1)]\sin^2\theta.$$

The integration in Eq. (18) is fulfilled in the same manner as in Eq. (13). We apply to the Jacobi polynomial of Eq. (19) the recurrence relation

$$(1-x)P_m^{(\alpha,\beta)}(x) = [2/(2m+\alpha+\beta+1)][(m+\alpha)P_m^{(\alpha-1,\beta)}(x) - (m+1)P_{m+1}^{(\alpha-1,\beta)}(x)],$$

and then the recurrence relation

$$P_m^{(\alpha-1,\beta)}(x) = [1/(2m+\alpha+\beta)][(m+\alpha+\beta)P_m^{(\alpha,\beta)}(x) - (m+\beta)P_{m-1}^{(\alpha,\beta)}(x)].$$

Further, we use the orthogonality relation for Jacobi polynomials. After some simplifications we obtain that for $(m_1',m_2') = (m_1+1,m_2-1)$,

$$K = \frac{n[(m_1-k_2+n-1)(k_1-m_2+n-1)(m_1+n-1)(-m_2+n-1)(k_2-m_2+1)(m_1-k_1+1)]^{1/2}}{(n-1)(m_1-m_2+n-1)(m_1-m_2+n)[(m_1+1)(-m_2+1)]^{1/2}},$$

for $(m_1',m_2') = (m_1-1,m_2+1)$,

$$K = \frac{n[(k_1-m_2+n-2)(m_1-k_2+n-2)(m_1-k_1)(k_2-m_2)m_1(-m_2)]^{1/2}}{(n-1)(m_1-m_2+n-2)(m_1-m_2+n-1)[(m_1+n-2)(-m_2+n-2)]^{1/2}},$$

and for $(m_1',m_2') = (m_1,m_2)$,

$$K = 1 - \frac{n(k_1-m_2+n-1)(m_1-k_1+1)}{(n-1)(m_1-m_2+n-1)(m_1-m_2+n)} - \frac{n(m_1-k_2+n-2)(k_1-m_2+n-2)}{(n-1)(m_1-m_2+n-2)(m_1-m_2+n-1)}.$$

Now, according to Eq. (25) of Ref. 13 one has

$$\begin{aligned}\pi^{\sigma q}(E_{nn})|m_1,m_2;\alpha\rangle &= \frac{[(m_1-k_2+n-1)(k_1-m_2+n-1)(m_1-k_1)(k_2-m_2)]^{1/2}}{(m_1-m_2+n-2)[(m_1-m_2+n-1)(m_1-m_2+n-3)]^{1/2}}\\ &\quad\times\frac{n}{n-1}(\sigma-m_1+m_2-2n+2)|m_1-1,m_2+1;\alpha\rangle\\ &\quad+\frac{n[(m_1-k_2+n-1)(k_1-m_2+n-1)(m_1-k_1+1)(k_2-m_2+1)]^{1/2}}{[(m_1-m_2+n-1)(m_1-m_2+n+1)]^{1/2}(m_1-m_2+n)(n-1)}\\ &\quad\times(\sigma+m_1-m_2)|m_1+1,m_2-1;\alpha\rangle+\left[1-\frac{n(k_1-m_2+n-1)(m_1-k_1+1)}{(n-1)(m_1-m_2+n-1)(m_1-m_2+n)}\right.\\ &\quad\left.-\frac{n(m_1-k_2+n-2)(k_1-m_2+n-2)}{(n-1)(m_1-m_2+n-2)(m_1-m_2+n-1)}\right](\sigma-n)|m_1,m_2;\alpha\rangle. \qquad (20)\end{aligned}$$

The other infinitesimal operators can be found by making commutation of $\pi^{\sigma q}(E_{nn})$ with infinitesimal operators of the representations of U(n).

[1] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk USSR **71**, 825 (1950).
[2] I. M. Gel'fand and M. L. Zetlin, Dokl. Akad. Nauk USSR **71**, 1017 (1950).
[3] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[4] A. U. Klimyk and B. Gruber, J. Math. Phys. **23**, 1399 (1982).
[5] A. U. Klimyk and B. Gruber, J. Math. Phys. **20**, 2011 (1979).
[6] A. U. Klimyk and B. Gruber, J. Math. Phys. **20**, 1995 (1979).
[7] A. U. Klimyk, *Matrix Elements and Clebsch–Gordan Coefficients of Group Representations* (Naukova Dumka, Kiev, 1979) (in Russian).
[8] B. Gruber and A. U. Klimyk, J. Math. Phys. **22**, 2762 (1981).
[9] A. U. Klimyk and I. I. Kachurik, *Computation Methods in Theory of Group Representations* (Vyshcha Shkola, Kiev, 1986) (in Russian).
[10] N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations* (Am. Math. Soc., Providence, RI, 1968).
[11] A. U. Klimyk, J. Math. Phys. **24**, 224 (1983).
[12] E. Chacon, M. Ciftan, and L. C. Biedenharn, J. Math. Phys. **13**, 577 (1972).
[13] A. U. Klimyk, J. Phys. A **15**, 3009 (1982).
[14] A. U. Klimyk and A. M. Gavrilik, J. Math. Phys. **20**, 1624 (1979).

Some isoscalar factors for $SO_N \supset SO_{N-1}$ and state expansion coefficients for $SO_N \supset SO_{N-1}$ in terms of $SU_N \supset SU_{N-1} \supset SO_{N-1}$

Feng Pan
Department of Physics, Liaoning Normal University, Dalian, People's Republic of China
Yu-Fang Cao
Department of Physics, East China Normal University, Shanghai, People's Republic of China

(Received 28 January 1988; accepted for publication 20 July 1988)

General analytical expressions of state expansion coefficients for $SO_N \supset SO_{N-1}$ in terms of $SU_N \supset SU_{N-1} \supset SO_{N-1}$ are derived. Analytical expressions of isoscalar factors for $SO_N \supset SO_{N-1}$ for coupling $(w_1w_2) \times (w_30)$ to $(w_1'w_2')$ for the SO_N $(N > 4)$ irrep and $(v_1v_2) \times (v_30)$ to $(v_1'v_2')$ for the SO_{N-1} irrep with $w_1 + w_2 + w_3 = w_1' + w_2'$ and $v_1 + v_2 + v_3 = v_1' + v_2'$ are obtained by using these coefficients and isoscalar factors for $SU_N \supset SU_{N-1}$.

I. INTRODUCTION

Isoscalar factors for $SO_N \supset SO_{N-1}$ are very useful in many physical problems, for example, in IBM-2, isoscalar factors for coupling $(w_1w_2) \times (w_30)$ to $(w_1'w_2')$ for SO_6 irreps in the chain of $SO_6 \supset SO_5$ are important in deriving wave functions and calculating electromagnetic transition rates and nucleon-transfer intensities in the SO_6 limit case.[1]

Some isoscalar factors for low N, especially for $SO_4 \supset SO_3$ and $SO_3 \supset SO_2$ [which are simply Clebsch–Gordan (CG) coefficients] have been given by many authors.[2,3] Isoscalar factors for $SO_N \supset SO_{N-1}$ for coupling $l_1 \times l_2$ to (L_1L_2) with all three irreps symmetric (i.e., with $L_2 = 0$) were considered by Gavrilik[4] and Kildyushov and Kuznetsov,[5] but only Norvaisas and Alisauskas observed that such isoscalar factors are the analytical continuation of semistretched isoscalar factors of $Sp_4 \supset SU_2 \times SU_2$ of the second kind, introduced by Alisauskas and Jucys.[6] The substitution group technique of Sp_4 (SO_5)[7] allowed them to derive expressions for isoscalar factors of $SO_N \supset SO_{N-1}$ for the coupling $l_1 \times l_2$ to (L_1L_2) with its resulting subgroup irrep $(L_1'L_2')$.[8–10] But these results are very complicated, especially for L_2 and $L_2' \neq 0$. These isoscalar factors are expressed by the sums over a series of SU_2 $6j$ and $9j$ coefficients, and can only be simplified in a few special cases.[10] However, the polynomial-type expressions (with as few summation signs as possible) are useful in several aspects (cf. Ref. 11); most notably, by their use one can obtain exact values of the quantities that they represent in a much shorter computing time when compared with the time spent by the standard crude expressions.

In Ref. 12, in order to calculate the reduced matrix elements under the $SO_6 \supset SO_5$ group chain, the basis vectors for $SO_6 \supset SO_5$ symmetric irreps are expanded in terms of $SU_5 \supset SO_5$ basis vectors; this method can easily be extended to the general $SO_N \supset SO_{N-1}$ case. This method along with the isoscalar factors for $SU_N \supset SU_{N-1}$ given by Ref. 13 make it possible to derive the analytical expressions of isoscalar factors for $SO_N \supset SO_{N-1}$ for the coupling $(w_1w_2) \times (w_30)$ to $(w_1'w_2')$ for SO_N irrep and $(v_1v_2) \times (v_30)$ to $(v_1'v_2')$ for SO_{N-1} irrep with $w_1 + w_2 + w_3 = w_1' + w_2'$ and $v_1 + v_2 + v_3 = v_1' + v_2'$. The results are simple and will be expressed in the polynomial-type forms.

II. EXPANSION COEFFICIENTS

First, we use boson creation (annihilation) operators $l_\mu^+, \tilde{l}_\mu\ [= (-)^{l+\mu} l_{-\mu}]$, $\mu = -l, -l+1, \ldots, l$, and s^+, s to construct SU_{2l+2} generators, they are

$$\begin{aligned}&(l^+l)^{(k)}_q, \quad k = 1,2,\ldots,2l, \quad q = -k, -k+1,\ldots,k,\\ &(l^+s)^{(l)}_q, (s^+1)^{(l)}_q, \quad q = -l, -l+1,\ldots,l.\end{aligned} \tag{2.1}$$

Obviously, SU_{2l+2} has the following decompositions:

$$\begin{tikzcd} & SU_{2l+1} \arrow[dr] & \\ SU_{2l+2} \arrow[ur] \arrow[dr] & & SO_{2l+1}, \\ & SO_{2l+2} \arrow[ur] & \end{tikzcd} \tag{2.2}$$

then we introduce the SU(1,1) group generators

$$\begin{aligned}S_+ &= l^+ \cdot l^+/2 - s^+s^+/2,\\ S_- &= l \cdot l/2 - ss/2,\\ S_0 &= \frac{1}{4}\sum_m (l_m^+ l_m + l_m l_m^+) + \frac{1}{4}(s^+s + ss^+),\end{aligned} \tag{2.3}$$

which satisfy the following commutation relations:

$$[S_0, S_\pm] = \pm S_\pm, \quad [S_+, S_-] = -2S_0. \tag{2.4}$$

Let the basis vectors of $SU_{2l+2} \supset SO_{2l+2} \supset SO_{2l+1}$ be $|Nwv\Omega\rangle$, where Ω stands for other quantum numbers; the matrix element of S_+S_- under this group chain is

$$\langle S_+S_- \rangle = S_0(S_0 - 1) - S(S-1). \tag{2.5}$$

The relations between S_0, S and N, w are

$$S_0 = N/2 + (l+1)/2, \quad S = w/2 + (l+1)/2; \tag{2.6}$$

then (2.5) can be rewritten as

$$\langle S_+S_-\rangle = \tfrac{1}{4}(N-w)(N+w+2l). \qquad (2.7)$$

In addition, let the basis vectors of $SU_{2l+2} \supset SU_{2l+1} \supset SO_{2l+1}$ be $|Nnv\Omega\rangle$. The matrix element of S_+S_- under this group chain is

$$\begin{aligned}\langle Nnv\Omega|S_+S_-|N'n'v'\Omega'\rangle &= \delta_{NN'}\delta_{vv'}\delta_{\Omega\Omega'}[(\tfrac{1}{4}(n-v)(n+v+2l-1)\\ &\quad +\tfrac{1}{4}(N-n)(N-n-1))\delta_{nn'}\\ &\quad -\tfrac{1}{4}[(n+v+2l-1)\\ &\quad \times(n-v)(N-n+2)(N-n+1)]^{1/2}\delta_{n'n-2}\\ &\quad -\tfrac{1}{4}[(n+v+2l+1)(n-v+2)(N-n)\\ &\quad \times(N-n-1)]^{1/2}\delta_{n'n+2}]. \qquad (2.8)\end{aligned}$$

Third, $|Nwv\Omega\rangle$ can be expanded in terms of $|Nnv\Omega\rangle$ as

$$|Nwv\Omega\rangle = \sum_n B^{wv}_{n(\bar{N})}|Nnv\Omega\rangle, \qquad (2.9)$$

where $\bar{N} = 2l+2$. Acting with S_+S_- on (2.9) and using the results of (2.7) and (2.8) (let $N = w$), we have

$$\begin{aligned}B^{wv}_{n(\bar{N})}&[(n-v)(n+v+\bar{N}-3)+(w-n)(w-n-1)]\\ &= B^{wv}_{n-2}(\bar{N})[(n+v+\bar{N}-3)\\ &\quad \times(n-v)(w-n+2)(w-n+1)]^{1/2}\\ &\quad + B^{wv}_{n+2(\bar{N})}[(n-v+2)(n+v+\bar{N}-1)\\ &\quad \times(w-n)(w-n-1)]^{1/2}. \qquad (2.10)\end{aligned}$$

By using (2.10), $B^{wv}_{n(\bar{N})}$ can be expressed as

$$B^{wv}_{n(\bar{N})} = B^{wv}_{0(\bar{N})} \times\left[\frac{(v+\bar{N}-3)!!w!}{(w-n)!(n+v+\bar{N}-3)!!(n-v)!!}\right]^{1/2}, \qquad (2.11)$$

and using the relation

$$\sum_n \frac{(w-v)!(n-v+1)!!}{(w-n)!(n+v+\bar{N}-3)!!(n-v+1)!} = \frac{(2w+\bar{N}-4)!!}{(w+v+\bar{N}-3)!}, \qquad (2.12)$$

we have

$$B^{wv}_{0(\bar{N})} = \left[\frac{(w-v)!(w+v+\bar{N}-3)!}{(v+\bar{N}-3)!!(2w+\bar{N}-4)!!w!}\right]^{1/2}, \qquad (2.13)$$

with the sum over n in Eq. (2.12) going over even or odd values depending on whether v is even or odd. The term $B^{wv}_{v(\bar{N})}$ is also valid when $\bar{N}$ is an odd integer, which will be proved in Sec. IV.

III. GENERAL METHODS OF EVALUATION

For the tow-rowed irrep of SU_N, we can also write the relation between the group chains $SU_N \supset SO_N \supset SO_{N-1}$ and $SU_N \supset SU_{N-1} \supset SO_{N-1}$ similar to Eq. (2.9),

$$|[N_1N_2](N_1N_2)(v_1v_2)\Omega\rangle = \sum_{\substack{n_1n_2\\ k}} B^{(N_1N_2)(v_1v_2)}_{(n_1n_2)k(N)}|[N_1N_2](n_1n_2)k(v_1v_2)\Omega\rangle, \qquad (3.1)$$

where $B^{(N_1N_2)(v_1v_2)}_{(n_1n_2)k(N)}$ is the expansion coefficients, and k is the multiplicity label, because $SU_{N-1} \supset SO_{N-1}$ is not fully reducible for tow-rowed irreps.

Next, coupling tow states given by Eq. (3.1), with one of them symmetric, and using the Racah factorization Lemma,[14] we have

$$\begin{aligned}&B^{(w_1w_2)(v_1v_2)}_{(n_1n_2)\bar{k}(N)}B^{w_3v_3}_{n_3}\left[\begin{array}{c|cc|c} SO_N & (w_1w_2) & (w_30) & (w_1'w_2')\\ SO_{N-1} & (v_1v_2) & (v_30) & l(v_1'v_2')\end{array}\right]\\ &= \sum_{n_1'n_2'k} B^{(w_1'w_2')(v_1'v_2')}_{(n_1'n_2')k(N)}\left[\begin{array}{c|cc|c} SU_N & (w_1w_2) & (w_30) & (w_1'w_2')\\ SU_{N-1} & (n_1n_2) & (n_30) & (n_1'n_2')\end{array}\right]\left[\begin{array}{c|cc|c} SU_{N-1} & (n_1n_2) & (n_30) & (n_1'n_2')\\ SO_{N-1} & (v_1v_2)\bar{k} & (v_30) & lk(v_1'v_2')\end{array}\right], \qquad (3.2)\end{aligned}$$

where

$$\left[\begin{array}{c|cc|c} G & (w_1w_2) & (w_30) & (w_1'w_2')\\ g & (v_1v_2) & (v_30) & l(v_1'v_2')\end{array}\right],$$

etc., are isoscalar factors for $G \supset g$, l is the multiplicity label for SO_{N-1}, and

$$\begin{aligned}&\sum_{\substack{v_1n_1n_2k\\ v_2n_1'n_2'}} B^{(w_1'w_2')(v_1'v_2')}_{(n_1'n_2')k(N)}\left[\begin{array}{c|cc|c} SU_N & (w_1'0) & (w_2'0) & (w_1'w_2')\\ SU_{N-1} & (n_10) & (n_20) & (n_1'n_2')\end{array}\right]\\ &\times\left[\begin{array}{c|cc|c} SU_{N-1} & (n_10) & (n_20) & (n_1'n_2')\\ SO_{N-1} & (v_10) & (v_20) & k(v_1'v_2')\end{array}\right]\left[\begin{array}{c|cc|c} SO_N & (w_10) & (w_20) & (w_1'w_2')\\ SO_{N-1} & (v_10) & (v_20) & (v_1'v_2')\end{array}\right]B^{w_1'v_1}_{n_1(N)}B^{w_2'v_2}_{n_2(N)} = 1. \qquad (3.3)\end{aligned}$$

In comparing with the condition

$$\sum_{n_1n_2k}(B^{(w_1w_2)(v_1v_2)}_{(n_1n_2)k(N)})^2 = 1, \qquad (3.4)$$

we obtain

$$B^{(w_1'w_2')(v_1'v_2')}_{(n_1'n_2')k(N)} = \sum_{\substack{n_1n_2\\ v_1v_2}} \left[\begin{array}{c|cc|c} \mathrm{SU}_N & (w_1'0) & (w_1'0) & (w_1'w_2') \\ \mathrm{SU}_{N-1} & (n_10) & (n_20) & (n_1'n_2') \end{array}\right] \left[\begin{array}{c|cc|c} \mathrm{SU}_{N-1} & (n_10) & (n_20) & (n_1'n_2') \\ \mathrm{SO}_{N-1} & (v_10) & (v_20) & k(v_1'v_2') \end{array}\right]$$

$$\times \left[\begin{array}{c|cc|c} \mathrm{SO}_N & (w_1'0) & (w_2'0) & (w_1'w_2') \\ \mathrm{SO}_{N-1} & (v_10) & (v_20) & (v_1'v_2') \end{array}\right] B^{w_1'v_1}_{n_1(N)} B^{w_2'v_2}_{n_2(N)} \; . \tag{3.5}$$

By using the state expansion coefficients given by Eqs. (2.11) and (2.13) together with Eq. (3.2) and isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$, some isoscalar factors for $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$ can be derived and the state expansion coefficients given by Eq. (3.5) can also be obtained for some special cases. Some results and a detailed evaluation will be given in the next section.

IV. THE METHOD EXEMPLIFIED

For the symmetric irrep of SU_N, $(w_1'w_2') = (w_1 + w_20)$, the coefficient $B^{(w_1+w_2)v}_{n(N)}$ is known,

$$\left[\begin{array}{c|cc|c} \mathrm{SU}_N & (w_10) & (w_20) & (w_1+w_20) \\ \mathrm{SU}_{N-1} & (w_10) & (w_20) & (w_1+w_20) \end{array}\right] = 1,$$

and inserting them in Eq. (3.2), we get

$$\left[\begin{array}{c|cc|c} \mathrm{SO}_n & (w_10) & (w_20) & (w_1+w_20) \\ \mathrm{SO}_{n-1} & (v_10) & (v_20) & (v0) \end{array}\right]$$

$$= (-)^b \left[\frac{(w_1+w_2+v+n-3)!(w_1+w_2-v)!(2w_1+n-4)!!(2w_2+n-4)!!w_1!w_2!v!(2v_1+n-3)(2v_2+n-3)}{2^{n-4}(n-5)!!(v+n-4)!(w_1+w_2)!(v_1+v_2+v+n-3)!!(v-v_1+v_2)!!(v_1-v_2+v)!!(v_1+v_2-v)!!}\right.$$
$$\left.\times\frac{(v_1+v_2+v+2n-8)!!(v-v_1+v_2+n-5)!!(v_1-v_2+v+n-5)!!(v_1+v_2-v+n-5)!!}{(2w_1+2w_2+n-4)!!(w_1-v_1)!(w_2-v_2)!(w_1+v_1+n-3)!(w_2+v_2+n-3)!}\right]^{1/2}, \tag{4.1}$$

where

$$b = \begin{cases} \frac{1}{2}(v - v_1 - v_2), & \text{for } n = 4, \\ 0, & \text{for } n > 4. \end{cases} \tag{4.2}$$

This formula is valid for $n \geqslant 4$ and the isoscalar factor for $\mathrm{SU}_{n-1} \supset \mathrm{SO}_{n-1}$ (see Ref. 15) has been used in deriving Eq. (4.1).

By using the special isoscalar factor for $\mathrm{SU}_{N-1} \supset \mathrm{SO}_{N-1}$,

$$\left[\begin{array}{cc|c} (n_1n_2) & (n_30) & (n_1'n_2') \\ (n_1n_2) & (n_30) & (n_1'n_2') \end{array}\right] = 1,$$

an important relation between the isoscalar factors for $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$ and $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ may be derived:

$$\left[\begin{array}{c|cc|c} \mathrm{SO}_N & (w_1w_2) & (w_30) & (w_1'w_2') \\ \mathrm{SO}_{N-1} & (v_1v_2) & (v_30) & (v_1'v_2') \end{array}\right]$$

$$= B^{(w_1'w_2')(v_1'v_2')}_{(v_1'v_2')(N)} \left(B^{(w_1w_2)(v_1v_2)}_{(v_1v_2)(N)} B^{(w_30)(v_30)}_{(v_30)(N)}\right)^{-1} \left[\begin{array}{c|cc|c} \mathrm{SU}_N & (w_1w_2) & (w_30) & (w_1'w_2') \\ \mathrm{SU}_{N-1} & (v_1v_2) & (v_30) & (v_1'v_2') \end{array}\right], \tag{4.3}$$

where the conditions $w_1 + w_2 + w_3 = w_1' + w_2'$ and $v_1 + v_2 + v_3 = v_1' + v_2'$ should be satisfied. In this case the resulting irrep $(w_1'w_2')$ or $(v_1'v_2')$ does not occur more than once; thus all the multiplicity labels can be omitted.

The expansion coefficient $B^{(w_1w_2)v}_{v(N)}$ can be derived from Eq. (4.3) by using the more special isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ and $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$,

$$\left[\begin{array}{cc|c} (w_10) & (w_20) & (w_1w_2) \\ (v0) & 0 & (v0) \end{array}\right]$$

(see Refs. 10 and 13); then $B^{(w_1w_2)(v0)}_{(v0)(N)}$ can be expressed as

$$B^{(w_1w_2)(v0)}_{(v0)(N)} = \left[\frac{(w_1+N-4)!(w_2+v+N-4)!(w_1+v+N-3)!(w_2+N-5)!}{(v+N-4)!(w_1+w_2+N-4)!(2v+N-3)!!(N-5)!!} \frac{1}{(2w_1+N-4)!!(2w_2+N-6)!!}\right]^{1/2}. \tag{4.4}$$

The expansion coefficient $B^{(w_1w_2)(v_1v_2)}_{(v_1v_2)(N)}$ can be obtained from the following equation:

$$B^{(w_1w_2)(v_1v_2)}_{(v_1v_2)(N)} = \left\{\sum_{n_1n_2} \left(B^{w_1n_1}_{n_1(N)} B^{w_2n_2}_{n_2(N)}\right)^{-2} \left[\begin{array}{c|cc|c} \mathrm{SU}_N & (w_10) & (w_20) & (w_1w_2) \\ \mathrm{SU}_{N-1} & (n_10) & (n_20) & (v_1v_2) \end{array}\right]^2\right\}^{-1/2} \tag{4.5}$$

and the property of the isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ given in the Appendix:

$$B^{(w_1 w_2)(v_1 v_2)}_{(v_1 v_2)(N)} = \left[\frac{(w_1 + v_2 + N - 4)!(w_2 + v_1 + N - 4)!(w_1 + v_1 + N - 3)!(w_2 + v_2 + N - 5)!}{(v_1 + v_2 + N - 4)!(w_1 + w_2 + N - 4)!(2v_1 + N - 3)!!(2v_2 + N - 5)!!} \right.$$
$$\left. \times \frac{1}{(2w_1 + N - 4)!!(2w_2 + N - 6)!!} \right]^{1/2}. \tag{4.6}$$

Equations (4.5) and (4.6) are valid when $N > 4$.

Theorem 2 of Ref. 9 shows that

$$\left[\begin{array}{c|cc|c} \mathrm{SO}_N & (w_1 0) & (w_2 0) & (w'_1 w'_2) \\ \mathrm{SO}_{N-1} & (v_1 0) & (v_2 0) & (v'_1 v'_2) \end{array} \right]$$
$$= \left[\begin{array}{c|cc|c} \mathrm{SO}_{N'} & w_1 + \frac{1}{2}(N - N') & w_2 + \frac{1}{2}(N - N') & (w'_1 + \frac{1}{2}(N - N') w'_2 + \frac{1}{2}(N - N')) \\ \mathrm{SO}_{N'-1} & v_1 + \frac{1}{2}(N - N') & v_2 + \frac{1}{2}(N - N') & (v'_1 + \frac{1}{2}(N - N') v'_2 + \frac{1}{2}(N - N')) \end{array} \right], \tag{4.7}$$

for $N' > 4$ and $N \geqslant N'$. Because the isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ are N independent, Eq. (4.7) requires that [see Eq. (4.3)]

$$B^{(w_1 + \frac{1}{2}(N - N') w_2 + \frac{1}{2}(N - N'))(v_1 + \frac{1}{2}(N - N') v_2 + \frac{1}{2}(N - N'))}_{(v_1 + \frac{1}{2}(N - N') v_2 + \frac{1}{2}(N - N'))(N')} = B^{(w_1 w_2)(v_1 v_2)}_{(v_1 v_2)(N)}, \tag{4.8}$$

and

$$B^{(w + \frac{1}{2}(N - N')0)(v + \frac{1}{2}(N - N')0)}_{(v + \frac{1}{2}(N - N')0)(N')} = B^{(w0)(v0)}_{(v0)(N)}. \tag{4.9}$$

These conditions are indeed satisfied; thus $B^{(w_1 w_2)(v_1 v_2)}_{(v_1 v_2)(N)}$ is valid for any integer N except for $N \leqslant 4$.

Using the isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ given by Ref. 13 and Eq. (4.3), isoscalar factors of the following type can be derived for $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$:

$$\left[\begin{array}{c|cc|c} \mathrm{SO}_N & (w_1 w_2) & (w_3 0) & (w'_1 w'_2) \\ \mathrm{SO}_{N-1} & (v_1 v_2) & (v_3 0) & (v'_1 v'_2) \end{array} \right]$$
$$= \left[\frac{(w'_1 + v'_2 + N - 4)!(w'_2 + v'_1 + N - 4)!(w'_1 + v'_1 + N - 3)!}{(w_1 + v_2 + N - 4)!(w_2 + v_1 + N - 4)!(w_1 + v_1 + N - 3)!} \right.$$
$$\times \frac{(w'_2 + v'_2 + N - 5)!(2w_1 + N - 4)!!(2w_2 + N - 6)!!(v_1 + v_2 + N - 4)!(w_1 + w_2 + N - 4)!(2v_1 + N - 3)!!}{(w_2 + v_2 + N - 5)!(2w'_1 + N - 4)!!(2w'_2 + N - 6)!!(v'_1 + v'_2 + N - 4)!(w'_1 + w'_2 + N - 4)!(2v'_1 + N - 3)!!}$$
$$\times \frac{(2v_2 + N - 5)!!(2v_3 + N - 3)!!(2w_3 + N - 4)!!(w_3 - v_3)!(v_1 - v_2 + 1)(w'_1 - w'_2 + 1)(v'_1 - v_1)!}{(2v'_2 + N - 5)!!(w_3 + v_3 + N - 3)!(v_1 - v'_2)!(w'_1 - w_1)!(w'_2 - w_2)!(w'_1 - w_2 + 1)!(w'_1 + 1)!}$$
$$\left. \times \frac{(v'_2 - v_2)!(v'_1 - v_2 + 1)!(w_1 - w'_2)!w_2!(w_1 + 1)!(w_1 - v_1)!(w_2 - v_2)!}{w'_2!(v_1 - w_2)!v_2!(v_1 + 1)!(w'_1 - v'_1)!(w'_2 - v'_2)!(w'_1 - v'_2 + 1)!} (w_1 - v_2 + 1)!(v'_1 - w'_2)!v'_2!(v'_1 + 1)! \right]^{1/2}$$
$$\times \sum_{xy} (-)^{x+y} \frac{(v_1 - v_2 + x - y + 1)(v_1 - w_2 + x)!(v_1 - v'_2 + x)!(w'_1 - v_1 - x)!(v_1 - v_2 - y)!(w'_1 - v_2 + 1 - y)!(w'_2 - v_2 - y)!}{x!y!(v_1 - v_2 + x + 1)!(w_1 - v_1 - x)!(v'_1 - v_1 - x)!(v_1 - w'_2 + x)!(w_1 - v_2 + 1 - y)!(w_2 - v_2 - y)!}$$
$$\times \frac{1}{(v_3 - v'_1 + v_1 - y)!(v_1 - v'_2 + v_3 + 1 - y)!}, \quad \text{for } N > 4, \quad w_1 + w_2 + w_3 = w'_1 + w'_2, \quad v_1 + v_2 + v_3 = v'_1 + v'_2. \tag{4.10}$$

V. CONCLUSION

Using the state expansion coefficients for $\mathrm{SU}_N \supset \mathrm{SO}_N \supset \mathrm{SO}_{N-1}$ basis vectors in terms of $\mathrm{SU}_N \supset \mathrm{SU}_{N-1} \supset \mathrm{SO}_{N-1}$ basis vectors and isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$, we obtain some isoscalar factors for $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$. The method outlined in this paper can also be extended to other group chains, for example, when the expansion coefficients for the basis vectors of the one group chain in terms of the other and the isoscalar factors for the one group chain are known, the isoscalar factors for the other group chain can be derived by using this method.

APPENDIX: SOME PROPERTIES OF ISOSCALAR FACTORS FOR $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ AND SOME SPECIAL ISOSCALAR FACTORS FOR $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ AND $\mathrm{SO}_N \supset \mathrm{SO}_{N-1}$

The isoscalar factors for $\mathrm{SU}_N \supset \mathrm{SU}_{N-1}$ can be expressed as[13]

$$\left[\begin{array}{c|cc|c} \mathrm{SU}_N & (w_1 w_2) & (w_3 0) & (w'_1 w'_2) \\ \mathrm{SU}_{N-1} & (n_1 n_2) & (n_3 0) & (n'_1 n'_2) \end{array} \right]$$
$$= \delta_{w_1 + w_2 + w_3\, w'_1 + w'_2} \delta_{n_1 + n_2 + n_3\, n'_1 + n'_2}$$
$$\times \left[\frac{(w_3 - n_3)!(n_1 - n_2 + 1)(w'_1 - w'_2 + 1)(n'_1 - n_1)!(n'_2 - n_2)!(n'_1 - n_2 + 1)!(w_1 - w'_2)!w_2!(w_1 + 1)!}{(n_1 - n'_2)!(w'_1 - w_1)!(w'_2 - w_2)!(w'_1 - w_2 + 1)!(w'_1 + 1)!w'_2!(n_1 - w_2)!n_2!(n_1 + 1)!} \right.$$

$$\times\frac{(w_1-n_1)!(w_2-n_2)!(w_1-n_2+1)!(n_1'-w_2')!n_2'!(n_1'+1)!}{(w_1'-n_1')!(w_2'-n_2')!(w_1'-n_2'+1)!}\Bigg]^{1/2}$$

$$\times\sum_{xy}(-)^{x+y}\frac{(w_1'-n_1-x)!(n_1-n_2'+x)!(n_1-w_2+x)!(n_1-n_2+x-y-1)}{x!y!(n_1-n_2+x+1)!(w_1-n_1-x)!(n_1'-n_1-x)!(n_1-w_2'+x)!}$$

$$\times\frac{(n_1-n_2-y)!(w_1'-n_2+1-y)!(w_2'-n_2-y)!}{(w_1-n_2+1-y)!(w_2-n_2-y)!(n_3+n_1-n_1'-y)!(n_1-n_2'+n_3+1-y)!}. \tag{A1}$$

Equation (A1) can be simplified when $w_2=0$ or n_2 and $n_2'=0$, and it can be expressed by the following SU_2 CG coefficients:

$$\left[\begin{array}{c|cc|c} SU_N & (w_1w_2) & (w_30) & (w_1'w_2') \\ SU_{N-1} & (n_10) & (n_30) & (n0) \end{array}\right]$$

$$=\delta_{w_1+w_2+w_3\,w_1'+w_2}\delta_{n_1+n_3n}\left[\frac{(w_3-w_2')!(w_1'-w_2'+1)(w_1-w_2')!}{(w_1'-w_3)!(w_1'-w_1)!(w_1+w_3-n+1)}\right]^{1/2}$$

$$\times(-)^{n_2}\left[\begin{array}{cc|c} \frac{1}{2}(n-w_2) & \frac{1}{2}(w_1'-w_2') & \frac{1}{2}(w_1+w_3-n) \\ \frac{1}{2}(n_1-n_3-w_2) & \frac{1}{2}(w_3-w_1+w_2) & \frac{1}{2}(n_1-n_3-w_1+w_3) \end{array}\right], \tag{A2}$$

and

$$\left[\begin{array}{c|cc|c} SU_N & w_1 & w_2 & (w_1'w_2') \\ SU_{N-1} & n_1 & n_2 & (n_1'n_2') \end{array}\right]$$

$$=\delta_{w_1+w_2\,w_1'+w_2'}\delta_{n_1+n_2\,n_1'+n_2'}$$

$$\times\left[\frac{(w_1'-w_2'+1)}{(w_1'+w_2'-n_1-n_2+1)}\right]^{1/2}(-)^{n_2-n_2'}\left[\begin{array}{cc|c} \frac{1}{2}(n_1'-n_2') & \frac{1}{2}(w_1'-w_2') & \frac{1}{2}(w_1'+w_2'-n_1'-n_2') \\ \frac{1}{2}(n_1-n_2) & \frac{1}{2}(w_2-w_1) & \frac{1}{2}(n_1-n_2-w_1+w_2) \end{array}\right]. \tag{A3}$$

From Eqs. (A2) and (A3), an important property of isoscalar factors for $SU_N \supset SU_{N-1}$ can be derived:

$$\left[\begin{array}{c|cc|c} SU_N & w_1 & w_2 & (w_1'w_2') \\ SU_{N-1} & n_1 & n_2 & (n_1'n_2') \end{array}\right]=\left[\begin{array}{c|cc|c} SU_N & w_1-n_2' & w_2-n_2' & (w_1'-n_2'\;w_2'-n_2') \\ SU_{N-1} & n_1-n_2' & n_2-n_2' & (n_1'-n_2'0) \end{array}\right]. \tag{A4}$$

Using Eqs. (A4) and (4.5), we obtain

$$B^{(w_1w_2)(v_1v_2)}_{(v_1v_2)(N)}=B^{(w_1-v_2w_2-v_2)(v_1-v_20)}_{(v_1-v_20)(N+2v_2)}. \tag{A5}$$

The following expressions for the special isoscalar factors are of importance:

$$\left[\begin{array}{c|cc|c} SU_N & w_1 & w_2 & w_1+w_2 \\ SU_{N-1} & n_1 & n_2 & n_1+n_2 \end{array}\right]=\left[\frac{w_1!w_2!(w_1+w_2-n_1-n_2)!(n_1+n_2)!}{(w_1-n_1)!n_1!(w_2-n_2)!n_2!(w_1+w_2)!}\right]^{1/2}, \tag{A6}$$

$$\left[\begin{array}{c|cc|c} SU_N & w_1 & w_2 & (w_1'w_2') \\ SU_{N-1} & n & 0 & n \end{array}\right]=\left[\frac{n!w_2!(w_1'-n)!(w_1-w_2')!(w_1'-w_2'+1)}{w_2'!(n-w_2')!(w_1'+1)!(w_1-n)!(w_2-w_2')!}\right]^{1/2}, \tag{A7}$$

$$\left[\begin{array}{c|cc|c} SO_N & w_1 & w_2 & (w_1'w_2') \\ SO_{N-1} & n & 0 & n \end{array}\right]$$

$$=\Bigg[\frac{(w_1'+N-4)!(w_2'+n+N-4)!(w_1'+n+N-3)!(w_2'+N-5)!}{(w_1'+w_2'+N-4)!(2w_1'+N-4)!!(2w_2'+N-6)!!(n-w_2')!}$$

$$\times\frac{(2w_1+N-4)!!(2w_2+N-4)!!(N-3)n!w_2!(w_1'-n)!(w_1-w_2')!(w_1'-w_2'+1)}{(n+N-4)!(w_1+n+N-3)!(w_2+N-3)!w_2'!(w_1'+1)!(w_1-n)!(w_2-w_2')!}\Bigg]^{1/2},$$

$$\text{for } N>4 \text{ and } w_1+w_2=w_1'+w_2'. \tag{A8}$$

[1] P. Van Isacker, K. Heyde, J. Jolie, and A. Sevrin, Ann. Phys. (NY) **171**, 253 (1986).

[2] L. C. Biedenharn, J. Math. Phys. **2**, 433 (1961).

[3] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, NJ, 1957), and references therein.

[4] A. M. Gavrilik, Kiev preprint ITP-73-104R, 1973.

[5] M. S. Kildyushov and G. I. Kuznetzov, Institute of Atomic Energy preprint, 1973.

[6] S. J. Alisauskas and A. P. Jucys, J. Math. Phys. **12**, 594 (1971).

[7] S. J. Alisauskas and A. P. Jucys, J. Math. Phys. **10**, 2227 (1969).

[8] E. Z. Norvaisas and S. J. Alisauskas, Liet. Fiz. Rinkinys **14**, 443 (1974).

[9] S. J. Alisauskas, Part. Nucl. **14**, 1336 (1983).

[10] S. J. Alisauskas, J. Phys. A: Math. Gen. **20**, 35 (1987).

[11] J. A. Castilho Alicaras and V. Vanagas, J. Math. Phys. **28**, 1995 (1987).

[12] A. Arima and F. Iachello, Ann. Phys. (NY) **123**, 468 (1979).

[13] E. Chacon, M. Ciftan, and L. C. Biedenharn, J. Math. Phys. **13**, 577 (1972).

[14] B. G. Wybourne, *Classical Groups for Physicists* (Wiley, New York, 1974).

[15] S. J. Alisauskas, J. Phys. A: Math. Gen. **17**, 2927 (1984).

Discrete supergroups and super Riemann surfaces

Paolo Teofilatto
Department of Mathematics, King's College London, Strand, London WC2R 2LS, England

(Received 7 July 1987; accepted for publication 15 June 1988)

The properties of discrete supergroups that represent the homotopy groups of super Riemann surfaces are emphasized. After the study of the superforms invariant with respect to these supergroups, the set of superconformal differentials are defined and an example is given of those by defining Poincaré's super theta functions. Some remarks on a possible definition of the Jacobi variety of a super Riemann surface are given.

I. INTRODUCTION

In the covariant quantization of bosonic string theory the multiloop amplitudes are expressed in terms of integrals over the spaces of genus g conformally inequivalent Riemann surfaces.[1] These spaces, called moduli spaces, have an extensive mathematical literature based on the early works of Riemann, Fricke, and Teichmuller.

The analysis of the dependence of the string amplitude on the moduli coordinates leads us to recognize that the bosonic string amplitude is divergent. To avoid this problem, a supersymmetrized version of the theory (superstring theory) has been proposed.[2]

A geometrical formulation of superstring theory requires new structures with an anticommuting coordinate, which simulate the properties of ordinary Riemann surfaces (RS's). These new objects are called super Riemann surfaces (SRS's). The supersymmetric string amplitude depends again on the space of superconformally inequivalent SRS's called supermoduli space. Several interesting results have been achieved on the analysis of supermoduli.[3–6]

In this paper we go over some of the fundamentals of the mathematical theory of super Riemann surfaces. In particular, the analysis of the properties of the transformation supergroups involved in the theory enables us to give a fairly detailed description of the representations of the homotopy groups of SRS's, and to define the category of marked SRS's. Once the notion of marking is clarified, the uniformization procedures for SRS's are straightforward.

In particular, via the Schottky uniformization, we can define the space of superconformal differentials. That leads to a possible definition of the Jacobi variety of a SRS that could be a useful tool in the supermoduli analysis.

This paper is organized as follows: notations and reviews of some of the mathematical apparatus needed later on are given in Sec. II. The topological properties of the supergroups of transformation are analyzed and the meaning of the discrete supergroup is clarified in Sec. III. The definition of SRS is recalled and the uniformization process for compact SRS, defined in Ref. 3, is gone through in some detail in Sec. IV. As a result of that, we get an explicit realization of homotopy groups of compact SRS's. That allows us to define an alternative uniformization process by means of Schottky supergroups. In Sec. V we use the results of Sec. III to define automorphic superforms and, via the Schottky uniformization, we are able to show the space of superconformal differentials of a SRS. Finally, we give some remarks on a possible definition of the Jacobi variety of a SRS.

II. NOTATIONS AND DIFFERENTIAL CALCULUS ON $\mathbb{C}$

In this section we define the notation and recall some standard results on supermanifold theory. All the details can be found in Refs. 7 and 8.

In dealing with supersymmetry we need a space where commutative and anticommutative coordinates are treated on the same ground. A good model space is a Grassmann algebra with a certain number L of generators β_i, $\mathbb{K}_L = \mathbb{K}\{1,(\beta_i),(\beta_i \wedge \beta_J),\ldots,\beta_1 \wedge \cdots \wedge \beta_L\}$, where $\mathbb{K}$ is a field and $\wedge$ is the Grassmann product. Here $\mathbb{K}_L$ is a $\mathbf{Z}_2$-graded algebra where the even (odd) sector $\mathbb{K}_L^{1,0}$ ($\mathbb{K}_L^{0,1}$) is that generated by combinations of an even (odd) number of generators, so any point of $\mathbb{K}_L$ can be regarded as a couple (z,ϑ), that is, $\mathbb{K}_L = \mathbb{K}_L^{1,0} \oplus \mathbb{K}_L^{0,1}$.

The projection map to the zero degree element of the Grassmann algebra $\varepsilon: \mathbb{K}_L \to \mathbb{K}$ is called the body map and it is an algebra homomorphism; moreover, ker ε is the ideal $\overline{\mathbb{K}}_L$ of nilpotent elements (called souls) generated by the β_i's defining a filtration of $\mathbb{K}_L$ by ideals $\overline{\mathbb{K}}^{(n)}$. For each $z \in \mathbb{K}_L$ we will call the set $\varepsilon^{-1}\varepsilon(z)$ the ε fiber on z. Two different topologies can be defined on $\mathbb{K}_L$.

(i) The de Witt topology is the weakest topology such that ε is continuous, that is, $\widetilde{U}$ is de Witt open on $\mathbb{K}_L$ iff $\widetilde{U} = \varepsilon^{-1}(U)$ for some open set U on $\mathbb{K}$.

(ii) The Rogers topology is defined by the l_1 norm on the coefficients of the basis of $\mathbb{K}_L$.

That obviously extends to the product

$$\mathbb{K}_L^{m,n} = \underbrace{\mathbb{K}_L^{1,0} \times \cdots \times \mathbb{K}_L^{1,0}}_{m} \times \underbrace{\mathbb{K}_L^{0,1} \times \cdots \times \mathbb{K}_L^{0,1}}_{n}.$$

In what follows we will be mainly interested in $\mathbb{C}_L$ (i.e., $\mathbb{K} = \mathbb{C}$) and we choose the value L in the $L \to \infty$ limit, regarded as a direct limit of spaces (as in Ref. 8). We will write $\mathbb{C}_L = \mathbb{C}^{1,1}$. A function $\widetilde{\omega}$ on $\mathbb{C}^{1,1}$ to $\mathbb{C}^{1,1}$ is said to be superanalytic if it is polynomial on ϑ, that is,

$$\widetilde{\omega}(z,\vartheta) = \widetilde{\omega}_0(z) + \vartheta\widetilde{\omega}_1(z), \tag{2.1}$$

and each coefficient $\widetilde{\omega}_0$, $\widetilde{\omega}_1$, is uniquely defined by an analytic complex function, say ω_0, ω_1, via the expansion

$$\widetilde{\omega}_i(z) = \omega_i(\varepsilon z) + \omega_i'(\varepsilon z)(z - \varepsilon z) + \tfrac{1}{2}\omega_i''(\varepsilon z)(z - \varepsilon z)^2 + \cdots \quad (i = 0,1). \tag{2.2}$$

A (m,n) supermanifold is a topological manifold on $\mathbb{C}^{m,n}$ with superanalytic transition functions. Since two topologies are available on $\mathbb{C}^{m,n}$ we will have de Witt or Rogers supermanifolds. (See Fig. 1.) Any de Witt supermanifold M with atlas $\{(\widetilde{U}_i,\widetilde{\varphi}_i)\}$ admits a body M_0 which is the m-dimension-

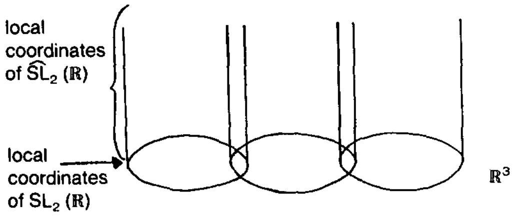


FIG. 1. Local coordinates of a de Witt supergroup.

al complex manifold with atlas $\{(\varepsilon\widetilde{\varphi}_i(\widetilde{U}_i), \varphi_i = \varepsilon\circ\widetilde{\varphi}_i)\}$ and if the transition functions of M are φ_{ij}, those of M_0 are the φ_{ij} in the sense of (2.2). Vice versa if an m-dimensional complex manifold M_0 is given, one can define an (m,n)-dimensional de Witt supermanifold as follows: the charts are the inverse image via ε of those of M_0, and the transition functions are $\widetilde{\varphi}_{ij}$ given by those φ_{ij} of M via (2.2).

The following equivalence relation on $\widetilde{M}_0 = \cup_i \widetilde{U}_i \subset \mathbb{C}^{m,0}$ is defined: $(z,\varepsilon(z)) \sim (z',\varepsilon(z'))$ iff $\varepsilon(z) = \varepsilon(z')$, $\widetilde{\varphi}_{iJ}(z) = z'$, $z\in\widetilde{U}_i$, $z'\in\widetilde{U}_J$. The quotient $M = \widetilde{M}_0/\sim \times \mathbb{C}^{0,n}$ gives a (m,n)-de Witt supermanifold with body M_0. Although this construction is a great source of examples of supermanifolds, it is important to note that not all the (complex) de Witt supermanifolds can be achieved by this way.[9] The relations between ordinary manifolds and supermanifolds is, in the Rogers context, far more intricate.[10] Finally, we recall the definition of integration on $\mathbb{C}^{1,1}$. First, an ordinary contour integral is defined for ρ^1 paths of $\mathbb{C}^{1,0}$: let γ: $[a,b] \to \widetilde{U} \subset \mathbb{C}^{1,0}$ be a ρ^1 function (path) and let f: $\widetilde{U} \to \mathbb{C}^{1,1}$ be continuous, then

$$\int_\gamma f\,dz = \int_a^b f(\gamma(t))\gamma'(t)dt.$$

The Cauchy theorem holds, namely, if f is superanalytic on and inside a closed path of $\mathbb{C}^{1,0}$, γ, then $\oint_\gamma f\,dz = 0$; moreover if f is superanalytic but with poles inside γ, then

$$\oint_\gamma f\,dz = 2\pi i \text{ (sum of residues at the poles).}$$

Note that by (2.2) the singularities of f are only dependent on the body values $\varepsilon(z)$. Then the singularities of f relative to the path γ of $\mathbb{C}^{1,0}$ are in fact relative to the projected path $\varepsilon\gamma$ of $\mathbb{C}$.[7] However, the residues (2.3) can have values in $\mathbb{C}^{1,1}$. The ordinary contour integral on $\mathbb{C}^{1,0}$ has to be patched with the Berezin integration rules on $\mathbb{C}^{0,1}$:

$$\int d\vartheta = 0, \quad \int d\vartheta\,\vartheta = 1.$$

For any superanalytic function $\widetilde{\omega}$, the result is

$$\oint_\gamma dz\,d\vartheta(\widetilde{\omega}_0 + \vartheta\widetilde{\omega}_1) = \oint_\gamma dz\,\widetilde{\omega}_1(z). \tag{2.4}$$

Note that the Berezin integral is defined without ambiguity because the body manifold is compact.

III. DISCRETE SUPERGROUP

A super-Lie group (supergroup) is a superanalytic (topological) supermanifold that is a group with superanalytic (continuous) operations. We recall the construction of a super-Lie group starting from a Lie algebra.[11] In fact, the underlying linear structure of a super-Lie group is not a Lie algebra but a graded Lie algebra, that is a $\mathbb{Z}_2$-graded complex algebra $\mathscr{G} = \mathscr{g}_0 \oplus \mathscr{g}_1$ with a product $\langle\cdot,\cdot\rangle$ which verifies a generalized Jacobi identity.

The Grassmann envelope of the graded Lie algebra $\mathscr{G} = \mathscr{g}_0 \oplus \mathscr{g}_1$ is the space $\widetilde{\mathscr{G}} = \mathbb{C}^{1,0} \otimes \mathscr{g}_0 \oplus \mathbb{C}^{0,1} \otimes \mathscr{g}_1$ that turns out to be a Lie algebra with respect to the Lie product $[xg,yh] = xy \otimes \langle g,h\rangle$. A topology can be defined on $\widetilde{\mathscr{G}}$ requiring the projection map φ: $\widetilde{\mathscr{G}} \to \mathbb{C}^{m,n}$ to be a homomorphism (here m,n are the complex dimension of the vector spaces $\mathscr{g}_0$, $\mathscr{g}_1$, respectively). An exponential map is defined on $\widetilde{\mathscr{G}}$ which locally defines a super-Lie group $\widetilde{G}$ with the product given by the Campbell–Hausdorff formula.

The charts of $\widetilde{G}$ are $\{(\widetilde{U},\varphi\circ\exp^{-1})\}$ and its topological structure is given by requiring the maps $\varphi\circ\exp^{-1}$ to be homomorphisms. Then the topology of $\widetilde{G}$ depends on the topology defined on $\mathbb{C}^{m,n}$: if there is the de Witt (Rogers) topology, $\widetilde{G}$ will be a de Witt (Rogers) super-Lie group. In both cases the transition functions and the group operations are superanalytic.

Moreover, by construction, $\widetilde{G}$ admits a body G, locally defined by the Lie algebra $\mathscr{g}_0$. Here G is an analytic Lie group with $\mathscr{g}_0$ as its Lie algebra.

The whole picture is given by the following diagram:

$$\begin{tikzcd}
\widetilde{G} \arrow[r, "\exp^{-1}"] \arrow[d, swap, "E"] & \widetilde{\mathscr{G}} \arrow[r, "\widetilde{\varphi}"] \arrow[d, swap, "\varepsilon"] & \mathbb{C}^{m,n} \arrow[d, "\varepsilon"] \\
G \arrow[r, swap, "\exp^{-1}"] & \mathscr{g}_0 \arrow[r, swap, "\varphi"] & \mathbb{C}^m
\end{tikzcd} \quad . \tag{3.1}$$

As an example, take the following graded Lie algebra with commutation-anticommutation rules:

$$\begin{aligned}
&[L_m,L_n] = (m-n)L_{(m+n)\text{mod }2},\\
&\mathscr{G} = \mathbb{C}\{L_{-1},L_0,L_1\} \oplus \mathbb{C}\{G_{-1/2},G_{1/2}\},\\
&[L_m,G_r] = (m/2-r)G_{(m+r)\text{mod }2},\\
&\{G_r,G_s\} = 2L_{r+s}.
\end{aligned}$$

This is a subalgebra of the Neveu–Schwartz algebra,[2] and its even sector $\mathscr{g}_0$ is the Lie algebra of $\mathrm{SL}_2(\mathbb{C})$. The super-Lie group obtained by exp of the Grassmann envelope $\widetilde{\mathscr{G}} = \mathbb{C}^{1,0}\{L_{-1},L_0,L_1\} \oplus \mathbb{C}^{0,1}\{G_{\pm 1/2}\}$ is $\widehat{\mathrm{S}\,\mathrm{L}}_2(\mathbb{C})$, which has local coordinates on open sets of $\mathbb{C}^{3,2}$, namely,

$$g(a,b,c,d,\gamma,\delta) = \begin{pmatrix} a & b & b_\gamma - a\delta \\ c & d & d_\gamma - c\delta \\ \gamma & \delta & 1 + \delta\gamma/2 \end{pmatrix}, \quad ad - bc = 1, \tag{3.2}$$

with g a complex (2/1) supermatrix.[12] We stress again that two possible topologies can be given to $\widetilde{G} = \widehat{\mathrm{S}\,\mathrm{L}}_2(\mathbb{C})$. In the case of de Witt topology, it is natural to define a discrete supergroup of $\widetilde{G}$ as the following.

Definition 3.1: A subgroup $\widetilde{\Gamma}$ of a de Witt super-(Lie) group $\widetilde{G}$ is discrete iff the image of $\widetilde{\Gamma}$ via the map E of (3.1), which we will call $\varepsilon\widehat{\Gamma}$, is a discrete subgroup of the body of $\widetilde{G}$, G.

It follows that the points of $\widetilde{\Gamma}$ have as local coordinates the ε inverse of the coordinates of isolated points of the Lie group G (see Fig. 2).

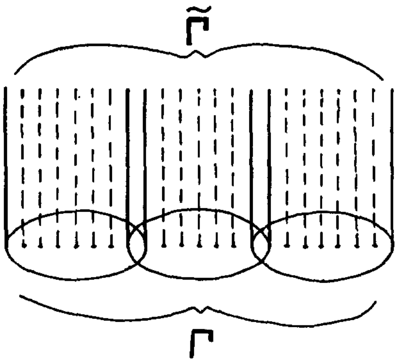


FIG. 2. Local coordinates of a de Witt discrete supergroup.

We have by definition the following proposition.

Proposition 3.1: Any discrete supergroup $\tilde{\Gamma}$ of $\tilde{G}$ has a body which is discrete with respect to G.

Now we want to know what are the conditions to get a de Witt supermanifold as a quotient of a de Witt supermanifold M with respect to the discontinuous action of a discrete supergroup acting on it.

We recall that a de Witt supermanifold M is a bundle over its body M_0, and the local representative of the bundle projection map is ε. Therefore the action of a topological transformation group of M has to preserve the ε fibers, that is,

$$\varepsilon(g\cdot p) = \varepsilon(g)\cdot\varepsilon(p). \tag{3.3}$$

Hence the Rogers topology is not in general an admissible topology for a transformation group $\tilde{G}$ of a de Witt supermanifold M: $\tilde{G}$ has to be a de Witt supergroup. Now we have the following proposition.

Proposition 3.2: Let M be a de Witt supermanifold, $\tilde{G}$ be a super-Lie group of transformations of M, and $\tilde{\Gamma}$ a discrete subgroup of $\tilde{G}$ with discontinuous and fixed point-free action on M, then $M/\tilde{\Gamma}$ is a de Witt supermanifold.

Proposition 3.2 (whose proof is standard[13]) is not completely satisfactory because we do not know how to check the conditions on the action of $\tilde{\Gamma}$ on M, mainly the discontinuity, and we need a more concrete condition. We note that if $\tilde{\Gamma}$ is discontinuous on M, then its body $\varepsilon\tilde{\Gamma}$ is discontinuous on the body of M, M_0. The converse is almost true, if $\epsilon\tilde{\Gamma}$ is discontinuous on M_0, the set $\{\epsilon\gamma\colon \epsilon\gamma(\epsilon p) = \epsilon p\}$ is finite for each point ϵp on M_0, and if the action of $\epsilon\tilde{\Gamma}$ is fixed point-free, the previous set is the identity. Because of (3.3) an element of $\tilde{\Gamma}$ cannot move along the ε fiber on εp, so the action of $\tilde{\Gamma}$ on M would be discontinuous and fixed point-free. But elements γ of $\tilde{\Gamma}$ such that $\varepsilon\gamma = 1$ are admissible topological transformations of a de Witt supermanifold and they act only along the ε fibers. However in the following sections we will be interested in discrete supergroups $\tilde{\Gamma}$, which, for geometrical reasons, are isomorphic to their bodies $\varepsilon\tilde{\Gamma}$. Equivalently the $\tilde{\Gamma}$'s will be isomorphic to their quotients with respect to the equivalence relation induced on $\tilde{\Gamma}$ by ε, that is, $\gamma_1 \sim \gamma_2 \Leftrightarrow \varepsilon\gamma_1 = \varepsilon\gamma_2$. With this condition the above ambiguity disappears (in particular the unity $1_{\tilde{\Gamma}}$ of $\tilde{\Gamma}$ corresponds to the unity 1_{Γ} of Γ), then one has the following proposition.

Proposition 3.3: Let M be a de Witt supermanifold and $\tilde{\Gamma}$ a de Witt discrete supergroup such that $\tilde{\Gamma} \cong \varepsilon\tilde{\Gamma}$. If $\varepsilon\tilde{\Gamma}$ is discontinuous and fixed point-free on M, then $M/\tilde{\Gamma}$ is a de Witt supermanifold.

Now if $\mathscr{H}$ is the complex upper half-plane, let $\tilde{\mathscr{H}} = \varepsilon^{-1}(\mathscr{H})$ be the complex super upper half-plane. A subgroup $\tilde{\Gamma}$ of $\hat{S}L_2(\mathbb{R})$ acts on $\tilde{\mathscr{H}}$ by the usual Möbius action,

$$\begin{pmatrix} a & b & b\gamma - a\delta \\ c & d & d\gamma - c\delta \\ \gamma & \delta & 1 + \delta\gamma/2 \end{pmatrix} \begin{pmatrix} z \\ \vartheta \end{pmatrix} = \left(\frac{az + b + \vartheta(a\delta - b\gamma)}{cz + d + \vartheta(c\delta - d\gamma)}, \frac{\gamma z + \delta + \vartheta(1 + \delta\gamma/2)}{cz + d + \vartheta(c\delta - d\gamma)} \right). \tag{3.4}$$

The classical theory tells us that the action of any discrete subgroup of $\hat{S}L_2(\mathbb{R})$ on $\mathscr{H}$ is discontinuous,[13] so in this case Proposition 3.3 reads as follows.

Proposition 3.4: Given $\tilde{\mathscr{H}}$ and a discrete subgroup $\tilde{\Gamma}$ of $\hat{S}L_2(\mathbb{C})$, if $\varepsilon\tilde{\Gamma}$ is fixed point-free on $\mathscr{H}$ (so $\mathscr{H}/\varepsilon\tilde{\Gamma}$ is a compact Riemann surface), then $\tilde{\mathscr{H}}/\tilde{\Gamma}$ is a de Witt supermanifold.

In Sec. IV we will see that $\tilde{\mathscr{H}}/\tilde{\Gamma}$ is in fact a super Riemann surface (SRS) and any SRS with a compact body is equivalent to $\tilde{\mathscr{H}}/\tilde{\Gamma}$ via some $\tilde{\Gamma} < \hat{S}L_2(\mathbb{R})$.

IV. SUPER RIEMANN SURFACES AND SUPERMODULI

Following Ref. 3, super Riemann surfaces are superanalytic supermanifolds M on $\mathbb{C}^{1,1}$, such that their tangent bundles TM are spanned by the basis $D = \partial/\partial\vartheta + \vartheta(\partial/\partial z)$, $D^2 = \partial/\partial z$; this restricts the transition functions of M to be of the form

$$\tilde{z} = \tilde{f} + \vartheta\psi\sqrt{\tilde{f}}, \quad \tilde{\vartheta} = \psi + \vartheta\sqrt{\tilde{f}' + \psi\psi'}, \tag{4.1}$$

which also implies that D transforms homogeneously,

$$D = (D\tilde{\vartheta})\tilde{D}, \tag{4.2}$$

and the "metric" $dz + \vartheta\, d\vartheta$ transforms homogeneously as well. Here M is also required to be de Witt type; this will be assumed from now on. The body of M, M_0, is a Riemann surface with transition functions induced by $\tilde{f}$ via (2.1). A spin structure is induced on M_0 by the consistent way to choose the square root implicit in (4.1). On the other hand, given a Riemann surface M_0 with a spin structure, one can define, by the methods of Sec. II, a SRS M with transition functions:

$$\tilde{z} = \tilde{f}(z), \quad \tilde{\vartheta} = \vartheta\sqrt{\tilde{f}'(z)}, \tag{4.3}$$

where $\tilde{f}$ is the expansion (2.2) of the transition functions of M_0, f. Such SRS's are called split.

We recall the definition of three split SRS's $\mathbb{C}^{1,1}$, $\tilde{\mathscr{H}}$, $\mathbb{C}P^{1,1}$ whose importance relies on the result of Ref. 3.

Proposition 4.1: The following three split SRS's are the unique ones having simply connected RS's as bodies.

(a) $\mathbb{C}^{1,1}$ and $\tilde{\mathscr{H}}$: Their bodies are the complex plane and the upper half-plane, respectively, which have unique spin structure and transition functions $f(z) = z$. Therefore the superconformal transition functions on $\mathbb{C}^{1,1}$, $\tilde{\mathscr{H}}$ are $(\tilde{z},\tilde{\vartheta}) = (z,\vartheta)$.

(b) $\mathbb{C}P^{1,1}$: Its superconformal structure is induced by the unique spin structure on the Riemann sphere $\mathbb{C}P^1$ and the superconformal transition functions are $(\tilde{z},\tilde{\vartheta}) = (1/z, \vartheta/z)$.

We can properly call $\mathbb{C}P^{1,1}$ the super Riemann sphere because a superanalog of the stereographic projection exists. Let us introduce it briefly: the supersphere $S^{2,2}$ is given by the equation on $\mathbb{R}^{3,2}$

$$x_1^2 + x_2^2 + x_3^2 + 2\vartheta_1\vartheta_2 = 1. \tag{4.4}$$

By the implicit function theorem,[14] it turns out that the space of solutions of this equation is a (2,2)-dimensional de Witt supermanifold with the body the ordinary sphere S^2.

We define a projection of $S^{2,2}$ to $\mathbb{C}P^{1,1}$: let N be the north pole of S^2, we call $\mathscr{F}(N)$ the ε fiber of N, and consider the map s: $S^{2,2} - \mathscr{F}(N) \rightarrow \mathbb{R}^{2,2}$ (noninvertible elements), defined by

$$S(u_1,u_2,u_3,\vartheta_1,\vartheta_2) = \left(\frac{u_1}{1-u_3}, \frac{u_2}{1-u_3}, \frac{\vartheta_1}{1-u_3}, \frac{\vartheta_2}{1-u_3}\right).$$

This map has its values on $\mathbb{R}^{2,2}$ (noninvertible elements), because we have taken off $\mathscr{F}(N)$, then $\varepsilon(u_3) \neq 1$ and $\varepsilon(u_1)$ or $\varepsilon(u_2)$ are different from zero.

The inverse map is

$$s^{-1}(x_1,x_2,\vartheta_1,\vartheta_2) = \left(\frac{2x_1}{1+x_1^2+x_2^2+2\vartheta_1\vartheta_2}, \frac{2x_2}{1+x_1^2+x_2^2+2\vartheta_1\vartheta_2}, \frac{x_1^2+x_2^2+2\vartheta_1\vartheta_2-1}{1+x_1^2+x_2^2+2\vartheta_1\vartheta_2}, \frac{2\vartheta_1}{1+x_1^2+x_2^2+2\vartheta_1\vartheta_2}, \frac{2\vartheta_2}{1+x_1^2+x_2^2+2\vartheta_1\vartheta_2}\right).$$

Once the identification $\mathbb{R}^{2,2} \equiv \mathbb{C}^{1,1}$ has been done, the map s gives a correspondence s: $S^{2,2} \rightarrow \mathbb{C}P^{1,1}$, where the ε fiber $\mathscr{F}(N)$, which is isomorphic to $\mathbb{C}^{1,1} - \mathbb{C}$, and points of $\mathbb{C}P^{1,1}$ with coordinates $[z_0, z_1, \vartheta]$, where z_0 is a noninvertible element of $\mathbb{C}^{1,0}$ (so z_1 has to be invertible), are left out. These points are in one to one correspondence with $\mathbb{C}^{1,1} - \mathbb{C}$ as easily seen by using affine coordinates $(z_0/z_1, \vartheta/z_1)$.

Then the identification "points at infinity" $\mathscr{F}(N)$ given by $(z_0/z_1, \vartheta/z_1) \mapsto (\bar{x}_1, \bar{x}_2, \vartheta_1, \vartheta_2) \mapsto (\bar{x}_1, \bar{x}_2, 1, \vartheta_1, \vartheta_2)$ completes the stereographic projection. The group of rotations of $S^{2,2}$, that is, the group that leaves invariant the quadratic form (4.4), is given in Ref. 15. If we project this group on $\mathbb{C}^{1,1}$ via s, that is,

$$(z,\vartheta) \overset{s^{-1}}{\longmapsto} (x_1,x_2,x_3,\vartheta_1,\vartheta_2) \overset{\text{Rot}}{\longmapsto} (x_1',x_2',x_3',\vartheta_1',\vartheta_2') \overset{s}{\longmapsto} (z',\vartheta'),$$

and if we allow the coefficients of the group to be complex, we get the group $\widehat{SL}_2(\mathbb{C})$. This group is just the group of superconformal automorphisms of $\mathbb{C}P^{1,1}$. The action of $\widehat{SL}_2(\mathbb{C})$ on $\mathbb{C}P^{1,1}$ in affine coordinates or equivalently the action of $\widehat{SL}_2(\mathbb{C})$ on $\mathbb{C}^{1,1}$ as the fractional linear transformation group is given by (3.4). The genus of a SRS is that of its body; Proposition 4.1 tells us that the genus zero SRS are only $\mathbb{C}^{1,1}$, $\mathscr{H}$, $\mathbb{C}P^{1,1}$. What happens to the higher genus? It is known that any RS M_0 can be regarded as a quotient of its covering space N_0 with respect to the group Γ of its covering transformations. Here Γ is isomorphic to the homotopy group of M_0, $\pi_1(M_0)$, and it is a discontinuous, discrete, and (if M_0 is compact) fixed point-free subgroup of the group of conformal automorphisms of N_0, $\text{Aut}(N_0)$. Here N_0 can be one of the genus zero RS $N_0 = \mathbb{C}, \mathscr{H}, \mathbb{C}P^1$: the analysis of the group $\text{Aut}(N_0)$ implies that $\pi_1(M_0)$ can be realized as the discontinuous subgroup of the Euclidean or hyperbolic isometries if $N_0 = \mathbb{C}$ or $\mathscr{H}$, respectively, and it is the identity if $N_0 = \mathbb{C}P^1$. This topological argument has to hold also in the context of SRS's. Note that, since SRS's are bundles over their bodies with contractible fiber, then the homotopy group of M is isomorphic to the homotopy group of its body M_0.

Proposition 4.1 implies that the covering space of any SRS can be $N = \mathbb{C}^{1,1}, \mathscr{H}, \mathbb{C}P^{1,1}$, then M is the quotient of N with respect to the group of its covering transformations $\tilde{\Gamma}$. Here $\tilde{\Gamma}$ is a subgroup of the group $\text{Aut}(N)$ of superconformal automorphisms of N, and it is isomorphic to the homotopy group of M, and hence to the homotopy group of the body of M, M_0.

As in the classical theory, we have to analyze separately the three cases $N = \mathbb{C}^{1,1}$, $\mathscr{H}$, $\mathbb{C}P^{1,1}$: the latter is again the simplest. Namely, $\text{Aut}(\mathbb{C}P^{1,1})$ is just $\widehat{SL}_2(\mathbb{C})$ whose transformations have one fixed point on $\mathbb{C}P^{1,1}$ at least. It follows that the only admissible $\tilde{\Gamma}$ is the identity, so any M with $\mathbb{C}P^{1,1}$ as the covering space is homeomorphic to $\mathbb{C}P^{1,1}$. As noted in Ref. 3, for the covering spaces $N = \mathbb{C}^{1,1}$, $\mathscr{H}$, the situation is more intricate. This is because the group of superconformal automorphisms of these spaces is not a subgroup of $\widehat{SL}_2(\mathbb{C})$, and one can find discontinuous, discrete, fixed point-free subgroups of $\text{Aut}(N)$ isomorphic to their bodies which are not subgroups of $\widehat{SL}_2(\mathbb{C})$.

But we know that the homotopy group of a SRS M with covering space $\mathbb{C}^{1,1}$ is isomorphic to the two-dimensional cyclic Abelian homotopy group of its body which is a complex torus T. The homotopy group of T has an explicit realization in the group of the Euclidean isometries of $\mathbb{C}$. The generalization of the Euclidean metric on $\mathbb{C}^{1,1}$ is the metric $|dz + \vartheta\, d\vartheta|$,[2] and the aim is to represent the homotopy group of M as a group of isometries of $(\mathbb{C}^{1,1}, |dz + \vartheta\, d\vartheta|)$.[3]

Therefore we restrict ourselves to consider $\mathbb{C}$ (and later $\mathscr{H}$) with its superprojective structure rather than its superconformal structure. Roughly, this is what happens in the uniformization procedures for complex manifolds of higher dimensions.[16] As it happens $\mathbb{C}^{1,1}$ has several similarities with $\mathbb{C}^2$: two (super) complex coordinates, (super) analytic transformations for each coordinate, a notion of (super) conformality which looks like the notion of (quasi) conformality of $\mathbb{C}^2$: the maps that transform homogeneously the metric $dz + \vartheta\, d\vartheta$.

An explicit description of the representation

$$\rho: \pi_1(M) \rightarrow \{\text{super-Euclidean isometries of } \mathbb{C}^{1,1}\}$$

has been given in Ref. 3 (M is a genus 1 SRS).

An analogous direct realization of the group $\pi_1(M)$ is possible also in the higher genus. (See Fig. 3.) Namely, a compact (i.e., with compact body) SRS of genus $g > 1$ can be realized as quotient $\mathscr{H}/\tilde{\Gamma}$, with $\tilde{\Gamma}$ the discrete subgroup of $\widehat{SL}_2(\mathbb{R})$. We can think of $\tilde{\Gamma}$ as a subgroup of the "superhyperbolic isometries" of $\mathscr{H}$ (that is, transformations which

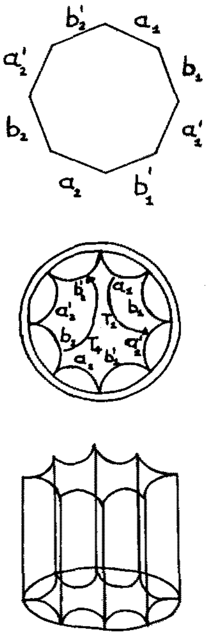


FIG. 3. (a) Fundamental polygon of a genus two Riemann surface. (b) Fundamental domain of the representation group Γ of $\pi_1(M)$. (c) Fundamental domain of the representation group $\widetilde{\Gamma}$ of $\pi_1(\widetilde{M})$.

preserve the metric $|dz + \vartheta\, d\vartheta\,|/|\mathrm{Im}\, z + \vartheta\bar{\vartheta}/2|$) once a marking on M has been chosen. A marking on M means a choice of a map ρ such that in

$$\pi_1(M_0) \xrightarrow{\rho} \hat{S}L_2(\mathbb{R}) \xrightarrow{\varepsilon} SL_2(\mathbb{R}) \tag{4.5}$$

the map $\varepsilon \circ \rho$ is a monomorphism.

Then the generators $\widetilde{T}_i$ of the group $\widetilde{\Gamma} = \rho(\pi_1(M_0))$ correspond, by ε, to the $2g$ generators T_i of $\Gamma = \varepsilon \cdot \rho(\pi_1(M_0))$ and if D_0 is the fundamental domain of the group Γ (the $2g$ hyperbolic polygon), then ε^{-1} (D_0) is the fundamental domain of $\widetilde{\Gamma}$.

In order to get the representation

$$\pi_1(M) \to \{\text{superhyperbolic isom.}\}$$

we still have to impose the commutative condition on the generators

$$[\widetilde{T}_{2g},\widetilde{T}_g] \cdots [\widetilde{T}_{g+1},\widetilde{T}_1] + 1.$$

Because each of the $\widetilde{T}_1$ is defined by a point of $\mathbb{R}^{3,2}$ then the group

$$\rho(\pi_1(M)) = \widetilde{\Gamma} = \langle \widetilde{T}_1,\ldots,\widetilde{T}_{2g.}\rangle$$

depends on $(6g,4g)$ parameters minus the commutative condition which imposes three even and two odd conditions: then $\widetilde{\Gamma}$ depends in fact on $(6g-3, 4g-2)$ parameters. Note that if one operates a conjugation by an element T of $\hat{S}L_2(\mathbb{R})$ on all the generators of the group $\widetilde{\Gamma}$, that is, $\widetilde{S} = T\widetilde{T}_i T^{-1}$ for each $i = 1,\ldots,2g$, the resultant group $\widetilde{S} = \langle \widetilde{S}_1,\ldots,\widetilde{S}_{2g} \rangle$ gives the same SRS M, up to a superconformal diffeomorphism.

Therefore to compute the number of SRS's up to superconformal diffeomorphisms (supermoduli), we have to subtract this freedom by performing an overall conjugation. That implies three even and two odd conditions on the parameters of the basis $\widetilde{T}_i$.

By varying with continuity the set of parameters of $\widetilde{\Gamma}$, one describes a $(6g-6, 4g-4)$ real de Witt supermanifold which is the super analog of the Fricke space.[17] In fact, the ambiguity of the overall conjugation implies that there exist two different spaces of supermoduli related one to the other by the inversion I (see Ref. 3); that is, the Fricke supermanifold has a double covering. This formulation of supermoduli parameters as (real) coordinates of the Fricke supermanifold has been given in Ref. 3.

An interesting way to recognize the complex structure of the covering space of supermoduli space is given in Ref. 5. There the Fricke supermanifold is shown to be related to the set Hom(Γ,Scf) of the homomorphisms from the group $\Gamma \cong \pi_1(M)$ to the group of superconformal automorphisms of $\mathscr{H}$, Scf.

The marking is $\Gamma \xrightarrow{\rho} \mathrm{Scf} \xrightarrow{\varepsilon} SL(2,\mathbb{R})$ with $\varepsilon \circ \rho$ injective, and the Fricke supermanifold results isomorphic to the quotient of the group Hom(Γ,Scf) with respect to the conjugate action of Scf on the space of homorphisms, that is Hom(Γ,Scf)/Scf. This space is shown to be a complex supermanifold of dimension $(3g-3, 2g-2)$ by the following cohomological approach which we briefly describe (all the details are in Refs. 5 and 6).

All the homomorphisms $\Gamma \to \mathrm{Scf}$ with a fixed body are classified by the cohomology group $H^1(\Gamma,\mathrm{Scf}_1)$ defined by the (non-Abelian) action of Γ on the group $\mathrm{Scf}_1 = \ker \varepsilon$. One can actually compute $H^1(\Gamma,\mathrm{Scf}_1)$ by noting that Scf_1 has a natural filtration given by $\mathrm{Scf}_1^{(n)} = \{(f,\psi)$ as in (4.1) such that $(f,\sqrt{f'},\psi)$ is congruent to $(z,1,0) \bmod \bar{\mathbb{C}}^n\}$.

The cohomology of the quotient groups of the filtration $\mathrm{Scf}_1^{(n)}/\mathrm{Scf}_1^{(n+1)}$ is known because basically these groups are defined on the body, and one has

$$H^1(\Gamma,\mathrm{Scf}_1^{(n)}/\mathrm{Scf}_1^{(n+1)}) = \begin{cases} H^1(M_0,\sqrt{K^{-1}}) \otimes \bar{\mathbb{C}}^n/\bar{\mathbb{C}}^{n+1} = (2g-2) \underset{\mathbb{C}}{\otimes} \bar{\mathbb{C}}^n/\bar{\mathbb{C}}^{n+1} \ (n \text{ odd}) \\ H^1(M_0,K^{-1}) \otimes \bar{\mathbb{C}}^n/\bar{\mathbb{C}}^{n+1} = (3g-3) \underset{\mathbb{C}}{\otimes} \bar{\mathbb{C}}^n/\bar{\mathbb{C}}^{n+1} \ (n \text{ odd}) \end{cases}$$

(K is the cotangent bundle of the body M_0). Then we have the desired result by passing to the direct limit $\mathrm{Scf}_1 = \lim_n \mathrm{Scf}_1/\mathrm{Scf}_1^{(n)}$.

V. AUTOMORPHIC SUPERFORMS AND SUPERCONFORMAL DIFFERENTIALS

In this section we will define the space of automorphic superforms of weight h and in particular the space $\widetilde{\Omega}$ of automorphic superforms of weight 1, that is, superconformal differentials. Moreover, via the Schottky uniformization process, we can construct explicitly g linearly independent superconformal differentials that give a basis of $\widetilde{\Omega}$ for those SRS's which induce an even spin structure on the body.

We recall that, for SRS's, the analog of the canonical bundle is the (super) line bundle $\widetilde{K}$ defined by the cocycle

$$g_{\alpha\beta} = \text{sdet}\begin{pmatrix} \dfrac{\partial z_\alpha}{\partial z_\beta} & \dfrac{\partial z_\alpha}{\partial \vartheta_\beta} \\ \dfrac{\partial \vartheta_\alpha}{\partial z_\beta} & \dfrac{\partial \vartheta_\alpha}{\partial \vartheta_\beta} \end{pmatrix}$$

$$= \frac{\partial z_\alpha/\partial z_\beta}{\partial\vartheta_\alpha/\partial\vartheta_\beta - (\partial\vartheta_\alpha/\partial z_\beta)(\partial z_\alpha/\partial z_\beta)^{-1}(\partial z_\alpha/\partial\vartheta_\beta)}, \tag{5.1}$$

where $z_\beta, \vartheta_\beta) \to (z_\alpha, \vartheta_\alpha)$ are the transition functions of M in $U_\alpha \cap U_\beta$ (see, for instance, Ref. 18). In fact, $\widetilde{K}$ is the (super) line bundle defined by the class $[g_{\alpha\beta}]$ in the cohomology group $H^1(M_0, O_S^*)$, defined on the body with values on the sheaf of invertible superanalytic functions on M. The global sections of $\widetilde{K}$ are globally defined (1,1) superforms on $M, \tilde{\omega}\, dz\, d\vartheta$. Moreover, by uniformization, these sections can be regarded as superanalytic functions $\tilde{\omega}$ defined on $\widetilde{\mathscr{H}}$ and such that

$$\tilde{\omega}(\tilde{\gamma}(z,\vartheta)) = \tilde{\omega}(\tilde{z},\tilde{\vartheta})$$

$$= \frac{(cz+d)^2}{(1+\delta\gamma/2)(cz+d) + \vartheta(\gamma\delta - \delta c)}\tilde{\omega}(z,\vartheta) \tag{5.2}$$

for any $\tilde{\gamma}$ of the group $\widetilde{\Gamma}$, where $M = \widetilde{\mathscr{H}}/\widetilde{\Gamma}$. We call such an $\tilde{\omega}$ an automorphic superform of weight $m = 1$. An automorphic superform of weight $m \in \mathbb{Z}$ will be a section of $\widetilde{K}^m$. Formula (5.2) implies transformation properties on the components of $\tilde{\omega}$ in the expansion on ϑ: $\tilde{\omega} = \tilde{\omega}_0 + \vartheta\tilde{\omega}_1$. Namely, by expanding both sides of (5.2) in powers of ϑ and comparing the coefficients with the same degree, we have, for split SRS,

$$\tilde{\omega}_0(z) = (cz+d)^{-1}\tilde{\omega}_0(\tilde{z}),$$
$$\tilde{\omega}_1(z) = (cz+d)^{-2}\tilde{\omega}_1(\tilde{z}), \quad \tilde{z} = (az+b)/(cz+d), \tag{5.3}$$

where, for nonsplit SRS,

$$\tilde{\omega}_1(z)$$
$$= (cz+d)^{-2}[\tilde{\omega}_1(\tilde{z})(1+3\delta\gamma) + \tilde{\omega}_0(\tilde{z})(\delta c - \gamma\delta)], \tag{5.4}$$

with $\tilde{z}$ as in (3.4).

Transformations (5.3) mean that the weights of $\tilde{\omega}_0, \tilde{\omega}_1$ are $\frac{1}{2}$, 1, respectively, and we can think of $\tilde{\omega}$ as odd valued and with conformal weight $\frac{1}{2}$ (note that ϑ has conformal weight $-\frac{1}{2}$). The set $\widetilde{\Omega}$ of superconformal differentials splits into $\widetilde{\Omega}_0 + \widetilde{\Omega}_1$, where $\widetilde{\Omega}_0$ is in a one to one correspondence with the space of $\frac{1}{2}$-differentials on the body, that is, the space $H^0(M_0, O(\sqrt{K}))$, and $\widetilde{\Omega}_1$ is one to one with the one-differentials on M_0, that is, $\widetilde{\Omega}_1 \cong H^0(M_0, O(K))$.

So it is easy to know how many superconformal differentials there are on a split SRS: the dimension of $\widetilde{\Omega}_1$ is g (g = genus of M), thus there are g linearly independent odd superconformal differentials; the dimension of $\widetilde{\Omega}_0$ depends on the existence of harmonic spinors on the body. If the spin structure induced on the body is even, in general there are no harmonic spinors, so the dimension of $\widetilde{\Omega}_0$ is zero. If the spin structure is odd, $\widetilde{\Omega}_1$ has dimension 1 at least.

A more difficult question to answer is how many superconformal differentials there are on a nonsplit SRS. If the induced spin structure has no harmonic spinors, then Eq. (5.4) has solution of the type (5.3), certainly at first order with respect to the filtration of $\mathbb{C}$ and probably to all orders. In this case the number of superconformal differentials is the same as the split SRS's.[19] In the case of the existence of harmonic spinors, the space $\widetilde{\Omega}$ is not freely generated and it does not split into $\widetilde{\Omega}_0 + \widetilde{\Omega}_1$.[20]

We can give an explicit construction of the g superconformal differentials thus giving a basis of $\widetilde{\Omega}$ for the case of the spin structure with no harmonic spinors. In fact, we generalize the classical procedure where holomorphic differentials are generated by Poincaré theta functions. Poincaré theta functions are Poincaré series[21]

$$\sum_{\gamma\in\Gamma} \gamma'(z) f(\gamma(z)),$$

where $f(z)$ is the function $f(z) = 1/(z-a)$, which has poles at $z = a, \infty$. Therefore the series will have poles at the analog via Γ of these z values.

We call the Poincaré series with the previous choice of f, $\xi(z,a)$, and if T_i are the generators of the group Γ, it turns out that the g functions $\xi_i = \xi(z,a) - \xi(z,T_i a)$ are automorphic of weight 1, independent by the choice of a, without poles in the fundamental domain of Γ and linearly independent, and therefore they define a basis of the space of holomorphic differentials, provided that the series ξ is convergent.

The convergence of these series can be ensured if the group Γ is the Fuchsian group obtained by the Schottky uniformization of the Reimann surface. We spend few words on this uniformization which easily extends to SRS's. Roughly (all the details can be found in Ref. 22), the marking of a RS M_0 allows one to perform a dissection of M_0, mapped on a multiconnected planar region H which arises from cutting M_0 along the cycles $b_1,...,b_g$. To any b_i correspond two mutually disjoint "circles" (B_i, B'_i) on H. The $a_1,...a_g$ cycles correspond to the $T_1,...,T_g$ transformations of $\text{PL}(2,\mathbb{R})$ that map the circles B'_i onto B_i. The group generated by $T_1,...,T_g$ is called the Schottky group and has H as its fundamental region. As T ranges over $\Gamma = \langle T_1,...,T_g \rangle$, the regions $T(H)$ fill out the covering surface of M_0, N_0, without overlapping. Each of the T_i's is uniquely defined by its two complex fixed points and its multiplier k_i, that is, three complex parameters. Since N_0 is defined modulo a transformation of the complex plane, it is possible to exploit this freedom by fixing 3 of the $3g$ complex parameters of the elements of the basis, which will now depend on $3g-3$ complex parameters. For SRS's the choice of the homology basis $a_1,...,b_g$ defines g transformations of $\widehat{S L}_2(\mathbb{C})$ which generate a group $\widetilde{\Gamma} = \langle \widetilde{T}_1,...\widetilde{T}_g \rangle$ isomorphic via ε to the Schottky group of the body Γ. Any generator $\widetilde{T}_i$ is defined by a point of $\mathbb{C}^{3,2}$, subtracting an overall transformation of $\mathbb{C}^{1,1}$, that is, fixing the fundamental domain of $\widetilde{\Gamma}$, we get the $(3g-3, 2g-2)$ complex supermoduli. From the arguments of Sec. III, $\widetilde{\Gamma}$ has $\varepsilon^{-1}(H)$ as its fundamental domain. Note also that the choice of the $b_1,...,b_g$ cycles on M has to vary with the different induced spin structures. Now we define the Poincaré super theta function:

$$\Theta(z,\vartheta;a,\alpha) = \sum_{i\in\Gamma} \frac{D\vartheta_i}{z_i - a - \vartheta_i\alpha}\vartheta_i. \tag{5.5}$$

This function depends on a point (a,α) of $\mathbb{C}^{1,1}$ and it is possible to show that it is convergent on any domain D of $\mathbb{C}^{1,1}$ where the following points are excluded: (a) the point $(-d_i/c_i,0)$, namely, the points derivable from the infinity points by substitution of the group $\widetilde{\Gamma}$, including the infinity points themselves, (b) the infinities of $1/(z_i - a - \vartheta_i\alpha)$, that is, the point (a,α) and its analogs via $\widetilde{\Gamma}$. We prove (5.5) to be convergent with respect to the finest topology on $\mathbb{C}^{1,1}$ [that is, the Rogers one, see (ii)]. This means that we regard the series as a vector valued series (on the vector space $\mathbb{C}^m$, $m = 2^{L-1}$), thus showing the series to be convergent order by order. Using definition (3.4) and indexing $\widetilde{\Gamma}$ with i, we can rewrite (5.5) as

$$\sum_{i\in\widetilde{\Gamma}} \frac{\gamma_i z + \delta_i + \vartheta}{z_i - a - \vartheta_i\alpha} \frac{1}{(c_i z + d_i)^2}. \tag{5.6}$$

The coefficient of $(c_i z + d_i)^{-2}$ is a rational function of (z,ϑ), and we can say that for (z,ϑ) on D

$$\|(\gamma_i z + \delta_i)/(z_i - a - \vartheta_i\alpha)\|_{l_1} < \mathbf{M},$$

where $\|\cdot\|_{l_1}$ is the l_1 norm on $\mathbb{C}^{1,1}$, and $\mathbf{M}$ is a vector of $\mathbb{C}^m$. Therefore

$$\left|\left|\sum_i D\vartheta_i \frac{\vartheta_i}{z_i - a - \vartheta_i\alpha}\right|\right| < \mathbf{M}\left|\left|\sum_i (c_i z + d_i)^{-2}\right|\right|.$$

Then the convergence of (5.5) follows from the convergence of the series $\Sigma_i \operatorname{mod}(c_i z + d_i)^{-2}$, which can be proved by using for $\widetilde{\Gamma}$ the same arguments used to prove the convergence of the series $\Sigma_i \operatorname{mod}(c_i z + d_i)^{-2}$ related to the Schottky group Γ.[22]

The properties of the Poincaré super theta function are summarized in the following proposition.

Proposition 5.1: (i) Θ is automorphic of weight 1.

(ii) The difference $\Theta(z,\vartheta;a,\alpha) - \Theta(z,\vartheta;\widetilde{T}_i(a,\alpha))$ is independent of (a,α) for each $i = 1,\ldots,g$ and it is superanalytic on $\varepsilon^{-1}(H)$.

Proof: (i) Let $\widetilde{T}_J$ be any element of the group $\widetilde{\Gamma}$, then, writing the action of $\widetilde{T}_J$ by a subscript J, we have

$$\Theta(\widetilde{T}_J(z,\vartheta);a,\alpha) = \sum_{\sim} \frac{\widetilde{D}_J\,\tilde{\vartheta}_J}{\tilde{z}_J - a - \tilde{\vartheta}_J\alpha}\tilde{\vartheta}_J.$$

By (4.2) it follows that $\widetilde{D}_J = F_J\widetilde{D}$, where, F_J is just

$$\frac{(c_J z + d_J)^2}{(1 + \delta_i\,\gamma_J/2)(c_J z + d_J) + \vartheta(\gamma_J\,d_J - \delta_J\,c_J)}.$$

With a rescaling on the index of the series (written with a caret), we have the desired result:

$$\Theta(\widetilde{T}_J(z,\vartheta)) = F_J \sum_{J\in\widetilde{\Gamma}} \frac{D\tilde{\vartheta}}{\hat{z} - a - \tilde{\vartheta}\alpha}\hat{\vartheta} = F_J\Theta(z,\vartheta).$$

(ii) We write the calculation for the split case and give the result for the nonsplit one: the Θ for split SRS is

$$\Theta(z,\vartheta;a,\alpha) = \sum_{J\in\widetilde{\Gamma}} \frac{1}{(c_J z + d_J)^2} \frac{\vartheta}{(a_J z + b_J)/(c_J z + d_J) - a - \vartheta\alpha/(c_J z + d_J)}, \tag{5.7}$$

for any T of the form $\tilde{z} = (Az + B)/(Cz + D)$, $\tilde{\vartheta} = \vartheta/(Cz + D)$, we have

$$\begin{aligned}
\Theta(z,\vartheta;T(a,\alpha)) &= \sum_{J\in\widetilde{\Gamma}} \frac{1}{(c_J z + d_J)^2} \frac{\vartheta}{(a_J z + b_J)/(c_J z + d_J) - (A_a + B)/(C_a + D) - \vartheta\alpha/(c_J z + d_J)(C_a + D)} \\
&= \sum_{J\in\widetilde{\Gamma}} \frac{\vartheta(C_a + D)(c_J z + d_J)^{-1}}{(a_J z + b_J)(C_a + D) - (A_a + B)(c_J z + d_J) - \vartheta\alpha} \\
&= \sum_{s\in\widetilde{\Gamma}} \frac{\vartheta(C_a + D)[C(a_s z + b_s) + D(c_s z + d_s)]^{-1}}{a_s z + b_s - a(c_s z + d_s) - \vartheta_\alpha} \\
&\quad (\text{where } a_s = a_J D - Bc_J, \quad b_s = b_J D - Bd_J, \quad c_s = Ac_J - a_J C, \\
&\quad d_s = Ad_J - b_J C \Rightarrow c_J = Ca_s + Dc_s, \quad d_J = Cb_s + Dd_s) \\
&= \sum_s \frac{\vartheta}{(c_s z + d_s)^2} \frac{a + D/C}{(a_s z + b_s)/(c_s z + d_s) - a - \vartheta_\alpha(c_s z + d_s)} \frac{1}{(a_s z + b_s)/(c_s z + d_s) + D/C} \\
&= \Theta(z,\vartheta;a,\alpha) - \Theta(z,\vartheta;-D/C,0).
\end{aligned}$$

For nonsplit SRS the result is

$$\begin{aligned}
&\Theta(z,\vartheta;a,\alpha) - \Theta(z,\vartheta;T_i(a,\alpha)) \\
&\quad = \Theta(z,\vartheta;-d_i/c_i,-\delta_i/c_i) \\
&\quad = \Theta_i(z,\vartheta) \quad [T_i \text{ as in (A.4)}].
\end{aligned} \tag{5.8}$$

These functions are clearly superanalytic on the fundamental domain of $\widetilde{\Gamma}$, $\varepsilon^{-1}(H)$ [the poles being inside the circles $\varepsilon^{-1}(B_i)$]. We end up with some remarks about a possible definition of periodic matrix and Jacobi variety of a SRS. Formula (5.8) gives a basis of the space of superconformal differentials on M, if the spin structure induced on the body is even (with no harmonic spinors). For such SRS's it should be possible to define the concept of period matrix and Jacobi variety on the same lines of the RS case.

We can look at the "periods" of the super theta functions

$$A_{ik} = \int_{a_k} \Theta_i(z,\vartheta)dz\,d\vartheta, \quad B_{ik} = \int_{b_k} \Theta_i(z,\vartheta)dz\,d\vartheta. \tag{5.9}$$

These integrals are homological invariants for the SRS: by definition of the integral (see Sec. II), after the Berezin rule, the resulting line integral depends only on the homotopic relation of the curve with respect to the singularities of Θ_i and not on the specific curve. The periods (5.9) could be regarded as the entries of the (even valued) periodic matrix of M, P_M.

After normalization of such periods, the Jacobi variety of M would be the quotient of $\mathbb{C}^{g,0}$ with respect to the lattice generated by P_M. It is possible to make a definitive statement for the simple case of split SRS inducing an even spin structure: namely, after the Berezin rule, the periods are

$$A_{ik} = \int_{a_i} \tilde{\xi}_i(z)dz, \quad B_{ik} = \int_{k_i} \tilde{\xi}_i(z)dz,$$

where $\tilde{\xi}_i$ results the function defined by ξ_i via the formula (2.2). The singularities of $\tilde{\xi}_i$ are those of ξ_i and by the formula (2.3) we have that the lattice generated by P_M has only the body components latticized and the rest is continuous. That is the Jacobi variety of M: $J(M) = \mathbb{C}^{g,0}/P_M$ turns out to be a de Witt supermanifold with the body of the Jacobi variety of its body M_0 (as expected).

ACKNOWLEDGMENTS

L. Hodgkin and M. A. Sneider are warmly thanked for several discussions.

The Consiglio Nazionale delle Richerche (CNR) is acknowledged for financial support.

---

[1] A. Poliakov, Phys. Lett. B **103**, 211 (1981).

[2] *Superstrings*, edited by J. H. Schwartz (World Scientific, Singapore, 1985).

[3] M. Crane and J. Rabin, "S.R.S.'s: Uniformization and Teichmuller theory," Commun. Math. Phys. **113**, 601 (1988).

[4] J. Rabin, "Teichmuller deformations of S.R.S.'s," Phys. Lett. B **190**, 40 (1987).

[5] L. Hodgkin, "A direct calculation of super-Teichmuller space," Lett. Math. Phys. **14**, 74 (1987).

[6] L. Hodgkin, "On metrics and S.R.S.'s," Lett. Math. Phys. **14**, 156 (1987).

[7] B. de Witt, *Supermanifolds* (Cambridge U.P., Cambridge, 1984).

[8] A. Rogers, "Graded manifolds, supermanifolds and infinite dimensional Grassmann algebra," Commun. Math. Phys. **105**, 375 (1986).

[9] M. Rothstein, "Deformations of complex supermanifolds," Proc. Am. Math. Soc. **95**, 255 (1985).

[10] R. Catenacci, C. Reina, and P.Teofilatto, "On the body of super manifolds," J. Math. Phys. **26**, 671 (1985).

[11] P. Teofilatto, "Enlargeable graded Lie algebras of super symmetry," J. Math. Phys. **28**, 991 (1987).

[12] J. Arvis, "Classical dynamics of supersymmetric Liouville theory," Nucl. Phys. B **212**, 151 (1983).

[13] S. Lang, *Introduction to Modular Forms* (Springer, Berlin, 1976).

[14] J. Rabin and L. Crane, "Global properties of supermanifolds," Commun. Math. Phys. **100**, 141 (1985).

[15] A. Pais and V. Rittenberg, "Semisimple graded Lie algebras," J. Math. Phys. **16**, 2062 (1975).

[16] R. Gunning, *On Uniformization of Complex Manifolds*, Math. Notes 22 (Princeton U.P., Princeton, NJ, 1965).

[17] W. Abikoff, "Topics in the real analytic theory of Teichmuller space," *Lecture Notes in Mathematics*, Vol. 820 (Springer, New York, 1980).

[18] P. Teofilatto, "Line bundles and divisors on a S.R.S.," Lett. Math. Phys. **14**, 271 (1987).

[19] L. Hodgkin, "Problems on fields over S.R.S.'s," King's College, preprint, 1988.

[20] S. Giddings and P. Nelson, "Line bundles over S.R.S.'s," Harvard preprint, 1987; I thank the referee for pointing out this problem.

[21] L. Ford, *Automorphic Functions* (Chelsea, New York, 1951).

[22] H. Baker, *Abels Theorem and the Allied Theory* (Cambridge U.P., Cambridge, 1897).

Generalized Burgers equations and Euler–Painlevé transcendents. III

P. L. Sachdev, K. R. C. Nair, and V. G. Tikekar
Department of Applied Mathematics, Indian Institute of Science, Bangalore-560012, India

(Received 5 May 1987; accepted for publication 13 July 1988)

It was proposed earlier [P. L. Sachdev, K. R. C. Nair, and V. G. Tikekar, J. Math. Phys. **27**, 1506 (1986); P. L. Sachdev and K. R. C. Nair, *ibid.* **28**, 977 (1987)] that the Euler–Painlevé equations $y(d^2y/d\eta^2) + a(dy/d\eta)^2 + f(\eta)y(dy/d\eta) + g(\eta)y^2 + b(dy/d\eta) + c = 0$ represent generalized Burgers equations (GBE's) in the same way as Painlevé equations represent the Korteweg–de Vries type of equations. The earlier studies were carried out in the context of GBE's with damping and those with spherical and cylindrical symmetry. In the present paper, GBE's with variable coefficients of viscosity and those with inhomogeneous terms are considered for their possible connection to Euler–Painlevé equations. It is found that the Euler–Painlevé equation, which represents the GBE $u_t + u^\beta u_x = (\delta/2)g(t)u_{xx}$, $g(t) = (1+t)^n$, $\beta > 0$, has solutions, which either decay or oscillate at $\eta = \pm\infty$, only when $-1 < n < 1$. The solutions are shocklike when $n = 1$. On the other hand, they oscillate over the whole real line when $n = -1$. Furthermore, the solutions monotonically decay both at $\eta = +\infty$ and $\eta = -\infty$, that is, they have a single hump form if $\beta > \beta_n = (1-n)/(1+n)$. For $\beta < \beta_n$, the solutions have an oscillatory behavior either at $\eta = +\infty$ or at $\eta = -\infty$, or at $\eta = +\infty$ and $\eta = -\infty$. For $\beta = \beta_n$, there exists a single parameter family of exact single hump solutions, similar to those found for the nonplanar Burgers equations in Paper II. Thus the parametric value $\beta = \beta_n$ seems to bifurcate the families of solutions, which remain bounded at $\eta = \pm\infty$. Other GBE's considered here are also found to be reducible to Euler–Painlevé equations. The scope of these equations is broadened by relating them to a large number of nonlinear DE's selected from the compendia of Kamke [*Differential Gleichungen: Lösungsmethoden und Lösungen* (Akademische Verlagsgesellschaft, Leipzig, 1943)] and Murphy [*Ordinary Differential Equations and their Solutions* (Van Nostrand, Princeton, NJ, 1960)]. These latter equations arise from a wide range of physical applications and are of some historical interest as well. They are all special cases of a slightly generalized form of the Euler–Painlevé equation.

