Spectral theory for unitary operators on a quaternionic Hilbert space

C. S. Sharma and T. J. Coulson
Birkbeck College, University of London, 43 Gordon Square, London WC1H OPD, United Kingdom

(Received 4 November 1986; accepted for publication 29 April 1987)

The spectral theorem for a unitary operator on a quaternionic Hilbert space is established and a number of related results are proved.

I. INTRODUCTION

From some very reasonable physical assumptions it is possible to conclude[1,2] that the mathematical model of nonrelativistic quantum mechanics should be based on a σ-complete orthomodular lattice. The most familiar example of such a lattice is provided by the lattice of subspaces of a Hilbert space and traditionally quantum mechanics has been done in a Hilbert space on the complex field. Though the model is not bad, its shortcomings are becoming increasingly obvious. Two questions naturally arise: (i) can we find a better model by using an orthomodular lattice which is not isomorphic with the lattice of subspaces of a complex and separable Hilbert space, and (ii) can we find a better model by using a separable Hilbert space on a field other than the complex field? This work is motivated by a desire to find a definitive answer to the second question. In the present paper we sharpen many of the results proved in Ref. 3 and present a simpler and yet a more rigorous account of the spectral theory of unitary operators on a quaternionic Hilbert space. It should be noted that since observables are l-valued functions (cf. Ref. 1) on the Borel algebra of the reals, the field on which the Hilbert space is defined must include the reals as a subfield. Quaternions, strictly speaking, do not form a field but are an associative division algebra and according to a very old theorem of Frobenius there are only three associative division algebras on the reals: namely that of the real numbers, the complex numbers, and the quaternions.[4] The only additional division algebra over the real numbers is a Cayley algebra, which is nonassociative.[4] Nonassociativity makes it difficult to define something corresponding to the positive definite Hermitian product on a linear space over a Cayley algebra and it is the Hermitian product that provides a Hilbert space with its distinctive structure.

After some early work on quaternionic quantum mechanics by Jauch and co-workers,[3,5] there was a considerable revival of interest in the subject following in particular two notable papers by Horwitz and Biedenharn[6] and by Adler.[7] Thus the subject of our paper is of considerable topical interest.

II. QUATERNIONS

Quaternions, hereafter to be denoted by $\mathbb{H}$, form a normed associative algebra over $\mathbb{R}$ and are best defined with the help of three distinct linearly independent abstract square roots of -1, which are denoted by symbols i, j, and k, whose products are defined by

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k,$$
$$jk = -kj = i, \quad ki = -ik = j.$$

It is easy to verify that $\mathbb{H}$ is a four-dimensional vector space over $\mathbb{R}$. A conjugation is defined on $\mathbb{H}$ by

$$1^* = 1, \quad i^* = -i, \quad j^* = -j, \quad k^* = -k.$$

It is easy to verify that $\mathbb{H}$ is a normed algebra with the norm defined by

$$\|q\| = (q^*q)^{1/2}.$$

In addition to the axioms of the norm, the norm satisfies, as in the complex case,

$$\|q_1q_2\| = \|q_1\| \, \|q_2\|.$$

One of the easiest concrete realizations of $\mathbb{H}$ is the algebra of (2×2) complex matrices of the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

with matrix addition and multiplication. The zero in $\mathbb{H}$ corresponds to the zero matrix and that the matrix is nonzero implies that its determinant $|a|^2 + |b|^2$ is nonzero also and its inverse is another matrix of the same form. Now $\mathbb{C}$ (that is, the complex numbers), and therefore also $\mathbb{R}$, can be regarded as a subfield of $\mathbb{H}$ with the canonical embedding $e: \mathbb{C} \to \mathbb{H}$ given by

$$c \mapsto \begin{pmatrix} c & 0 \\ 0 & \bar{c} \end{pmatrix}. \tag{2.1}$$

In this realization $1, i, j$, and k are