I. INTRODUCTION

In Papers I and II (Refs. 1 and 2, respectively), we proposed that the generalized Burgers equations (GBE's) are characterized by a class of nonlinear ordinary differential equations (ODE's), which we called Euler–Painlevé equations, in the same manner as the Korteweg–de Vries equations are typified by the Painlevé equations. These ODE's result from the self-similar reduction of the GBE's. We studied, in particular, the equations

$$u_t + u^\beta u_x + \lambda u^\alpha = (\delta/2)u_{xx} \tag{1.1}$$

and

$$u_t + u^\alpha u_x + (ju/2t) = (\delta/2)u_{xx} . \tag{1.2}$$

The self-similar forms of Eqs. (1.1) and (1.2), after some further transformations, are special cases of the nonlinear ODE's,

$$yy'' + ay'^2 + f(x)yy' + g(x)y^2 + by' + c = 0 , \tag{1.3}$$

which we referred to as the Euler–Painlevé equation.

In the present paper we continue our study of GBE's and their connection with the Euler–Painlevé equation (1.3). We consider here the GBE with a variable coefficient of viscosity,

$$u_t + u^\beta u_x = (\delta/2)g(t)u_{xx} , \tag{1.4}$$

where β is real and positive and $g(t)$ is a smooth function. Scott[3] has studied a special case of (1.4) in the form

$$u_x - uu_t = g(x)u_{tt} , \tag{1.5}$$

where the roles of x and t are reversed. In (1.5) t is a retarded time, x is a rangelike variable, u is an acoustic variable, with the linear effects of changes in the duct area taken out, and $g(x)$ is a positive function that depends on the particular duct chosen. More precisely, if R is the range along the duct area, then

$$t = T - R/c_0 , \quad u = a^{1/2}v ,$$

$$x = \frac{(\gamma+1)}{2a_0^2}\int a^{-1/2}(R)\,dr ,$$

and

$$g(x) = \delta a^{1/2}(R)/(\gamma+1)c_0 ,$$

where c_0 is the sound speed, γ is the adiabatic exponent, and δ is the diffusivity of sound that measures the combined effects of the viscosity and the thermal conductivity of the medium. The boundary condition appropriate to a piston or loudspeaker in the duct is

$$u(0,t) = u_0(t) . \tag{1.6}$$

The problem (1.5) and (1.6) also occurs in the treatment of spherical and cylindrical nonlinear sound waves, in which $g(x)$ has the particular forms $g(x) = e^x$ and x, respectively. Scott studied, in particular, strictly self-similar solutions of the form $u = \Omega(t/x)$. He considered the intermediate asymptotic behavior of three kinds of self-similar solutions

0022-2488/88/112397-08$02.50

of (1.5), depending on the nature of the function $g(x)$. When $g(x) = x$, Eq. (1.5) is called cylindrical. In this case, the self-similar solution $\Omega(t/x)$ that tends to constant values as $t \to \pm\infty$ forms an intermediate asymptotic: the solution of (1.5) with $g(x) \sim x$ as $x \to \infty$ and with any continuous initial condition $u_0(t)$ having constant asymptotic values evolves towards the similarity solution for large x. Thus the solution $u = \Omega(t/x)$ is very stable. For the situation $g(x)/x \to \infty$ as $x \to \infty$, the supercylindrical situation such as in the spherically symmetric case, the limiting form of u is an error function that is a solution of the linearized problem. Finally, if $g(x)/x \to 0$ as $x \to \infty$, the case Scott has referred to as subcylindrical, the limiting nonlinear self-similar solution $u = E(t/x)$ is shown to be an expansion front, consisting of straight line segments. Scott gives a rigorous analytic proof to show that the above solutions form intermediate asymptotics (see Barenblatt[4] and Sachdev[5]).

We study Eq. (1.4) instead of Eq. (1.5) to conform with our earlier work and the notation used therein. In particular, we assume that $g(t) = (1+t)^n$. We find that the solutions which either decay or oscillate at $x = \pm\infty$ exist only in the range $-1 \leqslant n \leqslant 1$. More specifically, the solutions decay at $x = +\infty$ and $x = -\infty$ if $\beta \geqslant \beta_n = (1-n)/(1+n)$, while they have an oscillatory behavior either at $x = +\infty$ or $x = -\infty$ or both at $x = +\infty$ and $x = -\infty$ if $\beta < \beta_n$. For $\beta = \beta_n$, there exists a single parameter family of exact single hump solutions, similar to those found for the nonplanar Burgers equation in Paper II. Thus the parametric value $\beta = \beta_n$ seems to bifurcate the families of solutions that remain bounded at $x = \pm\infty$.

Here again we find that the inverse function H [see Eq. (2.3)] is governed by a special case of the Euler–Painlevé equation (1.3). Scott's cylindrical solution is identified as a special case of our Eq. (1.4). In the terminology of Scott, the solutions of Eq. (1.5) with a general nonlinear convective term exist for the subcylindrical case only. It may be emphasized that while Scott's solutions are strictly self-similar, our solutions decay explicitly with time except when $n = 1$. As pointed out by Scott, the self-similar form of supercylindrical solutions derives from the linearized equation; therefore, the correspondence to the Euler–Painlevé equation does not exist in this case.

We also give here two other GBE's: the inhomogeneous Burgers equation and the GBE with a variable coefficient of viscosity, depending exponentially on time. While both equations can be reduced to the Euler–Painlevé form, physically relevant solutions exist only for the former. These solutions represent sawtooth form.

In this paper, we also refer to a set of ODE's in the compendia of Kamke[6] and Murphy[7] that are either special cases of the Euler–Painlevé equation (1.3) directly or are special cases of a generalized form of (1.3), in which the coefficients are made to vary with the independent variable. This brings Euler–Painlevé equations in contact with a much larger class of DE's which have appeared in diverse applications quite fortuitously.

The scheme of the paper is as follows: In Sec. II, we derive the nonlinear ODE for self-similar solutions, analyze it, and find its special exact solutions; we also pose the connection problem for the same. In Sec. III, we solve the connection problem numerically. In Sec. IV, we discuss the evolution of the single hump initial profile under the governance of the nonlinear partial differential equation to its intermediate asymptotic form. In Sec. V, we consider other GBE's and their (possible) connection to Euler–Painlevé transcendents. In Sec. VI, we give the equation numbers of ODE's related to Euler–Painlevé transcendents included in the compendia of nonlinear ODE's by Kamke[6] and Murphy.[7] Finally, the conclusions of this study are contained in Sec. VII.

II. ANALYSIS OF SELF-SIMILAR SOLUTIONS–EULER–PAINLEVÉ TRANSCENDENTS

As in Paper I, we seek self-similar forms of solutions of Eq. (1.4) with $g(t) = (1+t)^n$ (n is a parameter) in the form

$$u = (1+t)^{a_1} F(\xi), \quad \xi = (1+t)^{b_1} x, \tag{2.1}$$

where a_1 and b_1 are real constants, and determine the values of the parameters β and n for which self-similar solutions of Eq. (1.4) exist satisfying certain asymptotic conditions at $x = \pm\infty$. Substitution of (2.1) into (1.4) shows that, for the similarity form, $a_1 = -(1-n)/2\beta$ and $b_1 = -(1+n)/2$. With the scaling $f = \delta^{1/2\beta} F$ and $\eta = \delta^{-1/2}\xi$, Eq. (1.4) reduces to

$$f'' - 2f^\beta f' + (1+n)\eta f' + [(1-n)/\beta] f = 0, \tag{2.2}$$

where a prime denotes differentiation with respect to η. The inverse function

$$H = f^{-\beta} \tag{2.3}$$

satisfies the equation

$$HH'' - [(\beta+1)/\beta] H'^2 + (1+n)\eta HH' - (1-n)H^2 - 2H' = 0. \tag{2.4}$$

Equation (2.4) is a special case of (1.3) with $a = -(1+\beta)/\beta$, $f(x) = (1+n)x$, $g(x) = -(1-n)$, $b = -2$, and $c = 0$. We find the range of the parameter n, for which single hump type solutions exist for Eq. (2.2). At the maximum of the hump, $f' = 0$ and $f'' < 0$. Therefore, a necessary condition for the single hump solutions to exist is that $n < 1$. We derive later in this section a more restricted range of the parameter n for which solutions satisfying asymptotic conditions exist. We first seek some exact solutions of (2.2). For

$$\beta = (1-n)/(1+n) = \beta_n, \tag{2.5}$$

say, Eq. (2.2) assumes the form

$$f'' - 2f^\beta f' + [2/(1+\beta)](\eta f' + f) = 0. \tag{2.6}$$

Integrating (2.6) once and using vanishing conditions at $\eta \to \pm\infty$, we get

$$f' - [2/(1+\beta)] f^{1+\beta} + [2/(1+\beta)]\eta f = 0.$$

Integrating again, we have

$$f(\eta) = e^{[-1/(1+\beta)]\eta^2} h^{-1/\beta}(\eta), \tag{2.7}$$

where

$$h(\eta) = A - (2m)^{1/2}\,\mathrm{erf}[(m/2)^{1/2}\eta] \,, \quad A = f^{-\beta}(0) \quad \text{and} \quad m = 2\beta/(1+\beta) \,. \tag{2.8}$$

Correspondingly,

$$H(\eta) = e^{[\beta/(1+\beta)]\eta^2} h(\eta) \tag{2.9}$$

and

$$u(x,t) = \delta^{1/2\beta}(1+t)^{-1/(1+\beta)} e^{[-1/(1+\beta)]\eta^2} h^{-1/\beta}(\eta) \,, \tag{2.10}$$

$$\eta = \delta^{-1/2}(1+t)^{-1/(1+\beta)} x \,. \tag{2.11}$$

For general n and β we seek the Taylor series solution for Eq. (2.4), namely,

$$H(\eta) = \sum_{r=0}^{\infty} a_r \, \eta^r \,. \tag{2.12}$$

The coefficients a_{r+2}, $r = 1,2,\ldots$, are found by substituting (2.12) into (2.4),

$$\begin{aligned} a_{r+2} = \frac{1}{(r+1)(r+2)a_0}\Bigg[&\left(\frac{\beta+1}{\beta} a_1 + 2\right)(1+r)a_{r+1} \\ &+ (1-n)a_0 a_r + \sum_{i=1}^{r} (r+1-i)a_{r+1-i}\Bigg\{\frac{\beta+1}{\beta} \\ &\times (r+i-2)a_{r+i-2} - (1+n)a_{i-1}\Bigg\} \\ &- \sum_{i=1}^{r} a_i\{(r+2-i)(r+1-i)a_{r+2-i} \\ &- (1-n)a_{r-i}\}\Bigg] \,, \end{aligned} \tag{2.13a}$$

$$a_2 = (1/2a_0)\,[(\{[(\beta+1)/\beta]a_1 + 2)a_1 + (1-n)a_0^2] \,. \tag{2.13b}$$

Thus we have a two-parameter a_0, a_1 family of solutions. For $m = 1 - n = 2\beta/(1+\beta)$, the parameter $a_1 = -m$. This special choice corresponds to the exact solution (2.9). The free parameter a_0 gives a single parameter family of solutions. This could either be the amplitude parameter or the Reynolds number

$$R = \frac{1}{\delta_-}\int_{\infty}^{\infty} u \, dx \,,$$

which is the ratio of the area under the profile to the coefficient of diffusivity of sound.

We now investigate the linear behavior of Eq. (2.2) under the conditions $f, f' \to 0$ as $\eta \to \pm\infty$ and pose a connection problem over the interval $-\infty < \eta < \infty$. The linearized form of Eq. (2.2) is

$$f'' + (1+n)\eta f' + [(1-n)/\beta] f = 0 \,. \tag{2.14}$$

The change of variables

$$f(\eta) = y(z) \,, \quad z = -[(1+n)/2]\eta^2 \tag{2.15}$$

transforms (2.14) into

$$zy'' + (\tfrac{1}{2} - z)\,y' - \alpha y = 0 \,, \tag{2.16}$$

where $\alpha = (1/2\beta)(1-n)/(1+n)$. The solution of Eq. (2.16) is the confluent hypergeometric function $\phi(\alpha,\frac{1}{2};z)$. It vanishes as $\eta \to \pm\infty$ asymptotically only when $\alpha > 0$. This requires that $|n| < 1$. For $\beta = (1-n)/(1+n)$, $\alpha = \frac{1}{2}$, $\phi(\frac{1}{2},\frac{1}{2};z) = e^z = e^{[-(1+n)/2]\eta^2}$ is an exact solution of Eq. (2.14). We pose the connection or boundary value problem for Eq. (2.2), namely,

$$\begin{gathered} f'' - 2f^{\beta}f' + (1+n)\eta f' + [(1-n)/\beta] f = 0 \,, \\ f \sim A\phi(\alpha,\tfrac{1}{2}; -[(1+n)/2]\eta^2) \,, \quad \text{as} |\eta| \uparrow \infty \,, \\ |f| < \infty \,, \quad -\infty < \eta < \infty \,. \end{gathered} \tag{2.17}$$

It is interesting to observe the behavior of the solutions of Eqs. (2.2) and (2.4) for the limiting values -1 and $+1$ of the parameter n. For $n = 1$, Eqs. (2.2) and (2.4) assume the forms

$$f'' - 2f^{\beta}f' + 2\eta f' = 0 \,, \tag{2.18}$$

$$HH'' - [(1+\beta)/\beta]H'^2 + 2\eta HH' - 2H' = 0 \,, \tag{2.19}$$

where $f = A$, a constant, and $H = A^{-1/\beta}$ are special solutions of Eqs. (2.18) and (2.19), respectively. The linearized form of Eq. (2.18), namely,

$$f'' + 2\eta f' = 0 \,, \tag{2.20}$$

has the solution

$$f(\eta) = (2B/\pi^{1/2})\,\mathrm{erf}\,\eta \,, \tag{2.21}$$

where $B = f(0)$ is the amplitude parameter. For $n = -1$, Eq. (2.2) becomes

$$f'' - 2f^{\beta}f' + (2/\beta)f = 0 \,. \tag{2.22}$$

The linearized form

$$f'' + (2/\beta)f = 0 \tag{2.23}$$

has an oscillatory type of solution

$$f = B\cos(2/\beta)^{1/2}\eta \,. \tag{2.24}$$

The parametric value $\beta = \beta_n$ [see Eq. (2.5)] seems to bifurcate the types of solutions of Eq. (2.2) for $-1 < n < 0$. The transformation

$$f(\eta) = e^{[-(1+n)/4]\eta^2} s(\eta) \tag{2.25}$$

reduces the linear equation (2.14) to the form

$$s'' + \tfrac{1}{4}q(\eta)s = 0 \,, \tag{2.26}$$

where

$$q(\eta) = (4/\beta)(1-n) - 2(1+n) - (1+n)^2\eta^2 \,. \tag{2.27}$$

(i) For $\beta = \beta_n = (1-n)/(1+n)$, the right-hand side of Eq. (2.27) becomes

$$2(1+n) - (1+n)^2\eta^2 = r(\eta) \,, \tag{2.28}$$

say. In the linear regime $|\eta| > \eta_s$, where $\eta_s = [2/(1+n)]^{1/2}$, $r(\eta) < 0$, and therefore, $s(\eta)$ decays exponentially as $\eta \to \pm\infty$.

(ii) For $\beta > \beta_n$, $q(\eta) - r(\eta) = 4[((1-n)/\beta) - (1+n)] < 0$.

Therefore, $q(\eta) < r(\eta) < 0$ for $\eta < \eta_s$. Again in this case, $s(\eta)$ decays exponentially as $\eta \to \pm\infty$.

(iii) For $\beta < \beta_n$, $q(\eta) > (1+n)[2 - (1+n)\eta^2] > 0$ for $|\eta| < \eta_s < \infty$.

This suggests oscillatory solution of Eq. (2.2) for $|\eta| < \eta_s$. [In Sec. III, we confirm these results by numerical integration of Eq. (2.2).] In fact, it is interesting to note that for the limiting case $n = -1$, $\beta_n = \infty$, the solution of Eq. (2.2) is purely oscillatory in the interval $-\infty < \eta < \infty$.

We now summarize the similarity solutions found by previous investigators[8,9] for some specific values of the parameters n and β. Sachdev[5] discussed the self-similar solution of the equation

$$u_t + uu_x = (\delta/2)(1+t)u_{xx}\,. \tag{2.29}$$

The self-similar form of Eq. (2.29) was sought in the form

$$u = f(\xi)\,,\quad \xi = x/(1+t)\,. \tag{2.30}$$

The resulting equation after suitable transformations was reduced to a second-order linear ODE whose solution in the physically interesting case was

$$\begin{gathered} u = \pm\{(2\delta)[-S + S_0 e^{2(S-S_0)}]\}^{1/2} + \xi\,, \\ \xi = \gamma + \left(\frac{\delta}{2}\right)^{1/2} \int_{S_0}^{S} \frac{1}{H(S)}\,ds\,,\quad S \gg S_0\,, \end{gathered} \tag{2.31}$$

where S_0 is a parameter. The solution (2.31) has two constant parameters, γ and S_0. While γ can assume any value, the parameter S_0 can vary only between $-\frac{1}{2}$ and $+\infty$.

III. NUMERICAL SOLUTION OF CONNECTION PROBLEM (2.17)

The connection problem (2.17) was solved by numerically integrating the first equation in (2.17) starting from $\eta = \eta_s$ such that f and f' are small $O(10^{-3})$ and continuing the solution to $\eta \to -\infty$ for all $\beta > 0$ and $|n| < 1$. There are three varieties of solutions: (i) pure single humps, (ii) single humps ending with an oscillatory tail at one end, and (iii) solutions with oscillatory tails at both ends. Initial conditions were obtained from an asymptotic form of the confluent hypergeometric function for large η. In the linear regime, $\eta > \eta_s$, the behavior of the nonlinear equation (2.2) was compared with that of the linear equation (2.14) to assess the validity of the asymptotic solution. For $\eta > \eta_s$, the asymptotic solution of the confluent hypergeometric function agrees very closely with the numerical solution of the linear equation (2.14) and with that of the nonlinear equation (2.2) (see Table I). The solution of the linear equation continues to agree with that of the nonlinear one for $\eta < \eta_s$ and then begins to depart from it (see Fig. 1). Numerical solution of the nonlinear equation was checked with exact analytic solution (2.7) for special values of the parameters. The values of H and H' at η_s were also calculated from a known linear solution of Eq. (2.14) and the relation (2.3).

TABLE I. Comparison of solutions of linear equation (2.14) and nonlinear equation (2.2), and the asymptotic form of Eq. (2.16) for $n = -0.75$, $\beta = 3$.

η	Linear equation (2.14)	Asymptotic equation (2.16)	Nonlinear equation (2.2)
2.33	0.008 978	0.008 978	0.008 978
2.23	0.057 410	0.057 409	0.057 410
2.13	0.108 274	0.108 273	0.108 271
2.03	0.161 272	0.161 271	0.161 250
1.93	0.216 063	0.216 062	0.215 974
1.83	0.272 272	0.272 272	0.271 995
1.73	0.329 489	0.329 489	0.328 778
1.63	0.387 272	0.387 272	0.385 694
1.53	0.445 153	0.445 154	0.442 013
1.43	0.502 647	0.502 647	0.496 918
1.33	0.559 250	0.559 251	0.549 527
1.23	0.614 452	0.614 453	0.598 963
1.13	0.667 741	0.667 742	0.644 280
1.03	0.718 610	0.718 611	0.684 792
0.93	0.766 563	0.766 564	0.719 870

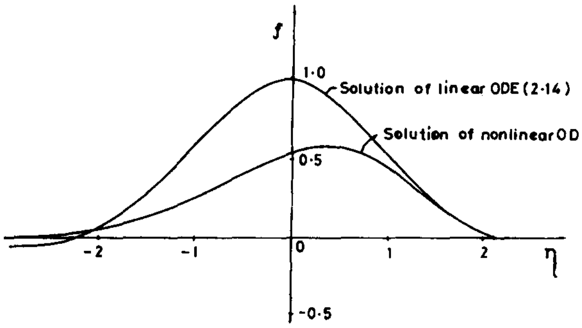


FIG. 1. Solutions of linear equation (2.14) and nonlinear equation (2.2) for $\beta = 1$, $n = -0.1$.

The series (2.12) for H was then summed up in the interval $-\infty < \eta < \eta_s < \infty$. Corresponding values of f were retrieved using (2.3) and compared with the numerical solution of the connection problem and the exact solution (2.7)–(2.11). All these solutions agree very well (see Table II for the three solutions for $n = -0.6$, $\beta = 4$). For the parametric ranges $0 < n < 1$ and $\beta \gg 1$, Eq. (2.17) has single hump solutions either vanishing at $\eta = \pm\infty$ or vanishing at $\eta = +\infty$ and tending to a nonzero constant at $-\infty$ [see Fig. 2 for solution of Eq. (2.2) for $\beta = 1$, $n = 0$, 0.25, 0.5, 0.9]. These types of solutions were also found in the range $-1 < n < 0$ for $\beta \gg \beta_n$ [see Fig. 3 for solution of Eq. (2.17) for $n = -0.8$, $\beta = 11$]. As predicted in Sec. II, the linear solution for $\beta < \beta_n$ and $-1 < n < 0$ has an oscillatory tail either at one or at both ends characteristic of Eq. (2.26) when $q(\eta) > 0$. The amplitude of the oscillatory tail increases and approaches that of the main hump as $n \to 1$ (see Figs. 4 and 5). The series solution for this limiting case does not converge at points where $H \to \infty$ corresponding to $f \to 0$.

TABLE II. Exact solution, numerical solution, and series solution of Eq. (2.2) for $n = -0.6$, $\beta = 4$, $\delta = 0.01$, $t = 5$.

η	Exact solution $f(\eta)$	Numerical solution $f(\eta)$	Series solution $H(\eta)$	$f(\eta)$
5.0	0.000 6738	0.000 6739	∞	0.000 6739
4.5	0.001 7423	0.001 7424	∞	0.001 7424
4.0	0.004 0764	0.004 0766	∞	0.004 0766
3.5	0.008 6297	0.008 6301	∞	0.008 6301
3.0	0.016 5306	0.016 5311	13 390 217.00	0.016 5311
2.5	0.028 6517	0.028 6526	1 483 698.00	0.028 6526
2.0	0.044 9347	0.044 9361	245 256.10	0.044 9360
1.5	0.063 7653	0.063 7672	60 480.08	0.063 7671
1.0	0.081 8758	0.081 8781	22 249.98	0.081 8781
0.5	0.095 1249	0.095 1275	12 211.70	0.095 1275
0.0	0.100 0002	0.100 0028	9 998.87	0.100 0028
-0.5	0.095 1213	0.095 1238	12 213.60	0.095 1238
-1.0	0.081 8707	0.081 8727	22 255.83	0.081 8727
-1.5	0.063 7606	0.063 7621	60 499.27	0.063 7620
-2.0	0.044 9312	0.044 9323	245 339.67	0.044 9322
-2.5	0.028 6494	0.028 6500	1 484 229.00	0.028 6500
-3.0	0.016 5293	0.016 5296	13 395 316.00	0.016 5296
-3.5	0.008 6290	0.008 6291	∞	0.008 6292
-4.0	0.004 0761	0.004 0761	∞	0.004 0761
-4.5	0.001 7421	0.001 7421	∞	0.001 7421
-5.0	0.000 6738	0.000 6737	∞	0.000 6738

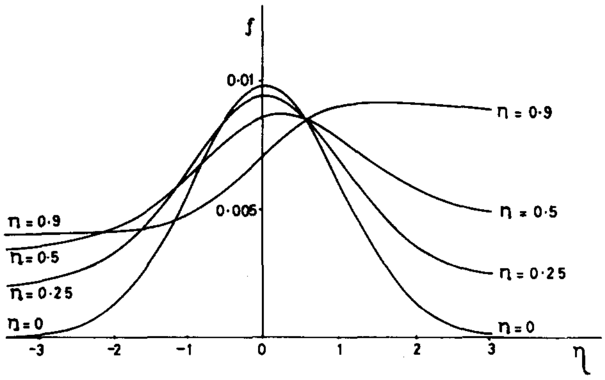


FIG. 2. Solution of the connection problem (2.17) for $\beta = 1$, $n = 0,0.25,0.5,0.9$.

Analytic continuation at such points was not feasible as the coefficients (2.13) of the series (2.12) become very large resulting in overflow of the partial sums of the series (2.12). This happens whenever f crosses the η axis. At other points the series (2.12) converges to the exact solution of Eq. (2.4).

Equations (2.18) and (2.22) were also solved for limiting values $n = 1$ and $n = -1$, respectively. For the former, starting with the initial condition $f = \text{const}$ at $\eta = \eta_s$, monotonic shocklike solutions were obtained. For the latter, starting with initial condition (2.24) at $\eta = \eta_s$, pure oscillatory solutions were obtained (see Fig. 5).

IV. NUMERICAL SOLUTION OF THE GBE (1.4) AND INTERMEDIATE ASYMPTOTICS

Equation (1.4) was solved subject to the initial condition

$$u(x,t_i) = s(x)\,, \quad -\infty < x < \infty\,, \tag{4.1}$$

where the function $s(x)$ is a smooth single hump. We used the implicit predictor–corrector scheme since the initial profile is smooth. The details of the scheme were reported extensively in Paper I. We consider the evolution of the initial profile into self-similar form for the following ranges of the parameters n and β.

(i) $-1 < n < 0$, $\beta \geqslant \beta_n$. In this case, the initial profile evolves to diffuse and goes into the self-similar form. Figures 6(a) and 6(b) show the evolution of the initial profile under the governance of Eq. (1.4), whereas Figs. 7(a) and 7(b) show the evolution of the initial profile into self-similar form for the parametric values $n = -0.6$, $\beta = 2$ and $n = -0.5$, $\beta = 4$.

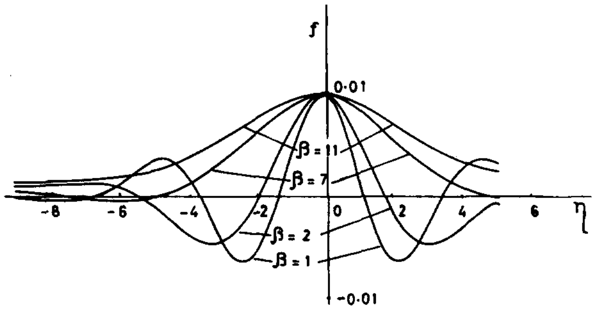


FIG. 3. Solution of the connection problem (2.17) for $n = -0.8$, $\beta = 1,2,7,11$.

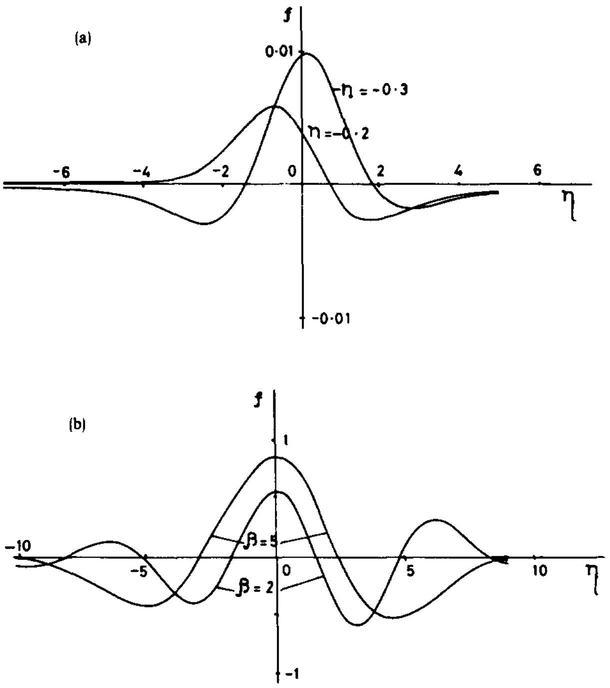


FIG. 4. Solution of the connection problem (2.17) for (a) $\beta = 1$, $n = -0.2, -0.3$; (b) $n = -0.9, \beta = 2,5$.

(ii) $0 < n < 1$, $\beta > 1$. The evolution of the single hump initial profile into its self-similar form takes place as in case (i) above.

(iii) $-1 < n < 0, \beta < \beta_n$. In this case, the solution of Eq. (1.4) starting from the initial profile (4.1) breaks at the front [see Fig. 6(c)]. The similarity solution (2.1) does not form an intermediate asymptotic; in this case, it may be recalled, the solution of the connection problem (2.17) contains oscillatory tails and the series (2.12) does not converge near zeros of $f(\eta)$.

V. OTHER GBE's, THEIR SELF-SIMILAR FORMS, AND SOLUTIONS

We consider two other GBE's, namely,

$$u_t + u^\alpha u_x = (\delta/2)e^{mt}u_{xx}\,, \quad \alpha > 0\,, \tag{5.1}$$

and

$$u_t + u^\beta u_x = (\delta/2)u_{xx} + \tfrac{1}{4}t^\gamma g[x(2\delta t)^{-1/2}]\,, \quad \beta > 0\,, \tag{5.2}$$

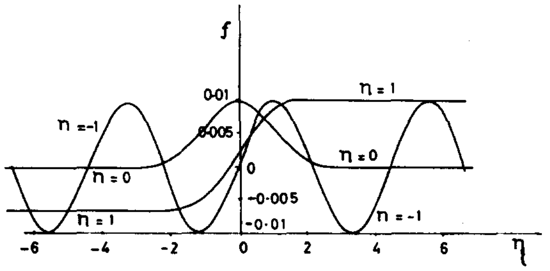


FIG. 5. Solution of the connection problem (2.17) for $\beta = 1$, $n = -1,0,1$.

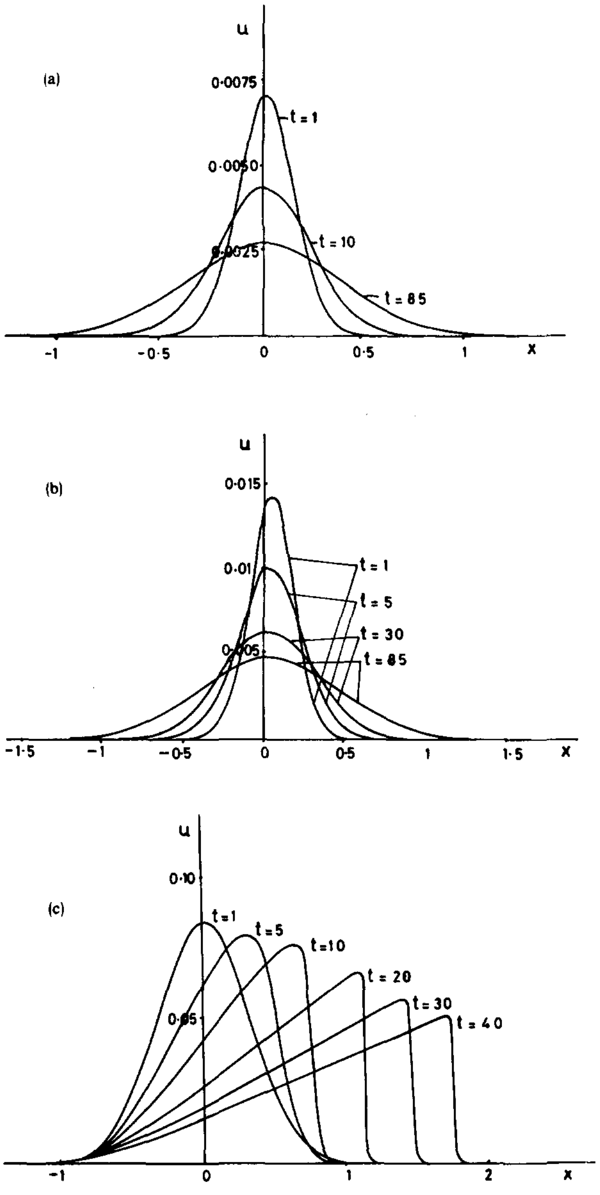


FIG. 6. Solution of the nonlinear PDE (1.10) for (a) $n = -0.6, \beta = 2, \delta = 0.01$; (b) $n = -0.5, \beta = 4, \delta = 0.01$; (c) $n = -0.9, \beta = 1, \delta = 0.01$.

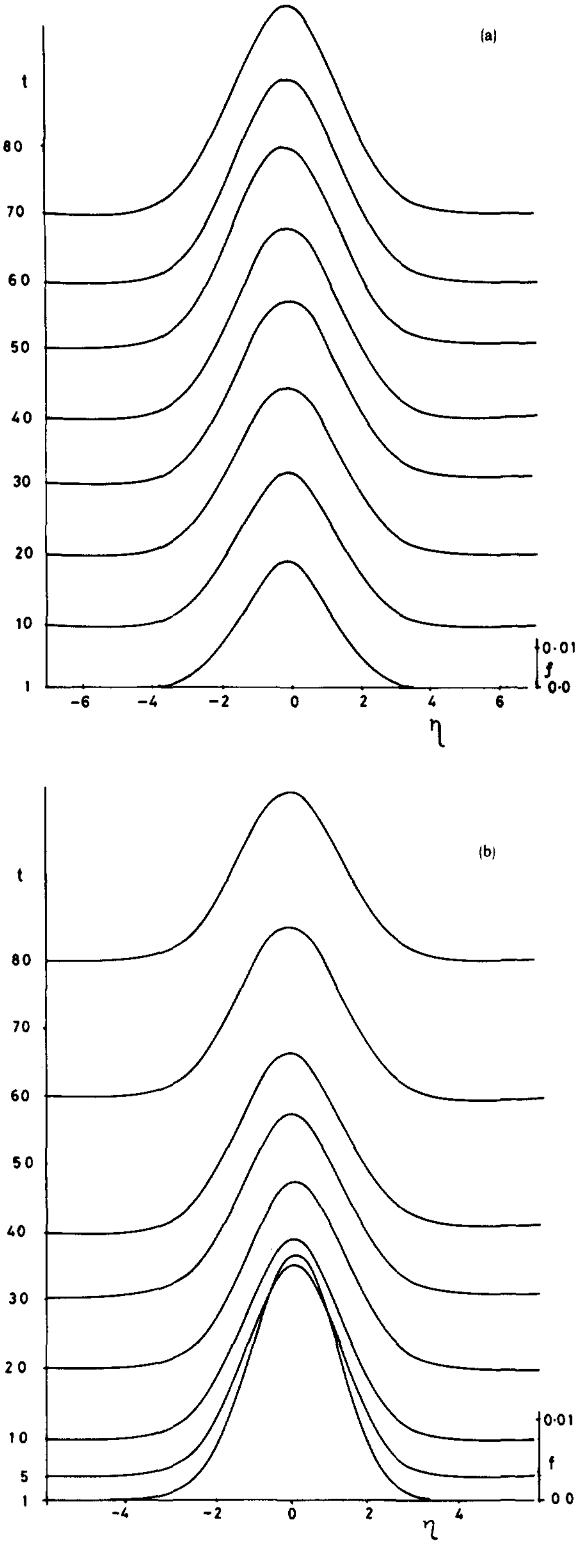


FIG. 7. Solution of the nonlinear ODE (2.17) for (a) $n = -0.6, \beta = 2$; (b) $n = -0.5, \beta = 4$.

and study their self-similar solutions and (possible) connection to the Euler–Painlevé transcendents. For the former, the similarity transformation corresponding to (2.1) is

$$u = (\delta e^{mt})^{1/2\alpha} f(\eta) , \quad \eta = x(\delta e^{mt})^{-1/2} . \tag{5.3}$$

The resulting equation

$$f'' - 2f^{\alpha} f' + m\eta f' - (m/\alpha) f = 0 \tag{5.4}$$

has a single hump type of solution if, at the maximum,

$$f'' = (m/\alpha) f < 0 . \tag{5.5}$$

A necessary condition for this is that $m < 0$, since we have assumed that $\alpha > 0$. We now look into the linear behavior of (5.4) for large η. The linearized form

$$f'' + m\eta f' - (m/\alpha) f = 0 \tag{5.6}$$

has the confluent hypergeometric function $\phi(-1/2\alpha, \frac{1}{2}; z)$ as one solution, where $z = -(m/2)\eta^2$. For a solution vanishing at $\eta = \pm \infty$, α should be negative. This contradicts the requirement for a single hump type of solution discussed above. However, the transformation

$$H = f^{-\alpha} \tag{5.7}$$

does give a representation of Eq. (5.4) in the Euler–Painlevé form, namely,

$$HH'' - [(\alpha+1)/\alpha]H'^2 - m\eta HH' - 2H' + mH^2 = 0 . \tag{5.8}$$

The inhomogeneous Burgers equation (5.2) has a similarity form

$$u = t^{-1/2\beta} f(\eta) , \quad \eta = x(2\delta t)^{-1/2} , \tag{5.9}$$

provided $\gamma = -(1 + 1/2\beta)$. The corresponding equation for f is

$$f'' - 4(2\delta)^{-1/2} f^\beta f' + 2\eta f' + (2/\beta) f + g(\eta) = 0 . \tag{5.10}$$

The inverse function

$$H = \delta^{1/2} f^{-\beta} \tag{5.11}$$

satisfies the equation

$$HH'' - [(\beta+1)/\beta]H'^2 + 2\eta HH' - 2^{3/2}H' - 2H^2 - \beta\delta^{-1/2\beta} g(\eta) H^{(1+1/\beta)} = 0 . \tag{5.12}$$

Equation (5.12) does not belong to the class of Euler–Painlevé equations, since the last term in it contains a higher degree term in H. However, we shall present here some exact solutions of (5.2) found earlier (see Sachdev[10]) for some special choice of the parameters and the function g in Eq. (5.2). For this purpose we scale out δ and write Eq. (5.2) in the simpler form (with $\beta = 1$)

$$u_t + uu_x = u_{xx} + \tfrac{1}{4} t^\gamma g(x,t) . \tag{5.13}$$

We consider the following cases.

$$\text{(i) } \gamma = -\tfrac{3}{2}, g = -\eta, \eta = xt^{-1/2}. \tag{5.14}$$

With the transformations

$$u = t^{-1/2} f(\eta) \tag{5.15}$$

and

$$H = f^{-1} , \tag{5.16}$$

the equations corresponding to (5.10) and (5.12) are

$$f'' - ff' + \tfrac{1}{2}\eta f' + \tfrac{1}{2} f - \tfrac{1}{4}\eta = 0 \tag{5.17}$$

and

$$HH'' - 2H'^2 + \tfrac{1}{2}\eta HH' - H' - \tfrac{1}{2}H^2 + \tfrac{1}{4}\eta H^3 = 0 . \tag{5.18}$$

Equation (5.13) has exact solutions

$$u = x/2t \tag{5.19}$$

and

$$u = (x/2t) - (2/x) . \tag{5.20}$$

Solution (5.19) is the sawtooth form shown in Fig. 5 of Benton and Platzman[11] (BP) and (5.20) corresponds to (3.5) of BP. Corresponding solutions of Eqs. (5.17) and (5.18) may be found. Thus

$$f = \tfrac{1}{2}\eta \tag{5.21a}$$

and

$$f = \tfrac{1}{2}\eta - 2/\eta \tag{5.21b}$$

are the solutions of Eq. (5.17), and

$$H = 2/\eta \tag{5.22a}$$

and

$$H = 2\eta/(\eta^2 - 4) \tag{5.22b}$$

are those of Eq. (5.18).

$$\text{(ii) } \gamma = -\tfrac{3}{2}, \quad g \equiv R(\eta), \quad \eta = x/2t^{1/2} . \tag{5.23}$$

The transformation

$$u = t^{-1/2}(f + \eta) \tag{5.24}$$

changes Eq. (5.13) into

$$f'' - 2ff' + R(\eta) + 2\eta = 0 . \tag{5.25}$$

This equation can be put in the standard Riccati form

$$\frac{df}{d\eta} - f^2 = -\int (2\eta + R(\eta))\, d\eta . \tag{5.26}$$

These similarity solutions correspond to (3.1′) of BP, shown in their Figs. 9–11 and describe solitary compression pulses (Lighthill[12]).

VI. EQUATIONS RELATED TO EULER–PAINLEVÉ TRANSCENDENTS—COMPENDIA OF KAMKE AND MURPHY

It is interesting to note that there is a large number of nonlinear DE's of second order, listed in Kamke[6] and Murphy,[7] which form special cases of (1.3), directly or after some simple transformations. We list here the equation numbers of these and other DE's which are special cases of (1.3) if we allow the coefficients a,b,c to vary with x. Kamke has appended some historical notes with each of these equations and has also given their geometrical or physical origin. Here, we content ourselves with listing the equation numbers. We note that a few common features characterize this set of equations: the equations are either autonomous or are linearizable by a logarithmic or a power law transformation, or they may be reducible to first-order equations such as Riccati and Bernoulli. The equations may be solved in closed form in terms of a quadrature or treated in the phase plane.

The following DE's are from Kamke[6]: 6.104–11, 6.117, 6.122, 6.124–27, 6.129, 6.131, 6.133–34, 6.136–39, 6.150–52, 6.155–58, 6.164, 6.166, 6.168–70, 6.173–79 (41 equations).

The following DE's are from Murphy[7]: 129–30, 133, 138, 140, 142, 150, 190, 195, 199, 201, 203–4, 219–22, 227–31, 233–34 (24 equations).

VII. CONCLUSIONS

The present paper extends our earlier studies on generalized Burgers equations in Papers I and II to encompass GBE's with variable coefficient of viscosity and inhomogeneous Burgers equations. The latter equations are also shown to reduce to Euler–Painlevé form. The GBE with the (time) power law coefficient exhibits new behavior not found earlier for nonplanar GBE or GBE with damping. There exist solutions in addition to single hump type, which oscillate either at $x = +\infty$ or at $x = -\infty$ or at $x = +\infty$ and $x = -\infty$. The transition to oscillatory behavior takes place when the parameter β, the degree of nonlinearity in the convective term, assumes a definite value equal to $(1-n)/(1+n)$ [see Eq. (2.5)]. We also find some exact solutions for special cases of the GBE's considered in the present paper. The scope of Euler–Painlevé equations is considerably enlarged by juxtaposing them with a large number of nonlinear DE's garnered by Kamke[6] and Murphy[7] from different sources and applications. It would seem, therefore, that the generalized Euler–Painlevé equations [Eq. (1.3)] have a larger role to play in a variety of applications than would be suggested by GBE's alone.

ACKNOWLEDGMENT

Part of this work was carried out under the DST (India) Project AM:PLS:129.

[1] P. L. Sachdev, K. R. C. Nair, and V. G. Tikekar, J. Math. Phys. **27**, 1506 (1986).

[2] P. L. Sachdev and K. R. C. Nair, J. Math. Phys. **28**, 997 (1987).

[3] J. F. Scott, Proc. R. Soc. London, Ser. A **373**, 443 (1981).

[4] G. I. Barenblatt, *Similarity, Self-Similarity and Intermediate Asymptotics* (Consultants Bureau, New York, 1979).

[5] P. L. Sachdev, *Nonlinear Diffusive Waves* (Cambridge U.P., London, 1987).

[6] E. Kamke, *Differential Gleichungen: Lösungsmethoden und Lösungen* (Akademische Verlagsgesellschaft, Leipzig, 1943).

[7] G. M. Murphy, *Ordinary Differential Equations and their Solutions* (Van Nostrand, Princeton, NJ, 1960).

[8] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics*, English translation by R. T. Beyer (Consultants Bureau, New York, 1977).

[9] T. H. Chong and L. Sirovich, J. Fluid Mech. **58**, 53 (1973).

[10] P. L. Sachdev, Quart. Appl. Math. **34**, 118 (1976).

[11] E. R. Benton and G. W. Platzman, Quart. Appl. Math. **30**, 195 (1972).

[12] M. J. Lighthill, in *Surveys in Mechanics*, edited by G. K. Batchelor and R. M. Davies (Cambridge U.P., Cambridge, 1956), pp. 258–351.

Some unbounded commutants of a set of operators on a partial inner product space

J. Shabani[a]
International Centre for Theoretical Physics, Trieste, Italy

(Received 10 December 1987; accepted for publication 29 June 1988)

This paper continues our systematic study of unbounded commutants within the framework of operators on a partial inner product (PIP) space. The most general commutant of a set of operators on a PIP space is considered and its behavior with respect to a topology finer than the weak and quasiweak *-topologies used in previous investigations is studied. The relationship between a bicommutant introduced by Shabani and a bicommutant introduced by Araki and Jurzak for closed Op*-algebras satisfying some countability conditions is given.

I. INTRODUCTION

In recent years unbounded commutants were analyzed by many authors,[1–11] both from the mathematical point of view along the lines of the usual theory of W^*- and Op*-algebras, and for their applications in quantum field theories.

In Refs. 10 and 11 a general theory of unbounded commutants is developed using the framework of operators on a partial inner product (PIP) space[12,13] and some of the results of Refs. 1–9 are extended to their general case.

This paper continues the systematic study of unbounded commutants started in Refs. 10 and 11. Let V be a PIP space and m be an *-invariant subset of the space Op V of all operators on V. We will define the commutant of m as being the subset m' of all operators of Op V that commute with m (see Sec. II for definitions). This commutant is the most general one for a set of elements of Op V. We will also be concerned with the bicommutant m'' and we will study topological properties of both m' and m'' with respect to a topology finer than the weak and quasiweak *-topologies used in Refs. 10 and 11.

The paper is organized as follows. In Sec. II, following Refs. 12–15 we recall briefly the basic properties of PIP spaces and operators on them. Let m be an *-invariant subset of Op V. We define on the space Lm of all left multipliers of m (resp. the space Rm of all right multipliers of m) the locally convex topologies $t^r(m)$ and $t^r_*(m)$ [resp. $t^l(m)$ and $t^l_*(m)$] and prove some sequential completeness properties of Op V, Lm, and Rm with respect to these topologies. In this section we also define the commutant and bicommutant we will study in this paper and compare them with those introduced in Refs. 1–11.

In Sec. III after noting that the four topologies defined on Lm and Rm coincide on the commutant m' [we denote this topology by $t(m)$], we investigate the topological properties of m' (resp. m'') with respect to the $t(m)$-topology [resp. the $t(m')$-topology]. In particular, we show that m' is closed in $Lm \cap Rm$ with respect to the $t(m)$-topology and similarly for m'' in $Lm' \cap Rm'$ with respect to the $t(m')$-topology. Thus here these topologies seem to play the role of the weak topology for bounded operators. In this section we also prove under the assumption of reflexivity of the dual pair $\langle V^{\#}, V \rangle$ that m' is $t(m)$ sequentially complete.

In Sec. IV we answer the following question: When is our bicommutant the closure of the original set of operators with respect to the $t(m')$-topology? We give a sufficient condition yielding this result.

Section V provides a comparison of our bicommutant with a bicommutant (to be denoted by m'_A) introduced in Ref. 4 for closed Op*-algebras satisfying some countability conditions. We show that our bicommutant is contained in the bicommutant of Ref. 4.

An analogous comparison for commutants was carried out in Ref. 10 where the equality $m' = m'_A = m'_C = m'_\sigma$ was obtained. Here m_C denotes the strong unbounded commutant,[1,3] whereas m'_σ stands for the weak unbounded commutant.[5,8]

II. THE PARTIAL *-ALGEBRA OF OPERATORS ON A PARTIAL INNER PRODUCT (PIP) SPACE

A. Abstract partial *-algebra (Refs. 16 and 17)

Definition 2.1: A *partial *-algebra* is a complex vector space U with an antilinear involution $x \mapsto x^*$ and a subset $\Gamma \subset U \times U$ such that:

(i) $(x,y) \in \Gamma$ iff $(y^*,x^*) \in \Gamma$;

(ii) if $(x,y) \in \Gamma$ and $(x,z) \in \Gamma$, then $(x, \lambda y + \mu z) \in \Gamma$, for all $\lambda, \mu \in \mathbb{C}$;

(iii) whenever $(x,y) \in \Gamma$, there exists an element $xy \in U$ with the usual properties of product:

$$x(y + \lambda z) = xy + \lambda(xz), \quad (xy)^* = y^* z^* .$$

Definition 2.2: Let $m \subset U$. We define the set of *left multipliers* of m by

$$Lm = \{x \in U \mid (x,y) \in \Gamma, \text{ for all } y \in m\}.$$

Similarly the set of right multipliers of m is given by

$$Rm = \{x \in U \mid (y,x) \in \Gamma, \text{ for all } y \in m\}.$$

In particular, for single elements we have

$$L(z) \equiv L\{z\} \text{ and } R(z) \equiv R\{z\}.$$

This suggests the simpler notation,

$$(x,y) \in \Gamma \Leftrightarrow x \in L(y) \Leftrightarrow y \in R(x).$$

Definition 2.3: A vector subspace $m \subset U$ is called a **-subalgebra* of the partial *-algebra U if the following conditions are satisfied:

(i) m contains the identity;

(ii) m is *-invariant, i.e., $x \in m$ implies $x^* \in m$;

---

[a] On leave of absence from the Department of Mathematics, University of Burundi, B.P. 2700 Bujumbura, Burundi.

0022-2488/88/112405-06$02.50

(iii) if $x,y \in m$ and $x \in L(y)$, then $xy \in m$.

As pointed out in Ref. 18, for partial *-algebras, the usual definition of associativity is rarely realized in practice. However, for most purposes, a weaker concept is sufficient.

Definition 2.4: The partial *-algebra U is called *semiassociative* if for any $x,y,z \in U$ such that $y \in R(x)$ and $z \in RU$ the following conditions are satisfied:

(i) $yz \in R(x)$;

(ii) $(xy)z = x(yz)$.

B. Operators on a PIP space

1. Basic properties

A PIP space[12–15] consists of a complex vector space V, a nondegenerate Hermitian form $\langle \cdot | \cdot \rangle$, and a family of vector subspaces $\{V_r, r \in I\}$ satisfying the following conditions.

(i) The family $\mathscr{S} = \{V_r, r \in I\}$ covers V and is an involutive lattice with respect to set intersection, vector sum, and involution $\#$: $V_r \leftrightarrow V_{\bar{r}}$. Besides elements of $\mathscr{S}$, we consider also the extreme spaces

$$V^{\#} = \bigcap_{r \in I} V_r \quad \text{and} \quad V = \bigcup_{r \in I} V_r .$$

(ii) The Hermitian form $\langle \cdot | \cdot \rangle$, called the partial inner product, is defined on $\bigcup_{r \in I} V_r \times V_{\bar{r}}$.

(iii) V possesses a central Hilbert space, i.e., there exists an element $o = \bar{o}$ in I such that $V_o = V_{\bar{o}} \equiv \mathscr{H}$ is a Hilbert space with respect to $\langle \cdot | \cdot \rangle$.

The assumption of nondegeneracy $(V^{\#})^{\perp} = \{o\}$ implies that every pair $\langle V_r, V_{\bar{r}} \rangle$ as well as $\langle V^{\#}, V \rangle$ is a dual pair with respect to $\langle \cdot | \cdot \rangle$. Therefore each V_r may be endowed with its canonical Mackey topology $\tau(V_r, V_{\bar{r}})$ and similarly for $V^{\#}, V$. This choice implies the following.

(i) Whenever $V_p \subset V_q$, the embedding E_{qp}: $V_p \hookrightarrow V_q$ is continuous and has dense range.

(ii) $V^{\#}$ is dense in every V_r and every V_r is dense in V.

An *operator*[13] on the PIP space V is a map A: $\mathscr{D}(A) \to V$, where $\mathscr{D}(A)$ is the largest union of subspaces V_r such that the restriction of A to any of them is linear and continuous into V. Such operators may be extremely singular, since the range of $A|_{V^{\#}}$ may be much larger than the central Hilbert space $\mathscr{H}$. Yet every operator A has an adjoint $A^{\times}$, which is also an operator on V, and the correspondence $A \leftrightarrow A^{\times}$ is an involution on the set Op V of all operators on V. The set Op V is a vector space but not an algebra (it is a partial $^{\times}$-algebra[16,17]); two operators A and B may always be added, but their product AB is defined only if there is a continuous factorization through some V_q,

$$V^{\#} \xrightarrow{B} V_q \xrightarrow{A} V.$$

An operator $A \in \mathrm{Op}\, V$ is called *regular*[19] if $\mathscr{D}(A) = \mathscr{D}(A^{\times}) = V$; equivalently, if A maps both $V^{\#}$ and V into themselves continuously. It is well known that equipped with the involution $A \leftrightarrow A^*$, where A^* is the restriction to $V^{\#}$ of the adjoint operator $A^{\times}$, the set Reg V of all regular operators on V is a *-algebra, isomorphic to an Op*-algebra,[20] i.e., a *-sublagebra with unit of the algebra $L^{+}(V^{\#})$ of all closable operators on $\mathscr{H}$, which together with their (Hilbertian) adjoint leave $V^{\#}$ invariant. The space Op V contains another remarkable subset, namely,

$$C(V^{\#},\mathscr{H}) = \{A \text{ closable in } \mathscr{H} | V^{\#} \subset \mathscr{D}(A) \cap \mathscr{D}(A^*)\}$$
$$= \{A \in \mathrm{Op}\, V | A, A^*\colon\ V^{\#} \to \mathscr{H}\}.$$

We have Reg $V \subseteq C(V^{\#},\mathscr{H}) \subseteq \mathrm{Op}\, V$.

We will assume that V is quasicomplete in its Mackey topology. This implies in particular that Reg V is isomorphic to $L^{+}(V^{\#})$; see Ref. 19, Proposition 2.5. The condition of Mackey quasicompleteness of V is actually satisfied in almost all examples; the only known exceptions are pathological.[21,22]

Proposition 2.5: The space Op V is semiassociative.

Proof: Let $A,B,C \in \mathrm{Op}\, V$ be such that $B \in R(A)$ and $C \in R\, \mathrm{Op}\, V$. We recall that $R\, \mathrm{Op}\, V = \{X \in \mathrm{Op}\, V|$ X: $V^{\#} \to V^{\#}\}$. Let us show that $BC \in R(A)$ and $(AB)C = A(BC)$.

For all $f \in V^{\#}$ we have

$$[A(BC)]f = A(BC)f = A(BCf) = (AB)Cf.$$

The product $(AB)C$ is well defined since C: $V^{\#} \to V^{\#}$ and $B \in R(A)$. Therefore the product $A(BC)$ exists, i.e., $BC \in R(A)$ and $A(BC) = (AB)C$.

2. Topologies on the spaces of multipliers

Let m be a $^{\times}$-invariant subset with unit of Op V. Then m generates two locally convex topologies on Rm defined by the following family of seminorms:

$$t^{l}(m)\colon\ B \in Rm \mapsto |\langle (AB)f,g \rangle| ,$$

$$t^{l}_{*}(m)\colon\ B \in Rm \mapsto |\langle (AB)f,g \rangle| + |\langle (AB)^{\times}f,g \rangle| ,$$

$$\forall f,g \in V^{\#} \text{ and } A \in m .$$

Similarly one may define on Lm the following topologies:

$$t^{r}(m)\colon\ C \in Lm \mapsto |\langle (CA)f,g \rangle| ,$$

$$t^{r}_{*}(m)\colon\ C \in Lm \mapsto |\langle (CA)f,g \rangle| + |\langle (CA)^{\times}f,g \rangle|,$$

$$\forall f,g \in V^{\#} \text{ and } A \in m .$$

In general we have $t^{l}(m) < t^{l}_{*}(m)$ and $t^{r}(m) < t^{r}_{*}(m)$; where $<$ means "weaker than." If $m = \{1\}$, then $t^{l}(1) = t^{r}(1) \equiv t(1)$ is the weak topology considered in Ref. 10 (also called $V^{\#}$-weak topology), whereas $t^{l}_{*}(1) = t^{r}_{*}(1) \equiv t_{*}(1)$ coincides with the quasiweak *-topology introduced in Ref. 11.

These topologies are related in the following way:

$$\begin{array}{ccccccc} t^{r}(m) & > & t^{r}(1) & = & t^{l}(1) & > & t^{l}(m) , \\ \wedge & & \wedge & & \wedge & & \wedge \\ t^{r}_{*}(m) & > & t^{r}_{*}(1) & = & t^{l}_{*}(1) & < & t^{l}_{*}(m) . \end{array}$$

Proposition 2.6: Let V be a PIP space. If $\langle V^{\#}, V \rangle$ is a reflexive dual pair, then Op V is $t(1)$-sequentially complete.

Proof: Let $\{T_n\}$ be a $t^{l}(1)$-Cauchy sequence in Op V, i.e., for all $f \in V^{\#}$, $\{T_n f\}$ is a $t^{l}(1)$-Cauchy sequence in V. Since $\langle V^{\#}, V \rangle$ is reflexive, it follows that $V^{\#}$ and V are quasicomplete with respect to the weak topologies $\sigma(V^{\#}, V)$ and $\sigma(V, V^{\#})$. Therefore $V^{\#}$ and V are weakly sequentially complete, i.e.,

$$\text{w-lim}_{n \to \infty} T_n f = Tf \in V.$$

This means that T is a map from $V^{\#}$ into V. In order to prove that T is continuous, one uses the dual mapping theory.[23]

Remark 2.7: Actually the condition of reflexivity of the dual pair $\langle V^{\#}, V\rangle$ is weak enough to cover most of the spaces of practical interest, in particular all spaces of distributions.[15]

Proposition 2.8: Let m be a ${}^{\times}$-invariant subset with unit of Op V. If for every $r \in I$, $\langle V_r, V_{\bar{r}}\rangle$ is a reflexive dual pair, then Rm is $t^l(m)$-sequentially complete. Similarly Lm is $t^r(m)$-sequentially complete.

Proof: Let $\{B_n\} \subset Rm$ be a $t^l(m)$-Cauchy sequence. Since the $t^l(1)$-topology is weaker than the $t^l(m)$-topology [we write $t^l(1) < t^l(m)$], it follows that $\{B_n\}$ is also a $t^l(1)$-Cauchy sequence. Moreover, since Op V is $t^l(1)$-sequentially complete, there exists a $t^l(1)$-limit B of $\{B_n\}$ such that $B \in$ Op V. The sequence $\{B_n\}$ is a $t^l(1)$-Cauchy sequence, which means that for every $A \in m$, $\{AB_n\}$ is a $t^l(m)$-Cauchy sequence and again the fact that $t^l(1) < t^l(m)$ implies that $\{AB_n\}$ is also a $t^l(1)$-Cauchy sequence, i.e., $t^l(1)\text{-}\lim_{n\to\infty} AB_n = Q \in \text{Op } V$, which means that for all $f,g \in V^{\#}$ we have

$$\lim_{n\to\infty} \langle AB_n f, g\rangle = \langle Qf, g\rangle .$$

The sequence $\{B_n f\}$ is a $t^l(1)$-Cauchy sequence in $\mathscr{D}(A)$ and since for every $r \in I$, $\langle V_r, V_{\bar{r}}\rangle$ is a reflexive dual pair, it follows that $\mathscr{D}(A)$ is $t^l(1)$-sequentially complete. Therefore

$$t^l(1)\text{-}\lim_{n\to\infty} B_n f = Bf \in \mathscr{D}(A)$$

and hence

$$t^l(1)\text{-}\lim_{n\to\infty} AB_n f = ABf = Qf,$$

which means that $Q = AB$, i.e., $B \in Rm$.

3. Commutants and bicommutants

(a) Commutants: Let m be a ${}^{\times}$-invariant subset of Op V. In Ref. 10 the following commutant was introduced:

$$m' = \{X \in \text{Op } V \mid X \in Lm \cap Rm \equiv M(m),$$
$$XA = AX,\ \forall A \in m\} .$$

It was pointed out there that m' is a vector subspace of $M(m)$. Moreover, it is ${}^{\times}$-invariant and contains the identity. But nothing more is known about this commutant and our aim in this paper is to perform a systematic analysis of m' along the lines of the usual theory of von Neumann algebras.

The regular part of m', i.e.,

$$m'_c \equiv m' \cap L^{+}(V^{\#}) = \{X \in L^{+}(V^{\#}) \mid AX = XA,\ \forall A \in m\}$$

is the strong unbounded commutant studied in Ref. 10. It is an Op*-algebra on $V^{\#}$.

If m is an Op*-algebra on $V^{\#}$, then the condition $X \in M(m)$ in the definition of m' is automatically satisfied and in this case m' coincides with the commutant m'_0 considered in Ref. 10.

Furthermore, if m is an Op*-algebra on $V^{\#}$, then one may define the following (weak) commutant is Op V:

$$m'_{\breve{\sigma}} = \{X \in \text{Op } V \mid \langle Xf, A^{*}g\rangle = \langle Af, X^{*}g\rangle,$$
$$\forall f,g \in V^{\#} \text{ and } A \in m\} .$$

We note that $m'_\sigma \equiv m'_{\breve{\sigma}} \cap C(V^{\#}, \mathscr{H})$ is the weak unbounded commutant introduced in Ref. 5 (see also Ref. 8).

From the above discussion, it follows that the different commutants introduced in this section are related in the following way:

$$m'_c \subseteq m'_\sigma \subseteq m' = m'_0 = m'_{\breve{\sigma}} .$$

Indeed we have the following proposition.

Proposition 2.9: If m is an Op*-algebra on $V^{\#}$, then $m'_0 = m'_{\breve{\sigma}}$.

Proof: Let $X \in m'_0$, i.e., for all $f,g \in V^{\#}$ and $A \in m$ we have

$$\langle (XA)f, g\rangle = \langle (AX)f, g\rangle .$$

Then

$$\langle (XA)f, g\rangle = \langle Af, X^{*}g\rangle = \langle f, A^{*}X^{*}g\rangle .$$
$$= \langle f, X^{*}A^{*}g\rangle = \langle Xf, A^{*}g\rangle .$$

Thus if $X \in m'_0$ then $\langle Af, X^{*}g\rangle = \langle Xf, A^{*}g\rangle$, i.e., $X \in m'_{\breve{\sigma}}$.

(ii) Let now $X \in m'_{\breve{\sigma}}$. Then for all $f,g \in V$ and $A \in m$ we have $\langle (XA)f, g\rangle = \langle Af, X^{*}f\rangle = \langle Xf, A^{*}g\rangle = \langle (AX)f, g\rangle$, i.e., $X \in m'_0$.

At this stage, the natural question which arises is the following: When do the five commutants introduced above coincide. Let m be an Op*-algebra on $V^{\#}$. Then m defines on $V^{\#}$ a locally convex topology t_m (called the m topology) by the family of seminorms $f \to \|Af\|$; $f \in V^{\#}$, $A \in m$. This topology is the coarsest locally convex topology on $V^{\#}$ such that every $A \in m$ is continuous from $V^{\#}[t_m]$ into $\mathscr{H}$ endowed with the usual Hilbert space norm topology.

Definition 2.10: The Op*-algebra m is said to be closed if $V^{\#}[t_m]$ is complete.

A comparison of the commutants m'_c, m'_σ, and m'_0 was done in Ref. 10 for closed Op*-algebra on $V^{\#}$ satisfying the condition I_0 (Ref. 4) (i.e., m contains a generating monotone increasing sequence $A_n \geqslant 1$ such that $A_n V^{\#} = V^{\#}$). In this case we obtain the equality $m'_c = m'_\sigma = m' = m'_0 = m'_{\breve{\sigma}}$.

If m is a ${}^{\times}$-invariant subset of Op V, then one may also define the following commutant:

$$m'_{\breve{c}} = \{X \in L^{+}(V^{\#}) \mid \langle Xf, A^{*}g\rangle = \langle Af, X^{*}g\rangle;$$
$$\forall f,g \in V^{\#} \text{ and } A \in m\}.$$

Following Proposition 2.9, one can easily prove that

$$m'_{\breve{c}} = m'_c \subseteq m' .$$

As pointed out in Ref. 10, in general, m' is not a *-subalgebra of Op V.

Proposition 2.11: Let m be an Op*-algebra on $V^{\#}$. Then m' is a *-subalgebra of Op V.

Proof: Let $X, Y \in m'$ and assume the product XY is defined.

We want to show that $XY \in m'$, i.e., for all $f,g \in V^{\#}$ and $A \in m$, the following relation holds:

$$\langle Af, (XY)^{*}g\rangle = \langle (XY)f, A^{*}g\rangle .$$

We have

$$\langle Af,(XY)^*g\rangle = \langle Af,Y^*X^*g\rangle = \langle YAf,X^*g\rangle = \langle AYf,X^*g\rangle$$
$$= \langle Yf,A^*X^*g\rangle = \langle Yf,X^*A^*g\rangle$$
$$= \langle (XY)f,A^*g\rangle, \text{ i.e., } XY\in m'.$$

(b) Bicommutants: In this paper we will be concerned with the following bicommutant.

Let m be a $^\times$-invariant subset of Op V. Then we define

$$m'' = \{Y\in \text{Op } V \mid Y\in M(m'), \quad YX = XY, \ \forall X\in m'\} .$$

This bicommutant is related to $m''_{c0} \equiv \{Y\in\text{Op } V \mid YX = XY, \ \forall X\in m'_c\}$ (cf. Ref. 10) in the following way:

$$m'' \subseteq m''_{c0} = (m'_c)' \subset \text{Op } V .$$

III. TOPOLOGICAL PROPERTIES OF THE COMMUTANT AND BICOMMUTANT

Since $m' \subset Lm \cap Rm$, it follows that one can consider on m' either of the four topologies defined on the spaces of multipliers of m. Clearly, since the elements of m and m' commute, the four topologies coincide on m' and we will write simply $t(m)$.

On m'' we will consider the topology $t(m')$ given by the seminorms,

$$t(m') \colon B\in m'' \mapsto |\langle (AB)f,g\rangle| = |\langle (BA)f,g\rangle| ,$$
$$\forall f,g\in V^\# \text{ and } A\in m' ,$$

and on $m \subseteq m''$, the $t(m')$-topology inherited from m''.

It is well known that for the algebra $B(\mathscr{H})$ of bounded operators, the usual commutant and bicommutant are closed in the weak (and *a fortiori* the strong) topology.[24]

Let m be a $^\times$-invariant subset of Op V. In this section, replacing the weak topology, respectively, by $t(m)$ and $t(m')$, we extend the above property of bounded commutant and bicommutant to m' and m'', respectively. Furthermore, we describe the relation between the commutant of m and that of its $t(m')$ closure and we show, under the assumption of reflexivity of the dual pair $\langle V^\#,V\rangle$, that m' is $t(m)$-sequentially complete.

Proposition 3.1: If m is a $^\times$-invariant subset with unit of Op V, then m' is closed in $M(m)$ with respect to the $t(m)$-topology.

Proof: Let $A\in$Op V be the limit of a $t(m)$ converging net $\{A_\alpha\}\subset m'$, i.e., for all $f,g\in V^\#$ and $B\in m$ we have

$$\lim_\alpha \langle BA_\alpha f,g\rangle = \lim_\alpha \langle A_\alpha Bf,g\rangle = \langle ABf,g\rangle ,$$

which implies that

$$\langle ABf,g\rangle = \lim_\alpha \langle A_\alpha Bf,g\rangle = \lim_\alpha \langle BA_\alpha f,g\rangle$$
$$= \langle BAf,g\rangle, \quad \text{i.e., } A\in m' .$$

Corollary 3.2: If m is a $^\times$-invariant subset of Op V, then m'' is closed in $M(m') \equiv Lm' \cap Rm'$ with respect to the $t(m')$-topology.

Proposition 3.3: If m is a $^\times$-invariant subset of Op V, then the commutant of m is equal to the commutant of its $t(m')$ closure, i.e.,

$$m' = (\overline{m}^{t(m')})' .$$

Proof: The inequality $(\overline{m}^{t(m')})' \subset m'$ follows from the fact that $m \subset \overline{m}^{t(m')}$. Let us now prove the opposite inclusion.

Let $B\in\overline{m}^{t(m')}$, i.e., there exists a net $\{B_\alpha\}\subset m$ such that $t(m')\text{-lim}_\alpha B_\alpha = B$. Let $X\in m'$, i.e., $XB_\alpha = B_\alpha X$. Then, for all $f,g\in V^\#$, we have

$$\langle XBf,g\rangle = \lim_\alpha \langle XB_\alpha f,g\rangle = \lim_\alpha \langle B_\alpha Xf,g\rangle = \langle BXf,g\rangle ,$$

i.e., $X\in(\overline{m}^{t(m')})'$.

Corollary 3.4: Let m_1 and m_2 be two $^\times$-invariant subsets of Op V, such that $m_1\subset m_2$ and m_1 is $t(m'_1)$ dense in m_2. Then $m'_1 = m'_2$.

Proof: The inclusion $m'_2 \subset m'_1$ follows from the fact that $m_1\subset m_2$. Now since m_1 is $t(m'_1)$ dense in m_2, i.e., $m_2\subset\overline{m}_1^{t(m'_1)}$ we get $(\overline{m}_1^{t(m'_1)})'\subset m'_2$. From the proposition 3.3, we know that $(\overline{m}_1^{t(m'_1)})' = m'_1$, which implies that $m'_1 \subset m'_2$, and hence the equality $m'_1 = m'_2$.

Proposition 3.5: Let m be a $^\times$-invariant subset with unit of Op V. If $\langle V^\#,V\rangle$ is a reflexive dual pair, then the commutant m' is $t(m)$-sequentially complete.

Proof: Let $\{X_n\}\subset m'$ be a $t(m)$-Cauchy sequence and consequently a $t(1)$-Cauchy sequence. Since Op V is $t(1)$-sequentially complete, for every $A\in m$, the following $t(1)$-limits exist in Op V:

$$X_n \to X ,$$
$$AX_n \to Ax ,$$
$$X_nA \to XA ,$$

which implies that $t(m)\text{-lim}_{n\to\infty} X_n = X$. Moreover $AX = XA$, which means that $X\in m'$.

IV. BICOMMUTANT AND THE $t(m')$ CLOSURE OF m

Since m'' is closed with respect to the $t(m')$-topology, the natural question to ask is whether it coincides with the $t(m')$ closure of m. In this section we give a sufficient condition which guarantees this result.

Following Ref. 25, we will say that a subspace W of a PIP space V is orthocomplemented in V, if W is the range of an orthogonal projection P, i.e., $W = PV$.

Proposition 4.1: Let m be an Op*-algebra on $V^\#$. If for all $f\in V^\#$ and $C\in m'$, the $\sigma(V,V^\#)$ closure $W = \overline{mCf}^\sigma = \overline{Cmf}^\sigma$ of mCf is orthocomplemented in V then $m'' = \overline{m}^{t(m')}$.

Proof: The inclusion $\overline{m}^{t(m'')}\subset m''$ follows from the fact that m'' is closed with respect to the $t(m')$ topology.

Let us now prove the opposite inclusion.

(a) Let $f\in V^\#$, $C\in m'$, and P_w be the orthogonal projection on $W = \overline{mCf}^\sigma$. Since m is an Op*-algebra it is $\sigma(V,V^\#)$ continuous and therefore it leaves W invariant, i.e., $P_w\in m'$.

Now take $Y\in m''$ and any $g\in V^\#$. Then

$$\langle f,C^*Y^*(1-P_w)g\rangle = \langle Cf,Y^*(1-P_w)g\rangle$$
$$= \langle Cf,(1-P_w)Y^*g\rangle$$
$$= \langle (1-P_w)Cf,Y^*g\rangle = 0 ,$$

i.e., $YCf = P_w YCf$.

We then conclude that given $Y \in m''$, for every $\epsilon > 0$, $f \in V^{\#}$ there exists $M \in m$ such that

$$|\langle (Y - M) Cf, h\rangle < \epsilon; \quad h \in V^{\#} .$$

Let us show that $m'' \subset \overline{m}^{t(m')}$. We recall that zero neighborhoods in the $t(m')$-topology are of the form

$$\mathscr{V}_{f_1,\ldots,f_n;h_1,\ldots,h_n;C_1,\ldots,C_n;\epsilon}(0)$$
$$= \{A \in M(m') \| \langle AC_1 f_1, h_1\rangle | < \epsilon, \ldots, |\langle AC_n f_n, h_n\rangle| < \epsilon\} ,$$

for any finite sequences $f_i, h_i \in V^{\#}$, $C_i \in m'$; $i = 1,\ldots,n$.

It is sufficient to prove that if $Y \in m''$, $f_i, h_i \in V^{\#}$, $C_i \in m'$, $i = 1,2$, and $\epsilon > 0$ then there exists $M \in m$ such that $(Y - M) \in \mathscr{V}_{f_1,f_2;h_1,h_2;C_1,C_2;\epsilon}(0)$.

For this consider the PIP space $V \oplus V$ with central Hilbert space $\mathscr{H} \oplus \mathscr{H}$, and the subalgebra $m \oplus m$ of $L^+(V^{\#} \oplus V^{\#})$. Every $M \in m$ gives rise to a regular operator

$$\widetilde{M} = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix} .$$

We denote by $\widetilde{m}$ the set of such operators, i.e.,

$$\widetilde{m} = \left\{ \widetilde{M} = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix} | M \in m \right\} .$$

We can compute explicitly the unbounded commutant and bicommutant of $\widetilde{m}$ and we get, respectively,

$$\widetilde{m}' = \left\{ X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} | X_{ij} \in m'; \quad i,j = 1,2 \right\} ,$$

$$\widetilde{m}'' = \left\{ Y = \begin{pmatrix} Y & 0 \\ 0 & Y \end{pmatrix} | Y \in m'' \right\} = \widetilde{m''} .$$

Now, applying the results of part (a) of this proof to $\widetilde{m}''$ we get that $\forall \widetilde{Y} \in \widetilde{m}''$, $\forall \widetilde{f}, \widetilde{h} V^{\#} \oplus V^{\#}$ and $\epsilon > 0$ there exists $\widetilde{M} \in \widetilde{m}$ such that

$$|\langle (\widetilde{Y} - \widetilde{M}) \subset f, \widetilde{h}\rangle| < \epsilon .$$

Thus is $\widetilde{f} = \binom{f_1}{f_2}$ and $\widetilde{h} = \binom{h_1}{h_2}$ it follows that

$$|\langle (Y - M) Cf_1, h_1\rangle| < \epsilon/2$$

and

$$|\langle (Y - M) Cf_2, h_2\rangle| < \epsilon/2 .$$

(c) In part (b) we have used the fact that the subspace $\overline{\widetilde{m}\binom{Cf_1}{Cf_2}}^{\sigma}$ is left invariant by $\widetilde{m}$. Now, since $C_1, C_2 \in m'$ implies $\begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix} \in \widetilde{m}'$ and $\overline{\widetilde{m}\binom{C_1 f_1}{C_2 f_2}}^{\sigma}$ is also invariant under $\widetilde{m}$, it follows that $\forall Y \in m''$ and $\forall \epsilon > 0$, there exists $M \in m$ such that

$$|\langle (Y - M) C_1 f_1, h_1\rangle| < \epsilon/2$$

and

$$|\langle (Y - M) C_2 f_2, h_2\rangle| < \epsilon/2 .$$

Thus $(Y - M) \in \mathscr{V}_{f_1,f_2;h_1,h_2;C_1C_2;\epsilon}(0)$. Since that is true for any neighborhood, we have that $Y \in \overline{m}^{t(m')}$.

V. CLOSED Op*-ALGEBRAS AND THE ARAKI–JURZAK BICOMMUTANT

In Ref. 10 it was shown that if m is a closed Op*-algebra satisfying the condition I_0, then $m'_A = m'$. Actually, in this case all the commutants considered in this paper coincide, i.e.,

$$m'_A = m'_c = m'_\sigma = m'_0 = m'_{\bar{\sigma}} = m' \subset L^+(V) .$$

In this section, following the strategy of Ref. 10, we study the relationship between the bicommutants m''_A and m''.

Let m be a closed Op*-algebra on $V^{\#}$ and denote by $B(V^{\#}[t_m], V^{\#}[t_m])$ the set of all sesquilinear forms which are jointly continuous in the m-topology, i.e., for all $\beta \in B(V^{\#}[t_m], V^{\#}[t_m])$ there exists an $A \in m$ such that for some constant M and all $f,g \in V^{\#}$, we have

$$|\beta(f,g)| \leqslant M \|Af\| \, \|Ag\| .$$

One may define the following commutant and bicommutant[4]:

$$m'_A = \{\beta \in B(V^{\#}[t_m], V^{\#}[t_m]) | \beta(Af,g) = \beta(f, A^* g);$$
$$\forall f,g \in V^{\#} \text{ and } A \in m\} ,$$

$$m''_{AA} = \{\gamma \in B(V^{\#}[t_{m'_A}], V^{\#}[t_{m'_A}]) | \gamma(Cf,g) = \gamma(f, C^* g);$$
$$\forall f,g \in V^{\#} \text{ and } C \in m'_A\} .$$

Proposition 5.1 (Ref. 4): If m is a closed Op*-algebra on $V^{\#}$ satisfying the condition I_0, then

(i) m'_A is an Op*-algebra on $V^{\#}$ satisfying I_0, but m'_A is not closed (which implies that in general m''_{AA} is not an Op*-algebra).

(ii) The m topology is metrizable; it is given by the seminorms,

$$f \to \|A_n f\|; \quad n \in N, \quad f \in V^{\#} .$$

Let m be a closed Op*-algebra on $V^{\#}$ satisfying I_0. First of all we know that m''_{AA} is contained in $B(V^{\#}[t_{m'_A}], V^{\#}[t_{m'_A}])$, whereas m'' belongs to Op V which is isomorphic to the space $\widetilde{B}(V^{\#}[\tau], V^{\#}[\tau])$ of all Mackey separately continuous sesquilinear forms on $V^{\#} \times V^{\#}$.[13] Therefore, m''_{AA} will coincide with m'' if, in particular

$$B(V^{\#}[t_{m'_A}], V^{\#}[t_{m'_A}]) \equiv \widetilde{B}(V^{\#}[\tau], V^{\#}[\tau]) .$$

Since the PIP space V posesses a central Hilbert space $\mathscr{H}$, the topologies $t_{m'_A}$ and $\tau(V^{\#}, V)$ are comparable.

In general, the Mackey topology is strictly finer than the m'_A-topology and we have the following situation (where $V_{m'_A}$—the dual of $V^{\#}[t_{m'_A}]$ and V_τ—the dual of $V^{\#}[\tau]$ are equipped with their Mackey topologies and each arrow denotes a continuous embedding with dense range):

$$V^{\#}[\tau] \hookrightarrow V^{\#}[t_{m'_A}] \hookrightarrow \mathscr{H} \hookrightarrow V_{m'_A} \hookrightarrow V_\tau .$$

According to Proposition 5.1, the m'_A-topology is metrizable and this implies that (Ref. 26, Proposition 36.3) on the incomplete space $V^{\#}$ the $t_{m'_A}$-topology coincides with the Mackey topology $\tau(V^{\#}, V)$.

However, an element of $\widetilde{B}(V^{\#}[\tau], V^{\#}[\tau])$ need not be jointly continuous. Therefore we only have the function

$$\widetilde{B}(V^{\#}[\tau], V^{\#}[\tau]) \subset B(V^{\#}[t_{m'_A}], V^{\#}[t_{m'_A}])$$

which implies that $m'' \subset m''_{AA}$. We summarize this analysis in the following proposition.

Proposition 5.2: If m is a closed Op*-algebra satisfying the condition I_0, then $m'' \subset m''_{AA}$.

ACKNOWLEDGMENTS

This work was begun during a stay at the University of Palermo.

It is a pleasure to thank Professor J-P. Antoine, Professor G. Epifanio, Dr. F. Mathot, and Dr. C. Trapani for many fruitful discussions. I should also like to thank Professor Abdus Salam, the International Atomic Energy Agency, and the UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste, where this work was completed.

I gratefully acknowledge the kind hospitality offered to me at the Istituto di Fisica, as well as financial support from AGCD and FNRS (Belgium) and CNR (Italy).

[1] R. T. Powers, Commun. Math. Phys. **21**, 85 (1971).

[2] A. Gudder and W. Scruggs, Pac. J. Math. **70**, 369 (1977).

[3] A. Inoue, Proc. Am. Math. Soc. **69**, 97 (1978).

[4] H. Araki and J-P. Jurzak, Publ. RIMS, Kyoto Univ. **18**, 1013 (1982).

[5] G. Epifanio and C. Trapani, J. Math. Phys. **25**, 2633 (1984).

[6] A. Voronin, V. N. Sushko, and S. S. Khoruzhii, Theor. Math. Phys. **59**, 335 (1984); **60**, 849 (1984).

[7] K. Schmüdgen, "Unbounded commutants and intertwining spaces of unbounded symmetric operators and *-representations," Preprint Karl-Marx-Universität, Leipzig, 1985.

[8] F. Mathot, J. Math. Phys. **26**, 1118 (1985).

[9] J-P. Antoine, F. Mathot, and C. Trapani, Ann. Inst. H. Poincaré **46**, 325 (1987).

[10] J. Shabani, J. Math. Phys. **25**, 3204 (1984).

[11] J. Shabani, J. Math. Phys. **28**, 2046 (1987).

[12] J-P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 369 (1976).

[13] J-P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 379 (1976).

[14] J-P. Antoine, J. Math. Phys. **21**, 268 (1980).

[15] J-P. Antoine, J. Math. Phys. **21**, 2067 (1980).

[16] H-J. Borchers, in RCP 25 (Strasbourg) **22**, 26 (1975); Acta Phys. Austriaca Suppl. **16**, 15 (1976).

[17] J-P. Antoine and W. Karwowski, Publ. RIMS, Kyoto Univ. **21**, 205 (1985); **22**, 507 (A) (1986).

[18] J-P. Antoine and F. Mathot, Ann. Inst. H. Poincaré **46**, 299 (1987).

[19] J-P. Antoine and F. Mathot, Ann. Inst. H. Poincaré **37**, 29 (1982).

[20] G. Lassner, Rep. Math. Phys. **3**, 279 (1972).

[21] M. Friedrich and G. Lassner, Wiss. Z. Karl-Marx-Univ. Leipzig Math. Natur. Reihe **R27**, 245 (1978).

[22] K-D. Kürsten, "On topological properties of domains of unbounded operator algebras," in *Proceedings of the II International Conference on Operator Algebras, Ideals and their Applications in Theoretical Physics*, Leipzig, 1983, edited by H. Baumgärtel, G. Lassner, A. Pietsch, and A. Uhlmann (Teubner, Leipzig, 1984).

[23] G. Köthe, *Topological Vector Spaces I* (Springer, Berlin, 1969).

[24] J. Dixmier, *les Algèbres d'opérateurs dans l'espace Hilbertien* (Gauthier-Villars, Paris, 1962).

[25] J-P. Antoine and A. Grossmann, J. Math. Phys. **19**, 329 (1978).

[26] F. Treves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967).

Semilinear operators

C. S. Sharma and D. F. Almeida
Mathematical Research Unit, Birkbeck College, University of London, 43 Gordon Square, London WC1H OPD, England

(Received 26 February 1988; accepted for publication 20 July 1988)

Semilinear operators on a complex Hilbert space are studied in a part of a program that aims to develop the theories of additive operators on complex and quaternionic Hilbert spaces for application to problems in mathematical physics. The more notable among the new results proved on the eigenvalue problem for semilinear operators are the following: (i) if α is an eigenvalue of a semilinear operator then so also is any complex number which has the same modulus as α; (ii) if a normal semilinear operator has two eigenvectors belonging to different eigenvalues, then either the two eigenvectors are orthogonal or two eigenvalues have the same moduli; and (iii) a normal semilinear operator has a complete set of eigenvectors if and only if it is self-adjoint. Further, it is shown that there exists a norm-preserving semilinear isomorphism between the spaces of bounded linear and semilinear operators on a complex Hilbert space. Finally it is demonstrated how the theory of semilinear operators can be exploited to solve the problems of finding three involutive mutually anticommuting self-adjoint two-by-two matrices and four four-by-four matrices with the same properties: the unusual and remarkably easy solution of this old familiar exercise establishes the relevance of the theory being developed here to physics.

I. INTRODUCTION

A semilinear transformation was first defined by Segrè[1] about 100 years ago. According to Segrè's definition, a semilinear transformation is a pair $\sigma = (\sigma', \sigma'')$ of mappings between linear spaces $\mathscr{V}$ and $\mathscr{W}$ over the fields F and $\mathbb{G}$, respectively, where σ' is an isomorphism between the additive group $\mathscr{V}$ and the additive group $\mathscr{W}$, and σ'' is an isomorphism between fields $\mathbb{F}$ and $\mathbb{G}$ subject to the condition

$$\sigma(\alpha u) = \sigma''(\alpha)\sigma'(u), \quad \forall \alpha \in \mathbb{F} \text{ and } \forall u \in \mathscr{V}. \qquad (1.1)$$

The concept was used to obtain a number of useful results in projective geometry culminating in the first fundamental theorem of projective geometry.[2]

Jacobson[3] modified the definition so that σ' was merely a homomorphism between the additive groups in the two linear spaces and used the concept to obtain a penetrating result relating isomorphisms of rings of linear transformations on vector spaces to isomorphisms of the vector spaces—a result that is of fundamental importance in the representation theory of a simple ring. According to Jacobson's definition both linear and antilinear maps between complex vector spaces are semilinear.

For maps between complex vector spaces, many later authors (see, for example, Lang[4]) used the term semilinear to mean the same thing as what physicists call antilinear or conjugate linear. In this work, as in our earlier works,[5–13] we use the term in the sense described in the preceding sentence, that is, in a sense synonymous with antilinear and conjugate linear. It should be noted that maps semilinear according to our definition are also semilinear according to Jacobson's definition, but maps semilinear according to Jacobson's definition[3] need not be semilinear according to our definition.

Our original interest in semilinearity arose when we observed that though practically every branch of modern mathematics was used in quantum theory, the functional calculus used in quantum theory was somewhat primitive and messy. We identified that at the root of this malady lies the fact that while modern calculus on Banach space is lopsided in the sense that it relies too much on linearity, all the most important functionals in quantum theory have both linearity and semilinearity in equal amounts, which in turn is a consequence of the fact that all observed values must be real. With an aim to rectify this situation, in Ref. 6 we developed a new calculus on Banach space in which the novelty was that we abandoned the requirement that the derivative must be a linear map in favor of the requirement that it must be a direct sum of a linear and a semilinear map. When the derivative is bounded this is equivalent to abandoning linearity in favor of additivity.[9] A group of Italian physicists have developed our calculus further and published five more papers[14–18] on it. Even though over a dozen papers now exist on this calculus, it is still in its infancy and only the simplest problems in the calculus of variations on a complex Hilbert space have been solved by this calculus. Both the Italian group and ourselves are engaged in further work on the subject which will be reported in due course.

Semilinearity in quantum theory makes its first appearance in the Riesz representation theorem: the isomorphism between a complex Hilbert space and its dual is semilinear. As is well-known, Riesz representation theorem lies at the heart of the spectacularly successful Dirac formalism. Though the overwhelming majority of operators appearing in quantum theory are linear, at least two, namely the time reversal operator and the charge conjugation operator, both of which are of the greatest fundamental importance, are semilinear. The time reversal operator was introduced into quantum theory by Wigner.[19,20] He also proved[19] that a Hilbert space operator which preserves the modulus of the inner

0022-2488/88/112411-10$02.50

product must be either unitary or semiunitary (antiunitary). A related result, which asserts that a transformation that preserves convex combinations of quantum mechanical states (in this context a state is a non-negative self-adjoint operator of unit trace and a convex combination is a linear combination with positive coefficients) is either unitary or semiunitary, was proved by Kadison.[21] Yet the literature on semilinear operators is sketchy and unsatisfactory: the most comprehensive account, to the best of our knowledge, is to be found in the admirable book by Messiah.[22] More recently several attempts have been made to develop the theory of an operator algebra on a quaternionic Hilbert space with the aim of finding a more satisfactory description of quantum theory: the two more notable studies are by Horwitz and Biedenharn[23] and Adler.[24] In trying to unravel some of the complexities (or maybe hypercomplexities) of operators on a quaternionic Hilbert space, we observed[12,13] that semilinearity plays a fundamental role in the study of such operators. We then used the ideas developed in Refs. 12 and 13 to prove[25] an important theorem on the algebra of additive operators on a complex Hilbert space.

The term algebra is used by algebraists to describe a ring that is also a vector space. If used in this sense, the term algebra cannot be used to describe the collection of bounded linear operators on a quaternionic Hilbert space or the collection of bounded semilinear operators on a complex Hilbert space. We observed[12,13,25] that the smallest algebras containing these collections are the algebras of additive operators on a quaternionic and a complex Hilbert space, respectively. Our current interest in semilinear operators owes its origin to our desire to develop a rigorous mathematical theory of the algebras of bounded additive operators on complex and quaternionic Hilbert spaces. In this paper we consider only semilinear operators on a complex Hilbert space and in our next paper[26] we consider additive functionals and operators on a quaternionic Hilbert space using similar methods.

In Sec. II we develop the definitions and the notations that we use in this work, in Sec. III we collect all the results proved in earlier works that are relevant to our present study, in Sec. IV we present all the new results of the present work, in Sec. V we demonstrate the relevance of our study to the theory of spinors, and in Sec. VI we conclude with a few closing remarks.

II. FORMALITIES

We denote the field of real and complex numbers by $\mathbb{R}$ and $\mathbb{C}$, respectively, and the skew field of quaternionic numbers by $\mathbb{H}$.

Let $\mathscr{H}$ be a vector space over $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. We define a positive definite Hermitian form on $\mathscr{H}$ by

$$\langle\ ,\ \rangle: \mathscr{H} \times \mathscr{H} \rightarrow \mathbb{F},$$

$$\langle \alpha u, \beta v\rangle = \alpha \langle u,v\rangle \beta^*, \tag{2.1}$$

$$\langle u+v, w\rangle = \langle u,w\rangle + \langle v,w\rangle, \tag{2.2}$$

$$\langle u,v\rangle^* = \langle v,u\rangle, \tag{2.3}$$

$$\langle u,u\rangle = 0, \quad \text{only if } u = 0, \tag{2.4}$$

where $\alpha^* = \alpha$ if $\mathbb{F}$ is real, $\alpha^* =$ complex conjugate of α if $\mathbb{F}$ is complex, and $\alpha^* =$ quaternionic conjugate of α if $\mathbb{F}$ is quaternionic.

Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be Hilbert spaces over $\mathbb{F}$. We say that a map $L: \mathscr{H}_1 \rightarrow \mathscr{H}_2$ is *additive* if and only if for all $u,v \in \mathscr{H}_1$,

$$L(u+v) = L(u) + L(v). \tag{2.5}$$

Thus an additive map is a homomorphism of the additive group in a vector space. If, in addition, the additive map L satisfies

$$L(\alpha u) = \alpha L(u), \tag{2.6}$$

for all $\alpha \in \mathbb{F}$ and all $u \in \mathscr{H}_1$, then it is called *linear*. If, on the other hand, the additive map L satisfies

$$L(\alpha u) = \alpha^* L(u), \tag{2.7}$$

for all $\alpha \in \mathbb{F}$ and all $u \in \mathscr{H}_1$, then it is called *semilinear* or *antilinear*. It was actually shown in Ref. 13 that for a quaternionic Hilbert space a semilinear map defined in this way leads to a contradiction and, therefore, does not exist. There, given any particular choice of i, j, and k, it is necessary to define three different kinds of semilinearities called i, j, and k semilinearity (see Ref. 13). In this work we do not propose to deal with the quaternionic case and so we do not define these semilinearities. For our present purposes we concentrate our attention to the complex case. From now on, unless otherwise stated, $\mathscr{H}$, $\mathscr{H}_1$, $\mathscr{H}_2$ denote Hilbert spaces over $\mathbb{C}$.

An additive map $L: \mathscr{H}_1 \rightarrow \mathscr{H}_2$ is said to be *prelinear* if and only if

$$\|L(\alpha u)\| = |\alpha|\ \|Lu\| \tag{2.8}$$

for all $\alpha \in \mathbb{F}$ and all $u \in \mathscr{H}_1$. Note that both linear and semilinear maps are prelinear.

The collection of all bounded linear maps from $\mathscr{H}$ to $\mathbb{C}$ is a complex vector space that is called the *dual* of $\mathscr{H}$ and is denoted by $\tilde{\mathscr{H}}$.

The collection of all bounded semilinear maps from $\mathscr{H}$ to $\mathbb{C}$ is a complex vector space that is called the *semidual* of $\mathscr{H}$ and is denoted by $\tilde{\mathscr{H}}_s$.

The collection of all bounded additive maps from $\mathscr{H}$ to $\mathbb{C}$ is a complex vector space that is called the *additive dual* or *addual* (for short) of $\mathscr{H}$ and is denoted by $\tilde{\mathscr{H}}_a$.

The collection of all bounded linear maps from $\mathscr{H}_1$ to $\mathscr{H}_2$ is a complex vector space that is denoted by $\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$. Note that

$$\mathscr{L}(\mathscr{H},\mathbb{C}) = \tilde{\mathscr{H}}. \tag{2.9}$$

The collection of all bounded semilinear maps from $\mathscr{H}_1$ to $\mathscr{H}_2$ is a complex vector space that is denoted by $\mathscr{S}\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$. Note that

$$\mathscr{S}\mathscr{L}(\mathscr{H},\mathbb{C}) = \tilde{\mathscr{H}}_s. \tag{2.10}$$

The collection of all bounded additive maps from $\mathscr{H}_1$ to $\mathscr{H}_2$ is a complex vector space that is denoted by $\mathscr{A}(\mathscr{H}_1,\mathscr{H}_2)$. Note that

$$\mathscr{A}(\mathscr{H},\mathbb{C}) = \tilde{\mathscr{H}}_a. \tag{2.11}$$

The reader will have noticed that our convention is that vector spaces are denoted by letters in capital script, maps by capital italics, vectors by lowercase italic letters, and scalars by lowercase Greek letters. With these conventions writing

Lu for $L(u)$ will not give rise to ambiguity and in the future we shall do so.

Let A be either a linear or a semilinear map from $\mathscr{H}_1$ to $\mathscr{H}_2$. A norm of A denoted by $\|A\|$ is defined by the formula

$$\|A\| = \sup_{\|x\|=1} \|Ax\|. \tag{2.12}$$

This definition turns $\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$ and $\mathscr{S}\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$ into normed spaces.

Let $\mathscr{N}_1 \oplus \mathscr{N}_2$ be the direct sum of two normed spaces $\mathscr{N}_1$ and $\mathscr{N}_2$. An element of $\mathscr{N}_1 \oplus \mathscr{N}_2$ is a pair (x_1,x_2) with $x_1 \in \mathscr{N}_1$ and $x_2 \in \mathscr{N}_2$. A norm of (x_1,x_2) denoted by $\|(x_1,x_2)\|$ is defined by the formula

$$\|(x_1,x_2)\| = \|x_1\| + \|x_2\|. \tag{2.13}$$

We shall now define adjoints. These will be defined for two kinds of maps: (i) maps between different spaces and (ii) maps from a space to itself—such maps will be described as *operators*. Case (ii) is a special case of (i), but the definitions of adjoints look rather different, at least superficially. In what follows when a map is regarded as belonging to case (i) it will carry a hat and operators will be hatless.

Let $\hat{A}$: $\mathscr{H}_1 \to \mathscr{H}_2$ be a linear map. The adjoint $\hat{A}^*$: $\tilde{\mathscr{H}}_2 \to \tilde{\mathscr{H}}_1$ is a linear map defined by the property

$$f(\hat{A}u) = (\hat{A}^* f)(u), \tag{2.14}$$

for all $u \in \mathscr{H}_1$ and all $f \in \tilde{\mathscr{H}}_2$.

Let $\hat{A}$: $\mathscr{H}_1 \to \mathscr{H}_2$ be an additive map. The adjoint $\hat{A}^*_a$: $\tilde{\mathscr{H}}_{2a} \to \tilde{\mathscr{H}}_{1a}$ is a linear map defined by the property

$$f(\hat{A}u) = (\hat{A}^*_a f)(u), \tag{2.15}$$

for all $u \in \mathscr{H}_1$ and all $f \in \tilde{\mathscr{H}}_{2a}$.

Let A: $\mathscr{H} \to \mathscr{H}$ be a bounded linear operator. The adjoint A^*: $\mathscr{H} \to \mathscr{H}$ is a bounded linear operator with the property that

$$\langle Au,v\rangle = \langle u,A^*v\rangle, \tag{2.16}$$

for all $u,v \in \mathscr{H}$.

Let A: $\mathscr{H} \to \mathscr{H}$ be a bounded semilinear operator. The adjoint A^*: $\mathscr{H} \to \mathscr{H}$ is a bounded semilinear operator with the property that

$$\langle Au,v\rangle = \langle A^*v,u\rangle, \tag{2.17}$$

for all $u,v \in \mathscr{H}$.

Since an operator A can be regarded as a special case of a map between two spaces, we have here two kinds of adjoints for A: A^*, which is an operator on the same space, and $\hat{A}^*$, which is an operator on the dual of the space on which A acts. Note that there is no definition of an adjoint of a semilinear map corresponding to the definition (2.14) for linear maps. An easy computation shows that an attempt to define the adjoint of a semilinear map as a semilinear map with a property analogous to (2.14) leads to disaster. However, a semilinear map is also additive and definition (2.15) is a generalization of (2.14) and works equally well for linear, semilinear, and additive maps, but the adjoints thus defined are all linear. The linear adjoint defined in this way for a semilinear operator was reconciled with the semilinear adjoint defined through (2.17) in Ref. 13.

We now have a further collection of useful definitions, where we use the not uncommon convention that unless otherwise stated the domain of an operator is the whole of $\mathscr{H}$ and indeed in this work we do not use a single operator whose domain is not the whole space.

Let A be a linear or a semilinear operator on $\mathscr{H}$. A subspace $\mathscr{M}$ of $\mathscr{H}$ is said to be *invariant* under A if $A(\mathscr{M}) \subset \mathscr{M}$, or in other words if $u \in \mathscr{M}$ implies $Au \in \mathscr{M}$.

Let A be a linear or a semilinear operator on $\mathscr{H}$ and let $\mathscr{M}$ be a subspace of $\mathscr{H}$ with the property that both $\mathscr{M}$ and $\mathscr{M}^\perp$ are invariant under A, then $\mathscr{M}$ is said to *reduce A*.

Spaces of bounded linear, semilinear, and additive operators on $\mathscr{H}$ will be denoted by $\mathscr{L}(\mathscr{H})$, $\mathscr{S}\mathscr{L}(\mathscr{H})$, and $\mathscr{A}(\mathscr{H})$, respectively.

A bounded linear or semilinear operator A on $\mathscr{H}$ is called *normal* if and only if

$$AA^* = A^*A. \tag{2.18}$$

A bounded linear or semilinear operator A on $\mathscr{H}$ is called *self-adjoint* if and only if

$$A = A^*. \tag{2.19}$$

A bounded normal operator U on $\mathscr{H}$ satisfying

$$UU^* = I, \tag{2.20}$$

where I is the identity map on $\mathscr{H}$ and is called (i) *unitary* if it is linear or (ii) *semiunitary* if it is semilinear.

Note that it follows from their definitions that self-adjoint (whether linear or semilinear), unitary, and semiunitary operators are all normal.

Let $\mathscr{V}$ be a vector space on a field **F**. Let a product be defined on $\mathscr{V}$ in such a way that it turns $\mathscr{V}$ into a ring, then $\mathscr{V}$ with this additional operation is called an *algebra*. If such an algebra in addition satisfies the property that

$$(\alpha u)(\beta v) = \alpha\beta(uv), \tag{2.21}$$

then it is called a *K algebra*. It turns out that $\mathscr{L}(\mathscr{H})$ with the product defined by composition of maps is a K algebra, $\mathscr{A}(\mathscr{H})$ with the product defined in the same way is an algebra but not a K algebra, and $\mathscr{S}\mathscr{L}(\mathscr{H})$ with the product defined by composition of maps does not satisfy the requirements of being an algebra.

Let A be an operator or a matrix. Then A is said to be *involutive*[27] if and only if

$$A^2 = I. \tag{2.22}$$

Let $\mathscr{A}$ be an algebra. An *involution* $*$ on $\mathscr{A}$ is an involutive operator on $\mathscr{A}$ that takes A to A^* and satisfies the following properties: (i) $*$ is a homomorphism of the additive group in the algebra, that is,

$$(A+B)^* = A^* + B^*, \tag{2.23}$$

for all $A,B \in \mathscr{A}$; (ii) $*$ is product reversing, that is,

$$(AB)^* = B^*A^*, \tag{2.24}$$

and being involutive, of course, means that it satisfies

$$A^{**} = A, \tag{2.25}$$

for all $A \in \mathscr{A}$. This definition is a generalization of the one given by Rudin.[28] Here, unlike Rudin, we do not require $*$ to be semilinear but we require it to be additive.

Let A be an operator on a complex Hilbert space $\mathscr{H}$. Then a nonzero vector v satisfying

$$Av = \alpha v, \tag{2.26}$$

for some complex number α, is called an eigenvector of A. Eigenvectors of A belonging to the eigenvalue $+1$ are called *fixed points* of A.

III. SOME KNOWN RESULTS

In this section we state without proof results that are known and are relevant to the further development of our theory. We, however, give references to the works where the proofs can be found except in cases where the proof is immediately obvious.

Proposition 3.1: The product (by composition) of two linear operators or two semilinear operators is linear and the product of a linear operator and a semilinear operator is semilinear.

Proof: Obvious.

Proposition 3.2: Let A be a semilinear operator on a complex Hilbert space $\mathscr{H}$. Then the following statements about A are equivalent: (a) A is continuous at a point $x_0 \in \mathscr{H}$; (b) A is bounded; (c) A is continuous at every point $x \in \mathscr{H}$.

Proof: See Ref. 6 after noting that both semilinear and linear operators are prelinear and that $\mathscr{H}$ is a normed space.

Remark: In making use of Ref. 6 note that Proposition [P2] there is false in that the space of prelinear operators is not a vector space because the sum of a linear operator that takes the vector αu to αv and the semilinear operator that takes the same vector to $\alpha^* v$ is clearly not prelinear. However, for each of the spaces $\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$, $\mathscr{S}\mathscr{L}(\mathscr{H}_1,\mathscr{H}_2)$, and $\mathscr{A}(\mathscr{H}_1,\mathscr{H}_2)$ Proposition [P2] in Ref. 6 appropriately modified is valid with the proof given there being correct.

Proposition 3.3: The space of bounded additive operators on a complex Hilbert space $\mathscr{H}$ is the direct sum of spaces of bounded linear and semilinear operators on $\mathscr{H}$.

Proof: See Ref. 9.

Proposition 3.4: Let y be any element of a complex Hilbert space $\mathscr{H}$. Let Φ_y be the semilinear functional on $\mathscr{H}$ defined by

$$\Phi_y(x) = \langle y,x \rangle. \tag{3.1}$$

The correspondence $y \mapsto \Phi_y$ is a norm-preserving linear isomorphism from $\mathscr{H}$ to $\tilde{\mathscr{H}}_s$.

Proof: See Ref. 13.

Proposition 3.5: If the semidual is identified with the original space by the isomorphism of Proposition 3.4 and the dual is identified with the original space by a similar but semilinear isomorphism, then the adjoint of a semilinear map defined by (2.15) becomes identical with that defined by (2.17).

Proof: See Ref. 13.

Proposition 3.6: The algebra of bounded additive operators on a complex Hilbert space $\mathscr{H}$ is the smallest algebra, that is, a vector space in which a ring structure is defined on the set of vectors, containing both linear and semilinear bounded operators on $\mathscr{H}$. Furthermore, the algebra is normed and the correspondence between an operator and its adjoint is a norm-preserving involution on this algebra.

Proof: See Ref. 25.

IV. SOME NEW RESULTS

The end of a proof is marked by a □.

Proposition 4.1: Let A be a semilinear operator on a complex Hilbert space $\mathscr{H}$. Let u be an eigenvector of A belonging to the eigenvalue $\alpha \in \mathbb{C}$. Then every complex number γ satisfying

$$|\gamma| = |\alpha| \tag{4.1}$$

is also an eigenvalue of A and every vector v in the one-dimensional subspace spanned by u is also an eigenvector belonging to some eigenvalue satisfying (4.1).

Proof: Every γ satisfying (4.1) can be written as

$$\gamma = |\alpha| e^{i\theta}, \tag{4.2}$$

with $\theta \in [0,2\pi[$, and in particular α itself can be written as

$$\alpha = |\alpha| e^{i\phi}, \tag{4.3}$$

for some $\phi \in [0,2\pi[$.

Now

$$Au = \alpha u = |\alpha| e^{i\phi} u, \tag{4.3'}$$

hence

$$Ae^{-i(\theta-\phi)/2}u = e^{i(\theta-\phi)/2}Au = e^{i(\theta-\phi)/2}|\alpha|e^{i\phi}u = |\alpha|e^{i\theta}e^{-i(\theta-\phi)/2}u. \tag{4.4}$$

Thus $\gamma = |\alpha|e^{i\theta}$ is an eigenvalue.

If v belongs to the one-dimensional subspace spanned by u, then

$$v = \xi u = |\xi| e^{i\rho} u, \tag{4.5}$$

for some $\rho \in [0,2\pi[$. Then

$$\begin{aligned} Av = |\xi| e^{-i\rho} Au &= |\xi|\,|\alpha| e^{-i(\rho-\phi)} u \\ &= |\alpha| e^{-i(2\rho-\phi)} |\xi| e^{i\rho} u \\ &= |\alpha| e^{-i(2\rho-\phi)} v, \end{aligned} \tag{4.6}$$

which shows that v is an eigenvector of A belonging to the eigenvalue $|\alpha|e^{-i(2\rho-\phi)}$, which, of course, satisfies (4.1). □

Remark: An eigenvector as in the linear case spans an eigenspace $\mathscr{E}$, but for semilinear operators $\mathscr{E}$ does not correspond to a fixed eigenvalue, rather it is characterized by an eigencircle in the complex plane. Each vector in $\mathscr{E}$ is an eigenvector belonging to an eigenvalue somewhere on the circle and each point on the eigencircle has infinitely many eigenvectors in $\mathscr{E}$ (we find one such eigenvector v by Proposition 4.1 and then note that ξv with ξ real belongs to the same eigenvalue). Here we have a forerunner of what happens in the quaternionic case[12,29] and here, too, we can define an equivalence class of eigenvalues and call it an eigenclass. However, as was shown in Ref. 12, defining such an equivalence class does not seem to serve any very useful purpose.

Proposition 4.2: Let A be a semilinear operator on a complex Hilbert space $\mathscr{H}$. Let $\mathscr{M}$ be a subspace of $\mathscr{H}$ invariant under A. Then $\mathscr{M}^\perp$ is invariant under A^*.

Proof: Let $u \in \mathscr{M}$ and let $v \in \mathscr{M}^\perp$. Then invariance of $\mathscr{M}$ under A implies that $Au \in \mathscr{M}$. Thus

$$\langle Au,v \rangle = 0 = \langle A^*v,u \rangle, \tag{4.7}$$

for every $u \in \mathscr{M}$, which shows that $A^*v \in \mathscr{M}^\perp$. □

Corollary 4.2.1: Let A be a semilinear operator on a complex Hilbert space $\mathscr{H}$. Let $\mathscr{M}$ be a subspace of $\mathscr{H}$ invariant under both A and A^*. Then $\mathscr{M}$ reduces A.

Proof: It is given that $\mathscr{M}$ is invariant under both A and A^* and invariance of $\mathscr{M}$ under A^* implies by the main Proposition that $\mathscr{M}^\perp$ is invariant under A. □

Proposition 4.3: Let A be a normal semilinear operator on a complex Hilbert space $\mathscr{H}$. Let $\mathscr{M}$ be a subspace of $\mathscr{H}$ spanned by the eigenvectors of A. Then $\mathscr{M}$ reduces A and the restriction of A to $\mathscr{M}$ is self-adjoint.

Proof: We shall first show that if u is an eigenvector of A belonging to the eigenvalue α, then u is also an eigenvector of A^* belonging to the same eigenvalue. This is done by computing $\|(A^* - \alpha)u\|^2$ as follows:

$$\begin{aligned}\langle (A^* - \alpha)u, (A^* - \alpha)u\rangle &= \langle A^*u, A^*u\rangle - \alpha\langle u, A^*u\rangle \\ &\quad - \alpha^*\langle A^*u, u\rangle + \alpha\alpha^*\langle u,u\rangle \\ &= \langle AA^*u,u\rangle - \alpha\langle u,Au\rangle - \alpha^*\langle Au,u\rangle + \alpha\alpha^*\langle u,u\rangle \\ &= \langle A^*Au,u\rangle - \alpha\langle u,Au\rangle - \alpha^*\langle Au,u\rangle + \alpha\alpha^*\langle u,u\rangle \\ &= \langle Au,Au\rangle - \alpha\langle u,Au\rangle - \alpha^*\langle Au,u\rangle + \alpha\alpha^*\langle u,u\rangle \\ &= \langle (A-\alpha)u,(A-\alpha)u\rangle = 0. \end{aligned} \tag{4.8}$$

Hence

$$(A^* - \alpha)u = 0 \tag{4.9}$$

and u is an eigenvector of A^* belonging to the same eigenvalue α. Thus in $\mathscr{M}$, A and A^* have the same action on members of a basis and therefore on the whole subspace and thus the restriction to $\mathscr{M}$ of both A and A^* are identical. Finally it is evident that $\mathscr{M}$ is invariant under both A and A^* and therefore by Corollary 1 $\mathscr{M}$ reduces A (and also A^*). □

Corollary 4.3.1: Let A be a normal semilinear operator on a complex Hilbert space $\mathscr{H}$. If there exists a basis in $\mathscr{H}$ each of whose members is an eigenvector of A, then A is self-adjoint.

Proof: The proof follows from Proposition 4.3 by noting that the span of the eigenvectors is the whole space $\mathscr{H}$. □

Remark: In a one-dimensional space every semilinear operator is a scalar multiple of complex conjugation and therefore has an eigenvector. Since the span of its eigenvectors reduces A in an n-dimensional space if it has $(n-1)$ linearly independent eigenvectors, then it must also have n linearly independent eigenvectors. These facts can be summarized in the following corollary.

Corollary 4.3.2: Let A be a normal semilinear operator on an n-dimensional complex Hilbert space $\mathscr{H}$. Then A is either self-adjoint or the number of linearly independent eigenvectors of A does not exceed $(n-2)$.

Remark: Proposition 4.3 shows that every normal semilinear operator A on $\mathscr{H}$ has a decomposition $A = B + C$ in which B and C have invariant essential supports (the essential support of an operator is the orthogonal complement of its kernel) in orthogonal subspaces and B is self-adjoint. This may suggest that C is anti-self-adjoint, that is, $C = -C^*$, but the example constructed in Proposition 4.6 shows that this is not necessarily true.

Proposition 4.4: Let A be a semilinear self-adjoint operator on a finite dimensional Hilbert space $\mathscr{H}$. Then there exists at least one eigenvector of A in $\mathscr{H}$.

Proof: Here A is semilinear and self-adjoint implies that A^2 is linear and self-adjoint. Hence by the spectral theorem for linear self-adjoint operators A^2 has a complete set of eigenvectors and in particular it has an eigenvector u belonging to a real eigenvalue γ. We consider the following possibilities.

(i) $\gamma = 0$. Then it is easy to see that u is an eigenvector of A belonging to the eigenvalue 0.

(ii) $\gamma \neq 0$. Then an easy calculation shows that either iu or $\sqrt{\gamma}u + Au$ is an eigenvector of A belonging to the eigenvalue $\sqrt{\gamma}$. □

Proposition 4.5: Let A be a semilinear self-adjoint operator on a finite dimensional complex Hilbert space $\mathscr{H}$. Then A has a complete set of eigenvectors.

Proof: Let $\dim \mathscr{H} = n$. By Proposition 4.4 we find one eigenvector u of A and let $\mathscr{M}$ be the one-dimensional subspace of $\mathscr{H}$ spanned by u. Then by Proposition 4.3 $\mathscr{M}$ reduces A which means that the restriction of A to $\mathscr{M}^\perp$ is a self-adjoint semilinear operator on an $(n-1)$-dimensional Hilbert space. By repeated application of Propositions 4.4 and 4.3 we get n mutually perpendicular eigenvectors of A [actually if we use Corollary 4.3.2 it is sufficient to have obtained $(n-1)$ such eigenvectors] and since $\dim \mathscr{H} = n$, these eigenvectors must span $\mathscr{H}$. □

Remark: The Dirac formalism suggests that it could be possible to generalize Proposition 4.5 to the infinite dimensional case. A rigorous generalization in terms of spectral measures rather than eigenvectors is under construction and will be presented in a forthcoming paper. We have now seen that normal semilinear operators that have a complete set of eigenvectors are self-adjoint and we have also seen that self-adjoint semilinear operators have a complete set of eigenvectors. The question that arises next is the following: is there a normal semilinear operator that is not self-adjoint or, in other words, which does not have a complete set of eigenvectors? We present the answer in the following proposition.

Proposition 4.6: On every complex Hilbert space of dimension greater than one there exists a normal semilinear operator which is neither self-adjoint nor anti-self-adjoint.

Proof: Let $\dim \mathscr{H} = 2$. Let u_1 and u_2 be an orthonormal basis in $\mathscr{H}$. Define A by semilinearity and

$$Au_1 = (1/\sqrt{2})(u_1 + iu_2), \tag{4.10}$$

and

$$Au_2 = -(1/\sqrt{2})(iu_1 + u_2). \tag{4.11}$$

Then from (2.17) it follows that

$$A^*u_1 = (1/\sqrt{2})(u_1 - iu_2) \tag{4.12}$$

and

$$A^*u_2 = (1/\sqrt{2})(iu_1 - u_2). \tag{4.13}$$

Now

$$\begin{aligned} AA^*u_1 &= (1/\sqrt{2})(Au_1 + iAu_2) \\ &= \tfrac{1}{2}(u_1 + u_1 + iu_2 - iu_2) = u_1 \end{aligned} \tag{4.14}$$

and

$$A^*Au_1 = (1/\sqrt{2})(A^*u_1 - iA^*u_2)$$
$$= \tfrac{1}{2}(u_1 + u_1 - iu_2 + iu_2) = u_1. \quad (4.15)$$

A similar calculation shows that

$$AA^*u_2 = u_2 = A^*Au_2. \quad (4.16)$$

Thus A is normal and

$$AA^* = I, \quad (4.17)$$

where I is the identity operator on $\mathscr{H}$. Here, A thus defined is normal but neither self-adjoint nor anti-self-adjoint. Thus there exists a normal semilinear operator with the required properties on a two-dimensional Hilbert space. We can construct a semilinear operator on any Hilbert space of dimension greater than 2 with the required properties by taking the direct sum of the operator constructed above on a two-dimensional subspace with a self-adjoint semilinear operator on the orthogonal complement of the subspace. ☐

Remark: Since in Proposition 4.6 with A defined on a two-dimensional space, $A \neq A^*$, A cannot have any eigenvectors because if it had an eigenvector u, then span $\{u\}$ will reduce A and then since $(\text{span}\{u\})^{\perp}$ is a one-dimensional, A will have an eigenvector there also and A will have a complete set of eigenvectors which would imply $A = A^*$, a contradiction. Since every Hilbert space of dimension greater than 1 contains a two-dimensional subspace, this construction shows that in general normal semilinear operators need not have a complete set of eigenvectors.

Proposition 4.7: Let A be a normal semilinear operator on a complex Hilbert space $\mathscr{H}$. Let u and v be eigenvectors of A belonging to eigenvalues α and β, respectively. Then either

$$|\alpha| = |\beta| \quad (4.18)$$

or

$$\langle u,v\rangle = 0. \quad (4.19)$$

Proof: We have

$$Au = \alpha u \quad (4.20)$$

and

$$Av = \beta v. \quad (4.21)$$

Equation (4.21) implies by Proposition 4.3 that

$$A^*v = \beta v. \quad (4.22)$$

Hence

$$\alpha\langle u,v\rangle = \langle Au,v\rangle = \langle A^*v,u\rangle = \beta\langle v,u\rangle \quad (4.23)$$

or

$$|\langle u,v\rangle|(|\alpha| - |\beta|) = 0, \quad (4.24)$$

which implies that either (4.18) or (4.19) is true. ☐

Remark: Let a normal operator A have two linearly independent eigenvectors u and v belonging to the same eigenvalue. Then their linear combinations, if A is semilinear, will not in general be eigenvectors of A, though linear combinations involving only real coefficients will be eigenvectors of A. In such a situation things can always be so arranged that A has two mutually perpendicular eigenvectors belonging to the same eigenvalue. This is proved in the proposition that follows.

Proposition 4.8: Let u and v be two linearly independent eigenvectors of a normal semilinear operator A on a complex Hilbert space $\mathscr{H}$ belonging to the same eigenvalue α. Then there exist two mutually perpendicular eigenvectors of A belonging to the eigenvalue α.

Proof: We have

$$Au = \alpha u \quad (4.25)$$

and

$$Av = \alpha v. \quad (4.26)$$

Assume that $\alpha \neq 0$, then the following calculation using (2.17) and Proposition 4.3 shows that $\langle u,v\rangle$ is real:

$$\alpha\langle u,v\rangle = \langle Au,v\rangle = \langle A^*v,u\rangle = \alpha\langle v,u\rangle. \quad (4.27)$$

Hence Gram–Schmidt orthogonalization involves only real coefficients and therefore orthonormal vectors obtained by the Gram–Schmidt process will be eigenvectors of A belonging to α.

Finally assume $\alpha = 0$. In this case span $\{u,v\}$ belongs to the kernel of A and every nonzero vector in the kernel is an eigenvector of A belonging to the eigenvalue 0. So again the Gram–Schmidt process will provide two orthonormal eigenvectors belonging to the eigenvalue 0. ☐

Corollary 4.8.1: Let $u_1,\ldots,u_n$ be n linearly independent eigenvectors of a normal semilinear operator A on a complex Hilbert space $\mathscr{H}$ belonging to the same eigenvalue α. Then there exist n orthonormal eigenvectors of A belonging to the eigenvalue α.

Proof: A calculation similar to that in Proposition 4.8 shows that if $\alpha \neq 0$, then each inner product involved in the Gram–Schmidt process is real and if $\alpha = 0$, the span of eigenvectors is in the kernel and as before every nonzero vector in the kernel is an eigenvector belonging to the eigenvalue 0. ☐

The following proposition is a variant of Wigner's theorem.[19]

Proposition 4.9: Let U be a semilinear operator on a complex Hilbert space $\mathscr{H}$. Then the following assertions about U are equivalent:

(i) for every $x \in \mathscr{H}$,

$$\|Ux\| = \|x\|; \quad (4.28)$$

(ii) for every $(x,y) \in \mathscr{H} \times \mathscr{H}$,

$$\langle Ux,Uy\rangle = \langle y,x\rangle; \quad (4.29)$$

(iii) $U^*U = I.$ (4.30)

Proof: We shall prove that (i) ⇒ (ii) ⇒ (iii) ⇒ (i). Suppose (i) is true. Then

$$\|U(x-y)\|^2 = \|x-y\|^2, \quad (4.31)$$
$$\|Ux\|^2 = \|x\|^2, \quad (4.32)$$
$$\|Uy\|^2 = \|y\|^2, \quad (4.33)$$
$$\|U(x-iy)\|^2 = \|x-iy\|^2. \quad (4.34)$$

Hence

$$\langle Ux,Uy\rangle + \langle Uy,Ux\rangle = \langle x,y\rangle + \langle y,x\rangle \quad (4.35)$$

and

$$\langle Ux,Uy\rangle - \langle Uy,Ux\rangle = -\langle x,y\rangle + \langle y,x\rangle. \quad (4.36)$$

Therefore

$$\langle Ux,Uy\rangle = \langle y,x\rangle. \tag{4.37}$$

Suppose (ii) is true. Then

$$\langle U^*Uy,x\rangle = \langle y,x\rangle, \tag{4.38}$$

which implies that

$$U^*U = I. \tag{4.39}$$

Finally suppose that (iii) is true. Then

$$\|x\| = \langle U^*Ux,x\rangle^{1/2} = \langle Ux,Ux\rangle^{1/2} = \|Ux\|. \tag{4.40}$$

□

Proposition 4.10: Let U be a semiunitary self-adjoint operator on a finite-dimensional complex Hilbert space $\mathscr{H}$. Then U has a complete orthonormal set of fixed points.

Proof: Let α be an eigenvalue of U. From Proposition 4.9 it follows that

$$|\alpha| = 1. \tag{4.41}$$

From Proposition 4.1 it follows that one itself is an eigenvalue and then the corresponding eigenvector is, by definition, a fixed point. From Proposition 4.5 it follows that it has a complete set of orthonormal eigenvectors and it follows from Proposition 4.1 that each eigenvector after multiplication by a suitable phase factor belongs to the eigenvalue 1. Multiplication by phase factors does not affect the orthonormality of a set. Hence U has a complete set of orthonormal fixed points. □

Remark: Given a complete orthonormal set in a complex Hilbert space $\mathscr{H}$, we can define a semiunitary operator U by taking each member of the complete orthonormal set to be fixed points of U and extending it to $\mathscr{H}$ by semilinearity. Such a semiunitary operator is clearly self-adjoint, that is,

$$U = U^*, \tag{4.42}$$

or in other words

$$U^2 = I, \tag{4.43}$$

so that it is involutory also.

Proposition 4.11: There exists a semilinear norm preserving isomorphism between $\mathscr{L}(\mathscr{H})$ and $\mathscr{S}\mathscr{L}(\mathscr{H})$.

Proof: Let U be any self-adjoint semiunitary operator on $\mathscr{H}$. Define L_U: $\mathscr{L}(\mathscr{H}) \to \mathscr{S}\mathscr{L}(\mathscr{H})$ by

$$L_UA = UA. \tag{4.44}$$

The calculation

$$\begin{aligned} L_U(\alpha A + \beta B) &= U(\alpha A + \beta B) = \alpha^* UA + \beta^* UB \\ &= \alpha^* L_U A + \beta^* L_U B \end{aligned} \tag{4.45}$$

shows that L_U is semilinear.

Define $\hat{L}_U$: $\mathscr{S}\mathscr{L}(\mathscr{H}) \to \mathscr{L}(\mathscr{H})$ by

$$\hat{L}_U A = UA. \tag{4.46}$$

Then

$$\hat{L}_U \circ L_U A = UUA = IA = A \tag{4.47}$$

shows that $\hat{L}_U$ is the inverse of L_U. Hence L_U is bijective.

Finally the calculation

$$\|L_UA\| = \sup_{\|x\|=1} \|UAx\| = \sup_{\|x\|=1} \|Ax\| = \|A\| \tag{4.48}$$

shows that L_U is norm preserving. □

Remark: The Remark preceding Proposition 4.11 shows that given a complete orthonormal set in $\mathscr{H}$ there exists a self-adjoint semiunitary operator on $\mathscr{H}$ such that members of the orthonormal set are its fixed points. There are infinitely many different complete orthonormal sets in $\mathscr{H}$: it follows therefore from the proof of Proposition 4.11 that there are infinitely many norm-preserving bijections between $\mathscr{L}(\mathscr{H})$ and $\mathscr{S}\mathscr{L}(\mathscr{H})$.

V. SOME HEURISTIC CONSIDERATIONS

The study of semilinear operators is interesting in its own right, but our motivation for the study comes from theoretical physics where its relevance is conclusively demonstrated by the work of Wigner[19,20] and we hope that our results, too, will find many applications in physics. The real success of our results will come only after we have solved an unsolved new problem in physics with their help. However, we can demonstrate the relevance of our work to physics by providing astonishingly simple solutions to a couple of old problems whose solutions are known to every physicist. We shall show that the insight gained in our earlier work[13] on this subject enables us to find three involutory two-by-two complex matrices that anticommute with each other and to find four involutory four-by-four matrices with the same property and our solution is achieved without performing a single matrix multiplication. All we need to know to solve these problems are the following: (i) how to multiply a complex number by i, (ii) how to interchange two complex numbers, (iii) how to change the sign of a complex number (by putting a "minus" sign in front of it, of course), (iv) how to combine these operations, and (v) in order to get the matrix representations of our operations we need to know also that a linear transformation which takes the jth member of an orthonormal basis to the ith one and every other member to zero is represented by a matrix which has 1 in the ith row of the jth column and zeroes everywhere else and that all matrices are linear combinations of such elementary matrices.

In Ref. 13 we saw that a complex vector space $\mathscr{V}$ can be regarded as a real vector space $\mathscr{V}_r$ of twice the original dimension. When this is done multiplication by i in $\mathscr{V}$ is replaced by a linear operator $\mathbf{i}$ (denoted by a bold letter) in $\mathscr{V}_r$ and only those linear operators on the real space $\mathscr{V}_r$ that commute with $\mathbf{i}$ continue to be linear operators on the original space $\mathscr{V}$ when we revert to the original complex structure. It is immediately obvious that semilinear operators on $\mathscr{V}$ are linear operators on $\mathscr{V}_r$ and only those linear operators on $\mathscr{V}_r$ that anticommute with $\mathbf{i}$ become semilinear operators on $\mathscr{V}$ when we revert to the original space with its complex structure.

In a one-dimensional complex space $\mathscr{V}$ (each such space is isomorphic with $\mathbb{C}$ as a vector space and we can, without loss of generalization, take $\mathscr{V} = \mathbb{C}$) the spaces of both linear and semilinear operators are one-dimensional, which means that any linear operator is a scalar multiple of the identity operator I and any semilinear operator is likewise a scalar multiple of complex conjugation $\mathbf{c}$.

Let us regard $\mathscr{V} = \mathbb{C}$ as a two-dimensional real space $\mathscr{V}_r$. We choose the natural basis in $\mathscr{V}_r$, so that coordinates of a complex number α are (α_1,α_2), where α_1 is the real part of a and α_2 is the imaginary part. What does the linear operator **i** corresponding to i do in this space? We know that if

$$a = \alpha_1 + i\alpha_2, \tag{5.1}$$

then

$$ia = -\alpha_2 + i\alpha_1, \tag{5.2}$$

in other words **i** takes (α_1,α_2) into $(-\alpha_2,\alpha_1)$. Complex conjugation **c** is also a linear operator in $\mathscr{V}_r$; we know that if

$$a = \alpha_1 + i\alpha_2, \tag{5.3}$$

then

$$a^* = \alpha_1 - i\alpha_2, \tag{5.4}$$

in other words **c** takes (α_1,α_2) into $(\alpha_1,-\alpha_2)$. We know from our general considerations that

$$\mathbf{i}^2 = -I \tag{5.5}$$

and

$$\mathbf{c}^2 = I. \tag{5.6}$$

We now use the first part of the following lemma.

Lemma 5.1: Let A and B be two anticommuting linear operators. Then AB anticommutes with both A and B and any linear operator C that anticommutes with both A and B commutes with AB.

Proof: Obvious. ☐

It follows from the lemma that **i**, **c**, and **ic** = **t** (say) are three mutually anticommuting linear operators on $\mathscr{V}_r = \mathbb{R}^2$. We next compute

$$\mathbf{t}^2 = \mathbf{icic} = -i^2c^2 = I. \tag{5.7}$$

So all we have to do now to get our solution for anticommuting and involutive two-by-two complex matrices is to find the real matrices that represent **i**, **c**, and **t** and multiply the matrix of **i** by the pure imaginary number i so that it too becomes involutive. What does **t** do to (α_1,α_2)? Complex conjugation takes (α_1,α_2) to $(\alpha_1,-\alpha_2)$ and **i** takes $(\alpha_1,-\alpha_2)$ to (α_2,α_1). With the help of the rule on matrix representation mentioned in the opening paragraph of this section, we can now immediately write down the three matrices we need, namely, i times the matrix of **i** and matrices of **t** and **c**, which are

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

They look, apart from the order in which they are written (which is of no consequence), familiar enough and as promised we have not done a single matrix multiplication.

Next, to find four involutive four-by-four matrices that anticommute we do a similar analysis on a two-dimensional complex vector space $\mathscr{V} = \mathbb{C}^2$. The space $\mathscr{V}$ can be regarded as a four-dimensional real vector space $\mathscr{V}_r$ and we choose a basis in $\mathscr{V}_r$ in such a way that if a vector u in $\mathscr{V}$ has coordinates (α_1,α_2) with $\alpha_1 = \beta_1 + i\beta_4$ and $\alpha_2 = \beta_2 + i\beta_3$ then its coordinates in $\mathscr{V}_r$ are $(\beta_1,\beta_2,\beta_3,\beta_4)$. The choice may seem a little odd: a more natural first choice could be $\alpha_1 = \beta_1 + i\beta_2$ and $\alpha_2 = \beta_3 + i\beta_4$. However, the choice can be made in several different but equivalent ways and any choice will lead to a good answer. The reason for our particular choice will become clear in due course. Multiplication by i takes α_1 into $-\beta_4 + i\beta_1$ and α_2 into $-\beta_3 + i\beta_2$, or, in other words, the linear operator **i** in $\mathscr{V}_r$ corresponding to multiplication by i in $\mathscr{V}$ takes

$$(\beta_1,\beta_2,\beta_3,\beta_4) \text{ into } (-\beta_4,-\beta_3,\beta_2,\beta_1).$$

Conjugation takes α_1 into $\beta_1 - i\beta_4$ and α_2 into $\beta_2 - i\beta_3$, or in other words the linear operator **c** in $\mathscr{V}_r$ corresponding to complex conjugation in $\mathscr{V}$ takes

$$(\beta_1,\beta_2,\beta_3,\beta_4) \text{ into } (\beta_1,\beta_2,-\beta_3,-\beta_4).$$

Conjugation followed by multiplication by i will take α_1 into $\beta_4 + i\beta_1$ and α_2 into $\beta_3 + i\beta_2$, or in other words **t** = **ic** takes

$$(\beta_1,\beta_2,\beta_3,\beta_4) \text{ into } (\beta_4,\beta_3,\beta_2,\beta_1).$$

As before **i**, **c**, and **t** are mutually anticommuting and i **i**, **c**, and **t** are involutive. However, we need four such operators, but Lemma 5.1 tells us that no operator can anticommute with all three and therefore we can make use of only two of the three operators we have found so far. If we can find two nonsingular operators A and B that anticommute with each other and which commute with both **i** and **c**, then it is immediately obvious that each of the sets $\{\mathbf{i}A,\mathbf{i}B,\mathbf{c},\mathbf{t}\}$ and $\{\mathbf{c}A,\mathbf{c}B,\mathbf{i},\mathbf{t}\}$ consist of four mutually anticommuting operators and all we will have to do is to normalize the set to get the answer we want—a task made easier by the knowledge that **c** and **t** are already involutive.

Thus our problem is now reduced to finding two nonsingular operators A and B on $\mathscr{V}_r$ that anticommute and both of which commute with both **i** and **c**. This may come as a surprise to the reader, but we have already found two such operators: we have seen that

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

anticommute and we regard them as linear operators on our two-dimensional complex space $\mathscr{V}$; then their representations in $\mathscr{V}_r$ as four-by-four matrices are bound not only to do the same but also to commute with the representation of **i**. Furthermore, the first one clearly represents interchange of coordinates and it makes no difference in the final outcome whether we interchange two complex numbers first and then take their complex conjugates or take the complex conjugates of the two numbers and then interchange them: thus it commutes with **c** also. The second one represents a change of sign of the second coordinate and a change of sign clearly commutes with complex conjugation: thus it also commutes with **c**. Since the four-dimensional representation of both these operators commute with both **i** and **c**, their product will do likewise and we are free to choose any two from the three operators found in this way. (It is interesting to summarize what we have done to find these three four-by-four matrices that anticommute with each other and commute with multiplication by i and complex conjugation if the four-dimensional real space on which they operate is regarded as a two-dimensional complex space: for complex numbers operations of multiplication by i and complex conjugation anticommute with each other and the product of the two operations anticommutes with each: thus we have a set of three mutually anticommuting operators on the one-dimensional complex space $\mathbb{C}$; two of these are semilinear and one linear. When $\mathbb{C}$ is regarded as a two-dimensional real space these operators all become real linear and have representations as two-by-two real matrices that anticommute. These matrices can be regarded as linear operators on $\mathbb{C}^2$ and will have representations as four-by-four matrices if $\mathbb{C}^2$ is regarded as a four-dimensional real space. A representation cannot change their inherent properties and so the four-by-four matrices we find in this way will continue to anticommute: since they are linear operators on $\mathbb{C}^2$ they commute with multiplication by i and therefore with its representation as a four-by-four matrix. Furthermore as operators on $\mathbb{C}^2$ they involve only interchange of coordinates and change of sign of coordinates that obviously commute with the operation of taking complex conjugates of all complex coordinates.)

Before proceeding further we note that we are completely spoiled for choice: First, we could choose the coordinates of the four-dimensional real space corresponding to $\mathbb{C}^2$ in a variety of ways; second, we could choose either $\{\mathbf{i}A,\mathbf{i}B,\mathbf{c},\mathbf{t}\}$ or $\{\mathbf{c}A,\mathbf{c}B,\mathbf{i},\mathbf{t}\}$ as our four anticommuting operators and finally we can take A and B to be the four-dimensional representation of any two of the following three matrices:

$$\begin{bmatrix}0&-1\\1&0\end{bmatrix},\quad\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad\begin{bmatrix}0&1\\1&0\end{bmatrix},$$

regarded as linear operators on $\mathbb{C}^2$ and then $\mathbb{C}^2$ regarded as a four-dimensional real space. The various choices mentioned above are not exhaustive, for example, we can change the sign of any of these operators and have yet another choice. However, it is an inherent property of the problem that changing the sign of any matrix in the set of matrices constituting a solution is also a solution and different solutions obtained in this way clearly belong to the same equivalence class.

We choose the set $\{\mathbf{i}A,\mathbf{i}B,\mathbf{c},\mathbf{t}\}$ and the first two of the three matrices of the preceding paragraph as our A and B, respectively.

The matrix $[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}]$ interchanges the coordinates and then changes the sign of the first coordinate, in other words, regarded as an operator on $\mathbb{C}^2$ it takes (α_1,α_2) into $(-\alpha_2,\alpha_1)$. As we have taken this matrix to be our operator A, we now compute the action of $\mathbf{i}A$ on (α_1,α_2): (α_1,α_2) obviously goes into $(-i\alpha_2,i\alpha_1)$, or in other words $\mathbf{i}A$ takes

$$(\beta_1,\beta_2,\beta_3,\beta_4)\text{ into }(\beta_3,-\beta_4,\beta_1,-\beta_2),$$

where as defined earlier $(\beta_1,\beta_2,\beta_3,\beta_4)$ is the coordinate representation in $\mathbb{R}^4$ of (α_1,α_2) when $\mathbb{C}^2$ is regarded as a real vector space.

The matrix $[\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}]$ takes (α_1,α_2) into $(\alpha_1,-\alpha_2)$. As we have taken this matrix to be our operator B, we can now compute the action of $\mathbf{i}B$ on (α_1,α_2): (α_1,α_2) obviously goes into $(i\alpha_1,-i\alpha_2)$, or in other words $\mathbf{i}B$ takes

$$(\beta_1,\beta_2,\beta_3,\beta_4)\text{ into }(-\beta_4,\beta_3,-\beta_2,\beta_1).$$

We know that

$$\mathbf{c}^2=\mathbf{t}^2=I$$

and since A and B are four-dimensional representations of $\mathbf{i}$ and $\mathbf{c}$ of our earlier problem, we must have

$$-A^2=B^2=I$$

so that

$$(\mathbf{i}A)^2=-(\mathbf{i}B)^2=I.$$

Thus we have to multiply the matrix representation of $\mathbf{i}B$ by i to make all our matrices involutive. We can now write down the matrices of $\mathbf{i}A$, i times $\mathbf{i}B$, $\mathbf{c}$, and $\mathbf{t}$ because we know that they take $(\beta_1,\beta_2,\beta_3,\beta_4)$ into $(\beta_3,-\beta_4,\beta_1,-\beta_2)$, $(-i\beta_4,i\beta_3,-i\beta_2,i\beta_1)$, $(\beta_1,\beta_2,-\beta_3,-\beta_4)$, and $(\beta_4,\beta_3,\beta_2,\beta_1)$:

$$\begin{bmatrix}0&0&1&0\\0&0&0&-1\\1&0&0&0\\0&-1&0&0\end{bmatrix},\quad\begin{bmatrix}0&0&0&-i\\0&0&i&0\\0&-i&0&0\\i&0&0&0\end{bmatrix},$$
$$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix},\quad\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}.$$

Again the final answer is familiar and again we have nowhere done any matrix multiplication (anyway, not by the usual rules). The various choices we made at various stages were made in such a way as to ensure that we ended up with the familiar answer, but any choice at any stage would have led to an equally valid answer.

All our matrices are self-adjoint and traceless. None of this is accidental. At every stage we have used transformations on real spaces that take orthonormal bases into orthonormal bases, so all these transformations are not only nonsingular but also orthogonal. They were all chosen in such a way that their squares were either $+I$ or $-I$, which means that they are all either symmetric or antisymmetric. Symmetric real matrices are Hermitian when they are regarded as complex ones and all the matrices whose squares were $-I$ were multiplied by i and we know that antisymmetric real matrices multiplied by i are Hermitian. Finally self-adjoint involutive matrices are unitary, being unitary their eigenvalues have unit moduli and since these matrices are also self-adjoint their eigenvalues are all real. Hence $+1$ and -1 are the only possible eigenvalues of these matrices. Further the four-by-four matrices A and B were obtained by treating linear operators on $\mathbb{C}^2$ as linear operators on a four-dimensional real space. We have seen in Ref. 13 that in such situations all eigenvalues must be evenly degenerate: so A (or B) can have eigenvalues that are all $+1$ or all -1 or two of them are $+1$ and two of them -1. In the first two cases A is either $+I$ or $-I$ and in neither case can it anticommute with any operator: since it does anticommute with other operators these possibilities are excluded so it must have two eigenvalues that are $+1$ and two that are -1. Thus both A and B must be traceless. To prove tracelessness of $\mathbf{c}$ and $\mathbf{t}$ we first consider the two-by-two case. They are involutive, anticommuting, and self-adjoint. For reasons already stated possible eigenvalues are $+1$ and -1 and if both eigenvalues have the same sign the operator is either $+I$ or $-I$ and in either of these cases the operator cannot anticommute with any other operator. Hence one eigenvalue of each of these operators is $+1$ and the other -1, so that they are traceless. For the four-by-four case $\mathbf{c}$ and $\mathbf{t}$ can be obtained, with appropriate choice of coordinates, by taking direct sums of two identical copies of the corresponding operators on $\mathbb{C}$

regarded as a two-dimensional real space: for an arbitrary choice of coordinates these are unitarily equivalent to the direct sums. Hence **c** and **t** in the four-dimensional case must also be traceless.

VI. CONCLUDING REMARKS

We are aware that some readers will say that our analysis in the preceding section is more tortuous than a mere multiplication of matrices, but the power of the analysis is demonstrated by the fact that it quite naturally leads us to other possible answers also.

We have here only solved a problem whose solution is far too well known, but the novelty of our approach and the ease with which it led us in the right direction, we hope, demonstrates the relevance of our method to the study of spinors and conjugations.

The results we have developed so far are obviously relevant for the further development of the functional calculus where derivatives are required to be merely additive and not necessarily linear and also for the further development of the algebras of additive operators on both complex and quaternionic Hilbert spaces. We believe that the new calculus and the new operator algebras have important roles to play in modeling physical phenomena. It is manifestly clear that our work is directly relevant to the study of spin, charge conjugation, time reversal, and other similar topics in gauge field theories and we hope this has been adequately demonstrated by our heuristic discussion in the preceding section. Further work is in progress and will be reported in due course.

ACKNOWLEDGMENT

It is a pleasure to acknowledge numerous helpful suggestions from the referee, but we do claim absolute ownership of all the shortcomings that survived his scrutiny.

[1] C. Segrè, Rend. Palermo **5**, 192 (1891).

[2] R. Baer, *Linear Algebra and Projective Geometry* (Academic, New York, 1965).

[3] N. Jacobson, *Structure of Rings* (Am. Math. Soc., Providence, RI, 1956).

[4] S. Lang, *Algebra* (Addison–Wesley, Reading, MA, 1965).

[5] C. S. Sharma and I. Rebelo, Phys. Lett. A **44**, 29 (1973).

[6] C. S. Sharma and I. Rebelo, Int. J. Theor. Phys. **13**, 323 (1975).

[7] J. Pian and C. S. Sharma, Phys. Lett. A **76**, 365 (1980).

[8] J. Pian and C. S. Sharma, J. Phys. A **14**, 1261 (1981).

[9] J. Pian and C. S. Sharma, Int. J. Theor. Phys. **22**, 107 (1983).

[10] J. Pian and C. S. Sharma, Lett. Nuovo Cimento **40**, 126 (1984).

[11] J. Pian and C. S. Sharma, Nuovo Cimento B **94**, 193 (1986).

[12] C. S. Sharma and T. J. Coulson, J. Math. Phys. **28**, 1941 (1987).

[13] C. S. Sharma, J. Math. Phys. **29**, 1069 (1988).

[14] G. Fonte, Nuovo Cimento B **49**, 200 (1979).

[15] G. Fonte, J. Math. Phys. **21**, 800 (1980).

[16] G. Fonte and G. Schifferer, Nuovo Cimento B **74**, 1 (1983).

[17] G. Fonte, Lett. Nuovo Cimento **41**, 577 (1984).

[18] G. Fonte, Nuovo Cimento B **94**, 149 (1986).

[19] E. P. Wigner, Göttingen Nachrichten, Math.-Phys. **31**, 546 (1932).

[20] E. P. Wigner, *Group Theory* (Academic, New York, 1959).

[21] R. V. Kadison, Ann. Math. **54**, 325 (1951).

[22] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1964), Vol. II.

[23] L. P. Horwitz and L. C. Biedenharn, Ann. Phys. (NY) **157**, 432 (1984).

[24] S. Adler, Commun. Math. Phys. **104**, 611 (1986).

[25] C. S. Sharma, Nuovo Cimento B **100**, 291 (1987).

[26] C. S. Sharma and D. F. Almeida, submitted to J. Math. Phys.

[27] P. M. Cohn, Algebra 1 (Wiley, Chichester, 1982).

[28] W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973).

[29] D. Finkelstein, J. M. Jauch, and D. Speiser, in *The Logico-Algebraic Approach to Quantum Mechanics, Volume II: Contemporary Consolidation*, edited by C. A. Hooker (Reidel, Dordrecht, 1979).

Proper conformal symmetries of conformal symmetric space-times

Ramesh Sharma
Department of Mathematics, Michigan State University, East Lansing, Michigan 48824

(Received 30 March 1988; accepted for publication 29 June 1988)

It is shown that conformal symmetric space-times admitting an infinitesimal conformal symmetry are either of type O or N. Those of type N represent plane gravitational waves with parallel rays, provided the Einstein tensor is invariant of the infinitesimal conformal symmetry.

I. INTRODUCTION

Most of the well-known explicit solutions of the Einstein field equations admit one or more Killing vector fields (infinitesimal isometries). Solutions admitting proper conformal vector fields (infinitesimal conformal symmetries) are scarce. Unlike isometries and homothetic symmetries, the conformal symmetries do not preserve the Einstein tensor and place severe restrictions on the space-times. Collinson and French[1] proved that a vacuum space-time with a proper conformal symmetry must be of type N. Recently, Eardley *et al.*[2] have proved that the only vacuum solutions with a conformal symmetry are everywhere locally flat space-times (i.e., of type O), or of type N representing certain plane wave solutions. As established by Eardley *et al.*[2] and Garfinkle,[3] asymptotically flat space-times with reasonable energy conditions and a proper conformal symmetry are locally flat (type O). More recently, Garfinkle and Tian[4] have shown that vacuum space-times with cosmological constant and proper conformal symmetry are of constant curvature; i.e., are of type O and represent de Sitter and anti-de Sitter universes.

Generalizing these results for nonvacuum and non-Einstein space-times, we show that conformal symmetric space-times (in particular, symmetric ones) admitting a proper conformal symmetry are either of type O or N. Those of type N represent plane-fronted gravitational waves with parallel rays provided the Einstein tensor remains invariant of the conformal symmetry. Our result exhibits the fact that the existence of a proper conformal symmetry restricts the conformal symmetric space-times (which are in general, of type O, N, or D as shown by McLenaghan and Leroy[5]) to be of type O or N only.

II. PRELIMINARIES

We denote by M a four-dimensional space-time manifold, its Lorentzian metric tensor by g_{ab}, Christoffel symbols by Γ^a_{bc}, covariant derivative operator by ∇_a, Riemann curvature tensor by $R^{\ a}_{bcd}$, Ricci curvature tensor by R_{ab}, scalar curvature by R, and the Weyl conformal curvature tensor by $C^{\ a}_{bcd}$.

A space-time is said to be of type D if the Weyl conformal tensor has two pairs of coincident principal null directions, of type N if it has four distinct principal null directions, and of type O (conformally, flat) if it does not single out any principal null directions. For details on this classification we refer to Kramer *et al.*[6]

McLenaghan and Leroy[5] defined M to be complex recurrent if

$$\nabla_e \overset{+}{C}_{abcd} = K_e \overset{+}{C}_{abcd}, \tag{1}$$

where $\overset{+}{C}_{abcd}$ denotes the self-dual part of C_{abcd}, given by

$$\overset{+}{C}_{abcd} = \tfrac{1}{2}(C_{abcd} - i{*}C_{abcd}), \tag{2}$$

with $*$ denoting the dual operator, ${*}\,C_{abcd} = \frac{1}{2}\epsilon_{abef}\,C^{ef}_{\ \ cd}$, and $i = \sqrt{-1}$. The recurrence vector field K_e is, in general, complex. For $K_e = 0$, M is called conformal symmetric space-time (Chaki and Gupta[7]) defined by

$$\nabla_e C_{abcd} = 0. \tag{3}$$

Obviously, the symmetric space-times ($\nabla_e R_{abcd} = 0$) are conformal symmetric; but the converse need not be true. It is known[5] that complex recurrent space-times (in particular conformal symmetric space-times) are of type O, N, or D. (Note that McLenaghan and Leroy[5] assumed the recurrent space-times in their definition not to be conformally flat, whereas we relax this condition.)

A space-time M is said to admit a conformal vector field (infinitesimal conformal symmetry) ξ if there exists a smooth scalar field σ on M such that

$$\pounds_\xi g_{ab} = 2\sigma g_{ab}, \tag{4}$$

where $\pounds_\xi$ stands for the Lie-derivative operator along ξ. For σ nonzero constant, ξ is called a homothetic vector field (infinitesimal homothetic symmetry) and for $\sigma = 0$, ξ is called a Killing vector field (infinitesimal isometry). A conformal vector field is known to satisfy the following[8]:

$$\pounds_\xi \Gamma^a_{bc} = \delta^a_b \sigma_c + \delta^a_c \sigma_b - g_{bc}\sigma^a, \tag{5}$$

$$\begin{aligned}\pounds_\xi R^{\ a}_{bcd} = {} & -\delta^a_c \nabla_d \sigma_b + \delta^a_d \nabla_c \sigma_b \\ & - (\nabla_c \sigma^a) g_{bd} + (\nabla_d \sigma^a) g_{bc},\end{aligned} \tag{6}$$

$$\pounds_\xi R_{ab} = -2\nabla_a \sigma_b - (\Box\sigma) g_{ab}, \tag{7}$$

$$\pounds_\xi R = -2\sigma R - 6\Box, \tag{8}$$

$$\pounds_\xi G_{ab} = -2\nabla_a \sigma_b + 2(\Box\sigma) g_{ab}, \tag{9}$$

$$\pounds_\xi C^{\ a}_{bcd} = 0, \tag{10}$$

where $\sigma^a = \nabla^a \sigma$, $\Box\sigma = \nabla_a \nabla^a \sigma$, and $G_{ab} \equiv$ Einstein tensor $= R_{ab} - \frac{1}{2} R g_{ab}$.

III. CONFORMAL SYMMETRIC SPACE-TIMES WITH PROPER CONFORMAL SYMMETRY

We state and prove our main result as follows.

Theorem: Let a conformal symmetric space-time M admit a proper conformal vector field ξ. Then M is either of type O or N. In case M is of type N and the Einstein tensor is invariant under ξ, M represents plane-fronted gravitational waves with parallel rays.

Proof: Let us consider the commutation formula[8]

$$\begin{aligned}\pounds_\xi \nabla_a C^{b}{}_{cde} - \nabla_a \pounds_\xi C^{b}{}_{cde} \\ = (\pounds_\xi \Gamma^{b}_{af}) C^{f}{}_{cde} - (\pounds_\xi \Gamma^{f}_{ac}) C^{b}{}_{fde} \\ - (\pounds_\xi \Gamma^{f}_{ad}) C^{b}{}_{cfe} - (\pounds_\xi \Gamma^{f}_{ae}) C^{b}{}_{cdf}.\end{aligned} \tag{11}$$

By hypothesis, M is conformal symmetric and hence we have (3). Moreover, ξ being conformal, we must have (5) and (10). Consequently, Eq. (11) assumes the form

$$\begin{aligned}(\delta^b_a \sigma_f + \delta^b_f \sigma_a - g_{af}\sigma^b) C^{f}{}_{cde} - (\delta^f_a \sigma_c + \delta^f_c \sigma_a - g_{ac}\sigma^f) C^{b}{}_{fde} \\ - (\delta^f_a \sigma_d + \delta^f_d \sigma_a - g_{ad}\sigma^f) C^{b}{}_{cfe} \\ - (\delta^f_a \sigma_e + \delta^f_e \sigma_a - g_{ae}\sigma^f) C^{b}{}_{cdf} = 0.\end{aligned} \tag{12}$$

A straightforward contraction at a and b yields

$$\sigma_a C^{a}{}_{bcd} = 0. \tag{13}$$

Use of (13) back in (12) implies

$$\sigma^a \sigma_a C^{b}{}_{cde} = 0.$$

It shows that, either $C^{b}{}_{cde} = 0$ (that is, M is of type O) or $\sigma^a \sigma_a = 0$. Since ξ is a proper conformal vector field, $\sigma^a \neq 0$ and hence must be null. This fact, together with (13), proves that M is of type N and the quadruply repeated principal null direction of the Weyl tensor is given by σ^a.

Now, if M is of type N and $\pounds_\xi G_{ab} = 0$, then it follows after a little computation that $\nabla_a \nabla_b \sigma = 0$. Hence σ_a generates a nonrotating, shear-free, divergence-free, null geodesic congruence. Thus M represents a plane-fronted gravitational wave with parallel rays. This completes the proof.

Remark 1: The metric of a conformal symmetric space-time of type N can be written (in local coordinates) as[5]

$$\begin{aligned}ds^2 = -2\{(1+e)x^2 + (e-1)y^2\} du^2 \\ - 2\, du\, dr + dx^2 + dy^2,\end{aligned} \tag{14}$$

where $e = e(u)$ is an arbitrary real function.

Remark 2: The condition "Einstein tensor is invariant under ξ" stated in the hypothesis of the theorem means that ξ defines a natural symmetry ($\pounds_\xi G_{ab} = 0$) of Einstein's field equations and is equivalent to $\nabla_a \nabla_b \sigma = 0$ [as can be observed in view of Eq. (9)]. For a conformal vector ξ, this is further equivalent to $\pounds_\xi R^{a}{}_{bcd} = 0$, which defines a fundamental symmetry of the space-times, called curvature collineation. (See Katzin *et al.*[9] for a detailed treatment of this symmetry.)

IV. SYMMETRIC SPACE-TIMES WITH CONFORMAL SYMMETRY

The theorem given in the previous section holds for symmetric space-times that are special cases of conformal symmetric ones. To characterize symmetric space-times we assume that the Einstein tensor is invariant under the proper conformal vector field ξ. Then, as mentioned in Remark 2 above, we have $\nabla_a \nabla_b \sigma = 0$ and $\pounds_\xi R^{a}{}_{bcd} = 0$. Using the last equation, Eq. (5), and $\nabla_a R^{b}{}_{cde} = 0$, in the following commutation formula[8]:

$$\begin{aligned}\pounds_\xi \nabla_a R^{b}{}_{cde} - \nabla_a \pounds_\xi R^{b}{}_{cde} \\ = (\pounds_\xi \Gamma^{b}_{af}) R^{f}{}_{cde} - (\pounds_\xi \Gamma^{f}_{ac}) R^{b}{}_{fde} \\ - (\pounds_\xi \Gamma^{f}_{ad}) R^{b}{}_{cfe} - (\pounds_\xi \Gamma^{f}_{ae}) R^{b}{}_{cdf},\end{aligned}$$

we can show that σ^a is a null vector field and $R = 0$. Moreover, as $\nabla_a \nabla_b \sigma = 0$, we also have σ^a as a covariant constant. From the obvious relations $\sigma_a R^{a}{}_{bcd} = 0$ and $\sigma_a C^{a}{}_{bcd} = 0$, it follows that $\sigma_a R_{bc} = \sigma_b R_{ac}$. We therefore observe that $R_{ab} = \lambda \sigma_a \sigma_b$ ($\lambda = \text{const}$). Thus these space-times represent plane-fronted gravitational waves with parallel rays and have the metric (14) with $e = \text{const}$. The Einstein–Maxwell equations for these space-times have their solutions as null electromagnetic fields. Other solutions of these space-times are directed massless radiation that may be considered as incoherent superposition of waves with random phases and polarizations but the same propagation direction. (See Kramer *et al.*[6])

V. CONCLUDING REMARK

As pointed out in Sec. II, conformal symmetric space-times are of type O, D, or N. We have shown that the existence of a proper conformal symmetry restricts those space-times to be of type O or N only. This supports the fact that conformal symmetry imposes severe restrictions on the space-times.

ACKNOWLEDGMENTS

I am grateful to Lord Sri Satya Sai Baba whose grace made this work possible. I am thankful to the referee for his valuable comments and suggestions towards the improvement of this work.

[1] C. D. Collinson and D. C. French, J. Math. Phys. **8**, 701 (1967).

[2] D. Eardley, J. Isenberg, J. Marsden, and V. Moncrief, Commun. Math. Phys. **106**, 137 (1986).

[3] D. Garfinkle, J. Math. Phys. **28**, 28 (1987).

[4] D. Garfinkle and Q. Tian, Class. Quantum Gravit. **4**, 137 (1987).

[5] R. G. McLenaghan and J. Leroy, Proc. R. Soc. London Ser. A **327**, 229 (1972).

[6] D. Kramer, H. Stephani, E. Herlt, and M. MacCallum, *Exact Solutions of Einstein's Field Equations* (Cambridge U.P., Cambridge, 1980).

[7] M. C. Chaki and B. Gupta, Indian J. Math. **5**, 113 (1963).

[8] K. Yano, *The Theory of Lie Derivatives and its Applications* (Interscience, New York, 1957).

[9] G. H. Katzin, J. Levine, and W. R. Davis, J. Math. Phys. **10**, 617 (1969).

Factorization of zero curvature representations

R. K. Dodd[a] and S. K. Paul
Research Institute for Theoretical Physics, Hiroshima University, Takehara, Hiroshima 725, Japan

(Received 23 February 1988; accepted for publication 13 July 1988)

The factorization of the L operator in a Lax pair $\{L,A\}$ has been extremely useful in the theory of integrable systems. For instance, the Miura transformation is obtained by factorizing the Schrödinger operator; $-\partial_x^2+u-k^2=(-\partial_x+a)(\partial_x+b)$ requires that $u=k^2-a_x+a^2$ and $a=b$. In a recent paper an analogous procedure for the factorization of zero curvature representations was presented. In this paper the theory of this method is developed and applied to the third-order scalar Lax equation.

I. INTRODUCTION

In Ref. 1, henceforth referred to as I, a method was presented for factorizing zero curvature representations and obtaining new integrable equations that could be associated with the original integrable system. In this Introduction we review this material and develop it further. The theory is then applied in Sec. II to the third-order scalar Lax equation.

Let

$$Y_x=P(k)Y,\quad Y_t=Q(k)Y \tag{1.1}$$

be a zero curvature representation of the integrable system[2]

$$P_t(k)-Q_k(k)+[P(k),Q(k)]=0\,. \tag{1.2}$$

The $P(k)$, $Q(k)$ are matrix valued functionals, which we assume belong to $\mathrm{gl}(n,\mathbb{C})$, of the variables in the solvable equation as well as depending upon the isospectral parameter k. Given the zero curvature representation (1.1), it is convenient to denote the solvable equation (1.2) by $\{P,Q\}$. If (P_i,Q_i,Y_i) satisfy (1.1), then we write $(P_i,Q_i)\in\{P,Q\}$.

In particular, assume that (1.2) admits an auto-Bäcklund transformation (ABT) that can be obtained from a k-dependent gauge transformation of (1.1). Let (P_i,Q_i,Y_i) and T_j denote, respectively, the solutions to (1.1) and the gauge transformations $i=0,\dots,N$, $j=0,\dots,N-1$ with (P_0,Q_0,Y_0) denoting a given "seed" or initial solution. Then

$$Y_{i+1}(k)=T_i(k)Y_i(k)\,,\quad 0\leqslant i\leqslant N-1\,, \tag{1.3}$$

and the solution $(P_{i+1},Q_{i+1})\in\{P,Q\}$ is given by

$$P_{i+1}T_i=T_{i,x}+T_iP_i,\quad Q_{i+1}T_i=T_{i,t}+T_iQ_i\,. \tag{1.4}$$

In principle, there is considerable freedom in the choice of T_i both as a function of k and in terms of the arbitrary parameters which it contains. It is convenient to think of T_i as adding in one soliton although this is not necessary to the method. For the equations considered in this section this requires T_i to be linear in k.

Let T be one of the gauge transformations introduced above; then in general $T\in\mathrm{GL}(n,\mathbb{C})$. In this case we can apply the Gauss decomposition

$$T=\Delta^-\Delta^+\,, \tag{1.5}$$

where Δ^+ is upper unipotent (1's on the diagonal) and Δ^- is lower triangular. The decomposition (1.5) is unique if T is nonsingular. If T is singular then Δ^+ is unique. It is sometimes useful to use a weaker form of decomposition in which Δ^+ is no longer unipotent so that the decomposition is not unique. This corresponds to the Bruhat decomposition of T which we write as

$$T=\Delta^-\pi\Delta^+\,, \tag{1.6}$$

where π is a permutation matrix and Δ^+ (Δ^-) are upper (lower) triangular matrices. We can generally take $\pi=1$. Finally, we mention another decomposition that is also useful in the general theory. It gives the usual form of the factorizations associated with the modified Korteweg–de Vries (MKdV) and sine–Gordon (SG) equations,

$$T=U\Delta^+\,, \tag{1.7}$$

where U is a unitary matrix and Δ^+ is upper triangular.

Factorization problem: *to determine the conditions that T must satisfy to ensure the existence of an intermediate equation given by the Gauss decomposition (1.5).*

Before considering this in detail we introduce some further notation used extensively throughout the paper. An upper index on a quantity labels the intermediate equation; thus $(P_i^j,Q_i^j)\in\{P^j,Q^j\}$ refers to solutions that belong to the jth intermediate equation with $j=0$ the initial or seed equation.

The factorization problem, if solvable, determines at the ith step the ith intermediate equation $\{P^i,Q^i\}$ through the diagram

$$\begin{tikzcd}(P_j^{i-1},Q_j^{i-1}) \arrow[r, "\Delta_j^{+\,i-1}"] & (P_j^i,Q_j^i) \arrow[r, "\Delta_j^{-\,i-1}"] & (P_{j+1}^{i-1},Q_{j+1}^{i-1})\,.\end{tikzcd}$$

In all diagrams arrows will indicate maps between the fundamental solutions of the zero curvature representations.

Consider the factorization of the ith transform of the zeroth equation. For simplicity we deal only with the P equation—analogous statements about the Q transform are obtained by the replacements $P\to Q$, $x\to t$:

$$P_{i+1}^0T_i^0=T_{i,x}^0+T_i^0P_i^0\,, \tag{1.8}$$

$$\begin{aligned}P_i^1\Delta_i^{+0}&=\Delta_{i,x}^{+0}+\Delta_i^{+0}P_i^0\,,\\ \Delta_i^{-0}P_i^1&=-\Delta_{i,x}^{-0}+P_{i+1}^0\Delta_i^{-0}\,.\end{aligned} \tag{1.9}$$

There is a distinct difference between these two sets of transforms. First, we consider the ABT for the integrable equation $\{P^0,Q^0\}$ defined by (1.9).

(i) The integrable equation is k independent.

(ii) Constraints are imposed on T_i^0 to ensure that

---

[a] Permanent address: School of Mathematics, Trinity College, Dublin 2, Ireland.

0022-2488/88/112423-07$02.50

(P^0_{i+1},Q^0_{i+1}) satisfy the linear system (1.1) defined by $(P^0_{i+1},Q^0_{i+1},Y^0_{i+1})$.

(iii) The complex k-analytic structure of T^0_i determines the ABT.

(iv) In general, T^0_i belongs to the gauge group (GL(n,C)).

Usually point (ii) is hidden in the *a priori* assumption that (P^0_{i+1},Q^0_{i+1}) satisfy (1.1) and the constraints are not explicitly derived. The following example will clarify these points. Let $E_{i,j}:=(\delta_{ui}\delta_{vj})_{l<u,v<n}$ be the $n\times n$ matrix with 1 in the ij th position and zero elsewhere.

A. The AKNS system

The AKNS system (Ref. 3) involves sl(2,C). Put $h:=E_{11}-E_{22}$, $e:=E_{12}$, $f:=E_{21}$. Then

$$P^0_i(k)=kh+q_ie+r_if,$$
$$Q^0_i(k)=A_i(k)h+B_i(k)e+C_i(k)f,$$

and (1.2) gives, for $\{P^0,Q^0\}$,

$$A_x+rB-qC=0\,,$$
$$q_t-B_x+2kB-2qA=0\,,$$
$$r_t-C_x-2kC+2rA=0\,.$$

Then as is well known the integrable equations are obtained by expanding A, B, and C as Laurent series in k whose coefficients are functionals of q and r. Viable reductions of the system occur for

(i) $r=-1$,

(ii) $r=\epsilon q=\epsilon\bar{q}$,

(iii) $r=\epsilon q,\quad \epsilon=\pm 1$.

In particular, the following forms of well-known equations are obtained:

$$q_t+6qq_x+q_{3x}=0\quad\text{(the KdV equation)},$$
$$r=-1,\quad A=-4k^3-2kq-q_x,$$
$$B=-4k^2q-2kq_x-2q^2-q_{2x}\,,$$
$$C=4k^2+2q;$$
$$q_t-6q^2q_x+q_{3x}=0\quad\text{(the MKdV equation)}\,,$$
$$r=-q=-\bar{q},\quad A=-4k^3-2kq^2,$$
$$B=-4k^2q-2kq_x-q_{2x}-2q^3\,,$$
$$C=4k^2q-2kq_x+q_{2x}+2q^3;$$
$$u_{xt}=\sin u\quad\text{(the SG equation)}\,,$$
$$-r=q=\bar{q}=-\tfrac{1}{2}u_x,\quad A=(1/4k)\cos u,$$
$$B=(1/4k)\sin u,\quad C=(1/4k)\sin u\,.$$

Let $T:=T^0_i$ be linear in k; then in order for it to define an ABT of the system $\{P^0,Q^0\}$, the k-analytic structure of (1.8) requires T to satisfy the following conditions:

$$[h,\dot{T}(0)]=0\,,$$
$$P^0_{i+1}(0)\dot{T}(0)=\dot{T}(0)_x+[T(0),h]+\dot{T}(0)P^0_i(0)\,,$$
$$P^0_{i+1}(0)T(0)=T(0)_x+T(0)P^0_0(0)\,,$$

where $\dot{T}(k)=\partial_k T(k)$ and $P^0_{i+1}(0)$ is interpreted as an off-diagonal matrix. This system can be solved to obtain the ABT, and $T(k)$ can be written in terms of the ith and $(i+1)$th variables. Under the conditions $q_i,r_i\rightarrow 0$, as $x\rightarrow+\infty$, we obtain

$$v_1q_{i,x}-v_2q_{i+1,x}$$
$$=q_{i+1}p_2-q_ip_1+(q_{i+1}v_2+q_iv_1)J_{i+1,i}\,,$$
$$v_1r_{i+1,x}-v_2r_{i,x}$$
$$=r_{i+1}p_1-r_ip_2-(r_{i+1}v_1+r_iv_2)J_{i+1,i}\,,$$

where

$$J_{i+1,i}:=\int_x^\infty(q_{i+1}r_{i+1}-q_ir_i)dx$$

and v_j,p_j are constants. The transformation T^0_i can be written as

$$T^0_i:=[v_1k+\tfrac{1}{2}(p_1-v_1J_{i+1,i})]E_{11}+\tfrac{1}{2}(q_iv_1-q_{i+1}v_2)E_{12}$$
$$+\tfrac{1}{2}(r_{i+1}v_1-r_iv_2)E_{21}$$
$$+[v_2k+\tfrac{1}{2}(p_2+v_2J_{i+1,i})]E_{22}\,.$$

If we consider (1.9), then in general (P^1_i,Q^1_i) are functions of k so that if the intermediate equation exists it is also a function of k. If Δ_i^{-0} is nonsingular, then (P^1_i,Q^1_i) is obtained from (1.9) without any constraints. From Eqs. (1.8) and (1.9) we get, without assuming the transformations are nonsingular, that

$$\{P^0_{i+1,t}-Q^0_{0+1,x}+[P^0_{i+1},Q^0_{i+1}]\}T^0_i$$
$$=T^0_i\{P^0_{i,t}-Q^0_{i,x}+[P^0_i,Q^0_i]\}\,,\tag{1.10}$$

$$\begin{aligned}&\{P^1_{i,t}-Q^1_{i,x}+[P^1_i,Q^1_i]\}\Delta_i^{+0}\\&\quad=\Delta_i^{+0}\{P^0_{i,t}-Q^0_{i,x}+[P^0_i,Q^0_i]\}\,,\\&\Delta_i^{-0}\{P^1_{i,t}-Q^1_{i,x}+[P^1_i,Q^1_i]\}\\&\quad=\{P^0_{i+1,t}-Q^0_{i+1,x}+[P^0_{i+1},Q^0_{i+1}]\}\Delta_i^{-0}\,.\end{aligned}\tag{1.11}$$

For $(P^0_i,Q^0_i)\in\{P^0,Q^0\}$, $(P^0_{i+1},Q^0_{i+1})\in\{P^0,Q^0\}$ provided T^0_i satisfies the constraint conditions or, equivalently, the solutions of $\{P^0,Q^0\}$ corresponding to (P^0_i,Q^0_i), (P^0_{i+1},Q^0_{i+1}) are connected by an ABT. In this case $(P^0_{i+1},Q^0_{i+1})\in\{P^0,Q^0\}$ even if T^0_i is singular. Suppose T^0_i is singular at $k=k_c$, $\det T^0_i(k_c)=0$; then we define $S^0_i(k_c):=T^0_i(k_c)$ and usually write $S^0_i(k)$ for simplicity. If T^0_i is nonsingular then Eqs. (1.11) show that $\{P^1,Q^1\}$ is identically zero. Therefore in order to obtain a nontrivial intermediate equation it is necessary that T^0_i be singular. Assume $S^0_i(k)$ exists; then $S^0_i=\Delta_i^{-0}\Delta_i^{+0}$ and Δ_i^{-0} is singular. In this case the second equation in (1.9) imposes constraints on (P^1_i,Q^1_i) that define Bäcklund transformations between $\{P^0,Q^0\}$ and $\{P^1,Q^1\}$. Equations (1.11) show that this constrained $(P^1_i,Q^1_i)\in\{P^1,Q^1\}$, which is nontrivial when $(P^0_i,Q^0_i)\in\{P^0,Q^0\}$.

Proposition 1: (a) Let $\{P^0,Q^0\}$ be a given integrable equation that admits an ABT defined by the gauge transformation T^0,

$$(P^0_i,Q^0_i)\xrightarrow{T^0_i}(P^0_{i+1},Q^0_{i+1}).$$

If $T^0=\Delta^{-0}\Delta^{+0}$ is the Gauss decomposition, then a necessary condition for the existence of an intermediate equation $\{P^1,Q^1\}$ defined by

$$(P_i^0,Q_i^0) \xrightarrow{\Delta_i^{+0}} (P_i^1,Q_i^1) \xrightarrow{\Delta_i^{-0}} (P_{i+1}^0,Q_{i+1}^0),$$

$(P_i^1,Q_i^1)\in\{P^1,Q^1\}$, is that there exist $k=k_c$ for which $S^0(k_c):=T^0(k_c)$ is singular.

(b) For $k\neq k_c$, $T^0(k)\in\mathrm{GL}(n,C)$ and the Gauss decomposition is unique. The singular transformation $S^0(k)$ does not have a unique Gauss decomposition; however, Δ_+^0 is unique and belongs to the gauge group. Then $S^0(k_c)$ can be viewed as a limiting form of a family of nonsingular gauge transformations with unique Gauss decompositions.

As an example we consider the AKNS system.

B. The first intermediate equations for the AKNS system

Put $\Delta^+(a):=E_{11}+aE_{12}+E_{22}$, $\Delta^-(b,c,d):=bE_{11}+cE_{21}+dE_{22}$; then using T_0^0 defined above, the Gauss decomposition $T_0^0=\Delta_0^{-0}\Delta_0^{+0}$, where $\Delta_0^{+0}:=\Delta^+(a_0)$, $\Delta_0^{-0}:=\Delta^-(b_0,c_0,d_0)$, gives

$$a_0=\tfrac{1}{2}(v_1q_0-v_2q_1)/(kv_1+\tfrac{1}{2}(p_1-v_1J_{10})),$$
$$b_0=v_1k+\tfrac{1}{2}(p_1-v_1J_{10}),$$
$$c_0=\tfrac{1}{2}(v_1r_1-v_2r_0),$$
$$d_0=l(k)/(kv_1+\tfrac{1}{2}(p_1-v_1J_{10})),$$

where

$$l(k):=k^2v_1v_2+\tfrac{1}{2}k(v_1p_2+v_2p_1)+\tfrac{1}{4}p_1p_2\,.$$

In deriving this expression for d_0 we use the relationship

$$(v_1r_1-v_2r_0)(v_1q_0-v_2q_1)$$
$$=-(v_1p_2-v_2p_1)J_{10}-v_1v_2J_{10}^2$$

obtained from the ABT given above. It is clear that the critical values of k for which the transformation is singular are given by $l(k)=0$. Thus the factorization of $S_0^0(k)$ gives $\Delta_0^{+0}:=\Delta^+(a_0)$, $\Delta_0^{-0}:=\Delta^-(b_0,c_0,0)$. The second relation in (1.9) now imposes a constraint on P_0^1, and the Bäcklund transformations relating $\{P^0,Q^0\}$ and $\{P^1,Q^1\}$ are obtained from Eqs. (1.9):

$$P_0^1:=(k+a_0r_0)h+r_0f,$$
$$Q_0^1:=(A_0+a_0C_0)h+C_0f.$$

For the cases which we consider we have Table I.

TABLE I. Conditions for the existence of an intermediate equation.

$r=-1$	$r=-q=-\bar{q}$	BT
$q_0=2a_0k-a_0^2-a_{0,x}$	$q_0=(2a_0k-a_{0,x})/(1+a_0^2)$	$(P_0^0,Q_0^0)\xrightarrow{\Delta_0^{+0}}(P_0^1,Q_0^1)$
$q_1=2a_0k-a_0^2+a_{0,x}$	$q_1=(2a_0k+a_{0,x})/(1+a_0^2)$	$(P_0^1,Q_0^1)\xrightarrow{\Delta_0^{-0}}(P_1^0,Q_1^0)$

We use the facts that $c_0=\text{const}$ when $r=-1$ and $a_0=-c_0/b_0$ when $r=-q=-\bar{q}$ to obtain these results.

The first intermediate equations can now be obtained from (1.2) using the expressions for P_0^1,Q_0^1 derived above. For the cases of interest see Table II.

TABLE II. First intermediate equations.

Seed equation	First intermediate equation
$q_t+6qq_x+q_{3x}=0$	$a_t-6a^2a_x+12kaa_x+a_{3x}=0$
$q_t-6q^2q_x+q_{3x}=0$	$v_t+[v_{2x}+3k^2v^2+\tfrac{3}{2}(v^2v_x^2/(1-v^2))]_x=0$ $u_t+u_{3x}+\tfrac{1}{2}u_x^3+6k^2u_x\sin^2 u=0$ with $v=2a/(1+a^2)=\sin u$
$u_{xt}=\sin u$	$v_{xt}=(1-k^2v_t^2)^{1/2}\sin v$ with $a=\tan\tfrac{1}{2}v$

The intermediate equations for the MKdV and the SG are more naturally associated with the factorization (1.7) with

$$U=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

These Bäcklund transformations are Miura transformations. That is, the transformations are (for fixed k) 1–1 in one direction and 1–∞ in the reverse direction. They are called Bäcklund transformations of type II, in the classical literature,[4] whereas ABT's are of type III. The usefulness of Miura transformations and ABT's lies in the fact that the transformed equation does not involve the seed solution of the source equation. In general, though, this is not the case. In this paper we will only consider intermediate equations that admit ABT's and are derived from Miura transformations.

The derivation of the first intermediate equation given above is easily extended to define a hierarchy of intermediate equations associated with a given seed equation. At each step we obtain by this process a zero curvature representation and an ABT for the intermediate equation.

Consider the seed equation $\{P^0,Q^0\}$ and the family of transformations $\{T^0\}$ that satisfy Proposition 1. Let T_0^0,T_1^0,T_2^0 be the transformations defining the ABT's,

$$T_0^0:(P_0^0,Q_0^0)\rightarrow(P_1^0,Q_1^0),\quad T_1^0:(P_0^0,Q_0^0)\rightarrow(P_2^0,Q_2^0),$$
$$T_2^0:(P_2^0,Q_2^0)\rightarrow(P_3^0,Q_3^0).$$

Let $S_0^0=\Delta_0^{-0}\Delta_0^{+0}$, $S_2^0=\Delta_2^{-0}\Delta_2^{+0}$ be the corresponding singular transformations of T_0^0 and T_2^0. Then an ABT for $\{P^1,Q^1\}$, T_0^1: $(P_0^1,Q_0^1)\rightarrow(P_1^1,Q_1^1)$, is defined by

$$T_0^1:=(\Delta_2^{+0})T_1^0(\Delta_0^{+0})^{-1}.$$

The transformation factorizes and defines the second intermediate equation provided S_0^1 exists. From the definition this requires T_1^0 to be singular. This process can be repeated an arbitrary number of times and defines a hierarchy of intermediate equations. Zero curvature representations and ABT's for the equations are obtained at the same time. The process is depicted in Fig. 1.

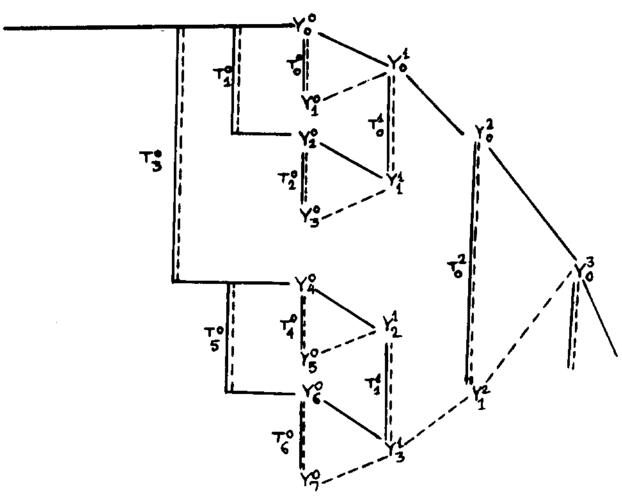


FIG. 1. The hierarchy of intermediate equations and Bäcklund transformations associated with a zero curvature representation $\{P^0,Q^0\}$. Moving between variables connected by an oblique line to the right corresponds to a transformation to the next intermediate equation, causing the previous transformation to become singular. This requires that the component of the transformation defining an ABT of the seed equation should be singular. The transformations defining ABT's of the seed equation are the vertical lines in the diagram and a singular transformation is denoted by a broken line.

In general, though, the hierarchy does not belong to the class we consider and usually only the first intermediate equation is of the required type. Thus the second intermediate equation associated with AKNS equations for which $q= -r= -\bar{r}$ is found to be derived from the type III transformation

$$a_{1,x} - 2ka_1 + [(2ka_0 - a_{0,x})/(1 + a_0^2)] + a_1(a_1 + 2a_0) = 0 ,$$

which is not a Miura transformation. The new equation involves both the variables a_1 and a_0.

Although this method provides a technique for determining the factorizations it is not very efficient because the determination of the gauge transformations T which define the ABT's and the singularity conditions involves considerable work. However, solution of the following problem provides an efficient algorithm for determining intermediate equations.

Miura problem: *Let* $\{P^0,Q^0\}$ *be an integrable system with zero curvature representation (1.1) and suppose it admits gauge transformations* $\{T^0(k)\}$ *which determine the ABT's. Then* $T^0 = \Delta^{-0}\Delta^{+0}$ *so that* Δ^{+0} *is unique even when* T^0 *is singular. Determine the Miura transformations defined by* $\{\Delta^{+0}\}$.

To see how this works consider

$$\Delta_0^{+0} := \Delta^+(a) = I + \sum_{j>i} a_{ij}E_{ij}, \quad P_0^0 = \sum_{i,j} p_{ij}e_{ij} ,$$

and for simplicity assume the nonzero p_{ij} are functionally independent. Then we get

$$P_0^1 = (\Delta_x^+(a) + \Delta^+(a)P_0^0)\Delta^+(a)^{-1}, \qquad (1.12)$$
$$Q_0^1 = (\Delta_t^+(a) + \Delta^+(a)Q_0^0)\Delta^+(a)^{-1} .$$

A Miura transformation is a special Bäcklund transformation of type II. It is defined by a set of relations which can be resolved for the nonzero p_{ij} into the form

$$p_{ij} = F_{ij}(a_{lm}, a_{lm,x}, \ldots, a_{lm,Nx}) , \quad N \geqslant 1 . \qquad (1.13)$$

The number of variables a_{lm} that appear as arguments of the F_{ij} is the same as the number of nonzero variables p_{ij}. The first relation in (1.12) defines first-order transformations of this type. They are obtained by equating to zero all the entries in P_0^1 that involve a derivative $a_{lm,x}$. This gives a set of linear equations for the p_{ij}. The Miura transformations are the consistent solutions of this system of equations. Since usually the p_{ij} are functionally related (i.e., the integrable equation is defined on a jet bundle) several cases arise. The first intermediate equation is determined from $(P_0^1,Q_0^1) \in \{P^1,Q^1\}$ or directly from (1.13) and $\{P^0,Q^0\}$. Note that Q_0^1 is obtained in the same way as P_0^1 from the second relation in (1.12). Usually, however, it is the P equation that contains the undifferentiated functions which define the Miura transformation. It is certainly true for the cases we consider. If the relations obtained from the second equation in (1.12) have the form ($q_{ij} \neq 0$)

$$q_{ij} = G_{ij}(a_{lm}, a_{lm,t}, \ldots, a_{lm,Mt}), \quad M \geqslant 1 , \qquad (1.14)$$

then since the q_{ij} are functionals of the p_{ij}, the Miura transformation can also be used to determine the first intermediate equation from (1.14).

C. AKNS ($r = -1$)

The relations (1.12) are, with $\Delta_0^{+0} = \Delta^+(a_0)$,

$$P_0^1 = (k - a_0)h + p_2e - f ,$$
$$Q_0^1(A_0 + a_0C_0)h + q_2e + C_0f ,$$

where

$$p_2 = a_{0,x} + q_0 - 2ka_0 + a_0^2,$$
$$q_2 = a_{0,t} + B_0 - 2A_0a_0 - a_0^2C_0 .$$

The Miura transformation is obtained from $p_2 = 0$. The equation $q_2 = 0$ gives, with the use of the Miura transformation, the first intermediate equation. This is the same as that obtained from $(P_0^1,Q_0^1) \in \{P^1,Q^1\}$ with $P_0^1 = (k - a_0)h - f$, $Q_0^1 = (A_0 + a_0C_0)h + C_0f$ given earlier.

Proposition 2: A solution of the Miura problem provides necessary and sufficient conditions for the existence of a solution to the factorization problem.

Proof: Let $(P_0^1,Q_0^1) \in \{P^1,Q^1\}$ be defined by a solution to the Miura problem for a given $(P_0^0,Q_0^0) \in \{P^0,Q^0\}$ and Δ^+. Suppose $(P_1^0,Q_1^0) \in \{P^0,Q^0\}$ is another solution and determine Δ^- algebraically (i.e., derivatives are treated as independent variables) as a solution of

$$\Delta^- P_0^1 = -\Delta_x^- + P_1^0\Delta^-, \quad \Delta^- Q_0^1 = -\Delta_t^- + Q_1^0\Delta^- .$$

Then if a solution exists $S = \Delta^-\Delta^+$ is singular by Proposition 1 (since all viable factorizations are singular). In particular, let $T_0^0(k):(P_0^0,Q_0^0) \rightarrow (P_1^0,Q_1^0)$ and let $T_0^0(k) = \Delta_0^{-0}\Delta_0^{+0}$ be its Gauss decomposition. Then a solution of the factorization problem exists if $\Delta_0^{+0},\Delta_0^{-0}$ can be made to satisfy the same conditions as Δ^+,Δ^-.

It was mentioned earlier that we obtain an equation of the required type only for low orders of factorization. However, in I we indicated a procedure whereby the class of integrable equations associated with a given seed equation by factorization can be enlarged.

Let $(P_0^0(k_i),Q_0^0(k_i))\in\{P^0,Q^0\}$, $i=1,\ldots,m$; that is, we choose a set of zero curvature representatives that give the same solution of $\{P^0,Q^0\}$. Then we can canonically embed this collection of representatives in a zero curvature representation associated with GL$(nm,\mathbb{C})$ by putting

$$Y_x=\Big(\bigoplus_{i=1}^{m} P_0^0(k_i)\Big)Y, \quad Y_t=\Big(\bigoplus_{i=1}^{m} Q_0^0(k_i)\Big)Y.$$

The additional degrees of freedom admitted by the larger problem allow new classes of equations (i.e., involving more than one parameter) connected to the seed equation by factorization to be found. A further extension would be to admit the derivative equations satisfied by $Y_{(j)}(k):=\partial^j/\partial k^j \times Y(k), j=1,\ldots,l$, into the scheme. We do not consider this aspect of the theory in this paper.

It is also possible to apply a suitably modified form of factorization to supersymmetric extensions of zero curvature representations. For example, the supersymmetric extension of the AKNS scheme has recently been developed by Gürses and Oğuz.[5]

Thus the Miura transformation for the super KdV (Ref. 6),

$$q_t = -q_{3x} - 6qq_x - 12(\epsilon\epsilon_x)_x,$$
$$\epsilon_t = -4\epsilon_{3x} - 6q\epsilon_x - 3q_x\epsilon,$$

is given by

$$q = -a_{1,x} - 2ka_1 + a_{2,x}a_2 - a_1{}^2,$$
$$\epsilon = -a_{2,\bar{x}}ka_2 - a_1a_2,$$

where a_1 is an even and a_2 an odd Grassman algebra valued variable. The next factorization also yields a transformation of Miura type, but the transformation cannot be brought into the canonical form (1.10).

The intermediate equations for the AKNS system were first derived by Chen[7] by introducing projective coordinates. If we consider the cases (i)–(iii) given above, then if $(y_1^1,y_2^1)^t$, $(y_1^0,y_2^0)^t$ are, respectively, the first column vectors of the solutions Y_0^1,Y_0^0, where $Y_0^1=\Delta^+(a_0)Y_0^0$, we get $y_1^1/y_2^1=(y_1^0/y_2^0+a_0)$. It is clear that $(y_1^1=0,y_2^1)^t$ is a solution of the zero curvature representation for the first intermediate equation for which the corresponding solution is $a_0=-y_1^0/y_2^0$.

The loop group associated with the zero curvature representation of the AKNS system is GL$(2,\mathbb{C}[k^{-1},k])$. We also have that $T_i^0\in$GL$(2,\mathbb{C}[k^{-1},k])$. However, any of the standard decomposition theorems for loop groups such as the Birkhoff or Bruhat decomposition are trivial because T_i^0 as an element of GL$(2,\mathbb{C}[k^{-1},k])$ is already "upper triangular." It is interesting to note that the factorizations $\Delta^+(a_0),\Delta^-(b_0,c_0,d_0)$ are elements of the completion of GL$(2,\mathbb{C}[k^{-1},k])$ since they involve formal series expansions of the type

$$(1-k)^{-1}=\sum_{i\in Z_+} k^i.$$

II. THIRD-ORDER SCALAR LAX EQUATIONS

Consider the third-order scalar Lax equation

$$L=\partial_x^3+u_1\partial_x+u_0. \tag{2.1}$$

If we choose

$$A=\partial_x^2+\tfrac{2}{3}u_1, \tag{2.2}$$

then the Lax equation $L_t=[A,L]$ gives the Boussinesq equation,[8] when $u_0=\frac{1}{2}v_x+w$, $u_1=v$,

$$v_{1,tt}=-\tfrac{1}{3}(v_{4x}+2(v^2)_{2x}).$$

In this case we will apply the factorization method to a first-order system associated with $L\psi=k^3\psi$,

$$Y_x=P_0^0Y,$$
$$P_0^0:=(u_0-k^3)E_{13}-E_{21}+u_1E_{23}-E_{32}. \tag{2.3}$$

This is the adjoint of the usual first-order system. Provided we allow arbitrary $T(k)$ to define an ABT gauge transformation and not just those linear in k, this system admits an ABT. In I we transformed the system to one that admitted a linear $T(k)$. Unfortunately this introduced cubic roots of unity into the problem and also led to some unnecessary restrictions.

We assume that (2.3) is part of a zero curvature representation of an integrable equation $\{P^0,Q^0\}$. The equation set is clearly nonempty since it contains the Boussinesq equation with Q_0^0 derived from the operator A above. The reason why we leave Q_0^0 free is because factorization leads to a number of canonical forms for P_0^0, and the Q_0^0 derived from A is too restrictive to apply to all cases.

We determine P_0^1 from the first equation in (1.9) using $\Delta^+(a):=I+a_1E_{12}+a_2E_{13}+a_3E_{23}$. This gives

$$\begin{aligned} P_0^1 = {} & -a_1E_{11}+p_2E_{12}+p_3E_{13}\\ & -E_{21}+(a_1-a_3)E_{22}+p_6E_{23}-E_{32}+a_3E_{33},\\ p_2 = {} & a_{1,x}-a_2+a_1^2, \qquad\qquad (2.4)\\ p_3 = {} & a_{2,x}+a_1u_1+u_0-k^3+a_1a_2,\\ p_6 = {} & a_{3,x}+u_1+a_2-u_3(a_1-a_3). \end{aligned}$$

If we apply the theory of the preceding section then $p_2=0$, $p_3=0$, $p_6=0$. For given sufficiently smooth functions u_0, u_1, an initial value problem for the resulting system of inhomogeneous equations has, for fixed t, a unique solution. Conversely the system can be resolved algebraically to give a unique solution for u_0 and u_1 and so defines a Miura transformation for the system

$$\begin{aligned} a_2 &= a_{1,x}+a_1^2,\\ u_1 &= -a_{3,x}-a_{1,x}-a_1^2-a_3^2+a_1a_3, \qquad\qquad (2.5)\\ u_0 &= k^3-a_{1,2x}-2a_1a_{1,x}+a_1a_{3,x}+a_1a_3^2-a_1^2a_3. \end{aligned}$$

If we put $a_1=y+z$, $a_3=y-z$, then this is the same transformation as obtained in Ref. 9 with $k=0$.

We observe that with this factorization the corresponding first intermediate function P_0^1 obtained from (2.3) does not depend explicitly on k. For this reason we use the less restrictive Bruhat decomposition of $T(k)$ and take $\Delta_B^+(a):=\mathrm{diag}(k^{-3},k^{-2},k^{-1})\Delta^+(a)$, $\Delta_B^-(a):=\Delta^-(a)$

$\times \operatorname{diag}(k^3,k^2,k)$. Then we obtain the same Miura transformation (2.3), but for P_0^1 we get

$$P_0^1 = -a_1E_{11} - kE_{21} + (a_1 - a_3)E_{22} - kE_{32} + a_3E_{33}\,. \tag{2.6}$$

We note that a restricted form of the Miura transformation is given by $k = 0$ in (2.5). In fact, this form of the Miura transformation can be obtained by a modification of our procedure without requiring $k = 0$ in P_0^1. The corresponding first intermediate zero curvature function is

$$P_{0(\mathrm{res})}^1 := P_0^1 - kE_{13}\,.$$

The second intermediate equations are associated with Gauss decomposition of ABT transformations $T(k)$ associated with zero curvature representations of $\{P^1,Q^1\}$ with P_0^1 defined by (2.6). In this case the resulting function P_0^2 still depends upon k, so that a more general factorization is not required. Put $\Delta_0^{+1} := \Delta^+(b)$ and we find in this case that

$$\begin{aligned} P_0^2 := {}& (-a_1 - b_1k)E_{11} - kE_{21} \\ & + (a_1 - a_3 - k(b_3 - b_1))E_{22} \\ & - kE_{32} + (a_3 + kb_3)E_{33}\,. \end{aligned}$$

The Miura transformation is determined from

$$\begin{aligned} & b_{1,x} + b_1(2a_1 - a_3) + k(b_1^2 - b_2) = 0\,, \\ & b_{2,x} + b_2(a_1 + a_3) + kb_1b_2 = 0\,, \\ & b_{3,x} + b_3(2a_3 - a_1) + kb_2 + kb_3(b_3 - b_1) = 0\,. \end{aligned} \tag{2.7}$$

It follows from the theory of Sec. I that we obtain a Miura transformation when $b_2 = 0$. There are several degenerate cases in the resolution of (2.7) for a_1 and a_3. They are (i) $a_1 = 0, a_3 \neq 0$; (ii) $a_3 = 0, a_1 \neq 0$; and (iii) $a_1 = \alpha a_3$, α a constant. An examination of (2.3) shows that for each case we require $k = 0$ and then we find that (i) $u_0 = 0$, (ii) $u_0 = u_{1,x}$, (iii) $u_0 = \frac{1}{2}u_{1,x}$ and $\alpha = 1$. These are just the cases that arise in the factorization of the scalar Lax operator L.[9,10] The scalar operators corresponding to (i) and (ii) are adjoint and give rise to the same hierarchy of integrable equations. The operator is usually called the Sawada–Kotera operator after the lowest order nontrivial integrable equation obtained from a Lax pair.[11] Similarly the operator corresponding to (iii) is the Kupershmidt operator.

The Miura transformation for the system $\{P^2,Q^2\}$ obtained from (2.7) is

$$\begin{aligned} & a_3 - 2a_1 = b_{1,x}/b_1 + kb_1\,, \\ & a_1 - 2a_3 = b_{3,x}/b_3 + k(b_3 - b_1)\,. \end{aligned} \tag{2.8}$$

We note that Eqs. (2.5) and (2.8) are the Miura transformations for the first and second factorizations even when specialization occurs in the seed equation.

If

$$\begin{aligned} \Delta_0^{-0} = \Delta^-(c) := {}& c_1E_{11} + c_2E_{21} \\ & + c_3E_{22} + c_4E_{31} + c_5E_{31} + c_6E_{33}\,, \end{aligned}$$

then necessary and sufficient conditions for the existence of a solution to the first factorization problem are obtained from Eqs. (1.9). Thus from the $\Delta_{0,x}^{-0}$ equation we get $c_3 = c_5 = c_6 = 0$ and

$$\begin{aligned} & c_{1,x} + (\bar{u}_0 + k^3)c_4 - c_1a_1 = 0\,, \\ & c_{2,x} + c_1 - \bar{u}_1c_4 + a_1c_2 = 0\,, \\ & c_{4,x} + c_2 - c_4a_1 = 0\,, \end{aligned}$$

where $(\bar{u}_0,\bar{u}_1)$ correspond to a second solution $(P_1^0,Q_1^0) \in \{P^0,Q^0\}$. These equations together with the set obtained from the $\Delta_{0,t}^{-0}$ equation give the conditions for the existence of a solution to the factorization problem. Thus if

$$T_0^0\colon\ (P_0^0,Q_0^0) \to (P_1^0,Q_1^0) \in \{P^0,Q^0\}$$

then T_0^0 factorizes and defines an intermediate equation $\{P^1,Q^1\}$ provided there exist values $k = k_c$ such that $T_0^0(k_c) = S_0^0(k_c) = \Delta_0^{-0}\Delta_0^{+0}$ and the components of Δ_0^{-0} satisfy the above conditions.

We write out explicitly the equations obtained by factorization of the Boussinesq system. The Boussinesq equation corresponds to $u_0 = \frac{1}{2}v_x + w$, $u_1 = v$ in the system below.

Seed equation: Boussinesq equation:

$$\begin{aligned} P_0^0 = {}& (u_0 - k^3)E_{13} - E_{21} + u_1E_{23} - E_{32}\,, \\ Q_0^0 = {}& -\tfrac{2}{3}u_1E_{11} + (u_0 - \tfrac{2}{3}u_{1,x} - k^3)E_{12} \\ & + (u_{0,x} - \tfrac{2}{3}u_{1,2x})E_{13} + \tfrac{1}{3}u_1E_{22} \\ & + (u_0 - k^3 - \tfrac{1}{3}u_{1,x})E_{23} - E_{31} + \tfrac{1}{3}u_1E_{33}\,, \\ & u_{0,t} + \tfrac{2}{3}u_{1,3x} - u_{0,2x} + \tfrac{2}{3}u_1u_{1,x} = 0\,, \\ & u_{1,t} + u_{1,2x} - 2u_{0,x} = 0\,. \end{aligned}$$

First intermediate equation: Modified Boussinesq equation:

$$\begin{aligned} P_0^1 = {}& -a_1E_{11} - kE_{21} + (a_1 - a_3)E_{22} - kE_{32} + a_3E_{33}\,, \\ Q_0^1 = {}& (-\tfrac{2}{3}u_1 - a_2)E_{11} - a_3kE_{21} + (\tfrac{1}{3}u_1 + a_1a_3)E_{22} \\ & - k^2E_{31} + a_1kE_{32} + (\tfrac{1}{3}u_1 + a_2 - a_1a_3)E_{33}\,, \\ & a_{1,t} + \tfrac{1}{3}(-a_{1,x} + 2a_{3,x} - a_1{}^2 + 2a_3{}^2 - 2a_1a_3)_{,x} = 0\,, \\ & a_{3,t} + \tfrac{1}{3}(2a_1a_3 + a_{3,x} - 2a_{1,x} + a_3{}^2 - 2a_1{}^2)_{,x} = 0\,, \end{aligned}$$

where $a_2 = a_{1,x} + a_1{}^2$.

Second intermediate equation: Modified-modified Boussinesq equation:

$$\begin{aligned} P_0^2 = {}& (-a_1 - b_1k)E_{11} - kE_{21} + (a_1 - a_3 \\ & - k(b_3 - b_1))E_{22} - kE_{32} + (a_3 + kb_3)E_{33}\,, \\ Q_0^2 = {}& (-\tfrac{2}{3}u_1 - a_2 - a_3b_1k)E_{11} \\ & + (-a_3k - k^2b_3)E_{21} + (\tfrac{1}{3}u_1 + a_1a_3 + kb_3a_1 \\ & + kb_1a_3 + k^2b_1b_3)E_{22} - k^2E_{31} + (ka_1 + k^2b_1)E_{32} \\ & + (\tfrac{1}{3}u_1 + a_2 - a_1a_3 - ka_1b_3 - k^2b_1b_3)E_{33}\,, \\ & 3q_t - 2p_{2x} - q_{2x} - q_x^2 - 2p_xq_x - 2kq_xe^q \\ & \quad + 2k(p_x + q_x)e^p = 0\,, \\ & 3p_t + p_{2x} + 2q_{2x} + p_x^2 + 2p_xq_x - 2kp_xe^p \\ & \quad + 2k(q_x + p_x)e^q = 0\,, \end{aligned}$$

where $b_1 = e^p$, $b_3 = e^q$.

At the next factorization general Bäcklund transformations are obtained.

The factorization process associates one-parameter families of equations with the original seed equation. By using the embedding process we can increase the number of parameters. We consider the canonical embedding of the Lax operator in sl(6,$\mathbb{C}$). Let

$$P_0^0 := P_0^0(k_1) \oplus P_0^0(k_2)\,, \tag{2.9}$$

where $P_0^0(k_i)$ is defined in (2.3). Let W be the row vector formed from the rows of the upper triangular matrix

$$I + \sum_{j>i} f_{ij} E_{ij} \in GL(6,\mathbb{C})\,.$$

Relabel the elements f_{ij} in W sequentially $a_1,a_2,\ldots,a_{15}$ so that $a_3 = f_{14}$ and $a_{11} = f_{35}$, for instance. Denote the relabeled triangular matrix by $\Delta^+(a)$. Determine the Miura transformations associated with P_0^0 and $\Delta^+(a)$. Then it is easy to see that the transformations are only the canonical ones obtained previously since $u_1 \neq 0$ and $u_0 \neq 0$ (all the other a_j's are zero):

$$\begin{aligned}
&a_2 = a_{1,x} + a_1^2\,,\\
&u_1 = -a_{6,x} - a_{1,x} - a_1^2 - a_6^2 - a_1 a_6\,,\\
&u_0 = k_1^3 - a_{1,2x} - 2a_1 a_{1,x} + a_1 a_{6,x} + a_1 a_6^2 - a_1^2 a_6\,,\\
&a_{14} = a_{13,x} + a_{13}^2\,,\\
&u_1 = -a_{15,x} - a_{13,x} - a_{13}^2 - a_{15}^2 - a_{13}a_{15}\,,\\
&u_0 = k_2^3 - a_{13,2x} - 2a_{13}a_{13,x} + a_{13}a_{15,x} + a_{13}a_{15}^2 - a_{13}^2 a_{15}\,.
\end{aligned} \tag{2.10}$$

The second factorization contains new possibilities. We find that with $P_0^1 := P_0^1(k_1) \oplus P_0^1(k_2)$ and $\Delta_0^{+1} := \Delta^+(b)$,

$$\begin{aligned}
P_0^2 := {} & (-a_1 - b_1 k_1)E_{11} - k_1 E_{21} + (a_1 - a_6\\
& - k_1(b_6 - b_1))E_{22} - k_1 E_{32} + (a_6 + k_1 b_6)E_{33}\\
& + (-a_{13} - k_2 b_{13})E_{44} - k_2 E_{54} + (a_{13} - a_{15}\\
& - k_2(b_{15} - b_{13}))E_{55} - k_2 E_{65} + (a_{15} + k_2 b_{15})E_{66}\,.
\end{aligned} \tag{2.11}$$

The Miura transformation consisting of four independent equations is to be resolved from the following system:

$$\begin{aligned}
&b_{1,x} - k_1 b_2 + b_1(2a_1 - a_6) + k_1 b_1^2 = 0\,,\\
&b_{2,x} + b_2(a_6 + a_1) + k_1 b_1 b_2 = 0\,,\\
&b_{3,x} - k_2 b_4 + b_3(a_1 - a_{13}) + k_1 b_1 b_3 = 0\,,\\
&b_{4,x} - k_2 b_5 + b_4(a_{13} - a_{15} + a_1) + k_1 b_1 b_4 = 0\,,\\
&b_{5,x} + b_5(a_1 + a_{15}) + k_1 b_1 b_5 = 0\,,\\
&b_{6,x} + k_1 b_2 + b_6(2a_6 - a_1) + k_1 b_6^2 - k_1 b_1 b_6 = 0\,,\\
&b_{7,x} + k_1 b_3 - k_2 b_8 - b_7(a_1 - a_6 + a_{13}) + k_1 b_6 b_7\\
&\qquad - k_1 b_1 b_7 = 0\,,\\
&b_{8,x} + k_1 b_4 - k_2 b_9 + b_8(a_{13} - a_{15} - a_1 + a_6) + k_1 b_6 b_8\\
&\qquad - k_1 b_1 b_8 = 0\,,\\
&b_{9,x} + k_1 b_5 + b_9(a_{15} - a_1 + a_6) + k_1 b_6 b_9 - k_1 b_1 b_9 = 0\,,\\
&b_{10,x} - k_2 b_{11} + k_1 b_7 - b_{10}(a_6 + a_{13}) - k_1 b_6 b_{10} = 0\,,\\
&b_{11,x} - k_2 b_{12} + k_1 b_8 + b_{11}(a_{13} - a_{15} - a_6) - k_1 b_6 b_{11} = 0\,,\\
&b_{12,x} + k_1 b_9 + b_{12}(a_{15} - a_6) - k_1 b_6 b_{12} = 0\,,\\
&b_{13,x} - k_2 b_{14} + b_{13}(2a_{13} - a_{15}) + k_2 b_{13}^2 = 0\,,\\
&b_{14,x} + b_{14}(a_{15} + a_{13}) + k_2 b_{13} b_{14} = 0\,,\\
&b_{15,x} + k_2 b_{14} + b_{15}(2a_{15} - a_{13}) + k_2 b_{15}^2 - k_2 b_{13} b_{15} = 0\,.
\end{aligned} \tag{2.12}$$

An obvious solution of these equations corresponds to the canonical embedding. However, there are other solutions. For example,

$$\begin{aligned}
&(a_1 - 2a_6)b_6 = b_{6,x} + k_1 b_6^2\,,\\
&(a_6 + a_{13})b_{10} = b_{10,x} - k_2 b_{11} - k_1 b_{10} b_6\,,\\
&(a_6 - a_{13} + a_{15})b_{11} = b_{11,x} - k_2 b_{12} - k_1 b_6 b_{11}\,,\\
&(a_6 - a_{15})b_{12} = b_{12,x} - k_1 b_6 b_{12}\,,
\end{aligned} \tag{2.13}$$

with all other b_i's zero.

This Miura transformation defines a two-parameter equation associated with the Boussinesq equation. By the methods of this paper we also obtain a zero curvature representation for it in sl(6,$\mathbb{C}$).

ACKNOWLEDGMENTS

The authors would like to acknowledge the kindness and generosity of colleagues and students at the Research Institute of Theoretical Physics, Hiroshima University.

The work was carried out while the first author was on a JSPS fellowship and the second author on a Monbusho fellowship.

[1] R. K. Dodd, J. Phys. A **21**, 931 (1988).
[2] V. E. Zakharov and A. B. Shabat, Funct. Anal. Appl. **13**, 160 (1979).
[3] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1974).
[4] A. R. Forsythe, *Theory of Differential Equations* (Dover, New York, 1959).
[5] M. Gürses and O. Oğuz, Lett. Math. Phys. **11**, 235 (1986).
[6] B. A. Kupershmidt, Phys. Lett. A **102**, 213 (1984).
[7] H. H. Chen, Phys. Rev. Lett. **33**, 925 (1974).
[8] V. E. Zakharov and A. B. Shabat, Funct. Anal. Appl. **8**, 226 (1974).
[9] A. P. Fordy and J. Gibbons, J. Math. Phys. **22**, 1170 (1981).
[10] B. A. Kupershmidt and G. Wilson, Invent. Math. **62**, 403 (1981).
[11] R. K. Dodd and J. D. Gibbon, Proc. R. Soc. London Ser. A **358**, 287 (1977).

On the construction of Hamiltonians

C. A. Lucey
Department of Physics, Jamestown Community College, Jamestown, New York 14701
E. T. Newman
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 28 April 1988; accepted for publication 29 June 1988)

It is shown that, for the system of first-order ordinary differential equations of the form $dx^a/dt = f^a(x^b)$, it is always possible to construct a locally defined nondegenerate symplectic structure $\Omega_{ab}(x)$, and write the equations as the canonical equations for some Hamiltonian function $H(x)$ such that $f^a\Omega_{ab} = \nabla_b H$. Furthermore, a Lagrangian $L(x,\dot{x})$, which is linear in the velocities dx^a/dt, can be found, whose variation yields these same canonical equations. Finally we discuss, via the Dirac brackets, how the standard transition from a Lagrangian system to a Hamiltonian system works in this pathological case.

I. INTRODUCTION

We consider a system of ordinary first-order differential equations of the form

$$\frac{dx^a}{dt} \equiv \dot{x}^a = f^a(x^b) \tag{1.1}$$

on a $2N$-dimensional[1] manifold $\mathcal{M}$. There appears to be a common misconception in the literature,[2,3] namely, that systems of equations of the form (1.1), with dissipation, do not possess a Hamiltonian structure. It is one of the purposes of this paper to point out that this is not true; i.e., all systems of the form (1.1) do possess (locally or frequently globally on the original manifold or on the covering space) a Hamiltonian structure. Though this fact has already been pointed out,[4] the issue continues to be misunderstood in the discussions of chaotic systems. We thus felt it was appropriate to discuss it once more.

We will show that $\mathcal{M}$ can be given (locally) a symplectic structure and (1.1) can be written as the canonical equations of motion for some Hamiltonian $H(x)$ and the associated nondegenerate symplectic form Ω_{ab} so that it becomes

$$\dot{x}^a = \nabla_b H \Omega^{ba} . \tag{1.2}$$

Furthermore we show that there is a simple (but in general pathological) Lagrangian $L(x,\dot{x})$ given as a function on the tangent bundle of $\mathcal{M}$, which is linear in the velocities $\dot{x}^a$, whose Euler–Lagrange equations are essentially (1.2), namely,

$$\dot{x}^a\Omega_{ab} = \nabla_b H . \tag{1.3}$$

Note that *in this pathological situation* $\mathcal{M}$ *is now* playing the role of a $2N$-dimensional configuration space. In the attempt to go, in the standard fashion, from this Lagrangian based on the $2N$-dimensional configuration space to a Hamiltonian formalism, the new phase space should be $4N$ dimensional. We show that, using the Dirac[5,6] theory of constraints (the constraints arising from the pathology of the Lagrangian), the resulting phase space is again the $2N$-dimensional manifold $\mathcal{M}$ and the canonical equations are still (1.3).

We illustrate the above ideas with several examples including the case of the one-dimensional damped harmonic oscillator.

The problem discussed here arose in our attempts[7] to understand how to construct a quantum theory associated with classical equations of motion for which there does not exist a standard Lagrangian version. Physical examples of systems of this type abound, as for example, all Yang–Mills theories based on groups for which there does not exist a nondegenerate quadratic form. It becomes clear that even in these cases Hamiltonians do exist—but what to do with them is not so clear.

II. THE HAMILTONIAN AND SYMPLECTIC STRUCTURE

On the manifold $\mathcal{M}$, we wish to introduce a (local) symplectic structure [making $\mathcal{M}$ (locally) into a phase space] in the following fashion. We require a two-form Ω_{ab} to satisfy the following conditions.

(1) It must be Lie-dragged by the vector field f^a, i.e.,

$$\mathcal{L}_f \Omega_{ab} \equiv f^c\nabla_c\Omega_{ab} + \Omega_{ac}\nabla_b f^c + \Omega_{cb}\nabla_a f^c = 0 . \tag{2.1}$$

(2) It must be nondegenerate, i.e., it must have a unique inverse Ω^{ab} with

$$\Omega_{ab}\Omega^{bc} = \delta_a{}^c .$$

(3) It must be closed, i.e., $\nabla_{[c}\Omega_{ab]} = 0$, and hence be locally exact. Thus

$$\Omega_{ab} = \nabla_a\Theta_b - \nabla_b\Theta_a$$

for some one-form Θ_a which plays an important role in what follows.

It is easily seen that such a two-form does exist, at least locally. By introducing coordinates so that the integral curves of (1.1) are constant on $2N-1$ of them and on the last $x^1 = t$, the field f^a takes the simple form $f^a = \delta_1{}^a$. Equation (2.1) is then satisfied by a constant Ω and thus the remaining conditions are easily satisfied.

Once a symplectic form is found, a Hamiltonian $H(x)$ can be defined by

$$\nabla_b H = f^a\Omega_{ab} \tag{2.2}$$

or equivalently by

$$H = f^a\Theta_a .$$

0022-2488/88/112430-04$02.50

The integrability conditions are satisfied by (2.1) and the fact that Ω_{ab} is closed. We thus have

$$\dot{x}^a = \nabla_b H \Omega^{ba} . \tag{2.3}$$

[Note that though we are using the original coordinates x^a on $\mathscr{M}$ given with f^a, we could have switched to canonical coordinates $\hat{x}^a = (q^i, p_j) = \hat{x}^a(x)$. There is, however, no compelling reason to do so.]

III. THE LAGRANGIAN

Given a Hamiltonian, the standard variational method to obtain the canonical equations is to vary the p's and q's independently in the expression

$$\hat{L}(q,p) = \dot{q}^i p_i - H(q,p) . \tag{3.1}$$

If, in the canonical equations $\dot{q}^i = \partial H / \partial p_i$, one can solve for $p_i = p_i(q, \dot{q})$, then by substitution into (3.1) a standard Lagrangian $L(q, \dot{q})$ is produced. If, however, this cannot be done, we have a pathological situation analogous to the problem of passing from a Lagrangian to a Hamiltonian when $p_i = \partial L / \partial \dot{q}$ cannot be solved for $\dot{q}(q,p)$.

In general one expects this pathological situation. (For example, if the Hamiltonian is linear in one or more of the p's this problem occurs.)

Using our arbitrary coordinates x^a, our version of (3.1) is[8]

$$L(x, \dot{x}) = \dot{x}^a \Theta_a(x) - H(x) . \tag{3.2}$$

It is easily seen that the variation of the x^a independently in (3.2) yields Eq. (1.3), namely, our canonical equations

$$\dot{x}^a \Omega_{ab} = \nabla_b H . \tag{3.3}$$

As an alternative point of view we can think of (3.2), with $\Theta_a(x)$ and $H(x)$ as given functions of x, as a Lagrangian (linear in the velocities) on the tangent bundle of the *configuration* space $\mathscr{M}$ with standard Euler–Lagrange equations also given by (3.3). We can then ask the question of how do we pass from the Lagrangian system with a $2N$-dimensional configuration space $\mathscr{M}$ to the $4N$-dimensional Hamiltonian system—and what type of equations do we obtain?

IV. THE HAMILTONIAN AND THE DIRAC BRACKETS

Using (3.2) as a Lagrangian immediately produces a pathological Hamiltonian system[8] since the defining equations for the canonical momenta

$$p_a = \frac{\partial L}{\partial \dot{x}^a} = \Theta_a(x) \tag{4.1}$$

cannot be solved for the velocities, $\dot{x}^a$. In the language of Dirac's theory of constraints they become the primary constraints of the Hamiltonian system:

$$C_a = p_a - \Theta_a(x) \approx 0 . \tag{4.2}$$

Notice that, because the Poisson brackets between the constraints become

$$\{C_a, C_b\} = \Omega_{ab} \tag{4.3}$$

and because of the nondegeneracy of the Ω_{ab}, the constraints are all second class and hence can be eliminated, i.e., all p_a can be replaced by $\Theta_a(x)$ and the Poisson brackets replaced by Dirac brackets. (This reduces the $4N$-dimensional phase space by $2N$ dimensions to the constraint surface, which is our original manifold.) The Dirac brackets between the two functions A and B are defined (in this case) by the following expression:

$$\{A,B\}^* = \{A,B\} - \{A,C_a\}\Omega^{ab}\{C_b,B\} . \tag{4.4}$$

Notice that if B is a constraint, say, C_d, then

$$\begin{aligned}\{A,C_d\}^* &= \{A,C_d\} - \{A,C_a\}\Omega^{ab}\{C_b,C_d\} \\ &= \{A,C_d\} - \{A,C_a\}\Omega^{ab}\Omega_{bd} = 0 ,\end{aligned}$$

i.e., the constraints have a vanishing Dirac bracket with all functions, including the Hamiltonian.

Formally the new Hamiltonian becomes

$$\begin{aligned}H_T &= \dot{x}^a p_a - L(x^a, \dot{x}^a) \\ &= \dot{x}^a(p_a - \Theta_a(x)) + H(x) = \dot{x}^a C_a + H(x)\end{aligned} \tag{4.5}$$

and thus on the constraint surface

$$H_T = H(x) . \tag{4.6}$$

The canonical equations of motion are then, in the Dirac theory,

$$\begin{aligned}\dot{x}^d = \{x^d,H\}^* &= \{x^d,H\} - \{x^d,C_a\}\Omega^{ab}\{C_b,H\} \\ &= 0 - \delta^d{}_a \Omega^{ab} \frac{\partial H}{\partial x^b}\end{aligned}$$

or

$$\dot{x}^d = \nabla_b H \Omega^{bd} , \tag{4.7}$$

our equations of motion, (1.3). The symplectic structure associated with the Dirac bracket is thus our original symplectic structure Ω_{bd}.

V. EXAMPLES

A. The damped harmonic oscillator

The damped harmonic oscillator equation, $\ddot{x} + \alpha \dot{x} + \beta x = 0$, can be written as

$$\dot{x} = y \equiv f^1, \quad \dot{y} = -\alpha y - bx \equiv f^2 . \tag{5.1}$$

Equation (2.1) then becomes

$$y\Omega_{,x} - (\alpha y + \beta x)\Omega_{,y} = \alpha\Omega , \tag{5.2}$$

with $\Omega \equiv \Omega_{xy} = -\Omega_{yx}$. It can be easily solved by the method of characteristics yielding a (particular) solution

$$\Omega = [(\omega_1 x - y)/(\omega_1 - \omega_2)]^\nu , \tag{5.3}$$

with ω_1 and ω_2 defined by $\alpha = -(\omega_1 + \omega_2)$ and $\beta = \omega_1\omega_2$ and $\nu = \alpha/\omega_2$.

The equations [(2.2)] to determine the Hamiltonian are

$$\Omega y = \frac{\partial H}{\partial y} \quad \text{and} \quad \Omega(\alpha y + \beta x) = \frac{\partial H}{\partial x} ,$$

with a solution

$$H(x) = [\omega_2/(\omega_2 - \omega_1)](\omega_1 x - y)(\omega_2 x - y)\Omega . \tag{5.4}$$

In the limit of no damping, with $\alpha \to 0$ and $\omega_1 \to -\omega_2 \to i\omega$, we have that

$$\Omega \to 1 \quad \text{and} \quad H \to \tfrac{1}{2}(y^2 + \omega^2 x^2) , \tag{5.5}$$

the undamped symplectic form and Hamiltonian.

Finally the one-form Θ_a is found to be

$$\Theta_x = (\omega_2/2\omega_1)(y-\omega_1 x)\Omega + \Phi_{,x}\,,$$

$$\Theta_y = (\omega_2/2\omega^2{}_1)(y-\omega_1 x)\Omega + \Phi_{,y}\,,$$

with Φ an arbitrary function of x and y.

Indeed it is a straightforward matter to show that the general class of time-independent second-order equations of motion of the form

$$\ddot{x} = F(x,\dot{x}) \tag{5.6}$$

can be supplied trivially with a symplectic form and a Hamiltonian. Simply write the equation as a two-dimensional system:

$$\dot{x} = y = f^1\,, \tag{5.7}$$

$$\dot{y} = F(x,y) = f^2\,, \tag{5.8}$$

where their ratios for the orbit equation are

$$\frac{dy}{dx} = \frac{F(x,y)}{y}\,. \tag{5.9}$$

Assume the solution of the orbit equation is of the form $y = Y(K,x)$, where K is the constant of integration that specifies the orbit. Now substitute $y = Y(K,x)$ into (5.7) and integrate the resulting equation to obtain

$$x = X(K,t)\,, \tag{5.10}$$

$$y = Y(K,X(K,t))\,. \tag{5.11}$$

Consider these relations as describing a coordinate transformation on the orbit space from coordinates (x,y) to (q,p) where $q = t$ and $p = K$, so that

$$x = X(p,q), \tag{5.12}$$

$$y = Y(p,X(p,q))\,. \tag{5.13}$$

In these coordinates the equations of motion are now given by

$$\dot{p} = 0\,, \tag{5.14}$$

$$\dot{q} = 1\,. \tag{5.15}$$

The symplectic form is simply

$$\Omega_{ab}\, dx^a \wedge dx^b = dp \wedge dq\,, \tag{5.16}$$

and the Hamiltonian is $H = p$.

This argument is the two-dimensional version of the argument given earlier for the existence of the symplectic form.

B. Quadratically "damped" oscillator

Consider the oscillator with a quadratic velocity term

$$\ddot{x} + \alpha\dot{x}^2 + \beta x = 0\,. \tag{5.17}$$

Though the orbit equation can be solved explicitly we will turn to the equation for the symplectic form, (2.1), which becomes

$$y\frac{\partial\Omega}{\partial x} - (\alpha y^2 + \beta x)\frac{\partial\Omega}{\partial y} = 2\alpha y\Omega\,, \tag{5.18}$$

with solution

$$\Omega = -\exp 2\alpha x \tag{5.19}$$

and

$$H = \tfrac{1}{2}\exp(2\alpha x)(\beta/2\alpha^2 - y^2 - \beta x/\alpha)\,. \tag{5.20}$$

One can easily show from (5.20) that for positive H the orbits are periodic while for negative values, the solution is runaway.

C. The ice skate

Sorkin[9] has suggested that we use this technique to find a Hamiltonian for the equations of motion of a nonholonomically constrained system. He suggested that we consider the case of an ice skate, a system with three degrees of freedom. In this system, (x,y) locates the center of mass of the skate in the (x,y) plane and Θ determines its orientation with respect to the x axis. If I is the moment of inertia, m the mass, and λ a Lagrange multiplier, the equations of motion are

$$I\ddot{\Theta} = 0, \quad m\ddot{x} = -\lambda\sin\Theta, \quad m\ddot{y} = \lambda\cos\Theta\,, \tag{5.21}$$

with the nonholonomic constraint

$$\dot{x}\sin\Theta = \dot{y}\cos\Theta\,.$$

Eliminating the Lagrange multiplier [via $\lambda = m\dot{\Theta}(\dot{x}\cos\Theta + \dot{y}\sin\Theta)$] and rewriting the above equations, with $(x^1,x^2,x^3,x^4) = (\Theta,\dot{\Theta},x,\dot{x})$, we have

$$\dot{x}^1 = x^2, \quad \dot{x}^2 = 0, \quad \dot{x}^3 = x^4, \quad \dot{x}^4 = -x^2x^4\tan x^1\,.$$

The symplectic form is

$$\Omega = \begin{vmatrix} 0 & Sx^4/x^2 & 0 & -\sec x^1 \\ -Sx^4/x^2 & 0 & 1 & -T/x^2 \\ 0 & -1 & 0 & 0 \\ \sec x^1 & T/x^2 & 0 & 0 \end{vmatrix}\,,$$

with $S = (\sec^2 x^1 - \sec x^1 - x^1\sec x^1\tan x^1)$ and $T = (\tan x^1 - x^1\sec x^1)$ and

$$H = -x^2x^4\sec x^1\,.$$

D. Chaos

One of the underlying motivations for the work described here was eventually to try to study the relationship of a classical chaotic system with its "associated" quantum system. In order to do this we had hoped that it would be possible to take a relatively "simple" system, like the Lorenz attractor equations, and find its symplectic and Hamiltonian structure as a prelude to studying its quantum properties. However, we have been unable to explicitly integrate Eq. (2.1) for the symplectic form (aside from the special case of geodesics on compact, constant negative curvature surfaces[10] where the Hamiltonian structure is readily available). The basic impediment to this seems to be the impossibility of finding exact solutions to the equations of motion for the chaotic system. It thus appears as if the transformation to canonical coordinates from the "physical" coordinates is, itself, chaotic. It could well be that the classical and quantum Hamiltonian versions of most chaotic systems are simple and standard, with the *difficulties being in the relationship between the physical variables and the canonical variables.* It would thus seem that in the *definition* of a chaotic system, one must have a choice of coordinate chart.

VI. DISCUSSION

We have seen that equations of the form (1.1) always have a local symplectic and Hamiltonian structure. We do not know and have been unable to find any references on the conditions required to make the symplectic and Hamiltonian structure global.

An additional question pertains to the quantization of an arbitrary Hamiltonian system obtained in this way. In general, such Hamiltonians do not seem to be amenable to quantization. This question has not been carefully analyzed. These Hamiltonians are not, in general, interpretable as the energy (though they are conserved) and are usually nonpolynomial expressions.

We are extending these ideas to field theory and, in particular, to cases where there is no obvious Lagrangian, e.g., to self-dual Yang–Mills theory.

ACKNOWLEDGMENTS

We wish to thank M. Henneaux for valuable comments and for drawing our attention to his beautiful paper on the inverse problem of the calculus of variation where he deals with a similar problem to that addressed here. Above all we thank Abhay Ashtekar for pointing out the way for us.

This work was supported by the National Science Foundation under Grants No. Phy-8609897 and No. Phy-80023.

[1] If (1.1) is given for odd dimensions, one simply adds to the system another noninteracting equation, for example, $\dot{z} = 0$, in order to make it even dimensional.

[2] C. Grebogi, E. Ott, and J. A. Yorke, Science **238**, 632 (1987).

[3] J. P. Eckermann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[4] M. Henneaux, J. Math. Phys. **23**, 2101 (1982).

[5] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Yeshiva, New York, 1964); see also, Can. J. Math. **2**, 147 (1950).

[6] K. Sundermeyer, *Constrained Dynamics* (Springer, New York, 1982).

[7] A. Helfer, M. Hickman, C. Kozameh, C. Lucey, and E. T. Newman, Phys. Rev. D **36**, 1740 (1987).

[8] P. G. Bergmann and E. T. Newman, Phys. Rev. **99**, 587 (1955).

[9] R. Sorkin (private communication).

[10] M. C. Gutzwiller, Phys. Scr. **T9**, 184 (1985).

Wake-free cylindrical and spherical waves in inhomogeneous media

H. P. W. Gottlieb
School of Science, Griffith University, Nathan, Queensland 4111, Australia

(Received 2 September 1987; accepted for publication 13 July 1988)

Radius-dependent wave-speed functions for two- and three-dimensional inhomogeneous media are discovered that yield explicit, general, radially symmetric solutions that are wake-free (tailess). For the corresponding finite regions, the eigenfrequency spectra for radially symmetric vibrations with Dirichlet boundary conditions are shown to be exactly harmonic.

I. INTRODUCTION

It is well known[1–6] that the solutions to the ordinary wave equation in one and three dimensions possess basic differences from those in two dimensions. D'Alembert's general unidirectional traveling-wave solution in one dimension and the general spherically symmetric expanding progressive spherical-wave solution in three dimensions are wake-free (tailess, clean cut). If a source signal ceases, the disturbance at a finitely distanced point will also cease in a finite time, i.e., sharp signals remain sharp. In contrast, in two dimensions, for circular symmetry, an expanding circular wave (or cylindrical wave in three dimensions, independent of axial coordinate) does not have this property; there is an indefinitely prolonged tail (wake, reverberation, residual[7]).

This phenomenon can also be generalized in terms of the problem of radiation from a given spatially localized source, thereby relating it to the absence of an indefinitely prolonged residual effect at another point as a result of the source signal.[2–4,6,7] Thus traveling waves resulting from general localized sources in odd space dimensions are sharp, while those in even space dimensions are not.

The Hadamard conjecture[8] asserts that this overall situation also pertains for wave equations with nonconstant coefficients. Garabedian[9] has discussed the situation for odd and even n-dimensional spaces.

In this paper, we show that *for certain radius-dependent wave-speed functions in two dimensions, the general circularly symmetric expanding circular (cylindrical) wave solutions are wake-free.* A sharp symmetric signal at (a circle surrounding) the origin is transmitted as a sharp signal, with clean cutoff. For one inhomogeneous medium the solution is unattenuated, but there is radial spreading of the (outward) traveling waveform. For the other inhomogeneous medium, the solution also undergoes a change of magnitude, and wave fronts propagate to infinity in finite time.

These results might, at first sight, appear to be at variance with the above principles. However, the examples presented are solutions to wave equations equivalent via certain transformations to the one-dimensional constant-coefficient equation. Furthermore, although here we are concerned only with the explicit, general, radially symmetric solutions, these would not be the most general solutions to the corresponding full two-dimensional wave equation.

Despite the above-mentioned mathematical equivalence, the resulting radial wave equation formulation is, of course, physically distinct in that it represents a new two-dimensional medium. The solutions obtained are of considerable interest as a novel wave phenomenon for cylindrically symmetric systems.

A by-product of our approach is the discovery, also in the three-dimensional case, of a radial wave-speed function, of an *inhomogeneous* medium whose wave equation has a wake-free expanding spherical-wave solution, which may be unattenuated, and whose wave front propagates to infinity in a finite time.

This paper may therefore also be regarded as an investigation of operators that allow "relatively undistorted" progressing wave families, as defined by Courant.[10]

Finally, we discuss the eigenfrequencies for radially symmetric vibrations of the appropriate regions of finite extent in these inhomogeneous media with Dirichlet boundary conditions. These spectra are shown to be exactly harmonic.

II. WAKE-FREE CIRCULAR (CYLINDRICAL) WAVES

It is possible to find nonhomogeneous two-dimensional media in which circularly expanding waves are wake-free. To demonstrate this, we make use of the simple fact that solutions to the one-dimensional wave equation with spatially dependent wave speed,

$$c_1{}^2(x)u_{xx} = u_{tt}, \tag{2.1}$$

yield solutions to the circularly symmetric two-dimensional wave equation (with radially dependent wave speed)

$$c_2{}^2(\rho)[v_{\rho\rho} + (1/\rho)v_\rho] = v_{tt}, \tag{2.2}$$

through the relations

$$x/L = \ln(\rho/R), \tag{2.3a}$$

$$v(\rho) = u(x), \tag{2.3b}$$

$$c_2(\rho) = (\rho/L)\, c_1(x). \tag{2.3c}$$

Here, L and R are reference lengths included for dimensional reasons, and origins are chosen so that $\rho = R$ when $x = 0$.

Equation (2.2) also pertains to cylindrical waves in three dimensions with independence from axial coordinate z.

A. Unattenuated wake-free waves

For constant $c_1 = C_1$, Eq. (2.1) has the standard D'Alembert traveling-wave general solutions

$$u(x,t) = f(x - C_1 t) + g(x + C_1 t), \tag{2.4}$$

where f and g are arbitrary functions. Then, setting $\nu = C_1/L$ and using Eqs. (2.3), we obtain for Eq. (2.2) with wave-speed function

$$c_2(\rho) = \nu\rho, \tag{2.5}$$

the following general solution:

$$v(\rho,t) = \phi[\ln(\rho/R) - \nu t] + \psi[\ln(\rho/R) + \nu t], \qquad (2.6)$$

where ϕ and ψ are arbitrary functions, and R is retained for dimensional reasons.

The solution involving ϕ represents a general circularly symmetric expanding circular (or cylindrical) wave that is unattenuated and wake-free in this inhomogeneous medium. Circularly symmetric sharp signals emanating from a circle surrounding the origin would result in sharp signals at any receiver point. Of course, the result is not "reciprocal," in the sense that a sharp signal from (a circle surrounding) a point that is not at the origin of this medium would not be propagated in the form (2.6).

Equation (2.2) with (2.5) corresponds, for instance, to waves on a membrane with an inverse-square-radial areal density function

$$\sigma(\rho) \equiv \tau/c_2{}^2 = (\tau/\nu^2)/\rho^2, \qquad (2.7)$$

where τ is the (constant) membrane tension. Solutions (2.6) correspond to circle-front waves progressing with radial speed $c_2(\rho)$, Eq. (2.5). A signal emitted at $\rho = R$ takes time $(1/\nu)\ln(\rho/R)$ to reach radius $\rho > R$, and a sharp signal at $\rho = R$ has the same duration at any receiver point. There is no "tail" or residual disturbance.

The solution (2.6) may also be written in the form

$$v(\rho,t) = \phi_2(\rho e^{-\nu t}) + \psi_2(\rho e^{\nu t}). \qquad (2.6')$$

The waves are unattenuated in amplitude, but the outward-traveling waves (first term) are spread out radially as t increases, whereas the inward-travelling waves are compressed.

B. Another wake-free medium

As found by Synge[11] (cf. Lewis[12]) there is, in fact, one other special form of the wave speed in one dimension for which the general solution to (2.1) can be expressed in terms of arbitrary functions involving progressive waves. If

$$c_1(x) = (Ax + B)^2 \qquad (2.8)$$

(with $A \neq 0$), then (2.1) has general solution[11,13]

$$u(x,t) = (Ax + B)\{f[1/(Ax + B) + At] + g[1/(Ax + B) - At]\}. \qquad (2.9)$$

[Waves such as (2.9) with a specific factor outside the arbitrary functions have been termed "relatively undistorted."[10]] Hence here setting $\alpha = A\sqrt{L}$ and $\beta = B/\sqrt{L}$, we obtain [from Eqs. (2.3)] for Eq. (2.2) with wave-speed function

$$c_2(\rho) = \rho[\alpha \ln(\rho/R) + \beta]^2, \qquad (2.10)$$

the general solution

$$v(\rho,t) = [\alpha \ln(\rho/R) + \beta]\{\phi[1/(\alpha \ln(\rho/R) + \beta) + \alpha t] + \psi[1/(\alpha \ln(\rho/R) + \beta) - \alpha t]\}. \qquad (2.11)$$

The solution involving ϕ in (2.11) represents a relatively undistorted general circularly symmetric expanding circular wave of speed (2.10) that is wake-free. This heterogeneous medium corresponds to a membrane with an areal density function

$$\sigma(\rho) = (\tau/\rho^2)[\alpha \ln(\rho/R) + \beta]^{-4}. \qquad (2.12)$$

The solution (2.11) becomes indefinitely large as $\rho \to \infty$, just as for the Synge solution (2.9) for the heterogeneous one-dimensional (string) case[11] as $x \to \infty$. We find that a signal emitted at $\rho = R$ takes time

$$t_2 = (1/\beta)\ln(\rho/R)/[\alpha \ln(\rho/R) + \beta] \qquad (2.13)$$

to reach radius $\rho > R$, and the duration of a signal received at ρ is equal to that of a sharp signal at $\rho = R$. Furthermore, as $\rho \to \infty$, in (2.13), $t_2 \to 1/(\alpha\beta)$ which is finite. The wave front propagates to infinity in a finite time.

The solutions (2.6) and (2.11) obtained above for those special inhomogeneous media are in striking contrast to the general solutions of (2.2) for the homogeneous two-dimensional or cylindrical case $c_2 = C_2$, constant, which assume the "somewhat intractable" form of integrals[14]

$$v(\rho,t) = \int_0^\infty \phi(\rho \cosh \zeta - C_2 t)d\zeta + \int_0^\infty \psi(\rho \cosh \zeta + C_2 t)d\zeta, \qquad (2.14)$$

which imply the persistence of a "tail" even for a temporary source.

III. WAKE-FREE SPHERICAL WAVES IN AN INHOMOGENEOUS MEDIUM

The three-dimensional scalar wave equation for spherically symmetric waves with radius-dependent wave speed may be written as

$$c_3{}^2(r)\,(rw)_{rr} = (rw)_{tt}, \qquad (3.1)$$

whose solutions may be written down immediately from those of the one-dimensional wave equation (2.1) for inhomogeneous media by using the formal relations

$$rw(r) = u(r), \qquad (3.2a)$$

$$c_3(r) = c_1(r). \qquad (3.2b)$$

For constant $c_3 = C_3$, solutions (2.4) give the well-known attenuated progressive general spherical wave solutions

$$w(r,t) = (1/r)\,[F(r - C_3 t) + G(r + C_3 t)], \qquad (3.3)$$

where F and G are arbitrary functions.

One *inhomogeneous* three-dimensional medium may now be found that has general spherically symmetric wake-free spherical wave solutions. For the wave-speed function

$$c_3(r) = (Ar + B)^2 \qquad (3.4)$$

(with $A \neq 0$) analogous to (2.8), the Synge-type solutions (2.9) in the one-dimensional case give, via Eqs. (3.2), the general solution to (3.1):

$$w(r) = [A + B/r]\,\{F[1/(Ar + B) + At] + G[1/(Ar + B) - At]\}, \qquad (3.5)$$

where F and G are arbitrary functions. [Such a situation could be realized, for example, by $w = p$, the (excess) acoustic pressure in the linearized theory of sound propagation in inhomogeneous gases with radially varying sound speed.[15]]

The function F corresponds to a general spherically symmetric outwardly progressing wake-free wave of speed

(3.4). These waves, in general, undergo some distortion. However, in the case where $B = 0$, i.e., $c_3(r) \propto r^2$, we see that *these spherical waves are unattenuated throughout their propagation*.

A signal emitted at $r = a$ takes time

$$t_3 = (Aa + B)^{-1}(r - a)/(Ar + B) \tag{3.6}$$

to reach radius $r > a$. Sharp signals at $r = a$ are sharp at general radius r and are of the same duration. As $r \to \infty$ in (3.6), $t_3 \to [A(Aa + B)]^{-1}$, so the spherical wave front propagates to infinity in a finite time and, by (3.5), the amplitude remains finite.

IV. HARMONIC EIGENFREQUENCY SPECTRA

It is profitable also to investigate the eigenfrequency spectra that arise from imposing Dirichlet boundary conditions on a finite region, appropriate to the coordinates for the three inhomogeneous media discovered in the previous sections.

It is well known[16] that the frequency spectrum for radial vibrations of an annular membrane with fixed rims and inverse-square density [Eq. (2.7)] is exactly harmonic. This, in fact, follows directly from the standard constant-density string example and the transformations (2.3) with $c_1 = C_1$, constant.

It was shown by Borg[17] that the requirement of an exactly harmonic spectrum for a vibrating string with fixed ends requires (for suitably smooth functions) either constant wave speed (density) or squared-distance dependence of wave speed, Eq. (2.8), i.e., inverse fourth-power density.[18,19] Therefore, the radial vibrations of a fixed annular membrane with areal density function of the form (2.12) similarly possess an exactly harmonic spectrum. Explicitly, for a fixed annulus of inner radius R and other radius R_2 with wave speed (2.10), i.e., areal density function (2.12), the eigenfunctions are found to be given by

$$v_n(\rho,t) = \exp(-i\omega_n t)[\alpha \ln(\rho/R) + \beta] \times \sin\left[n\pi \frac{\ln(\rho/R)}{\ln(R_2/R)} \frac{\alpha \ln(R_2/R) + \beta}{\alpha \ln(\rho/R) + \beta}\right] \tag{4.1a}$$

($n = 1,2,\ldots$), with corresponding angular eigenfrequencies

$$\omega_n = n\pi\beta[\alpha \ln(R_2/R) + \beta]/\ln(R_2/R). \tag{4.1b}$$

Next, in three dimensions, it is a standard result that solutions of (3.1) with $c_3 = C_3$, constant, within a sphere on the surface of which $w = 0$, yield an exactly harmonic eigenspectrum. (This corresponds to a "pressure release" surface in the acoustical context.) It follows from the above-mentioned Borg result for wave speed (2.8) for the string and from Eq. (3.1) that the radial eigenfrequency spectrum for wave-speed function (3.4) with Dirichlet condition on the bounding spherical surface is also exactly harmonic.

The eigenfunctions for such a sphere of radius R_3 are found to be given explicitly by

$$w_n(r,t) = (A + B/r)\sin[n\pi(r/R_3)(AR_3 + B) \times (Ar + B)^{-1}]\exp(-i\omega_n t), \tag{4.2a}$$

with angular eigenfrequencies

$$\omega_n = n(\pi/R_3)B(AR_3 + B). \tag{4.2b}$$

V. CONCLUSIONS

We have found the solutions (2.6) and (2.11) that yield wake-free progressive circular (cylindrical) waves in inhomogeneous two- (three-) dimensional media. In contrast to the solutions (2.4), (2.6), and (3.3), the solutions (2.9) in one dimension, (2.11) in two dimensions, and the new wake-free three-dimensional solution (3.5) possess an overall multiplicative function factor outside the arbitrary functions that is the *general* radial solution to Laplace's equation for the correspondingly dimensioned radial symmetry.

Radial eigenfunctions and eigenfrequencies for Dirichlet boundary conditions with radial wave speeds (2.10) in two dimensions [annulus: Eqs. (4.1)] and (3.4) in three dimensions [sphere: Eqs. (4.2)] have been presented in Sec. IV: the spectra for these radially symmetric vibrations are harmonic. Thus the annulus with (2.10) is *isospectral,*[19] as far as its radial eigenfrequencies are concerned, with an annulus with wave speed (2.5) (for appropriate choice of parameters), and the sphere with (3.4) is radially isospectral with a sphere with constant wave speed.

Finally, it is intriguing to note that, as well as possessing the wake-free traveling-wave solutions (2.9) and a harmonic spectrum for a finite domain with Dirichlet boundary conditions as mentioned above, the squared-distance speed function (2.8) in the one-dimensional wave equation (2.1) is also singled out by possessing another property unique among space-dependent wave speeds, as shown by Bluman and Kumei[13]: the invariance group of the corresponding wave equation is infinite. In view of the transformations (2.3) and (3.2), respectively, and the findings of this paper, this unique infiniteness would also be the case for invariance groups for the two-dimensional radial wave equation (2.2) with wave speed (2.10) and the three-dimensional radial wave equation (3.1) with wave speed (3.4).

[1] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Part I, Secs. 6.1 and 7.3.

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Wiley, New York, 1962), Vol. II; R. Courant, *Partial Differential Equations* (Wiley, New York, 1962), Chaps. III and VI.

[3] P. R. Garabedian, *Partial Differential Equations* (Wiley, New York, 1964), Chap. 6.

[4] G. F. D. Duff and D. Naylor, *Differential Equations of Applied Mathematics* (Wiley, New York, 1966), Sec. 10.3.

[5] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974), Chap. 7.

[6] E. Zauderer, *Partial Differential Equations of Applied Mathematics* (Wiley, New York, 1983), Sec. 5.4.

[7] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Yale U.P., New Haven, 1923), Sec. 111.

[8] Reference 2, p. 765.

[9] Reference 3, Sec. 6.4.

[10] Reference 2, Chap. VI, Sec. 18.

[11] J. L. Synge, Q. Appl. Math. **39**, 292 (1981).

[12] J. T. Lewis, Q. Appl. Math. **38**, 461 (1981).
[13] G. Bluman and S. Kumei, J. Math. Phys. **28**, 307 (1987).
[14] H. Lamb, *Hydrodynamics* (Dover, New York, 1932), 6th ed., Sec. 302.
[15] Compare the formulation for the circular (cylindrical) case by J. E. Cole III, J. Acoust. Soc. Am. **81**, 222 (1987).
[16] A. B. Bassett, *An Elementary Treatise on Hydrodynamics and Sound* (Deighton Bell, Cambridge, 1900), p. 150.
[17] G. Borg, Acta Math. **78**, 1 (1946).
[18] W. Mothwurf, Mh. Math. **40**, 93 (1933).
[19] H. P. W. Gottlieb, J. Math. Anal. Appl. **132**, 123 (1988).

Supermultiplet of particles in a Coulomb potential: States and wave functions from the representation theory of OSp(2,1)

Louis Benoit and Luc Vinet
Laboratoire de Physique Nucléaire, Université de Montréal, C.P. 6128, Succ. A, Montréal, Québec H3C 3J7, Canada

(Received 2 April 1988; accepted for publication 20 July 1988)

A supermultiplet of two spin-0 and one spin-$\frac{1}{2}$ particle in a Coulomb potential has been shown recently to possess a dynamical OSp(2,1) symmetry algebra. The canonical chain of maximal subalgebras leads to a natural choice of quantum numbers and state vectors. The relevant representations of OSp(2,1) are constructed in this adapted basis, and the wave functions subsequently obtained by going to a coordinate realization.

I. INTRODUCTION

Over the last few years, the role of superalgebras as dynamical algebras in quantum mechanics has been well exemplified.[1] A case of special interest discussed by D'Hoker and Vinet[2] is that of two spin-0 and one spin-$\frac{1}{2}$ particle with electric charge $-\alpha/e$ in the field of dyons with electric charge e and magnetic charges, respectively, $(q\pm\frac{1}{2})/e$ and q/e. The Hamiltonian for such a system is given by

$$H_D=\frac{1}{2}\left[p_i-\left(q-\frac{1}{2}\Sigma\right)A_i^D\right]^2 -\frac{\alpha}{2r}+\frac{(\lambda-q)^2-q\Sigma+\frac{1}{4}\Sigma^2}{2r^2}-\frac{\lambda r^i\hat{S}^i}{r^3},\quad i=1,2,3, \tag{1.1}$$

where A_i^D is the vector potential for a magnetic monopole of unit strength,

$$\Sigma=\begin{pmatrix}\sigma_3 & 0\\ 0 & 0\end{pmatrix},\quad \hat{S}^i=\begin{pmatrix}0 & 0\\ 0 & \sigma^i/2\end{pmatrix}, \tag{1.2}$$

and λ and α are free parameters. Remarkably, as first demonstrated in Ref. 2, H_D admits an $\mathrm{OSp}(2,1)\oplus\mathrm{SU}(2)$ spectrum supersymmetry that allows for an algebraic resolution of the dynamics. The problem is to construct explicitly the relevant irreducible representations of the symmetry superalgebra. We address this question here. A natural representation basis has already been characterized in Ref. 2 by using, for quantum numbers, the eigenvalues of the Casimir operators associated to the following canonical chain of maximal subalgebras: $\mathrm{OSp}(2,1)\oplus\mathrm{SU}(2)\supset\mathrm{O}(2)\oplus\mathrm{O}(2,1)\oplus\mathrm{U}(1)\supset\mathrm{O}(2)\oplus\mathrm{O}(2)\oplus\mathrm{U}(1)$. In this paper, we complete the construction of the corresponding representations and obtain the action of the ladder operators in the given basis. We also compute the wave functions by going to a coordinate realization.

The paper is organized as follows. In Sec. II, we introduce a particular realization of $\mathrm{OSp}(2,1)\oplus\mathrm{SU}(2)$ in terms of differential operators in two complex variables. Using dimensional reduction techniques, we show in Sec. III how it arises as the spectrum-generating algebra for H_D. The symmetry-adapted state vectors are characterized in Sec. IV and the action of the dynamical group generators on these vectors are given in Sec. IV. The wave functions are the object of Sec. VI, and conclusions follow.

II. AN OSp(2,1) REALIZATION

The rank-2 OSp(2,1) superalgebra is eight dimensional. If we denote by R, $B_\pm$, and Y the four bosonic elements of a Cartan-like basis, the OSp(2,1) structure relation can be written as

$$[R,B_\pm]=\pm B_\pm,\quad [B_+,B_-]=-2R, \tag{2.1a}$$

$$[R,Y]=[B_\pm,Y]=0, \tag{2.1b}$$

$$\{F_\cdot^{L,R},F_\cdot^{L,R}\}=0,\quad \{F_\pm^L,F_\pm^R\}=B_\pm,$$
$$\{F_\pm^L,F_\mp^R\}=R\pm Y, \tag{2.1c}$$

$$[R,F_\pm^{L,R}]=\pm\tfrac{1}{2}F_\pm^{L,R},\quad [Y,F_\pm^L]=-\tfrac{1}{2}F_\pm^L,$$
$$[Y,F_\pm^R]=\tfrac{1}{2}F_\pm^R, \tag{2.1d}$$

$$[B_\pm,F_\pm^{L,R}]=0,\quad [B_\pm,F_\mp^{L,R}]=\mp F_\pm^{L,R}. \tag{2.1e}$$

The above algebra is realized as follows. Let z_a, $a=1,2$, denote two complex variables and introduce the 4×4 matrices

$$\eta_1=-\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1+\sigma_3\\ 1-\sigma_3 & 0\end{pmatrix},$$
$$\eta_2=-\frac{1}{\sqrt{2}}\begin{pmatrix}0 & \sigma_1+i\sigma_2\\ -\sigma_1-i\sigma_2 & 0\end{pmatrix}, \tag{2.2}$$

which satisfy

$$\{\eta_a,\eta_b\}=\{\eta_a^\dagger,\eta_b^\dagger\}=0,$$
$$\{\eta_a,\eta_b^\dagger\}=2\delta_{ab},\quad a,b=1,2. \tag{2.3}$$

(The symbol σ_i stands for the Pauli matrices.) Define

$$C=X+\Sigma, \tag{2.4a}$$

with

$$X=z_a\,\partial_a-\bar{z}_a\,\bar{\partial}_a\quad\text{and}\quad \Sigma=\begin{pmatrix}\sigma_3 & 0\\ 0 & 0\end{pmatrix}. \tag{2.4b}$$

It will also prove practical to use

$$\hat{S}^i=\frac{1}{4}\,\eta_a^\dagger\sigma_{ab}^i\eta_b=\begin{pmatrix}0 & 0\\ 0 & \sigma^i/2\end{pmatrix}. \tag{2.5}$$

Take

$$F_\pm^R=\mp(i/2)(\partial_a-\lambda\bar{z}_a/|z|^2\mp\bar{z}_a)\eta_a,$$
$$F_\pm^L=(F_\mp^R)^\dagger, \tag{2.6a}$$

$$R=\frac{1}{2}\left[-\partial_a\bar{\partial}_a+\frac{\lambda(\lambda-C)}{|z|^2}+z_a\bar{z}_a-\frac{2\lambda\bar{z}_a\sigma^i_{ab}z_b\hat{S}^i}{|z|^4}\right], \tag{2.6b}$$

$$B_\pm=\frac{1}{2}\left[(\partial_a\mp\bar{z}_a)(\bar{\partial}_a\mp z_a)-\frac{\lambda(\lambda-C)}{|z|^2}+\frac{2\lambda\bar{z}_a\sigma^i_{ab}z_b\hat{S}^i}{|z|^4}\right], \tag{2.6c}$$

$$Y=\tfrac{1}{2}(z_a\,\partial_a-\bar{z}_a\,\bar{\partial}_a)+\Sigma-\lambda, \tag{2.6d}$$

where λ is an arbitrary parameter. It is straightforward to verify that these operators obey the relations (2.1). This realization of OSp(2,1) can be further enlarged by observing that all the above charges are invariant under the SU(2) action generated by

$$J^i=-\tfrac{1}{2}(z_a\sigma^i_{ab}\,\partial_b-\bar{z}_a\sigma^i_{ab}\,\bar{\partial}_b)+\hat{S}^i,\quad i=1,2,3. \tag{2.7}$$

Note also that $C=z_a\,\partial_a-\bar{z}_a\,\bar{\partial}_a+\Sigma$ commutes with J^i and all the OSp(2,1) generators. We have thus obtained a realization of the direct sum algebra OSp(2,1) $\oplus$SU(2) $\oplus$U(1).

III. DIMENSIONAL REDUCTION AND THE COULOMB PROBLEM

We shall now establish the relation between the realization of Sec. II and the three-dimensional system whose dynamics is governed by the Hamiltonian H_D of Eq. (1.1). Introduce the coordinates

$$0<r\leqslant\infty,\quad 0\leqslant\theta<\pi,\quad 0\leqslant\phi<2\pi,\quad 0\leqslant\omega<4\pi \tag{3.1}$$

through

$$z_1=\sqrt{r}\cos(\theta/2)\exp[(i/2)\ (\omega-\phi)],$$
$$z_2=\sqrt{r}\sin(\theta/2)\exp[(i/2)\ (\omega+\phi)]. \tag{3.2}$$

In terms of these, the operator C takes the form

$$C=-i\frac{\partial}{\partial\omega}+\Sigma. \tag{3.3}$$

Consider now the eigenvalue equation

$$R\tilde{\Psi}=\alpha(-4E)^{-1/2}\tilde{\Psi}. \tag{3.4}$$

Since C is central it can be diagonalized simultaneously with R; this allows one to separate the variable ω. From the condition

$$C\tilde{\Psi}=2q\tilde{\Psi}, \tag{3.5}$$

we get

$$\tilde{\Psi}(r,\theta,\phi,\omega)=e^{(q-(1/2)\Sigma)\omega}\Psi(r,\theta,\phi). \tag{3.6}$$

Note that q must be an integer or a half-integer for $\tilde{\Psi}$ to be single valued. Using these wave functions in (3.4) and eliminating the ω dependence, one is left with the following eigenvalue problem:

$$\hat{R}\Psi=\alpha(-4E)^{-1/2}\Psi, \tag{3.7}$$

with

$$\hat{R}=\frac{1}{2}\left\{r\left[-i\frac{\partial}{\partial r^i}-\left(q-\frac{1}{2}\Sigma\right)A^D_i\right]^2+r+\frac{(\lambda-q)^2-q\Sigma+\frac{1}{4}\Sigma^2}{r}-\frac{2\lambda r^i\hat{S}^i}{r^2}\right\}. \tag{3.8}$$

The coordinates r^i in the above expression for $\hat{R}$ are given by

$$r^i=\bar{z}_a\sigma^i_{ab}\bar{z}_b; \tag{3.9}$$

they correspond to the standard Cartesian coordinates on $\mathbf{R}^3$. Indeed, observe that, from (3.2),

$$r_1=r\cos\phi\sin\theta,\quad r_2=r\sin\phi\sin\theta,$$
$$r_3=r\cos\theta. \tag{3.10}$$

Multiplying (3.7) with $1/r$ and rescaling according to $r_i\to(-E)^{1/2}r^i$, we finally find that Ψ satisfies the Schrödinger equation

$$H_D\Psi=E\Psi. \tag{3.11}$$

As already mentioned, this equation describes the quantum mechanics of two spin-0 and one spin-$\frac{1}{2}$ particles in dyon fields.

By construction, R possesses a dynamical OSp(2,1) $\oplus$SU(2) $\oplus$U(1) algebra. The projection of (3.4), which is achieved with the help of constraint (3.5), preserves these symmetries as C commutes with every generator. We may therefore conclude that the eigenfunctions Ψ of $\hat{R}$ belong to OSp(2,1) $\oplus$SU(2) representation spaces. These functions Ψ also happen to be eigenfunctions of H_D, which thus admits an OSp(2,1) $\oplus$SU(2) spectrum supersymmetry.

IV. THE QUANTUM NUMBERS AND THE BASIS STATES

The eigenstates of H_D can now be obtained by constructing bases for representations of OSp(2,1) $\oplus$SU(2). One choice of quantum numbers for the basis states of the irreducible representations of OSp(2,1) $\oplus$SU(2) is provided by the eigenvalues of the Casimir operators associated to the canonical chain of maximal subalgebras:

$$\underset{C_2,C_3}{\mathrm{OSp}(2,1)}\oplus\underset{\mathbf{J}^2}{\mathrm{SU}(2)}\supset\underset{Y}{\mathrm{O}(2)}\oplus\underset{C_0}{\mathrm{O}(2,1)}\oplus\underset{J_3}{\mathrm{U}(1)}$$
$$\supset\mathrm{O}(2)\oplus\underset{R}{\mathrm{O}(2)}\oplus\mathrm{U}(1).$$

The Casimir operators C_0 of O(2,1) are well known and given by

$$C_0=\tfrac{1}{2}(HK+KH)-D^2. \tag{4.1}$$

The quadratic and cubic Casimir operators of OSp(2,1) have, respectively, the following expression[3,4]:

$$C_2=C_0-Y^2-\tfrac{1}{2}[F^R_+,F^L_-]-\tfrac{1}{2}[F^L_+,F^R_-], \tag{4.2a}$$

$$C_3=Y(C_2-\tfrac{1}{4}[F^R_+,F^L_-]-\tfrac{1}{4}[F^L_+,F^R_-]-\tfrac{1}{2})-\tfrac{1}{4}[F^L_+,F^R_+]B_--\tfrac{1}{4}[F^L_-,F^R_-]B_+-\tfrac{1}{4}[F^R_+,F^L_-]R-\tfrac{1}{4}[F^R_-,F^L_+]R. \tag{4.2b}$$

In our realization, C_2 and C_3 are completely determined in terms of j, q, and λ so that at fixed angular momentum, the system is described by those representations of OSp(2,1) for which

$$C_2 = j(j+1) - j_0(j_0+1), \quad j_0 = |q| - \tfrac{1}{2}, \tag{4.3a}$$

$$C_3 = (q-\lambda)[j(j+1) - j_0(j_0+1)]. \tag{4.3b}$$

[As usual, the eigenvalues of $\mathbf{J}^2$ are written in the form $j(j+1)$, $j = j_0, j_0+1, \ldots$, and those of J_3 denoted by $m = -j, -j+1, \ldots, j$.] In order to characterize the states belonging to these irreducible representations, it is convenient to replace C_0 and Y by

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.4a}$$

and

$$\hat{A} = -\tfrac{1}{2}(1-\gamma^5)\operatorname{sgn}(2[F^R_+, F^L_-] + 2[F^L_+, F^R_-]) + \tfrac{1}{2}(1+\gamma^5)\Sigma, \tag{4.4b}$$

which represent an equivalent pair of labeling operators. It is not difficult to check that C_0 and Y can unambiguously be reconstructed in terms of γ^5 and $\hat{A}$, which both have their eigenvalues (χ and $\hat{\alpha}$) equal to ± 1. As a matter of fact,

$$C_0 = \{[J^2 - j_0(j_0+1) + (q-\lambda)^2]^{1/2} - \tfrac{1}{4}(1-\gamma^5)\hat{A}\}^2 - \tfrac{1}{4}, \tag{4.5a}$$

$$Y = \tfrac{1}{4}(1+\gamma^5)\hat{A} + (q-\lambda). \tag{4.5b}$$

Finally, from the representation theory of O(2,1), we know the eigenvalues of R to be given by

$$r_n = (\Delta_{j,\hat{\alpha},\chi} + n), \quad n = 0,1,2,\ldots, \tag{4.6}$$

if those of C_0 are written in the form

$$C_0 = \Delta_{j,\hat{\alpha},\chi}(\Delta_{j,\hat{\alpha},\chi} - 1). \tag{4.7}$$

Here,

$$\Delta_{j,\hat{\alpha},\chi} = [j(j+1) - j_0(j_0+1) + (q-\lambda)^2]^{1/2} - \tfrac{1}{4}(1-\chi)\hat{\alpha} + \tfrac{1}{2}. \tag{4.8}$$

In summary, we have the following eigenvalue equations to characterize the states of our system:

$$\mathbf{J}^2|j,m;\hat{\alpha},\chi,n\rangle = j(j+1)|j,m;\hat{\alpha},\chi,n\rangle, \tag{4.9a}$$

$$J_3|j,m;\hat{\alpha},\chi,n\rangle = m|j,m;\hat{\alpha},\chi,n\rangle, \tag{4.9b}$$

$$\hat{A}\,|j,m;\hat{\alpha},\chi,n\rangle = \hat{\alpha}|j,m;\hat{\alpha},\chi,n\rangle, \tag{4.9c}$$

$$\gamma^5|j,m;\hat{\alpha},\chi,n\rangle = \chi|j,m;\hat{\alpha},\chi,n\rangle, \tag{4.9d}$$

$$R\,|j,m;\hat{\alpha},\chi,n\rangle = (\Delta_{j,\hat{\alpha},\chi} + n)|j,m;\hat{\alpha},\chi,n\rangle. \tag{4.9e}$$

The spectrum is then easily derived from Eqs. (3.4) and (4.9e), and one finds

$$E_{j,\hat{\alpha},\chi,n} = -\alpha^2/4(\Delta_{j,\hat{\alpha},\chi} + n)^2. \tag{4.10}$$

V. ACTION OF THE LADDER OPERATORS

We will now determine the action of the remaining generators on the basis vectors of Sec. IV. From the structure relations (2.2), we find that $B_\pm$ act according to

$$B_\pm\,|j,m;\hat{\alpha},\chi,n\rangle = [(\Delta_{j,\hat{\alpha},\chi} + n)(\Delta_{j,\hat{\alpha},\chi} + n \pm 1) - \Delta_{j,\hat{\alpha},\chi}(\Delta_{j,\hat{\alpha},\chi} - 1)]^{1/2} \times |j,m;\hat{\alpha},\chi,n \pm 1\rangle. \tag{5.1}$$

The operators $F^{L,R}_\pm$ anticommute with γ^5 and therefore reverse the chirality χ. From their commutation relations with R and Y, one obtains

$$\begin{aligned} F^L_\pm\,|j,m;\hat{\alpha},\chi,n\rangle &= \Big[\sum_{\hat{\alpha}'=\pm 1} c^L_\pm(\hat{\alpha}')\Big|j,m;\hat{\alpha}',-1,n+\frac{\hat{\alpha}'}{2}\pm\frac{1}{2}\Big\rangle\,\delta_{\chi,1}\delta_{\hat{\alpha},1}\Big] \\ &\quad + d^L_\pm(\hat{\alpha})\Big|j,m;-1,1,n-\frac{\hat{\alpha}}{2}\pm\frac{1}{2}\delta_{\chi,-1}\Big\rangle, \end{aligned} \tag{5.2a}$$

$$\begin{aligned} F^R_\pm\,|j,m;\hat{\alpha},\chi,n\rangle &= \Big[\sum_{\hat{\alpha}'=\pm 1} c^R_\pm(\hat{\alpha}')\Big|j,m;\hat{\alpha}',-1,n+\frac{\hat{\alpha}}{2}\pm\frac{1}{2}\Big\rangle\,\delta_{\chi,1}\delta_{\hat{\alpha},-1}\Big] \\ &\quad + d^R_\pm(\hat{\alpha})\Big|j,m;1,1,n-\frac{\hat{\alpha}}{2}\pm\frac{1}{2}\delta_{\chi,-1}\Big\rangle. \end{aligned} \tag{5.2b}$$

What remains is to obtain the various constants $c^{L,R}_\pm(\hat{\alpha})$ and $d^{L,R}_\pm(\hat{\alpha})$. To this effect, note that

$$[F^L_\mp, F^R_\pm] = \tfrac{1}{4}(\mathscr{S} \pm \mathscr{A}), \tag{5.3}$$

with

$$\mathscr{S} = 2[F^L_-, F^R_+] - 2[F^R_-, F^L_+], \tag{5.4a}$$

and

$$\mathscr{A} = 2[F^L_-, F^R_+] + 2[F^L_-, F^L_+], \tag{5.4b}$$

and the anticommutator of $\mathscr{S}$ and $\mathscr{A}$ can be cast in the form

$$\{\mathscr{S}, \mathscr{A}\} = -32RY. \tag{5.5}$$

Now $\mathscr{A}$ can be expressed as follows:

$$\mathscr{A} = 4(C_2 - C_0 + Y^2). \tag{5.6}$$

Since $\mathscr{A}$, R, and Y are diagonal in our basis, we may thus write

$$\langle j,m;\hat{\alpha},\chi,n|[F^L_\mp, F^R_\pm]|j,m;\hat{\alpha},\chi,n\rangle = \langle j,m;\hat{\alpha},\chi,n|[-4\mathscr{A}^{-1}RY \pm \tfrac{1}{4}\mathscr{A}]|j,m;\hat{\alpha},\chi,n\rangle. \tag{5.7}$$

Recalling that $\{F^L_{\mp J}, F^R_\pm\} = R \mp Y$, we finally arrive at the following expression for the norms of $F^{L,R}_\pm|j,m;\hat{\alpha},\chi,n\rangle$:

$$\begin{aligned} \langle j,m;\hat{\alpha},\chi,n|F^L_\mp F^R_\pm\,|j,m;\hat{\alpha},\chi,n\rangle &= \Big(\sum_{\hat{\alpha}'} |c^R_\pm(\hat{\alpha}')|^2\Big)\delta_{\chi,1}\delta_{\hat{\alpha},-1} + |d^R_\pm(\hat{\alpha})|^2\delta_{\chi,-1} \\ &= \Big[\frac{1}{2} - \frac{\hat{\alpha}}{4}(1+\chi)\Big] \\ &\quad \times\Big[\Delta_{j,\hat{\alpha},\chi} + n \mp \Big(\frac{\hat{\alpha}}{4}(1+\chi) + q - \lambda\Big)\Big] \\ &\quad - \frac{\hat{\alpha}}{4}(1-\chi)\Big[(\Delta_{j,\hat{\alpha},\chi} + n)\frac{(q-\lambda)}{\Lambda} \mp \Lambda\Big], \end{aligned} \tag{5.8a}$$

$$\begin{aligned}\langle j,m;\hat{\alpha},\chi,n|F^{R}_{\mp}F^{L}_{\pm}\,|j,m;\hat{\alpha},\chi,n\rangle &= \Big(\sum_{\hat{\alpha}'}|c^{L}_{\pm}(\hat{\alpha}')|^2\Big)\delta_{\chi,1}\delta_{\hat{\alpha},1}+|d^{L}_{\pm}(\hat{\alpha})|^2\delta_{\chi,-1}\\ &=\Big[\frac{1}{2}+\frac{\hat{\alpha}}{4}(1+\chi)\Big]\\ &\quad\times\Big[\Delta_{j,\hat{\alpha},\chi}+n\pm\Big(\frac{\hat{\alpha}}{4}(1+\chi)+q-\lambda\Big)\Big]\\ &\quad+\frac{\hat{\alpha}}{4}(1-\chi)\Big[(\Delta_{j,\hat{\alpha},\chi}+n)\frac{(q-\lambda)}{\Lambda}\pm\Lambda\Big],\end{aligned}\tag{5.8b}$$

where $\Lambda=[C_2+(\frac{1}{2}C-\lambda)^2]^{1/2}$. The above relations immediately give the coefficients $d^{L,R}_{\pm}$ when χ is set equal to -1. When $\chi=+1$, one needs, in addition to (5.8), another equation in order to determine the coefficients $c^{L,R}_{\pm}$. Such a relation is provided by

$$\sum_{\hat{\alpha}}c^{L,R}_{\mp}(\hat{\alpha})d^{L,R}_{\pm}(\hat{\alpha})=0,\tag{5.9}$$

which is derived from (5.2) by exploiting the nilpotency of the operators $F^{L,R}_{\pm}$. In the end one finds

$$c^{L}_{\pm}(\hat{\alpha})=N_{-\hat{\alpha}}[\Lambda\mp\hat{\alpha}\Lambda+\tfrac{1}{2}\pm\tfrac{1}{2}+n]^{1/2},\tag{5.10a}$$

$$c^{B}_{\pm}(\hat{\alpha})=\hat{\alpha}N_{\hat{\alpha}}[\Lambda\mp\hat{\alpha}\Lambda+\tfrac{1}{2}\pm\tfrac{1}{2}+n]^{1/2},\tag{5.10b}$$

$$d^{L}_{\pm}(\hat{\alpha})=\hat{\alpha}N_{\hat{\alpha}}[\Lambda\pm\hat{\alpha}\Lambda+\tfrac{1}{2}(1-\hat{\alpha})+n]^{1/2},\tag{5.10c}$$

$$d^{R}_{\pm}(\hat{\alpha})=N_{-\hat{\alpha}}[\Lambda\pm\hat{\alpha}\Lambda+\tfrac{1}{2}(1-\hat{\alpha})+n]^{1/2},\tag{5.10d}$$

where $N_{\hat{\alpha}}=[[\Delta+\hat{\alpha}(q-\lambda)]/2\Lambda]^{1/2}$. Substitution in (5.2) explicitly gives the action of the supersymmetry generators on the basis states $|\,j,m;\hat{\alpha},\chi,n\rangle$. We have checked that the above coefficients could consistently be taken real.

We should point out that when the angular momentum takes its lowest possible value, $j=|q|-\frac{1}{2}$, the corresponding OSp(2,1) representation comprises only half the number of states found in the generic situation. The only positive chirality states $|\,j_0,m,\hat{\alpha},1,n\rangle$ that can be defined in this case are the ones for which $|q-\hat{\alpha}/2|=|q|-\frac{1}{2}$. Depending on the sign of q, this forces $\hat{\alpha}$ to take a definite value. The OSp(2,1) representation is then erected above $|\,j_0,m,\hat{\alpha},1,0\rangle$ by repeated application of the ladder operators whose action is still given by (5.2).

VI. WAVE FUNCTIONS

Now that we have found the action of the symmetry generators on our basis states, we can obtain the wave functions algebraically by going to a coordinate realization. As functions over $\mathsf{R}^{+}\times S^3$, the various basis states will read

$$\langle r,\theta,\phi,\omega|\,j,m;1,1,n\rangle \equiv\widetilde{\Psi}_{(1,1,n)}=\begin{bmatrix}\widetilde{\Psi}^{(j,m,1,1,n)}(r,\theta,\phi,\omega)\\0\\0\\0\end{bmatrix},\tag{6.1a}$$

$$\langle r,\theta,\phi,\omega|\,j,m;-1,1,n\rangle \equiv\widetilde{\Psi}_{(-1,1,n)}=\begin{bmatrix}0\\\widetilde{\Psi}^{(j,m,-1,1,n)}(r,\theta,\phi,\omega)\\0\\0\end{bmatrix},\tag{6.1b}$$

$$\langle r,\theta,\phi,\omega|\,j,m;\hat{\alpha},-1,n\rangle \equiv\widetilde{\Psi}_{(\hat{\alpha},-1,n)}=\begin{bmatrix}0\\0\\\widetilde{\Psi}^{(j,m,\hat{\alpha},-1,n)}(r,\theta,\phi,\omega)\\\widetilde{\Psi}_2{}^{(j,m,\hat{\alpha},-1,n)}(r,\theta,\phi,\omega)\end{bmatrix}.\tag{6.1c}$$

We consider first the case $j\neq j_0$. The wave functions $\widetilde{\Psi}_{(1,1,n)}$ are determined as follows. The ω dependence has already been fixed. [See Eq. (3.6).] The eigenvalue equations associated to $\mathbf{J}^2$ and J_3 further allow one to separate the variables θ and ϕ and to write $\widetilde{\Psi}^{(j,m,1,1,n)}(r,\theta,\phi,\omega)$ in the form

$$\widetilde{\Psi}^{(j,m,1,1,n)}(r,\theta,\phi,\omega)=e^{i(q-1/2)^{\omega}}\rho_n(r)\mathscr{Y}_{q-1/2,j,m}(\theta,\phi),\tag{6.2}$$

where the $\mathscr{Y}_{q-1/2,j,m}(\theta,\phi)$ are the monopole harmonics.[5] The radial equation satisfied by $\rho_0(r)$ is gotten from (4.9e), i.e., from $R\widetilde{\Psi}_{(1,1,0)}=\Delta_{j,1,1}\widetilde{\Psi}_{(1,1,0)}$, which, after separation of the angular variables, reduces to

$$\Big[r\frac{\partial^2}{\partial r^2}+2\frac{\partial}{\partial r}-\frac{\Lambda^2-\frac{1}{4}}{r}+2\Lambda+1-r\Big]\rho_0(r)=0.\tag{6.3}$$

Its normalized solution is given by

$$\rho_0(r)=[(2^{\Lambda+1/2}/\sqrt{\Gamma(2\Lambda+1)})e^{-r}r^{\Lambda-1/2}].\tag{6.4}$$

The wave function $\widetilde{\Psi}_{(1,1,n)}$ is obtained by successive application of the ladder operator B_{+}:

$$\begin{aligned}B_{+}\widetilde{\Psi}_{(1,1,n)}&=\Big(r-r\frac{r}{\partial r}-\Delta_{j,1,1}-1-n\Big)\widetilde{\Psi}_{(1,1,n)}\\&=(n+1)(2\Delta_{j,1,1}+n)\widetilde{\Psi}_{(1,1,n)}.\end{aligned}\tag{6.5}$$

A little algebra yields

$$\begin{aligned}\rho_n(r)&=(-1)^n[2^{2\Lambda+1}n!/\Gamma(2\Lambda+n+1)]^{1/2}\\&\quad\times e^{-r}r^{\Lambda-1/2}L^{2\Lambda}_{n}(2r).\end{aligned}\tag{6.6}$$

The wave functions that remain are found by first computing

$$\widetilde{\Psi}_{(1,-1,0)}=(\Lambda+\lambda-q)^{-1/2}\langle r,\theta,\phi,\omega|F^{L}_{-}\,|\,j,m{:}1,1,0\rangle,\tag{6.7a}$$

$$\widetilde{\Psi}_{(-1,1,0)}=(\Lambda-\lambda+q)^{-1/2}\langle r,\theta,\phi,\omega|F^{L}_{+}\,|\,j,m{:}1,-1,0\rangle,\tag{6.7b}$$

$$\begin{aligned}\widetilde{\Psi}_{(-1,-1,0)}&=(2\Lambda+1)^{-1/2}[N_{+}\langle r,\theta,\phi,\omega|F^{R}_{-}\,|\,j,m{:}1,1,0\rangle,\\&\quad-N_{-}\langle r,\theta,\phi,\omega|F^{R}_{-}\,|\,j,m{:}1,1,0\rangle],\end{aligned}\tag{6.7c}$$

and then applying B_{+} repeatedly on each of these functions in order to increase the value of n. Setting $\mathbf{x}=\mathbf{r}(-E)^{-1/2}$, we finally obtain for the eigenstates of the Hamiltonian H_D, when $j\neq j_0$,

$$\Psi^{(j,m,\hat{\alpha},1,n)}(\mathbf{x}) = \hat{\alpha}\rho_n^{j,\hat{\alpha},1}(|\mathbf{x}|)\mathscr{Y}_{q-\hat{\alpha}/2,j,m}(\theta,\phi), \tag{6.8a}$$

$$\begin{bmatrix}\Psi_1^{(j,m,\hat{\alpha},-1,n)}(\mathbf{x})\\ \Psi_2^{(j,m,\hat{\alpha},-1,n)}(\mathbf{x})\end{bmatrix} = i\rho_n^{j,\hat{\alpha},-1}(|\mathbf{x}|)\begin{bmatrix}\sqrt{[(j+m)/2j]}D_{\hat{\alpha}}\mathscr{Y}_{q,j-1/2,m-1/2}(\theta,\phi) - \sqrt{[(j-m+1)/(2j+2)]}\hat{\alpha}D_{-\hat{\alpha}}\mathscr{Y}_{q,j+1/2,m-1/2}(\theta,\phi)\\ \sqrt{[(j-m)/2j]}D_{\hat{\alpha}}\mathscr{Y}_{q,j-1/2,m-1/2}(\theta,\phi) + \sqrt{[(j+m+1)/(2j+2)]}\hat{\alpha}D_{-\hat{\alpha}}\mathscr{Y}_{q,j+1/2,m-1/2}(\theta,\phi)\end{bmatrix}, \tag{6.8b}$$

where

$$\rho_n^{j,\hat{\alpha},\chi}(|\mathbf{x}|) = (-1)^n[n!(-4E)^{\Delta}/\Gamma(2\Delta+n)]^{1/2}|\mathbf{x}|^{\Delta-1}e^{[-\sqrt{-E}|\mathbf{x}|]}L_n^{2\Delta-1}[\sqrt{-4E}|\mathbf{x}|], \tag{6.8c}$$

with $E = E_{j,\hat{\alpha},\chi,n}$ and $\Delta = \Delta_{j,\hat{\alpha},\chi}$. Recall that the wave functions $\widetilde{\Psi}_{(\hat{\alpha},1,n)}$ and $\widetilde{\Psi}_{(\hat{\alpha},-1,n)}$ are associated, respectively, to the spin-0 and spin-$\frac{1}{2}$ particles of our multiplet.

The case $j = |q| - \frac{1}{2}$ is treated analogously. The expressions for the wave function are simpler and read

$$\Psi^{(j_0,m,\hat{\alpha},1,n)}(\mathbf{x}) = \rho_n^{j_0,\hat{\alpha},1}(|\mathbf{x}|)\mathscr{Y}_{q-\hat{\alpha}/2,j_0,m}(\theta,\phi), \tag{6.9a}$$

$$\begin{bmatrix}\Psi_1^{(j_0,m,\hat{\alpha},-1,n)}(\mathbf{x})\\ \Psi_2^{(j_0,m,\hat{\alpha},-1,n)}(\mathbf{x})\end{bmatrix} = i\rho_n^{j,\hat{\alpha},-1}(|\mathbf{x}|)\begin{bmatrix}\sqrt{[(j-m+1)/(2j+2)]}\mathscr{Y}_{q,j_0+1/2,m-1/2}(\theta,\phi)\\ \sqrt{[(j+m+1)/(2j+2)]}\mathscr{Y}_{q,j_0+1/2,m-1/2}(\theta,\phi)\end{bmatrix}. \tag{6.9b}$$

On these lowest angular states the case $q = \lambda$ is pathological. In this instance, $\mathscr{A} = 1 + \gamma^5$ and $\hat{\alpha}$ is related to χ. This indicates that the supersymmetry cannot be implemented[6] and that the bosonic and fermionic sectors span separate O(2,1) representations.

VII. CONCLUSION

In summary we have seen that there exists an $N=2$ dynamical supersymmetry for the system consisting of two spin-0 particles and one spin-$\frac{1}{2}$ particle in a Coulomb potential. At fixed angular momentum, the entire dynamics is described by a single irreducible representation of OSp(2,1). The solutions to the corresponding Schrödinger and Pauli equations have been obtained by constructing these representations.

A special case covered by our analysis is that of the Pauli equation for a spin-$\frac{1}{2}$ particle in the field of a dyon. Indeed, for $\lambda = q$, the two lower components of H_D read as follows after a mass M has been reinstated:

$$H_{\rm P} = \frac{1}{2M}(\mathbf{p} - e\mathbf{A})^2 - \frac{\alpha}{|\mathbf{x}|} - \frac{e}{2M}\mathbf{B}\cdot\sigma, \tag{7.1}$$

where $\mathbf{B} = g(x/|\mathbf{x}|^3)$. The spectrum of $H_{\rm P}$ is obtained from Eq. (4.10),

$$\mathscr{E}_{j,n,\hat{\alpha}} = \frac{-\alpha^2 M}{2[\sqrt{j(j+1) - e^2g^2 + \frac{1}{4}} + \frac{1}{2}(1+\hat{\alpha}) + n]^2}, \tag{7.2}$$

and the wave functions are given by Eqs. (6.8) with $\mathbf{x}$ replaced by $\sqrt{2M}\,\mathbf{x}$.

ACKNOWLEDGMENTS

We thank Michel Mayrand for his help in the preparation of the manuscript.

This work has been supported in part through funds provided by the Natural Science and Engineering Research Council (NSERC) of Canada and the Fonds FCAR of the Québec Ministry of Education.

[1] For a review, see E. D'Hoker, V. A. Kostelecky, and L. Vinet, in *Dynamical Groups and Spectrum Generating Algebras*, edited by A. O. Barut, A. Bohm, and Y. Ne'eman (World Scientific, Singapore, in press).
[2] E. D'Hoker and L. Vinet, Nucl. Phys. **260**, 79 (1985).
[3] V. Kac, Funct. Anal. Appl. **9**, 263 (1975); Adv. Math. **26**, 8 (1977).
[4] M. Scheunert, W. Nahm, and V. Rittenberg, J. Math. Phys. **18**, 155 (1977).
[5] T. T. Wu and C. N. Yang, Nucl. Phys. B **107**, 365 (1976).
[6] E. D'Hoker and L. Vinet, Commun. Math. Phys. **97**, 341 (1985).

Correlated states and collective transition operators for multilevel atomic systems. II. Symmetry properties

Jagdish Rai
Department of Physics and Astronomy, University of Alabama, Tuscaloosa, Alabama 35487
C. L. Mehta
Department of Physics, Indian Institute of Technology, Delhi, New Delhi 110016, India
N. Mukunda
Center for Theoretical Studies, Indian Institute of Science, Bangalore 560012, India

(Received 4 April 1988; accepted for publication 13 July 1988)

The symmetry properties of the N-particle multilevel system are studied. For indistinguishable particles, the total wave function must be either symmetric or antisymmetric in the exchange of any two particles. The internal states of the multilevel system are combined with suitable spatial wave functions to produce a totally symmetric or antisymmetric wave function. The problem of degeneracy and the method of combining a given permutational symmetry of the internal wave function with the same or the conjugate symmetry for the spatial part are described. It is found that Young's diagram corresponding to the given irreducible representation of the permutation group plays an important role.

I. INTRODUCTION

In a recent paper,[1] we have defined the correlated states and collective transition operators for an assembly of N multilevel atoms interacting with common radiation fields. We found that the possible states are determined by various allowed Young diagrams, corresponding to various representations of the permutation group S_N. The collective transition operators were obtained as the generators of the group SU(n), where n is the total number of atomic levels. In this paper, we study the symmetry properties of these "correlated states." The total wave function of the assembly consists of an orbital part and a spin part (spin here refers to the internal coordinate or the excitation state of the atom). The total wave function must also be symmetric or antisymmetric under permutations of the atoms or molecules. For systems such as hydrogen and ammonia masers, the atoms are spatially separated, and there is no need for overall symmetrization of the wave function, while for lasers, the spatial correlation between molecules and the lasing mode is essential. This is very much analogous to how the Pauli principle can lead to magnetic effects even when there are no spin-dependent terms in the Hamiltonian. The Pauli exclusion principle requires that the total wave function Ψ change sign under the simultaneous interchange of both space and spin coordinates. There is thus a strict correlation between the spatial symmetry of the orbital part (spin-independent) and the total spin. In this paper, we propose a technique for constructing total wave functions of desired symmetry.

We first consider the gas or the interacting medium referred to as the molecular system to consist of a large number N of atoms or molecules each having two possible internal nondegenerate energy states. This system interacts with a resonant radiation field. The transition from one internal state of the molecular system to another is described in terms of angular momentum operators corresponding to the spin-$\frac{1}{2}$ system. Omitting the radiation field, the Hamiltonian of the molecular system can be written as[2]

$$\hat{H} = \hat{H}_0 + E \sum_{j=1}^{N} R_3^{(j)} = \hat{H}_0 + E\hat{R}_3 . \tag{1.1}$$

Here H_0 represents the translational and intermolecular interaction Hamiltonian, and it acts on the center of mass coordinates only, $E = \hbar\omega$ is the level separation (excitation energy of each molecule), and $\hat{R}_\alpha^{(j)}$ ($\alpha = 1,2,3$) are the Cartesian components of the spin-$\frac{1}{2}$ operators of the jth molecule. These operators satisfy the usual commutation relations

$$[R_\alpha^{(j)}, R_\beta^{(k)}] = i\delta_{jk}\epsilon_{\alpha\beta\gamma}\hat{R}_\gamma^{(j)} , \tag{1.2}$$

where δ_{jk} is the Kronecker δ symbol and $\epsilon_{\alpha\beta\gamma}$ is the completely antisymmetric Levi-Civita tensor. We also define the operators

$$\hat{R}_\alpha = \sum_{j=1}^{N} \hat{R}_\alpha^{(j)}, \quad \alpha = 1,2,3, \tag{1.3}$$

and

$$\hat{R}^2 = \hat{R}_1^2 + \hat{R}_2^2 + \hat{R}_3^2 . \tag{1.4}$$

A typical energy state of the N-molecule system is written as

$$\psi_{gm}^{(N)} = U_g^{(N)}(\underline{r}_1,\ldots,\underline{r}_N)\phi_m^{(N)} . \tag{1.5}$$

Here $U_g^{(N)}$ is the wave function describing the center of mass coordinates $\underline{r}_1,\ldots,\underline{r}_N$, and is an eigenstate of $\hat{H}_0$,

$$\hat{H}_0 U_g^{(N)} = E_g U_g^{(N)} . \tag{1.6}$$

The state

$$\phi_m^{(N)} = |+ - - + \cdots\rangle \tag{1.7}$$

is the internal energy state or what we shall call the spin state of the system and is an eigenstate of $\hat{R}_3^{(j)}$ (for all j), with an eigenvalue $\frac{1}{2}$ or $-\frac{1}{2}$ depending on whether there is a + or − sign at the jth place. There are obviously 2^N such states. If n_+ and n_- denote the number of + and − signs in $\phi_m^{(N)}$, then m is defined as

$$m = n_+ - n_- \,. \tag{1.8}$$

Further,

$$N = n_+ + n_- \tag{1.9}$$

is the total number of molecules. The total energy of the system is given by

$$\hat{H}\psi^{(N)}_{gm} = E_{gm}\psi^{(N)}_{gm} \,, \tag{1.10}$$

where

$$E_{gm} = E_g + \tfrac{1}{2}mE \,. \tag{1.11}$$

The $\psi^{(N)}_{gm}$ is, in general, degenerate. Both $U^{(N)}_g$ and $\phi^{(N)}_m$ contribute to the degeneracy. One may readily verify that the degeneracy of the state $\phi^{(N)}_m$ is given by

$$N!/((N+m)/2)!((N-m)/2)! \,. \tag{1.12}$$

If we take a suitable superposition of the states $\phi^{(N)}_m$ such that this superposition is an eigenstate of $\hat{R}^2$,

$$\hat{R}^2\phi^{(N)}_{rm} = r(r+1)\phi^{(N)}_{rm} \,, \tag{1.13}$$

then in analogy with the angular momentum states one finds that r is an integer with

$$|m| < r < N \,. \tag{1.14}$$

The degeneracy of $\phi^{(N)}_{rm}$ is reduced, but not completely removed. The state $\phi^{(N)}_{rm}$ has the degeneracy [cf. Eq. (2) of Ref. 2 and also Eq. (3.3) below]

$$N!(2r+1)/((N+2r+2)/2)!((N-2r)/2)! \,. \tag{1.15}$$

The state $\phi^{(N)}_{Nm}$ (corresponding to r having its maximum possible value $N/2$) is nondegenerate. Further, this state is completely symmetric under the permutation of internal spins. However, the state $\phi^{(N)}_{rm}$, for $r \neq N/2$, in general, does not have any definite symmetry under the permutation of the internal spins.

If the molecules are indistinguishable, the total wave function $\psi^{(N)}$ must be symmetric for Bose molecules and antisymmetric for Fermi molecules. These limitatations on account of symmetry have been completely ignored throughout the study of two-level molecules by Dicke as well as others.

The question naturally arises whether it is possible to choose a combination of the wave functions

$$\sum_{\Omega} U^{(N)}_{g\Omega}\phi^{(N)}_{rm\Omega}$$

such that the resultant state has the desired symmetry.[3–5] Here Ω is some parameter specifying the manner in which the states $U_{g\Omega}$ or $\phi_{rm\Omega}$ transform under permutations. We wish to explore this aspect of the problem in the present investigation.

Another generalization being considered recently is in treating the interaction problems. This is in regard to assuming that each molecule has three or, in general, n energy states. Transitions from one state to another are described in terms of the operators that are the generators of SU(n). Again, one would be interested in constructing wave functions that have a definite symmetry under permutations. We shall see that the Young diagrams corresponding to the given irreducible representation of the permutation group play an important role in the construction of such states.

In Sec. II, we obtain explicit expression for the three-particle states of the two-level system with given symmetry and quantum numbers r,m. The twofold degeneracy of the state $r = \frac{1}{2}$, $m = \frac{1}{2}$, gets removed. In Sec. III we generalize this to N-particle states of the two-level system. This is then further generalized to the n-level system in Sec. IV. In Sec. V we consider specific examples of two-particle and three-particle systems.

II. THREE-PARTICLE STATES WITH GIVEN SYMMETRY

In this section we consider a three-particle system of two-level "molecules" and give an explicit construction of states that are either symmetric or antisymmetric in the permutation of these particles. The example will serve as an illustration for more complex systems to be considered in later sections. The total wave function is written in the form

$$\psi = \sum_{\Omega} U^{(3)}_{g\Omega}(\underline{r}_1,\underline{r}_2,\underline{r}_3)\phi^{(3)}_{rm\Omega} \,, \tag{2.1}$$

where U is the wave function describing the center of mass coordinates of individual particles and ϕ describes the "spin" state of the system.

A. Spin-$\frac{3}{2}$ states

It is readily seen that the spin wave function

$$|+\ +\ +\rangle \tag{2.2}$$

corresponds to $r = \frac{3}{2}$, $m = \frac{3}{2}$, and is nondegenerate. It is symmetric in the permutation of the three particles.

Similarly the wave functions

$$(1/\sqrt{3})\{|+\ +\ -\rangle + |+\ -\ +\rangle + |-\ +\ +\rangle\}, \tag{2.3}$$

$$(1/\sqrt{3})\{|-\ -\ +\rangle + |-\ +\ -\rangle + |+\ -\ -\rangle\}, \tag{2.4}$$

and

$$|-\ -\ -\rangle \tag{2.5}$$

correspond to $r = \frac{3}{2}$ and $m = \frac{1}{2}$, $-\frac{1}{2}$, and $-\frac{3}{2}$, respectively. They are all nondegenerate and symmetric in the permutation of the three particles.

Hence if the three particles are bosons, the total wave function is required to be symmetric. This is achieved by taking U to be symmetric, i.e.,

$$U_{gs} = \frac{1}{\sqrt{6}} \sum_{P_\mu} U_{\mu_1}(\underline{r}_1)U_{\mu_2}(\underline{r}_2)U_{\mu_3}(\underline{r}_3) \,, \tag{2.6}$$

where the summation includes all six permutations (μ_1,μ_2,μ_3) of $(1,2,3)$. Written explicitly we have

$$\begin{aligned} U_{gs} = (1/\sqrt{6})\{ & U_1(\underline{r}_1)U_2(\underline{r}_2)U_3(\underline{r}_3) \\ & + U_1(\underline{r}_1)U_3(\underline{r}_2)U_2(\underline{r}_3) \\ & + U_2(\underline{r}_1)U_3(\underline{r}_2)U_1(\underline{r}_3) + U_2(\underline{r}_1)U_1(\underline{r}_2)U_3(\underline{r}_3) \\ & + U_3(\underline{r}_1)U_2(\underline{r}_2)U_1(\underline{r}_3) + U_3(\underline{r}_1)U_1(\underline{r}_2)U_2(\underline{r}_3)\} \,. \end{aligned} \tag{2.7}$$

Thus the total wave function

$$U_{gs}\phi_{3/2,m} \tag{2.8}$$

is automatically symmetric.

Similarly if the three particles are fermions, the total wave function is to be antisymmetric, and this is achieved by taking $U_{g\Omega}$ itself as antisymmetric, viz.,

$$U_{ga}=\frac{1}{\sqrt{6}}\sum_{P_\mu}\delta_{P_\mu}U_{\mu_1}(\underline{r}_1)U_{\mu_2}(r_2)U_{\mu_3}(\underline{r}_3)$$

$$=\frac{1}{\sqrt{6}}\begin{vmatrix} U_1(\underline{r}_1) & U_1(\underline{r}_2) & U_1(\underline{r}_3) \\ U_2(\underline{r}_1) & U_2(\underline{r}_2) & U_2(\underline{r}_3) \\ U_3(\underline{r}_1) & U_3(\underline{r}_2) & U_3(\underline{r}_3) \end{vmatrix}. \qquad (2.9)$$

Here δ_{P_μ} is ± 1 depending on whether the permutation P_μ is even or odd. The wave function

$$U_{ga}\phi_{3/2,m} \qquad (2.10)$$

is then antisymmetric.

It may be mentioned that the state U_{gs} [Eq. (2.6)] generates a one-dimensional irreducible representation of the symmetric group S_3 where each element of S_3 is represented by the number 1. This representation of S_3 is customarily denoted by the Young diagram

$$\square\square\square\;\cdot \qquad (2.11)$$

On the other hand the state U_{ga} [Eq. (2.9)] generates the other one-dimensional representation of S_3 where even permutations are represented by $+1$, and odd permutations by -1. Such a representation is denoted by the Young diagram

$$\begin{array}{c}\square\\ \square\\ \square\end{array}\cdot \qquad (2.12)$$

Each of the spin wave functions (2.2), (2.3), (2.4), or (2.5), i.e., ϕ_{rm}, with $r=\frac{3}{2}$, $m=\frac{3}{2},\frac{1}{2},-\frac{1}{2}$, or $-\frac{3}{2}$, also generates the one-dimensional irreducible representation (2.11) of S_3. Obviously no spin wave function generates the irreducible representation (2.12). In fact, only those irreducible representations that have at most two rows can be generated by spin-$\frac{1}{2}$ wave functions (cf. Wigner[6]). Each of the wave functions (2.8) generates the irreducible representation (2.11) of S_3, whereas each of the wave functions (2.10) generates the irreducible representation (2.12). This, in fact, is a consequence of the relations

$$\square\square\square \quad\otimes\quad \square\square\square \quad=\quad \square\square\square \qquad (2.13)$$

and

$$\square\square\square \quad\otimes\quad \begin{array}{c}\square\\ \square\\ \square\end{array} \quad=\quad \begin{array}{c}\square\\ \square\\ \square\end{array}\;. \qquad (2.14)$$

B. Spin-$\frac{1}{2}$ states

We now consider states with $r=\frac{1}{2}$, $m=\frac{1}{2}$. This state is twofold degenerate. We may choose these states explicitly as follows:

$$|a_1\rangle=\phi_{1/2,1/2,a_1}$$
$$=(1/\sqrt{6})$$
$$\times\{2|++-\rangle-|-++\rangle-|+-+\rangle\}, \qquad (2.15a)$$

$$|a_2\rangle=\phi_{1/2,1/2,a_2}=(1/\sqrt{2})\{|+-+\rangle-|-++\rangle. \qquad (2.15b)$$

These states are neither symmetric nor antisymmetric under the permutation of the three particles. Since all permutations can be obtained by repeated application of $\hat{P}_{12}$ and $\hat{P}_{13}$ (interchanges of particles 1 and 2 and of particles 1 and 3, respectively), we will restrict to symmetry under these operations only. We verify that the states $|a_i\rangle$ transform according to the following:

$$\begin{aligned}\hat{P}_{12}|a_1\rangle&=+|a_1\rangle,\\ \hat{P}_{12}|a_2\rangle&=-|a_2\rangle,\\ \hat{P}_{13}|a_1\rangle&=\tfrac{1}{2}\{-|a_1\rangle-3|a_2\rangle\},\\ \hat{P}_{13}|a_2\rangle&=\tfrac{1}{2}\{-\sqrt{3}|a_1\rangle+|a_2\rangle\}.\end{aligned} \qquad (2.16)$$

If we can construct a set of orbital wave functions also satisfying analogous transformations, viz.,

$$\begin{aligned}\hat{P}_{12}\chi_1&=\chi_1,\\ \hat{P}_{12}\chi_2&=-\chi_2,\\ \hat{P}_{13}\chi_1&=\tfrac{1}{2}(-\chi_1-\sqrt{3}\chi_2),\\ \hat{P}_{13}\chi_2&=\tfrac{1}{2}(-\sqrt{3}\chi_1+\chi_2),\end{aligned} \qquad (2.17)$$

then it may readily be verified that the state

$$\psi_s=\chi_1|a_1\rangle+\chi_2|a_2\rangle \qquad (2.18)$$

is symmetric under the operators $\hat{P}_{12}$ and $\hat{P}_{13}$ and hence under all permutations of the three particles. On the other hand, if we can construct states $\bar{\chi}_1$ and $\bar{\chi}_2$ such that they satisfy transformations analogous to Eq. (2.17) except for a minus sign,

$$\begin{aligned}\hat{P}_{12}\bar{\chi}_1&=-\bar{\chi}_1,\\ \hat{P}_{12}\bar{\chi}_2&=+\bar{\chi}_2,\\ \hat{P}_{13}\bar{\chi}_1&=-\tfrac{1}{2}(-\bar{\chi}_1-\sqrt{3}\bar{\chi}_2),\\ \bar{P}_{13}\bar{\chi}_2&=-\tfrac{1}{2}(-\sqrt{3}\bar{\chi}_1+\bar{\chi}_2),\end{aligned} \qquad (2.19)$$

then the state

$$\psi_a=\bar{\chi}_1|a_1\rangle+\bar{\chi}_2|a_2\rangle \qquad (2.20)$$

is antisymmetric under any transposition and hence under all odd permutations.

There are in fact two sets (χ_{11},χ_{21}) and (χ_{12},χ_{22}) satisfying Eq. (2.17), viz.,

$$\begin{aligned}\chi_{11}=(1/\sqrt{12})\{&2U_1(x_1)U_2(x_2)U_3(x_3)\\ &+2U_2(x_1)U_1(x_2)U_3(x_3)\\ &-U_3(x_1)U_2(x_2)U_1(x_3)-U_1(x_1)U_3(x_2)U_2(x_3)\\ &-U_3(x_1)U_1(x_2)U_2(x_3)-U_2(x_1)U_3(x_2)U_1(x_3)\},\end{aligned} \qquad (2.21a)$$

$$\begin{aligned}\chi_{21}=\tfrac{1}{2}\{&U_1(x_1)U_3(x_2)U_2(x_3)-U_3(x_1)U_2(x_2)U_1(x_3)\\ &+U_2(x_1)U_3(x_2)U_1(x_3)-U_3(x_1)U_1(x_2)U_2(x_3)\},\end{aligned} \qquad (2.21b)$$

$$\begin{aligned}\chi_{12}=\tfrac{1}{2}\{&U_1(x_1)U_3(x_2)U_2(x_3)+U_3(x_1)U_1(x_2)U_2(x_3)\\ &-U_3(x_1)U_2(x_2)U_1(x_3)-U_2(x_1)U_3(x_2)U_2(x_3)\},\end{aligned} \qquad (2.22a)$$

$$\begin{aligned}\chi_{22} = (1/\sqrt{12})\{&2U_1(x_1)U_2(x_2)U_3(x_3)\\ &-2U_2(x_1)U_1(x_2)U_3(x_3)\\ &+U_3(x_1)U_2(x_2)U_1(x_3)-U_2(x_1)U_3(x_2)U_1(x_3)\\ &+U_1(x_1)U_3(x_2)U_2(x_3)-U_3(x_1)U_1(x_2)U_2(x_3)\}\,.\end{aligned} \tag{2.22b}$$

Similarly there are two sets $(\bar{\chi}_{11},\bar{\chi}_{21})$ and $(\bar{\chi}_{12},\bar{\chi}_{22})$ satisfying Eq. (2.19). In fact, we find that we may choose

$$\begin{aligned}\bar{\chi}_{1i} &= \chi_{2i}\\ \bar{\chi}_{2i} &= -\chi_{1i}\end{aligned} \quad (i=1,2)\,. \tag{2.23}$$

Thus there are in fact two wave functions

$$\begin{aligned}\psi_{1s} &= \chi_{11}|a_1\rangle + \chi_{21}|a_2\rangle\,,\\ \psi_{2s} &= \chi_{12}|a_1\rangle + \chi_{22}|a_2\rangle\end{aligned} \tag{2.24}$$

with even symmetry and

$$\begin{aligned}\psi_{1a} &= \bar{\chi}_{11}|a_1\rangle + \bar{\chi}_{21}|a_2\rangle = \chi_{21}|a_1\rangle - \chi_{11}|a_2\rangle\,,\\ \psi_{2a} &= \bar{\chi}_{12}|a_1\rangle + \bar{\chi}_{22}|a_2\rangle = \chi_{22}|a_1\rangle - \chi_{12}|a_2\rangle\end{aligned} \tag{2.25}$$

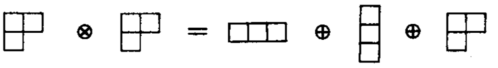
with odd symmetry. However, it may be noted that no permutation can transform one set of functions from (χ_{11},χ_{21}) into the other set (χ_{12},χ_{22}). They belong to different orthogonal spaces. Thus no interaction will connect ψ_{1s} to ψ_{2s} or ψ_{1a} to ψ_{2a}. One may interpret it to say that the degeneracy of the state (r,m) is completely removed.

Analogous considerations follow for the states with $r=\frac{1}{2}$, $m=-\frac{1}{2}$.

In terms of the irreducible representations of S_3, one finds that the wave functions (2.15) generate the two-dimensional irreducible representation

(2.26)

of S_3. The dimensionality of this representation is in fact equal to the degeneracy of the state $r=\frac{1}{2}$, $m=\frac{1}{2}$. The orbital wave functions (2.21) or (2.23) also generate the same irreducible representation (2.26). The orbital wave functions $\bar{\chi}_1,\bar{\chi}_2$ generate an irreducible representation conjugate to Eq. (2.26), which in this particular case happens to be Eq. (2.26) itself. The result that we could construct states (2.18) and (2.20), which generate the irreducible representations (2.11) and (2.12), is a consequence of the fact that the direct product of Eq. (2.26) with itself contains the representations (2.11) and (2.12), i.e.,



(2.27)

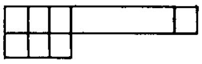
III. *N*-PARTICLE STATE WITH GIVEN SYMMETRY

In this section we wish to consider the N-particle system, where each particle has two levels, and hence behaves like a spin-$\frac{1}{2}$ system. As stated earlier, the general energy eigenstate is a direct product of orbital function and spin-wave function:

$$\psi_{gm}^{(N)} = U_g^{(N)}(\underline{r}_1,\underline{r}_2,\ldots,\underline{r}_N)\phi_{r,m}^{(N)}\,. \tag{3.1}$$

We wish to construct states with given (r,m) and with a given symmetry.

It is well known[6] that the spin wave functions generate an irreducible representation of S_N that corresponds to a Young diagram with at most two rows:



. (3.2)

The total number of squares in the diagram is N and the difference of the number of squares in the two rows is exactly r. The dimensionality of the irreducible representation (3.2) [cf. Ref. 7 and also Eq. (4.9) below] is given by

$$d = N!(2r+1)/((N+2r+2)/2)!((N-r)/2)!\,, \tag{3.3}$$

which is exactly the degeneracy of the state with a given (r,m), i.e., there are d states,

$$|a_i\rangle = \phi_{r,m,i}^{(N)}, \quad i=1,2,\ldots,d\,, \tag{3.4}$$

with a given (r,m). These states may be chosen to be orthogonal. However, these are neither symmetric nor antisymmetric under arbitrary permutations of the N particles but transform linearly among themselves:

$$\hat{P}_{1\alpha}|a_i\rangle = \sum_{j=1}^{d}\Gamma_{ij}^{(\alpha)}|a_j\rangle\,. \tag{3.5}$$

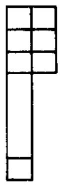
Since all permutations can be obtained by repeated applications of the transposition $\hat{P}_{1\alpha}$ ($\alpha=2,3,\ldots,N$), we have restricted the consideration under such transpositions only. We now assert that it is possible to construct a set of orbital wave functions χ_i and also $\bar{\chi}_i$, $i=1,2,\ldots,d$, such that they obey transformations

$$\hat{P}_{1\alpha}\chi_i = \Gamma_{ji}^{(\alpha)}\chi_j\,, \tag{3.6}$$

$$\hat{P}_{1\alpha}\bar{\chi}_i = -\Gamma_{ji}^{(\alpha)}\bar{\chi}_j\,. \tag{3.7}$$

Since a similar result holds for the more general case of n-level systems, we shall give the proof of this assertion in Sec. IV.

The functions χ_i generate the irreducible representation identical to Eq. (3.2), whereas the function $\bar{\chi}_i$ generates the irreducible representation conjugate to Eq. (3.2), viz., the one given by the Young diagram by interchanging rows and columns of the diagram (3.2):



. (3.8)

The wave function

$$\psi_s = \sum_{i=1}^{d}\chi_i|a_i\rangle \tag{3.9}$$

is then symmetric under all permutations, for

$$\begin{aligned}\hat{P}_{1\alpha}\psi_s &= \sum_{i,j,k}\Gamma_{ji}^{(\alpha)}\Gamma_{ik}^{(\alpha)}\chi_j|a_k\rangle\\ &= \Psi_s\,,\end{aligned} \tag{3.10}$$

where we have used the fact that $\hat{P}_{1\alpha}\hat{P}_{1\alpha}$ is an identity operation so that

$$\sum_{i=1}^{d}\Gamma_{ji}^{(\alpha)}\Gamma_{ik}^{(\alpha)} = \delta_{jk}\,. \tag{3.11}$$

Similarly the wave function

$$\Psi_a = \sum \bar{\chi}_i |a_i\rangle \tag{3.12}$$

is antisymmetric under all odd permutations.

Even though there is only one set of spin wave functions satisfying Eq. (3.5), there are in fact d sets of orbital wave functions (each set containing d wave functions), which obey Eq. (3.6) or Eq. (3.7), i.e.,

$$\hat{P}_{1\alpha}\chi_{ki} = \sum_{j=1}^{d} \Gamma_{ji}^{(\alpha)}\chi_{kj}\,, \tag{3.13}$$

$$\hat{P}_{1\alpha}\bar{\chi}_{ki} = -\sum_{j=1}^{d} \Gamma_{ji}^{(\alpha)}\bar{\chi}_{kj}\,. \tag{3.14}$$

Thus there are in fact d wave functions

$$\sum_{i=1}^{d} \chi_{ki}|a_i\rangle, \quad k = 1,2,\ldots,d, \tag{3.15}$$

which are symmetric under all permutations, and another d wave functions

$$\sum_{i=1}^{d} \bar{\chi}_{ki}|a_i\rangle, \quad k = 1,2,\ldots d, \tag{3.16}$$

which are antisymmetric under all odd permutations. However, no permutation can transform one set of wave functions χ_{ki} $(i = 1,2,\ldots,d)$ (or $\bar{\chi}_{ki}$) into any other set $\chi_{k'i}$ (or $\bar{\chi}_{k'i}$). Therefore no interaction will connect one symmetric (or antisymmetric) wave function to any other symmetric or antisymmetric wave function, and hence the apparent degeneracy of the state with a given (r,m) is completely removed.

IV. N-PARTICLE STATES OF THE n-LEVEL SYSTEM

In this section we consider the N-particle system, where each particle has n levels. We denote these n levels of the jth particle by the state

$$|\lambda_\alpha^{(j)}\rangle, \quad \alpha = 1,2,\ldots,n\,. \tag{4.1}$$

Transitions from state $|\lambda_\alpha^{(j)}\rangle$ to $|\lambda_\beta^{(j)}\rangle$ are described in terms of operators of the type

$$\hat{R}_{\alpha\beta}^{(j)} = |\lambda_\beta^{(j)}\rangle\langle\lambda_\alpha^{(j)}|\,. \tag{4.2}$$

We also have the completeness relation

$$\sum_{\alpha=1}^{n} |\lambda_\alpha^{(j)}\rangle\langle\lambda_\alpha^{(j)}| = 1\,. \tag{4.3}$$

We thus have $n^2 - 1$ such independent operators, which we may associate with the generators[8] of SU(n). In analogy with the Dicke operators of a two-level system, we may also define the generators of SU(n) for the total system,

$$\hat{R}_{\alpha\beta} = \sum_{j=1}^{N} \hat{R}_{\alpha\beta}^{(j)}\,. \tag{4.4}$$

A typical N-particle spin state could be expressed as

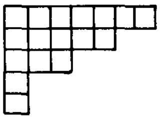
$$\phi^{(N)} = |\lambda_{\alpha_1}^{(1)}, \lambda_{\alpha_2}^{(2)}, \ldots, \lambda_{\alpha_j}^{(j)}, \ldots, \lambda_{\alpha_N}^{(N)}\rangle\,, \tag{4.5}$$

where the jth particle is in the α_jth state. There are obviously n^N such states and they generate an n^N-dimensional representation of the permutation group S_N. This representation is reducible. First, states with N_1 particles occupying the first level, N_2 particles occupying the second level, etc., will transform among themselves and hence they will generate a representation of S_N. The dimensionality of this representation is

$$N!/N_1!N_2!\cdots N_n! \tag{4.6}$$

and the number of such different representations is equal to the number of possibilities of $N_1, N_2, \ldots, N_n$ such that $N_1 + N_2 + \cdots + N_n = N$. Each of these representations of S_N is further reducible, since functions corresponding to given eigenvalues of the Casimir operators of SU(n) will transform among themeselves.

Another way of obtaining the reduction of the n^N dimensional representation of S_N with basis functions (4.5) is in terms of Young's diagrams. Only those diagrams with at most n rows are included in this representation. These representations are of the type

$\Gamma^{(\mu)} =$  , (4.7)

where there are k rows and the first row contains n_1 squares, the second row contains n_2 squares, etc., with the restriction

$$n_1 + n_2 + \cdots + n_k = N\,,$$
$$n_1 \geqslant n_2 \geqslant \cdots \geqslant n_k > 0, \quad k \leqslant n\,. \tag{4.8}$$

The dimensionality of this representation of S_N is given by[7]

$$d_\mu = d_{[n_k]} = \frac{N!}{l_1!l_2!\cdots l_k!} \prod_{1\leqslant i<j\leqslant k} (l_i - l_j)\,, \tag{4.9}$$

where

$$l_i = n_i + k - i \quad (i = 1,2,\ldots,k)\,. \tag{4.10}$$

These representations correspond to definite values of the Casimir operators of SU(n). One may also give an expression for the number of times the representation (4.7) occurs in the n^N-dimensional representation. This number is given by

$$\lambda_{[n_k]} = \prod_{i=1}^{k} \frac{(n_i + n - i)!}{l_i!(n-i)!} \prod_{1\leqslant i<j\leqslant k} (l_i - l_j)\,. \tag{4.11}$$

This in fact is the dimensionality of the irreducible representation of SU(n) corresponding to Young's diagram (4.7). Notice that Eq. (4.11) implies

$$\lambda_{[n_k]} = 0, \quad \text{if } k > n\,, \tag{4.12}$$

in conformity with the requirement that the Young diagrams with the number of rows greater than the number of levels are not allowed. One may also verify that

$$\sum_{[n_k]} d_{[n_k]}\lambda_{[n_k]} = n^N\,, \tag{4.13}$$

where summation is over all partitions $[n_k]$ of N satisfying Eq. (4.8).

We now wish to take the direct product of the set of spin wave functions that generate the irreducible representation (4.7) with a suitable set of orbital wave functions such that the combined wave function has a definite symmetry. However, before this we obtain an expression for the projection operator, which when acting on $\phi^{(N)}$ takes out the part that corresponds to the irreducible representation (4.7) of S_N.

Let $\hat{K}_{ij}^{(\mu)}$ be a linear combination of permutation operators defined by

$$\hat{K}_{ij}^{(\mu)} = \frac{d_\mu}{N!} \sum_{\hat{P}} \Gamma_{ij}^{(\mu)}(\hat{P}) \times \hat{P}, \tag{4.14}$$

where $\Gamma^{(\mu)}(P)$ are the matrices of the μth irreducible representation of S_N and the sum over $\hat{P}$ includes all permutations $\hat{P}$ of S_N. The subscripts (ij) on $\Gamma^{(\mu)}$ denote the (i,j)th matrix element ($1 < i, j < d$, where d is the dimensionality of the μth representation). We may readily verify the following properties of $\hat{K}_{ij}^{(\mu)}$:

$$\hat{K}_{ij}^{(\mu)} \hat{K}_{kl}^{(\nu)} = \delta_{\mu\nu} \delta_{jk} \hat{K}_{il}^{(\mu)}, \tag{4.15}$$

which follows from the basic orthogonality relations of the irreducible representations, viz.,

$$\sum_P \Gamma_{ij}^{(\mu)}(P) \Gamma_{kl}^{(\nu)}(P^{-1}) = \frac{N!}{d_\mu} \delta_{\mu\nu} \delta_{il} \delta_{jk}. \tag{4.16}$$

From Eq. (4.15) we also obtain a number of other relations, such as

$$\hat{K}_{ii}^{(\mu)} \hat{K}_{kl}^{(\nu)} = \delta_{\mu\nu} \delta_{ik} \hat{K}_{il}^{(\mu)}, \tag{4.17}$$

$$\hat{K}_{ij}^{(\mu)} \hat{K}_{kk}^{(\nu)} = \delta_{\mu\nu} \delta_{jk} \hat{K}_{ij}^{(\mu)}, \tag{4.18}$$

$$\hat{K}_{ij}^{(\mu)} \hat{K}_{jl}^{(\nu)} = \delta_{\mu\nu} \delta_{ij} \hat{K}_{ii}^{(\mu)}, \tag{4.19}$$

$$\hat{K}_{ij}^{(\mu)} \hat{K}_{ij}^{(\nu)} = \delta_{\mu\nu} \delta_{ij} \hat{K}_{ii}^{(\mu)}, \tag{4.20}$$

and

$$\hat{K}_{ii}^{(\mu)} \hat{K}_{jj}^{(\nu)} = \delta_{\mu\nu} \delta_{ij} \hat{K}_{ii}^{(\mu)}. \tag{4.21}$$

(No summation over repeated indices.)

The last relation (4.21) is of particular interest, since it shows that $\hat{K}_{ii}^{(\mu)}$ is a projection operator. Further on, setting $i = j$ in Eq. (4.14), summing over i and then over μ, and using the completeness property of the character of irreducible representations, we also obtain the completeness relation

$$\sum_\mu \sum_{i=1}^{d_\mu} \hat{K}_{ii}^{(\mu)} = 1. \tag{4.22}$$

We also find from definition (4.14) that

$$\begin{aligned} \hat{P}\hat{K}_{ij}^{(\mu)} &= \frac{d_\mu}{N!} \sum_{\hat{R}} \Gamma_{ij}^{(\mu)}(\hat{R}) \hat{P}\hat{R} \\ &= \frac{d_\mu}{N!} \sum_k \Gamma_{ik}^{(\mu)}(\hat{P}^{-1}) \sum_R \Gamma_{kj}^{(\mu)}(\hat{P}\hat{R}) \times (\hat{P}\hat{R}) \\ &= \sum_k \Gamma_{ik}^{(\mu)}(\hat{P}^{-1}) \hat{K}_{kj}^{(\mu)}. \end{aligned} \tag{4.23}$$

Thus the operator $\hat{K}_{ij}^{(\mu)}$ can be used to obtain the wave functions that would generate a given irreducible representation. We operate $\hat{K}_{ij}^{(\mu)}$ on the state $\phi^{(N)}$, Eq. (4.5), and obtain

$$\phi_{ij}^{(\mu)} = \hat{K}_{ij}^{(\mu)} \phi^{(N)}. \tag{4.24}$$

Because of Eq. (4.23) we find that any permutation of the N particles on this state gives a linear superposition of the state $\phi_{kj}^{(\mu)}$ (with j fixed):

$$\hat{P}\phi_{ij}^{(\mu)} = \sum_k \Gamma_{ik}^{(\mu)}(P^{-1}) \phi_{kj}^{(\mu)}. \tag{4.25}$$

Hence $\phi_{1j}^{(\mu)}$, $\phi_{2j}^{(\mu)}$,..., with a fixed j, generate the μth irreducible representation of S_N. One may readily verify that these functions are in fact orthogonal:

$$\langle \phi_{ij}^{(\mu)} | \phi_{kl}^{(\mu)} \rangle = \delta_{\mu\nu} \delta_{ik} (\phi^{(N)}, \phi_{jl}^{(N)}). \tag{4.26}$$

One point needs clarification at this stage, regarding the number of linearly independent sets of wave functions $(\phi_{1j}, \phi_{2j}, \ldots, \phi_{d\mu j})$ that can generate a given irreducible representation. Obviously the maximum number of such sets is d_μ corresponding to each different value of j, and this will happen if, and only if, each in Eq. (4.5) is different (which incidentally requires that the number of levels available is at least N). On the other hand, it may also happen that no set of ϕ's exists which can generate a given irreducible representation. This in fact will happen if the irreducible representation under consideration corresponds to a Young diagram having a number of rows greater than n (all $\phi_{ij}^{(\mu)}$ are zero in this case). The number of linearly independent sets of $(\phi_{1j}, \phi_{2j}, \ldots)$ generating a given irreducible representation is exactly the same as the number of times the given representation occurs in the $(N!/N_1!N_2!\cdots N_n!)$-dimensional space of wave functions with N_1 particles occupying the first level, N_2 particles occupying the second level, etc. [This number is different from the expression (4.11), which represents the number of times a given irreducible representation occurs in the n-dimensional space of wave functions of type (4.5) without any restriction on how many are occupying level 1, how many are occupying level 2, etc.]

We have discussed how we can generate a given irreducible representation of S_N using spin wave functions. Analogous considerations hold for the orbital wave functions. Here there is no restriction on the number of energy levels. A typical wave function is given by

$$U^{(N)} = U_1(x_1) U_2(x_2) \cdots U_N(x_N). \tag{4.27}$$

There are $N!$ such wave functions that can be obtained by various permutations of the N particles. This set of wave functions generates a $N!$-dimensional representation of S_N, which is obviously reducible. It follows fom the earlier considerations that the functions

$$U_{ij}^{(\mu)} = \hat{K}_{ij}^{(\mu)} U^{(N)}, \tag{4.28}$$

with a fixed j, will generate the μth irreducible representation of S_N.

Consider now the wave function

$$\Psi_S = \sum_i U_{ij}^{(\mu)} \phi_{im}^{(\mu)}. \tag{4.29}$$

One may readily verify that this state is symmetric under all permutations, for using Eq. (4.25),

$$\begin{aligned} \hat{P}\Psi_S &= \sum_i \{\hat{P}U_{ij}^{(\mu)}\}\{\hat{P}\phi_{im}^{(\mu)}\} \\ &= \sum_{i,k,l} \Gamma_{ik}^{(\mu)}(\hat{P}^{-1}) \Gamma_{il}^{(\mu)}(\hat{P}^{-1}) U_{kj}^{(\mu)} \phi_{lm} \\ &= \sum_k U_{kj}^{(\mu)} \phi_{km} = \Psi_S. \end{aligned} \tag{4.30}$$

We obtain different symmetric functions by taking different values of j and m. However, as remarked earlier, no permutation can connect one set to the other, and as such we may regard it as a nondegenerate state.

We also define a set of operators

$$\hat{\hat{K}}_{ij}^{(\mu)} = \frac{d\mu}{N!} \sum_{\hat{P}} \delta_P \Gamma_{ij}^{(\mu)}(\hat{P}) \times \hat{P}, \tag{4.31}$$

where δ_P is $+1$ or -1 according as P is an even or an odd permutation. Notice that

$$\hat{\hat{K}}_{ij}^{(\mu)} = \hat{K}_{ij}^{(\bar{\mu})},$$

where $\bar{\mu}$ is the conjugate representation, i.e., the one which is obtained by interchanging rows and columns of the corresponding Young's diagram.) The wave functions

$$\bar{U}_{ij}^{(\mu)} = \hat{\hat{K}}_{ij}^{(\mu)} U^{(N)}, \tag{4.32}$$

with a fixed j, will generate the $\bar{\mu}$th irreducible representation of S_N. The desired antisymmetric wave function is then given by

$$\Psi_A = \sum_i \bar{U}_{ij}^{(\mu)} \phi_{im}^{(\mu)}. \tag{4.33}$$

We have thus been able to obtain the N-particle symmetric or antisymmetric wave function for the n-level system corresponding to given values of the spin parameters. In Sec. V we consider special cases of Eqs. (4.29) and (4.33).

V. SPECIAL CASES

We now consider some special cases of Eqs. (4.29) and (4.33) as applied to two- and three-particle systems.

A. Two-particle system

In this case only two irreducible representations are applicable. These correspond to Young diagrams  and . The $\hat{K}$ operators are given by

$$\hat{K}_{11}^{(1)} = \tfrac{1}{2}[e + (12)], \tag{5.1}$$

$$\hat{K}_{11}^{(2)} = \tfrac{1}{2}[e - (12)]. \tag{5.2}$$

Notice that $\hat{\hat{K}}^{(1)} = \hat{K}^{(2)}$. Thus two totally symmetric and two totally antisymmetric functions are available:

$$\Psi_S = [\hat{K}^{(1)}|\lambda_1,\lambda_2\rangle \hat{K}^{(1)} U_1(\underline{r}_1) U_2(\underline{r}_2)] \tag{5.3}$$

or

$$= [\hat{K}^{(2)}|\lambda_1,\lambda_2\rangle \hat{K}^{(2)} U_1(\underline{r}_1) U_2(\underline{r}_2)]; \tag{5.4}$$

and

$$\Psi_A = [\hat{K}^{(1)}|\lambda_1,\lambda_2\rangle \hat{K}^{(2)} U_1(\underline{r}_1) U_2(\underline{r}_2)] \tag{5.5}$$

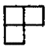
or

$$= [\hat{K}^{(2)}|\lambda_1,\lambda_2\rangle \hat{K}^{(1)} U_1(\underline{r}_1) U_2(\underline{r}_2)]. \tag{5.6}$$

If $\lambda_1 = \lambda_2$, then $\hat{K}^{(2)}|\lambda_1,\lambda_2\rangle = 0$, and we have only one function of each type.

B. Three-particle system

In this case we have three irreducible representations corresponding to the diagrams:

 ,  , and  . (5.7)

The $\hat{K}$ operators are given by

$$\hat{K}^{(1)} = \tfrac{1}{6}[e + (12) + (13) + (23) + (123) + (132)], \tag{5.8}$$

$$\hat{K}_{11}^{(2)} = \tfrac{1}{3}[e + (12) - \tfrac{1}{2}(13) - \tfrac{1}{2}(23) - \tfrac{1}{2}(123) - \tfrac{1}{2}(132)], \tag{5.9}$$

$$\hat{K}_{21}^{(2)} = (1/2\sqrt{3})[-(13) + (23) - (123) + (132)], \tag{5.10}$$

$$\hat{K}_{12}^{(2)} = (1/2\sqrt{3})[-(13) + (23) + (123) - (132)], \tag{5.11}$$

$$\hat{K}_{22}^{(2)} = \tfrac{1}{3}[e - (12) + \tfrac{1}{2}(13) + \tfrac{1}{2}(23) - \tfrac{1}{2}(123) - \tfrac{1}{2}(132)], \tag{5.12}$$

$$\hat{K}^{(3)} = \tfrac{1}{6}[e - (12) - (13) - (23) + (123) + (132)]. \tag{5.13}$$

Here $\hat{\hat{K}}^{(1)} = \hat{K}^{(3)}$ and $\hat{\hat{K}}^{(2)} = \hat{K}^{(2)}$ (actually $\hat{\hat{K}}_{11}^{(2)} = \hat{K}_{22}^{(2)}$, etc.). If we are considering a two-level system, then we will have spin wave functions of the type $|++-\rangle$, etc. In this case $\hat{K}^{(3)}$ operating on such wave functions will give zero. Operation of $\hat{K}^{(1)}$ will give functions of the type (2.2)–(2.5). The terms $\hat{K}_{11}^{(2)}$ and $\hat{K}_{21}^{(2)}$ operating on, say, $|++-\rangle$, will give functions (2.15a) and (2.15b), whereas $\hat{K}_{22}^{(2)}|++-\rangle$ and $\hat{K}_{12}^{(2)}|++-\rangle$ are both zero. One may readily verify that

$$\chi_{11} = \hat{K}_{11}^{(2)} U_1(x_1) U_2(x_2) U_3(x_3),$$

$$\chi_{21} = \hat{K}_{21}^{(2)} U_1(x_1) U_2(x_2) U_3(x_3),$$

etc.

The results of Sec. II are thus reproduced. Similar results may readily be obtained for the three-level systems. As the number of particles are increased, the explicit construction of wave functions with a given symmetry becomes more and more difficult. However, their existence is demonstrated and one may obtain some general properties of such states using the known results of the symmetry group.

ACKNOWLEDGMENT

This work was supported in part by a grant from the Alabama Research Institute.

[1] J. Rai, C. L. Mehta, and N. Mukunda, J. Math. Phys. **29**, 510 (1988).
[2] R. H. Dicke, Phys. Rev. **93**, 99 (1954).
[3] M. Gross and S. Haroche, Phys. Rep. **93**, 301 (1982), and references cited therein.
[4] J. Rai, C. L. Mehta, and N. Mukanda, *Proceedings of the Fifth Rochester Conference*, edited by L. Mandel and E. Wolf (Plenum, New York, 1984), p. 97.
[5] I. G. Kaplan, *Symmetry of Many Electron Systems* (Academic, New York, 1975).
[6] E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959).
[7] H. Boerner, *Group Representation* (Springer, Berlin, 1955).
[8] J. Rai and C. L. Mehta, Opt. Commun. **42**, 113 (1982); F. T. Hioe and J. H. Eberly, Phys. Rev. Lett. **47**, 838 (1981).

The unitarity relations for the four-body scattering amplitude

Yoshiko Matsui
Department of Applied Physics, Tokyo University of Agriculture and Technology, Naka-machi, Koganei-shi, Tokyo, Japan

(Received 10 September 1987; accepted for publication 13 June 1988)

A formal derivation of the general unitarity relation for the four-particle transition operator is given by generalizing the three-body formalism of Karlson and Zeiger to the four-body case. From this operator relation the on-shell unitarity relations for the amplitudes that describe elastic/rearrangement, partial breakup, and full breakup scattering processes are obtained.

I. INTRODUCTION

Faddeev[1] established the three-body scattering theory in which he gave the unitarity relation for the resolvent of the total Hamiltonian. Since his work, various approaches to derive the unitarity condition for the three-particle transition amplitudes have been proposed by many authors[2–5] in the formal sense and by a mathematical author[6] rigorously.

The scattering equations for the N-body problem have been presented by Yakubovskii,[7] who generalized the Faddeev treatment to the N-body case. Alt, Grassberger, and Sandhas[8] (AGS) succeeded in finding appropriate scattering equations. In this AGS approach, N-particle equations are written in an N-body matrix version of the corresponding two-body relations. Based on these two formalisms, Karlson and Zeiger[5] (KZ) constructed four-particle equations expressed in terms of singularity-free physical transition amplitudes that were obtained by a thorough singularity analysis of the Faddeev kernel. However, the study of the unitarity relations for those four-body scattering amplitudes has not received much attention.

When we embark on the numerical analysis of the scattering equations, the theory demands that we obtain the physical amplitudes which satisfy their unitarity conditions; this is the exact approach to the scattering problem. In fact, the unitarity relations play an important role in checking the calculating system, especially whether or not the normalization of the scattering equation is successful.[9]

Here I would like to derive the unitarity relation for the four-particle scattering amplitudes introduced by KZ, using Yakubovskii's equations. The present work will only give the procedure of the formal operational calculus without the rigorous argument of the problem, such as the proof that the boundary values of the scattering amplitudes exist at the right hand cut. Although the present paper has heuristic value only, it may give further stimulus to the perfect theory.

In Sec. II, the general unitarity relation in operator form is derived, using the discontinuities of the transition operators for the subsystems given in the Appendix. From this operator relation I obtain in Sec. III on-shell unitarity relations for the amplitudes that describe elastic/rearrangement, partial breakup, and full breakup processes.

II. THE GENERAL UNITARITY RELATION IN OPERATOR FORM

In this section I derive the general unitarity relation in operator form for the four-particle KZ operator starting at the Faddeev–Yakubovskii (FY) equation.

Throughout, the same notation for operators and kinematic variables as one finds in KZ will be used. Four-particle indices (lowercase letters: $a,b,\ldots$) will refer to different types of the seven possible partitions of four particles into two groups: (123)(4), (421)(3), (341)(2), (432)(1), (12)(34), (13)(24), (14)(23). These partitions denote (3 + 1) or (2 + 2) subsystems in the four-body system. The indices α, β, and γ run over all possible values: 12, 13, 23, 14, 24, 34. These pair indices will usually appear as subordinate indices, in the sense that they label interacting pairs within a certain channel of some type a. In such a case we write $\alpha \subset a$.

Let us denote four-particle operators by capital letters. From Faddeev's definition, the transition amplitude $M^{a}_{\beta\alpha}$ $(\beta,\alpha \subset a)$ for a (3 + 1) subsystem can be denoted by

$$M^{a}_{\beta\alpha} = \delta_{\beta\alpha} V_\beta - V_\beta G^a V_\alpha,$$

with pairwise interactions V_α and V_β and resolvent $G^a(z) = (H_a - z)^{-1}$ of the total Hamiltonian $H_a = H_0 + \Sigma_{\gamma\subset a} V_\gamma$. This amplitude obeys the equation

$$M^{a}_{\beta\alpha} = \delta_{\beta\alpha} T_\beta - T_\beta G_0 \sum_{\gamma\subset a} \bar{\delta}_{\beta\gamma} M^{a}_{\gamma\alpha}, \tag{1}$$

where $\bar{\delta}_{\beta\alpha} = 1 - \delta_{\beta\alpha}$, G_0 is the resolvent of H_0, and T_β is the T matrix defined as $T_\beta = V_\beta - V_\beta G_\beta V_\beta$ with the resolvent $G_\beta(z) = (H_\beta - z)^{-1}$. The three-body connected part of $M^{a}_{\beta\alpha}$ is defined as

$$W^{a}_{\beta\alpha} = M^{a}_{\beta\alpha} - \delta_{\beta\alpha} T_\beta,$$

and satisfies a system resembling (1),

$$W^{a}_{\beta\alpha} = T_\beta G_0 T_\alpha \bar{\delta}_{\beta\alpha} - T_\beta G_0 \sum_{\gamma\neq\beta} W^{a}_{\gamma\alpha}. \tag{2}$$

The formulation described above holds also for the (2 + 2) subsystem.

Generalizing the three-body formalism, Yakubovskii constructed a symmetric four-particle operator $M^{ba}_{\beta\alpha}$ $(\beta \subset b, \alpha \subset a)$,

$$M^{ba}_{\beta\alpha} = T_\beta G_0 T_\alpha \bar{\delta}_{\beta\alpha} \delta^{ba} + \sum_{\gamma\subset b} \sum_{\delta\subset a} T_\beta G_0 \bar{\delta}_{\beta\gamma} T^{\gamma,\delta} \bar{\delta}_{\delta\alpha} G_0 T_\alpha, \tag{3}$$

where $T^{\gamma,\delta} = V_\gamma \delta_{\gamma\delta} + V_\gamma G V_\delta$ and G is the resolvent of the four-body total Hamiltonian.

As in the three-body case, we note that $W^{ba}_{\beta\alpha} = M^{ba}_{\beta\alpha} - \delta^{ba} W^{a}_{\beta\alpha}$ is the four-body connected part of $M^{ba}_{\beta\alpha}$, and obeys the following equation:

0022-2488/88/112450-12$02.50

$$W^{ba}_{\beta\alpha} = \sum_{\gamma,\delta\subset b} \bar{\delta}^{ba}\bar{\delta}_{\gamma\delta} M^{b}_{\beta\gamma} G_0 W^{a}_{\delta\alpha}$$

$$+ \sum_{\substack{\gamma,\delta\subset b\\ c\supset\delta}} \bar{\delta}^{bc}\bar{\delta}_{\gamma\delta} M^{b}_{\beta\gamma} G_0 W^{ca}_{\delta\alpha}. \tag{4}$$

In order to obtain a more appropriate operator to define the physical amplitude, we proceed with the formalism of KZ.

Consider the Lovelace and AGS operators $U^{a(\pm)}_{\beta\alpha}$ and $U^{a}_{\beta\alpha}$, respectively, which are defined as follows:

$$U^{a(+)}_{\beta\alpha} = \sum_{\gamma\neq\beta} V_\gamma - \sum_{\gamma\neq\beta}\sum_{\delta\neq\alpha} V_\gamma G^{a} V_\delta,$$

$$U^{a(-)}_{\beta\alpha} = \sum_{\delta\neq\alpha} V_\delta - \sum_{\gamma\neq\beta}\sum_{\delta\neq\alpha} V_\gamma G^{a} V_\delta,$$

$$U^{a}_{\beta\alpha} = -\delta_{\beta\alpha}(H_\alpha - z) + U^{a(+)}_{\beta\alpha}$$

$$= -\delta_{\beta\alpha}(H_\beta - z) + U^{a(-)}_{\beta\alpha}.$$

Their equations are

$$\begin{aligned} U^{a(+)}_{\beta\alpha} &= \sum_{\gamma\neq\beta} V_\gamma - \sum_{\delta\neq\alpha} U^{a(+)}_{\beta\delta} G_0 T_\delta, \\ U^{a(-)}_{\beta\alpha} &= \sum_{\delta\neq\alpha} V_\delta - \sum_{\gamma\neq\beta} T_\gamma G_0 U^{a(-)}_{\gamma\alpha}, \\ U^{a}_{\beta\alpha} &= -\bar{\delta}_{\beta\alpha} G_0^{-1} - \sum_{\gamma\neq\beta} T_\gamma G_0 U^{a}_{\gamma\alpha}, \\ U^{a}_{\beta\alpha} &= -\bar{\delta}_{\beta\alpha} G_0^{-1} - \sum_{\gamma\neq\alpha} U^{a}_{\beta\gamma} G_0 T_\gamma. \end{aligned} \tag{5}$$

The AGS operator $U^{a}_{\beta\alpha}$ is related to $W^{a}_{\beta\alpha}$ by the equation

$$W^{a}_{\beta\alpha} = T_\beta G_0 U^{a}_{\beta\alpha} G_0 T_\alpha.$$

The four-particle AGS operator corresponding to $U^{a}_{\beta\alpha}$ is expressed by $U^{ba}_{\beta\alpha}$ and has the following relation with the FY operator $W^{ba}_{\beta\alpha}$:

$$W^{ba}_{\beta\alpha} = \sum_{\gamma\subset b}\sum_{\delta\subset a} W^{b}_{\beta\gamma} G_0 U^{ba}_{\gamma\delta} G_0 W^{a}_{\delta\alpha}.$$

Finally, defining our operator $H^{ba}_{\beta\alpha}$ by

$$H^{ba}_{\beta\alpha} = \widetilde{T}_\beta U^{ba}_{\beta\alpha} \widetilde{T}_\alpha, \quad \text{with } \widetilde{T}_\beta = G_0 T_\beta G_0,$$

we arrive at the system of equations for $H^{ba}_{\beta\alpha}$,

$$H^{ba}_{\beta\alpha} = \widetilde{T}_\beta \delta_{\beta\alpha} \bar{\delta}^{ba} + \sum_{c\supset\beta} \bar{\delta}^{bc} \sum_{\gamma\subset c} \widetilde{T}_\beta U^{c}_{\beta\gamma} H^{ca}_{\gamma\alpha}. \tag{6}$$

I would like to start with this system of equations to derive the unitarity relation in operator form.

Within this paragraph, it is convenient to make the no-

TABLE I. The order in which to arrange the 12 indices (a,α) of the elements of matrix A, with the number of rows or columns

No.	1	2	3	4	5	6	7	8	9	10	11	12
a	4	4	4	3	3	3	2	2	2	1	1	1
α	12	13	23	12	14	42	13	14	34	23	24	34

TABLE II. The order in which to arrange the six indices (k,κ) of the elements of matrix B, with number of rows.

No.	1	2	3	4	5	6
b	12	13	23	14	24	34
β	12	13	23	14	24	34

TABLE III. The elements of matrix T.

No.	1	2	3	4	5	6	7	8	9	10	11	12
1				$\widetilde{T}_{12}$								
2							$\widetilde{T}_{13}$					
3										$\widetilde{T}_{32}$		
4	$\widetilde{T}_{12}$											
5								$\widetilde{T}_{14}$				
6											$\widetilde{T}_{42}$	
7		$\widetilde{T}_{13}$										
8					$\widetilde{T}_{14}$							
9												$\widetilde{T}_{34}$
10			$\widetilde{T}_{23}$									
11						$\widetilde{T}_{24}$						
12									$\widetilde{T}_{34}$			

TABLE IV. The elements of matrix T'.

No.	1	2	3	4	5	6	7	8	9	10	11	12
1	$\widetilde{T}_{12}$			$\widetilde{T}_{12}$								
2		$\widetilde{T}_{13}$					$\widetilde{T}_{13}$					
3			$\widetilde{T}_{23}$							$\widetilde{T}_{32}$		
4					$\widetilde{T}_{14}$			$\widetilde{T}_{14}$				
5						$\widetilde{T}_{24}$					$\widetilde{T}_{24}$	
6									$\widetilde{T}_{34}$			$\widetilde{T}_{34}$

TABLE V. The elements of matrix T''.

No.	1	2	3	4	5	6
1	$\widetilde{T}_{12}$					
2		$\widetilde{T}_{13}$				
3			$\widetilde{T}_{23}$			
4	$\widetilde{T}_{12}$					
5				$\widetilde{T}_{14}$		
6					$\widetilde{T}_{42}$	
7		$\widetilde{T}_{13}$				
8				$\widetilde{T}_{14}$		
9						$\widetilde{T}_{34}$
10			$\widetilde{T}_{23}$			
11					$\widetilde{T}_{24}$	
12						$\widetilde{T}_{34}$

TABLE VI. The elements of matrix U.

No.	1	2	3	4	5	6	7	8	9	10	11	12
1	$U^4_{12,12}$	$U^4_{12,13}$	$U^4_{12,23}$									
2	$U^4_{13,12}$	$U^4_{13,13}$	$U^4_{13,23}$									
3	$U^4_{23,12}$	$U^4_{23,13}$	$U^4_{23,23}$									
4				$U^3_{12,12}$	$U^3_{12,14}$	$U^3_{12,24}$						
5				$U^3_{14,12}$	$U^3_{14,14}$	$U^3_{14,24}$						
6				$U^3_{24,12}$	$U^3_{24,14}$	$U^3_{24,24}$						
7							$U^2_{13,13}$	$U^2_{13,41}$	$U^2_{13,34}$			
8							$U^2_{14,13}$	$U^2_{14,41}$	$U^2_{14,34}$			
9							$U^2_{34,13}$	$U^2_{34,41}$	$U^2_{34,34}$			
10										$U^1_{23,23}$	$U^1_{23,42}$	$U^1_{23,34}$
11										$U^1_{42,23}$	$U^1_{42,42}$	$U^1_{42,34}$
12										$U^1_{34,23}$	$U^1_{34,42}$	$U^1_{34,34}$

TABLE VII. The elements of matrix U'.

No.	1	2	3	4	5	6
1	$U^{12}_{12,12}$					$U^{12}_{12,34}$
2		$U^{13}_{13,13}$			$U^{13}_{13,24}$	
3			$U^{23}_{23,23}$	$U^{23}_{23,14}$		
4			$U^{14}_{14,23}$	$U^{14}_{14,14}$		
5		$U^{24}_{24,13}$			$U^{24}_{24,24}$	
6	$U^{34}_{34,12}$					$U^{34}_{34,34}$

TABLE VIII. The elements of matrix Ω. All elements need to be multiplied by a factor of $\frac{1}{2}$.

No.	1	2	3	4	5	6	7	8	9	10	11	12
1	$-\tilde{T}_{12}^{-1}$			$\tilde{T}_{12}^{-1}$								
2		$-\tilde{T}_{13}^{-1}$					$\tilde{T}_{13}^{-1}$					
3			$-\tilde{T}_{23}^{-1}$							$\tilde{T}_{23}^{-1}$		
4	$\tilde{T}_{12}^{-1}$			$-\tilde{T}_{12}^{-1}$								
5					$-\tilde{T}_{14}^{-1}$			$\tilde{T}_{14}^{-1}$				
6						$-\tilde{T}_{24}^{-1}$					$\tilde{T}_{24}^{-1}$	
7		$\tilde{T}_{13}^{-1}$					$-\tilde{T}_{13}^{-1}$					
8					$\tilde{T}_{14}^{-1}$			$-\tilde{T}_{14}^{-1}$				
9									$-\tilde{T}_{34}^{-1}$			$\tilde{T}_{34}^{-1}$
10			$\tilde{T}_{23}^{-1}$							$-\tilde{T}_{23}^{-1}$		
11						$\tilde{T}_{24}^{-1}$					$-\tilde{T}_{24}^{-1}$	
12									$\tilde{T}_{34}^{-1}$			$-\tilde{T}_{34}^{-1}$

TABLE IX. The elements of matrix Ω'. All elements need to be multiplied by a factor of $\frac{1}{2}$.

No.	1	2	3	4	5	6	7	8	9	10	11	12
1	$\tilde{T}_{12}^{-1}$			$\tilde{T}_{12}^{-1}$								
2		$\tilde{T}_{13}^{-1}$					$\tilde{T}_{13}^{-1}$					
3			$\tilde{T}_{23}^{-1}$							$\tilde{T}_{23}^{-1}$		
4					$\tilde{T}_{14}^{-1}$			$\tilde{T}_{14}^{-1}$				
5						$\tilde{T}_{24}^{-1}$					$\tilde{T}_{24}^{-1}$	
6									$\tilde{T}_{34}^{-1}$			$\tilde{T}_{34}^{-1}$

tations more specific, denoting by e and h the different $(3+1)$ subsystems, i.e., the partition of the type $(\cdots)(\cdot)$, and by k and l the different $(2+2)$ subsystems, i.e., the partition $(\cdot\cdot)(\cdot\cdot)$. Then Eq. (6) can be separated into four groups, relating to the four types of amplitudes: $H^{he}_{\beta\alpha}$, $H^{ke}_{\kappa\alpha}$, $H^{hl}_{\beta\rho}$, and $H^{kl}_{\kappa\rho}$, respectively. Following Narodetsukii's[10] notation, we express these four groups by matrices A, B, C, and D, where $A=(H^{he}_{\beta\alpha})$, $B=(H^{ke}_{\kappa\alpha})$, etc. For simplicity, let the superscripts e and h denote the value $(\cdot)$ of the type $(\cdots)(\cdot)$, and let k and l denote either value $(\cdot\cdot)$ of the type $(\cdot\cdot)(\cdot\cdot)$. There are 12 possible sets of values for each pair of indices, (e,α) [or (h,β)] of $H^{he}_{\beta\alpha}$ because the pair has the condition $\alpha\subset e$ $(\beta\subset h)$. If we give a number to each pair of indices, (e,α) and (h,β), following the way given in Table I, the matrix A can be expressed in a square array of (12×12) elements $H^{he}_{\beta\alpha}$ arranged in 12 rows and 12 columns; thus

$$A=\begin{pmatrix} H^{4,4}_{12,12} & H^{4,4}_{12,13} & \cdots & H^{4,1}_{12,34} \\ H^{4,4}_{13,12} & H^{4,4}_{13,13} & \cdots & H^{4,1}_{13,34} \\ & \cdots & & \\ H^{1,4}_{34,12} & H^{1,4}_{34,13} & \cdots & H^{1,1}_{34,34} \end{pmatrix}.$$

While 12 columns of the (6×12) matrix B are arranged in the above order, six rows are arranged in the order given in Table II. The matrices C and D are also arranged in the same manner as A and B. If we show the matrices T,T',T'',U,U' in Tables III–VII, the system of equations (6) can be symbolically written in the following matrix form:

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix}=\begin{pmatrix} T & T'' \\ T' & 0 \end{pmatrix} +\begin{pmatrix} T & T'' \\ T' & 0 \end{pmatrix}\begin{pmatrix} U & 0 \\ 0 & U' \end{pmatrix}\begin{pmatrix} A & C \\ B & D \end{pmatrix}. \quad (7)$$

Following the AGS method suggested in Ref. 4, we can express the above equation in more simple matrix form,

$$\mathsf{H}=\mathsf{T}+\mathsf{T}\cdot\mathsf{U}\cdot\mathsf{H} \quad (8)$$

by introducing the matrices H, T, and U,

TABLE X. The elements of matrix Ω''. All elements need to be multiplied by a factor of $\frac{1}{2}$.

No.	1	2	3	4	5	6
1	$\widetilde{T}_{12}^{-1}$					
2		$\widetilde{T}_{13}^{-1}$				
3			$\widetilde{T}_{23}^{-1}$			
4	$\widetilde{T}_{12}^{-1}$					
5				$\widetilde{T}_{14}^{-1}$		
6					$\widetilde{T}_{24}^{-1}$	
7		$\widetilde{T}_{13}^{-1}$				
8				$\widetilde{T}_{14}^{-1}$		
9						$\widetilde{T}_{34}^{-1}$
10			$\widetilde{T}_{23}^{-1}$			
11					$\widetilde{T}_{24}^{-1}$	
12						$\widetilde{T}_{34}^{-1}$

$$\mathsf{H}=\begin{pmatrix} A & C \\ B & D \end{pmatrix},\quad \mathsf{T}=\begin{pmatrix} T & T'' \\ T' & 0 \end{pmatrix},\quad \mathsf{U}=\begin{pmatrix} U & 0 \\ 0 & U' \end{pmatrix}.$$

If we multiply by T^{-1} from the left and by H^{-1} from the right, Eq. (8) can be rewritten in the form

$$\mathsf{H}^{-1}(z)=\mathsf{T}^{-1}(z)-\mathsf{U}(z). \tag{9}$$

If we write Eq. (9) for $z=E+i0$ and $z=E-i0$, and get the discontinuity by subtraction and multiplication by $\mathsf{H}(E+i0)$, $\mathsf{H}(E-i0)$ from the left and right, we can get the following equation:

$$\begin{aligned}\mathsf{H}(E+i0)-\mathsf{H}(E-i0)&\\ =\mathsf{H}(E+i0)\{\mathsf{U}(E+i0)&-\mathsf{U}(E-i0)\}\mathsf{H}(E-i0)\\ -\mathsf{H}(E+i0)\{\mathsf{T}^{-1}&(E+i0)\\ -\mathsf{T}^{-1}(E-i0)\}&\mathsf{H}(E-i0).\end{aligned} \tag{10}$$

If we denote $A(E+i0)$ and $A(E-i0)$ by A_+ and A_- and likewise for B, C, D, U, and Ω, Eq. (10) can be expressed in terms of each element; for example,

$$\begin{aligned}A_+-A_-&\\ =A_+(U_+-U_-)A_-&+C_+(U'_+-U'_-)B_-\\ -A_+(\Omega_+-\Omega_-)A_-&-A_+(\Omega''_+-\Omega''_-)B_-\\ -C_+(\Omega'_+-\Omega'_-)A_-&-C_+(\Omega'''_+-\Omega'''_-)B_-.\end{aligned} \tag{11}$$

TABLE XI. The elements of matrix Ω'''. All elements need to be multiplied by a factor of $\frac{1}{2}$.

No.	1	2	3	4	5	6
1	$-\widetilde{T}_{12}^{-1}$					
2		$-\widetilde{T}_{13}^{-1}$				
3			$-\widetilde{T}_{23}^{-1}$			
4				$-\widetilde{T}_{14}^{-1}$		
5					$-\widetilde{T}_{24}^{-1}$	
6						$-\widetilde{T}_{34}^{-1}$

Here, we denote the inverse operator T^{-1} by

$$\mathsf{T}^{-1}=\begin{pmatrix} \Omega & \Omega'' \\ \Omega' & \Omega''' \end{pmatrix},$$

for each matrix Ω given in Tables VIII–XI. Similar expressions are available for B, C, and D. The first two terms of the right-hand side of Eq. (11) originate from the discontinuities of the transition operators for $(3+1)$ and $(2+2)$ subsystems. The last four terms originate from the discontinuity of the transition operator for the $(2+1+1)$ subsystem.

Let us again return to the first notation, in which the superscripts a, b, c, and d express different partitions of the two types: $(3+1)$ and $(2+2)$. Then Eq. (11) can be written in the more explicit form

$$\begin{aligned}&H^{ba}_{\beta\alpha}(E+i0)-H^{ba}_{\beta\alpha}(E-i0)\\ &\quad=\sum_{c}\sum_{\gamma\subset c}\sum_{\delta\subset c}H^{bc}_{\beta\gamma}(E+i0)\\ &\qquad\times\{U^{c}_{\gamma\delta}(E+i0)-U^{c}_{\gamma\delta}(E-i0)\}H^{ca}_{\delta\alpha}(E-i0)\\ &\qquad+\frac{1}{2}\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}H^{bc}_{\beta\kappa}(E+i0)\\ &\qquad\times\{\widetilde{T}_{\kappa}^{-1}(E+i0)-\widetilde{T}_{\kappa}^{-1}(E-i0)\}H^{da}_{\kappa\alpha}(E-i0),\end{aligned} \tag{12}$$

where η denotes the sign as follows: $\eta=1$ when $c=d$ and $\eta=-1$ when $c\neq d$. This equation expresses the discontinuity not only of A but also of B, C, and D. Hereafter we proceed with this notation.

Now, we may use the discontinuities of $U^{c}_{\gamma\delta}$ and $\widetilde{T}_{\kappa}^{-1}$ obtained in the Appendix to get

$$\begin{aligned}H^{ba}_{\beta\alpha}(E+i0)-H^{ba}_{\beta\alpha}(E-i0)=2\pi i\Big[&-\sum_{c}\sum_{\gamma\subset c}\sum_{\delta\subset c}\sum_{\mu\subset c}\sum_{\nu\subset c}\sum_{n}H^{bc}_{\beta\gamma}(E+i0)\bar{\delta}_{\gamma\mu}V_{\mu}\hat{\Delta}_0(E+E_{cn})\otimes J_{cn}V_{\nu}\bar{\delta}_{\nu\delta}H^{ca}_{\delta\alpha}(E-i0)\\ &-\sum_{c}\sum_{\gamma\subset c}\sum_{\delta\subset c}\sum_{\tau\subset c}\sum_{m}H^{bc}_{\beta\gamma}(E+i0)U^{c}_{\gamma\tau}(E+i0)\widetilde{\Delta}_{\tau m}(E)U^{c}_{\tau\delta}(E-i0)H^{ca}_{\delta\alpha}(E-i0)\\ &-\sum_{c}\sum_{\gamma\subset c}\sum_{\delta\subset c}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)\widetilde{\Delta}_0(E)U^{c(-)}_{0\delta}(E-i0)H^{ca}_{\delta\alpha}(E-i0)\\ &+\frac{1}{2}\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\sum_{m}H^{bc}_{\beta\kappa}(E+i0)\widetilde{T}_{\kappa}^{-1}(E+i0)\widetilde{\Delta}_{\kappa m}\widetilde{T}_{\kappa}^{-1}(E-i0)H^{da}_{\kappa\alpha}(E-i0)\end{aligned}$$

$$+\frac{1}{2}\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}H^{bc}_{\beta\kappa}(E+i0)\tilde{T}_{\kappa}^{-1}(E+i0)\tilde{\Delta}^{s}_{\kappa}\tilde{T}_{\kappa}^{-1}(E-i0)H^{da}_{\kappa\alpha}(E-i0)$$

$$-\frac{1}{2}\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}H^{bc}_{\beta\kappa}(E+i0)\tilde{T}_{\kappa}^{-1}(E+i0)\tilde{\Delta}_{0}\tilde{T}_{\kappa}^{-1}(E-i0)H^{da}_{\kappa\alpha}(E-i0)\Bigg]. \tag{13}$$

Because the third term of the right-hand side is expressed as the simple sum of the index c, we must further proceed to calculate so that the third term can be expressed in a double sum of indices c and d, which are the superscripts of the two groups of operators $H^{bc}_{\beta\gamma}U^{c(+)}_{\gamma 0}$ and $U^{d(-)}_{0\delta}H^{da}_{\delta\alpha}$.

In order to do so, note first that the third term in question is expressed by the difference

$$2\pi i\Bigg[-\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\delta\subset d}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)\tilde{\Delta}_{0}(E)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)$$

$$+\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\delta\subset d}\bar{\delta}^{cd}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)\tilde{\Delta}_{0}(E)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)\Bigg]. \tag{14}$$

Then it remains to calculate this second term, which we shall denote by J.

Next, we define an operator $F^{ca}_{\gamma\zeta}$ by

$$\sum_{\delta\subset b}\sum_{\gamma\subset b}\sum_{c\supset\gamma}\sum_{\zeta\supset a}\bar{\delta}^{bc}\bar{\delta}_{\delta\gamma}M^{b}_{\beta\delta}F^{ca}_{\gamma\zeta}W^{a}_{\zeta\alpha}=W^{ba}_{\beta\alpha}.$$

From the integral equation (4) for $W^{ba}_{\beta\alpha}$, we have for $F^{ba}_{\beta\alpha}$ the following equation:

$$F^{ba}_{\beta\alpha}=G_0\delta^{ba}\delta_{\beta\alpha}+G_0\sum_{\delta\subset b}\sum_{\gamma\subset b}\sum_{c\supset\gamma}\bar{\delta}^{bc}\bar{\delta}_{\delta\gamma}M^{b}_{\beta\delta}F^{ca}_{\gamma\alpha}.$$

From the definition of $F^{ba}_{\beta\alpha}$, we also have

$$-\tilde{T}_{\beta}U^{ba}_{\beta\alpha}G_0=\sum_{c\supset\beta}\bar{\delta}^{bc}F^{ca}_{\beta\alpha}.$$

Together with the relation

$$U^{d(-)}_{0\delta}(z)=\sum_{\mu\subset d}\sum_{\nu\subset d}\bar{\delta}_{\nu\delta}M^{d}_{\mu\nu}(z),$$

these results lead to

$$\sum_{\delta\subset d}G_0U^{d(-)}_{0\delta}H^{da}_{\delta\alpha}$$

$$=\sum_{\delta\subset d}\sum_{\mu\subset d}\sum_{\nu\subset d}\bar{\delta}_{\nu\delta}G_0M^{d}_{\mu\nu}\tilde{T}_{\delta}U^{da}_{\delta\alpha}\tilde{T}_{\alpha}=-\sum_{\delta\subset d}\sum_{\mu\subset d}\sum_{\nu\subset d}\sum_{e\supset\delta}\bar{\delta}_{\nu\delta}\bar{\delta}^{de}G_0M^{d}_{\mu\nu}F^{ea}_{\delta\alpha}T_{\alpha}G_0=-\sum_{\mu\subset d}\{F^{da}_{\mu\alpha}-G_0\delta^{da}\delta_{\mu\alpha}\}T_{\alpha}G_0. \tag{15}$$

We may now use Eq. (15) to eliminate $H^{da}_{\delta\alpha}$ in J in favor of $F^{da}_{\mu\alpha}$ as follows:

$$J=\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\delta\subset d}\bar{\delta}^{cd}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)G_0(E+i0)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)$$

$$-\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\delta\subset d}\bar{\delta}^{cd}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)G_0(E-i0)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)$$

$$=\sum_{c}\sum_{d}\sum_{\delta\subset d}\sum_{\mu\subset c}\bar{\delta}^{cd}\delta^{bc}\delta_{\beta\mu}\tilde{T}_{\beta}(E+i0)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)$$

$$-\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\mu\subset d}\bar{\delta}^{cd}\delta^{da}\delta_{\mu\alpha}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)\tilde{T}_{\alpha}(E-i0)$$

$$-\sum_{c}\sum_{d}\sum_{\delta\subset d}\sum_{\mu\subset c}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{bc}_{\beta\mu}(E+i0)U^{d(-)}_{0\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)$$

$$+\sum_{c}\sum_{d}\sum_{\gamma\subset c}\sum_{\mu\subset d}\bar{\delta}^{cd}H^{bc}_{\beta\gamma}(E+i0)U^{c(+)}_{\gamma 0}(E+i0)F^{da}_{\mu\alpha}(E-i0)T_{\alpha}(E-i0)G_0(E-i0)$$

$$=\sum_{c}\sum_{d}\sum_{\mu\subset c}\sum_{\lambda\subset d}\bar{\delta}^{cd}\delta^{bc}\delta_{\beta\mu}\delta^{da}\delta_{\lambda\alpha}\tilde{T}_{\beta}(E+i0)T_{\alpha}(E-i0)G_0(E-i0)$$

$$-\sum_{c}\sum_{d}\sum_{\mu\subset d}\sum_{\lambda\subset c}\bar{\delta}^{cd}\delta^{da}\delta_{\mu\alpha}\delta^{bc}\delta_{\beta\lambda}G_0(E+i0)T_{\beta}(E+i0)\tilde{T}_{\alpha}(E-i0)$$

$$
\begin{aligned}
&-\sum_{d}\sum_{\lambda\subset d}\bar{\delta}^{cd}\tilde{T}_{\beta}(E+i0)G_0^{-1}(E-i0)F^{da}_{\lambda\alpha}(E-i0)T_{\alpha}(E-i0)G_0(E-i0)\\
&+\sum_{c}\sum_{\lambda\subset c}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{bc}_{\beta\lambda}(E+i0)G_0^{-1}(E+i0)\tilde{T}_{\alpha}(E-i0)\\
&-\sum_{c}\sum_{\mu\subset c}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{bc}_{\beta\mu}(E+i0)T_{\alpha}(E-i0)G_0(E-i0)\\
&+\sum_{d}\sum_{\mu\subset d}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{da}_{\mu\alpha}(E-i0)T_{\alpha}(E-i0)G_0(E-i0)\\
&+\sum_{c}\sum_{d}\sum_{\mu\subset c}\sum_{\lambda\subset d}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{bc}_{\beta\mu}(E+i0)G_0^{-1}(E-i0)F^{da}_{\lambda\alpha}(E-i0)T_{\alpha}(E-i0)G_0(E-i0)\\
&-\sum_{c}\sum_{d}\sum_{\mu\subset d}\sum_{\lambda\subset c}\bar{\delta}^{cd}G_0(E+i0)T_{\beta}(E+i0)F^{bc}_{\beta\lambda}(E+i0)G_0^{-1}(E+i0)F^{da}_{\mu\alpha}(E-i0)T_{\alpha}(E-i0)G_0(E-i0)\\
&=2\pi i\sum_{c}\sum_{\lambda\subset a}\bar{\delta}^{ca}\delta^{bc}\delta_{\lambda\alpha}G_0(E+i0)T_{\beta}(E+i0)\widetilde{\Delta}_0(E)T_{\alpha}(E-i0)G_0(E-i0).
\end{aligned}
$$

Here, this last term must be again expressed in the term including $H^{da}_{\gamma\alpha}$. So, we sum Eq. (6) over c and μ,

$$
\sum_{c}\sum_{\mu\subset c}H^{ca}_{\mu\alpha}=\sum_{c}\sum_{\mu\subset c}\tilde{T}_{\mu}\delta_{\mu\alpha}\bar{\delta}^{ca}+\sum_{c}\sum_{\mu\subset c}\sum_{d\supset\mu}\bar{\delta}^{cd}\sum_{\delta\subset d}\tilde{T}_{\mu}U^{d}_{\mu\delta}H^{da}_{\delta\alpha}.
$$

This right-hand side can be further transformed into

$$
2\left[\sum_{\mu\subset a}\tilde{T}_{\mu}\delta_{\mu\alpha}\bar{\delta}^{ca}+\sum_{d}\sum_{\mu\subset d}\sum_{\delta\subset d}\tilde{T}_{\mu}U^{d}_{\mu\delta}H^{da}_{\delta\alpha}\right].
$$

Using the fact that $\widetilde{\Delta}_0(E)$ gives zero acting on $G_0^{-1}(E)$ multiplied by $H^{ca}_{\mu\alpha}$, one can finally express J as

$$
\begin{aligned}
J&=2\pi i\sum_{\lambda\subset a}\bar{\delta}^{ca}\delta_{\lambda\alpha}G_0(E+i0)T_{\beta}(E+i0)\widetilde{\Delta}_0(E)T_{\alpha}(E-i0)G_0(E-i0)-2\pi iG_0(E+i0)T_{\beta}(E+i0)\widetilde{\Delta}_0(E)G_0^{-1}(E-i0)\\
&\quad\times\left[\sum_{\mu\subset a}\tilde{T}_{\mu}(E-i0)\delta_{\mu\alpha}\bar{\delta}^{ca}+\sum_{d}\sum_{\mu\subset d}\sum_{\delta\subset d}\tilde{T}_{\mu}(E-i0)U^{d}_{\mu\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0)\right]\\
&=-2\pi iG_0(E+i0)T_{\beta}(E+i0)\widetilde{\Delta}_0(E)\sum_{d}\sum_{\mu\subset d}\sum_{\delta\subset d}T_{\mu}(E-i0)G_0(E-i0)U^{d}_{\mu\delta}(E-i0)H^{da}_{\delta\alpha}(E-i0).
\end{aligned}\tag{16}
$$

KZ defined the scattering amplitudes that describe physical four-body processes by taking appropriate matrix elements of the following operator:

$$
T^{ba}(z)=\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}V_{\xi}\bar{\delta}_{\xi\beta}H^{ba}_{\beta\alpha}(z)\bar{\delta}_{\alpha\zeta}V_{\zeta}.\tag{17}
$$

We can get the unitarity relations for the physical amplitudes once we obtain the discontinuity of this operator. The necessary equation is obtained by inserting Eq. (13) into Eq. (17) and using Eq. (14) with J expressed by Eq. (16),

$$
\begin{aligned}
T^{ba}(E+i0)-T^{ba}(E-i0)=2\pi i\Bigg[&-\sum_{c}\sum_{n}T^{bc}(E+i0)\widetilde{\Delta}_{cn}(E)T^{ca}(E-i0)\\
&-\sum_{c}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{\gamma\subset c}\sum_{\delta\subset c}\sum_{\tau\subset c}\sum_{m}V_{\xi}\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}(E+i0)\\
&\times U^{c}_{\gamma\tau}(E+i0)\widetilde{\Delta}_{\tau m}(E)U^{c}_{\tau\delta}(E-i0)H^{ca}_{\delta\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}\\
&-\sum_{c}\sum_{d}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{\gamma\subset c}\sum_{\delta\subset d}\sum_{\sigma\subset c}\sum_{\tau\subset c}\sum_{\mu\subset d}\sum_{\nu\subset d}V_{\xi}\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}(E+i0)\\
&\times\bar{\delta}_{\gamma\sigma}M^{c}_{\sigma\tau}(E+i0)\widetilde{\Delta}_0(E)M^{d}_{\mu\nu}(E-i0)\bar{\delta}_{\nu\delta}H^{da}_{\delta\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}\\
&+\sum_{d}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{\delta\subset d}\sum_{\mu\subset d}\sum_{\nu\subset d}V_{\xi}\bar{\delta}_{\xi\beta}G_0(E+i0)T_{\beta}(E+i0)\\
&\times\widetilde{\Delta}_0(E)M^{d}_{\mu\nu}(E-i0)\bar{\delta}_{\nu\delta}H^{da}_{\delta\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}\eta\sum_{\kappa}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{c\subset\kappa}\sum_{d\subset\kappa}\sum_{m}V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)\\
&\times U^{bc}_{\beta\kappa}(E+i0)\widetilde{\Delta}_{\kappa m}(E)U^{da}_{\kappa\alpha}(E-i0)\widetilde{T}_{\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}\\
&+\frac{1}{2}\eta\sum_{\kappa}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{c\subset\kappa}\sum_{d\subset\kappa}V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)\\
&\times U^{bc}_{\beta\kappa}(E+i0)\widetilde{\Delta}^{s}_{\kappa}(E)U^{da}_{\kappa\alpha}(E-i0)\widetilde{T}_{\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}\\
&-\frac{1}{2}\eta\sum_{\kappa}\sum_{\xi\subset b}\sum_{\zeta\subset a}\sum_{\beta\subset b}\sum_{\alpha\subset a}\sum_{c\subset\kappa}\sum_{d\subset\kappa}V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)\\
&\times U^{bc}_{\beta\kappa}(E+i0)\widetilde{\Delta}_{0}(E)U^{da}_{\kappa\alpha}(E-i0)\widetilde{T}_{\alpha}(E-i0)\bar{\delta}_{\alpha\zeta}V_{\zeta}\Bigg],
\end{aligned}
\tag{18}
$$

where

$$
\begin{aligned}
\widetilde{\Delta}_{cn}(E)&=\widehat{\Delta}_0(E+E_{cn})\otimes J_{cn}=\int|\mathbf{r}\Phi^{c}_{n}\rangle d^3r\,\delta(\tilde{r}^2-E-E_{cn})\langle\mathbf{r}\Phi^{c}_{n}|,\\
\widetilde{\Delta}_{\tau m}(E)&=\Delta_0(E+E_{\tau m})\otimes\hat{J}_{\tau m}=\int|\mathbf{rp}\varphi^{\tau}_{m}\rangle d^3r\,d^3p\,\delta(\tilde{r}^2+\tilde{p}^2-E-E_{\tau m})\langle\mathbf{rp}\varphi^{\tau}_{m}|,\\
\widetilde{\Delta}_0(E)&=\iiint|\mathbf{rpq}\rangle d^3r\,d^3p\,d^3q\,\delta(\tilde{r}^2+\tilde{p}^2+\tilde{q}^2-E)\langle\mathbf{rpq}|,\\
\widetilde{\Delta}^{s}_{\kappa}(E)&=[I-G_{\kappa}(E+i0)V_{\kappa}]\widetilde{\Delta}_0(E)[I-V_{\kappa}G_{\kappa}(E-i0)]\\
&=\iiint|\mathbf{rp}\psi^{-}_{\kappa\mathbf{q}}\rangle d^3p\,d^3p\,d^3q\,\delta(\tilde{r}^2+\tilde{p}^2+\tilde{q}^2-E)\langle\mathbf{rp}\psi^{-}_{\kappa\mathbf{q}}|.
\end{aligned}
$$

In the third and fourth terms of the right-hand side of Eq. (18), we use the following relations, respectively:

$$
U^{c(+)}_{\gamma0}(z)=\sum_{\sigma\subset c}\sum_{\tau\subset c}\bar{\delta}_{\gamma\sigma}M^{c}_{\sigma\tau}(z),
$$

$$
T_{\mu}(z)G_0(z)U^{d}_{\mu\delta}(z)=-\sum_{\nu\subset d}\bar{\delta}_{\nu\delta}M^{d}_{\mu\nu}(z).
$$

III. UNITARITY RELATION FOR THE AMPLITUDES WHICH DESCRIBE PHYSICAL FOUR-BODY PROCESSES

From the operator relation in the previous section, we shall derive in this section the unitarity relations for those on-shell amplitudes $\mathscr{H}^{ba}$, $\mathscr{F}^{ba}$, and $\mathscr{E}^{ba}$ that describe physical four-particle processes, elastic/rearrangement, partial breakup, and full breakup.

Following the treatment of KZ, in the case of a (3 + 1) subsystem we introduce the complete set of four-body channel eigenstates, $\{|\mathbf{r}\Phi^{(a)}_{n}\rangle,|\mathbf{r}\Psi^{(a)\pm}_{(\delta m)\mathbf{p}}\rangle,|\mathbf{r}\Psi^{(a)\pm}_{\mathbf{pq}}\rangle\}$, for all $\delta\subset a$, where $|\Phi^{(a)}_{n}\rangle$ is a three-body bound state (we assume several three-body bound states per channel) of energy $-E_{an}$, $|\Psi^{(a)+}_{(\delta m)\mathbf{p}}\rangle$ is the (outgoing wave) scattering state corresponding to an initial state of a bound pair (δm) and a third free particle with relative momentum **p**, and $|\Psi^{(a)+}_{\mathbf{pq}}\rangle$ is the (outgoing wave) scattering state corresponding to an initial state of three free particles of relative momenta **p** and **q**, while $\mathbf{r}_a$ is the momentum of the fourth particle relative to the center of mass of the other three.

For the case of a (2 + 2) subsystem, the complete set of channel eigenstates is given by $\{|\mathbf{s}\Phi^{(a)}_{n}\rangle$, $|\mathbf{s}\Psi^{(a)\pm}_{(\delta m)\mathbf{q}}\rangle$, $|\mathbf{s}\Psi^{(a)\pm}_{\mathbf{pq}}\rangle\}$. In this set, if $a=(12)(34)$ and $\delta=12$, $\gamma=34$, $|\mathbf{s}\Phi^{(a)}_{n}\rangle=|\mathbf{s}\varphi^{\delta}_{m}\varphi^{\gamma}_{m'}\rangle$ represents a state of two bound pairs moving with relative momentum **s**, $|\mathbf{s}\Psi^{(a)\pm}_{(\delta m)\mathbf{q}}\rangle=|\mathbf{s}\varphi^{\delta}_{m}\psi^{\pm}_{\mathbf{q}_{\gamma}}\rangle$ represents a state where the δ pair is bound, while the γ pair is in a scattering state of initial momentum $\mathbf{q}_{\gamma}$, and so on. In what follows, we will in general not treat the two kinds of indices a separately.

With these complete sets, KZ defined a fully-off-shell extension of the above three scattering amplitudes as

$$
\begin{aligned}
\mathscr{H}^{ba}_{n'n}(\mathbf{r};\mathbf{r}^{(0)};z)&=\langle\mathbf{r}\Phi^{(b)}_{n'}|T^{ba}(z)|\mathbf{r}^{(0)}\Phi^{(a)}_{n}\rangle,\\
\mathscr{F}^{b}_{(\epsilon m)}{}^{a}_{n}(\mathbf{rp};\mathbf{r}^{(0)};z)&=\langle\mathbf{r}\Psi^{(b)-}_{(\epsilon m)\mathbf{p}}|T^{ba}(z)|\mathbf{r}^{(0)}\Phi^{(a)}_{n}\rangle,\\
\mathscr{E}^{ba}_{n}(\mathbf{rpq};\mathbf{r}^{(0)};z)&=\langle\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}|T^{ba}(z)|\mathbf{r}^{(0)}\Phi^{(a)}_{n}\rangle.
\end{aligned}
$$

At the same time the amplitudes for the physical processes reverse to ones of $\mathscr{F}$ and $\mathscr{E}$ which are also defined as

$$
\begin{aligned}
\widetilde{\mathscr{F}}^{ba}_{n(\tau m)}(\mathbf{r};\mathbf{r}^{(0)}\mathbf{p}^{(0)};z)&=\langle\mathbf{r}\Phi^{(b)}_{n}|T^{ba}(z)|\mathbf{r}^{(0)}\Psi^{(a)+}_{(\tau m)\mathbf{p}^{(0)}}\rangle,\\
\widetilde{\mathscr{E}}^{ba}_{n}(\mathbf{r};\mathbf{r}^{(0)}\mathbf{p}^{(0)}\mathbf{q}^{(0)};z)&=\langle\mathbf{r}\Phi^{(b)}_{n}|T^{ba}(z)|\mathbf{r}^{(0)}\Psi^{(a)+}_{\mathbf{p}^{(0)}\mathbf{q}^{(0)}}\rangle.
\end{aligned}
$$

Next, let us derive the on-shell values of these amplitudes. Let all operators be taken on shell. There is a relationship between the three-body initial-state wave function and its Faddeev components,

$$
|\mathbf{r}\Phi^{(a)}_{\alpha;n}\rangle=G_0T_{\alpha}\sum_{\lambda\subset a}\bar{\delta}_{\alpha\lambda}G_0V_{\lambda}|\mathbf{r}\Phi^{(a)}_{n}\rangle.
\tag{19}
$$

If we now take the matrix element of $\mathscr{H}$ fully-on-shell, we can use Eq. (19) to obtain

$$\mathscr{H}^{ba}_{n'n}(\mathbf{r};\mathbf{r}^{(0)};E+i0)$$

$$=\sum_{\beta\subset b}\sum_{\alpha\subset a}\langle\mathbf{r}\Phi^{(b)}_{\beta;n'}|U^{ba}_{\beta\alpha}(E+i0)|\mathbf{r}^{(0)}\Phi^{(a)}_{\alpha;n}\rangle.$$

In order to get the on-shell value of $\mathscr{F}$, we need the expression for $|\mathbf{r}\Psi^{(b)-}_{(\epsilon m)\mathbf{p}}\rangle$ in terms of the initial state $|\mathbf{r}\mathbf{p}\varphi^{\epsilon}_{m}\rangle$, which is

$$\sum_{\lambda\subset b}\bar{\delta}_{\beta\lambda}V_{\lambda}|\mathbf{r}\Psi^{(b)-}_{(\epsilon m)\mathbf{p}}\rangle=U^{b}_{\beta\epsilon}(E-i0)|\mathbf{r}\mathbf{p}\varphi^{\epsilon}_{m}\rangle.$$

We then have for on-shell amplitude $\mathscr{F}$

$$\mathscr{F}^{b}{}_{(\epsilon m)}{}^{a}_{n}(\mathbf{r}\mathbf{p};\mathbf{r}^{(0)};E+i0)$$

$$=\sum_{\beta\subset b}\sum_{\alpha\subset a}\langle\mathbf{r}\mathbf{p}\varphi^{\epsilon}_{m}|U^{b}_{\epsilon\beta}(E+i0)$$
$$\times\widetilde{T}_{\beta}(E+i0)U^{ba}_{\beta\alpha}(E+i0)|\mathbf{r}\Phi^{(a)}_{\alpha;n}\rangle,$$

where $\tilde{r}^2+\tilde{p}^2-E_{\epsilon m}=\tilde{r}^{(0)2}-E_{bn}=E.$

Also in the case of the full breakup process, by using the relation

$$V_{\xi}|\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}\rangle=\sum_{\sigma\subset b}M^{b}_{\xi\sigma}(E-i0)|\mathbf{r}\mathbf{p}\mathbf{q}\rangle$$

the on-shell amplitude $\mathscr{C}$ can be written

$$\mathscr{C}^{ba}_{n}(\mathbf{r}\mathbf{p}\mathbf{q};\mathbf{r}^{(0)};E+i0)$$

$$=\sum_{\sigma\subset b}\sum_{\xi\subset b}\sum_{\beta\subset b}\sum_{\alpha\subset a}\langle\mathbf{r}\mathbf{p}\mathbf{q}|M^{b}_{\sigma\xi}(E+i0)\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)$$
$$\times U^{ba}_{\beta\alpha}(E+i0)|\mathbf{r}^{(0)}\Phi^{(a)}_{\alpha;n}\rangle.$$

Further, we recall from three-body theory that

$$T_{\beta}(E-i0)G_0(E-i0)\sum_{\lambda\subset b}\bar{\delta}_{\beta\lambda}V_{\lambda}|\mathbf{r}\Psi^{(b)-}_{(\epsilon m)\mathbf{p}}\rangle$$
$$=T_{\beta}(E-i0)G_0(E-i0)U^{b}_{\beta\epsilon}(E-i0)|\mathbf{r}\mathbf{p}\varphi^{\epsilon}_{m}\rangle$$
$$=-K^{b}_{\beta\epsilon}(E-i0)|\mathbf{r}\mathbf{p}\varphi^{\epsilon}_{m}\rangle,$$

$$\sum_{\xi\subset b}T_{\beta}(E-i0)G_0(E-i0)\bar{\delta}_{\beta\xi}V_{\xi}|\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}\rangle$$

$$=\sum_{\xi\subset b}T_{\beta}(E-i0)G_0(E-i0)\bar{\delta}_{\beta\xi}$$

$$\times\sum_{\lambda\subset b}M^{b}_{\xi\lambda}(E-i0)|\mathbf{r}\mathbf{p}\mathbf{q}\rangle$$

$$=-\sum_{\lambda\subset b}W^{b}_{\beta\lambda}(E-i0)|\mathbf{r}\mathbf{p}\mathbf{q}\rangle.$$

A. Elastic and rearrangement

Before starting to calculate the discontinuity of $\mathscr{H}^{ba}_{n'n}$, we must define some additional amplitudes:

$$\mathscr{C}^{d}{}_{(\kappa m)}{}^{a}_{n}(\mathbf{r}\mathbf{p};\mathbf{r}^{(0)};E+i0)=\sum_{\alpha\subset a}\langle\mathbf{r}\mathbf{p}\varphi^{\kappa}_{m}|U^{da}_{\kappa\alpha}(E+i0)|\mathbf{r}^{(0)}\Phi^{(a)}_{\alpha;n}\rangle,$$

$$\mathscr{D}^{d}{}_{(\kappa)}{}^{a}_{n}(\mathbf{r}\mathbf{p}\mathbf{q};\mathbf{r}^{(0)};E+i0)=\sum_{\alpha\subset a}\langle\mathbf{r}\mathbf{p}\psi^{-}_{\kappa\mathbf{q}}|U^{da}_{\kappa\alpha}(E+i0)|\mathbf{r}^{(0)}\Phi^{(a)}_{\alpha;n}\rangle,$$

$$\mathscr{M}^{da}_{\kappa n}(\mathbf{r}\mathbf{p}\mathbf{q};\mathbf{r}^{(0)};E+i0)=\sum_{\alpha\subset a}\langle\mathbf{r}\mathbf{p}\mathbf{q}|U^{da}_{\kappa\alpha}(E+i0)|\mathbf{r}^{(0)}\Phi^{(a)}_{\alpha;n}\rangle,$$

$$\widetilde{\mathscr{C}}^{b}{}_{n}{}^{c}{}_{(\kappa m)}(\mathbf{r};\mathbf{r}^{(0)}\mathbf{p}^{(0)};E+i0)=\sum_{\beta\subset b}\langle\mathbf{r}\Phi^{(b)}_{\beta;n}|U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}^{(0)}\mathbf{p}^{(0)}\varphi^{\kappa}_{m}\rangle,$$

$$\widetilde{\mathscr{D}}^{b}{}_{n}{}^{c}{}_{(\kappa)}(\mathbf{r};\mathbf{r}^{(0)}\mathbf{p}^{(0)}\mathbf{q}^{(0)};E+i0)=\sum_{\beta\subset b}\langle\mathbf{r}\Phi^{(b)}_{\beta;n}|U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}^{(0)}\mathbf{p}^{(0)}\psi^{+}_{\kappa\mathbf{q}^{(0)}}\rangle,$$

$$\widetilde{\mathscr{M}}^{bc}_{n\kappa}(\mathbf{r};\mathbf{r}^{(0)}\mathbf{p}^{(0)}\mathbf{q}^{(0)};E+i0)=\sum_{\beta\subset b}\langle\mathbf{r}\Phi^{(b)}_{\beta;n}|U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}^{(0)}\mathbf{p}^{(0)}\mathbf{q}^{(0)}\rangle.$$

The unitarity relation for the elastic and rearrangement transition amplitude is expressed as

$$\mathscr{H}^{ba}_{n'n}(\mathbf{r};\mathbf{r}^{(0)};E+i0)-\mathscr{H}^{ba}_{n'n}(\mathbf{r};\mathbf{r}^{(0)};E-i0)$$

$$=-2\pi i\sum_{c}\sum_{n''}\int\mathscr{H}^{bc}_{n'n''}(\mathbf{r};\mathbf{r}';E+i0)\delta(\tilde{r}'^2-E_{cn''}-E)\mathscr{H}^{ca}_{n''n}(\mathbf{r}';\mathbf{r}^{(0)};E-i0)d^3r'$$
$$-2\pi i\sum_{c}\sum_{\tau\subset c}\sum_{m}\int\mathscr{F}^{b}{}_{n'}{}^{c}{}_{(\tau m)}(\mathbf{r};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\tau m}-E)\mathscr{F}^{c}{}_{(\tau m)}{}^{a}_{n}(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'$$
$$-2\pi i\sum_{c}\sum_{d}\int\widetilde{\mathscr{C}}^{bc}_{n'}(\mathbf{r};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{C}^{da}_{n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'$$
$$+2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\sum_{m}\int\widetilde{\mathscr{C}}^{b}{}_{n'}{}^{c}{}_{(\kappa m)}(\mathbf{r};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\kappa m}-E)\mathscr{C}^{d}{}_{(\kappa m)}{}^{a}_{n}(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'$$
$$+2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\widetilde{\mathscr{D}}^{b}{}_{n'}{}^{c}{}_{(\kappa)}(\mathbf{r};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{D}^{d}{}_{(\kappa)}{}^{a}_{n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'$$
$$-2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\widetilde{\mathscr{M}}^{bc}_{n'\kappa}(\mathbf{r};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{M}^{da}_{\kappa n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'.$$

B. Partial breakup

As in Sec. III A, we define the following additional amplitudes:

$$\mathscr{k}^{b}_{(\mu m)}(\mathbf{p};\mathbf{p}'\mathbf{q}';E-\tilde{r}^2+i0)=\sum_{\beta\subset b}\langle\mathbf{p}\varphi^{\mu}_{m}|k^{b}_{\mu\beta}(E-\tilde{r}^2+i0)|\mathbf{p}'\mathbf{q}'\rangle,$$

$$\mathscr{P}^{b}_{(\mu m)}{}^{c}_{(\tau m')}(\mathbf{rp};\mathbf{r}'\mathbf{p}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\sum_{\gamma\subset c}\langle\mathbf{r}\Psi^{(b)-}_{(\mu m)\mathbf{p}}|V_{\xi}\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}(E+i0)U^{c}_{\gamma\tau}(E+i0)|\mathbf{r}'\mathbf{p}'\varphi^{\tau}_{m'}\rangle,$$

$$\mathscr{Q}^{b}_{(\mu m)}{}^{c}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\sum_{\gamma\subset c}\sum_{\sigma\subset c}\sum_{\tau\subset c}\langle\mathbf{r}\Psi^{(b)-}_{(\mu m)\mathbf{p}}|V_{\xi}\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}(E+i0)\bar{\delta}_{\gamma\sigma}M^{c}_{\sigma\tau}(E+i0)|\mathbf{r}'\mathbf{p}'\mathbf{q}'\rangle,$$

$$\mathscr{R}^{b}_{(\mu m)}{}^{c}_{(\kappa m')}(\mathbf{rp};\mathbf{r}'\mathbf{p}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{(\mu m)\mathbf{p}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\varphi^{\kappa}_{m'}\rangle,$$

$$\mathscr{S}^{b}_{(\mu m)}{}^{c}_{(\kappa)}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{(\mu m)\mathbf{p}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\psi^{-}_{\kappa\mathbf{q}'}\rangle,$$

$$\mathscr{X}^{b}_{(\mu m)}{}^{c}_{\kappa}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{(\mu m)\mathbf{p}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\mathbf{q}'\rangle.$$

We then obtain the discontinuity of $\mathscr{F}$,

$$\begin{aligned}
&\mathscr{F}^{b}_{(\mu m)}{}^{a}_{n}(\mathbf{rp};\mathbf{r}^{(0)};E+i0)-\mathscr{F}^{b}_{(\mu m)}{}^{a}_{n}(\mathbf{rp};\mathbf{r}^{(0)};E-i0)\\
&\quad=-2\pi i\sum_{c}\sum_{n'}\int\mathscr{F}^{b}_{(\mu m)}{}^{c}_{n'}(\mathbf{rp};\mathbf{r}';E+i0)\delta(\tilde{r}'^2-E_{cn'}-E)\mathscr{H}^{ca}_{n'n}(\mathbf{r}';\mathbf{r}^{(0)};E-i0)d^3r'\\
&\qquad-2\pi i\sum_{c}\sum_{\tau\subset c}\sum_{m'}\int\mathscr{P}^{b}_{(\mu m)}{}^{c}_{(\tau m')}(\mathbf{rp};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\tau m'}-E)\mathscr{F}^{c}_{(\tau m')}{}^{a}_{n}(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\\
&\qquad-2\pi i\sum_{c}\sum_{d}\int\mathscr{Q}^{b}_{(\mu m)}{}^{c}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{E}^{da}_{n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\qquad-2\pi i\sum_{d}\int\delta(\mathbf{r}-\mathbf{r}')\tilde{\mathscr{k}}^{b}_{(\mu m)}(\mathbf{p};\mathbf{p}'\mathbf{q}';E-\tilde{r}^2+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{E}^{da}_{n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\qquad+2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\sum_{m'}\int\mathscr{R}^{b}_{(\mu m)}{}^{c}_{(\kappa m')}(\mathbf{rp};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\kappa m'}-E)\mathscr{C}^{d}_{(\kappa m')}{}^{a}_{n}(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\\
&\qquad+2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\mathscr{S}^{b}_{(\mu m)}{}^{c}_{(\kappa)}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{D}^{d}_{(\kappa)}{}^{a}_{n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\qquad-2\pi i\eta\sum_{\kappa}\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\mathscr{X}^{b}_{(\mu m)}{}^{c}_{\kappa}(\mathbf{rp};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{M}^{da}_{\kappa n}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'.
\end{aligned}$$

C. Full breakup amplitude

To proceed as before, we define additional amplitudes by

$$\mathscr{w}^{b}(\mathbf{pq};\mathbf{p}'\mathbf{q}';E-\tilde{r}^2+i0)=\sum_{\lambda\subset b}\sum_{\beta\subset b}\langle\mathbf{pq}|w^{b}_{\lambda\beta}(E-\tilde{r}^2+i0)|\mathbf{p}'\mathbf{q}'\rangle,$$

$$\tilde{\mathscr{Q}}^{bc}_{(\tau m)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\sum_{\sigma\subset b}\sum_{\gamma\subset c}\sum_{\lambda\subset c}\langle\mathbf{rpq}|M^{b}_{\sigma\xi}(z)\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}\bar{\delta}_{\gamma\lambda}V_{\lambda}|\mathbf{r}^{(0)}\Psi^{(c)+}_{(\tau m)\mathbf{p}^{(0)}}\rangle,$$

$$\mathscr{Y}^{bc}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\sum_{\gamma\subset c}\sum_{\sigma\subset c}\sum_{\tau\subset c}\langle\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}|V_{\xi}\bar{\delta}_{\xi\beta}H^{bc}_{\beta\gamma}(E+i0)\bar{\delta}_{\gamma\sigma}M^{c}_{\sigma\tau}(E+i0)|\mathbf{r}'\mathbf{p}'\mathbf{q}'\rangle,$$

$$\mathscr{U}^{b}{}^{c}_{(\kappa m)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\varphi^{\kappa}_{m'}\rangle,$$

$$\mathscr{V}^{b}{}^{c}_{(\kappa)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\psi^{+}_{\kappa\mathbf{q}'}\rangle,$$

$$\mathscr{W}^{bc}_{\kappa}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)=\sum_{\beta\subset b}\sum_{\xi\subset b}\langle\mathbf{r}\Psi^{(b)-}_{\mathbf{pq}}|V_{\xi}\bar{\delta}_{\xi\beta}\widetilde{T}_{\beta}(E+i0)U^{bc}_{\beta\kappa}(E+i0)|\mathbf{r}'\mathbf{p}'\mathbf{q}'\rangle.$$

Finally the discontinuity of $\mathscr{E}$ is given as

$$\mathscr{C}_n^{ba}(\mathbf{rpq};\mathbf{r}^{(0)};E+i0)-\mathscr{C}_n^{ba}(\mathbf{rpq}\,;\mathbf{r}^{(0)};E-i0)$$

$$\begin{aligned}
&=-2\pi i\sum_c\sum_{n'}\int\mathscr{C}_{n'}^{bc}(\mathbf{rpq};\mathbf{r}';E+i0)\delta(\tilde{r}'^2-E_{cn'}-E)\mathscr{H}_{n'n}^{ca}(\mathbf{r}';\mathbf{r}^{(0)};E-i0)d^3r'\\
&\quad-2\pi i\sum_c\sum_{\tau\subset c}\sum_m\int\tilde{\mathscr{D}}^{bc}{}_{(\tau m)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\tau m}-E)\mathscr{F}^{c}{}_{(\tau m)}{}_n^a(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\\
&\quad-2\pi i\sum_c\sum_d\int\mathscr{Y}^{bc}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{C}_n^{da}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\quad-2\pi i\sum_d\int\delta(\mathbf{r}-\mathbf{r}')\tilde{\omega}^b(\mathbf{pq};\mathbf{p}'\mathbf{q}';E-\tilde{r}^2+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{C}_n^{da}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\quad+2\pi i\eta\sum_\kappa\sum_{c\supset\kappa}\sum_{d\supset\kappa}\sum_m\int\mathscr{U}^{bc}{}_{(\kappa m)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2-E_{\kappa m}-E)\mathscr{C}^{d}{}_{(\kappa m)}{}_n^a(\mathbf{r}'\mathbf{p}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\\
&\quad+2\pi i\eta\sum_\kappa\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\mathscr{V}^{bc}{}_{(\kappa)}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{D}^{d}{}_{(\kappa)}{}_n^a(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'\\
&\quad-2\pi i\eta\sum_\kappa\sum_{c\supset\kappa}\sum_{d\supset\kappa}\int\mathscr{W}_\kappa^{bc}(\mathbf{rpq};\mathbf{r}'\mathbf{p}'\mathbf{q}';E+i0)\delta(\tilde{r}'^2+\tilde{p}'^2+\tilde{q}'^2-E)\mathscr{M}_{\kappa n}^{da}(\mathbf{r}'\mathbf{p}'\mathbf{q}';\mathbf{r}^{(0)};E-i0)d^3r'\,d^3p'\,d^3q'.
\end{aligned}$$

APPENDIX: UNITARITY RELATIONS FOR THE AMPLITUDES $\tilde{T}_\alpha^{-1}, U_{\beta\alpha}$

In this Appendix, I shall derive discontinuities of the operators $\tilde{T}_\alpha^{-1}(z)$ and $U_{\beta\alpha}(z)$, following Lovelace's treatment[2] with his notation.

Three-particle operators are denoted by small italic letters, two-particle operators by putting a hat on them, and four-particle operators by capital italic letters, to avoid confusion later. Let $\hat{h}_0$ be the two-particle free Hamiltonian and $\hat{v}_\alpha$ be the two-particle interaction. We now consider the total two-particle Hamiltonian $\hat{h}_\alpha = \hat{h}_0 + \hat{v}_\alpha$ and its resolvent operator $\hat{g}_\alpha(z) = (\hat{h}_\alpha - z)^{-1}$. Under the relevant condition of the potential, $\hat{h}_\alpha$ and $\hat{g}_\alpha(z)$ have spectral representations as follows:

$$\begin{aligned}
\hat{h}_\alpha &= -\sum_n E_{\alpha n}\hat{J}_{\alpha n} + \int_0^\infty dE\,E\hat{\Delta}_\alpha(E),\\
\hat{g}_\alpha(z) &= -\sum_n\frac{\hat{J}_{\alpha n}}{E_{\alpha n}+s} + \int_0^\infty dE\frac{\hat{\Delta}_\alpha(E)}{E-s},
\end{aligned}\tag{A1}$$

where $\hat{\Delta}_\alpha(E)$ with $E\geqslant 0$ is the projection operator for the continuous eigenstates of $\hat{h}_\alpha$ on the positive real axis, and $\hat{J}_{\alpha n}$ are the projection operators for the bound states of binding energies $E_{\alpha n}$. As an exception we also use capital italic letters for these two- or three-particle projection operators. The projection operator $\hat{\Delta}_\alpha(E)$ is given by the discontinuity of $\hat{g}_\alpha(z)$ across the right-hand cut

$$\hat{\Delta}_\alpha(E) = (1/2\pi i)\{\hat{g}_\alpha(E+i0) - \hat{g}_\alpha(E-i0)\}.$$

For the bound-state projection operators, we have

$$\langle\mathbf{q}|\hat{J}_{\alpha n}|\mathbf{q}'\rangle = \sum_{j=1}^N\varphi_{nj}^\alpha(\mathbf{q})\varphi_{nj}^{\alpha}{}^*(\mathbf{q}').$$

All these relations are rigorously proved by many mathematical authors.[11,12]

Next we turn our attention to the spectral representation of the three-particle resolvent operator $g_\alpha(z) = (h_0 + v_\alpha - z)^{-1}$. Let us introduce the convolution formula[6] for the resolvent $R(z) = (H-z)^{-1}$ of an operator $H = h_1\otimes I + I\otimes h_2$ with variables separable. Here we will ignore temporarily the notational distinctions of three- or four-particle operators. If γ is a smooth contour separating $\sigma(h_1)$ (the spectral set of h_1) from $z-\sigma(h_2)$, and, say, coincides with the imaginary axis at large distances, then

$$R(z) = \frac{1}{2\pi i}\int_\gamma r_1(z-\zeta)\otimes r_2(\zeta)d\zeta.\tag{A2}$$

Applying this convolution formula to $h_\alpha = h_0 + v_\alpha$ and using (A1) gives the spectral representation of $g_\alpha(z)$,

$$g_\alpha(z) = -\sum_n\int_{-E_{\alpha n}}^\infty\frac{\Delta_{\alpha n}(\lambda)}{z-\lambda}d\lambda + \int_0^\infty\frac{\Delta_\alpha^s(\lambda)}{\lambda - z}d\lambda,$$

where

$$\Delta_{\alpha n}(\lambda) = \hat{\Delta}_0(\lambda + E_{\alpha n})\otimes\hat{J}_{\alpha n},$$

$$\hat{\Delta}_0(E) = (1/2\pi i)\{\hat{g}_0(E+i0) - \hat{g}_0(E-i0)\},$$

with $\hat{g}_0(z) = (\hat{h}_0 - z)^{-1}$, and

$$\Delta_\alpha^s(E) = [I - g_\alpha(E+i0)v_\alpha]\Delta_0(E)[I - v_\alpha g_\alpha(E-i0)].$$

Equation (A2) also gives the expression of the four-particle operator $G_\alpha = (H_0 + V_\alpha - z)^{-1}$,

$$G_\alpha(z) = \frac{1}{2\pi i}\int_\gamma\hat{g}_0(z-\zeta)\otimes g_\alpha(\zeta)d\zeta.$$

From this relation we wish to calculate the discontinuity of $G_\alpha(z)$ as follows:

$$\begin{aligned}
G_\alpha(E+i0) - G_\alpha(E-i0) &= \lim_{\epsilon\to 0}\frac{1}{2\pi i}\sum_n\int_\gamma\int_0^\infty\left\{\frac{1}{\mu+\zeta-E-i\epsilon} - \frac{1}{\mu+\zeta-E+i\epsilon}\right\}d\mu\int_{-E_{\alpha n}}^\infty\frac{\hat{\Delta}_0(\mu)\otimes\Delta_{\alpha n}(\lambda)}{\lambda-\zeta}d\lambda\,d\zeta\\
&\quad+\lim_{\epsilon\to 0}\frac{1}{2\pi i}\int_\gamma\int_0^\infty\left\{\frac{1}{\mu+\zeta-E-i\epsilon} - \frac{1}{\mu+\zeta-E+i\epsilon}\right\}d\mu\int_0^\infty\frac{\hat{\Delta}_0(\mu)\otimes\Delta_\alpha^s(\lambda)}{\lambda-\zeta}d\lambda\,d\zeta
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\pi i}\sum_n \int_\gamma \int_0^\infty 2\pi i \delta(\mu+\zeta-E)d\mu \int_{-E_{\alpha n}}^\infty \frac{\widehat{\Delta}_0(\mu)\otimes\Delta_{\alpha n}(\lambda)}{\lambda-\zeta}\,d\lambda\, d\zeta \\
&\quad + \frac{1}{2\pi i}\int_\gamma \int_0^\infty 2\pi i \delta(\mu+\zeta-E)d\mu \int_0^\infty \frac{\widehat{\Delta}_0(\mu)\otimes\Delta_\alpha^s(\lambda)}{\lambda-\zeta}\,d\lambda\, d\zeta \\
&= \sum_n \int_\gamma \int_{-E_{\alpha n}}^\infty \frac{\widehat{\Delta}_0(E-\zeta)\otimes\Delta_{\alpha n}(\lambda)}{\lambda-\zeta}\,d\lambda\, d\zeta + \int_\gamma\int_0^\infty \frac{\widehat{\Delta}_0(E-\zeta)\otimes\Delta_\alpha^s(\lambda)}{\lambda-\zeta}\,d\lambda\, d\zeta \\
&= \sum_n 2\pi i \int_{-E_{\alpha n}}^\infty \widehat{\Delta}_0(E-\lambda)\otimes\Delta_0(\lambda+E_{\alpha n})\otimes\hat{J}_{\alpha n}\, d\lambda + 2\pi i\int_0^\infty \widehat{\Delta}_0(E-\lambda)\otimes\Delta_\alpha^s(\lambda)d\lambda \\
&= 2\pi i \sum_n \widetilde{\Delta}_{\alpha n}(E) + 2\pi i \Delta_\alpha^s(E-\tilde{r}^2),
\end{aligned}
\tag{A3}
$$

where $\widetilde{\Delta}_{\alpha n}(E) = \Delta_0(E+E_{\alpha n})\otimes\hat{J}_{\alpha n}$.

Here we denote the four-particle projection operator by putting a tilde on the letter Δ.

Now we can get the discontinuity of the operator $\widetilde{T}_\alpha^{-1}(z)$. To do so, note first that

$$\widetilde{T}_\alpha(z) = -G_\alpha(z) + G_0(z).$$

Then we can easily derive the relation

$$
\begin{aligned}
&\widetilde{T}_\alpha^{-1}(E+i0) - \widetilde{T}_\alpha^{-1}(E-i0) \\
&\quad = -2\pi i \widetilde{T}_\alpha^{-1}(E+i0)\widetilde{\Delta}_0(E)\widetilde{T}_\alpha^{-1}(E-i0) \\
&\qquad + \widetilde{T}_\alpha^{-1}(E+i0)\{G_\alpha(E+i0) - G_\alpha(E-i0)\} \\
&\qquad \times \widetilde{T}_\alpha^{-1}(E-i0).
\end{aligned}
$$

Inserting Eq. (A3) to this relation, it follows that

$$
\begin{aligned}
&\widetilde{T}_\alpha^{-1}(E+i0) - \widetilde{T}_\alpha^{-1}(E-i0) \\
&\quad = 2\pi i \widetilde{T}_\alpha^{-1}(E+i0)\widetilde{\Delta}_{\alpha n}(E)\widetilde{T}_\alpha^{-1}(E-i0) \\
&\qquad + \widetilde{T}_\alpha^{-1}(E+i0)(I - G_\alpha(E+i0)V_\alpha)\widetilde{\Delta}_0(E) \\
&\qquad \times (I - V_\alpha G_\alpha(E-i0))\widetilde{T}_\alpha^{-1}(E-i0) \\
&\qquad - 2\pi i \widetilde{T}_\alpha^{-1}(E+i0)\widetilde{\Delta}_0(E)\widetilde{T}_\alpha^{-1}(E-i0).
\end{aligned}
$$

I conclude by obtaining the discontinuity of the operator $U_{\beta\alpha}^a(z)$. Consider the three-particle resolvent operator,

$$g^a(z) = \Big(h_0 + \sum_{\alpha\subset a} v_\alpha - z\Big)^{-1}.$$

By the spectral decomposition theorem established by Faddeev,

$$
\begin{aligned}
g^a(z) = &-\sum_n \frac{J_{an}}{z+E_{an}} + \sum_{\substack{\alpha\subset a\\ n}} \int_{-E_{\alpha n}}^\infty \frac{\Delta_{\alpha n,\bar{\alpha}}^s(\lambda)}{\lambda - z}\,d\lambda \\
&+ \int_0^\infty \frac{\Delta_a^s(\lambda)}{\lambda - z}\,d\lambda.
\end{aligned}
$$

In this representation the projection operators are defined by

$$
\begin{aligned}
J_{an} &= |\Phi_n^{(a)}\rangle\langle\Phi_n^{(a)}|, \\
\Delta_{\alpha m,\bar{\alpha}}(\lambda) &= \Big(I - \sum_{\gamma\neq\alpha} g^a(\lambda+i0)v_\gamma\Big) \\
&\quad \times \Delta_{\alpha n}(\lambda)\Big(I - \sum_{\delta\neq\alpha} v_\delta\, g^a(\lambda - i0)\Big) \\
&= \int |\Psi_{(\alpha n)\mathbf{p}}^{(a)-}\rangle d^3p\, \delta(\tilde{p}^2 - \lambda)\langle\Psi_{(\alpha n)\mathbf{p}}^{(a)-}|, \\
\Delta_a^s(\lambda) &= \Big(I - \sum_\gamma g^a(\lambda + i0)v_\gamma\Big) \\
&\quad \times \widetilde{\Delta}_0(\lambda)\Big(I - \sum_\delta v_\delta\, g^a(\lambda - i0)\Big) \\
&= \int |\Psi_{\mathbf{pq}}^{(a)-}\rangle d^3p\, d^3q\, \delta(\tilde{p}^2 + \tilde{q}^2 - \lambda)\langle\Psi_{\mathbf{pq}}^{(a)-}|.
\end{aligned}
$$

In order to proceed, we need the discontinuity of the four-particle operator $G^a(z)$,

$$G^a(z) = \frac{1}{2\pi i}\int_\gamma \hat{g}_0(z-\zeta)\otimes g^a(\zeta)d\zeta.$$

A similar procedure to that which leads to Eq. (A3) gives the required result

$$
\begin{aligned}
&G^a(E+i0) - G^a(E-i0) \\
&\quad = 2\pi i \sum_n \widehat{\Delta}_0(E+E_{an})\otimes J_{an} \\
&\qquad + 2\pi i \sum_{\substack{\alpha\subseteq a\\ m}} \Delta_{\alpha m,\bar{\alpha}}(E-\tilde{r}^2) + 2\pi i\Delta_a^s(E-\tilde{r}^2) \\
&\quad = 2\pi i \sum_n \widetilde{\Delta}_{an}(E) \\
&\qquad + 2\pi i \sum_{\substack{\alpha\subseteq a\\ m}} \Big(I - \sum_{\gamma\neq\alpha} G^a(E+i0)V_\gamma\Big) \\
&\qquad \times \widetilde{\Delta}_{\alpha m}(E)\Big(I - \sum_{\delta\neq\alpha} V_\delta G^a(E-i0)\Big) \\
&\qquad + 2\pi i\Big(I - \sum_\gamma G^\alpha(E+i0)V_\gamma\Big) \\
&\qquad \times \widetilde{\Delta}_0(E)\Big(I - \sum_\delta V_\delta G^\alpha(E-i0)\Big).
\end{aligned}
$$

Finally, from the definition of $U^{a}_{\gamma\delta}$ we can get

$$\begin{aligned} &U^{a}_{\gamma\delta}(E+i0) - U^{a}_{\gamma\delta}(E-i0) \\ &\quad = -2\pi i \sum_{n} \sum_{\mu\neq\gamma} \sum_{\nu\neq\delta} V_{\mu} \widetilde{\Delta}_{an}(E) V_{\nu} \\ &\qquad - 2\pi i \sum_{m} \sum_{\alpha\subset a} U^{a}_{\gamma\alpha}(E+i0) \widetilde{\Delta}_{\alpha m}(E) U^{a}_{\alpha\delta}(E-i0) \\ &\qquad - 2\pi i U^{a(+)}_{\gamma 0}(E+i0) \widetilde{\Delta}_{0}(E) U^{a(-)}_{0\delta}(E-i0), \end{aligned}$$

for the discontinuity of $U^{a}_{\gamma\delta}$ across the right-hand cut.

[1] L. D. Faddeev, *Mathematical Aspects of the Three-Body Problem in the Quantum Scattering Theory* (Davey, New York, 1965).

[2] C. Lovelace, in *Strong Interactions and High Energy Physics*, edited by R. G. Moorhouse (Oliver and Boyd, London, 1964); Phys. Rev. **135**, B1225 (1964).

[3] T. A. Osborn and D. Bolle, Phys. Rev. C **8**, 1198 (1973).

[4] E. Schmid and H. Ziegelman, *The Quantum Mechanical Three-Body Problem*, edited by H. Stumpf (Pergamon, New York, 1974).

[5] B. R. Karlson and E. M. Zeiger, Phys. Rev. D **9**, 1761 (1974); **11**, 939 (1975); **16**, 2553 (1977).

[6] J. S. Howland, J. Funct. Anal. **22**, 250 (1976).

[7] O. A. Yakubovskii, Yad. Fiz. **5**, 1312 (1967) [Sov. J. Nucl. Phys. **5**, 937 (1967)]; I. M. Norodetskii and O. A. Yakubovskii, Yad. Fiz. **14**, 315 (1971) [Sov. J. Nucl. Phys. **14**, 178 (1972)].

[8] E.O. Alt, P. Grassberger, and W. Sandhas, JINR Report E4-6688, 1972 (unpublished); in *Few Particle Problems in the Nuclear Interaction*, edited by I. Slaus, S. A. Moszkowski, R. P. Haddock, and W. T. H. van Oers (North-Holland, Amsterdam, 1972); P. Grassberger and W. Sandhas, Nucl. Phys. B **2**, 181 (1967).

[9] Y. Matsui, Phys. Rev. C **26**, 2620 (1982); **29**, 253 (1984); J. Math. Phys. **28**, 2677 (1987).

[10] I. M. Narodetskii, Yad. Fiz. **20**, 1063 (1974) [Sov. J. Nucl. Phys. **20**, 560 (1975)].

[11] T. Ikebe, Arch. Rational Mech. Anal. **5**, 1 (1960).

[12] T. Kato and S. T. Kuroda, Rocky Mountain J. Math. **1**, 127 (1971).

Invariant conformal vectors in space-times admitting a group of G_3 of motions acting on spacelike orbits S_2

C. Bona
Departament de Fisica, Universitat Illes Balears, 07071 Palma de Mallorca, Spain

(Received 20 February 1987; accepted for publication 22 June 1988)

The paper deals with four-dimensional space-times admitting locally a three-dimensional group of motions G_3 acting on two-dimensional spacelike orbits S_2. The local existence problem for conformal vectors invariant under G_3 is shown to be equivalent to the local existence problem for Killing vectors of a given two-dimensional pseudo-Riemannian metric g. This problem is explicitly solved in terms of the Gaussian curvature R of g and two of its scalar differential concomitants. The results are applied to the case of dust-filled space-times, where an exhaustive list of metrics has been obtained by using the algebraic computing language SMP. The metrics are either homogeneous, self-similar, or Friedmann models.

I. STATEMENT OF THE PROBLEM

Let us consider a four-dimensional pseudo-Riemannian manifold $(V_4,\hat{g})$. A vector field v is said to be conformal if

$$£_v(\hat{g}) = 2\phi\hat{g}\,, \tag{1}$$

where £ stands for the Lie derivative, and the conformal factor ϕ is a V_4 function. Conformal vector fields[1,2] are a well-known generalization of Killing vector fields, and the properties of space-times in which (1) admits nontrivial solutions have been recently studied.[3–8]

The purpose of this paper is to discuss the existence of solutions of (1) with the following two restrictions: (a) the manifold $(V_4,\hat{g})$ admits locally a three-dimensional group of motions G_3 acting on two-dimensional spacelike orbits S_2, and (b) the conformal vector v is invariant under the group G_3.

Condition (a) means that the two-surfaces S_2 are maximally symmetric, so that the metric $\hat{g}$ is conformal to the direct sum of a constant curvature metric h on S_2 and a metric g on the surfaces V_2 orthogonal to S_2; that is,

$$\hat{g} = Y^2(g\otimes h)\,. \tag{2}$$

The function Y is invariant by G_3, and it is defined up to a constant multiplicative factor. One can choose it so as to have

$$\mathrm{Ric}(h) = kh, \tag{3}$$

where Ric stands for the Ricci tensor, and $k = 1,0,\ -1$ is the (S_2,h) Gaussian curvature, corresponding to the cases of spherical, plane, or hyperbolic symmetry, respectively.

Condition (b) implies that the vector field v is tangent to V_2, and it leaves the metric h invariant.

The decomposition (2) allows a straightforward computation of the conformal tensor of $\hat{g}$, Conf($\hat{g}$). One gets in this way the following proposition.

Proposition 1: The necessary and sufficient condition for the conformal tensor Conf($\hat{g}$) of $(V_4,\hat{g})$ to be zero is that

$$R + k = 0\,, \tag{4}$$

where R is the Gaussian curvature of (V_2,g); that is,

$$\mathrm{Ric}(g) = Rg\,. \tag{5}$$

It follows also from (2) that every conformal vector v of $\hat{g}$ must be a conformal vector of $(g\otimes h)$ and vice versa. Allowing for the fact that h is left invariant by v, it follows that g must be also invariant by v as stated in the following proposition.

Proposition 2: In any space-time admitting a group G_3 of motions acting on spacelike orbits S_2, a vector field v invariant by G_3 is a conformal vector of $(V_4,\hat{g})$ if and only if it is a Killing vector of (V_2,g); that is,

$$£_v(g) = 0\,. \tag{6}$$

The conformal factor ϕ is given by

$$£_v(Y) = \phi Y\,. \tag{7}$$

Proposition 2 shows the relevance of the study of the Killing vectors of (V_2,g) to that of the invariant conformal vectors of $(V_4,\hat{g})$.

II. KILLING VECTORS IN (V_2,g)

Let us consider a generic two-dimensional pseudo-Riemannian manifold (V_2,g). The first set of integrability conditions of the Killing equation (6) is

$$£_v(R) = 0\,, \tag{8a}$$

and the remaining conditions are obtained by repeated application of the covariant derivative operator D (relative to g), namely,

$$£_v(dR) = 0\,, \tag{8b}$$

$$£_v(D\,dR) = 0\,, \tag{8c}$$

and so on. The dimension r of the linear space of solutions of (6) is known to be equal to $(3-q)$, where q is the rank of (8) considered as an algebraic linear system with the components of v and its first covariant derivatives (the antisymmetric part) as unknowns.[1]

It is well known that r may be either 3, 1, or 0, and that r is equal to 3 if and only if R is constant. It is also known[9] that Eq. (6) admits as a nontrivial solution a vector field v with zero norm (isotropic vector) if and only if (V_2,g) is flat $(R=0)$. It would be useful to set up an analogous condition to know whether r is equal to 1 so that the solutions of Eq. (6) are proportional to a single nonisotropic vector v.

0022-2488/88/112462-03$02.50

To provide one such criterion, let us construct the two following scalars:

$$s1 = \mathrm{tr}(dR \otimes dR), \quad s2 = \mathrm{tr}(D\,dR), \tag{9}$$

where tr is the trace operator relative to g. The condition can now be stated as follows.

Theorem: In a two-dimensional pseudo-Riemannian manifold (V_2,g) the solutions of the Killing equation (6) are proportional to a single nonisotropic Killing vector v if and only if $s1$ is not zero and both $s1$ and $s2$ are functions of the curvature R only, that is,

$$s1 \neq 0 \tag{10a}$$

$$ds1 \wedge dR = 0, \quad ds2 \wedge dR = 0, \tag{10b}$$

where $\wedge$ stands for the exterior product of forms, and the scalars $s1$ and $s2$ are defined in (9).

Proof of the necessary condition: Let us suppose that R is not constant (otherwise $r = 3$), and let v be a nonisotropic Killing vector, so that Eqs. (8) hold. It follows from (8a) that $s1$ cannot vanish. In addition, one gets easily from (8) that

$$\pounds_v(s1) = \pounds_v(s2) = \pounds_v(R) = 0, \tag{11}$$

so that Eqs. (10b) must also hold.

Proof of the sufficient condition: In (V_2,g), the conditions (10) imply the following tensor relationship:

$$D\,dR = A(R)\,dR \otimes dR + B(R)g, \tag{12}$$

where A and B are functions of the curvature R only. This implies that the rank q of the full algebraic system (8) is equal to that of the system of three equations (8a) and (8b). In addition, one gets from the first condition in (10b) that Eq. (8a) implies the relationship

$$\mathrm{tr}(dR \otimes \pounds_v(dR)) = \tfrac{1}{2}\pounds_v(s1) = 0, \tag{13}$$

so that the rank q of (8a) and (8b) is at most equal to 2, and there is at least one Killing vector v. Finally, condition (8a) ensures that R is not constant ($r \neq 3$) and v is nonisotropic.

III. THE SCALAR *R* AND THE FOUR-DIMENSIONAL RICCI TENSOR

The preceding sections show that the scalar R plays a crucial role in determining the conformally invariant properties of the metrics (2). It is the essential component of the conformal tensor $\mathrm{Conf}(\hat{g})$ (see Proposition 1), and it determines, together with its differential concomitants $s1$ and $s2$, the existence of invariant conformal motions.

The computation of R can be performed either from its definition (5) or from the four-dimensional Ricci tensor $\mathrm{Ric}(\hat{g})$. A partial result is stated in the following proposition.

Proposition 3: Let us assume that there exists a timelike unit vector u and two scalars μ and p such that $\mathrm{Ric}(\hat{g})$ admits the decomposition

$$\mathrm{Ric}(\hat{g}) = (\mu + p) u \otimes u + \tfrac{1}{2}(\mu - p)\hat{g} \tag{14}$$

(perfect fluid space-time). Then the following relationship holds:

$$R + k = 6M/Y - \mu Y^2, \tag{15}$$

where M is the Hernandez–Misner[10] scalar, defined as follows:

$$M = \tfrac{1}{2}Y[k - \mathrm{Tr}(dY \otimes dY)], \tag{16}$$

the operator Tr being the trace relative to the metric $\hat{g}$.

Proof: It follows from a straightforward computation of the Ricci tensor for a metric $\hat{g}$ of the form given in Eq. (2).

In the case in which $\mathrm{Ric}(\hat{g}) = 0$ (vacuum space-time), it is well known that the Hernandez–Misner scalar M must be constant. Allowing for Eq. (15), this means that either $M = 0$ (flat space-time) or the scalars R and Y are algebraically related one to another, so that it follows from Eqs. (7) and (8a) that any invariant conformal vector of $\hat{g}$ must be a Killing vector. This is in keeping with one recently published result of Garfinkle.[8]

The vacuum case, however, is trivial due to the well-known Birkhoff theorem.[11–13] All the possible metrics and their Killing structure are well known.[14] It is worthwhile to consider then a more general situation.

IV. APPLICATION TO DUST METRICS

Let us consider now the case in which $\mathrm{Ric}(\hat{g})$ admits the decomposition (14) with $p = 0$ but $\mu \neq 0$ (dust space-time). All dust metrics admitting a group G_3 of motions acting on spacelike orbits S_2 are known.[14] A systematic, computer-aided[15] application of the first condition of (10b) to all these metrics has shown that, in order to admit (at least) one invariant conformal vector v, the metric $\hat{g}$ must pertain to one of the three following families.[16]

(a) *Friedmann metrics,*[17] given by

$$\hat{g} = -dt \otimes dt + L^2(t)[(k - ar^2)^{-1}\,dr \otimes dr + r^2 h], \tag{17a}$$

where a is constant and the function $L(t)$ is defined by

$$t = \int^L \left[\frac{x}{b - ax}\right]^{1/2} dx, \tag{17b}$$

with b constant. In this case, one gets that

$$M = \tfrac{1}{2}br^3, \quad \mu = 3b/L^3, \quad R + k = 0, \tag{18}$$

so that there are three invariant conformal vectors, one of them being tangent to the t coordinate lines.

(b) *Self-similar metrics,*[18,19] given by

$$\hat{g} = -dt \otimes dt + (Y')^2(k - a)^{-1}\,dr \otimes dr + Y^2 h, \tag{19a}$$

where a is constant, the prime stands for r derivatives, and the function $Y(t,r)$ can be expressed as follows:

$$t = br + \int^Y \left[\frac{x}{2r - ax}\right]^{1/2} dx, \tag{19b}$$

with b constant. In this case, one gets that

$$M = r, \quad \mu = 2/(Y^2 Y'), \tag{20}$$

and $R + k$ can be expressed either as a function of Y/r or of t/r. In the generic case ($R + k$ not constant), the only invariant conformal vector is

$$v = t d_t + r d_r, \tag{21}$$

and it follows from Eqs. (7) and (19b) that it is a homothetic vector of $\hat{g}$ ($\phi = \mathrm{const}$). There are two additional invariant conformal vectors only when both a and b are zero in

(19), so that the metric is isometric to the $a = 0$ case of the Friedmann metrics (17).

(c) *Homogeneous metrics*,[20–22] given by

$$\hat{g} = -[t/(a-kt)]dt \otimes dt + [(a-kt)/t]\, Z^2(t) dr \otimes dr + t^2 h\,, \tag{22a}$$

where a is constant and Z is the following function of t:

$$Z = \int^t \left[\frac{x}{(a-kx)^3}\right]^{1/2} dx + b\,, \tag{22b}$$

with b constant. In this case, one gets that

$$M = \tfrac{1}{2}a, \quad \mu = (2/Z)t^{-2}[t/(a-kt)]^{1/2}\,, \tag{23}$$

and $R + k$ is a function of t. In the generic case ($R + k$ not constant), the only invariant conformal vector is

$$v = d_r\,, \tag{24}$$

and it follows from (22) that it is a Killing vector orthogonal to u. There are two additional invariant conformal vectors only when both k and b in (22) are zero, the metric being isometric to the case $k = 1$, $a = 0$ in the Friedmann metrics (17).

The results quoted here ensure that any dust[23] metric of the class considered admitting an invariant conformal vector v must pertain to one of the metric families (17), (19), or (22). The inverse result, that is, that every metric contained in (17), (19), or (22) admits such a vector v, is well known in the case of Friedmann metrics (17), and it can be verified by using the explicit expressions (21) and (24) of v in the other two cases. One arrives then at the following results.

Proposition 4: A dust metric admitting a three-dimensional group of motions G_3 acting on two-dimensional spacelike surfaces S_2 will admit at least one conformal vector v invariant by G_3 if and only if it is isometric to a Friedmann metric (17), to a self-similar metric (19), or to a homogeneous metric (22).

Proposition 5: A dust metric admitting a three-dimensional group of motions G_3 acting on two-dimensional spacelike surfaces S_2 will admit three invariant conformal vectors if and only if it is isometric to a Friedmann metric (17).

ACKNOWLEDGMENTS

The author is indebted to Dr. B. Coll for useful discussions and suggestions.

Financial support from the CAICYT of Spain under Project No. 1005/84 is also acknowledged.

[1] K. Yano, *The Theory of Lie Derivatives and Its Applications* (North-Holland, Amsterdam, 1955).

[2] D. R. Oliver and W. R. Davis, Gen. Relativ. Gravit. **8**,905 (1977).

[3] L. Herrera, J. Jiménez, L. Leal, J. Ponce de Leon, M. Esculpi, and V. Galina, J. Math. Phys. **25**, 3274 (1984).

[4] D. P. Mason and M. Tsamparlis, J. Math. Phys. **26**, 2881 (1985).

[5] R. Maartens and S. D. Maharaj, Class. Quantum Grav. **3**, 1005 (1986).

[6] M. Esculpi and L. Herrera, J. Math. Phys. **27**, 2087 (1986).

[7] R. Maartens, D. P. Mason, and M. Tsamparlis, J. Math. Phys. **27**, 2987 (1986).

[8] D. Garfinkle, J. Math, Phys. **28**, 28 (1987).

[9] L. P. Eisenhart, *Continuous Groups of Transformations* (Dover, New York, 1961).

[10] W. C. Hernández and C. W. Misner, Astrophys. J. **143**, 452 (1966).

[11] M. Cahen and R. Debever, C. R. Acad. Sci. **260**, 815 (1965).

[12] H. Goenner, Commun. Math. Phys. **16**, 34 (1970).

[13] A. Barnes, Commun. Math. Phys. **33**, 75 (1973).

[14] D. Kramer, H. Stephani, M. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations* (VEB, Berlin, 1980).

[15] The calculations have been performed using the symbolic manipulation language SMP running in the VAX 750 University computer.

[16] Local orthogonal coordinates are used so that the timelike unit vector u is tangent to the t coordinate lines.

[17] A. Friedmann, Z. Phys. **10**, 377 (1922); **21**, 326 (1924).

[18] M. E. Cahill and A. H. Taub, Commun. Math. Phys. **21**, 1 (1971).

[19] A. Taub, in *General Relativity* (papers in honor of J. L. Synge), edited by L. O'Raifeartaigh (Clarendon, Oxford, 1972), p. 133.

[20] A. Raychaudhuri, Proc. R. Soc. London **72**, 263 (1958).

[21] B. B. Robinson, Proc. Natl. Acad. Sci. USA **47**, 1852 (1961).

[22] R. Kantowski and R. K. Sachs, J. Math. Phys. **7**, 443 (1966).

[23] The vacuum case ($\mu = 0$) has been explicitly excluded.

Affine collineations in space-time

G. S. Hall and J. da Costa
Department of Mathematics, University of Aberdeen, Edward Wright Building, Dunbar Street, Aberdeen AB9 2TY, Scotland, United Kingdom

(Received 31 March 1988; accepted for publication 29 June 1988)

The existence of affine collineations in space-time is discussed and the types of space-time admitting proper affine collineations is displayed. The close connection between such space-times and their holonomy structure and local decomposability is established. Affine collineations with fixed points are also considered as is the problem of extending local affine collineations to the whole of space-time.

I. INTRODUCTION

Many of the techniques for finding exact solutions of Einstein's field equations in general relativity involve the assumption that certain symmetries exist in the space-time. A thorough discussion may be found in Ref. 1. Most of these assumptions require the existence of a certain number and type of Killing vector fields defined globally on space-time and lead to a local type of Lie group isometric action. Full details of local and global Lie group actions can be found in Ref. 2, while a brief summary has been given in Ref. 3. The purpose of this paper is to consider the case when a space-time admits local groups of affine collineations generated by global affine vector fields.

Section II gives a brief discussion of the interconnection between the metric and associated connection on space-time, the associated holonomy group, the existence of recurrent and covariantly constant tensors, and local and global decomposability. Some of the mathematical formalities of this section are merely noted in passing and are not required in any serious way in the remaining sections. Others, however, will be significant in what is to follow and all of them are, in their own way, helpful in understanding the problem. Local considerations are dealt with in Secs. III–V where the types of space-time admitting proper affine collineations will be displayed. The fixed point structure of affine collineations will be discussed in Sec. VI, while the global extension of local affine collineations will be covered in Sec. VII. A final summary and some examples will be given in Sec. VIII.

The notation is a standard one. The (connected) space-time manifold will be denoted by M and its Lorentz metric of signature $(-\ +\ +\ +)$ by g. The Riemann and Ricci tensors associated with g are denoted in local coordinates by R_{abcd} and $R_{ab} \equiv R^{c}{}_{acb}$, respectively, a covariant derivative with respect to g by a semicolon, and a partial derivative by a comma. The space-time manifold will always be assumed *simply connected* (and hence time orientable), although this is not necessary for some of the local considerations. This last assumption guarantees the existence of a global, nowhere-zero timelike vector field on M. Space-time will be assumed nonflat in the sense that the Riemann tensor does not vanish over an open subset of M. All structures on M will be assumed smooth.

II. HOLONOMY AND DECOMPOSABILITY

The metric connection of the space-time M leads naturally to the holonomy group of M, Φ_p at any $p \in M$. Full details of holonomy theory can be found in Ref. 4 (Vol. I). Since M is connected (and hence path connected) Φ_p and Φ_q are isomorphic for any $p,q \in M$ and one speaks of the holonomy group of M. This group is a connected (since M is simply connected) Lie group isomorphic to a connected subgroup of the (component of the identity of the) Lorentz group $\mathscr{L}_0$ and can thus be identified with one of the 15 types of subalgebra of the six-dimensional Lie algebra of $\mathscr{L}_0$. These subalgebras have been discussed elsewhere in the present context[5,6] and are labeled R_1–R_{15} according to Ref. 7. Here R_1 is the trivial case and R_{15} is the full Lorentz algebra (and R_5 is impossible for holonomy groups[5]). The holonomy groups arising from each of these subalgebras except R_{15} are *reducible* in the sense that if Φ_p is realized as a group of linear transformations from the tangent space T_pM to M at p onto itself, some nontrivial subspace of T_pM remains invariant under this group. Such a subspace then determines, by parallel transport, an integrable distribution on M and the resulting maximal integral submanifolds of this distribution constitute totally geodesic submanifolds of M whose nature (timelike, spacelike, null) is the same at each $p \in M$. The holonomy group of M is called *nondegenerately reducible* if some nontrivial, *non-null* subspace of T_pM is invariant under the holonomy in the above sense.

An important result was given by Wu[8] in the case that M is simply connected, (geodesically) complete, and nondegenerately reducible [a generalization of the well-known de Rham theorem for positive-definite manifolds—see, for example, Ref. 4 (Vol. I)]. In this case M is isometric to the metric product of the maximal integral submanifolds obtained as described above from a nontrivial, non-null, holonomy invariant subspace of T_pM and its orthogonal complement. If the above conditions on M, except completeness, are retained then M is necessarily locally decomposable[8] in the sense usually meant in general relativity.[1,9] The nontrivial nondegenerately reducible cases are the holonomy types R_3, R_6, and R_{10} $(1+3$ timelike$)$, R_{13} $(1+3$ spacelike$)$, R_4 $(1+1+2$ spacelike$)$, R_2 $(1+1+2$ timelike$)$, and R_7 $(2+2)$, where the description in parentheses refers in an

0022-2488/88/112465-08$02.50

obvious way to the nature of the decompostion. Thus the digit 1 refers to a non-null, holonomy invariant, one-dimensional distribution on M (and hence to an associated, global, nowhere-zero, covariantly constant, non-null vector field on M since M is simply connected) and the digits 2 and 3 refer to two- or three-dimensional, holonomy invariant timelike or spacelike (as indicated) distributions on M. These distributions span the totally geodesic, integral submanifolds of the holonomy and are themselves not nondegenerately reducible (but some of them are reducible).

The following results can now be stated.[10] To obtain maximum generality, M will be allowed to be flat in the results A, B, and C below, but elsewhere M will be assumed nonflat as stated earlier.

A: The following are equivalent for M.

(i) M is reducible.

(ii) M admits a global, nowhere zero recurrent vector field k (that is, in components, $k_{a;b} = k_a p_b$ for some global one-form field p).

(iii) M admits a global, nowhere-zero, second-order, symmetric recurrent tensor field S that is not proportional to the metric g.

B: The following are equivalent for M.

(i) M is nondegenerately reducible.

(ii) M admits a tensor field S satisfying the conditions of A(iii) above and also the condition that in every chart of M, $S^a{}_b S^b{}_c = S^a{}_c$. (This condition is given in Ref. 1.)

(iii) M admits a tensor field S satisfying the conditions of A(iii) and the condition that its Segrè type (necessarily the same everywhere) is $\{1,111\}$ or some degeneracy of this type [except, $\{(1,111)\}$].

In the statements B(ii) and B(iii) the tensor S may be chosen covariantly constant and so nondegenerately reducible space-times admit covariantly constant tensors S as above. A consideration of the holonomy types R_8 and R_{11} shows that the converse is false. The nondegenerately reducible types with R_7 excluded can be characterized by their admitting a global, nowhere zero, non-null covariantly constant vector field.

C: If M admits a tensor S satisfying the conditions A(iii) which is covariantly constant then all the eigenvalues of S are constants on M and one either has the type R_7 or else M admits a global, nowhere zero, covariantly constant vector field that may be chosen as an eigenvector field of S on M. The holonomy type of M is either R_1–R_4, R_6–R_8, R_{10}, R_{11}, or R_{13}.

The conditions of result C are equivalent to M admitting a Lorentz metric g' that is not a constant multiple of g but generates the same symmetric connection as g on M.[5]

III. AFFINE COLLINEATIONS

A bijection ψ: $M \to M$ such that ψ and ψ^{-1} are smooth and such that $\tilde{\psi}^* \omega = \omega$, where ω is the connection one-form on the frame bundle of M arising from the metric g on M, and $\tilde{\psi}$ the natural extension of ψ to the frame bundle is called an *affine transformation* of M. The question of the Lie group structure of the group of all such transformations is discussed in Refs. 3 and 4 (Vol. I). However, in this paper the more general, *local Lie groups of local affine transformations* (collineations) will be considered. These are characterized by a finite-dimensional Lie algebra of global *affine* vector fields on M. Such a vector field ξ on M can be characterized by the local coordinate condition

$$\xi_{a;bc} = R_{abcd} \xi^d. \tag{1}$$

A decomposition of $\xi_{a;b}$ into its symmetric and skew-symmetric parts and use of Eq. (1) gives (see, for example, Refs. 3 and 5)

$$\begin{aligned} &\text{(i) } \xi_{a;b} = h_{ab} + F_{ab} \quad (h_{ab} = h_{ba}, \quad F_{ab} = -F_{ba}),\\ &\text{(ii) } h_{ab;c} = 0, \\ &\text{(iii) } F_{ab;c} = R_{abcd} \xi^d \quad (\Rightarrow F_{ab;c} \xi^c = 0). \end{aligned} \tag{2}$$

The vector field ξ is *homothetic* (respectively, *Killing*) if $h_{ab} = \nu g_{ab}$ with $\nu = \text{const} \neq 0$ (respectively, $h_{ab} = 0$) and otherwise is called a *proper affine* vector field. In the last case one has a global, covariantly constant tensor field h on M as described in the last section. The global tensor F is called the *affine bivector*. The existence of the tensor h shows that if a *proper affine* is admitted the holonomy group of M is reducible[5]; in fact it is either nondegenerately reducible (and hence one has either a holonomy corresponding to the R_7 type or a covariantly constant, nowhere zero, non-null, global vector field on M) or else M admits a global, covariantly constant, nowhere zero null vector field (or both). The holonomy types are given in result C above. The allowed Petrov and Segrè types for the Weyl and energy-momentum tensors can be computed (and the table in Ref. 6 is sometimes useful but note this table is subject to the more restrictive conditions imposed in Ref. 6). As a result, and using Einsteins's equations with zero cosmological constant, many types of physical fields studied in general relativity theory are excluded.[5] For example, the existence of a proper affine eliminates all vacuum space-times except the *pp* waves, all perfect fluid space-times for which, in the usual notation, $0 \leqslant p \neq \rho \geqslant 0$, all non-null Einstein–Maxwell fields except the $2+2$ locally decomposable case (R_7), and all nonvacuum Einstein spaces again except the R_7 case.

It is convenient at this point to discuss the maximum number of global, independent, proper affine vector fields permitted. Let H denote the finite-dimensional vector space of global, covariantly constant, second-order, symmetric tensor fields on M and let $\dim H = n \geqslant 1$. Let K denote the subspace of H that consists of those members of H that arise from an affine vector field as described above and let $\dim K = m \leqslant n$. Finally suppose that M admits an r-dimensional Lie algebra of global affine vector fields. By taking appropriate linear combinations of these affine vector fields one can always arrange that $r - m$ of them are Killing. In fact, if $g \in K$ (equivalently M admits a global homothety), then one can arrange that $r - m + 1$ of these affine fields are homothetic ($r - m$ of which can be arranged to be Killing) and $m - 1$ of them are proper affine and if $g \notin K$ one can arrange that $r - m$ of them are Killing and m of them proper affine. The proof is straightforward. The vector spaces H and K may or may not be equal. In what is to follow, phrases like "the maximum number of independent proper affine vector fields" will always be taken in the above sense.

IV. THE 1+ 3 CASE

Consider first the 1 + 3 *spacelike* case corresponding to the holomony algebra R_{13}. Here one has a global, nowhere zero, covariantly constant, timelike vector field u uniquely determined by the coordinate relations $u_{a;b} = 0$, $u_a u^a = -1$ and the choice that u is future pointing, which is permissible since M is time orientable. The uniqueness of u follows from the holonomy condition and the only global independent, second-order, symmetric, covariantly constant tensors are g_{ab} and $u_a u_b$ as follows from result C. The totally geodesic submanifolds that arise from the holonomy are represented by the flow of u and the three-dimensional spacelike submanifolds orthogonal to u. The Ricci identity gives $R_{abcd}u^d = 0$ and hence $R_{ab}u^b = 0$ and so the Petrov type is I, D, or 0 (Ref. 11) and the Segrè type of the Ricci tensor is $\{1,111\}$ or one of its degeneracies.[12] Both types may vary over M.

Now the consequences of M admitting a global, affine vector field ξ will be evaluated. Of course, M already admits the affine vector field u. In local coordinates one has from (2),

$$\xi_{a;b} = \alpha g_{ab} + \beta u_a u_b + F_{ab} \quad (\alpha,\beta \text{ constants}). \tag{3}$$

The case when ξ is homothetic (respectively, Killing) corresponds to $\beta = 0$ (respectively, $\alpha = \beta = 0$). There is at most one global, independent, proper affine vector field in the sense of the previous section. If one defines a global, real-valued function κ on M given in local coordinates by $\kappa = u_a \xi^a$, then one easily finds $\kappa_{,ab} = 0$ and so $\kappa_a \equiv \kappa_{,a}$ is a constant multiple of u_a. In fact, (2) gives $\kappa_a u^a = \beta - \alpha$ and so $\kappa_a = (\alpha - \beta) u_a$ and then (3) gives $F_{ab}u^b = 0$. Hence F if nonzero is a simple, spacelike bivector and $\kappa = \text{const} \Leftrightarrow \alpha = \beta$ which, in turn, is equivalent to the Lie bracket $[\xi,u] = 0$. If $\kappa = \text{const}$ one can ensure it is nonzero by, if necessary, adding to ξ a constant multiple of u. This will not affect any of the above discussion and the decomposition (3) is left unchanged. With this assumed done, global vector fields k' and k can be constructed by the coordinate representations $k'^a = \kappa u^a$, $k^a = \xi^a + k'^a$ so that k is orthogonal to u, $k^a u_a = 0$, and $k' \not\equiv 0$. Then

$$k'_{a;b} = (\alpha - \beta) u_a u_b \quad [\Rightarrow k'^a{}_{;bc} = 0 \ (= R^a{}_{bcd} k'^d)], \tag{4}$$

$$k_{a;b} = \alpha (g_{ab} + u_a u_b) + F_{ab} \quad (\Rightarrow k^a{}_{;bc} = R^a{}_{bcd} k^d). \tag{5}$$

Hence k and k' are global, affine vector fields on M.

One can now consider the various cases obtained from the particular values of α, β, and F. If $\alpha \neq 0, \beta \neq 0$, then ξ and k are proper affine and k' is proper affine (respectively, Killing) if $\alpha \neq \beta$ (respectively, $\alpha = \beta$). If $\alpha \neq 0, \beta = 0$, then ξ is homothetic and k and k' are proper affine. If $\alpha = 0, \beta \neq 0$, then ξ and k' are proper affine and either F is identically zero on M, in which case k is identically zero on M (and so $\xi = -k'$), or F is not identically zero on M, in which case k is a Killing vector field on M. If $\alpha = \beta = 0$, then ξ is Killing on M and k' is a constant multiple of u. If also F is identically zero on M, then k is also and k' and $\xi (= -k')$ are constant multiples of u, whereas if F is not identically zero on M, k is a Killing vector on M. Whenever two or more proper affine vector fields occur they are not independent (in the sense of the previous paragraph), in that an appropriate linear combination of them is a global homothetic or Killing vector field [for example, if $0 \neq \alpha \neq \beta \neq 0$, $k + \alpha(\beta - \alpha)^{-1}k'$, $\xi + \beta(\beta - \alpha)^{-1}k'$, and $\xi - \beta\alpha^{-1}k$ are each homothetic].

Two further points may now be made. First, one always has a local, proper, affine vector field in the sense that for any $p \in M$ there exists a (contractible) neighborhood U of p and a real valued function w on U such that, on U, $u_a = w_{,a}$. Then wu^a is a proper, affine vector field on U satisfying (3) with $\alpha = 0, \beta = 1, F = 0$. This would produce a global, proper, affine vector field if M were contractible. Second, since the vector field k and the bivector field F are everywhere orthogonal to u they define in a natural way a global vector and bivector field, respectively, in each of the hypersurfaces orthogonal to u. The positive definite induced metric on each of these hypersurfaces is represented locally by $g_{ab} + u_a u_b$. Using standard coordinates from the local decomposition of M, the covariant constancy of u, and (5), one can show that either $k \equiv 0$ on M or else k determines a global nonidentically zero homothetic or Killing vector field with respect to the induced metric in each of these hypersurfaces. This follows since if the induced vector field from k is zero everywhere on a particular hypersurface then F is also zero on this hypersurface and $\alpha = 0$. Hence k is a Killing vector in M such that k^a and $k_{a;b}$ are simultaneously zero at some point of M and so $k \equiv 0$ on M since M is connected.

The conclusion here is that if M is to admit a global (or local) affine vector field not parallel to the covariantly constant vector field u, then the hypersurfaces orthogonal to u must themselves exhibit some homothetic or Killing symmetry and, in general, they will not, since no geometrical restriction of this nature is placed on these hypersurfaces.

In the case when M is *complete*, the nondegenerate reducibility of M and its simple connectedness give rise, as mentioned in Sec. II, to a global isometric decomposition $M = H \times \mathbb{R}$, where H is isometric to any of the submanifolds orthogonal to u. With the real valued function w representing the usual global chart on $\mathbb{R}$ in an obvious way, one sees that the proper affine vector field wu discussed earlier can now be taken as a global, proper affine vector field on M. Another global, proper, affine vector field on M not proportional to u will arise if and only if H admits a global homothetic or Killing vector field. Now since H is positive-definite, irreducible (by definition of the original holonomy) and complete (since M is complete and H totally geodesic) and since H cannot possess a nonempty open subset V in which the induced curvature tensor is zero (for then M would possess the flat open subset $V \times \mathbb{R}$), it follows that H admits no proper global homotheties [Ref. 4 (Vol. I)]. In this case one need only consider Killing symmetries on H and it follows that M admits no proper global homotheties. It also follows in this case from the completeness of M that all global affine vector fields that arise on M are complete vector fields [Ref. 4 (Vol. I)], and as a consequence of Palais' theorem[2] give rise to a global Lie group G action on M representing the affine transformations whose Lie algebra is isomorphic to the Lie algebra of global affine vector fields on M. Here, one can say more about G because any global Killing vector field in H with respect to the induced geometry in H gives rise in a natural way to a global Killing vector field on M as is easily checked using local coordinates adapted to the decomposition. It then follows that the Lie algebra of global, affine vector fields on M can be spanned by the vector fields u, wu and the global Killing vector fields of H. Further if ξ^1 represents either u or wu and ξ^2 is a global Killing vector field on M arising as described above from one in H then $[\xi^1,\xi^2]=0$ and so G is Lie isomorphic to $G_1 \times G_2$, where G_1 is the non-Abelian two-dimensional Lie group of affines spanned by u and wu and G_2 is the Lie group of isometries of H. Since $0 \leqslant \dim G_2 \leqslant 6$ one obtains $2 \leqslant \dim G \leqslant 8$.

If M is not complete then one still has a local metric product structure for M (local nondegenerate decomposability) such that each $p \in M$ may be assumed to lie in a contractible chart W which is isometric to $H' \times I$, where H' is a positive-definite three-dimensional submanifold of M everywhere orthogonal to u and I is an open interval of $\mathbb{R}$ with its usual induced metric. Again u and wu are affine vector fields on W, but the situation is different in two important respects from the complete case above. First, even though H' has no flat open subsets (for the same reason as above), there may now be homothetic or proper, affine vector fields defined on H' (with respect to the induced metric on H') and these can be extended naturally to affine vector fields defined on W. (The existence of a *proper*, affine vector field on H' is equivalent to the positive definite induced metric on H' having a reducible holonomy group [Ref. 4 (Vol. I)].) Second, one no longer has, in general, a Lie group of affine transformations on W, but only a local group action. The structure of the Lie algebra of affine vector fields on W may change if p and W are changed. In certain cases local affine vector fields many be extended to global ones and this will be considered in Sec. VII. However, whether M is complete or not, there may be local affine vector fields on some open subset W of M that are not globally extendible to M and the *local* affine structure of M may not be independent of W in this sense.

The general ideas given above apply to the 1 + 3 timelike cases also and so this case will not be discussed in detail. There are, however, some differences caused by the fact that the three-dimensional totally geodesic submanifolds obtained from the holonomy are now orthogonal to a global, covariantly constant, unit spacelike vector field y (Ref. 5) (and hence have an induced metric of Lorentz signature) and these will be briefly pointed out. The Ricci identity gives $R_{abcd}y^d=0$ and $R_{ab}y^b=0$. In the R_{10} case there are no restrictions on the Petrov or Segrè types[11,12] and the three-dimensional holonomy submanifolds mentioned above are irreducible. In the R_6 case these submanifolds are reducible but not nondegenerately so because M admits a global, recurrent, null vector field everywhere tangent to them. The Ricci identity can then be used to show that the Petrov type is N or III at points where the Ricci scalar $R=0$ and II or D at points where $R \neq 0$ (see, for example, Ref. 6). The Segrè type (which may vary over M) is either $\{(1,1)(11)\}$, $\{(31)\}$, $\{2(11)\}$, or $\{(211)\}$ as can be deduced from the table in Ref. 6 (but note the more restrictive conditions imposed in this reference). The corresponding affine bivector now satisfies $F_{ab}y^d=0$ and so if nonzero is simple but may be timelike, spacelike, or null. Finally, those results appealed to in the previous case and which relied upon the three-dimensional holonomy submanifolds inducing a positive-definite metric are no longer applicable.

The 2 + 2 (R_7) case will be discussed in the next section. Concerning the remainder of the nondegenerately decomposable cases: 1 + 3 timelike (R_3), 1 + 1 + 2 timelike (R_2), and 1 + 1 + 2 spacelike (R_4), their treatment is similar to that given above. For example, in the 1 + 1 + 2 spacelike case one has, on M, global, nowhere zero, unit, covariantly constant timelike and spacelike vector fields given in components by v^a and z^a, respectively, which together span the flat timelike submanifolds obtained from the holonomy. There are exactly four independent, global, second order, symmetric tensor fields whose components are $v_a v_b$, $z_a z_b$, $v_{(a}z_{b)}$, and g_{ab} and hence a maximum of three independent (in the sense used earlier) global, proper affine vector fields on M. A general global affine vector field ξ projects onto the spacelike submanifolds determined by the holonomy to give an affine vector field on M which is a global homothetic vector field in the induced geometry of these submanifolds. The affine bivector of ξ is in general nonsimple and its canonical blades are tangent to the above pair of holonomy submanifolds at each point (cf. the 2 + 2 case in Sec. V). The projection of ξ onto the timelike holonomy submanifold is a linear combination of the six independent affine vector fields admitted locally by this flat two-space and represented by the affine vector field $(a+bv+cz)v^a+(d+ev+fz)z^a$, where $v_a=v_{,a}z_a=z_{,a}$ and $a,b,\ldots,f$ are six arbitrary constants. Thus M always admits the local, independent *proper* affines vv^a, zz^a, and vz^a+zv^a (and these lead to global affines if M is contractible or complete). In either of the 1 + 1 + 2 cases, if M is complete, a Lie group G of affine transformations arises which satisfies $6 \leqslant \dim G \leqslant 9$. That the maximum number can be achieved follows from a consideration of the Lorentz manifold $S^2 \times M^2$, where S^2 has its usual meaning and (positive definite) metric and M^2 is two-dimensional Minkowski space. Further details can be found elsewhere.[13] It turns out, in fact, that 9 is the maximum dimension for the Lie algebra of global affine vector fields on any nonflat space-time as follows from the case by case study presented in this and the next section. An easier proof in a more general context will be given elsewhere.[14]

V. THE 2+ 2 CASE

Now consider the 2 + 2 case with holonomy algebra R^7 where the totally geodesic submanifolds determined by the holonomy constitute two orthogonal families of two-dimensional submanifolds, one spacelike and one timelike. Here one may introduce locally a null tetrad field l,n,x,y (with $l^a n_a = x^a x_a = y^a y_a = 1$, and all other inner products zero) such that l,n and x,y span the totally geodesic submanifolds at each point. In this case there are exactly two independent, covariantly constant, second-order, symmetric tensor fields defined globally on M.[5] They can be given locally by $P_{ab} = 2l_{(a}n_{b)}$ and $Q_{ab} = x_a x_b + y_a y_b$ and they are related to the metric by the completeness relation $g_{ab} = P_{ab} + Q_{ab}$. The Weyl tensor is either type D (with l and n spanning its principal null directions) or 0 and the Ricci tensor is either of Segrè type $\{(1,1)(11)\}$ (with l,n and x,y spanning the degenerate eigenspaces) or its degeneracy. Again the types may vary over M. The null vectors l and n can be globally defined on M (because M is simply connected) and are recurrent, while the bivectors represented locally by $M_{ab} = 2l_{[a}n_{b]}$ and $\overset{*}{M}_{ab} = 2x_{[a}y_{b]}$ can also be extended to global nowhere zero covariantly constant bivectors on M. [Here, an asterisk denotes the usual duality operator and $\overset{*}{M}_{ab}$ can be globally extended on M since M is orientable (because it is simply connected).] The space-time M admits no global covariantly constant, nowhere zero vector fields.[5] The Riemann tensor takes the local form[9]

$$R_{abcd} = aM_{ab}M_{cd} + b\overset{*}{M}_{ab}\overset{*}{M}_{cd}, \tag{6}$$

where a (respectively, b) depends only on the coordinates of the timelike (respectively, spacelike) totally geodesic submanifold in the usual locally decomposable charts.

Now suppose a global affine collineation ξ exists on M satisfying (1), (2) and hence the local relation

$$\begin{aligned}\xi_{a;b} &= \alpha g_{ab} + 2\beta l_{(a}n_{b)} + F_{ab}\\ &= (\alpha+\beta)P_{ab} + \alpha Q_{ab} + F_{ab},\end{aligned} \tag{7}$$

where α and β are constants and F is the affine bivector. Define global vector fields by the local relations $k^a = P^a{}_b\xi^b$ and $k'^a = Q^a{}_b\xi^b$, so that $\xi = k + k'$. Then for k one obtains from (7),

$$k_{a;b} = (\alpha+\beta)P_{ab} + P_{ac}F^c{}_b \tag{8}$$

and (2iii), (6), and (8) show that k is an affine vector field

$$k_{a;bc} = P_{ad}R^d{}_{bce}\xi^e = aP_{ad}M^d{}_bM_{ce}k^e = R_{abce}k^e. \tag{9}$$

Equation (9) shows that $k_{(a;b)}$ is covariantly constant and so $k_{(a;b)} = \mu g_{ab} + \nu P_{ab}$ (μ,ν constant). Inserting this into (8) and contracting with the local vector field x shows that $\mu = 0$, $\nu = (\alpha+\beta)$, and

$$F_{ab}x^b = -\rho y_a, \quad F_{ab}y^b = \rho x_a, \tag{10}$$

for some real-valued function ρ. Equation (10) shows that F if nonzero is a non-null bivector whose canonical pair of blades is spanned by the pairs l,n and x,y at each point and so

$$F_{ab} = \sigma M_{ab} + \sigma'\overset{*}{M}_{ab} \tag{11}$$

for real-valued functions σ and σ'. Then (8) gives

$$k_{a;b} = (\alpha+\beta)P_{ab} + \sigma M_{ab}. \tag{12}$$

In (12), k lies in the blade of the simple bivector M and so is *hypersurface orthogonal* in the sense that $k_{[a;b}k_{c]} = 0$. It is also noted here that Eqs. (2iii) and (12) show that $\sigma_{,a}$ (respectively, $\sigma'_{,a}$) lies everywhere in the blade of M_{ab} (respectively, $\overset{*}{M}_{ab}$) and that $\xi^a\sigma_{,a} = \xi^a\sigma'_{,a} = 0$.

Similar calculations show that the global vector field k' is affine,

$$k'_{a;bc} = R_{abcd}k'^d, \tag{13}$$

$$k'_{a;b} = \alpha Q_{ab} + \sigma'\overset{*}{M}_{ab}. \tag{14}$$

Now k' lies everywhere in the blade of $\overset{*}{M}_{ab}$ and is *hypersurface orthogonal* in the above sense.

As before, various cases can be distinguished depending on the values of α, β, and F. If $\alpha \neq 0$, $\beta \neq 0$, and $\alpha+\beta \neq 0$, then ξ, k, and k' are proper affine vector fields on M, and k and k' uniquely define homothetic Killing vector fields in their respective totally geodesic submanifolds with respect to the induced metrics. If $\alpha \neq 0$ and $\alpha+\beta = 0$, then ξ and k' are proper affines and k is Killing on M. Also k' is homothetic in its totally geodesic submanifold and either $k \equiv 0$ on M or else defines a nonidentically zero Killing field in each of its totally geodesic submanifolds. If $\alpha = 0$, $\beta \neq 0$ (respectively, $\alpha \neq 0$, $\beta = 0$), then ξ is proper affine (respectively, homothetic) on M, k is proper affine on M, and k' is Killing (respectively, proper affine) on M, while in their respective totally geodesic submanifolds k is homothetic and k' either Killing or identically zero (respectively, homothetic) on M. If $\alpha = \beta = 0$, then since M admits no global covariantly constant vector fields, any of ξ, k, and k', if not identically zero, is Killing on M with similar comments applying to k and k' with respect to their respective totally geodesic submanifolds.

The $2+2$ (R_7) case is thus qualitatively similar to the previous case. Consider the case when M is complete. Then each of the two-dimensional totally geodesic submanifolds determined by the holonomy is complete and M is isometric to $H_1 \times H_2$, where H_1 and H_2 are isometric to the above holonomy submanifolds. Suppose that H_1 is timelike and H_2 spacelike. The manifolds H_1 and H_2 are irreducible (by the holonomy condition), simply connected (since M is), and complete (since they are totally geodesic). Palais' theorem then shows that the affine vector fields on M (which are now complete since M is [Ref. 4 (Vol. I)]) lead to a Lie group G of affine transformations on M which, as before, is of the form $G = G_1 \times G_2$, where G_1 and G_2 are the Lie groups of homotheties on H_1 and H_2, respectively. (In fact, G_2 will be the Lie group of *isometries* on H_2. This follows because H_2 is positive definite, complete, and irreducible and hence possesses no global proper homotheties [Ref. 4 (Vol. I)]. It follows that $0 \leqslant \dim G \leqslant 6$. If H_1 possesses a global, proper homothety then it can possess at most one global Killing vector field and so in this case $\dim G_1 \leqslant 2$ and $\dim G \leqslant 5$. This is because if H_1 possessed two independent, global Killing vector fields, then since it is simply connected, it must possess a third (see Sec. VII). This gives $\dim G_1 = 4$ and so H_1 is flat, in contradiction to its irreducibility.)

If M is not complete, the discussion presented earlier in the incomplete $1+3$ case applies with obvious modifications. One major difference in both the complete and incomplete cases is that here one does not necessarily have a global or local proper affine vector field existing. In fact, there is at most one independent, global, proper affine vector field in the sense used in this paper.

Finally, in the case of interest for global, affine collineations when M is not nondegenerately reducible, one has a unique (up to a constant factor), global, covariantly constant, nowhere zero vector field l on M and it is null. The holonomy types here are R_8 and R_{11}. One can analyze these cases in the way that was done for the others, but less information was obtained. If one supposes that ξ is a global, affine vector field on M, then (1) and (2) hold with $h_{ab} = \alpha g_{ab} + \beta l_a l_b$ (α,β constants), and it can be shown that the affine bivector satisfies $F_{ab}l^b = \gamma l_a$ with γ constant. The constant γ need not be zero as can be seen by choosing ξ to be the standard homothety of the vacuum or generalized plane wave solution in general relativity.[15] If $\kappa = \xi^a l_a$, then $\xi + \kappa l$ is also affine and there is at most one independent, global affine vector field on M in the sense described earlier. One always has the local relation $l_a = \psi_{,a}$ for some real-valued function ψ, and so the local, proper affine vector field $\psi\psi_{,a}$ is always admitted which is global if M is contractible.

VI. FIXED POINTS OF AFFINE COLLINEATIONS

The subset of points of M which remain fixed by every member of the local group of local affine collineations generated by a nonidentically zero global affine vector field ξ on M is either empty, discrete, or is such that each of its components is a totally geodesic submanifold of M [Ref. 4 (Vol. II)]. If ξ is Killing such components are either empty, single points, or two-dimensional and if ξ is homothetic they are either empty, single points, or (part of) a null geodesic in M.[15] In this section a brief discussion will be given of the (local) fixed point structure of local one-parameter groups of local, proper affine collineations on M represented in a standard notation (Ref. 4, Vol. I) by local maps ϕ_t for t in some interval about $0 \in \mathbb{R}$. Such a local group may admit no fixed points. It admits a fixed point $p \in M$ if and only if $\xi(p) = 0$. Suppose p is a fixed point of all the ϕ_t and let ψ be the usual exponential diffeomorphism from some open neighborhood U' of $0 \in T_P M$ to some open neighborhood U of p. Then U and U' may be chosen so that $\psi \circ \phi_t{}^* = \phi_t \circ \psi$ holds where it makes sense [Ref. 4 (Vol. I)]. One may then follow the method used in Ref. 15 to study the fixed point structure because, at p, $\phi_t{}^* = \exp[t\xi^a{}_{;b}] = \exp[t(h^a{}_b + F^a{}_b)]$. Fixed points of all the ϕ_t correspond under ψ to members of $T_p M$ that are eigenvectors of the matrix $\xi^a{}_{;b}$ at p with zero eigenvalue and so one is left with an algebraic study of h and F at p. It turns out that either p is isolated or the component of the set of fixed points containing p is of dimension 1, 2, or 3. Further details of the possibilities can be found in Ref. 13.

More can be said about the conditions at a fixed point p of the ϕ_t if one recalls that for any affine vector field ξ on M (and denoting the Lie derivative along ξ by $\mathscr{L}_\xi$)

$$\mathscr{L}_\xi \operatorname{Ricc} = 0 \quad (\Leftrightarrow \phi_t{}^* \operatorname{Ricc} = \operatorname{Ricc}), \tag{15}$$

where now it is convenient to use the index free symbol Ricc for the Ricci tensor. Equation (15) and the condition on $\mathscr{L}_\xi g$ obtained directly from [2(i)] then supply important information about the algebraic structure of Ricc at p.[10] For example, consider the $1 + 3$ spacelike case. Here one has from (3)

$$\phi_t{}^* g = e^{-2\alpha t} g + \gamma u \otimes u, \tag{16}$$

where $\gamma = \text{const}$. From Sec. IV, Ricc is diagonalizable (over $\mathbb{R}$) everywhere (Segrè type $\{1,111\}$ or one of its degeneracies) and u is an eigenvector everywhere with zero eigenvalue. Following the argument given in Ref. 10 (see also a similar one in Refs. 3 and 15), let ω be any other eigenvector of Ricc at p with eigenvalue λ which without any loss of generality for the following argument may be taken to be orthogonal to u at p. Then for any $\omega' \in T_p M$, Ricc $(\omega,\omega') = \lambda g(\omega,\omega')$ and so from (15) and (16) one has

$$\begin{aligned} \phi_t{}^* \operatorname{Ricc}(\omega,\omega') &= \lambda\, [e^{2\alpha t}\phi_t{}^* g - \gamma e^{2\alpha t} u \otimes u](\omega,\omega') \\ &\Rightarrow \operatorname{Ricc}(\phi_t{}^*\omega,\phi_t{}^*\omega') \\ &= \lambda e^{2\alpha t} g(\phi_t{}^*\omega,\phi_t{}^*\omega') \end{aligned}$$

and so $\phi_t{}^*\omega$ is an eigenvector of Ricc at p with eigenvalue $\lambda e^{2\alpha t}$. Since there are at most four distinct eigenvalues of Ricc at p, one has $\alpha = 0$ or $\lambda = 0$. If $\alpha \neq 0$ at p, then all Ricci eigenvalues are zero at p, and because of the diagonalizability of Ricc, Ricc $= 0$ at p. It should also be noted that a zero of the vector field ξ is also a zero of the vector fields k and k' in the notation of Sec. IV and this fact may be used to give a variation of the above discussion of the Ricci eigenvalues at p. The other cases may be discussed similarly, but it should be noted that the vanishing of all the Ricci eigenvalues at some $p \in M$, in the general case, implies that either Ricc $= 0$ or has Segrè type $\{(211)\}$ with zero eigenvalue or Segrè type $\{(31)\}$ with zero eigenvalue at p.

VII. GLOBAL EXTENSIONS OF LOCAL AFFINE VECTOR FIELDS

Sections III–V discussed the consequences for a space-time M which admits a global, proper affine vector field. The work in these sections also showed how one could construct local affine vector fields on M, in the nondegenerately decomposable case, directly from the local decomposition by taking appropriate combinations of local affine vector fields (if any exist) in the submanifolds of decomposition. This is a method quite likely to occur in practice and so the question then arises whether or not these local affine vector fields can be extended to global affine vector fields on M. In general, the answer to this question is no. However, with one extra assumption on M, which is perhaps not unreasonable from the physical viewpoint, such an extension is always possible. Two proofs of this fact will be briefly mentioned. One is an extension of a proof due to Nomizu,[16] while the other relies on a covering space argument.[17]

Let U be a connected, coordinate domain of M and ξ a (local) affine vector field of M defined everywhere on U and hence satisfying (2) on U. If $p \in U$ and c is any curve in U through p with tangent w at p, then one has

$$\begin{aligned} &\xi_{a;b} w^b = h_{ab} w^b + F_{ab} w^b, \quad F_{ab;c} w^c = R_{abcd} w^c \xi^d, \\ &h_{ab;c} w^c = 0. \end{aligned} \tag{17}$$

It follows that if ξ_a, F_{ab}, and h_{ab} (that is, ξ_a and $\xi_{a;b}$) are given at p, then ξ is uniquely determined throughout U. Now let $\Omega_p M$ (respectively, $S_p M$) denote the vector space of all second-order, covariant, skew-symmetric (respectively, symmetric) tensors at p. Then let V_p be the direct sum $V_p = T_p M \oplus \Omega_p M \oplus S_p M$. If $A(U)$ is the Lie algebra of affine vector fields on U there is an obvious natural linear mapping $A(U) \to V_p$ given by $\xi \to (\xi^a(p), F_{ab}(p), h_{ab}(p))$, and which from the above discussion is one-to-one. Hence $\dim A(U) \leqslant \dim V_p = 20$. For each $p \in M$ let $A_p{}^*$ denote the vector space of germs of affine vector fields defined on some open neighborhood of p, that is, the vector space of equivalence classes of affine vector fields under the equivalence relation of equality on some open neighborhood of p. As Nomizu showed, one may always choose U such that $A(U)$

is naturally isomorphic to $A_p{}^*$ and such a U is called a *special* neighborhood of p. The theorem may now be stated.

Theorem: Let M be a (connected, smooth, Hausdorff) simply connected space-time manifold with smooth Lorentz metric g such that $\dim A_p{}^*$ is the same for each $p \in M$. Then if ξ is a smooth, affine vector field on some connected, open subset U of M it may be extended to a global, smooth, affine vector field on M.

Proof: One may prove the theorem by first using Eq. (17) and the above discussion to modify Nomizu's technique of prolongation. The details follow those given by Nomizu fairly closely and need not be repeated. They can be found in Ref. 13. An alternative proof has been given[17] which starts by choosing an open covering $\{U_\alpha\}$ of M consisting of all *special* coordinate neighborhoods of M. Each U_α carries an m-dimensional Lie algebra of affine vector fields, where $m = \dim A_p{}^*$, which is completely determined by the values of these vector fields and their first covariant derivatives at any point of U_α. Roughly speaking, one builds the union $\cup U_\alpha^\beta$ (where U_α^β denotes a choice of m independent members of this Lie algebra, labeled by β, on U_α) for all possible choices of α and β. By a process of identification and choice of (connected) component, the latter depending on the original given local vector field ξ, this set gives rise to a connected manifold M' that is a covering space of M, π: $M' \to M$ and admits a Lorentz metric $\pi^* g$ and m global affine vector fields. Since M is simply connected, π is an isometry $M' \to M$ and thus M admits m global affine vector fields, one of which, in fact, extends the original, local affine vector field ξ.

This result in no way depends on the dimension of M or the signature of g. As stated here, it applies to affine vector fields but it can be shown to apply equally well to homothetic, Killing, and conformal Killing vector fields.[17] It may also be used to establish a result mentioned without proof towards the end of Sec. V, namely, that if a simply connected two-dimensional manifold admits two independent, global Killing vector fields then a third, global Killing vector field is admitted. The existence of a third *local* Killing vector field in some neighborhood of each point of the manifold is, of course, well-known and so one immediately sees that the dimension of the vector space of germs of Killing vector fields associated with any $p \in M$ is equal to 3. The theorem (applied in the case of Killing vector fields) then provides the third, independent global Killing vector field on the manifold.

VIII. EXAMPLES AND CONCLUSIONS

This paper has attempted to present a complete discussion of the existence of affine collineations in (simply connected) space-times. The results indicate that they are of limited use in general relativity because of the restrictions their existence imposes on the space-time manifold. However, a number of standard examples admit groups or local groups of affine collineations and some of these can now be briefly discussed.

The Einstein static universe: Here the manifold is $\mathbb{R} \times S^3$ and is thus simply connected. A local form of the metric is

$$ds^2 = -dt^2 + dr^2 + \sin^2 r(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{18}$$

This space-time admits seven global Killing vector fields one of which may be taken timelike, covariantly constant, and equal to $\partial/\partial t$ when restricted to any of the above charts. The other six may be assumed to lie in the complete, spacelike submanifolds of constant positive curvature orthogonal to the above Killing vector [$(1+3)$ spacelike case]. The space-time is complete and so a seven-dimensional (transitive) Lie group of isometries arises. The results of Sec. IV show that an eight-dimensional (transitive) Lie group of affine transformations arises whose independent, proper affine members (in the sense described earlier) may be assumed generated by the global, proper affine vector field represented by $t(\partial/\partial t)$ in each of the above charts. This space-time admits no global, proper homothetic vector fields (or even local ones) as follows from the results of Sec. IV, the homogeneity of the space-time, and the results of Sec. VII.[17]

The Gödel universe: Here the manifold is $\mathbb{R}^4$ and is thus simply connected. The metric in a global chart is

$$ds^2 = -dt^2 + dx^2 + dz^2 - \tfrac{1}{2}e^{2\sqrt{2}\omega x}\, dy^2 - 2e^{\sqrt{2}\omega x}\, dt\, dy, \tag{19}$$

where ω is a positive constant. This manifold admits five global Killing vector fields and being complete, a five-dimensional transitive Lie group of isometries. The global vector field $(\partial/\partial z)$ is spacelike and covariantly constant and the $1+3$ timelike case results. The results of Sec. IV show that there exists a six-dimensional transitive Lie group of affine transformations and $z(\partial/\partial z)$ is a global, proper affine vector field. It can be shown that no local or global homothetic vector fields are admitted.

Another example of the $1+3$ timelike case is the manifold $\mathbb{R}^4$ together with the global metric

$$ds^2 = e^{ux}\, du\, dv + dx^2 + dy^2. \tag{20}$$

This metric was given in Ref. 18 and discussed further in Ref. 15. Here $\partial/\partial y$ is a global, covariantly constant spacelike vector field and there are two global independent Killing vector fields that may be taken as $\partial/\partial y$ and $\partial/\partial v$. In this case, however, there is a global, proper homothetic vector field given by

$$3v\frac{\partial}{\partial v} - u\frac{\partial}{\partial u} + x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}.$$

The three vector fields given above together with the global, proper affine vector field $y(\partial/\partial y)$ are all complete vector fields and so by Palais' theorem one has a four-dimensional (nontransitive) Lie group of affine transformations.

The Bertotti–Robinson metric provides an example of the $2+2$ case, but since such a space-time is the metric product of two two-dimensional spaces of constant curvature, inclusion in the present $2+2$ case would require both curvature constants to be nonzero. As a result neither of the two-dimensional manifolds can admit a homothety and from the results of Sec. V one sees that no global, proper homothetic and hence no global, proper affine vector fields occur in this case. Nevertheless, examples of the $2+2$ case admitting global, proper affine vector fields can be constructed.

Of those metrics admitting covariantly constant, global null vector fields the generalized plane waves provide examples where the manifold is $\mathbb{R}^4$ (hence simply connected) and the space-time complete. They admit a Lie group of isometries of dimension 5 or 6 (respectively, 6 or 7) in the type N (respectively, type 0) case. A global, proper homothetic vector field is always admitted[15,19] as is the global, proper affine vector field given at the end of Sec. V. Thus there is a Lie group of affine transformations of dimension 7 or 8 (respectively, 8 or 9) in the type N (respectively, type 0) case. They may be intrasitive (of dimension 7 or 8) or transitive (of dimension 8 or 9). Thus, again, the maximum dimension 9 is achieved (see the end of Sec. IV).

As a final remark, the scarcity of proper affine collineations can be made precise in a topological sense using a Whitney topology argument. It is discussed in Ref. 5 and is based on a theorem in Ref. 20.

ACKNOWLEDGMENT

One of the authors (J. da Costa) gratefully acknowledges the award of an I.N.I.C. Grant from the Portuguese Government.

[1] D. Kramer, H. Stephani, M. A. H. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations* (Deutscher Verlag der Wissenschaften, Berlin, 1980).

[2] R. S. Palais, Mem. Am. Math. Soc. **22** (1957).

[3] G. S. Hall, in *Relativity Today*, edited by Z. Perjes (World Scientific, Singapore, 1988).

[4] S. Kobayaski and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963, Vol. I, 1969, Vol. II).

[5] G. S. Hall, Gen. Relativ. Gravit. **20**, 399 (1988).

[6] G. S. Hall and W. Kay, J. Math. Phys. **29**, 428 (1988).

[7] J. F. Schell, J. Math. Phys. **2**, 202 (1961).

[8] H. Wu, Illinois J. Math. **8**, 291 (1964).

[9] G. S. Hall and W. Kay, J. Math. Phys. **29**, 420 (1988).

[10] G. S. Hall, University of Aberdeen preprint, 1987.

[11] C. B. G. McIntosh and E. H. van Leeuwen, J. Math. Phys. **23**, 1149 (1982).

[12] G. S. Hall, J. Phys. A **9**, 541 (1976).

[13] J. da Costa, Ph.D. thesis, University of Aberdeen.

[14] G. S. Hall and J. Steele, University of Aberdeen preprint, 1988.

[15] G. S. Hall, Gen. Relativ. Gravit. **20**, 671 (1988).

[16] K. Nomizu, Ann. Math. **72**, 105 (1960).

[17] G. S. Hall, to be published in Classical Quantum Gravity.

[18] J. K. Beem, Lett. Math. Phys. **2**, 317 (1978).

[19] L. Defrise-Carter, Commun. Math. Phys. **40**, 273 (1975).

[20] A. D. Rendall, J. Math. Phys. **29**, 1569 (1988).

Static spherically symmetric space-times with six Killing vectors

Asghar Qadir[a] and M. Ziad[b]
Department of Mathematics, Quaid-i-Azam University, Islamabad, Pakistan

(Received 31 March 1988; accepted for publication 29 June 1988)

It had been proved earlier that spherically symmetric, static space-times have ten, seven, six, or four independent Killing vectors (KV's), but there are no cases in between. The case of six KV's is investigated here. It is shown that the space-time corresponds to a hyperboloid cross a sphere, reminiscent of Kaluza–Klein theory, with a compactification from four down to two dimensions. In effect, there is a unique metric for this space-time corresponding to a uniform mass distribution over all space.

I. INTRODUCTION

Since Einstein presented his field equations there has been much work done on the classification of space-times according to their symmetries. In fact there is a complete classification of Einstein spaces,[1] and a great deal is known about exact solutions of the Einstein field equations for given stress-energy tensors.[2] In the first case one is limited to stress-energy tensors proportional to the metric tensor, while in the latter case many other stress-energy tensors are considered. An alternate approach would be to require a given space-time symmetry for arbitrary stress-energy tensors possessing the required symmetry. While following the spirit of Petrov's work on Einstein spaces, it allows for all stress-energy tensors as occur in the "exact solutions" approach. This approach was adopted for static spherically symmetric space-times.[3] Later it was noticed that more spherically symmetric space-times were possible.[4]

It was found[4] that the most general static spherically symmetric space-times admit ten, seven, six, or four Killing vectors (KV's). If there are ten KV's, the metric is either de Sitter, Minkowski, or anti-de Sitter. If there are seven KV's, the metric is either Einstein or anti-Einstein. In the case of four KV's (e.g., for Schwarzschild or Reissner–Nordstrom metrics), there is no further restriction on the metric. The six KV's occur with a metric of the form given by the equation

$$ds^2 = e^{\nu(r)}\,dt^2 - e^{\lambda(r)}\,dr^2 - a^2(d\theta^2 + \sin^2\theta\,d\phi^2), \qquad (1)$$

satisfying the differential constraint involving an arbitrary constant α,

$$\tfrac{1}{2}(\nu' e^{-\lambda/2})' = -\alpha e^{-\nu+\lambda/2} \quad (\alpha \gtreqless 0). \qquad (2)$$

In this paper we investigate the properties of the space-time with six KV's. It appeared as if Eq. (1), along with the differential constraint given by Eq. (2), represents a class of metrics. As we show, it represents a unique metric with the radial coordinate transformed arbitrarily. (This includes the Bertotti–Robinson metric.[2])

II. DISCUSSION

The reason why the metric given by Eq. (1) appeared to represent a class of space-times was that there were two arbitrary functions related by one second-order differential constraint. However, our choice of the radial coordinate is now arbitrary, as the coefficient of the "solid angle element" is a^2, instead of r^2 in the usual spherically symmetric metric. We are at liberty, therefore, to define a new "radial" coordinate x by

$$dx = e^{\lambda(r)/2}\,dr. \qquad (3)$$

Since there is only a simple integration involved, x is well defined by Eq. (3). In these coordinates Eq. (1) becomes

$$ds^2 = e^{\nu(x)}\,dt^2 - dx^2 - a^2(d\theta^2 + \sin^2\theta\,d\phi^2), \qquad (4)$$

and correspondingly Eq. (2) reduces to

$$\nu'' = -2\alpha e^{-\nu}. \qquad (5)$$

Taking logarithms and differentiating Eq. (5) with respect to x gives

$$\nu''' = -\nu'\nu'' = -\tfrac{1}{2}[(\nu')^2]'. \qquad (6)$$

Integrating Eq. (6) twice we get

$$e^{\nu} = \cos^2(A + \sqrt{\alpha}x) \qquad (\alpha > 0), \qquad (7)$$

$$e^{\nu} = e^{A+Bx} \qquad (\alpha = 0), \qquad (8)$$

$$e^{\nu} = \cosh^2(A + \sqrt{-\alpha}x) \qquad (\alpha < 0), \qquad (9)$$

where A and B are arbitrary constants.

We see that in each of the case $\alpha \lesseqgtr 0$, Eq. (1), along with Eq. (7), (8), or (9), respectively, represents a unique metric. In these cases the six KV's for $\alpha \gtreqless 0$, respectively, are

$$\underline{K} = \left[C_0 - \frac{\nu'(x)}{2\sqrt{\alpha}}(C_1 \cosh\sqrt{\alpha}t + C_2 \sinh\sqrt{\alpha}t)\right]\frac{\partial}{\partial t} + (C_1 \sinh\sqrt{\alpha}t + C_2 \cosh\sqrt{\alpha}t)\frac{\partial}{\partial x} + (C_3\cos\phi + C_4\sin\phi)\frac{\partial}{\partial\theta} + [\cot\theta(-C_3\sin\phi + C_4\cos\phi) + C_5]\frac{\partial}{\partial\phi}, \qquad (10)$$

$$\underline{K} = \left[C_0 - \frac{B}{2}\left(C_1 t + C_2\frac{t^2}{2}\right) + \frac{C_2}{B}e^{-\nu(x)}\right]\frac{\partial}{\partial t} + (C_1 + C_2 t)\frac{\partial}{\partial x} + (C_3\cos\phi + C_4\sin\phi)\frac{\partial}{\partial\theta} + [\cot\theta(-C_3\sin\phi + C_4\cos\phi) + C_5]\frac{\partial}{\partial\phi}, \qquad (11)$$

---

[a] Also, Center of Basic Sciences, UGC, H-9, Islamabad, Pakistan.

[b] Also, Department of Mathematics, Government Postgraduate College, Asghar Mall, Rawalpindi, Pakistan.

0022-2488/88/112473-02$02.50

$$
\underset{-}{K}=\left[C_0-\frac{\nu'(x)}{2\sqrt{\alpha}}(C_1\sin\sqrt{\alpha}t-C_2\cos\sqrt{\alpha}t)\right]\frac{\partial}{\partial t}
+(C_1\cos\sqrt{\alpha}t+C_2\sin\sqrt{\alpha}t)\frac{\partial}{\partial x}
+(C_3\cos\phi+C_4\sin\phi)\frac{\partial}{\partial\theta}
+[\cot\theta(-C_3\sin\phi+C_4\cos\phi)+C_5]\frac{\partial}{\partial\phi}. \quad (12)
$$

The three KV's corresponding to the parameters C_3, C_4, and C_5 satisfy an SO(3) Lie algebra, while the other three KV's (corresponding to C_0, C_1, and C_2) satisfy an SO(1,2) Lie algebra. Thus the symmetry of the space-time has the local structure of SO(3) ⊗ SO(1,2). The SO(3) acts on a spacelike S^2 (a sphere of radius a), while the SO(1,2) is the equivalent of the de Sitter symmetry for one time and one space dimension on a hyperboloid.

So far we have not used the Einstein field equations at all. It is not generally regarded as very meaningful to take a metric, insert it into the field equations, and obtain a stress-energy tensor, and hence the matter–energy distribution in the space-time. However, it is useful for our purposes to do so.

The Ricci tensor corresponding to the metric given by Eq. (1) is

$$
\begin{aligned}
&R^0_0=(e^\nu/2)(\nu''+\nu'^2/2)=R^1_1,\\
&R^2_2=-1/a^2=R^3_3,\quad R^\mu_\nu=0\quad(\mu\neq\nu).
\end{aligned}\quad (13)
$$

Thus the stress-energy tensor is diagonal and is given by

$$
T^0_0=T^1_1=1/\kappa a^2,\quad T^2_2=T^3_3=|\alpha|/4\kappa\quad(\alpha\neq0), \quad (14)
$$

$$
T^0_0=T^1_1=1/\kappa a^2,\quad T^2_2=T^3_3=-B^2/4\kappa\quad(\alpha=0). \quad (15)
$$

Thus the mass density in the space-time is constant.

For completeness we give the geodesics in this space-time for $\alpha=0$:

$$
\theta=\cos^{-1}[\theta_0\cos(\phi-\phi_0)], \quad (16)
$$

$$
x=x_0+B^{-1}\ln[1+B^2(t-t_0)^2/4], \quad (17)
$$

where ϕ_0, θ_0, x_0, and t_0 are the constants giving the specific geodesic. In the cases $\alpha \gtrless 0$ we get integrals for the equation relating x and t:

$$
t-t_0=\int\frac{k\sec^2(\sqrt{\alpha}x+A)dx}{\sqrt{k^2\sec^2(\sqrt{\alpha}x+A)-1}} \quad (18)
$$

and

$$
t-t_0=\int\frac{k\,\mathrm{sech}^2(\sqrt{-\alpha}x+A)dx}{\sqrt{k^2\,\mathrm{sech}^2(\sqrt{-\alpha}x+A)-1}}. \quad (19)
$$

It is interesting to note that the space-time is like the compactified Kaluza–Klein space-times,[5] except that here we have compactification from four down to two space-time dimensions. Presumably an analog of this compactification from ten down to four could be of interest in superstring theory.[6] It is also interesting to note that for $\alpha=0$ and $B<0$ there is a finite total mass,

$$
M=(1/B)e^{A/2}. \quad (20)
$$

Hence in some sense it corresponds to a finite "universe," since the volume element decreases as x increases.

ACKNOWLEDGMENTS

We would like to thank Faheem Hussain and Amjad Pervez for some useful discussions. We also acknowledge some useful techniques conveyed to us by Hassan Azad.

[1] A. Z. Petrov, *Einstein Spaces* (Pergamon, New York, 1969).
[2] D. Kramer, H. Stephani, M. MacCallum, and E. Herlt, *Exact Solutions of Einstein Field Equations* (Cambridge U.P., Cambridge, 1980).
[3] A. H. Bokhari and A. Qadir, J. Math. Phys. **28**, 1019 (1987).
[4] A. H. Bokhari and A. Qadir, "Killing vectors of static, spherically symmetric metrics," Department of Mathematics, Quaid-i-Azam University preprint, 1987.
[5] Th. Kaluza, Sitzunger. Preuss. Akad. Wiss. Berlin, Math. Phys. **K1**, 966 (1921); O. Klein, Z. Phys. **37**, 895 (1926); T. Applequist, A. Chodos, and P. G. O. Freund, *Modern Kaluza-Klein Theories* (Addison-Wesley, Reading, MA, 1987).
[6] M. B. Green and G. T. Horowitz, Phys. Lett. B **149**, 117 (1984); P. Candelas, G. T. Horowitz, A. Strominger, and E. Witten, Nucl. Phys. B **258**, 46 (1985); J. H. Schwarz, *Superstrings* (World Scientific, Singapore, 1985), Vols. I and II.

Birkhoff and Taub theorems generalized to metrics with conformal symmetries

Kuang zhi-quan
Center of Theoretical Physics, CCAST (World Lap.), Beijing, China and Institute of Mathematics, Academia Sinica, Beijing, China
Liang can-bin
Center of Theoretical Physics, CCAST (World Lap.), Beijing, China and Department of Physics, Beijing Normal University, Beijing, China

(Received 13 January 1988; accepted for publication 29 June 1988)

The Birkhoff theorem for spherically symmetric vacuum solutions and the Taub theorem for plane-symmetric vacuum solutions are both generalized to vacuum solutions with conformal symmetries, i.e., it is proved in this paper that any conformally spherically (resp. plane-) symmetric vacuum solution to Einstein equations must be the Schwarzschild (resp. either Taub–Kasner or flat) solution.

I. INTRODUCTION

Two theorems concerning symmetric vacuum solutions to Einstein equations are well known: The Birkhoff theorem states that any spherically symmetric vacuum solution must be the Schwarzschild solution (including the flat solution as a special case), while the Taub theorem[1,2] states that any plane-symmetric vacuum solution must be either the solution represented by

$$dS^2 = \pm z^{-1/2}(-dt^2 + dz^2) + z(dx^2 + dy^2) \qquad (1)$$

or the flat solution. Metric (1) with the plus sign is referred to as the Taub solution while that with the minus sign is referred to as the Kasner solution.[3] We will refer to (1) as the Taub–Kasner solution in this paper.

A question of interest is as follows: can these two theorems be generalized to vacuum solutions with conformal symmetries, i.e., is it true that any vacuum solution conformal to a spherically (resp. plane-) symmetric metric have to be Schwarzschild (resp. Taub–Kasner or flat)?

The issue of conformal mapping between two Einstein spaces was first investigated by Brinkmann.[4] A space-time is said to be an Einstein space if

$$R_{ab} = Rg_{ab}/n,$$

where g_{ab} and R_{ab} are the metric and Ricci tensor, respectively, $R \equiv g^{ab}R_{ab}$ and n the dimension of the space-time. A metric satisfying this condition is referred to as an Einstein metric. Vacuum spaces are obviously Einstein spaces. A main result of Brinkmann is that if two vacuum solutions are conformal to each other, the conformal factor must be a constant unless both of them are *pp*-wave metrics.[5] Our problem involves a vacuum solution $\hat{g}_{ab}$ and a conformally related spherically (resp. plane-) symmetric metric g_{ab} (i.e., $\hat{g}_{ab} = \Omega^2 g_{ab}$ for some smooth, nowhere vanishing function Ω). If g_{ab} is also a vacuum solution, then Brinkmann's result together with the Birkhoff (resp. Taub) theorem immediately requires that $\hat{g}_{ab}$ be Schwarzschild (resp. Taub–Kasner or flat). However, there exist plenty of spherically (resp. plane-) symmetric metrics g_{ab} that are not vacuum solutions (not even Einstein metrics). Thus the following question still remains to be answered: Does a vacuum solution $\hat{g}_{ab}$ conformal to a spherically (resp. plane-) symmetric non-Einstein metric g_{ab} have to be Schwarzschild (resp. Taub–Kasner or flat)? The main purpose of the present paper is to prove an affirmative answer to this question.

By virtue of the fact that any two-dimensional metric is locally conformally flat, all spherically (plane-) symmetric metrics can be written in the forms

$$dS_S^2 = \Omega^2(t,r)[H(t,r)(-dt^2 + dr^2) + d\vartheta^2 + \sin^2\vartheta\, d\phi^2]$$
$$\equiv \Omega^2(t,r)dS_{HS}^2$$

for spherically symmetric metrics, and

$$dS_P^2 = \Omega^2(t,z)[H(t,z)(-dt^2 + dz^2) + dx^2 + dy^2]$$
$$\equiv \Omega^2(t,z)dS_{HP}^2$$

for plane-symmetric metrics. Since the metrics dS_{HS}^2 and dS_{HP}^2 (referred to as the "H-type metrics" hereafter) are much simpler than dS_S^2 and dS_P^2, and since we are concerned only with conformal symmetries, we prefer to deal with dS_{HS}^2 and dS_{HP}^2 instead of dS_S^2 and dS_P^2. The metrics dS_{HS}^2 and dS_{HP}^2 can be taken as two extreme cases ($\zeta = 1$ and $\zeta = 0$) of the following metric:

$$dS_H^2 = H(t,r)(-dt^2 + dr^2) + d\vartheta^2$$
$$+ (1 - \zeta\cos^2\vartheta)d\phi^2 \quad (\zeta \text{ a const, } 0 \leqslant \zeta \leqslant 1),$$

which can alternatively be expressed as

$$dS_H^2 = (1-\zeta)[H(t,r)(-dt^2 + dr^2) + d\vartheta^2 + d\phi^2]$$
$$+ \zeta\,[H(t,r)(-dt^2 + dr^2)$$
$$+ d\vartheta^2 + \sin^2\vartheta\, d\phi^2] = (1-\zeta)dS_{HP}^2 + \zeta dS_{HS}^2,$$

and it is also interesting to ask whether there is any Ricci-flat solution conformal to dS_H^2 with $\zeta \neq 0,1$. A negative answer will be given in Sec. III.

II. GENERAL *H*-TYPE METRICS

Given an H-type metric dS_H^2 with $\zeta \in [0,1]$, we want to know if there is any function $\Omega(t,r,\vartheta,\phi)$ such that $\widehat{dS}^2 = \Omega^2\, dS_H^2$ is Ricci-flat. The nonvanishing Christoffel sym-

0022-2488/88/112475-04$02.50

bols and the Ricci tensor components of dS_H^2 are

$$\Gamma^r_{rr} = \Gamma^r_{tt} = \Gamma^t_{rt} = \frac{1}{2}\frac{\partial \ln|H|}{\partial r},$$

$$\Gamma^t_{rr} = \Gamma^t_{tt} = \Gamma^r_{rt} = \frac{1}{2}\frac{\partial \ln|H|}{\partial t},$$

$$\Gamma^\phi_{\vartheta\phi} = \zeta \sin\vartheta \cos\vartheta/(1-\zeta\cos^2\vartheta),$$

$$\Gamma^\vartheta_{\phi\phi} = -\zeta\sin\vartheta\cos\vartheta;$$

$$R_{rr} = -R_{tt} = \frac{1}{2}\left(\frac{\partial^2 \ln|H|}{\partial t^2} - \frac{\partial^2 \ln|H|}{\partial r^2}\right),$$

$$R_{\vartheta\vartheta} = 1 + (\zeta - 1)(1-\zeta\cos^2\vartheta)^{-2},$$

$$R_{\phi\phi} = (1-\zeta\cos^2\vartheta) + (\zeta-1)(1-\zeta\cos^2\vartheta)^{-1}.$$

It will be proved in the Appendix that (a) any nonflat H-type metric dS_H^2 with $\zeta \neq 1$ is not an Einstein metric; and (b) any H-type metric with $\zeta = 1$ is not an Einstein metric if it is to be conformal to a vacuum solution. Thus all dS_H^2 dealt with in this paper are not Einstein metrics and the answer to our question is beyond the scope of Brinkmann's results.

Using the relation between the Ricci tensors of two conformally related metrics[6] one has

$$\Omega^{-1}R_{\mu\nu} + 2(\Omega^{-1})_{;\mu\nu} - \Psi g_{\mu\nu} = 0, \tag{2}$$

where

$$\Psi \equiv \Omega^{-3}(\Omega^2)_{;\rho\sigma}g^{\rho\sigma}/2. \tag{3}$$

Formula (2) represents ten equations restricting the unknown function Ω,

$$\frac{\partial^2\Omega^{-1}}{\partial r\,\partial\vartheta} = \frac{\partial^2\Omega^{-1}}{\partial r\,\partial\phi} = \frac{\partial^2\Omega^{-1}}{\partial t\,\partial\vartheta} = \frac{\partial^2\Omega^{-1}}{\partial t\,\partial\phi} = 0, \tag{4}$$

$$\frac{\partial^2\Omega^{-1}}{\partial\vartheta\,\partial\phi} - \zeta\sin\vartheta\cos\vartheta(1-\zeta\cos^2\vartheta)^{-1}\frac{\partial\Omega^{-1}}{\partial\phi} = 0, \tag{5}$$

$$\frac{\partial^2\Omega^{-1}}{\partial r\,\partial t} - \Gamma^r_{rt}(\Omega^{-1})_{,r} - \Gamma^t_{rt}(\Omega^{-1})_{,t} = 0, \tag{6}$$

$$\Omega^{-1}R_{rr} + 2(\Omega^{-1})_{;rr} - \Psi H = 0, \tag{7}$$

$$\Omega^{-1}R_{tt} + 2(\Omega^{-1})_{;tt} + \Psi H = 0, \tag{8}$$

$$\Omega^{-1}R_{\vartheta\vartheta} + 2(\Omega^{-1})_{;\vartheta\vartheta} - \Psi = 0, \tag{9}$$

$$\Omega^{-1}R_{\phi\phi} + 2(\Omega^{-1})_{;\phi\phi} - \Psi(1-\zeta\cos^2\vartheta) = 0. \tag{10}$$

It follows directly from Eqs. (4) that

$$\Omega^{-1} = L(t,r) + S(\vartheta,\phi),$$

where L and S are arbitrary functions of the arguments shown.

Introducing coordinates $u = t - r$, $v = t + r$, it follows from Eqs. (6)–(8) that

$$\frac{\partial^2 L}{\partial u^2} - \frac{\partial \ln|H|}{\partial u}\frac{\partial L}{\partial u} = 0, \quad \frac{\partial^2 L}{\partial v^2} - \frac{\partial \ln|H|}{\partial v}\frac{\partial L}{\partial v} = 0,$$

which can be integrated to yield

$$H = \lambda(v)\frac{\partial L}{\partial u} = \mu(u)\frac{\partial L}{\partial v}, \tag{11}$$

where $\lambda(v)$ and $\mu(u)$ are arbitrary functions of their own arguments. Introducing new coordinates (u^*,v^*) and (t^*,r^*) such that

$$\frac{du^*}{du} = \frac{\mu(u)}{2}, \quad \frac{dv^*}{dv} = \frac{\lambda(v)}{2},$$

$$t^* = (v^* + u^*)/2, \quad r^* = (v^* - u^*)/2,$$

one obtains

$$\frac{\partial L}{\partial u^*} = \frac{\partial L}{\partial v^*},$$

hence

$$\frac{\partial L}{\partial r^*} = 0,$$

or

$$L = L(t^*).$$

Equation (11) then yields

$$H = \frac{dL}{dt^*}\frac{du^*}{du}\frac{dv^*}{dv}.$$

The preceding argument shows that there always exists a coordinate system (t^*,r^*,ϑ,ϕ) in which the Ricci-flat metric conformal to an H-type metric can be expressed as

$$d\hat{S}^2 = \Omega^2\Big[\frac{dL}{dt^*}(-dt^{*2} + dr^{*2}) + d\vartheta^2 + (1-\zeta\cos^2\vartheta)d\phi^2\Big]$$

and hence $H(u^*,v^*) = dL/dt^*$. From now on we will stick to this coordinate system and drop the asterisks.

Equation (5) can be integrated to give

$$S(\vartheta,\phi) = f(\phi)(1-\zeta\cos^2\vartheta)^{1/2} + h(\vartheta), \tag{12}$$

where $f(\phi)$ and $h(\vartheta)$ are arbitrary functions of their own arguments. This, together with Eqs. (9) and (10), gives

$$P_1(\vartheta) - f''(\phi) - f(\phi)P_2(\vartheta) = 0, \tag{13}$$

where

$$P_1(\vartheta) \equiv h''(\vartheta)(1-\zeta\cos^2\vartheta)^{1/2} - h'(\vartheta)\zeta\sin\vartheta\cos\vartheta(1-\zeta\cos^2\vartheta)^{-1/2},$$

$$P_2(\vartheta) \equiv -\zeta[1 - 2\sin^2\vartheta/(1-\zeta\cos^2\vartheta)]. \tag{14}$$

We now have to discuss the cases $\zeta\in(0,1)$, $\zeta = 0$, and $\zeta = 1$ separately.

III. *H*-TYPE METRICS WITH $\zeta\in(0,1)$

Differentiating (13) with respect to ϑ and ϕ, respectively, one gets

$$f'(\phi)P_2'(\vartheta) = 0,$$

which implies that either $P_2'(\vartheta) = 0$ or $f'(\phi) = 0$. The first choice is, however, unacceptable since, on account of Eq. (14), it is true only for some fixed values of ϑ when $\zeta \neq 0,1$; thus one has $f'(\phi) = 0$. On the other hand, Eqs. (8) and (9) yield

$$L(t)A(\vartheta) + S(\vartheta,\phi)B(t) + Q(t) + P(\vartheta,\phi) = 0, \tag{15}$$

where

$$A(\vartheta) \equiv (\zeta-1)(1-\zeta\cos^2\vartheta)^{-2}, \quad B(t) \equiv 1 + R_{tt}H^{-1},$$

$$Q(t) \equiv L(t)B(t) + 2L_{;tt}H^{-1}, \tag{16}$$

$$P(\vartheta,\phi) \equiv S(\vartheta)A(\vartheta) + 2\frac{\partial^2 S}{\partial\vartheta^2}.$$

Differentiation of (15) with respect to ϑ and t, respectively, yields

$$\frac{dL}{dt}\frac{dA}{d\vartheta}+\frac{dB}{dt}\frac{dS}{d\vartheta}=0.$$

Since $dL/dt=H\neq 0$, and $dA/d\vartheta\neq 0$ for $\zeta\neq 0,1$, one has $dS/d\vartheta\neq 0$ and hence

$$\frac{dA}{d\vartheta}\left(\frac{dS}{d\vartheta}\right)^{-1}=\frac{dB}{dt}\left(\frac{dL}{dt}\right)^{-1}=\text{const},$$

but $dA/d\vartheta$ and $dS/d\vartheta$ can be calculated, respectively, from (16) and (12) (with f a constant) and the ratio of them is not a constant, showing that the choice $f'(\phi)=0$ is also unacceptable. Therefore we conclude that there exists no Ricci-flat metric conformal to an H-type metric with $\zeta\neq 0,1$.

IV. *H*-TYPE METRICS WITH $\zeta=1$ and $\zeta=0$

An H-type metric with $\zeta=1$ (resp. $\zeta=0$) is a spherically (resp. plane-) symmetric one. In these two cases $P_2=\zeta$ and

$$P_1(\vartheta)=h''(\vartheta)(1-\zeta\cos^2\vartheta)^{1/2}-h'(\vartheta)\zeta\cos\vartheta;$$

thus Eq. (13) becomes

$$h''(\vartheta)(1-\zeta\cos^2\vartheta)^{1/2}-h'(\vartheta)\zeta\cos\vartheta$$
$$-f''(\phi)-\zeta f(\phi)=0,$$

implying

$$h''(\vartheta)(1-\zeta\cos^2\vartheta)^{1/2}-h'(\vartheta)\zeta\cos\vartheta=\rho\ \ (\text{const}), \tag{17}$$

which in turn gives

(A) for $\zeta=1$, $h(\vartheta)=-\rho\sin\vartheta+b\cos\vartheta+a$, a,b const,

hence

$$S(\vartheta,\phi)=[f(\phi)-\rho]\sin\vartheta+b\cos\vartheta+a,$$

and

$$\frac{\partial^2 S}{\partial\vartheta^2}=-S+a; \tag{18}$$

(B) for $\zeta=0$, $h''(\vartheta)=a$,

and hence

$$\frac{\partial^2 S}{\partial\vartheta^2}=a. \tag{18'}$$

Expressions (18) and (18') can be written in a unified form,

$$\frac{\partial^2 S}{\partial\vartheta^2}=-\zeta S+a. \tag{18''}$$

Equations (15) and (16) then lead to

$$S(R_{tt}H^{-1}-\zeta)+L(R_{tt}H^{-1}+\zeta)+2a+2L_{;tt}H^{-1}=0. \tag{15'}$$

To make further discussion using this equation we note that for each case ($\zeta=1$ or $\zeta=0$) one has to distinguish three subcases characterized by the quantity $R_{tt}H^{-1}$, and it can be shown that this quantity is invariant under coordinate transformations involving t and r, hence the division into these subcases is coordinate independent.

Subcase (1): $R_{tt}H^{-1}=\zeta$ everywhere.

It is easily checked that the Weyl tensor of dS_H^2 vanishes in this subcase, implying that dS_H^2 and hence $d\hat{S}^2$ are conformally flat. But $d\hat{S}^2$ is Ricci-flat as well, and therefore it must be flat. [In the case of $\zeta=0$, the conformal flatness of $d\hat{S}^2$ can easily be seen without recourse to the Weyl tensor: since $2R_{tt}=-\partial^2\ln|H|/\partial u\,\partial v$, $R_{tt}=0$ implies $H=P(u)Q(v)$, where $P(u)$ and $Q(v)$ are arbitrary functions. Therefore $dS_H^2=-P(u)Q(v)du\,dv+d\vartheta^2+d\phi^2$ is flat and hence $d\hat{S}^2$ is conformally flat.]

Subcase (2): $R_{tt}H^{-1}\neq\zeta$ everywhere.

In this subcase Eq. (15') leads to $S=\text{const}$, and hence can be absorbed into $L(t)$ to give $\Omega^{-1}=L(t)$. The unknown function Ω is originally restricted by ten equations: (4)–(10). Equations (4) are fulfilled by setting $\Omega^{-1}=L(t,r)+S(\vartheta,\phi)$, Eq. (5) is satisfied by writing S in form (12) (including $S=0$), Eq. (6) is satisfied by setting $L=L(t)$, $H=dL/dt$, and Eq. (9) is equivalent to (10) for $S=0$. What is left then is three equations—(7)–(9) containing Ψ. It follows by direct calculation from (3) that

$$\Psi=-\frac{3}{LH}\left(\frac{dL}{dt}\right)^2+\frac{1}{H}\frac{d^2L}{dt^2}, \tag{19}$$

hence Eq. (9) becomes

$$L\frac{d^2L}{dt^2}-3\left(\frac{dL}{dt}\right)^2-\zeta L^2\frac{dL}{dt}=0. \tag{9'}$$

Introducing a new function

$$F=L^{-2}\frac{dL}{dt}=\frac{-d\Omega}{dt},$$

one has

$$d\hat{S}^2=-F^{-1}d\Omega^2+F\,dr^2$$
$$+\Omega^2[d\vartheta^2+(1-\zeta\cos^2\vartheta)d\phi^2],$$

and Eq. (9') is equivalent to

$$\frac{dF}{d\Omega}=-\frac{F+\zeta}{\Omega},$$

which can be integrated to yield

$$F=-\zeta+\alpha\Omega^{-1},\quad \alpha\ \text{const}. \tag{20}$$

It is straightforward to check that (20) also satisfies Eqs. (7) and (8), and consequently is the general solution to the ten original equations. The metric $d\hat{S}^2$ now can be expressed as follows.

(A) $\zeta=1$:

$$d\hat{S}^2=(1-\alpha/\Omega)^{-1}d\Omega^2-(1-\alpha/\Omega)dr^2$$
$$+\Omega^2(d\vartheta^2+\sin^2\vartheta\,d\phi^2),$$

which can be written in the standard Schwarzschild form by setting $T=r$ and $R=\Omega$:

$$d\hat{S}^2=-(1-\alpha/R)dT^2+(1-\alpha/R)^{-1}dR^2$$
$$+R^2(d\vartheta^2+\sin^2\vartheta\,d\phi^2).$$

(B) $\zeta=0$:

$$d\hat{S}^2=-\alpha^{-1}\Omega\,d\Omega^2+\alpha\Omega^{-1}dr^2+\Omega^2(d\vartheta^2+d\phi^2).$$

There are still two sub-subcases to be distinguished.

(a) $\alpha>0$: setting $T^{1/2}=(4\alpha)^{-1/3}\Omega$, $Z=(\alpha/2)^{1/3}r$, $X=(4\alpha)^{1/3}\vartheta$, $Y=(4\alpha)^{1/3}\phi$, one obtains

$$d\hat{S}^2 = T^{-1/2}(-dT^2 + dZ^2) + T(dX^2 + dY^2),$$

which is the Kasner metric.

(b) $\alpha < 0$: setting $Z^{1/2} = (-4\alpha)^{-1/3}\Omega$, $T = (-\alpha/2)^{1/3}r$, $X = (-4\alpha)^{1/3}\vartheta$, $Y = (-4\alpha)^{1/3}\phi$, one obtains

$$d\hat{S}^2 = Z^{-1/2}(-dT^2 + dZ^2) + Z(dX^2 + dY^2),$$

which is the Taub metric.

Subcase (3): $R_{tt}H^{-1} = \zeta$ somewhere while $R_{tt}H^{-1} \neq \zeta$ elsewhere.

Let p be a point where $R_{tt}H^{-1} \neq \zeta$. The required smoothness of the metric ensures that there exists a neighborhood U of p such that $R_{tt}H^{-1} \neq \zeta$ in U. Thus dS_H^2 is Schwarzschild (or Taub–Kasner) in U. The smoothness of the metric then requires that dS_H^2 be Schwarzschild with the same mass (or Taub–Kasner) everywhere.

Combining the results of subcases (1)–(3) we arrive at the following theorem.

Theorem: Any conformally spherically (resp. plane-) symmetric solution to the vacuum Einstein equations must be the Schwarzschild (resp. either Taub–Kasner or flat) metric.

A direct corollary of this theorem is that any vacuum solution $\hat{g}_{ab}$ conformal to the Schwarzschild (resp. Taub–Kasner) solution g_{ab} must be the Schwarzschild (resp. Taub–Kasner) solution, in accordance with Brinkmann's theorems, which applied to this case imply that the conformal factor must be a constant and hence $\hat{g}_{ab}$ Schwarzschild (or Taub–Kasner).

As a final remark, we note that once the result $\Omega^{-1} = L(t)$ is obtained, it follows that $d\hat{S}^2$ is spherically (resp. plane-) symmetric and hence, by virtue of the Birkhoff (resp. Taub) theorem, must be the Schwarzschild (resp. Taub–Kasner or flat) solution. We prefer, however, to prove this conclusion along the line of the present paper for completeness and for offering a somehow different (conformal transformation based) proof of these two theorems.

ACKNOWLEDGMENTS

We would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Center for Theoretical Physics, Trieste, where this work was done. We would also like to thank Dr. O. Moreschi for drawing our attention to Brinkmann's work.

APPENDIX: PROOF OF TWO ASSERTIONS

We now prove the following assertions.

(a) Any nonflat H-type metric dS_H^2 with $\zeta \neq 1$ is not an Einstein metric.

(b) Any H-type metric with $\zeta = 1$ is not an Einstein metric if it is conformal to a vacuum solution.

Since the components of the Ricci tensor and the metric tensor of dS_H^2 satisfy

$$\frac{R_{rr}}{g_{rr}} = \frac{R_{tt}}{g_{tt}} = \frac{1}{2H}\left(\frac{\partial^2 \ln|H|}{\partial t^2} - \frac{\partial^2 \ln|H|}{\partial r^2}\right),$$

$$\frac{R_{\vartheta\vartheta}}{g_{\vartheta\vartheta}} = \frac{R_{\phi\phi}}{g_{\phi\phi}} = 1 + \frac{(\zeta - 1)}{(1 - \zeta\cos^2\vartheta)^2},$$

dS_H^2 is an Einstein metric if and only if

$$1 + \frac{(\zeta - 1)}{(1 - \zeta\cos^2\vartheta)^2} = \frac{1}{2H}\left(\frac{\partial^2 \ln|H|}{\partial t^2} - \frac{\partial^2 \ln|H|}{\partial r^2}\right) = K, \tag{A1}$$

where K is a constant. If $\zeta \neq 1$, then $K \neq 1$ and (A1) leads to

$$(1 - \zeta\cos^2\vartheta)^2 = (\zeta - 1)/(K - 1),$$

which in turn implies a contradiction $\cos\vartheta = \text{const}$ unless $\zeta = 0$. In the case $\zeta = 0$, Eq. (A1) yields

$$\frac{\partial^2 \ln|H|}{\partial t^2} - \frac{\partial^2 \ln|H|}{\partial r^2} = 0,$$

which implies that dS_{HP}^2 is flat.

We now take up the case $\zeta = 1$. Suppose dS_{HS}^2 is conformal to a vacuum solution $d\hat{S}^2$,

$$d\hat{S}^2 = \Omega^2\, dS_{HS}^2,$$

then, as shown in the text, in a suitable coordinate system we have $L = L(t)$ and $H = dL/dt$. If dS_{HS}^2 is an Einstein metric, then Eq. (A1) holds and it reduces to

$$\frac{d^2 \ln|dL/dt|}{dt^2} = 2\frac{dL}{dt},$$

hence

$$\frac{d^2L}{dt^2} = \frac{2LdL}{dt} + A\frac{dL}{dt}, \quad A \text{ const}, \tag{A2}$$

and

$$\frac{dL}{dt} = L^2 + AL + B, \quad B \text{ const}. \tag{A3}$$

Since

$$\frac{R_{tt}}{H} = -\frac{R_{tt}}{g_{tt}} = -\frac{R_{\phi\phi}}{g_{\phi\phi}} = -1 \neq \zeta,$$

Eq. (9′) in the text holds and takes the form

$$L\frac{d^2L}{dt^2} - 3\left(\frac{dL}{dt}\right)^2 - L^2\frac{dL}{dt} = 0. \tag{A4}$$

Substitution of Eqs. (A2) and (A3) into (A4) then yields $L = \text{const}$, a contradiction.

[1] A. H. Taub, Ann. Math. **53**, 472 (1951).

[2] D. Kramer, H. Stephani, E. Herlt, and M. MacCallum, *Exact Solutions of Einstein's Field Equations* (Cambridge U.P., Cambridge, 1980).

[3] E. Kasner, Am. J. Math. **43**, 217 (1921).

[4] H. W. Brinkmann, Math. Ann. **94**, 119 (1925).

[5] J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).

[6] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U.P., Cambridge, 1973).

Solutions of Einstein's equations relevant to the description of "bubbles" in the early universe

J. Ponce de Leon[a)]
Department of Physics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

(Received 2 March 1988; accepted for publication 15 June 1988)

A new class of spherically symmetric solutions of the Einstein field equations is presented. Their main features are that (i) the azimuthal metric coefficient depends on time only, and (ii) they possess similarity symmetry. The physical motivation for the study of such a class of solutions is that according to recent investigations [R. N. Henriksen, A. G. Emslie, and P. S. Wesson, Phys. Rev. D **27**, 1219 (1983); P. S. Wesson, Phys. Rev. D **34**, 3925 (1986)], they can be relevant to "bubbles" of new phases, in phase transitions typical of inflationary universe models. The solutions have shear, are inhomogeneous, and may be interpreted as "mixtures" of perfect fluids. They have some adjustable parameters which can be used to assure the fulfillment of the energy conditions. There are two different types of solutions. One of them has similarity symmetry of the first kind and negative total pressure. These models can be used in a classical description of particle production phases in the early universe. The other type of solution has similarity of the second kind, i.e., it represents models with dimensional constraints. Explicit solutions representing mixtures of fluids with equations of state $p = n\rho$ and $\rho = p$ of this type are given. They may be useful for cosmological models in closed universes. The dimensional constraints are found to be due to the "boundary conditions" in such universes. The specific characteristics of both types of solutions suggest that a transition from a particle-production phase to a radiation-dominated era can be described by means of bubbles of a "broken-symmetry" phase with positive pressure growing into a region of "unbroken-symmetry" phase with negative pressure.

I. INTRODUCTION

In recent years there has been considerable interest in cosmological models capable of describing both the particle creation in the early universe, predicted by some quantum field theories,[1] and the phase transitions typical of inflationary universe models.[2,3] In this paper I present a class of exact solutions of the Einstein field equations relevant to the particle creation as well as to the description of phase changes.

Before discussing the specific details of the solutions I would first like to summarize briefly some previous results in the literature that motivate this work.

Some time ago Henriksen, Emslie, and Wesson[4] found a class of solutions of the Einstein field equations with nonvanishing cosmological constant that describes the outward motion of a spherical disturbance in an asymptotically homogeneous and isotropic universe. Their solutions possess similarity symmetry of the second kind and have the remarkable features that (a) the matter pressure is negative, (b) the disturbance grows in an exponentially expanding universe, and (c) the metric in the region of the disturbance reduces to one belonging to the Kantowski–Sachs[5] class of cosmological models. The first feature means that the solutions can describe particle creation in the early history of the universe. The other two features are of particular relevance to the present work, since they suggest that the "bubbles" associated with the phase transitions of inflationary models can be described by solutions of the Einstein field equations, which are of the "Kantowski–Sachs type."

In a recent paper Wesson[6] showed that the equations describing the above-mentioned disturbance have the feature that they generate (besides the solution by Henriksen, Emslie, and Wesson) a new class of solutions that (a) also have similarity symmetry but of the first kind, and (b) represent an extension of the Kantowski–Sachs models to the case of finite negative pressure. He pointed out that such properties of the solutions may be exploited to describe phase transitions involving the disappearance of the cosmological constant. In this way, this new class of solutions gives some additional stimulus to the viewpoint that such solutions of the Kantowski–Sachs type can be relevant to phase changes.

Motivated by the above discussion I study here, in some detail, the class of perfect-fluid solutions of the Einstein field equations, which (i) as in the Kantowski–Sachs models, have an azimuthal metric coefficient that depends on time only, and (ii) possess similarity symmetry. It should be noted that neither the kind of symmetry nor the equation of state is *a priori* restricted to having some specific form. My primary interest in this work is to obtain the exact analytic form as well as the general properties of the models. We will see that the conditions (i) and (ii) lead to two types of shear-anisotropic solutions, and that each of them has different physical and mathematical properties. Both types of solutions depend on certain constants, which can be chosen in such a way as to satisfy the usual conditions imposed upon the energy-momentum tensor, viz., the "weak," "dominant," and "strong" energy conditions.[7] Solutions of one of the types continually expand and have negative pressure; so they can be used to describe particle-creation phases in the

---

[a)] Permanent address: Apartado 2816, Caracas 1010-A, Venezuela.

0022-2488/88/112479-07$02.50

early universe. This type of model has homothetic symmetry and generalizes a family of solutions recently discussed by Wesson.[6] The other type contains similarity solutions with positive pressure. They have the interesting property that, contrary to what happens in the homothetic case,[8–12] the dimensionless quantities in the theory are not scale-free, but they depend on a dimensional parameter. As a consequence of the dimensional constraints the simple symmetry under the "scaling" group is broken, and the solutions of this type admit a homothetic Killing vector in the hypersurfaces $t = \text{const}$ only. The solutions can be useful for cosmological models in closed universes because the additional dimensional parameter may be related to the existence of "boundaries" in such universes.

I will show that the models can be interpreted as mixtures of perfect fluids. We will also discuss the possibility of transitions from an "unbroken-symmetry" phase (i.e., with homothetic symmetry) to a "broken-symmetry" phase (i.e., with partial homothetic symmetry only). The paper is organized as follows. The model as well as the conventions used are described in Sec. II. In Sec. III we obtain the general solutions and display some particular cases. In Sec. IV the "multifluid" interpretation of the solutions is given. Section V is a summary and discussion.

II. THE MODEL

The assumption of spherical symmetry implies that there are coordinates (t,r,θ,ϕ) in which the line element takes the form

$$ds^2 = c^2 e^{\nu(r,t)}\, dt^2 - e^{\lambda(r,t)}\, dr^2 - R^2(r,t) d\Omega^2 , \tag{1}$$

where

$$d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2 ,$$

c is the velocity of light, and ν, λ, and R are, in general, arbitrary functions of r and t.

As is known, the Kantowski–Sachs metric is obtained from Eq. (1) by assuming that the space-time admits, besides the three Killing vectors related to the spherical symmetry, an additional Killing vector along the r direction. This assumption requires that $\nu' = \lambda' = R' = 0$. Then rescaling the time and denoting the coefficients of dr and $d\Omega$ by $X^2(t)$ and $Y^2(t)$, respectively, the Kantowski–Sachs metric takes the form

$$ds^2 = c^2\, dt^2 - X^2(t) dr^2 - Y^2(t) d\Omega^2 , \tag{2}$$

from which it follows that

$$\underset{\xi^{\mu}_{(4)}}{L}\, g_{\mu\nu} = 0 , \tag{3}$$

for $\xi^{\mu}_{(4)} = \text{const} \times \delta^{\mu}_{1}$.

An interesting property of the metric (2) is that the azimuthal metric coefficient Y depends on t only. As discussed in Sec. I, this property of the Kantowski–Sachs models may be relevant to the description of phase transitions. Thus the metrics we consider are of the following form:

$$ds^2 = c^2 e^{\nu(r,t)}\, dt^2 - e^{\lambda(r,t)}\, dr^2 - R^2(t) d\Omega^2 . \tag{4}$$

It is obvious that these metrics differ from that of Kantowski–Sachs in that they do not admit the spacelike Killing vector given by Eq. (3).

I shall further assume that the coordinate system in Eq. (4) is comoving with the matter. This means that $T^0{}_1 = T^1{}_0 = 0$. Then it follows from the Einstein field equations that ν is a function of t only. Redefining the time t one can put $\nu = 0$, and the remaining field equations for a perfect fluid reduce to

$$\frac{8\pi G\rho}{c^4} = \frac{1}{R^2}\left\{1 + \frac{(\dot R^2 + \dot\lambda \dot R R)}{c^2}\right\} , \tag{5}$$

$$\frac{8\pi G p}{c^4} = -\frac{1}{R^2}\left\{1 + \frac{(\dot R^2 + 2\ddot R R)}{c^2}\right\} , \tag{6}$$

$$\frac{8\pi G p}{c^4} = -\frac{1}{c^2 R^2}\left\{\frac{\dot\lambda \dot R R}{2} + \frac{\ddot\lambda R^2}{2} + \frac{\dot\lambda^2 R^2}{4} + \ddot R R\right\} , \tag{7}$$

where ρ is the energy density, p is the pressure, and a dot denotes differentiation with respect to t. It should be noted at this point that a class of solutions to Eqs. (5)–(7) has recently been discussed, in another context, by Wesson.[13]

Equations (5) and (6) show that, in the model under consideration, the pressure depends only on time, and that the energy density is not spatially homogeneous but depends also on the radial coordinate. Moreover, it is easy to see from Eq. (4) that the expansion and shear of the fluid are also functions of t and r.

The second feature that may be relevant to phase transitions is that the bubbles of new phases can be described by solutions with similarity symmetry. Thus we assume that the metric coefficients are essentially functions of the single dimensionless variable ξ defined as

$$\xi \equiv F(t)/Q(r) , \tag{8}$$

where, for generality, F and Q are two unknown functions of their arguments.

Thus we may write

$$R^2 = r^2 S(\xi), \quad \lambda = \lambda(\xi) . \tag{9}$$

Since $R' = 0$, it follows from Eq. (9) that $Q(r) = Cr^{2/\omega}$ and $S(\xi) = \overline{C}\xi^{\omega}$, where C, $\overline{C}$, and ω are constants of integration. Redefining both the coordinate r and the similarity variable ξ we can put, without loss of generality, $R^2 = r^2\xi$ and $Q(r) = r^2$. (This redefinition leaves the comoving nature of the coordinate system unchanged, and therefore the pressure and density remain invariant.) The corresponding field equations with similarity are then as follows:

$$\eta \equiv \frac{8\pi G r^2}{c^4}\rho = \frac{1}{\xi}\left\{1 + \frac{\dot F^2}{c^2 F}\left(\frac{1}{4} + \frac{\xi\lambda_{,\xi}}{2}\right)\right\} , \tag{10}$$

$$P \equiv \frac{8\pi G r^2}{c^4} p = \frac{1}{\xi}\left\{-1 + \frac{\dot F^2}{c^2 F}\left(\frac{1}{4} - \frac{\ddot F F}{\dot F^2}\right)\right\} , \tag{11}$$

$$\frac{4}{F} + \frac{2\ddot F}{c^2 F} = [3\xi\lambda_{,\xi} + 2\xi^2\lambda_{,\xi\xi} + (\xi\lambda_{,\xi})^2]\frac{\dot F^2}{c^2 F^2} + \frac{2\xi\lambda_{,\xi}}{c^2}\left(\frac{\dot F}{F}\right)^{\bullet} , \tag{12}$$

where η and P are dimensionless quantities, and the subscript ξ following a comma denotes differentiation with respect to ξ. Moreover, the equation specifying the bubble may be written as

$$\xi = \xi_1 ,$$

where ξ_1 is a constant. (In the model of Henriksen, Emslie, and Wesson $\xi_1 = 1$. For details see Ref. 4.) In general the space-time region of the new phase will be given by

$$0<\xi_0<\xi<\xi_1 ,$$

where ξ_0 is another constant. Then the full picture describing the motion of the bubble is obtained by fitting a solution of Eqs. (10)–(12) to another known cosmological solution of the Einstein field equations across a hypersurface $\xi = \text{const}$. (The technique of fitting a similarity solution to another solution of the field equations across such hypersurfaces has been treated by Cahill and Taub,[8] and we will not discuss it here.)

III. THE SOLUTIONS

We will see in this section that (a) the above assumptions entirely define two different families of solutions of the field equations, and (b) both families have reasonable physical properties.

Inspection of Eq. (12) shows that the assumed symmetry is maintained in two cases only. The first case is when

$$(\dot F/F)^{\cdot} = k(\dot F/F)^2 , \tag{13}$$

where k is a dimensionless constant. The second case arises when

$$\xi\lambda_{,\xi} = a = \text{const.} \tag{14}$$

We now proceed to discuss these cases separately.

A. Case I

The integration of Eq. (13) gives

$$F = B\,[kt + A\,]^{-1/k} , \tag{15}$$

where A and B are constants of integration. Substituting this expression into Eq. (12) we obtain an equation that separates only if

$$k = -\tfrac{1}{2}, \tag{16a}$$

$$2\xi^2\lambda_{,\xi\xi} + (\xi\lambda_{,\xi})^2 + 2\xi\lambda_{,\xi} = 1 + 4c^2/B . \tag{16b}$$

To integrate the last equation we introduce two new variables u and y,

$$e^{\lambda} = u^2, \quad \xi = e^{y} , \tag{17}$$

in terms of which Eq. (16b) becomes

$$4\frac{d^2u}{dy^2} = a^2 u , \tag{18}$$

where $a^2 \equiv 1 + 4c^2/B$. The general solution of this equation is given by

$$u = Ce^{ay/2} + De^{-ay/2} , \tag{19}$$

where C and D are constants of integration. Finally, using Eq. (17) we get

$$e^{\lambda} = [C\xi^{a/2} + D\xi^{-a/2}]^2 . \tag{20}$$

From Eqs. (10), (11), (15), (16a), and (20) we obtain the explicit form of the solution as follows:

$$\eta = \frac{a}{(a^2-1)\xi}\,\frac{[(a+2)C\xi^a + (a-2)D\,]}{[C\xi^a + D\,]} , \tag{21}$$

$$P = -a^2/(a^2-1)\xi , \tag{22}$$

$$ds^2 = c^2\,dt^2 - [C\xi^{a/2} + D\xi^{-a/2}]^2\,dr^2 - r^2\xi\,d\Omega^2 , \tag{23}$$

where the similarity variable ξ is

$$\xi = c^2(t-t_0)^2/(a^2-1)r^2 , \tag{24}$$

where we have redefined the constant A, viz., $A \equiv t_0/2$. The generator of the corresponding Lie group symmetry is given by

$$\xi^{\mu} = (t - t_0, r, 0, 0) , \tag{25}$$

and the Lie derivative of the metric is

$$\underset{\xi^{\mu}}{L}\; g_{\mu\nu} = 2g_{\mu\nu} . \tag{26}$$

Thus the solution (21)–(24) admits a homothetic Killing vector.

The above solution depends on four parameters, viz., t_0, a, C, and D, which can be chosen in such a way as to assure the positiveness of ρ and $(-g_{22})$ as well as the fulfillment of other physical restrictions usually imposed upon density and pressure, viz., the dominant and the strong energy conditions.

It is seen from Eq. (22) that for such values of the parameters the pressure is always negative. At this point it is worthwhile to recall that nonequilibrium states with negative pressures can exist in what are called metastable states. For example, a superheated liquid may have negative pressure.[14] In general relativity, solutions of the Einstein field equations with negative pressure have recently been discussed[4,6,15] in connection with the classical description of the particle-production phases, in the early universe, predicted by some symmetry-breaking particle theories.[1] In particular, if we put

$$(t - t_0)/\sqrt{a^2-1} = \bar t$$

and $D = 0$ in Eqs. (21)–(24), then our solution reduces to a solution recently obtained by Wesson.[6]

To conclude the discussion of case I it should be noted that $k \neq 0$ in Eq. (13). In fact, for $k = 0$ this equation has no solutions compatible with both the perfect fluidity condition and the similarity condition. What this means is that our solutions might be relevant to models of bubbles whose rate of expansion is not exponential but rather relatively slow, viz., $R \propto t$. It has been argued[16,17] that such kinds of bubbles in rapidly expanding universes may lead to the solution of some problems in cosmology.

B. Case II

From Eq. (14) we find

$$e^{\lambda} = C\xi^{a} . \tag{27}$$

From this equation it could be thought that this case follows from case I [see Eq. (23)] with $D = 0$. However, in general this is not so, since in the case under consideration Eqs. (12) and (27) define a more general class of functions F than in case I. In fact, from (12) and (27) we get

$$F\ddot F + (a/2)\dot F^2 + 2c^2F/(1-a) = 0, \quad a \neq 1 . \tag{28}$$

The first integral of this equation is given by

$$\dot{F}^2 = c^2[4F + F^{-a}(r_0)^{2(a+1)}]/(a^2-1)\,, \qquad (29)$$

where r_0 is a constant of integration with the dimensions of length. For $r_0 = 0$ we recover the solutions of case I, given by Eq. (23), with $D = 0$. For $r_0 \neq 0$, however, we obtain new solutions.

Substituting Eq. (29) into (10) and (11) we obtain the final form of the solution as follows:

$$\eta = \frac{a(a+2)}{(a^2-1)\xi} + \frac{(2a+1)}{4(a^2-1)\xi^{a+2}}\left(\frac{r_0}{r}\right)^{2(a+1)}\,, \qquad (30)$$

$$P = -\frac{a^2}{(a^2-1)\xi} + \frac{(2a+1)}{4(a^2-1)\xi^{a+2}}\left(\frac{r_0}{r}\right)^{2(a+1)}\,, \qquad (31)$$

and

$$ds^2 = c^2\,dt^2 - C\xi^a\,dr^2 - r^2\xi\,d\Omega^2\,. \qquad (32)$$

We see that, in general, the dimensionless quantities η and P depend not only on the similarity variable ξ but also on r; in other words, they are not scale-free. In fact, only for $r_0 = 0$ or $a = -\frac{1}{2}$ we have $\eta = \eta(\xi)$ and $P = P(\xi)$, and also only for $r_0 = 0$ the distribution is invariant under the scaling group. As a result of the scale restrictions the homothetic symmetry holds on the spacelike hypersurfaces orthogonal to t only, namely,

$$\underset{\xi^{\mu}_{(1)}}{L}\, h_{\mu\nu} = 2h_{\mu\nu}\,, \qquad (33)$$

where $h_{\mu\nu} \equiv g_{\mu\nu} - U_\mu U_\nu$ and

$$\xi^{\mu}_{(1)} = (2(F/\dot{F}), r, 0, 0)\,. \qquad (34)$$

This means that as one moves along the orbits of $\xi^{\mu}_{(1)}$, only the spatial lengths increase at the same rate. The r-t hypersurfaces, defined by $d\theta = d\phi = 0$, admit the conformal Killing vector

$$\xi^{\mu}_{(2)} = (\alpha F^{a/2}, \beta r^a, 0, 0)\,, \qquad (35)$$

where α and β are constants, namely,

$$\underset{\xi^{\mu}_{(2)}}{L}\, g_{\mu\nu} = \Psi g_{\mu\nu}\,, \qquad (36)$$

with $\Psi = \alpha a F^{(a-2)/2}\dot{F}$. Thus the solutions (30)–(32) are partially self-similar and partially conformal in the sense of Tomita.[18–20]

In the present case the constants can be chosen to satisfy the standard energy conditions, namely, $\rho \geqslant 3p \geqslant 0$ or, if it is desired, $\rho \geqslant p \geqslant 0$. However, in general the function F from Eq. (29) is given in terms of elliptic integrals. We shall show here two simple particular solutions arising from (29)–(32).

Let us first consider the case $a = -\frac{1}{2}$. With the transformation

$$R = \sqrt{F}\,, \qquad (37)$$

Eq. (29) becomes

$$\dot{R}^2/c^2 = -\tfrac{4}{3} - r_0/3R\,, \qquad (38)$$

from which it follows that r_0 must be negative. Putting $r_0 = -3R_0$ ($R_0 > 0$), the solution of this equation, in parametric form, is given by

$$R \equiv \sqrt{F} = (3R_0/8)[1-\cos\tau],$$

$$\pm c(t-t_0) = (3\sqrt{3}R_0/16)[\tau - \sin\tau]\,, \qquad (39)$$

where t_0 is a constant of integration. The equation of state corresponding to this solution is that of radiation, viz., $\rho = 3p > 0$. Moreover, η and P are given by

$$\eta = 3P = 1/\xi\,, \qquad (40)$$

$$\xi = \tfrac{9}{64}(R_0/r)^2[1-\cos\tau]^2\,. \qquad (41)$$

Another interesting solution arises for $a = 0$. In this case[21]

$$\eta = P = (R_0/r)^2/\xi^2\,, \qquad (42)$$

where

$$\xi = [R_0{}^2 - c^2(t-t_0)^2]/r^2\,, \qquad (43)$$

and R_0 and t_0 are constants of integration. In this latter case the equation of state is the stiff equation $\rho = p > 0$.

Thus the above examples illustrate the fact that regardless of the likeness between Eqs. (23) and (32), the solutions of case II differ from those of case I not only in the choice of the similarity variable ξ, but also in their physical properties.

IV. PROPERTIES OF THE SOLUTIONS

Thus we have succeeded in finding all the similarity solutions under the assumptions of Sec. II. Important features of the solutions are their simplicity and the fact that they contain a number of parameters [viz., (a,C,D) in case I, and (a,r_0) in case II], which may be used to satisfy physical requirements.

In this section I will first discuss the relation between our model and other models in the literature, and then I will show that the source of the gravitational field, associated with our solutions, may be represented as a noninteracting mixture of perfect fluids having different equations of state.

A. Case I

We immediately see that if neither C nor D is equal to zero, the r dependence of g_{rr} in the metric (23) cannot be eliminated by means of coordinate transformations. Therefore, in general, the energy density ρ, as obtained from Eq. (21), is a function of t and r. Consequently, in general, Eqs. (21)–(24) represent a family of solutions that not only have nonvanishing shear but also are spatially inhomogeneous.

The distribution admits an "equation of state" $\eta = \eta(P)$, which is given in parametric form by Eqs. (21) and (22). (Notice that this equation relates the dimensionless quantities η and P rather than the density and pressure.) Since the distribution is inhomogeneous, the relation between ρ and p is more complicated because it should also include r, viz., $\rho = \rho(p,r)$.

In general $\eta = \eta(P)$ is not a linear function, but in the case where either C or D is equal to zero it reduces to the well-known linear equation of state $\eta = \alpha P$, so that $\rho = \alpha p$, where α is a constant. In such a case the energy density and pressure are functions of t only, and the r dependence in the metric can be eliminated by a simple rescaling of r.

Thus if either C or D is equal to zero, our inhomogeneous solutions (21)–(24) reduce to homogeneous solutions of the Kantowski–Sachs (KS) type with $\rho = \alpha p$. It was already mentioned that a specific family of KS solutions of this kind has recently been discovered by Wesson.[6] Therefore the solutions of case I with $C \neq 0$ and $D \neq 0$ generalize Wesson's solutions not only in a pure mathematical sense but also in their physical content (for further discussion see Sec. V).

Now I would like to point out that the energy density and pressure associated with Eqs. (21) and (22) can be separated into two components, namely,

$$\rho = \rho_v + \rho_{\text{dust}}, \quad p = p_v + p_{\text{dust}}, \tag{44}$$

where ρ_v and p_v represent, respectively, the positive energy density and the pressure of a perfect fluid that obey the equation of state of "false vacuum," viz.,[22]

$$\rho_v = -p_v, \tag{45}$$

with

$$\rho_v = (c^2/8\pi G)a^2/t^2, \tag{46}$$

where we have set $t_0 = 0$. Moreover,

$$\rho_{\text{dust}} = \frac{ac^2}{4\pi G t^2}\left[\frac{\overline{C}t^{2a} - \overline{D}r^{2a}}{\overline{C}t^{2a} + \overline{D}r^{2a}}\right], \qquad p_{\text{dust}} = 0, \tag{47}$$

where $\overline{C} = C(c)^{2a}$ and $\overline{D} = D(a^2-1)^a$.

Thus we conclude that the homothetic solutions of case I may be considered as describing a noninteracting mixture of (inhomogeneous) "dust" ($p = 0$) and a "vacuum fluid" ($\rho = -p$) whose energy density varies with time. I believe these are the first known solutions of this kind in the literature.

At this point it is worthwhile to recall that ρ_v is allowed to vary, for example, if it is regarded as the energy-momentum tensor ($T_{\mu\nu} = \rho_v g_{\mu\nu}$), associated with some scalar field, as in many inflationary universe scenarios.

B. Case II

In this case the distribution is homogeneous, and a suitable redefinition of the r coordinate renders the metric (32) in the Kantowski–Sachs form. Now contrary to what happened in case I, a simple relation between η and P, in general, does not exist, but it also contains r. However, for each specific value of the constants a and r_0 there is an equation of state $\rho = \rho\ (p)$ whose parametric form is obtained from Eqs. (30) and (31). For arbitrary values of a, the shape of this function is complicated, but what is important is that the ratio p/ρ, in general, varies with time. Consequently, without loss of generality, we can interpret the solutions of case II (for arbitrary values of a) as representing a single homogeneous perfect fluid obeying the equation of state $p = n\rho$, where n, which is the velocity of sound squared, is really a function of t.

Another possible interpretation of case II arises from Eqs. (30) and (31) when the source is written as a mixture of two fluids, namely,

$$\rho = \bar{\rho} + \rho_{\text{stiff}}, \quad p = \bar{p} + p_{\text{stiff}}, \tag{48}$$

obeying the equations of state

$$\bar{p} = n\bar{\rho}, \quad 0 < n < 1, \tag{49}$$

and

$$p_{\text{stiff}} = \rho_{\text{stiff}}, \tag{50}$$

where

$$n = -a/(a+2) \tag{51}$$

and

$$\bar{p} = (c^4/8\pi G)a(a+2)/(a^2-1)F, \tag{52}$$

$$\rho_{\text{stiff}} = \frac{c^4}{32\pi G}\frac{2a+1}{a^2-1}(r_0)^a\left(\frac{r_0}{F}\right)^{a+2}. \tag{53}$$

Thus solutions of case II can be interpreted as two-perfect-fluid models composed of "stiff matter" ($\rho = p$) and a fluid obeying the equation of state $p = n\rho$. (Notice that for $-1 < a < 0, \bar{\rho}$ and $\bar{p}$ are positive and $0 < n < 1$.)

It is now clear that the null-fluid solution, given by Eqs. (37)–(41), as well as the stiff-matter solution, given by Eqs. (42) and (43), are the only two possible limiting cases where the mixture reduces to a single fluid because one of its ingredients vanishes identically. In fact for $a = -\frac{1}{2}$, by Eqs. (49)–(53), $\bar{\rho} = 3\bar{p}$ and $\rho_{\text{stiff}} = p_{\text{stiff}} = 0$. For $a = 0$, by Eqs. (51) and (52), $\bar{\rho} = \bar{p} = 0$, and the mixture reduces to a fluid with $\rho = p$ (recall that $r_0^2 = -4R_0^2$ [see Eqs. (42) and (43)]).

However, for any value of a district from $a = -\frac{1}{2}$ and $a = 0$, the solutions of case II have a much richer material content consisting of a mixture of two perfect fluids. As far as I know, these solutions are new in the literature.

For any given value of a, the integral of Eq. (29) [or (60)] is, in general, given either in parametric form or in terms of elliptic functions. However, it is easy to see that the global properties of the motion for $-1 < a < 0$ are essentially the same as those of the null fluid discussed in Sec. III.

To finish, I would like to point out the likeness between the multifluid interpretation of solutions I and II. To this end, notice that Eqs. (30) and (31) can be written as

$$\rho = \rho_v + \rho_{\text{dust}} + \rho_{\text{stiff}}, \qquad p = p_v + p_{\text{dust}} + p_{\text{stiff}}, \tag{54}$$

with

$$\rho_v = -p_v = (c^4/8\pi G)a^2/(a^2-1)F, \tag{55}$$

$$\rho_{\text{dust}} = (c^4/4\pi G)a/(a^2-1)F, \qquad p_{\text{dust}} = 0, \tag{56}$$

$$\rho_{\text{stiff}} = p_{\text{stiff}} = \frac{c^4}{32\pi G}\frac{(2a+1)}{(a^2-1)}(r_0)^a\left(\frac{r_0}{F}\right)^{a+2}. \tag{57}$$

These equations show that the solutions of case II may also be interpreted as a noninteracting mixture of a vacuum fluid, dust, and stiff matter.

If one uses $[(a^2-1)F]^{1/2}/c$ as a new time coordinate, the expressions for ρ_v and ρ_{dust}, given by Eqs. (55) and (56), become formally identical to Eqs. (46) and (47) with $D = 0$. Therefore, from Eqs. (44) and (54), one may conclude that the difference noted between solutions of cases I and II is essentially due to the additional stiff-matter component in case II. In particular, this explains the symmetry breaking that takes place in case II. (I shall return to this point in Sec. V.)

V. SUMMARY AND DISCUSSION

We have solved here the Einstein field equations with spherical symmetry and perfect fluid in the case where (i) $R' = 0$ and (ii) the metric coefficients are essentially functions of a dimensionless variable ξ only. We have seen that the solutions of case I admit a homothetic Killing vector, while in the case II the self-similarity property holds in the hypersurfaces $t = \text{const}$ only. Some comments regarding this symmetry are in order at this point.

Following the classical notion of similarity, the homothetic symmetry is usually expected to arise in systems without temporal or spatial characteristic scales. In such systems, by a suitable transformation of coordinates all dimensionless quantitites can be put in a form in which they are functions of $\xi = ct/r$ only.

In applications to cosmology this symmetry may be important in open universes[23] without cosmological constant. This is suggested by the fact that they contain no fundamental scales, have no boundaries, and expand continuously. More than this, even if such universes for some reason[24] had some characteristic scales at early times, it is still reasonable to expect that such universes will develop as scale-free asymptotically, i.e., at times and distances large compared with the initial scales. In other words, one can expect that as time proceeds, an open universe with $\Lambda = 0$ will "forget" any possible initial scale, due to its continuous expansion. Therefore self-similarity is a desirable property in theoretical models attempting to describe open universes.

However, the situation in closed universes is radically different because they have intrinsic scale restrictions even in the absence of a cosmological constant. In fact, in a finite period of time, such universes expand up to some maximum value and then recollapse to a singularity. Then the age of the universe (i.e., the time needed for a full cycle) as well as its maximum "radius" impose important scale restrictions in the form of dimensional boundary conditions. Therefore one should not expect the closed universes to be described by scale-invariant models. Instead, exact models of such universes should contain dimensional parameters that may be related to the boundary conditions.

Thus one expects that the physical difference between open and closed universes may lead to some differences in their mathematical description. The solutions of cases I and II clearly illustrate the above discussion. In fact, although similarity symmetry has been imposed at the outset, the physical and mathematical properties of the solutions obtained in cases I and II are not the same.

The solutions of case I are "open" solutions, i.e., they continuously expand[23] (or contract). This can be seen by a direct calculation of the expansion Θ, namely,

$$\Theta = \frac{\dot{\lambda}}{2} + 2\frac{\dot{R}}{R} = \frac{\dot{F}}{F}\left(\frac{\xi\lambda_{,\xi}}{2} + 1\right). \qquad (58)$$

Now substituting F and λ as given by Eqs. (23) and (24) and then using (21) and (22), we find

$$\Theta = (4\pi G/c^2)(t - t_0)(\rho + p) + 2/(t - t_0). \qquad (59)$$

Since $(\rho + p) > 0$, if $t - t_0 > 0$ (< 0), then the fluid expands (contracts) forever. Thus the solutions of case I have neither boundaries nor scale restrictions. Therefore they are consistent with the classical notion of similarity and admit a homothetic Killing vector.

The solutions of case II with $|a| > 1$ are also open solutions. In fact, Eq. (29) can be written as

$$4\dot{R}^2 = [c^2/(a^2 - 1)][4 + (r_0/R)^{2(a+1)}], \qquad (60)$$

where $R \equiv \sqrt{F}$. From this we see that for $|a| > 1$ and $r_0 > 0$, $\dot{R}$ can never become zero. Therefore if $\dot{R} > 0$ at some initial time, then R increases continuously as time increases. The characteristic length given by r_0 is introduced by a change in the sign of the total pressure at[25]

$$R = r_0[(2a + 1)/4a^2]^{1/2(a+1)}. \qquad (61)$$

Therefore, although such models have no boundaries, they cannot be described by exact self-similar solutions. However, for $R \gg r_0$ the second time in Eq. (60) is negligible, which means that the models develop as scale-free asymptotically. In that limit they admit the homothetic Killing vector given by Eqs. (25) and (26), as was discussed before.

The solutions of case II with $|a| < 1$ are "closed." This fact is illustrated by the null-fluid solution given by Eqs. (37)–(41). This universe emerges from a cigar singularity at $t = t_0$ and R increases from zero reaching its maximum value $R_{\max} = 3R_0/4$ within a finite time $(t - t_0) = (3\pi\sqrt{3}/16) \times R_0/c$. After this R again goes back to zero. Since the lifetime of such universes is finite, one cannot in a strict sense speak about the "asymptotic behavior" of such universes, i.e., they will never "forget" their boundary conditions (scale restrictions), as opposed to what happened in solutions of case II with $a > 1$ (see Ref. 26).

If $r_0 = 0$, the solutions of case II reduce to those with the usual symmetry under the scaling group. Specifically, they reduce to the subset of solutions of case I defined by $D = 0$.

If $r_0 \neq 0$ and $D = 0$, the solutions of both cases have the interesting property that their line elements take essentially the same form, viz.,

$$ds^2 = c^2\, dt^2 - \text{const} \times \xi^a\, dr^2 - r^2 \xi\, d\Omega^2. \qquad (62)$$

(However, it should be emphasized that the matter distribution in each case is different, since they involve different choices of ξ.) This line element shows that there is a notable likeness between the similarity solutions of type II and those of type I with $D = 0$. In this way it suggests that the symmetries studied here may represent successive phases of a matter distribution. Specifically, it suggests that a solution of case I (an unbroken-symmetry phase with

$$\underset{\xi}{L}\, g_{\mu\nu} = 2g_{\mu\nu})$$

with particle creation ($p < 0$) can undergo a phase transition in which the pressure becomes positive, and thereafter is described by a solution of case II (a broken-symmetry phase with

$$\underset{\xi}{L}\, g_{\mu\nu} \neq 2g_{\mu\nu}, \quad \underset{\xi}{L}\, h_{\mu\nu} = 2h_{\mu\nu}).$$

In fact, one could expect that a bubble of the broken-symmetry phase ($p > 0$) could nucleate in the symmetric phase ($p < 0$) due to some fluctuations. Then as a result of the difference of pressures such a bubble starts to grow, converting the symmetric phase of negative pressure and particle creation into a broken-symmetry phase of positive pressure, described by some of the cosmological solutions of case II. At this point it is worthwhile to recall that according to the inflationary universe models, the period of exponential expansion is followed by a period of reheating with a tremendous amount of particle production, which ends in a radiation-dominated era. Thus our solutions suggest a useful mechanism to overcome the problem of the return to a radiation-dominated era.

It should be noted that Wesson[6] recently discussed a similar mechanism for the disappearance of the cosmological constant. However, he considered only distributions with negative pressure, which are a subset of our solutions of case I. Consequently the solutions obtained in this work extend the previous discussion by Wesson[6] in two senses, namely, (a) to a more general class of solutions with negative pressure, and (b) to include possible transitions involving phases with positive pressure.

According to Tabensky and Taub,[27] an irrotational fluid with equation of state $p = \rho$ represents a source equivalent to a massless scalar field. Thus by Eq. (48) the solutions of case II can be interpreted as universes containing a scalar field and a perfect fluid obeying the equation of state $p = n\rho$. The symmetry breaking in these solutions is then explained as due to the nonvanishing values of the scalar field. Conversely, the symmetry is restored whenever the scalar field vanishes.

The solutions of case I as well as those of case II (with $r_0 \neq 0$) that have negative pressure can describe particle-creation phases in the early universe. The other solutions with positive pressure can be applied to the early universe after the epoch of creation has terminated, and can be relevant to theories of galaxy formation. In particular, the solutions with $-1 < a < 0$ can be useful for cosmological models of closed universes containing a scalar field minimally coupled to gravitation and a perfect fluid obeying the equation state $p = n\rho$. The rich matter content and the fact that these solutions have shear and admit a homothetic Killing vector, in the hypersurfaces $t = \text{const}$, as well as a conformal Killing vector in the subspace $d\theta = d\phi = 0$, may be regarded as an asset for the study of such universes.

ACKNOWLEDGMENTS

I would like to thank P. S. Wesson for the encouragement and many useful discussions.

This work was supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada (P. S. Wesson).

[1] R. Brout, F. Englert, and E. Gunzig, Ann. Phys. (NY) **115**, 78 (1978).

[2] A. H. Guth, Phys. Rev. D **23**, 347 (1981).

[3] A. H. Guth and E. J. Weinberg, Phys. Rev. D **23**, 876 (1981).

[4] R. N. Henriksen, A. G. Emslie, and P. S. Wesson, Phys. Rev. D **27**, 1219 (1983).

[5] R. Kantowski and R. K. Sachs, J. Math. Phys. **7**, 443 (1966).

[6] P. S. Wesson, Phys. Rev. D **34**, 3925 (1986).

[7] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge U.P., London, 1973), p. 85.

[8] M. E. Cahill and A. H. Taub, Commun. Math. Phys. **21**, 1 (1971).

[9] D. M. Eardley, Commun. Math. Phys. **37**, 287 (1974).

[10] G. V. Bicknell and R. N. Henriksen, Astrophys. J. **219**, 1043 (1978).

[11] R. H. Henriksen and P. S. Wesson, Astrophys. Space Sci. **53**, 429 (1978).

[12] G. V. Bicknell and R. N. Henriksen, Astrophys. J. **225**, 237 (1978).

[13] P. S. Wesson, J. Math. Phys. (to be published).

[14] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part 1, 3rd ed., revised and enlarged, Vol. 5 of *Course of Theoretical Physics* (Pergamon, London, 1980), pp. 44 and 260.

[15] P. S. Wesson, Astron. Astrophys. **151**, 276 (1985).

[16] A. H. Guth and S. H. Tye, Phys. Rev. Lett. **44**, 631, 963 (1980).

[17] J. Preskill, Phys. Rev. Lett. **43**, 1365 (1979).

[18] K. Tomita, Prog. Theor. Phys. **66**, 2055 (1981).

[19] K. Tomita, Prog. Theor. Phys. Suppl. **70**, 286 (1981).

[20] K. Tomita, Prog. Theor. Phys. Lett. **70**, 886 (1983).

[21] This solution is better obtained by solving Eq. (28) with $a = 0$ and then substituting the result directly into (10) and (11).

[22] Ya. B. Zel'dovich, Usp. Fiz. Nauk **95**, 209 (1968) [Sov. Phys. Usp. **11**, 381 (1968)].

[23] In the models under consideration the three-curvature scalar is positive for all times. However, there are both kind of solutions, viz., "open" and "closed," i.e., ever-expanding and recollapsing solutions, respectively.

[24] According to the inflationary scenario our universe in its very early history underwent a phase transition producing a vacuum-dominated inflationary era. This era lasted from the GUT time $t_G = \sqrt{3/\Lambda} = 1.0 \times 10^{-35}$ sec to $t = 1.3 \times 10^{-33}$ sec. Therefore this transition introduces some characteristic scales in the very early universe.

[25] According to the multifluid interpretation, given by Eqs. (54)–(57), this occurs at the value of R for which the "repulsion" produced by the negative pressure associated with the vacuum fluid is balanced by the "attraction" produced by the positive matter pressure.

[26] A null-fluid solution is expected to be useful in an early epoch of the universe only. But even in that limit ($\tau \ll 1$) the metric coefficients in (37)–(41) contain the dimensional parameter r_0.

[27] R. T. Tabensky and A. H. Taub, Commun. Math. Phys. **29**, 61 (1973).

Super-Selberg trace formula from the chaotic model

Shozo Uehara[a)]
Department of Mathematics, King's College, Strand, London WC2R 2LS, England
Yukinori Yasui
Department of Physics, Osaka City University, Osaka 558, Japan

(Received 8 April 1988; accepted for publication 13 July 1988)

A nonrelativistic superparticle moving freely on the super-Riemann surface of genus $g \geqslant 2$ is investigated. The classical motion is characterized by the super-Fuchsian group. The periodic orbits are unstable and the system is classically chaotic. Quantization is performed in the path integral method. The kernel function is explicitly given in the semiclassical approximation. The quantized energy spectrum is related to the length spectrum of the classical periodic orbits, which is a superanalog of Selberg's trace formula.

I. INTRODUCTION

Recently we investigated the system of a nonrelativistic superparticle moving freely on the "super"-Poincaré upper half plane $\mathrm{SH} \equiv \{(z,\theta) \,|\, \mathrm{Im}\, z > 0\}$.[1] Here we proceed to a superparticle on the super-Riemann surface (SRS) of genus $g \geqslant 2$, which is represented by SH/SΓ with super-Fuchsian group SΓ.[2] This is a supersymmetric extension of a particle on the Riemann surface. The latter is a prototype of a classical chaotic system[3] and it helps us to understand quantum chaos.[4,5] Energy is the only conserved quantity there and a quantized energy sum rule, which can be evaluated in the semiclassical approximation,[6] is nothing but Selberg's trace formula.[7] It plays an important role in bosonic string theory.[8] The supersymmetric extension should be of great interest. It will be a prototype of the classically chaotic supersymmetric system and it will also implement a geometric formulation of superstring. In this paper we briefly discuss the classical motion of a superparticle on SRS first and then we proceed to quantum mechanics via the path integral formulation. We give the kernel function explicitly in the semiclassical approximation, which leads to a superanalog of Selberg's trace formula.[9]

II. SUPERPARTICLE ON SH

We begin with a brief summary of Ref. 1. The Lagrangian of a superparticle on SH is given by the line element with the metric being a superconformally flat extension of the flat extension of the Poincaré metric:

$$L = \frac{m}{2}\left(\frac{ds}{dt}\right)^2 , \tag{1}$$

where

$$\begin{aligned} ds^2 &= dq^A g_{AB}\, dq^B \\ &= Y^{-2}\{dz\, d\bar{z} - i\theta\, dz\, d\bar{\theta} - i\bar{\theta}\, d\theta\, d\bar{z} \\ &\quad - (2Y + \theta\bar{\theta}) d\theta\, d\bar{\theta}\} , \\ &(q^z, q^{\bar{z}}, q^\theta, q^{\bar{\theta}}) = (z, \bar{z}, \theta, \bar{\theta}) , \end{aligned} \tag{2}$$

$$Y \equiv \mathrm{Im}\, z + \tfrac{1}{2}\theta\bar{\theta} . \tag{3}$$

Note that ds^2 can be written with the basis of the tangent space E^A's obtained from the flat ones by super-Weyl transformations[10] with the parameter Y^{-1} (Ref. 1),

$$ds^2 = E^z E^{\bar{z}} - 2E^\theta E^{\bar{\theta}} , \tag{4}$$

and the metric g_{AB} is "super"-Kähler.[1] The Lagrangian is invariant under the super-Möbius transformation SPL$(2,\mathbb{R})$,[2]

$$\begin{aligned} A(z) &= \frac{az + b}{cz + d} + \theta \frac{\alpha z + \beta}{(cz + d)^2} , \\ A(\theta) &= \frac{\alpha z + \beta + \theta(1 + \frac{1}{2}\beta\alpha)}{cz + d} , \quad ad - bc = 1 , \end{aligned} \tag{5}$$

where a, b, c, and d are real Grassmann even parameters and α and β are Grassmann odd ones with $\bar{\alpha} = i\alpha$, $\bar{\beta} = i\beta$. The quantum mechanical Hamiltonian H_Q is the "super"-Laplace–Beltrami operator,

$$H_Q = [(-)^a/2m] g^{-1/4} p_A g^{1/2} g^{AB} p_B g^{-1/4} , \tag{6}$$

where

$$[p_A, q^B\} = -i\hbar \delta_A{}^B , \tag{7}$$

$$g \equiv |\mathrm{sdet}\, g_{AB}| = 1/(4Y^2) , \tag{8}$$

$$[g^{AB}]: \quad \text{the inverse to } [g_{AB}] , \tag{9}$$

and the sign factor $(-)^a$ means that $+1$ for A (or a) $= z, \bar{z}$ or -1 for $A = \theta, \bar{\theta}$.

The Euler–Lagrange equations from L in (1) are

$$\ddot{z} + i\frac{\dot{z}^2}{Y} + \frac{\bar{\theta}\dot{z}\dot{\theta}}{Y} = 0 , \quad \ddot{\theta} + i\frac{\dot{z}\dot{\theta}}{Y} = 0 , \tag{10}$$

and the complex conjugated ones.[11] The solutions are given by

$$\begin{aligned} &z(t) = c_1\{\tanh \omega(t + t_0) + i\, \mathrm{sech}\, \omega(t + t_0)\} + c_2 , \\ &\theta(t) = \varepsilon_1 z(t) + \varepsilon_2 , \quad \bar{z}(t) = \overline{z(t)} , \quad \bar{\theta}(t) = \overline{\theta(t)} , \end{aligned} \tag{11}$$

where the constants of integration c_1, c_2 ω, and t_0 are Grassmann even real numbers, while ε_1 and ε_2 are Grassmann odd ones with $\bar{\varepsilon}_1 = ie^{i\phi}\varepsilon_1$, $\bar{\varepsilon}_2 = ie^{i\phi}\varepsilon_2$, where ϕ is a Grassmann even real one (the imaginary unit "i" is attached for convenience). The Grassmann even part of the solutions, i.e., $z(t)$,

---

[a)] On leave of absence from Research Institute for Theoretical Physics, Hiroshima University, Takehara, Hiroshima 725, Japan.

0022-2488/88/112486-05$02.50

$\bar{z}(t)$, take the same form as the solutions in the purely bosonic case.[3] As a result of the (nonzero) phase factor ϕ the supersymmetries are broken in (11).[1] Actually there are other solutions;[1] however, they do not yield periodic orbits on SRS. Using these solutions in (11) we define the "super"-hyperbolic distance between two points $q_i = q(t_i)$ $(i = 1,2)$,

$$d(q_1,q_2) = \int_{t_1}^{t_2} \sqrt{\left(\frac{ds}{dt}\right)^2}\, dt\,, \tag{12}$$

which becomes $\omega(t_2 - t_1)$. This is rewritten by

$$\cosh[d(q_1,q_2)] = 1 + \tfrac{1}{2}R(q_1,q_2) - 2r(q_1,q_2)\,, \tag{13}$$

where[9]

$$R(q_1,q_2) = \frac{|z_1 - z_2 - \theta_1\theta_2|^2}{Y_{(1)}Y_{(2)}}\,, \tag{14}$$

$$\begin{aligned} r(q_1,q_2) = {} & \frac{2\theta_1\bar{\theta}_1 + i(\theta_2 - i\bar{\theta}_2)(\theta_1 + i\bar{\theta}_1)}{4Y_{(1)}} \\ & + \frac{2\theta_2\bar{\theta}_2 + i(\theta_1 - i\bar{\theta}_1)(\theta_2 + i\bar{\theta}_2)}{4Y_{(2)}} \\ & + \frac{(\theta_2 + i\bar{\theta}_2)(\theta_1 + i\bar{\theta}_1)\mathrm{Re}(z_1 - z_2 - \theta_1\theta_2)}{4Y_{(1)}Y_{(2)}}\,. \end{aligned} \tag{15}$$

All these two-point quantities d, R, and r enjoy the following property:

$$\begin{aligned} &\cdot(A(q_1),A(q_2)) = \cdot(q_1,q_2) = \cdot(q_2,q_1), \quad A\in\mathrm{SPL}(2,\mathbf{R})\,. \\ &\cdot\,\cdot \end{aligned} \tag{16}$$

III. CLASSICAL MOTION ON SRS

Now we examine the classical motion on SRS.[12] Since the Lagrangian (1) [Hamiltonian (16)] is invariant under SPL(2,R), it can be regarded as the one of the superparticles on SRS. Hence we can deduce the classical motion on SRS from that on SH by taking the quotient by SΓ. First we notice that every element γ $(\neq 1)$ of SΓ has two fixed points, (u,μ) and (v,ν), where $\mathrm{Im}\, u = \mathrm{Im}\, v = 0$, $\bar{\mu} = i\mu$, $\bar{\nu} = i\nu$, and γ and $\gamma'\in$SΓ have the same fixed points iff there exists such $\gamma_0\in$SΓ such that $\gamma = \gamma_0^n$, $\gamma' = \gamma_0^m$, $n,m\in Z$. On the other hand, there exists a unique geodesic curve or classical path that connects those fixed points. In fact, the classical solutions with the constants of integration

$$\begin{aligned} c_1 &= \frac{v-u}{2}\,, & c_2 &= \frac{u+v}{2}\,, \\ \varepsilon_1 &= \frac{\mu - \nu}{u - v}\,, & \varepsilon_2 &= \frac{u\nu - v\mu}{u - v} \end{aligned} \tag{17}$$

define the geodesic curve connecting two points $(u,\mu) = q(t = -\infty)$, $(v,\nu) = q(t = +\infty)$, $\omega > 0$. Since this geodesic curve is γ_0 invariant, it becomes a periodic orbit on SRS. Furthermore we can see that two geodesic curves that connect the fixed points of γ and those of its conjugate $k\gamma k^{-1}$ $(k\in$S$\Gamma)$, respectively, become the same periodic orbit on SRS. Therefore each primitive inconjugate element γ_0 in SΓ corresponds to a periodic orbit on SRS. The "length" of a periodic orbit is given by

$$l(\gamma_0) \equiv d(q,\gamma_0(q)) \quad (q\text{: any point on the orbit})\,. \tag{18}$$

Since $l(\gamma_0)$ depends only on the conjugacy class of γ_0 in SΓ [or SPL(2,R)], we can take a magnification as γ_0,[9]

$$\begin{aligned} \gamma_0(z) &= N_0 z\,, \\ \gamma_0(\theta) &= \pm N_0^{1/2}\theta \equiv \chi(\gamma_0) N_0^{1/2}\theta\,, \quad N_0 > 1\,, \end{aligned} \tag{19}$$

and then

$$l(\gamma_0) = \ln N_0\,. \tag{20}$$

Any geodesic curve (on SH) that does not connect the fixed points of an element $\gamma\in$SΓ becomes a nonperiodic orbit and such geodesic curves are dense on SH, which implies that the classical motion on SRS is chaotic. In fact, the classical motion has the Anosov property[13] and the Kolmogorov–Sinai entropy[14] is given by the integral constant ω in (10).[2,12]

IV. QUANTIZATION

Next we discuss quantization in the path integral method. Let $|q\rangle$ be the eigenstate of the coordinate operator $\hat{q}$ with eigenvalue q. The inner product and the identity operator are given by (cf. Ref. 15)

$$\langle q|q'\rangle = g^{-1/4}(q)g^{-1/4}(q')\delta(q - q')\,, \tag{21}$$

$$\mathrm{Id} = \int dq\, g^{1/2}(q)|q\rangle\langle q|\,, \tag{22}$$

with

$$\begin{aligned} dq &\equiv d(\mathrm{Re}\, z)d(\mathrm{Im}\, z)d\theta\, d\bar{\theta}\,, \\ \delta(q - q') &\equiv \delta^2(z - z')(\bar{\theta} - \bar{\theta}')(\theta - \theta')\,, \end{aligned} \tag{23}$$

$$\int d(\mathrm{Re}\, z)d(\mathrm{Im}\, z)\delta^2(z) = \int d\theta\, \theta = \int d\bar{\theta}\bar{\theta} = 1\,. \tag{24}$$

In the q representation the canonical momentum operator $\hat{p}$ becomes

$$\hat{p}_A = -i\hbar g^{-1/4}\partial_A g^{1/4}\,. \tag{25}$$

The kernel function $K(q',t';q,t)$ may be expressed by

$$\begin{aligned} K(q',t';q,t) = {} & \int \prod_{k=1}^{N-1} g^{1/2}(q_k)dq_k \\ & \times \langle q',t'|q_{N-1},t_{N-1}\rangle \cdots \langle q_1,t_1|q,t\rangle\,, \end{aligned} \tag{26}$$

where we have subdivided the time interval $t' - t$ into N small intervals of equal duration δt and $t_k = t + k\delta t$. Since

$$\begin{aligned} \langle q',\delta t\,|q,0\rangle &= \langle q'|e^{-(i\delta t/\hbar)H_Q}|q\rangle \\ &\simeq \langle q'|q\rangle - (i\delta t/\hbar)\langle q'|H_Q|q\rangle\,, \end{aligned} \tag{27}$$

we shall calculate the matrix element $\langle q'|H_Q|q\rangle$ of H_Q in (6). We find (cf. Ref. 15)

$$\begin{aligned} \langle q'|H_Q|q\rangle = {} & \pi^{-2}g^{-1/4}(q')g^{-1/4}(q)\int dp\, e^{i(q'-q)\cdot p/\hbar} \\ & \times \left[\frac{1}{2m}g^{AB}(\bar{q})p_B p_A + \Delta V(\bar{q})\right], \end{aligned} \tag{28}$$

where $\bar{q} = (q' + q)/2$, $dp \equiv d(\mathrm{Re}\, p_z)d(\mathrm{Im}\, p_z)dp_\theta\, dp_{\bar{\theta}}$, and the additional potential $\Delta V(\bar{q})$ is given by

$$\begin{aligned} \Delta V = (\hbar^2/8m)\{ & g^{AB}\Gamma_B\Gamma_A + 2(-)^a\partial_A(g^{AB}\Gamma_B) \\ & + (-)^{a+b}\partial_B\partial_A g^{AB}\}\,, \end{aligned} \tag{29}$$

$$\Gamma_A \equiv \tfrac{1}{2}\partial_A (\ln g)\,, \tag{30}$$

which is a quantum correction of order $\hbar^2$. This vanishes exactly with our metric g_{AB} in (2).[1] Equations (26)–(28) yield

$$K(q',t';q,t) = [g(q')g(q)]^{-1/4}\int \prod_{k=1}^{N-1} dq_k \prod_{l=1}^{N} \frac{dp^{(l)}}{\pi^2} \times \exp\left\{\frac{i\delta t}{\hbar}\sum_{n=1}^{N} S_0(n,n-1)\right\}, \tag{31}$$

where

$$S_0(n,n-1) = [(q_n - q_{n-1})/\delta t] \cdot p^{(n)} - (1/2m) g^{AB}(\bar{q}_n) p_B^{(n)} p_A^{(n)}\,, \tag{32}$$

$$\bar{q}_n = (q_n + q_{n-1})/2\,, \quad q_0 \equiv q, \quad q_N \equiv q'\,. \tag{33}$$

The integration with respect to $p^{(k)}$ can be performed and we get

$$K(q',t';q,t) = [g(q')g(q)]^{-1/4}\int \prod_{k=1}^{N-1} dq_k \prod_{l=1}^{N} \times\left[\left(-\frac{g^{1/2}(\bar{q}_l)}{\pi}\right)\exp\left\{\frac{i\delta t}{\hbar} S(l,l-1)\right\}\right], \tag{34}$$

where

$$S(l,l-1) = (m/2)(\delta t)^{-2}(q_l - q_{l-1})^A \times g_{AB}(\bar{q}_l)(q_l - q_{l-1})^B\,. \tag{35}$$

In the semiclassical approximation the kernel function becomes

$$\widetilde{K}(q',t',q,t) = C[g(q')g(q)]^{-1/4} \times\left\{\mathrm{Sdet}\left[(-)^b \frac{\partial^2 S_{\mathrm{cl}}}{\partial q^A\,\partial q'^B}\right]\right\}^{1/2} e^{(i/\hbar)S_{\mathrm{cl}}}, \tag{36}$$

where C is a normalization constant and S_{cl} is the action integral along the classical path,

$$S_{\mathrm{cl}}(q',q;t'-t) = \int_t^{t'} L(q,\dot{q})dt\,. \tag{37}$$

Actually we can see that Eq. (36) satisfies the Schrödinger equation up to a remainder of order $\hbar^2$. We find

$$S_{\mathrm{cl}} = (m/2)[d(q',q)]^2/(t'-t)\,, \tag{38}$$

where d is the "length" given in (13).[16] Note that the classical path between any two points on SH is uniquely given, so that the summation with respect to the classical path is unnecessary in (36). Plugging (38) into (36) we finally obtain the kernel function on SH in the semiclassical approximation:

$$\widetilde{K}(q',q|t'-t) \equiv \widetilde{K}(q',t';q,t) = -\left\{\frac{2\tanh(d/2)}{\pi^2 d}\right\}^{1/2} \times\exp\left\{\frac{imd^2}{2\hbar(t'-t)}\right\}. \tag{39}$$

Then the kernel function on SRS in the semiclassical approximation is given by[9]

$$\widetilde{K}_{\mathrm{SRS}}(q',q|T) = \sum_{\gamma\in S\Gamma} \widetilde{K}(q',\gamma(q)|T)\,. \tag{40}$$

This yields a superanalog of Selberg's trace formula, which we shall see in Sec. V.

V. SUPERANALOG OF SELBERG'S TRACE FORMULA

We calculate the supertrace of the kernel function in (40), which is defined by

$$\mathrm{Str}\,\widetilde{K}_{\mathrm{SRS}} = \int_F dq\, g^{1/2}(q)\widetilde{K}_{\mathrm{SRS}}(q,q|T)\,, \tag{41}$$

where F is a fundamental domain of SH/SΓ. Equations (40) and (41) yield (cf. Ref. 17)

$$\mathrm{Str}\,\widetilde{K}_{\mathrm{SRS}} = A_0 + \sum_{\substack{\text{inconjugate}\\ \text{primitive } \gamma_0}} \sum_{n=1}^{\infty} A_n(\gamma_0)\,, \tag{42}$$

where

$$A_n(\gamma_0) = \int_{F_{\gamma_0}} dq\, g^{1/2}(q)\widetilde{K}(q,\gamma_0^n(q)|T)\,, \tag{43}$$

F_{γ_0}: a fundamental domain for the centralizer

$$Z(\gamma_0) = \{k\,|\,k\gamma_0 k^{-1} = \gamma_0\,,\ k\in S\Gamma\}\,, \tag{44}$$

$$A_0 = A_n(1) \quad (F_1 = F)\,. \tag{45}$$

The first term A_0 is the contribution from the classical periodic orbit with zero length and it becomes a constant dependent only on the genus g of SRS,

$$A_0 = \mathrm{Area}(F)\cdot\widetilde{K}(0,0,|T) = 1 - g\,. \tag{46}$$

Next we shall evaluate the second term. Since γ_0 can be a magnification (19), we take[9,17]

$$F_{\gamma_0} = \{q\,|-\infty < \mathrm{Re}\,z < +\infty\,,\ 1 \leqslant \mathrm{Im}\,z \leqslant N_0\}\,. \tag{47}$$

Then we find

$$A_n(\gamma_0) = \int_1^{N_0} d(\mathrm{Im}\,z)\int_{-\infty}^{\infty} d(\mathrm{Re}\,z)\int d\theta\, d\bar{\theta}\,\frac{-1}{2Y} \times\left\{\frac{2\tanh(d(\gamma)/z)}{\pi^2 d(\gamma)}\right\}^{1/2}\exp\left\{\frac{im}{2\hbar T}d^2(\gamma)\right\}$$
$$= \frac{-\ln N_0}{2\pi(N^{1/2} - N^{-1/2})}\int_{-\infty}^{\infty} dx \int d\zeta\, d\bar{\zeta} \times\left\{\frac{\tanh(d(\gamma)/2)}{d(\gamma)/2}\right\}^{1/2}\exp\left\{\frac{\mathrm{im}}{2\hbar T}d^2(\gamma)\right\}, \tag{48}$$

where

$$d(\gamma) \equiv d(q,\gamma(q)) = \cosh^{-1}\left[\frac{x^2}{2} + \frac{N+N^{-1}}{2} - \frac{N^{1/2}+N^{-1/2}}{2} \times(N^{1/2}+N^{-1/2} - 2\chi(\gamma)\zeta\bar{\zeta})\right], \tag{49}$$

$$\gamma = (\gamma_0)^n, \quad N = (N_0)^n, \quad \chi(\gamma) = \{\chi(\gamma_0)\}^n\,, \tag{50}$$

and we have taken the convenient variables[9]

$$\begin{aligned} \mathrm{Re}\, z &= [Y/(N^{1/2}-N^{-1/2})]x, \\ \theta &= (\mathrm{Im}\, z)^{1/2}\zeta, \quad \overline{\theta} = (\mathrm{Im}\, z)^{1/2}\overline{\zeta}. \end{aligned} \tag{51}$$

We shall integrate x, ζ, and $\overline{\zeta}$ in the semiclassical approximation: Expanding $d(\gamma)$ around the extremum $(x^2 = \zeta\overline{\zeta} = 0)$ corresponding to a classical periodic orbit,

$$\begin{aligned} d(\gamma) = l(\gamma) + \hbar\bigg\{ & \frac{N}{N^2-1}\left(\frac{x^2}{\hbar}\right) \\ & - \frac{N+1-2\chi(\gamma)N^{1/2}}{N-1}\left(\frac{\zeta\overline{\zeta}}{\hbar}\right)\bigg\} + O(\hbar^2), \end{aligned} \tag{52}$$

$$l(\gamma) = \ln N = nl(\gamma_0), \tag{53}$$

we get

$$\begin{aligned} A_n(\gamma_0) &= \left(\frac{-im}{2\pi\hbar T}\right)^{1/2} l(\gamma_0) \\ &\quad \times \left\{\coth\left(\frac{l(\gamma)}{2}\right) - \chi(\gamma)\operatorname{csch}\left(\frac{l(\gamma)}{2}\right)\right\} e^{(im/z\hbar T)l^2(\gamma)}. \end{aligned} \tag{54}$$

Note that

$$\begin{aligned} \sum_{\substack{\text{inconjugate} \\ \text{primitive } \gamma_0}} \sum_{n=1}^{\infty} A_n(\gamma_0) &= \sum_{\substack{\text{periodic} \\ \text{orbit}}} \sum_{n=1}^{\infty} \left(\frac{-im}{2\pi\hbar T}\right)^{1/2} 2l \left\{\coth\left(\frac{nl}{2}\right)\right. \\ &\quad \left. - (\text{sign})^n \operatorname{csch}\left(\frac{nl}{2}\right)\right\} e^{(im/2\hbar T)(nl)^2}, \end{aligned} \tag{55}$$

where l is the length of a classical periodic orbit and (sign) means $\chi(\gamma_0)$.

The supertrace of the exact kernel function may be written by

$$\begin{aligned} \mathrm{Str}\, K_{\mathrm{SRS}} &= \mathrm{Str}[e^{-(iT/\hbar)H_Q}] \\ &= \sum_{n=0}^{\infty} (e^{-(iT/\hbar)E_n^{\mathrm{B}}} - e^{-(iT/\hbar)E_n^{\mathrm{F}}}), \end{aligned} \tag{56}$$

where H_Q is the quantum mechanical Hamiltonian in (6) and E_n^{B} and E_n^{F} are "energy" eigenvalues for bosonic and fermionic states, respectively. Equating the supertrace of the exact kernel function with that in the semiclassical approximation, we get

$$\begin{aligned} \sum_{n=0}^{\infty} (e^{-(iT/\hbar)E_n^{\mathrm{B}}} - e^{-(iT/\hbar)E_n^{\mathrm{F}}}) &= 1 - g + \sum_{\substack{\text{periodic} \\ \text{orbit}}} \sum_{n=1}^{\infty} \left(\frac{2m}{i\pi\hbar T}\right)^{1/2} l \left\{\coth\left(\frac{nl}{2}\right)\right. \\ &\quad \left. - (\text{sign})^n \operatorname{csch}\left(\frac{nl}{2}\right)\right\} e^{(im/2\hbar T)(nl)^2}, \end{aligned} \tag{57}$$

which is exactly a superanalog of Selberg's trace formula.[9] In fact, since H_Q is proportional to the "super"-Laplace–Beltrami operator Δ_{SLB},

$$\begin{aligned} H_Q &= -(\hbar^2/2m)\Delta_{\mathrm{SLB}} \\ &= -(\hbar^2/2m)[2Y(\partial_\theta + \theta\,\partial_z)(-\partial_{\overline{\theta}} + \overline{\theta}\,\partial_{\overline{z}})]^2, \end{aligned} \tag{58}$$

Eq. (57) is rewritten by[18]

$$\begin{aligned} \mathrm{Str}[e^{t\Delta_{\mathrm{SLB}}}] &= 1 - g \\ &\quad + \sum_{\substack{\text{inconjugate} \\ \text{primitive } \gamma_0}} \sum_{n=1}^{\infty} (4\pi t)^{-1/2} \frac{\ln N_0}{N^{1/2} - N^{-1/2}} \\ &\quad \times (N^{1/2} + N^{-1/2} - 2\chi(\gamma))e^{-(\ln N)^2/4t}. \end{aligned} \tag{59}$$

Equation (57) gives the response function $g(E)$ of a quantum system to a continuing external stimulus of a given frequency $\omega = E/\hbar$,

$$\begin{aligned} g(E) &\equiv \frac{1}{i\hbar}\int_0^{\infty} dT\, e^{iET/\hbar}\, \mathrm{Str}\, K_{\mathrm{SRS}} \\ &\quad \left[= \sum_{n=0}^{\infty}\left(\frac{1}{E - E_n^{\mathrm{B}}} - \frac{1}{E - E_n^{\mathrm{F}}}\right)\right] \\ &= \frac{1-g}{E} + \sum_{\substack{\text{periodic} \\ \text{orbit}}} \sum_{n=1}^{\infty} \frac{2ml}{i\hbar p} \\ &\quad \times \left[\coth\left(\frac{nl}{2}\right) - (\text{sign})^n \operatorname{csch}\left(\frac{nl}{2}\right)\right] e^{inlp/\hbar}, \end{aligned} \tag{60}$$

where

$$p \equiv (2mE)^{1/2}. \tag{61}$$

VI. SUMMARY AND COMMENTS

We have investigated the free motion of a nonrelativistic superparticle on SRS on genus $g \geqslant 2$. This should be a good example of classically chaotic supersymmetric models. Energy is the only quantum number and the "super"-sum (supertrace) of the energy eigenvalues (Hamiltonian) was evaluated in the semiclassical approximation, which led to a superanalog of Selberg's trace formula. The "super"-sum over energy eigenvalues, which is a quantum mechanical quantity, is determined from information about the classical periodic orbits. Physically this would be because the system is chaotic. This model will also help in understanding SRS itself, which is important in superstring theory.

Finally we make some comments.

(i) The fact that the rhs of (57) or (60) does not vanish means that supersymmetry is broken due to the boundary conditions.

(ii) In the bosonic case,[6] i.e., a particle on the Riemann surface ($g \geqslant 2$), the correspondent to the rhs of (57) or (60) is affected by a quantum correction because ΔV corresponding to (29) gives a nonvanishing constant.[19] (This correction should be made by hand when equating the exact result to the approximated one.)

(iii) In Eq. (59), when $t \to 0^+$ the first term from the periodic orbit with zero length becomes dominant, which can be evaluated in the heat kernel expansion,[20]

$$K(q,q|T) \sim \sum_{n=0}^{\infty} B_n(q) t^n, \quad t \to 0^+ . \tag{62}$$

Then Eq. (59) implies that all the coefficients B_n except B_0 vanish.

ACKNOWLEDGMENTS

One of the authors (Y.Y.) is grateful to the theoretical particle physics group at King's College for their hospitality. He is indebted to the Ministry of Education, Science, and Culture of Japan for financial support.

---

[1] S.Uehara and Y. Yasui, Phys. Lett. B **202**, 530 (1988).

[2] D. Friedan, in the *Proceedings of the Workshop on Unified String Theories*, edited by D. Gross and M. Green (World Scientific, Singapore, 1986); L. Crane and J. M. Rabin, Chicago preprint EFI 87-25, 1986; J. M. Rabin, Phys. Lett. B **190**, 40 (1987).

[3] N. L. Balazs and A. Voros, Phys. Rep. **143**, 109 (1986).

[4] M. C. Gutzwiller, Phys. Scr. **9**, 184 (1985).

[5] F. Steiner, DESY preprint DESY 87-022, 1987; Preprint ISSN 0418-9833, 1987.

[6] M. C. Gutzwiller, J. Math. Phys. **12**, 343 (1971); Phys. Rev. Lett. **45**, 150 (1980); Contemp. Math. **53**, 215 (1986).

[7] A. Selberg, J. Indian Math. Soc. **20**, 47 (1956).

[8] E. D'Hoker and D. H. Phong, Phys. Rev. Lett. **56** 912 (1986); Commun. Math. Phys. **104**, 537 (1986); Nucl. Phys. B **269**, 205 (1986).

[9] M. A. Baranov, Yu. I. Manin, I. V. Frolov, and A. S. Schwarz, Sov. J. Nucl. Phys. **43**, 670 (1986); Commun. Math. Phys. **111**, 373 (1987); K. Aoki, Commun. Math. Phys. **117**, 405 (1988).

[10] P. Howe, J. Phys. A **12**, 393 (1979).

[11] Note that $\overline{X+Y} = \overline{X} + \overline{Y}$, $\overline{XY} = \overline{Y}\,\overline{X}$.

[12] S. Matsumoto and Y. Yasui, Prog. Theor. Phys. **79**, 1022 (1988).

[13] D. V. Anosov, Proc. Steklov Inst. Math. **90** (1967).

[14] Ya. G. Sinai, *Introduction to Ergodic Theory* (Princeton U.P., Princeton, NJ, 1976).

[15] M. Omote, Nucl. Phys. B **120**, 325 (1977).

[16] This expression (38) is valid for other classical solutions than those in (11) as well as for the solutions in (11).

[17] H. P. McKean, Commun. Pure Appl. Math. **25**, 47 (1972).

[18] In Ref. 3 the correspondent to the first term on the rhs should read

$$i(g-1)\int_{-\infty}^{\infty} dr \exp\{t(\tfrac{1}{2}+ir)^2\}\tanh \pi r ,$$

which can be evaluated to be $1-g$.

[19] C. Grosche and F. Steiner, DESY preprint 87-014.

[20] See, for example, P. B. Gilkey, "The Index Theorem and the Heat Equation," Math. Lect. Series, Vol. 4.

Limit theorems for spin systems with an Abelian discrete symmetry

Nestor Caticha
Instituto de Fisica e Quísica de São Carlos, Departamento de Fisica e Ciência dos Materiais, Universidade de São Paulo, Caixa Postal 369, 13.560 São Carlos, São Paulo, Brazil

(Received 23 February 1988; accepted for publication 20 July 1988)

Exponential convergence in the number of states is proved for the pressure of a class of discrete Abelian spin systems, to the O(2) invariant model's pressure in d dimensions. Analogous results are obtained for the convergence of the expectation values of a class of functions under more stringent conditions.

I. INTRODUCTION

Due to the possibility that experimental realizations of the $Z(p)$ and O(2) invariant models may exist, it is of interest to study the relation between these models. In particular, we are interested in the high p limit in which the discrete symmetry approaches the continuous one. Several authors have dealt with the phase structure of those models especially in two dimensions.[1,2] Results concerning the intermediate massless phase were first obtained by Elitzur *et al.*[1] using duality arguments and correlation inequalities. Their results suggest that the temperature of symmetry breaking decreases as $1/p^2$. Renormalization group arguments and Monte Carlo simulations[3] support this suggestion. A $1/p^2$ lower bound to this temperature can be obtained by a Peierls–Chessboard argument.[4] Rigorous upper bounds are not known. The existence of the intermediate phase was proved by Frölich and Spencer.[5] The decrease of the temperature of symmetry breaking is a consequence of the decrease of the mass gap of the discrete system as the number of states increases, which is due to the better approximations of a continuous symmetry by a discrete one. We thus want to control the convergence of thermodynamic quantities as p increases. Theorem 1, in Sec. II, proves the exponential convergence of the pressure. Its proof is based on properties of the *a priori* measures. For temperatures where the expectation values of a class of functions of the O(2) are Holder continuous, exponential convergence is also proved.

Aside from its theoretical interest, this type of result may be useful to judge numerical methods of simulation of spin and gauge systems, where a discrete symmetry is used to approximate a continuous one in order to increase speed in computations.

II. THE MODEL AND MAIN RESULT

We first introduce the notation and the model.

Consider a $Z(p)$ invariant model in a volume $\Lambda \subset Z^d$, a variable θ_i at each site $i \in \Lambda$, and Hamiltonian

$$H_\Lambda(\Phi_x) = \sum_{\{\theta_i\} \in x \subset \Lambda} \Phi_x(\{\theta\})$$

periodic in each of its variables θ_i, and *a priori* measure

$$d\nu_p = \prod_{i \in \Lambda} d\mu_p(\theta_i),$$

$$d\mu_p(\theta) = \frac{d\theta}{p} \sum_{n=0}^{p-1} \delta\left(\theta - \frac{2\pi}{p} n\right).$$

Call the O(2) invariant measure

$$d\nu = \prod_{i \in \Lambda} d\mu(\theta_i), \quad d\mu(\theta) = \frac{d\theta}{2\pi},$$

the finite volume partition functions

$$Z_\Lambda^{Z(p)} = \int d\nu_p \, e^{-\beta H_\Lambda}, \quad Z_\Lambda^{O(2)} = \int d\nu \, e^{-\beta H_\Lambda},$$

and the pressures

$$P^{Z(p),O(2)} = \lim_{\Lambda \to \infty} P_\Lambda^{Z(p),O(2)},$$

where

$$P_\Lambda^{Z(p),O(2)} = (1/|\Lambda|) \ln Z_\Lambda^{Z(p),O(2)}.$$

We will prove the following theorem.

Theorem 1: If

(i) $\Phi_x \in B_1 = \left\{ \Phi \,\middle|\, \|\Phi\|_1 = \sum_{0 \in x} \sup|\Phi_x| \text{ is finite} \right\};$

(ii) Φ_x is periodic in each $\theta_i \in R$; and (iii) Φ_x is analytic for $|\mathrm{Im}(\theta_i)| \leqslant A$, and for $0 < \alpha \leqslant A$; there is $0 < g(\alpha)$ such that for each $i \in \Lambda$,

$$|\mathrm{Re}(\Phi_x(\theta_i + i\alpha, \tilde{\theta}))| \leqslant |\Phi_x(\theta_i, \tilde{\theta})| g(\alpha),$$

where $\tilde{\theta}$ denotes all other variables θ other than θ_i, then there is a bounded function $C(\beta)$ such that for fixed β

$$\begin{aligned} 1 - 2C(\beta)e^{-p\alpha}/(1 - e^{-p\alpha}) & \\ \leqslant \exp[P^{Z(p)} - P^{O(2)}] & \\ \leqslant 1 + 2C(\beta)e^{-p\alpha}/(1 - e^{-p\alpha}), & \qquad (1) \end{aligned}$$

with $C(\beta)$ given by

$$C(\beta) = \exp[\beta \|\Phi\|_1 (1 + g(\alpha))]$$

and $g(\alpha)$ is a monotonic increasing function of α.

Remarks: The difference in pressures is exponentially small in the number of states p for sufficiently large p, as can be seen by taking the logarithm of Eq. (1) and expanding, therefore there is no $1/p$ expansion for the $Z(p)$ pressure around the O(2)'s. For the Villain model in two dimensions[4] a nonrigorous argument suggests a Gaussian convergence; however, this cannot be general since in one dimension the calculation can be done explicitly and the convergence proved to be only exponential. The result of Theorem 1 is independent of the lattice dimension.

This result can also be expanded to include lattice $Z(p)$ and U(1) gauge theories.

It improves a result of Simon[6] using Bishop–de Leew order, who proves the convergence to be at least as fast as 1/

0022-2488/88/112491-03$02.50

p^2. Condition (iii) is satisfied automatically for finite n if Φ has only up to n-body interactions.

As a final remark we stress that the theorem is a consequence of the exponential convergence of Riemann sums of periodic functions to the Riemann integral, in the number of terms in the sum.

III. PROOF OF THEOREM 1

We state without proof some preliminary standard results.

Proposition 1: Let $f(\theta)$: $C \to R$ be a periodic positive function, analytic for $|\text{Im}(\theta)| < A$, with Fourier coefficients

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{ik\theta} \, d\theta$$

and $0 < \alpha < A$, then there is constant $C_1 > 0$ (which depends on α but not on k) such that

$$\text{(a)} \quad -C_1 f_0 e^{-k\alpha} < f_k \leqslant C_1 f_0 e^{-k\alpha}$$

and with $d\mu_p$ and $d\mu$ as in Sec. II;

$$\text{(b)} \int d\mu f(\theta)\left[1 - \frac{2C_1 e^{-p\alpha}}{1 - e^{-p\alpha}}\right] < \int d\mu_p f(\theta) < \left[1 + \frac{2C_1 e^{-p\alpha}}{1 - e^{-p\alpha}}\right] \int d\mu f(\theta).$$

Part (a) is used to prove part (b) which shows the well known result of exponential convergence of the R-sums to the R-integral.

We first assume that Φ has only up to n-body interactions, then consider the Boltzmann factor $\exp(-\beta H_\Lambda)$ and a site $i \in \Lambda$. Since as a function of θ_i it satisfies the conditions of Proposition 1, we have

$$\left[1 - \frac{2C(\tilde{\theta}) e^{-p\alpha}}{1 - e^{-p\alpha}}\right] \int d\mu_i(\theta) e^{-\beta H_\Lambda(\theta_i, \tilde{\theta})},$$
$$\leqslant \int d\mu_p(\theta_i) e^{-\beta H_\Lambda(\theta_i, \tilde{\theta})}$$
$$\leqslant \left[1 + \frac{2C(\tilde{\theta}) e^{-p\alpha}}{1 - e^{-p\alpha}}\right] \int d\mu_i \, e^{-\beta H_\Lambda(\theta_i, \tilde{\theta})}.$$

It is straightforward to show that a $\tilde{\theta}$-independent bound can be obtained for $C(\tilde{\theta})$,

$$C(\tilde{\theta}) \leqslant \exp[\|\Phi\|_1 (1 + g(\alpha))] = C(\beta).$$

For infinite-body interactions condition (iii) is imposed, leading to the same bound for $C(\tilde{\theta})$. Calling

$$K_\pm = 1 \pm 2C(\beta) e^{-p\alpha} / (1 - e^{-k\alpha})$$

and integrating sequentially every site $i \in \Lambda$, one obtains a cascade of inequalities, which lead to

$$K_-^{|\Lambda|} Z_\Lambda^{O(2)} \leqslant Z_\Lambda^{Z(p)} \leqslant K_+^{|\Lambda|} Z_\Lambda^{O(2)}. \tag{2}$$

Since the bound on $C(\tilde{\theta})$ is valid in the infinite volume, the bound of Eq. (2) is seen to be uniform in the volume, thus completing the proof.

IV. CONVERGENCE OF EXPECTATION VALUES

By adding a source term to the Hamiltonians and taking derivatives one defines the expectation values of functions in a translationally invariant state,

$$H_\Lambda(\Phi_x + t\psi^A) = H(\Phi_x) + t \sum_{(i+x) \subset \Lambda} \tau_i A,$$

where $\tau_i A$ are the translates of A,

$$\langle A \rangle_\Phi = -\lim_{\Lambda \to \infty} \frac{1}{\beta} \frac{d}{dt} P_\Lambda(\Phi_x + t\psi^A)|_{t=0}.$$

We want to study under which additional conditions the correlations of the $Z(p)$ models converge to those of the O(2), thus we have the following theorem.

Theorem 2: If (i) $\exp\text{-}\beta H_\Lambda(\Phi + t\psi)$ satisfies the conditions of Theorem 1; (ii) $P_\Lambda^{Z(p)}(\Phi + t\psi)$ and $P_\Lambda^{O(2)}(\Phi + t\psi)$ are differentiable w.r.t. t in a neighborhood of $t = 0$; and (iii) $(d/dt) P_\Lambda^{O(2)}(\Phi + t\psi)$ is Holder continuous, then for p sufficiently large there exist positive constants α' and $C'(\beta)$ such that

$$|\langle A \rangle_\Phi^{Z(p)} - \langle A \rangle_\Phi^{O(2)}| < C'(\beta) \exp - p\alpha'.$$

The proof of the theorem follows from the next result.

Proposition 2: Let $F_n(z)$ be a sequence of convex functions that converge pointwise to $F(z)$, and C, M, and α be positive constants such that (i) $|F_n(z) - F(z)| < Ce^{-p\alpha}$ for any z; (ii) $F_n(z)$ and $F(z)$ are differentiable in a neighborhood of x; (iii) the derivative $DF(x)$ is Holder continuous of order γ; then $|DF_n(x) - DF(x)| < (2C + M) e^{-n\alpha[\gamma/(\gamma+1)]}$.

Proof: Define

$$D^+ F_n(x) = \lim_{\theta \uparrow 1} \frac{F_n(\theta x + (1 - \theta) y) - F_n(x)}{(1 - \theta)(y - x)},$$
$$D^- F_n(y) = \lim_{\theta \downarrow 1} \frac{F_n(y) - F_n(\theta x + (1 - \theta) y)}{\theta (y - x)},$$

by convexity

$$D^- F_n(y) \geqslant \frac{F_n(y) - F_n(x)}{x - y} \geqslant D^+ F_n(x).$$

$F(z)$ is also convex and the convergence is uniform in compact subsets, so similar formulas hold. Consider a sequence $y_m = x + e^{-mA}$, m and A to be chosen later, then

$$D^+ F_n(x) - D^- F(y_m) \leqslant 2Ce^{-n\alpha + mA},$$

and we can drop the superscripts $\pm$ since F_n, F are differentiable in a neighborhood of x.

Holder continuity states that

$$|DF(y) - DF(x)| < M |y - x|^\gamma,$$

thus

$$|DF(x) - DF_n(x)|$$
$$< |DF(x) - DF(y_m)| + |DF(y_m) - DF_n(x)|$$
$$< Me^{-mA\gamma} + 2Ce^{-n\alpha + mA},$$

choosing $A = \alpha/(\gamma + 1)$ and $m = n$,

$$|DF(x) - DF_n(x)| < (2C + M) e^{-n\alpha[\gamma/(\gamma+1)]}.$$

Observing that $P_\Lambda^{Z(p)}$ are convex and since a bound uniform in $|\Lambda|$ can be obtained using Proposition 2, Theorem 2 is proved.

ACKNOWLEDGMENTS

The author would like to thank B. Simon for suggesting the problem and for discussions and V. Rittenberg for encouragement. This work was partially supported by FAPESP.

[1] S. Elitzur, P. Pearson, and J. Shigemitsu, Phys. Rev. D **19**, 3968 (1979).

[2] F. C. Alcaraz and R. Köberle, J. Phys. A **13**, L153 (1980).

[3] F. C. Alcaraz and J. Jacobs, Nucl. Phys. B **210**, 246 (1982).

[4] N. Caticha, Ph.D. thesis, Caltech, 1984 (unpublished).

[5] J. Frölich and T. Spencer, Commun. Math. Phys. **81**, 527 (1981).

[6] B. Simon, *The Statistical Mechanics of Lattice Gases* (Princeton U.P., Princeton, NJ, to be published).

Anomalous congruences

H. W. Braden[a)]
Institute of Field Physics, Department of Physics and Astronomy, and Department of Mathematics, University of North Carolina, Chapel Hill, North Carolina 27514

(Received 19 April 1988; accepted for publication 29 June 1988)

Congruences between different Chern classes and hence anomaly coefficients are observed. These are determined in general for SU(n) vector bundles over spheres. The results are of use in calculating global anomaly freedom for gauge theories as well as other index theory calculations.

I. INTRODUCTION

A problem that regularly arises for both mathematicians and physicists is the evaluation of characteristic classes of an associated vector bundle E_λ to some principal G-bundle P over M. Here G is a compact, semisimple Lie group and λ a representation of this group. The problem may conveniently be divided into two portions. The first is a Lie algebraic one. Recall we may identify $P(L)^G$, the G-invariant polynomial functions on the Lie algebra L of G, with cohomology classes of $H^*(BG,R)$. Further, by a theorem of Chevalley,[1] $P(L)^G$ is isomorphic with $P(H)^W$, the algebra of Weyl-invariant polynomials on H, a Cartan subalgebra of L. The first aspect of the problem is then to express the λ-dependent polynomial representing a given characteristic class in terms of some basis of $P(H)^W$. Typically, this basis is chosen so as to facilitate the second and remaining part of the problem, which is to evaluate the integrals of the forms appearing over M. For the cases most often dealt with M is a sphere and the basis chosen reflects the homotopy generators of G. In the ensuing paper we will focus on this first aspect, expressing a Weyl-invariant polynomial in terms of a fixed basis of $P(H)^W$. In particular we shall observe some congruences that arise in this procedure when $G = \mathrm{SU}(n)$; this will yield congruences amongst Chern classes of E_λ. These congruences exist for other Lie groups as well but we will not examine these here. These congruences are of particular relevance when studying the anomalies of a given gauge theory.

To be more specific, suppose $\pi = (p_1^{\pi_1}, p_2^{\pi_2}, \dots, p_r^{\pi_r})$ is a partition of k; then we seek the coefficients $A_\pi(\lambda,\nu)$ for two representations λ and ν of G defined by (for $x \in L$)

$$\mathrm{Tr}_\lambda\, x^k = \sum_{|\pi| = k} A_\pi(\lambda,\nu) \prod_{i=1}^{r} (\mathrm{Tr}_\nu\, x^{p_i})^{\pi_i}.$$

In determining $A_\pi(\lambda,\nu)$ we may take $x \in H$. Further definitions are given in Sec. II where we describe how to calculate the coordinates $A_\pi(\lambda,\nu)$ of a polynomial in $P(H)^W$ by restricting our attention to the weight spaces of the above equation. Upon defining

$$f_k(\mu) = \frac{1}{|\mathrm{Stab}\,\mu|} \sum_{\sigma\in W} (\sigma\mu)^k,$$

we may express this Weyl-invariant polynomial in the form

$$f_k(\mu) = \sum_{|\pi| = k} a_\pi(\mu,\nu) \prod_{i=1}^{r} f_{p_i}(\nu)^{\pi_i}.$$

We then have

$$A_\pi(\lambda,\nu) = \sum_{\mu\in\Pi^{+}(\lambda)} m_\lambda(\mu) a_\pi(\mu,\nu),$$

where $m_\lambda(\mu)$ are the multiplicities of the weight μ in the representation λ. Properties of $A_\pi(\lambda,\nu)$ can then be obtained from those of $a_\pi(\lambda,\nu)$. In particular we have the following lemma when $A_\pi(\lambda,\nu)$ and $a_\pi(\lambda,\nu)$ are integers.

Lemma 1:

$A_\pi(\lambda,\nu) \equiv A_{\pi'}(\lambda,\nu) \bmod q$ (for all $\lambda\in\Pi^+$) if and only if $a_\pi(\lambda,\nu) \equiv a_{\pi'}(\lambda,\nu) \bmod q$ (for all $\lambda\in\Pi^+$).

We report and prove here the following theorem for $G = \mathrm{SU}(n)$. Let λ_1 be the defining or n-dimensional representation of G and set $A_{(k)}(\lambda,\lambda_1) \equiv A_{(k)}(\lambda)$ and $a_{(k)}(\lambda,\lambda_1) = a_{(k)}(\lambda)$.

Theorem 1: Let $k,k' \leqslant n$. Then for SU(n) we have, for all representations μ,

$$a_{(k)}(\mu) \equiv a_{(k')}(\mu) \mod q$$

if and only if

(i) $m^k \equiv m^{k'} \bmod q$, for all $m \in Z$,

(ii) $p^{k-1} \equiv p^{k'-1} \bmod q$, for primes $p < n$.

Corollary 1: With the same conditions as above we find $A_{(k)}(\mu) \equiv A_{(k')}(\mu) \bmod q$ for all representations μ.

The corollary follows immediately from Theorem 1 and Lemma 1. We note that the integers $A_{(k)}(\lambda)$ give us the Chern class c_k of E_λ over S^{2k}, thus we have the following corollary.

Corollary 2: Let c_k ($c_{k'}$) be the Chern class of an SU(n) bundle E_μ over S^{2k} ($S^{2k'}$) that corresponds to the generator of $\pi_{2k-1}(G)$ [$\pi_{2k'-1}(G)$] with $k,k' \leqslant n$. Then

$$c_k \equiv c_{k'} \mod q,$$

for all bundles E_μ with q determined in Theorem 1.

Various divisibility properties are known[2,3] about Chern classes for fixed k; we are unaware of where these congruences between different Chern classes have been discussed. For convenience we tabulate the number q for various low k and k' in Table I. The number theoretic calculation is discussed in Ref. 4. The first line of this table displays the denominators of the Bernoulli numbers when k' is even.

We remark that similar congruences exist for the other

[a)] Address after 1 January 1988: Department of Mathematical Sciences, University of Durham, Durham DH1 3LE, England.

0022-2488/88/112494-05$02.50

TABLE I. The number q given by Theorem 1 for $k<k'<n$.

k/k'	2	3	4	5	6
2	∞	2	6	2	30
3		∞	12	12	2
4			∞	2	24
5				∞	2
6					∞

Lie groups but as yet we have not been able to determine these in general. Similarly, congruences exist out of the stable range. After some preliminaries (Sec. II) our approach is to give in Sec. III an explicit expression (Lemma 2) for $f_k(\mu)$ when $G=\mathrm{SU}(n)$ and hence $a_\pi(\mu)$ (Lemma 3). Although these expressions are of some interest to those wishing to calculate anomaly coefficients we present them here only in so far as they enable us to prove Theorem 1. A brief conclusion and list of outstanding problems is given in Sec. V.

The above theorem has been used to show the global anomaly freedom of certain G_2, SU(3), and SU(2) associated vector bundles over S^6. For a description of these and related matters see Refs. 4 and 5.

A helpful example to end this Introduction with is that of SU(2). Although $k,k'\geqslant 2$ for this example corresponds to being in the unstable range for which Theorem 1 is not directly applicable, we can nonetheless see the appearance of the above mentioned congruences very simply. Let $\mu = n\lambda_1$ be a nonzero weight, then we easily see

$$\begin{aligned} f_k(n\lambda_1) &= (n^k/2)[1+(-1)^k]\lambda_1^k \\ &= n^k f_k(\lambda_1) = n^k[f_2(\lambda_1)]^{k/2} \end{aligned}$$

and this vanishes unless k is even. In particular

$$a_{(2^{k/2})}(n\lambda_1) = n^k,$$

$$A_{(2^{k/2})}(n\lambda_1) = 2^{-k/2}\sum_{j=0}^{n}(n-2j)^k.$$

By requiring that $a_{(2^{k/2})}(\mu)\equiv a_{(2^{k'/2})}(\mu) \mod q$ holds for all $\mu = n\lambda_1$ we obtain that $n^k \equiv n^{k'} \mod q$ for all $n\in Z$, analogous to (i) of Theorem 1. This example also shows it is often easier to deal with the reduced coefficients a_π rather than A_π.

II. PRELIMINARIES

To be more precise we must introduce some notation. For this we will largely follow Ref. 1. Let Φ be a root system of the rank l Lie algebra L and denote a basis of Φ by Δ. Fix an ordering of the simple roots $\{\alpha_1,\ldots,\alpha_l\}$ of Δ and let $\{\lambda_1,\ldots,\lambda_l\}$ be the dual basis of the weight lattice Λ with respect to Φ. Denote by Λ^+ the set of dominant weights. Now let V be an L-module and denote by $\pi(V)$ the set of all of its weights. We will be working with finite-dimensional modules $V=V(\lambda)$ of highest weight λ; in this case will also denote $\pi(V)=\pi(\lambda)$ and set $\pi^+(\lambda)=\pi(\lambda)\cap\Lambda^+$. Let $m_\lambda(\mu)$ be the multiplicity of the weight μ in $V(\lambda)$ and denote by Stab μ the stabilizer of the weight μ in W. Because W is a finite reflection group, $P(H)^{\mathrm{W}}$ is finitely generated with l generators of known degrees m_j. Let $d(L)=\{m_j: j=1,\ldots,l\}$. It will be convenient to introduce the polynomials $f_k(\mu)$ in $P(H)^{\mathrm{W}}$ defined for each $\mu\in\Lambda$ and $k\in Z^+$ by

$$f_k(\mu)=\frac{1}{|\mathrm{Stab}\,\mu|}\sum_{\sigma\in\mathrm{W}}(\sigma\mu)^k. \tag{1}$$

This is clearly a homogeneous polynomial of degree k in the λ, with integer coefficients. It suffices to take $\mu\in\Lambda^+$. (This polynomial is referred to as $\mathrm{Sym}\,\mu^k$ in Ref. 1.) If ϕ: $L\to gl[V(\lambda)]$ is a representation of L we will denote the trace polynomial (for $x\in L$) $\mathrm{Tr}[\phi(x)^k]$ by $\mathrm{Tr}_\lambda\, x^k$. We note the trace polynomials generate $P(L)^G$. For $h\in H$ we then have

$$\mathrm{Tr}_\lambda\, h^k=\sum_{\mu\in\Pi^+(\lambda)} m_\lambda(\mu)f_k(\mu)[h]. \tag{2}$$

Let

$$\pi=(\gamma_1,\gamma_2,\ldots,\gamma_q)=(p_1{}^{\pi_1},\ldots,p_r{}^{\pi_r}),$$

where

$$\gamma_1\geqslant\gamma_2\geqslant\cdots\geqslant\gamma_q>0,\quad p_1>p_2>\cdots>p_r>0,$$

be a partition of length $l(\pi)=q$ and weight $|\pi|=\gamma_1+\cdots+\gamma_q$. For each partition π, $x\in L$ and $\mu\in\Lambda^+$, we define polynomials $P_\pi(\mu,x)$ by

$$P_\pi(\mu,x)=(\mathrm{Tr}_\mu\, x^{p_1})^{\pi_1}\cdots(\mathrm{Tr}_\mu\, x^{p_r})^{\pi_r}. \tag{3}$$

Finally let us call $F(L)$ the set of those $\mu\in\Lambda^+$ such that $\{f_k(\mu),\, k\in d(L)\}$ constitutes a Q basis of $P(H)^{\mathrm{W}}$.

For a given $\nu\in F(L)$ our first problem becomes that of computing the coefficients $A_\pi(\lambda,\nu)$ in

$$\mathrm{Tr}_\lambda\, x^k=\sum_{|\pi|=k}A_\pi(\lambda,\nu)P_\pi(\nu,x). \tag{4}$$

Clearly these basis coefficients $A_\pi(\lambda,\nu)$ are determined by their restriction to H. Although in principle known, efficient algorithms for their computation remain another matter. One approach to their evaluation is via the Weyl character formula. Rather than following this we note that if we are able to express the Weyl-invariant polynomials $f_k(\mu)$ in terms of the basis constructed from $\{f_{m_j}(\nu): j=1,\ldots,l\}$ we can reconstruct the coefficients $A_\pi(\mu,\nu)$. That is, if

$$f_k(\mu)=\sum_{|\pi|=k}a_\pi(\mu,\nu)\prod_i(f_{p_i}(\nu))^{\pi_i}, \tag{5}$$

then

$$A_\pi(\lambda,\nu)=\sum_{\mu\in\pi^+(\lambda)}m_\lambda(\mu)a_\pi(\mu,\nu). \tag{6}$$

This follows by restricting our attention to the weight spaces appearing in (4). We note further that because $m_\lambda(\mu)$ may be viewed as a lower triangular matrix with integer entries and $m_\lambda(\lambda)=1$ this may be inverted over Z to yield

$$a_\pi(\mu,\nu)=\sum m_\lambda(\mu)^{-1}A_\pi(\lambda,\nu). \tag{7}$$

If both $a_\pi(\lambda,\nu)$ and $A_\pi(\lambda,\nu)$ are integral then this last remark provides us with Lemma 1.

In what ensues we will choose a basis such that $a_\pi(\lambda,\nu)$ and $A_\pi(\lambda,\nu)$ are integral.

The approach we adopt here is to calculate $A_\pi(\lambda,\nu)$ by first calculating $a_\pi(\mu,\nu)$. This method depends on knowing the multiplicities $m_\lambda(\mu)$; these may be evaluated by several known recursions[1] and have been extensively tabulated.[6]

III. SU(n)

We will now fix our attention on the group $G = \mathrm{SU}(n)$ and find for a particular basis the coefficients we have described in general in Sec. II. Here we view the weight space of L as the subspace of R^n orthogonal to $\eta = x_1 + x_2 + \cdots + x_n$, where $\{x_1,\ldots,x_n\}$ form an orthonormal basis of R^n. The Weyl group acts naturally via

$$\sigma_i(x_j) = \begin{cases} x_{i+1}, & i \equiv j, \\ x_i, & i = j-1, \\ x_j, & i \neq j, j-1, \end{cases} \tag{8}$$

and we have $\Phi = \{\alpha_i\colon \alpha_i = x_i - x_{i+1},\ 1 \leqslant i \leqslant n-1\}$. The elementary weights are (with $\bar\eta = \eta/n$)

$$\lambda_k = x_l + \cdots + x_k - k\bar\eta. \tag{9}$$

We may associate to any dominant weight μ (and hence irreducible representation) a Young diagram as follows. Let

$$\mu = \sum_{i=1}^{n-1} n_i\lambda_i = \sum_{j=1}^{n-1} l_j x_j - c(\mu)\bar\eta, \tag{10}$$

with

$$l_j = \sum_{i \leqslant j} n_i, \quad c(\mu) = \sum_{j=1}^{n-1} j n_j.$$

Then μ is associated with the diagram whose jth row has length l_j.

With this notation we find $f_k(\lambda_1)$ defined by (1) to be

$$\begin{aligned} f_k(\lambda_1) &= \lambda_1{}^k + (\lambda_1 - \lambda_2)^k + \cdots + (\lambda_n - \lambda_{n-1})^k \\ &\quad + (-\lambda_n)^k \\ &= x_1{}^k + (x_2 - \bar\eta)^k + \cdots + (x_n - \bar\eta)^k. \end{aligned}$$

Since we will be dealing with $L = \mathrm{SU}(n)$, i.e., the $\bar\eta = 0$ subspace of R^n, we will simply set $\bar\eta = 0$ in the above and identify $f_k(\lambda_1)$ with the power symmetric functions $s_k = \Sigma_{i=1}^n x_i{}^k$.

Because of the homotopy equivalence between $\mathrm{GL}(n,C)$ and $\mathrm{U}(n)$ the basis $\{f_k(\lambda_1),\ 2 \leqslant k \leqslant n\}$ for $P(L)^G$ is very convenient for actual calculations. In particular we shall see the coefficients that appear in (5) are integers. Our intent in the remainder of this section is to evaluate these coefficients $a_\pi(\mu) \equiv a_\pi(\mu,\lambda_1)$.

For the case at hand, $P(L)^G$ is the graded ring of symmetric polynomials with integer coefficients, $P(L)^G = Z[x_1,\ldots,x_n]^{S_n}$. We wish to express the homogeneous symmetric polynomial $f_k(\mu)$ in terms of (the *a priori* Q basis) power sum symmetric functions. To do this, first recall[7] the definition of the symmetric monomials m_π and augmented symmetric monomials n_π associated with the partition π [of length $l(\pi) \leqslant n$],

$$m_\pi(x_1,x_2,\ldots,x_n) = \sum_{\substack{\text{distinct} \\ \text{selections}}} x_1{}^{\gamma_1}\cdots x_k{}^{\gamma_k},$$

$$n_\pi(x_1,x_2,\ldots,x_n) = \sum_{\substack{\text{all} \\ \text{selections}}} x_1{}^{\gamma_1}\cdots x_k{}^{\gamma_k}.$$

If we define $g(\pi) = \pi_1!\cdots\pi_r!$, then

$$n_\pi(x) = g(\pi) m_\pi(x).$$

Thus $s_r = m_{(r)}$ and we let $s_\pi = \Pi_i s_{\gamma_i}$.

Let us now write

$$\mu = \sum_{j=1}^{n} l_j x_j, \tag{11}$$

where l_j can be nonzero for $\mathrm{U}(n)$. Using the multinomial expansion we obtain

$$\mu^k = \sum_{|\pi| = k} \frac{k!}{\gamma_1!\cdots\gamma_k!} m_\pi(l_1x_1,\ldots,l_nx_n).$$

Utilizing elementary properties of symmetric functions we may then show [with $(l) = (l_1,l_2,\ldots,l_n)$, etc.] the following lemma.

Lemma 2:

$$f_k(\mu) = \frac{|\mathrm{Orbit}\,\mu|}{n!} \sum_{|\pi| = k} \frac{k!(n - l(\pi))!}{\gamma_1!\cdots\gamma_k!} m_\pi(l) n_\pi(x). \tag{12}$$

Finally let $n_{\pi\pi'}$ be the transition functions between the augmented symmetric monomials $n_\pi(x)$ and the symmetric powers $s_{\pi'}$:

$$n_\pi(x) = \sum_{\pi'} n_{\pi\pi'} s_{\pi'}. \tag{13}$$

(The tabulation and calculation of these transition functions is treated, for example, in Ref. 7.) Utilizing (5), (12), and (13) we obtain the following lemma.

Lemma 3: Let $k \leqslant n$. Then

$$f_k(\mu) = \sum_{|\pi'| = k} a_{\pi'}(\mu) s_{\pi'}, \tag{14}$$

where

$$a_{\pi'}(\mu) = \frac{|\mathrm{Orbit}\,\mu|}{n!} \sum_{|\pi| = k} n_{\pi\pi'} \frac{k!(n - l(\pi))!}{\gamma_1!\cdots\gamma_k!} m_\pi(l) \tag{15}$$

and $a_{\pi'}(\mu) \in Z$.

Remarks: (a) The constraint $k \leqslant n$ corresponds to being in the stable range for $\mathrm{SU}(n)$'s homotopy. Here it means we can use s_π as a basis for $P(L)^G$ when $l(\pi) \leqslant n$. When $l(\pi) > n$ there are further relations amongst the monomials to be accounted for.

(b) Suppose the Young diagram corresponding to μ has only r rows; then $m_\pi(l) = 0$ whenever $r < l(\pi)$.

(c) For $\mathrm{SU}(n)$ we have $s_1 = 0$ and $a_\pi = 0$ whenever $l_n \geqslant 1$.

Lemma 3 provides us with the explicit coordinates of $f_k(\mu)$ for the choice of basis $\{f_k(\lambda_1),\ 2 \leqslant k \leqslant n\}$ and $k \leqslant n$. We will now specialize (15) to the partition $\pi' = (k)$. (In the physics literature we are calculating the "leading anomaly.") In this case we have the particularly simple identity

$$n_{(k)\pi} = (-1)^{l(\pi)-1}(l(\pi)-1)! \tag{16}$$

and we may rewrite (15) as

$$a_{(k)}(\mu) = \frac{|\text{Orbit}\,\mu|}{n!}\sum_{j=1}^{k}(-1)^{j-1}(j-1)!(n-j)! \times F_N(k,j)[\mu], \tag{17}$$

where we introduce the symmetric polynomial $F_N(k,j)$ in N variables defined by

$$F_N(k,j) = \sum_{\substack{|\pi|=k\\ l(\pi)=j}} \frac{k!}{\gamma_1!\cdots\gamma_k!}\, m_\pi. \tag{18}$$

Here m_π is the symmetric monomial $m_\pi(x_1,...,x_N)$. If $j>N$ no partition of $\{x_1,...,x_N\}$ exists with $l(\pi)=j$ and in this case we take $F_N(k,j)=0$. Similarly $F_N(k,j)=0$ if $j>k$. By (b), the value of N is only required to be larger than the number of rows of the Young diagram associated with μ. An alternate description of $F_N(k,j)$ will prove useful. Straightforward use of combinatorics yields

$$F_N(k,j) = \sum_{p=0}^{j-1}(-1)^p\binom{N-j+p}{p} \times \sum (x_1+\cdots+x_{j-p})^k, \tag{19}$$

where the second summation is over distinct selections of $j-p$ variables from $\{x_1,...,x_N\}$. If $j>N$, then $F_N(k,j)=0$ is manifest; when $j>k$, (19) also gives $F_N(k,j)=0$.

We note with the form (19) it is easy to calculate $a_{(k)}(\mu)$ and thus $A_{(k)}(\mu)$. In particular we obtain for the elementary representations $\lambda_r = (1^r)$.

Lemma 4:

$$a_{(k)}(1^r) = \sum_{j=1}(-1)^{j-1}\binom{n-j}{r-j}S(k,j). \tag{20}$$

Proof: This follows from

$$F_N(k,j)[1^r] = j!\binom{r}{j}S(k,j), \tag{21}$$

where $S(k,j)$ are the ubiquitous Stirling numbers of the second kind, the number of ways of putting k distinct objects into j identical boxes allowing no box to be empty.

Using the recurrence relations of the Stirling numbers and the relation with the Eulerian numbers

$$a_{n,l} = \sum_k (-1)^{k-l}\binom{k}{l}(n-k)!S(n,n-k),$$

we may reexpress (20) in the following equivalent forms:

$$a_{(k)}(1^r) = \sum_{l=0}^{k-1}(-1)^{l-1}l!\binom{n-1-l}{r-1-l}S(k-1,l) \tag{22a}$$

$$= \sum_{l=0}(-1)^{r+l-1}(r-l)^{k-1}\binom{n}{l} \tag{22b}$$

$$= \sum_{s=0}(-1)^s\binom{n-k}{r-1-s}a_{k-1,s}. \tag{22c}$$

Remark: Expressions (22b) and (22c) are given in Refs. 8 and 9 and (22c) in given in Ref. 10.

IV. PROOF OF THEOREM 1

The necessity part of Theorem 1 is easy and was mentioned in Ref. 4. By homogeneity $a_{(k)}(m) = m^k$, where $\mu = m\lambda_1$ corresponds to the Young diagram of one row with length m. Likewise, the necessity of (ii) follows from expression (22b) for $a_{(k)}(1^r)$ by induction on $r<n$. Because we are considering SU(n) then $a_{(k)}(1^n)=0$ and so we only have restrictions on primes less than n.

Showing the sufficiency of conditions (i) and (ii) is a little lengthier. Let $\mu = (l_1,l_2,...,l_r) = (a_1^{\mu_1},...,a_s^{\mu_s})$ be the weight whose diagram has row lengths l_j. Then

$$|\text{Orbit}\,\mu| = \frac{n!}{(n-r)!}\frac{1}{\mu_1!\cdots\mu_s!}. \tag{23}$$

We will show $a_{(k)}(\mu) \equiv a_{(k')}(\mu) \bmod q$ by showing they agree term by term when expressed as a sum. First let $M = \max\{k,k'\}$. Because $F_r(k,j)=0$ when $j>k$, we then have, from (17) and (18),

$$a_{(k)}(\mu) = \sum_{j=1}^{M}\sum_{p=0}^{j-1}(-1)^{p+j-1}N(p,j,k)[\mu], \tag{24}$$

where

$$N(p,j,k)[\mu] = \frac{(j-1)!(n-j)!}{(n-r)!\mu_1!\cdots\mu_s!}\binom{r-j+p}{p} \times \sum (l_1+\cdots+l_{j-p})^k. \tag{25}$$

Likewise $a_{(k')}(\mu)$ differs only by replacing k with k' in the summand.

We remark in passing that when $\mu_i = 1$ ($1 \leqslant i \leqslant s$) condition (i) means $N(p,j,k)$ and $N(p,j,k')$ are equivalent (mod q); thus additional constraints are required only for diagrams possessing rows of equal lengths.

To proceed we note

$$\sum_{\substack{\text{distinct}\\ \text{selections}}}(l_1+\cdots+l_{j-p})^k = \sum_{\substack{m_i\\ \Sigma m_i = j-p}}\binom{\mu_1}{m_1}\cdots\binom{\mu_s}{m_s}(m_1a_1+\cdots+m_sa_s)^k.$$

Upon using $\Sigma\mu_i = r$ and $\Sigma(\mu_i - m_i) = r-j+p$ we may combine the several factorials in $N(p,j,k)[\mu]$ to give

$$N(p,j,k)[\mu] = \sum_{\substack{m_i\\ \Sigma m_i = j-p}}\frac{1}{j}\binom{j}{p}\binom{n-j}{r-j} \times \binom{j-p \qquad\qquad r-j+p}{m_1\cdots m_s \quad \mu_1-m_1\cdots\mu_s-m_s} \times (m_1a_1+\cdots+m_sa_s)^k.$$

Two possibilities now occur: either j divides the factorials multiplying $(m_1a_1 + \cdots + m_sa_s)^k$ or j divides this number. In the first instance condition (i) is sufficient to guarantee that expressions (24) for $a_{(k)}(\mu)$ and $a_{(k')}(\mu)$ agree $(\bmod q)$ term by term. In the second instance, noting $1 \leqslant j \leqslant n$ and by considering the prime factors of both j and $m_1a_1 + \cdots + m_sa_s$ condition (ii) again gives us equality term by term. Actually, because we are considering $\mathrm{SU}(n)$ rather than $\mathrm{U}(n)$ we need only consider primes $p < n$ by remark (c) following Lemma 3. We have then the sufficiency of conditions (i) and (ii).

V. DISCUSSION

In this paper we have proved a theorem relating the leading anomaly coefficients or Chern classes of $\mathrm{SU}(n)$ vector bundles over spheres; in particular, the congruences for the stable range have been found. The utility of these congruences in the calculation of global anomalies and index calculations has been discussed in Ref. 5. Indeed, as the study of global anomalies[11,12] often reduces to calculating an index mod m (for some integer m) it suffices to know this number is congruent mod q (where $m|q$) to some other more readily computed quantity such as the second Dynkin index. This was the strategy advocated in Refs. 4 and 5 and we have shown here how q may be determined in general for $\mathrm{SU}(n)$ in the stable range.

Several questions have been raised by this work and we conclude by mentioning them. First, in reference to the extension to arbitrary Lie groups G, similar congruences exist both in and out of the stable ranges for arbitrary G: is there a simpler (perhaps cohomological) prescription for finding these? Second, we have left unanswered the nature of the set $F(L)$ of generators of $P(H)^W$ described in Sec. II; can one give a simple description of the exceptional representations not included in this set?

ACKNOWLEDGMENTS

I have enjoyed many discussions with J. Stasheff.

This work was partially supported by the U.S. Department of Energy Grant No. DE-FG05-85ER-40219.

[1] J. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, New York, 1979).

[2] R. Bott, "The space of loops on a Lie group," Mich. Math. J. **5**, 35 (1958).

[3] R. Bott, "The stable homotopy of the classical groups," Ann. Math. **70**, 313 (1959).

[4] H. W. Braden, "Anomaly coefficients: Their calculation and congruences," J. Math. Phys. **29**, 727 (1988).

[5] H. W. Braden, "On the absence of global anomalies," Phys. Lett. B **200**, 91 (1988).

[6] M. R. Brenner, R. V. Moody, and J. Patera, *Tables of Dominant Weight Multiplicities for Representations of Simple Lie Algebras* (Marcel Dekker, New York, 1985).

[7] F. N. David, M. G. Kendall, and D. E. Barton, *Symmetric Function and Allied Tables* (Cambridge U.P., Cambridge, 1966).

[8] P. H. Frampton and T. W. Kephart, "Analysis of anomalies in higher space-time dimensions," Phys. Rev. D **28**, 1010 (1983).

[9] P. H. Frampton and T. W. Kephart, "Explicit evaluation of anomalies in higher dimensions," Phys. Rev. Lett. **50**, 1347 (1983).

[10] S. Okubo and J. Patera, "General indices of simple Lie algebras and symmetrized product representations," J. Math. Phys. **24**, 2722 (1983).

[11] E. Witten, "An SU(2) anomaly," Phys. Lett. B **11**, 324 (1982).

[12] S. Elitzur and V. P. Nair, "Nonperturbative anomalies in higher dimensions," Nucl. Phys. B. **243**, 205 (1984).

Supersymmetric extension of the Korteweg–de Vries equation

Pierre Mathieu
Département de Physique, Université Laval, Québec G1K 7P4, Canada

(Received 17 September 1987; accepted for publication 13 July 1988)

It is shown that among a one-parameter family of supersymmetric extensions of the Korteweg–de Vries equation, there is a special system that has an infinite number of conservation laws, which can be formulated in the second Hamiltonian structure, and which has a nontrivial Lax representation. Its modified version is also discussed.

I. INTRODUCTION

This work is concerned with the study of the supersymmetric Korteweg–de Vries (KdV) equation:

$$u_t = -u''' + 6uu'. \tag{1.1}$$

The KdV equation is a completely integrable nonlinear evolution equation for a bosonic (or equivalently a commuting) field.[1] Its supersymmetric extension refers to a system of coupled equations for a bosonic and a fermionic field which reduces to the KdV equation in the limit where the fermionic field is set equal to zero. In a classical context, a fermionic field is described by an anticommuting field. However, supersymmetry entails more than a mere coupling of a bosonic field to a fermionic field. It also implies the existence of a transformation relating these two fields that leaves the system invariant. Roughly speaking, the generator of a space and/or time supersymmetry transformation is the square root of the generator of space and/or time translation. Here one considers merely the space supersymmetric extension of the KdV equation: a double application of the supersymmetry generator is proportional to ∂_x.

By using superspace and superfield techniques, one constructs a general (space) supersymmetric extension of the KdV equation, which turns out to contain one free parameter. The integrability of the resulting system is then studied from the point of view of integrable deformations and the Lax representation. The Hamiltonian structure of the generic system is also analyzed, which is particularly relevant in view of the relation found in Ref. 2 between the super-Virasoro algebra and a particular form of the supersymmetric KdV equation. These analyses are preceded by a brief survey of some relevant results on the KdV equation.

II. SURVEY OF SELECTED RESULTS ON THE KdV EQUATION

This section collects a certain number of well-known results on the KdV equation, which will be generalized to the supersymmetric case in the following sections.

An important early discovery on the study of the KdV equation (1.1) is that this equation can be formulated under the Lax representation[3,4]

$$L_t = [P_3, L], \tag{2.1}$$

where

$$L = \partial_x^2 - u, \tag{2.2a}$$

$$P_3 = -4\,\partial_x^3 + 6u\,\partial_x + 3u' = -4\,L_+^{3/2}. \tag{2.2b}$$

(The square root of a differential operator is, in general, a pseudodifferential operator, involving negative powers of ∂_x, whose differential part is denoted by $+$.) The Lax representation implies that the Schrödinger-like equation

$$Ly = \lambda y \tag{2.3}$$

is isospectral (i.e., $\lambda_t = 0$) when u satisfies the KdV equation.[5] The linear eigenvalue problem (2.3) was first reached by a linearization of the Miura transformation[6]

$$u = v' + v^2, \tag{2.4}$$

which maps a solution of the modified KdV (mKdV) equation

$$v_t = -v''' + 6v^2v' \tag{2.5}$$

into a solution of the KdV equation. The linearization is achieved by means of the Cole–Hopf transformation,

$$v = \partial_x \ln y, \tag{2.6}$$

which yields $(\partial_x^2 - u)y = 0$. The substitution $\partial_x{}^2 \rightarrow \partial_x{}^2 - \lambda$ leads to (2.3).[7]

A generalization of the Miura transformation was proposed by Gardner as the starting point of a simple constructive proof of the existence of an infinite number of conservation laws for the KdV equation. A conservation law is an expression of the form $H_n = \int dx\, h_n$, where h_n is a polynomial of u and its space derivatives, which is such that $\partial_t h_n$ is a total derivative. The Gardner transformation is

$$u = w + \varepsilon w' + \varepsilon^2 w^2 \tag{2.7}$$

and it maps a solution of the Gardner equation

$$w_t = -w''' + 6ww' + 6\varepsilon^2 w^2 w' \tag{2.8}$$

into a solution of the KdV equation. Since w_t can be written as a total derivative, the inversion of the map (2.7) with $\varepsilon \rightarrow 0$ yields

$$\int w\, dx = \sum_{n>0} \varepsilon^n \int dx\, h_n[u], \tag{2.9}$$

and the set $\{h_n\}$ provides an infinite number of conservation laws, out of which only those with n even are nontrivial.[8] (A trivial conservation law is such that h_n is a total derivative.) From (2.4) and (2.7) it follows that there is also an infinite number of conservation laws for both the mKdV equation and the Gardner equation.

It was observed that the conservation laws for the KdV equation are all homogeneous with respect to the grading $\deg(\partial_x^k u) = 2 + k$ and that there exists one and only one

0022-2488/88/112499-08$02.50

nontrivial conservation law, $\int h_{n-2}\, dx$, for each even value n of the degree.[9] The first few conservation laws are

$$\begin{aligned} H_0 &= \frac{1}{2}\int dx\, u, \quad H_2 = \frac{1}{2}\int dx\, u^2 , \\ H_4 &= \frac{1}{2}\int dx[(u')^2 + 2u^3] , \qquad (2.10) \\ H_6 &= \frac{1}{2}\int dx[(u'')^2 + 10u(u')^2 + 5u^4] . \end{aligned}$$

These can also be obtained as $\int dx$ Res $L^{(2n+1)/2}$ where Res picks up the coefficient of the ∂_x^{-1} term.[1,10]

Another important property of the KdV equation is that it can be formulated as a Hamiltonian system, that is, under the form $u_t = \{u, \mathscr{H}\}$. Actually this can be done in two different ways, either with the Poisson bracket

$$\{u(x),u(y)\} = \delta'(x-y) \quad (' = \partial_x) \qquad (2.11)$$

and the Hamiltonian $\mathscr{H}^{(1)} = H_4$—the first Hamiltonian structure[11]—or with the Poisson bracket

$$\begin{aligned} \{u(x),u(y)\} = &-\delta'''(x-y) + 4u(x)\delta'(x-y) \\ &+ 2u'(x)\delta(x-y) \qquad (2.12) \end{aligned}$$

and the Hamiltonian $\mathscr{H}^{(2)} = H_2$—the second Hamiltonian structure.[12] The existence of the second Hamiltonian structure can be traced back to the existence of both the natural Hamiltonian structure for the mKdV equation and the Miura map.[13] For the mKdV equation, the natural Hamiltonian structure is defined by the Poisson bracket

$$\{v(x),v(y)\} = \delta'(x-y) \qquad (2.13)$$

and the Hamiltonian $\mathscr{H}^{(2)}$ with u replaced by (2.4). Then the existence of the second Hamiltonian structure can be viewed as being a consequence of the identity

$$\mathscr{D}\,\partial_x \mathscr{D}^\dagger = -\partial_x^3 + 4u\,\partial_x + 2u' , \qquad (2.14)$$

where $\mathscr{D}$ is the Fréchet derivative of u with respect to v,

$$\mathscr{D} = \partial_x + 2v \qquad (2.15)$$

and $\mathscr{D}^\dagger$ is its adjoint, $-\partial_x + 2v$. On the other hand, given the existence of the second Hamiltonian structure for the KdV equation, the identity (2.14) can be interpreted as expressing the canonical character of the Miura map when the mKdV and the KdV equation are formulated, respectively, in the natural and the second Hamiltonian structure. The map is canonical in the sense that it preserves the structure of the corresponding Poisson brackets.

Another important result is that the conservation laws are all in involution[11] (i.e., $\{H_{2n}, H_{2m}\} = 0$) with respect to both Hamiltonian structures. This suggests a relation between an action-angle formalism, a relation that has actually been found by Zakharov and Faddeev.[14] Hence the KdV equation is a completely integrable Hamiltonian system.

Finally, one recalls that the KdV equation is the first nontrivial equation among an infinite set of equations, the KdV hierarchy, which can be defined either via the Lax representation (2.1) with P_3 replaced by $P_{2k+1} = (L^{(2k+1)/2})_+$ or as $u_t = \{u, H_{2m}\}$, where the Poisson bracket can be either (2.11) or (2.12).

III. SUPERSYMMETRIC KdV EQUATION AND HAMILTONIAN STRUCTURES

One now turns to an explicit construction of an extension of the KdV equation to a system invariant under a supersymmetric transformation. As mentioned previously, only a space supersymmetric invariance is considered. In a superspace approach the variable x is then extended to a doublet (x,θ), where θ is an anticommuting variable $(\theta^2 = 0)$. The new system will be constructed in a way that will make it invariant under the transformation $x \to x - \eta\theta$ and $\theta \to \theta + \eta$, where η is an anticommuting parameter. This transformation is generated by the operator

$$Q = \partial_\theta - \theta\,\partial_x . \qquad (3.1)$$

In a superspace formalism, the usual field $u(x)$ is replaced by a superfield that has a further dependence upon θ and is written as $\Phi(x,\theta)$. (The time dependence is implicit everywhere.) This superfield is chosen to be fermionic. This is required in order to have a nontrivial extension of the KdV equation—see below. The expansion of $\Phi(x,\theta)$ in terms of θ yields

$$\Phi(x,\theta) = \xi(x) + \theta u(x) , \qquad (3.2)$$

where $\xi(x)$ is an anticommuting field. Under a supersymmetry transformation $\delta\Phi = \eta Q\Phi$, or equivalently

$$\delta u(x) = \eta\xi'(x), \qquad (3.3a)$$

$$\delta\xi(x) = \eta u(x). \qquad (3.3b)$$

In addition to the superfield, one introduces the covariant derivative

$$D = \theta\,\partial_x + \partial_\theta, \qquad (3.4)$$

which has the property of anticommuting with Q. (Notice also that $D^2 = \partial_x$.) Expressions written in terms of the covariant derivative and the superfield are manifestly invariant under the supersymmetry transformation (3.3).

A natural way to extend the KdV equation to a supersymmetric system is simply to rewrite the KdV equation in terms of the superfield and the covariant derivative. In other words, one multiplies the KdV equation by θ and supersymmetrizes the result. For instance, θu_t is then transformed into Φ_t. However, such a superfield generalization is not always unique: indeed $\theta(6uu')$ is the fermionic part [when $\xi(x) = 0$] of either $3D^2(\Phi D\,\Phi)$ or $6D\Phi D^2\Phi$. Therefore, in order to construct the most general possible supersymmetric extension of the KdV equation, one must consider a linear combination of all the possible terms. This leads to the one-parameter family of equations

$$\Phi_t = -D^6\Phi + aD^2(\Phi D\,\Phi) + (6-2a)D\Phi D^2\Phi , \qquad (3.5)$$

which will be referred to as the sKdV-a equation. In terms of the component fields, (3.5) is equivalent to

$$u_t = -u''' + 6uu' - a\xi\xi'' , \qquad (3.6a)$$

$$\xi_t = -\xi''' + (6-a)\xi'u + a\xi u' . \qquad (3.6b)$$

This system represents the most general nontrivial supersymmetric extension (with only one fermionic field and with no inverse power of the fields) of the KdV equation. [The case $a = 0$ represents an instance where the supersymmetric extension is trivial. Another trivial case is obtained if, instead of a fermionic superfield, one considers a bosonic superfield of the form $\Omega(x,\theta) = u(x) + \theta\xi(x)$. The superfield extension of the KdV equation would then be $\Omega_t = -D^6\Omega + 6\Omega D^2\Omega$, or equivalently $u_t = -u''' + 6uu'$ and $\xi_t = -\xi''' + 6(\xi u)'$. The triviality in these examples comes from the fact that the resulting system is linear in the fermionic field and, as a result, one of the two equations is simply the KdV equation itself.]

Let us see whether (3.5) can be formulated as a Hamiltonian system. To investigate this point, consider first the supersymmetric extension of the first Hamiltonian structure. The extension of the corresponding Hamiltonian is

$$\bar{\mathscr{H}}^{(1)} = \frac{1}{2}\int [2\Phi(D\Phi)^2 + D^2\Phi D^3\Phi]\, d\mu, \tag{3.7}$$

where $d\mu = dx\, d\theta$, $\int d\theta = 0$, and $\int \theta\, d\theta = 1$. (A bar over a quantity will be used to indicate its supersymmetric generalization.) It is easily verified that $\bar{\mathscr{H}}^{(1)}$ is a conserved quantity for the sKdV-a equation. [As usual the expression for a conserved quantity is unique modulo total derivative terms. Notice that here such terms are of the form $D(\cdot)$, which contains as a special case the form $D^2(\cdot)$. For definiteness, one assumes that the fields vanish sufficiently fast at infinity or are periodic functions of x.] The superfield version of the Poisson bracket (2.11) is

$$\{\Phi_1,\Phi_2\} = -D_1\Delta, \tag{3.8}$$

where the subscript i refers to the set of variables (x_i,θ_i) [that is, $\Phi_1 = \Phi(x_1,\theta_1)$] and

$$\Delta = (\theta_1 - \theta_2)\delta(x_1 - x_2) \tag{3.9}$$

is the superdelta function, which satisfies

$$\int F(x_1,\theta_1)\Delta\, d\mu_1 = F(x_2,\theta_2)\,. \tag{3.10}$$

With (3.7) and (3.8) one finds

$$\Phi_t = \{\Phi,\bar{\mathscr{H}}^{(1)}\} = -D^6\Phi + 4D\Phi D^2\Phi + 2\Phi D^3\Phi\,. \tag{3.11}$$

Comparison of (3.11) with (3.5) shows that the first Hamiltonian structure exists only for the case $a = 2$.

On the other hand, the superfield extension of the Hamiltonian $\mathscr{H}^{(2)}$ is

$$\bar{\mathscr{H}}^{(2)} = \frac{1}{2}\int \Phi D\Phi\, d\mu \tag{3.12}$$

and it is also a conserved quantity for the sKdV-a equation. The Poisson bracket (2.12) generalizes to

$$\begin{aligned}\{\Phi_1,\Phi_2\} = {} & + D_1^5\,\Delta - c\Phi_1 D_1^2\,\Delta - (4-c)(D_1\Phi_1)D_1\Delta \\ & - 2(D_1^2\Phi_1)\Delta\,,\end{aligned} \tag{3.13}$$

where c is an undetermined constant. This constant can be fixed uniquely by imposing the Jacobi identity. This gives $c = 3$. For this value, (3.12) and (3.13) yield

$$\Phi_t = \{\Phi,\bar{\mathscr{H}}^{(2)}\} = -D^6\Phi + 3D^2(\Phi D\,\Phi)\,, \tag{3.14}$$

and this is equivalent to Eq. (3.5) for $a = 3$.

Therefore, there is no value of a for which the system (3.5) has a bi-Hamiltonian structure. Furthermore each possible Hamiltonian structure singles out a particular value of a. This is a first indication that the sKdV-a equation might not be completely integrable for all values of a. Also at this point one can expect that if there exists a supersymmetric extension of the Miura map, the value $a = 3$ should be singled out. This will be considered in the next section. On the other hand, it is shown in Appendix A that it is possible to construct a fermionic extension of the KdV equation that has a bi-Hamiltonian structure. However, the resulting system is not invariant under supersymmetry.

IV. SUPERSYMMETRIC MODIFIED KORTEWEG–de VRIES EQUATION AND THE SUPER-MIURA TRANSFORMATION

The simplest way of establishing the existence of an infinite number of conservation laws for the sKdV-a equation, presumably for only a few particular values of a, is to search for an integrable deformation generalizing the Gardner transformation. Since the Gardner transformation contains in a particular limit the Miura transformation (when $\varepsilon \to \infty$, with $\varepsilon w = v$), it is natural to consider first the supersymmetric extension of the latter. But for this the superextension of the mKdV equation is required. Introducing the fermionic superfield $\psi(x,\theta)$, defined in terms of field components as

$$\psi(x,\theta) = \theta v(x) + \xi(x) \tag{4.1}$$

[where $\xi(x)$ is an anticommuting field, the superpartner of $v(x)$], it is simple to see that the generic form of the supersymmetric mKdV equation is given by

$$\begin{aligned}\psi_t = {} & -D^6\psi + 3(D\psi)D^2(\psi D\psi) \\ & + (3-b)(D\psi)D(\psi D^2\psi)\,,\end{aligned} \tag{4.2}$$

where b is an arbitrary constant. This equation will be called the smKdV-b equation. On the other hand, the Miura transformation has a unique superfield extension, which is

$$\Phi = D^2\psi + \psi D\psi\,. \tag{4.3}$$

Consider now whether this transformation maps solutions of the smKdV-b equation into solutions of the smKdV-a equation for all or for some values of a and b. For this, the result of the application of the operator $\bar{\mathscr{D}}$, the Fréchet derivative of Φ with respect to ψ, which is given by

$$\bar{\mathscr{D}} = D^2 + \psi D + (D\psi)\,, \tag{4.4}$$

on the smKdV-b equation must be compared with (3.5) when Φ is replaced by (4.3). A straightforward calculation shows that the resulting two expressions are equal only if $a = b = 3$. Hence, as expected, the super-Miura transformation singles out the sKdV-3 equation. On the other hand, one can now expect that the smKdV-b equation has a supersymmetric natural Hamiltonian structure only for the case $b = 3$. The corresponding Hamiltonian is given by $\bar{\mathscr{H}}^{(2)}$ with Φ replaced by (4.3), that is,

$$\bar{\mathscr{H}}^{(N)} = \frac{1}{2}\int [\psi(D\psi)^3 + D^2\psi D^3\psi]d\mu\,. \tag{4.5}$$

With the Poisson bracket

$$\{\psi_1,\psi_2\} = -D_1\Delta\,, \tag{4.6}$$

it follows that

$$\psi_t = \{\psi,\bar{\mathscr{H}}^{(N)}\} = -D^6\psi + 3(D\psi)D^2(\psi D\psi)\,, \tag{4.7}$$

which indeed corresponds to the smKdV-3 equation. Actually $\overline{\mathscr{H}}^{(N)}$ is a conservation law only for this particular value of b. Finally a direct consequence of the identity

$$\overline{\mathscr{D}}D\overline{\mathscr{D}}^{\dagger} = -D^5 + 3\Phi D^2 + (D\Phi)D + 2(D^2\Phi)\,, \quad (4.8)$$

where $\overline{\mathscr{D}}^{\dagger}$ is the adjoint of $\overline{\mathscr{D}}$,

$$\overline{\mathscr{D}}^{\dagger} = -D^2 - \psi D + 2(D\psi)\,, \quad (4.9)$$

is that the super-Miura map is canonical.

From these results it already appears that the sKdV-3 and smKdV-3 equation are the natural supersymmetric generalizations of the corresponding bosonic equations in the sense that, for these specific cases, the supersymmetrization preserves most of the important properties of the original equations. This will be further confirmed in the following two sections.

For completeness, the sKdV-3 and the smKdV-3 equations, their corresponding Hamiltonian structures, and the super-Miura map are written explicitly in terms of the component fields.

For sKdV-3:

the evolution equations are

$$u_t = -u''' + 6uu' - 3\xi\xi'', \quad (4.10a)$$

$$\xi_t = -\xi''' + 3(\xi u)'; \quad (4.10b)$$

the Hamiltonian is

$$\overline{\mathscr{H}}^{(2)} = \frac{1}{2}\int dx(u^2 - \xi\xi'); \quad (4.11)$$

and the Poisson brackets are

$$\{u(x),u(y)\} = -\delta'''(x-y) + 4u(x)\delta'(x-y) + 2u'(x)\delta(x-y)\,, \quad (4.12a)$$

$$\{u(x),\xi(y)\} = 3\xi(x)\delta'(x-y) + \xi'(x)\delta(x-y), \quad (4.12b)$$

$$\{\xi(x),\xi(y)\} = \delta''(x-y) - u(x)\delta(x-y)\,. \quad (4.12c)$$

For smKdV-3:

the evolution equations are

$$v_t = -v''' + 6v^2v' - 3\zeta(v\zeta')', \quad (4.13a)$$

$$\zeta_t = -\zeta''' + 3v(v\zeta)'; \quad (4.13b)$$

the Hamiltonian is

$$\overline{\mathscr{H}}^{(N)} = \frac{1}{2}\int dx[(v')^2 + v^4 - \zeta'\zeta'' - 3v^2\zeta\zeta']; \quad (4.14)$$

and the Poisson brackets are

$$\{v(x),v(y)\} = \delta'(x-y), \quad (4.15a)$$

$$\{v(x),\zeta(y)\} = 0, \quad (4.15b)$$

$$\{\zeta(x),\zeta(y)\} = -\delta(x-y). \quad (4.15c)$$

Finally, the super-Miura transformation is

$$u = v' + v^2 - \zeta\zeta', \quad (4.16a)$$

$$\xi = \zeta' + v\zeta. \quad (4.16b)$$

After a suitable rescaling, Eq. (4.10) can be shown to be equivalent to the sKdV equation obtained in Ref. 15 by reduction from the super-Kadomtsev–Petviashvili hierarchy. However, they differ from the "super"-KdV equations of Kupershmidt[16] [cf. Eqs. (A5)], which are not supersymmetric.

V. SUPER-GARDNER TRANSFORMATION AND CONSERVATION LAWS

The superfield extension of the Gardner transformation (2.7) is

$$\Phi = \chi + \varepsilon D^2\chi + \varepsilon^2\chi D\chi, \quad (5.1)$$

where χ is a new fermionic superfield given by

$$\chi(x,\theta) = \theta w(x) + \sigma(x)\,. \quad (5.2)$$

It is straightforward to verify that this transformation maps a solution of the super-Gardner equation

$$\chi_t = -D^6\chi + 3D^2(\chi D\chi) + \epsilon^2 3(D\chi)D^2(\chi D\chi) \quad (5.3)$$

into a solution of the sKdV-3 equation. The form of the super-Gardner equation can be uniquely determined from the requirement that it must interpolate between the particular sKdV and smKdV equations related by the super-Miura transformation, a special case of (5.1). *En passant*, one notices that (5.3) does not have a natural Hamiltonian structure, with $\{\chi_1,\chi_2\} = -D_1\Delta$; this follows from the fact that for the sKdV-3 equation, the first Hamiltonian structure is absent (cf. the analysis of Ref. 17 for the bosonic case).

The expansion of (5.1) and (5.3) in terms of θ yields

$$u = w + \varepsilon w' + \varepsilon^2(w^2 - \sigma\sigma')\,, \quad (5.4a)$$

$$\xi = \sigma + \varepsilon\sigma' + \varepsilon^2\sigma w\,, \quad (5.4b)$$

and

$$w_t = -w''' + 6ww' - 3\sigma\sigma'' + \varepsilon^2[6w^2w' - 3\sigma(\sigma'w)']\,, \quad (5.5a)$$

$$\sigma_t = -\sigma''' + 3(\sigma w)' + \varepsilon^2[3w(w\sigma)']\,. \quad (5.5b)$$

Since $\sigma\sigma'' = (\sigma\sigma')'$ and $\sigma(\sigma'w)' = (\sigma\sigma'w)'$, w_t can be written as a total derivative so that $\int w\,dx$ is a conservation law. By inverting the super-Gardner map, one then finds

$$\int dx\, w = \sum_{n>0}\varepsilon^n\int dx\,\bar{h}_n[u,\xi] \quad (5.6)$$

and the quantities $\bar{h}_n$ yield an infinite number of conservation laws for the sKdV-3 equation. As a corollary, the smKdV-3 equation and the super-Gardner equation (5.3) also have an infinite number of conservation laws.

As in the bosonic case, the conservation laws with n odd are trivial. This can be seen as follows. First notice that $\int dx\,w = \int \chi\,d\mu$. Then one writes $\chi = \chi^+ + \chi^-$ with

$$\chi^+ = \sum_{n>0}\varepsilon^{2n}\Upsilon_{2n}\,, \quad (5.7a)$$

$$\chi^- = \sum_{n>0}\varepsilon^{2n+1}\Upsilon_{2n+1}\,, \quad (5.7b)$$

where Υ_n is a fermionic superfield whose θ part is $\bar{h}_n$. These expressions are then substituted into (5.1). Since the left-hand side is invariant under the change $\varepsilon \to -\varepsilon$, the part of the resulting equation that is odd in ε must vanish. This leads to

$$\chi^- = -\varepsilon^{-1} D \ln(1+\varepsilon^2 D\chi^+) - \varepsilon^2 (1+\varepsilon^2 D\chi^+)^{-1} \chi^+ D\chi^- , \tag{5.8}$$

which is solved iteratively. The first term is a total derivative while the second term generates an infinite number of terms, all of the form $\chi^+ (D\chi^+)^m D^3 \chi^+$. But these terms can also be written as total derivatives, i.e.,

$$(m+1)\chi^+ (D\chi^+)^m D^3\chi^+ = D\,[D(\chi^+ (D\chi^+)^{m+1}) - (m+2)^{-1} (D\chi^+)^{m+2}] . \tag{5.9}$$

This establishes the triviality of the conservation laws for n odd. For n even, one has

$$\Phi \simeq \chi^+ + \varepsilon^2 \chi^+ D\chi^+ + \cdots , \tag{5.10}$$

so that

$$\chi^+ \simeq \sum_{n>0} \varepsilon^{2n} (\Phi (D\Phi)^n + \cdots). \tag{5.11}$$

Since the first term cannot be written in the form of a total derivative, the conservation laws for n even are necessarily nontrivial. The explicit expression of the first few conservation laws is

$$\begin{aligned}
\overline{H}_0 &= \frac{1}{2}\int \Phi\, d\mu = \frac{1}{2}\int dx\, u ,\\
\overline{H}_2 &= \frac{1}{2}\int \Phi D\Phi\, d\mu = \frac{1}{2}\int dx (u^2 - \xi\xi') ,\\
\overline{H}_4 &= \frac{1}{2}\int (2\Phi(D\Phi)^2 + D^3\Phi D^2\Phi) d\mu\\
&= \frac{1}{2}\int dx (2u^3 + (u')^2 - \xi'\xi'' - 4u\xi\xi') ,\\
\overline{H}_6 &= \frac{1}{2}\int (5\Phi(D\Phi)^3 + 10\Phi(D^3\Phi)^2\\
&\quad - 6\Phi D^2\Phi D^4\Phi + D^4\Phi D^5\Phi) d\mu\\
&= \frac{1}{2}\int dx (5u^4 + 10u(u')^2 - (u'')^2 - \xi''\xi'''\\
&\quad + 2u\xi'\xi'' + 8u\xi\xi''' - 15u^2\xi\xi').
\end{aligned} \tag{5.12}$$

In principle these conservation laws can be obtained by the inversion of the super-Gardner transformation. However, this is certainly not the most efficient approach. At this point, the simplest way is to start from the bosonic version of the conservation laws and to supersymmetrize them. But since the superfield generalization may not be unique, some coefficients remain undetermined. Then one adds all the other possible distinct terms that have the same degree but vanish after the θ integration when $\xi = 0$. (For instance, this is the case for the third term in the superfield expression of $\overline{H}_6$.) Obviously the coefficient of these terms is also undetermined. Then the undetermined coefficients are fixed by the condition $d\overline{H}_n/dt = 0$, where the time derivative is eliminated by means of the field equation.[9] [In this context, two terms are distinct if they are not equal modulo a total derivative term. Also notice that $\deg D^k\Phi = \frac{3}{2} + k/2$ (and $\deg \theta = -\frac{1}{2}$).]

It should be pointed out that all the conservation laws obtained above are bosonic. It is natural to ask whether there are also fermionic conservation laws (with half-integer grading). However, notice that the super-Gardner deformation does not imply the existence of an infinite number of fermionic conservation laws because σ_t in (5.5b) cannot be written in the form of a total derivative. This is probably an indication that only a finite number of them exist. The only fermionic conservation law found for the sKdV-3 equation is $\int \xi\, dx$, although an extensive search for them has not been performed.

As a final comment on the conservation laws, we mention that the absence of a bi-Hamiltonian structure implies the absence of a supersymmetric version of the Lenard recursion formula.

VI. ASSOCIATED LINEAR PROBLEM

As in the bosonic case, one now tries to linearize the super-Miura transformation in order to relate the sKdV-3 equation to a linear eigenvalue problem and from them obtain a Lax representation. The component version of the super-Miura relation (4.16) can be linearized by means of the substitution

$$v = y'/y, \quad \zeta = \phi/y , \tag{6.1}$$

where ϕ is an anticommuting field. This leads to

$$(\partial_x^2 - u)y = -\xi\phi , \tag{6.2a}$$

$$\phi' = \xi y . \tag{6.2b}$$

Differentiation of the second equation followed by the substitution $\partial_x^2 \to \partial_x^2 - \lambda$ yield

$$(\partial_x^2 - u - \lambda)y = -\xi\phi , \tag{6.3a}$$

$$(\partial_x^2 - \lambda)\phi = (\xi y)' . \tag{6.3b}$$

This is our candidate for the linear eigenvalue problem associated to the sKdV-3 equation.

The superfield expression of (6.1) is

$$\psi = D \ln \Lambda , \tag{6.4}$$

where Λ is a bosonic superfield, whose expansion in terms of θ is

$$\Lambda(x,\theta) = y(x) + \theta\phi(x) . \tag{6.5}$$

Equation (6.4) is nothing but a super-Cole–Hopf transformation. In terms of Λ, Φ, and D, the linear eigenvalue problem (6.3) can be written in the form

$$\overline{L}\Lambda = \lambda\Lambda , \tag{6.6}$$

where

$$\overline{L} = D^4 - (D\Phi) + \Phi D . \tag{6.7a}$$

Hence $\overline{L}$ should be the first member of the Lax pair of the sKdV-3 equation. The second member should then be $\overline{P}_3 = -4(\overline{L})^{3/2}_{+}$. To calculate $\overline{P}_3$ one introduces the formal square root of $\overline{L}$, defined by the series expansion

$$J = (\overline{L})^{1/2} = \sum_{i=-\infty}^{2} C_{-i} D^i , \tag{6.8}$$

where the coefficients C_i are fermionic (resp. bosonic) if i is odd (resp. even). The rules for operating with D^{-n} follow directly from the formula[1,10]

$$\partial_x^{-n} f = \sum_{j=0}^{\infty} (-)^j \binom{n+j-1}{j} f^{(j)} \partial_x^{-n-j} \quad (n>0) \tag{6.9}$$

(where $f^{(j)} = \partial_x^j f$), and the identities $D^{-2n} = \partial_x^{-n}$ and $D^{-2n-1} = D\partial_x^{-n-1}$.

The coefficients C_i are determined by setting $J^2 = \bar{L}$. This gives, for the first few coefficients,

$$\begin{aligned} &C_{-2} = 1, \quad C_{-1} = C_0 = 0, \quad C_1 = \Phi/2, \\ &C_2 = -D\Phi/2, \quad C_3 = -D^2\Phi/4, \quad C_4 = D^3\Phi/4. \end{aligned} \tag{6.10}$$

With these values, one finds that $\bar{P}_3 = -4(J\bar{L})_+$ is given by

$$\begin{aligned} \bar{P}_3 = &-4D^6 + 3(D^3\Phi) - 3(D^2\Phi)D \\ &+ 6(D\Phi)D^2 - 6\Phi D^3. \end{aligned} \tag{6.7b}$$

Then a straightforward calculation shows that the Lax equation

$$\bar{L}_t = [\bar{P}_3, \bar{L}] \tag{6.11}$$

is indeed equivalent to the sKdV-3 equation. Notice that the Lax pair can also be written in the form

$$\bar{L} = D\cdot(D^3 - \Phi), \tag{6.12a}$$

$$\bar{P}_3 = D\cdot(-4D^5 + 3D^2\cdot\Phi + 3\Phi D^2), \tag{6.12b}$$

where $D\cdot\Phi = (D\Phi) - \Phi D$. The expressions in parentheses in (6.12) are the direct superfield extensions of the corresponding bosonic operators [cf. Eq. (2.2)].

As in the bosonic case one can expect that the Lax representation will imply the existence of an infinite number of conservation laws that, intuitively, should be given by $\int \overline{\text{Res}}\, J^k\, d\mu$, where $\overline{\text{Res}}$ is the coefficient of the D^{-1} term. This is indeed the case and, since $d\mu$ and $\overline{\text{Res}}\, J^k$ are both fermionic, the resulting conservation laws are bosonic; in fact they agree with those obtained previously via the super-Gardner deformation. To prove that $\int \overline{\text{Res}}\, J^k\, d\mu$ is a conservation law, one first notices that (6.11) can be written in the form

$$J_t^k = [-4J_+^3, J^k]. \tag{6.13}$$

Therefore it follows that

$$\partial_t\, \overline{\text{Res}}\, J^k = \overline{\text{Res}}\, J_t^k = \overline{\text{Res}}[-4J_+^3, J^k]. \tag{6.14}$$

Then the result is a consequence of the general fact that in the expression of the commutator of two pseudodifferential operators, the coefficient of the D^{-1} term can be written under the form of a total superderivative. This is demonstrated by direct calculation, and the proof is contained in Ref. 15. In fact, at this point the present work makes contact with the one of Manin and Radul[15] on the super-Kadomtsev–Petviashvili (sKP) hierarchy. They constructed the sKP hierarchy via Lax equations for operators with superderivatives, and since the hierarchy contains the sKdV-3 equation as a special case, their results can be applied directly to the present case.

By a straightforward extension of the argument presented in Ref. 10, one can prove that the flows defined by the different $\bar{L}^{k/2}$ all commute with each other (see Appendix B). This automatically implies the involutivity of the conservation laws $\bar{H}_n$ with respect to the second Hamiltonian structure.

We now comment on the possibility of representing the sKdV-a equation for $a \neq 3$ by a Lax pair. To investigate this point one considers a generic Lax pair with the same structure as (6.7), but with each term multiplied by an undetermined coefficient. The consistency of the Lax equation fixes the possible values of these coefficients. In this way one finds only two possible solutions, namely the Lax pair (6.7), associated with the sKdV-3 equation,[18] and the Lax pair

$$\tilde{L} = D^4 - (D\Phi), \tag{6.15a}$$

$$\tilde{P}_3 = -4D^6 + 6(D\Phi)D^2 + 3(D^3\Phi) = -4\tilde{L}_+^{3/2}, \tag{6.15b}$$

which furnishes a representation of the sKdV-0 equation. However, the triviality of this system has already been pointed out. Here, it is reflected in the fact that in the expression for $\tilde{L}^{1/2}$, all the C_i's with i odd are identically zero. As a result, $\overline{\text{Res}}\, \tilde{L}^{k/2} = 0$ so that this particular Lax representation does not imply the existence of an infinite number of conservation laws, i.e., it is trivial. (However, notice that for the sKdV-0 equation there exists an infinite sequence of conservation laws: these are just the KdV conservation laws. But obviously they are not invariant under supersymmetry.) Therefore the sKdV-a system has a nontrivial Lax representation only for the case $a = 3$.

Finally notice that the operator $\bar{L}$ can be factorized under the form

$$\bar{L} = (D^2 + D\cdot\psi)(D^2 - D\cdot\psi), \tag{6.16}$$

where ψ is related to Φ by the super-Miura transformation (4.3). This factorization suggests that the smKdV-3 equation should be associated to the eigenvalue problem

$$(D^2 + D\cdot\psi)\Lambda_1 = \sqrt{\lambda}\,\Lambda_2, \tag{6.17a}$$

$$(D^2 - D\cdot\psi)\Lambda_2 = \sqrt{\lambda}\,\Lambda_1, \tag{6.17b}$$

where the Λ_i are bosonic superfields.

VII. CONCLUSION

By considering a fermionic extension of the KdV equation, invariant under a supersymmetry transformation, we have obtained a one-parameter family of coupled equations for a commuting and an anticommuting field. This system has been shown to be completely integrable only for a specific value of the parameter (i.e., the sKdV-3 equation). Specifically, for this particular case it has been shown that (1) there exists an infinite number of conservation laws; (2) there exists a Hamiltonian structure, with respect to which the conservation laws are in involution; and (3) there is a Lax representation. These results are actually true for an infinite sequence of equations, the sKdV hierarchy, in which the mth equation can be written as

$$\Phi_t = \{\Phi, \bar{H}_{2m}\}, \tag{7.1}$$

with the Poisson bracket (3.13) with $c = 3$ or equivalently as

$$\bar{L}_t = [\bar{L}_+^{(2m+1)/2}, \bar{L}]. \tag{7.2}$$

From this hierarchy, a new integrable hierarchy can be defined that also has the above properties. The mth equations in the two hierarchies are related by the supersymmetric extension of the Miura map given by (4.3).

Notice that the super-KdV equation (4.10) is not the only integrable simple fermionic extension of the KdV equation. (A simple fermionic KdV extension refers to a commuting KdV field coupled to a single anticommuting field.) Another such system is the one proposed by Kupershmidt [cf. Eqs. (A5)]. However, as already mentioned, this system is not invariant under a supersymmetric transformation. A Painlevé analysis[19] suggests that among all systems of the form

$$u_t = -u''' + 6uu' - 3\xi\xi'' , \tag{7.3a}$$

$$\xi_t = -c\xi''' + b\xi' u + a\,\xi u' , \tag{7.3b}$$

$c \neq 0$, only the systems (4.10) and (A5) are integrable. It is remarkable that these two equations can be formulated as Hamiltonian systems with the Poisson brackets (4.12) of the second Hamiltonian structure. In Ref. 20, it was shown that via the second Hamiltonian structure, the KdV equation can be related to the Virasoro algebra. A similar relation holds between (4.10) and (A5) and the super-Virasoro algebra.[2] This connection comes as follows. One considers the fields to be periodic functions of x and expands them in Fourier series. The substitution of these series into the various field Poisson brackets yields a well-defined algebra for the Fourier components. This algebra turns out to be exactly the (super-) conformal algebra with a central extension, realized in terms of Poisson brackets.

More complicated integrable fermionic extensions of the KdV equation (with or without an extended supersymmetry invariance) are considered in Refs. 21–23.

ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

APPENDIX A: A NONSUPERSYMMETRIC BI-HAMILTONIAN SYSTEM

In this appendix it is shown that by relaxing the condition of invariance under supersymmetry, one can find a system of coupled equations for a commuting and an anticommuting field for which a bi-Hamiltonian structure exists. Consider first a general fermionic extension of the Hamiltonian associated with the second Hamiltonian structure:

$$\tilde{\mathscr{H}}^{(2)} = \frac{1}{2}\int dx(u^2 + c_0\xi\xi') , \tag{A1}$$

where c_0 is a numerical constant, and $\tilde{\mathscr{H}}^{(2)}$ is homogeneous with respect to the grading $\deg \partial_x^k u = 2 + k$ and $\deg \partial_x^k \xi = \frac{3}{2} + k$. As a complement to the Poisson bracket (2.12), explicit expressions for the Poisson brackets $\{u(x),\xi(y)\}$ and $\{\xi(x),\xi(y)\}$ are required. Compatibility with the above grading fixes their form to be

$$\{u(x),\xi(y)\} = a_1\xi(x)\delta'(x-y) + a_2\xi'(x)\delta(x-y) , \tag{A2a}$$

$$\{\xi(x),\xi(y)\} = a_3 u(x)\delta(x-y) + a_4\delta''(x-y) . \tag{A2b}$$

Substitution of these expressions into the various Jacobi identities yields $a_1 = 3$, $a_2 = 1$, $a_3 = -a_4 = K$, where K is arbitrary. [For $K = -1$, one recovers (4.12).] The grading also fixes the general form of the fermionic extension of the Hamiltonian $\mathscr{H}^{(1)}$:

$$\tilde{\mathscr{H}}^{(1)} = \frac{1}{2}\int dx[(u')^2 + 2u^3 + c_1\xi'\xi'' + c_2 u\xi\xi'] . \tag{A3}$$

The corresponding Poisson brackets are (2.11) and

$$\{u(x),\xi(y)\} = 0, \quad \{\xi(x),\xi(y)\} = c_3\delta(x-y) . \tag{A4}$$

The various constants are determined by equating the evolution equations obtained by these two different Hamiltonian structures. One gets in this way $K = 4/c_0$, $c_1 = 4c_0$, $c_2 = 6c_0$, and $c_3 = 1/c_0$. Clearly c_0 can be absorbed in a redefinition of the anticommuting field. Setting then $c_0 = 1$, one finally obtains

$$\begin{aligned} u_t &= -u''' + 6uu' + 3\xi\xi'' , \\ \xi_t &= -4\xi''' + 6\xi' u + 3\xi u' . \end{aligned} \tag{A5}$$

The system (A5) was first described by Kupershmidt[16] (who also proved its complete integrability) and rederived from a different point of view by Gürses and Oguz.[21]

APPENDIX B: COMMUTATIVITY OF THE FLOWS

Introducing an infinite number of time variables $(t_1,t_3,\ldots)$, defined by

$$\partial_{t_k}\bar{L} = [\bar{L}^{k/2}_{+},\bar{L}] , \tag{B1}$$

one proves the commutativity of the flows defined by $\bar{L}^{k/2}$ (and hence the involutivity of the corresponding Hamiltonians) by establishing the equality[10]

$$\partial_{t_k}\partial_{t_l}\bar{L} = \partial_{t_l}\partial_{t_k}\bar{L} . \tag{B2}$$

For this one uses the fact that (B1) is equivalent to

$$\partial_{t_k}J^m = [J^k_+ ,J^m] , \tag{B3}$$

where $J = \bar{L}^{1/2}$. Therefore it follows that

$$\begin{aligned} \partial_{t_l}\partial_{t_k}\bar{L} &= [\partial_{t_l}J^k_+ ,\bar{L}] + [J^k_+ ,\partial_{t_l}\bar{L}] \\ &= [[J^l_+ ,J^k]_+,\bar{L}] + [J^k_+ ,[J^l_+ ,\bar{L}]] . \end{aligned} \tag{B4}$$

Using the Jacobi identity, one then finds

$$\partial_{t_l}\partial_{t_k}\bar{L} - \partial_{t_k}\partial_{t_l}\bar{L} = [B,\bar{L}] , \tag{B5}$$

where

$$B = [J^l_+ ,J^k]_+ - [J^k_+ ,J^l]_+ + [J^k_+ ,J^l_+] . \tag{B6}$$

Now, since

$$[J^l_+ ,J^k]_+ = [J^k,J^l_-]_+ = [J^k_+ ,J^l_-]_+ , \tag{B7}$$

where $J^l_- = J^l - J^l_+$, B becomes

$$\begin{aligned} B &= [J^k_+ ,J^l_-]_+ - [J^k_+ ,J^l]_+ + [J^k_+ ,J^l_+] \\ &= -[J^k_+ ,J^l_+]_+ + [J^k_+ ,J^l_+] = 0 . \end{aligned} \tag{B8}$$

This proves (B2).

[1] For a review, see Yu. I. Manin, J. Sov. Math. **11**, 1 (1979).
[2] P. Mathieu, Phys. Lett. B **203**, 287 (1988).
[3] P. Lax, Commun. Pure Appl. Math. **21**, 467 (1968).
[4] I. M. Gelfand and L. A. Dickii, Funct. Anal. Appl. **10**, 259 (1976).
[5] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. **19**, 1095 (1967).
[6] R. M. Miura, J. Math. Phys. **9**, 1202 (1968).

[7] Originally (2.3) was obtained by the substitution $u \rightarrow u + \lambda$ which is allowed by the Galilean invariance of the KdV equation. However, this invariance does not hold for the higher analog of the KdV equation to be introduced below.

[8] R. M. Miura, C. S. Gardner, and M. D. Kruskal, J. Math. Phys. **9**, 1204 (1968).

[9] M. D. Kruskal, R. M. Miura, C. S. Gardner, and N. J. Zabusky, J. Math. Phys. **11**, 952 (1970).

[10] G. Wilson, Math. Proc. Camb. Philos. Soc. **86**, 131 (1979).

[11] C. S. Gardner, J. Math. Phys. **12**, 1548 (1971).

[12] F. Magri, J. Math. Phys. **19**, 1156 (1978).

[13] B. A. Kupershmidt and G. Wilson, Invent. Math. **62**, 403 (1981).

[14] V. E. Zakharov and L. D. Faddeev, Funct. Anal. Appl. **5**, 280 (1971).

[15] Yu. I. Manin and A. O. Radul, Commun. Math. Phys. **98**, 65 (1985).

[16] B. A. Kupershmidt, Phys. Lett. A **102**, 213 (1984).

[17] B. A. Kupershmidt, J. Math. Phys. **22**, 449 (1981).

[18] Actually, the Lax representation for the sKdV-3 equation is degenerate since by using $\overline{L} = D^4 - \Phi D$ one also recovers (4.10). This is the representation given in Ref. 15. It can be obtained from the super-Miura map with the transformation $\psi = D \ln D\Xi$, where Ξ is a fermionic superfield.

[19] P. Mathieu, Phys. Lett. A **128**, 169 (1988).

[20] J. L. Gervais and A. Neveu, Nucl. Phys. B **209**, 125 (1982); J. L. Gervais, Phys. Lett. B **160**, 277, 279 (1985).

[21] M. Gürses and O. Oguz, Lett. Math. Phys. **11**, 235 (1986).

[22] B. A. Kupershmidt, J. Phys. A **17**, L869 (1984).

[23] C. A. Laberge and P. Mathieu, "$N = 2$ superconformal algebra and integrable O(2) fermionic extensions of the Korteweg–de Vries equation," preprint.

Coherent-state path-integral *S*-matrix formalism applied to the Lee model

C. L. Hammer
Ames Laboratory and the Department of Physics, Iowa State University, Ames, Iowa 50011
J. E. Shrauner
Arthur Holly Compton Laboratory of Physics, Washington University, St. Louis, Missouri 63130

(Received 31 March 1988; accepted for publication 13 July 1988)

In this paper it is shown that a coherent-state path-integral treatment significantly simplifies the analysis of the Lee model, including higher sector calculations. The method may be applicable to more general field theories. However, the special properties of the Lee model allow the results to be displayed in closed form.

I. INTRODUCTION

The application of coherent-state and path-integral techniques to problems with large quantum numbers is well known in many fields of physics. In quantum field theory these methods are especially useful in integrating certain subsets of dynamical field variables within detailed microscopic theories to obtain effective action formulations at more phenomenological levels.[1]

We have developed a general formulation of quantum field theory in terms of coherent-state path integrals for the elements of the U and S matrices in a coherent-state basis. A detailed description of this formulation was developed in Ref. 2. Because this formulation is primarily in terms of the elements of the U and S matrices, rather than the Green's functions (as has been most widely used previously) it lends a powerful formalism for dealing with spontaneous symmetry breaking in quantum field theories. We have previously emphasized the convenience of this formulation in treating quantum field theories with classical coherent-state condensates (solitons, vortices, etc.) in Ref. 3.

We illustrated this formulation with its application to the quantum field theory of the scalar meson interacting with heavy nucleons in Ref. 2. The renormalization and scattering amplitudes were calculated very straightforwardly to obtain agreement with well-known results.

Extension to the next simplest model quantum field theory, the Lee model (Ref. 4), involves considerably more subtle complications. The scalar meson theory shares with QED the characteristic that the only conserved charge is fermion conservation, which is not carried by the meson field in the Abelian theory. On the other hand, the Lee model shares, along with QCD and chiral field theories, the additional characteristic that the meson fields carry non-Abelian charges.

It is this additional complication that causes the extensive algebraic difficulty in analyzing the Lee model, especially when one discusses the higher sectors of the model.

In this paper we show that the coherent-state path-integral treatment simplifies the analysis of the Lee model. We are able to identify an algorithm that associates with every fully dressed V-particle line a renormalization factor Z and that associates with every fully dressed internal V particle line a factor $Z Z_R(e)$, where e is the energy denominator as determined by old-fashioned perturbation theory comprised of fully renormalized physical-particle energies. This allows us to use a simple diagrammatic approach in the analysis.

In Sec. II we recapitulate the construction of the coherent-state basis necessary to establish notational formalism. In Sec. III we construct the matrix elements of the U and S matrices and determine (a) the mass and wave-function renormalizations, (b) the N-θ scattering amplitude, and (c) the coupling constant renormalization. In Sec. IV we discuss higher sectors of the Lee model and the relevance of certain aspects of our treatment to further problems.

II. COHERENT STATES OF THE MESON FIELD

The Lee model has a nucleonlike fermion field of two states, N and V, interacting with a mesonlike boson field θ as $V \leftrightarrow N + \theta$. We will be interested in coherent states of this meson field that are eigenstates of the pure annihilation operator part of the $\theta(x,t)$ field operator:

$$\theta_+(x,t)|\phi(t)\rangle = \phi(x,t)|\phi(t)\rangle, \tag{1}$$

where $\phi(x,t)$ is a c-number coherent-state wave function that is well behaved with respect to the inner products that will be required of it. The coherent states are generated by the operator $\exp \Sigma_\phi$ as

$$\begin{aligned} &e^{\Sigma_\phi(t)}|\Omega\rangle = |\phi(t)\rangle, \\ &\langle\Omega(t)|e^{-\Sigma_\phi(t)} = \langle\phi(t)|, \end{aligned} \tag{2}$$

in which $|\Omega\rangle$ is the Fock-space vacuum state. The operator Σ_ϕ is constructed from the dynamical field operators θ as

$$\Sigma_\phi(t) = (\theta(t),\phi(t)) - (\phi(t),\theta(t)), \tag{3}$$

in which the inner product of θ with ϕ is

$$(\phi(t),\theta(t)) = i\int d^3x\, \phi^*(x,t)\overleftrightarrow{\partial}_0\, \theta(x,t). \tag{4}$$

Coherent states of the θ field constructed this way satisfy

$$\begin{aligned} \langle\phi(t)|\chi(t)\rangle = \exp[&-\tfrac{1}{2}(\phi(t),\phi(t)) - \tfrac{1}{2}(\chi(t),\chi(t)) \\ &+ (\phi(t),\chi(t))]. \end{aligned} \tag{5}$$

They form an overcomplete set on which the resolution of unity is a functional integral

$$1 = \int \frac{D^2\phi}{\pi}\, |\phi(t)\rangle\langle\phi(t)|, \tag{6}$$

in which the differential element has the forms

$$\frac{D^2\phi}{\pi}=\frac{D\,\mathrm{Re}\,\phi}{\sqrt{\pi}}\frac{D\,\mathrm{Im}\,\phi}{\sqrt{\pi}}=\frac{D\phi}{\sqrt{2\pi}}\frac{D\phi^*}{\sqrt{2\pi}}. \tag{7}$$

In terms of the coefficients β_k for expansion of ϕ in a complete basis

$$\frac{D\phi}{\sqrt{\pi}}=\prod_k\frac{(d\beta_k)}{\sqrt{\pi}}=\frac{d\beta_{k1}}{\sqrt{\pi}}\frac{d\beta_{k2}}{\sqrt{\pi}}\frac{d\beta_{k3}}{\sqrt{\pi}}\cdots. \tag{8}$$

All the relations for the coherent states mentioned above apply for any operator-valued field θ satisfying canonical equal-time commutation relations and the associated conserved currents. We will be interested in coherent states of free asymptotic fields; in such a case the inner products [Eq. (4)] are based on conserved currents of free fields and thus are independent of time.

Choosing a complete set of positive-energy free-particle normal modes $\{\phi_k\}$, orthonormal as $(\phi_k,\phi_q)=\delta^3(k,q)$, to expand both θ as

$$\theta_{k,\mathrm{in}}=(\phi_k,\theta_{\mathrm{in}})$$

and ϕ as

$$\beta_k(t)=(\phi_k(t),\phi(t)), \tag{9}$$

the coherent-state generator $\Sigma_{\phi\,\mathrm{in}}$ is expanded as

$$\Sigma_{\phi\,\mathrm{in}}(t)=(\theta_{\mathrm{in}}(t),\phi(t))-(\phi(t),\theta_{\mathrm{in}}(t))$$

$$=\sum_k[\theta^\dagger_{k\,\mathrm{in}}\,\beta_k(t)-\beta^*_k(t)\theta_{k\,\mathrm{in}}]. \tag{10}$$

The coherent states generated are

$$|\phi(t)\rangle_{\mathrm{in}}=e^{\Sigma_{\phi\,\mathrm{in}}(t)}|\Omega\rangle_{\mathrm{in}}$$

$$=\prod_k e^{\theta^\dagger_{k\,\mathrm{in}}\beta_k(t)}e^{-(1/2)|\beta_k(t)|^2}|\Omega\rangle_{\mathrm{in}}$$

$$=\prod_k\left[\sum_{n_k}\frac{\beta_k(t)^{n_k}}{\sqrt{n_k!}}|n_k\rangle_{\mathrm{in}}\right]e^{-(1/2)|\beta_k(t)|^2}. \tag{11}$$

The usual Fock-space in states are obtained from these coherent states as

$$|n_{k_1},\ldots,n_k,\ldots\rangle_{\mathrm{in}}=\prod_k\left(\frac{\partial}{\partial\beta_k}\right)^{n_k}\frac{1}{\sqrt{n_k!}}e^{(1/2)|\beta_k(t)|^2}|\phi\rangle_{\mathrm{in}}|_{\substack{\phi=0\\=\beta_k}}, \tag{12}$$

where n_k is the number of quanta in the kth mode, and the t dependence has been suppressed, since it vanishes after the $\phi\to0,\beta_k\to0$ limit. The usual S-matrix element between such states is similarly

$$_{\mathrm{out}}\langle n_{k_1},\ldots,n_k,\ldots|n_{q_1},\ldots,n_q,\ldots\rangle_{\mathrm{in}}$$

$$=\prod_{k,q}\left(\frac{\partial}{\partial\beta'^*_k}\right)^{n_k}\left(\frac{\partial}{\partial\beta_q}\right)^{n_q}\sqrt{n_k!n_q!}$$

$$\times e^{(1/2)(|\beta'_k|^2+|\beta_q|^2)}\,{}_{\mathrm{out}}\langle\phi'|\phi\rangle_{\mathrm{in}}|_{\substack{\phi=0\\=\phi'}}, \tag{13}$$

with $\beta'_k(t)=(\phi_k(t),\phi'(t))$.

The coherent-state wave functions that we are interested in all satisfy free-field wave equations and we have defined all the inner products in terms of free-field conserved currents. This ensures that the S-matrix element in the coherent state basis is the limit

$$\lim_{\substack{t\to-\infty\\t'\to\infty}}{}_{\mathrm{out}}\langle\phi'(t')|\phi(t)\rangle_{\mathrm{in}}=\lim_{\substack{t\to-\infty\\t'\to\infty}}\langle\phi'(0)|U(t',t)|\phi(0)\rangle, \tag{14}$$

where

$$U(t',t)=e^{iH_0t'}e^{-iH(t'-t)}e^{-iH_0t}, \tag{15}$$

with $H=H_0+H'$ the full interacting Hamiltonian and H_0 the free-field Hamiltonian. See Ref. 2 for a careful and thorough discussion of the time dependence of operators with the use of asymptotic free fields and their currents in construction of coherent-state S-matrix elements. In terms of the above limits the S-matrix element is

$$_{\mathrm{out}}\langle n_{k_1},\ldots,n_k,\ldots|n_{q_1},\ldots,n_q,\ldots\rangle_{\mathrm{in}}$$

$$=\lim_{\substack{t\to-\infty\\t'\to\infty}}\prod_{k,q}\left(\frac{\partial}{\partial\beta'^*_k}\right)^{n_k}\left(\frac{\partial}{\partial\beta_q}\right)^{n_q}\frac{1}{\sqrt{n_k!n_q!}}$$

$$\times e^{(1/2)(|\beta'_k|^2+|\beta_q|^2)}\langle\phi'(t')|\phi(t)\rangle|_{\phi=0=\phi'}. \tag{16}$$

III. REPRESENTATION OF *U* AND *S* MATRICES IN COHERENT-STATE BASIS

The dynamical field operators of which H, H_0, and H' are constructed are assumed separated into pure creation and annihilation parts as $\theta=\theta_++\theta_-$. A normal ordering is defined so that in coherent-state matrix elements the annihilation parts act to the right on their eigenstates as $\theta_+(x,t)|\phi(0)=\theta(x,t)|\phi(0)\rangle$ and the creation parts act correspondingly to the left on their eigenstates. If there is no non-Albelian conserved current in the interaction Hamiltonian this leads straightforwardly to infinitesimal-time-interval U-matrix elements in which the exponential operators can be replaced by a c-number functional as

$$\langle\phi'(0)|e^{-i\epsilon H'(\theta_-(t'),\theta_+(t))}|\phi(0)\rangle$$

$$=\langle\phi'(0)|\phi(0)\rangle e^{-i\epsilon H'(\phi'(t'),\phi(t))}. \tag{17}$$

Because this matrix element is first order in the infinitesimal ϵ we have some liberty as to specifying its time arguments; for example, it might be convenient to label them $H'(\phi'(t+\epsilon),\phi(t))$ in the exponent, or $H'(\phi'(t),\phi(t))$ also. This hardly made the previous treatment of the scalar meson QFT different in this aspect from the familiar treatment of the external-current bremsstrahlung problem.

The Lee model[4] has a conserved current in the interaction H' that makes it quite different from the previous simple cases. In the Lee model the Hamiltonian is

$$H=H_0+H',$$

$$H_0=\sum_q\{m_vV^\dagger_qV_q+m_NN^\dagger_qN_q\}+\sum_ka^\dagger_ka_k\omega_k,$$

$$H'=\sum_{k,p}\left\{\frac{\lambda f(k)}{\sqrt{2\omega_k}}a_kV^\dagger_pN_{p-k}+\frac{\lambda f(k)^*}{\sqrt{2\omega_k}}a^\dagger_kN^\dagger_pV_{p+k}\right\}$$

$$+\sum_p\{\delta m_VV^\dagger_pV_p+\delta m_NN^\dagger_pN_p\}, \tag{18}$$

where the momentum dependence of energies of the heavy V and N fermions is ignored and the θ meson energies are $\omega_k=\sqrt{k^2+\mu^2}$. The physical masses are m_V, m_N, and μ, while δm_V and δm_N are physical mass minus bare mass correction terms. The coupling constants and regularizing form factor $f(k)$ are presumed chosen so that no ghost state results.[4] This ensures that the Hamiltonian remains Hermitian and the S matrix unitary.[4] The nonvanishing equal-time commutation relations are

$$\{N_q,N_{q'}^\dagger\} = \{V_q,V_{q'}^\dagger\} = [a_q,a_{q'}^\dagger] = \delta^3(q,q'). \tag{19}$$

It is convenient to introduce an (incomplete) isospin-operator notation with $\psi = \binom{V}{N}$ and $\tau_\pm$ the usual raising/lowering operators whose action and normalization are described by $\tau_+\tau_-\psi = V$ and $\tau_-\tau\psi = N$.

With this notation

$$\begin{aligned} H_0 &= \sum_p \psi_p^\dagger m \psi_p + \sum_k \omega_k a_k^\dagger a_k, \\ H' &= -\sum_p \psi_p^\dagger \delta m \psi_p \\ &\quad + \sum_p \sum_k \{\psi_{p-k}^\dagger \bar{g}_k a_k \psi_p + \psi_{p+k}^\dagger \bar{g}_k^\dagger a_k^\dagger \psi_p\}, \end{aligned} \tag{20}$$

where

$$\begin{aligned} m &= m_N\tau_-\tau_+ + m_V\tau_+\tau_-, \\ \delta m &= \delta m_N \tau_-\tau_+ + \delta m_V \tau_+\tau_- \\ \bar{g}_k &= \lambda f(k)\tau_+/\sqrt{2\omega_k}\,. \end{aligned} \tag{21}$$

The time dependence of the in-state field operators is

$$a_k(t) = a_k(0)e^{-i\omega_k t}, \quad N_p(t) = N_p(0)e^{-im_N t},$$
$$V_p(t) = V_p(0)e^{-im_V t},$$

with physical masses m_N and m_V.

The most expeditious way to construct the U- and S-matrix elements in the coherent-state basis is to expand the operator $T\exp[-i\int dt\, H'(t)]$, insert complete sets of coherent states between each interaction, and integrate all ordered time partitions. It is useful to define

$$g_k(t_l) = \bar{g}_k\, e^{i\mathbf{k}\cdot\mathbf{x} - i\tilde{\omega}_k t_l}, \tag{22}$$

where $\tilde{\omega}_k = \omega_k + m_N - m_V$. Then the matrix element of H' between coherent states of the θ field and single fermion field labeled by their momenta p_l and p_r are

$$\begin{aligned} &\langle p_r\phi_r|H'(t_i)|p_l\phi_l\rangle \\ &\quad = \int \frac{d^3x}{(2\pi)^3 Z} e^{i\mathbf{x}\cdot(\mathbf{p}_l - \mathbf{p}_r)} \langle\phi_r|\phi_l\rangle \\ &\qquad \times\Big[-\delta m + \sum_k (g_k(t_i)\beta_{lk} + g_l^\dagger(t_i)\beta_{rk}^*)\Big], \end{aligned} \tag{23}$$

where the β_{lk}'s are the expansion coefficients of the coherent-state wave function ϕ_l in the basis of θ-particle in-states ϕ_k, as in Eq. (9).

The S-matrix element that determines the wave function and mass renormalizations is

$$\begin{aligned} \delta^3_{\mathbf{p}',\mathbf{p}} &= \langle p'|p\rangle = \langle p'\phi'|U(t',t)|p\phi\rangle|_{\phi=0=\phi'} \\ &= \frac{\delta^3_{p'p}}{Z} - i\int_{t\to-\infty}^{t'\to\infty} dt_1 \langle p'\phi'|H_1(t_1)|p\phi\rangle|_{\theta=0=\phi'} \\ &\quad + (-i)^2 \int_{t\to-\infty}^{t'\to\infty} dt_2 \int_{t\to-\infty}^{t_2} dt_1 \int \frac{D^2\phi_1}{\pi} \\ &\quad \times \sum_{p_1} \langle p'\phi'|H'(t_2)|p_1\phi_1\rangle \\ &\quad \times \langle p_1\phi_1|H'(t_1)|p\phi\rangle|_{\theta=0=\phi'} + \cdots, \\ &= \frac{\delta^3_{p'p}}{Z}\Big\{1 + i\delta m(t'-t) + \frac{1}{2}(i\delta m(t'-t))^2 \\ &\quad - \sum_k g_k g_k^\dagger \Big[\frac{t'-t}{i\tilde{\omega}_k} - (e^{-i\tilde{\omega}_k(t'-t)} - 1)\Big] + \cdots\Big\}. \end{aligned} \tag{24}$$

Terms like $\exp(-i\tilde{\omega}_k(t'-t))$ are dropped by the Reimann–Lebesgue lemma as the averaging to zero of rapid oscillations in large-time limit. Identifying δm as second order in the coupling constant λ, and g as first order, only terms of even order survive in the above matrix element when $\phi = 0 = \phi'$. Within the large brackets we have[5]

$$\begin{aligned} &1 + i(t'-t)\Big[\delta m + \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k}\Big] - \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k^2} \\ &\quad - \frac{(t'-t)^2}{2}\Big[\delta m^2 + \delta m \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k} \\ &\quad + \Big(\sum_R \frac{g_k g_k^\dagger}{\tilde{\omega}_k}\Big)\delta m + \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k} \sum_q \frac{g_q g_q^\dagger}{\tilde{\omega}_q}\Big] \\ &\quad - i2\delta m \sum_k g_k g_k^\dagger \Big[\frac{i}{\tilde{\omega}_k^3} + \frac{t'-t}{\tilde{\omega}_k^2}\Big] \\ &\quad + \sum_k g_k g_k^\dagger \sum_q g_q g_q^\dagger \Big[\frac{2}{\tilde{\omega}_q\tilde{\omega}_k^3} + \frac{1}{\tilde{\omega}_k^2\tilde{\omega}_q^2} + \frac{2(t'-t)}{i\tilde{\omega}_q\tilde{\omega}_k^2}\Big] \\ &\quad + \cdots. \end{aligned} \tag{25}$$

The leading powers of $(t'-t)$ in every order of λ sum into

$$\exp\Big[i(t'-t)\Big(\delta m + \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k}\Big)\Big],$$

which we set to unity by requiring the mass renormalization condition

$$\delta m = -\sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k}. \tag{26}$$

This requirement, in fact, also cancels all dependence on $(t'-t)$ in the remaining terms in every order of the coupling; leaving only the terms independent of $(t'-t)$. In the $2n$th order term of our matrix element the time- independent term is $(-\Sigma_k g_k g_k^\dagger/\tilde{\omega}_k^2)^n$. These terms sum to give the wave-function renormalization constant:

$$Z = \sum_{n=0}\Big(-\sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k^2}\Big)^n = \Big[1 + \sum_k \frac{g_k g_k^\dagger}{\tilde{\omega}_k^2}\Big]^{-1}. \tag{27}$$

These are the familiar renormalization conditions obtained by other more familiar methods. In our isospin notation $g_k g_k^\dagger$ is proportional to the V-particle projection operator $\tau_+\tau_-$. Thus $Z_N = 1$, $\delta m_N = 0$, and only the V particle is dressed by the interaction and renormalizes.

Scattering of a θ particle in momentum state k by an N particle in momentum state p into final states of momenta k' and p', respectively, is determined from the S-matrix element

$$\begin{aligned}\langle k'p'|S|kp\rangle &= \left\{\frac{\partial}{\partial\beta_k'^*}\frac{\partial}{\partial\beta_k}e^{(1/2)(|\beta'_{k'}|^2+|\beta_k|^2)}\right.\\ &\quad\left.\times\langle p'\phi'|\tau_-\tau_+U(t'-t)\tau_-\tau_+|p\phi\rangle\Big|_{\substack{t'\to\infty\\ t\to-\infty}}\right\}\Bigg|_{\phi'=0=\phi}\\ &= \langle p_1k_1|\tau_-\tau_+U(t'-t)\tau_-\tau_+|p,k\rangle\Big|_{\substack{t'\to\infty\\ t\to-\infty}}.\end{aligned}\tag{28}$$

The development of the S-matrix element for N-θ scattering can be traced easily through diagrams because of the severe constraints of current conservation in this model. After mass and wave-function renormalization, the Born approximation matrix element is obtained from

$$\begin{aligned}&-\langle p'\phi'|\tau_-\tau_+\int_t^{t'}dt_2\int_t^{t_2}dt_1\,H'(t_2)H'(t_1)\tau_-\tau_+|p\phi\rangle\\ &\quad= -\int\frac{d^3x}{(2\pi)^3}e^{i\mathbf{x}\cdot(\mathbf{p}-\mathbf{p}')}\int_t^{t'}dt_2\\ &\qquad\times\int_t^{t_2}dt_1\langle\phi'|\phi\rangle\bar{u}_{p'}g_{k'}^\dagger(t_2)g_k(t_1)\beta_{k'}'^*\beta_k u_p\\ &\quad= 2\pi i\delta(\omega_{k'}-\omega_k)\delta^3_{p+k,p'+k'}\\ &\qquad\times\bar{u}_{p'}g_{k'}^\dagger\frac{1}{\tilde{\omega}_k}g_k\langle\phi'|\phi\rangle\beta_{k'}'^*\beta_k u_p\end{aligned}\tag{29}$$

by the process of Eq. (13) as

$$\begin{aligned}&\langle p'k'|S|p,k\rangle_{\text{Born}}\\ &\quad= i2\pi\delta(\omega_{k'}-\omega_k)\delta^3_{p+k,p'+k'}\bar{u}_{p'}g_{k'}^\dagger(1/\tilde{\omega}_k)g_k u_p,\end{aligned}\tag{30}$$

where $g_k = g_k(t=0)$. The diagram for this process is Fig. 1(a).

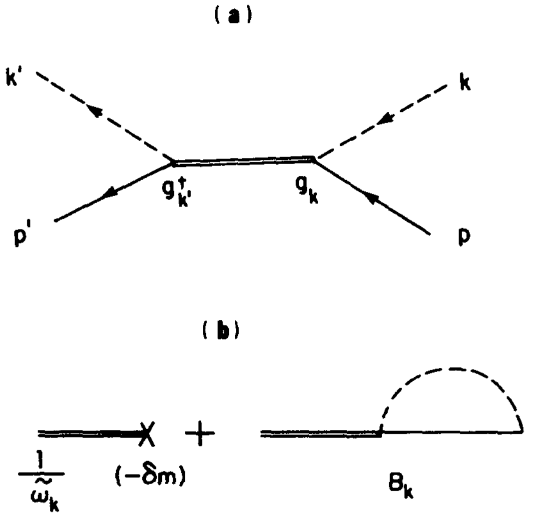


FIG. 1. (a) This diagram can be viewed as the unrenormalized Born amplitude diagram for N-θ scattering where the dashed lines represent θ particles, the single straight lines represent N particles, and the double straight lines represent the unrenormalized V particle. This same diagram can be viewed as the renormalized skeleton diagram for V-θ scattering in which the double straight lines represent fully dressed renormalized V particles. The dressing of the V-particle line is accomplished by the sum of all possbile corrections of the kinds shown in Fig. 1(b). (b) The mass corrections and N-θ bubble insertions which, when inserted in a V-particle line and summed over all possible combinations, yield the fully dressed V-particle line.

Current conservation restricts the action of H' in higher-order corrections to the insertion of all possible combinations of δm and N-θ bubbles in the internal V-particle line of the Born term. Because δm has been identified to have a definite order of $N\theta V$ coupling strength, viz., $\delta m = -\Sigma_k g_k g_k^\dagger/\tilde{\omega}_k$, the sum of all possible combinations of δm and the N-θ bubble $B_k = \Sigma_q g_q g_q^\dagger/((\tilde{\omega}_k - \tilde{\omega}_q)\tilde{\omega}_k)$ in a V-particle line of momentum $k+p$ of total order n is

$$\sum_{l=0}^{n}\binom{n}{l}(-\delta m)^l B^{n-l} = \left(-\frac{1}{\tilde{\omega}_k}\delta m + B_k\right)^n.\tag{31}$$

The $l/\tilde{\omega}_k$ with δm is the associated propagator as in Fig. 1(b), which B_k already has. The N-θ bubble and δm combine as

$$\begin{aligned}-\frac{1}{\tilde{\omega}_k}\delta m + B_k &= \sum_q g_q g_q^\dagger\left[\frac{1}{\omega_k-\omega_q}+\frac{1}{\tilde{\omega}_q}\right]\frac{1}{\tilde{\omega}_k}\\ &= \sum_q g_q g_q^\dagger\frac{1}{\tilde{\omega}_q(\omega_k-\omega_q)}.\end{aligned}\tag{32}$$

The sum of all orders N then gives for the total S-matrix element

$$\begin{aligned}\langle k'p'|S|k,p\rangle &= i2\pi\delta(\omega_{k'}-\omega_k)\delta^3_{p+k,p'+k'}\bar{u}_{p'}g_{k'}^\dagger\\ &\quad\times\sum_{n=0}\left(-\frac{1}{\tilde{\omega}_k}\delta m + B_k\right)^n\frac{1}{\tilde{\omega}_k}g_k u_p\\ &= i2\pi\delta(\omega_{k'}-\omega_k)\delta^3_{p+k,p'+k'}\bar{u}_{p'}g_{k'}^\dagger\\ &\quad\times\frac{1}{\tilde{\omega}_k}Z(\tilde{\omega}_k)g_k u_p,\end{aligned}\tag{33}$$

where

$$Z^{-1}(\tilde{\omega}_k) = 1 + \sum_q\frac{g_q g_q^\dagger}{\tilde{\omega}_q(\omega_q-\omega_k)}.\tag{34}$$

The S-matrix element for N-θ scattering in Eq. (33) is of the form of the Born amplitude times $Z(\tilde{\omega}_k)$, which contains a logarithmic divergence. This divergence is removable by coupling-constant renormalization in which the bare coupling constant is replaced by $\lambda\sqrt{Z} = \lambda_R$ everywhere and a subtraction replaced the logarithmic divergent expression in Eq. (34) by the convergent expression (since $g_q \propto 1/\sqrt{\omega_q}$)

$$\begin{aligned}Z^{-1}(\tilde{\omega}_k) &= Z^{-1}\left[1 + \tilde{\omega}_k Z\sum_q\frac{g_q g_q^\dagger}{\tilde{\omega}_q^2(\omega_q-\omega_k)}\right]\\ &= Z^{-1}Z_R^{-1}(\tilde{\omega}_k).\end{aligned}\tag{35}$$

We have chosen the subtraction point to be where $Z^{-1}(\tilde{\omega}_k = 0) = Z^{-1}$, thereby avoiding the necessity of introducing any new quantities that require phenomenological identification. Again, these results are the same as those obtained by usual methods.

Because of the severely constrained structure of the Lee model and the uniform structure of our expansions we are able to identify a very useful algorithm for extending these renormalization results to more complicated processes. The algorithm associates with every fully dressed external V-particle line a factor Z and with every fully dressed internal V-particle line a factor $ZZ_R(e)$ in which e is the energy denominator associated with the internal V line, as by (old-fashioned) perturbation theory comprised of fully renormalized, physical-particle energies. All energy denominators are determined this way in the skeleton diagrams built of these fully dressed particle lines.

Our presumed condition on the coupling constants and the form factor ensure that $Z^{-1}(e)$ never vanishes. Therefore only a single, physical V-particle state exists (no ghost states) and the S matrix is unitary.

A similar perturbation analysis in terms of fully renormalizable quantities was developed by Ruijgrok[6] for a more sophisticated model that includes the Lee model as a special case.

IV. APPLICATION OF THE ALGORITHM TO OTHER PROCESSES AND OTHER SECTORS

The treatment of the Lee model in the N-θ sector by our path integral in the coherent-state basis has led to an algorithm with which the familiar results in that sector are corroborated. This algorithm can now be extended to processes in the V-θ and V-N sectors. There are two processes of interest in the V-θ sector: V-θ elastic scattering and θ production in $V\theta \rightarrow N\theta\theta$. In the V-N sector there are also two processes of interest: V-N elastic scattering and θ production in $VN \rightarrow NN\theta$.

Our path integral for U- and S-matrix elements in the meson coherent-state basis gives sums of terms that correspond to the old-fashioned perturbation theory diagrams for the expansion of the U- or S-matrix operator in ordered time integrations of powers of the interaction Hamiltonian. The charge conservation restrictions and the extreme nonrelativistic fermions of the model limit the classes of diagrams to manageable sets. In the N-θ sector the sum of all allowed subdiagrams inserted into external V particle lines just leads to wave-function and mass renormalization of these lines. The sum of all allowed subdiagrams inserted into the internal V-particle line in the N-θ scattering S-matrix element leads to coupling-constant renormalization and multiplication of the internal V line by the finite factor $Z_R(\tilde{\omega}_k)$ described in Eqs. (33)–(35). With suitable generalization and care this algorithm can be extended to the S-matrix elements for processes in the V-θ and V-N sectors.

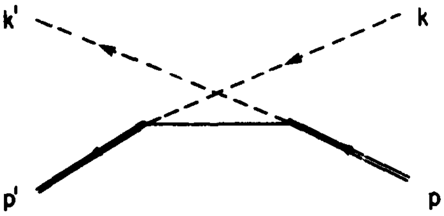


FIG. 2. Renormalized skeleton Born diagram for V-θ scattering. The V-particle lines represent fully dressed and renormalized V-particle lines as described in Sec. III.

Since neither θ nor N dress or renormalize in the Lee model there is only one class of interactions contributing to V-θ scattering, which is represented by the single diagram in Fig. 2 and its iterations. The V-particle lines in Fig. 2 represent the completely dressed physical particles. Each of these lines represents the sum of all possible mass correction and N-θ bubble diagram insertions, as occurred in dressing and renormalization of the stable single physical V-particle state above. The summation of all possible iterations of the above renormalized skeleton diagram gives the complete V-θ scattering matrix element. For $V\theta \rightarrow N\theta\theta$ each diagram of V-θ scattering has the final V dissociate into $N\theta$, and no other diagrams contribute. The external V-particle lines each dress to give a factor of Z, half of which is canceled by the factor $1/\sqrt{Z}$ associated with the asymptotic in-field operator. The remaining uncanceled $\sqrt{Z}$ factor renormalizes the coupling constant at the end of the iterative chain. The internal V lines each get the factor $Z_R(e)$ between renormalized couplings. The first few terms in this sequence of skeleton diagrams of fully dressed particle lines are shown in Fig. 3. These dressed-particle skeleton diagrams represent terms of the T-matrix element

$$\langle k'p'|S-1|kp\rangle = 2\pi i\delta(\omega-\omega')\delta^3_{k+p,k'+p'}T(k',k), \tag{36}$$

where

$$T(k',k) = g_{kR}\frac{1}{-\tilde{\omega}}g^{\dagger}_{k'R} + \sum_{k_1}\frac{g_{k_{1R}}Z_R(\tilde{\omega}-\tilde{\omega}_1)g^{\dagger}_{k'R}g_{kR}g^{\dagger}_{k_{1R}}}{\tilde{\omega}_1^2(\tilde{\omega}-\tilde{\omega}_1)}$$
$$+\sum_{k_1k_2}\frac{g_{k_2R}Z_R(\tilde{\omega}-\tilde{\omega}_2)g^{\dagger}_{k'R}g_{k_1R}Z_R(\tilde{\omega}-\tilde{\omega}_1)g^{\dagger}_{k_2R}g_{kR}g^{\dagger}_{k_1R}}{-\tilde{\omega}_2(\tilde{\omega}-\tilde{\omega}_2)(\tilde{\omega}-\tilde{\omega}_1-\tilde{\omega}_2)(\tilde{\omega}-\tilde{\omega}_1)(-\tilde{\omega}_1)} + \cdots. \tag{37}$$

Here $\tilde{\omega}' = \tilde{\omega}_{k'}$, $\tilde{\omega}_1 = \tilde{\omega}_{k_1}$, and so on. This sequence can be summed into a closed finite expression, as first done by Amado in Ref. 7 and amplified by Pagnamenta in Ref. 8 and others as

$$T(k',k)|_{\omega=\omega'} = g_{kR}\frac{Z_R(\tilde{\omega})}{\tilde{\omega}}g^{\dagger}_{k'R}\frac{1+C(\tilde{\omega})}{1-C(\tilde{\omega})}, \tag{38}$$

where[5,6]

$$C(\tilde{\omega}) = \tilde{\omega}Z_R^{-1}(\tilde{\omega})\sum_{k'}\frac{g_{k'R}g^{\dagger}_{k'R}|Z_R(\tilde{\omega}')|^2Z_R(\tilde{\omega}-\tilde{\omega}')}{\tilde{\omega}'^2(\tilde{\omega}-\tilde{\omega}')}. \tag{39}$$

Pagnamenta obtains $T(k',k)$, as in Eq. (38), as a solution of the integral equation

$$T(k',k) = g_{kR}\frac{Z_R(\tilde{\omega}-\tilde{\omega}')}{-\tilde{\omega}}g^{\dagger}_{k'R} + Z_R(\tilde{\omega}-\tilde{\omega}')$$
$$\times\sum_{k_1}\frac{g_{k_1R}g^{\dagger}_{k'R}T(k_1,k)}{(\tilde{\omega}-\tilde{\omega}_1)(\tilde{\omega}-\tilde{\omega}_1-\tilde{\omega}')}, \tag{40}$$

which is easily seen to give our expansion of Eq. (37) by self-iteration of Eq. (40) in a Neumann–Liouville series. The zeroth–order term of the iteration is the first term on the rhs

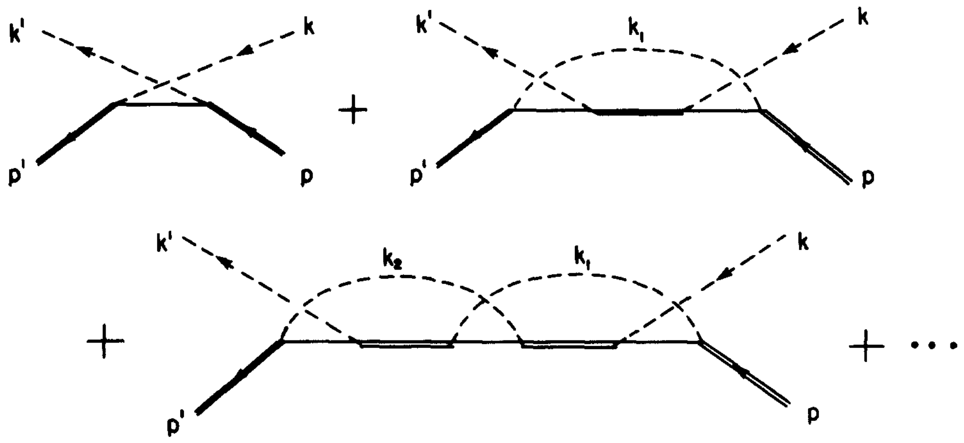


FIG. 3. Renormalized skeleton diagrams for the complete V-θ scattering amplitude. All the higher order expansion diagrams are fully dressed renormalized skeleton diagrams.

of Eq. (40) with $\tilde{\omega} = \tilde{\omega}'$, $Z_R(0) = 1$.

The T-matrix element for $V\theta \rightarrow N\theta\theta$ is

$$T(k'',k',k) = g^{\dagger}_{k''R} \frac{Z_R(\tilde{\omega} - \tilde{\omega}')}{(\tilde{\omega} - \tilde{\omega}')} T(k',k), \qquad (41)$$

where $T(k' - k)$ is that of Eqs. (37)–(40) which, as far as we know, may be a new result.

Our algorithm can be applied to N-V scattering just as directly as in the above cases. However, it would require a more cumbersome notation and a generalization of our treatment above to include momentum-dependent fermion energies.

ACKNOWLEDGMENTS

The authors wish to thank John R. Klauder for many useful discussions.

This work was supported by the U.S. Department of Energy, Contract No. W-7405-Eng-82, Office of Energy Research (KA-01-01), and by Contract No. DE-AC02-78-ER04915, Division of High Energy and Nuclear Physics.

[1] J. Cornwall, R. Jackiw, and E. Tomboulis, Phys. Rev. D **10**, 2428 (1974); R. W. Haymaker, T. Matsuki, and F. Cooper, *ibid.* **32**, 2049 (1985); **33**, 1137 (1986); T. Morozumi, Kyoto preprint RIFP-671, 1986; H. Kleinert, Phys. Lett. B **62**, 429 (1976); E. Shrauner, Phys. Rev. D **16**, 1877 (1977); T. Kugo, Phys. Lett. B **76**, 625 (1978); I. J. R. Aitchison and C. M. Frazer, *ibid.* **146**, 63 (1984); P. Simic, Rockefeller University preprint RU/85/124, 1985.

[2] C. L. Hammer, E. Shrauner, and B. DeFacio, Phys. Rev. D **18**, 373 (1978); *Coherent States Applications in Physics and Mathematical Physics*, edited by J. R. Klauder and B.-S. Skagerstam (World Scientific, Singapore, 1985), p. 885.

[3] C. L. Hammer, E. Shrauner, and B. DeFacio, Phys. Rev. D **19**, 667 (1979); C. L. Hammer and E. Shrauner, Phys. Rev. B **26**, 1297 (1982); **29**, 232, 5107 (1984); **32**, 7308 (1985).

[4] T. D. Lee, Phys. Rev. **95**, 1329 (1954); G. Kallen and W. Pauli, Dgl. Danske Videnskab. Selskab Mat.-Fys. Medd. **30**, #7 (1955).

[5] Here and throughout, all energy denominators are understood so that the paths of energy integrations are always the limit from below the real energy axis, i.e., $\Sigma_k A(k)(\tilde{\omega}_k - a)^{\prime} = \lim_{\epsilon \rightarrow 0} \Sigma_k A(k)(\tilde{\omega}_k - a - i\epsilon)^{-\prime}$.

[6] Th. W. Ruijgrok, Physica **25**, 357 (1959).

[7] R. Amado, Phys. Rev. **122**, 697 (1961).

[8] A. Pagnamenta, J. Math. Phys. **6**, 955 (1965).

ERRATUM

Erratum: Factorization of the wave equation in a nonplanar stratified medium [J. Math. Phys. 29, 36 (1988)]

Vaughan H. Weston
Department of Mathematics, Purdue University, West Lafayette, Indiana 47907

(Received 2 August 1988; accepted for publication 17 August 1988)

On p. 37, right column, line 14, replace "$C(\mathscr{S})\times C[0,T]$ into itself" by "$C(\mathscr{S})\times C^{(0,\lambda)}[0,T]$ into $C(\mathscr{S})\times C[0,T]$," and on lines 15 and 28 replace "$C(\mathscr{S})\times C^{1}(0,T]$" by "$C(\mathscr{S})\times C^{(1,\lambda)}[0,T]$." For additional proof of corresponding statements see Lemma 1 of Appendix C of Ref. 1.

On p. 40, right column, line 14, replace "compact" by "bounded."

[1] V. H. Weston, "Factorization of the dissipative wave equation and inverse scattering," J. Math. Phys. **29**, 2205 (1988).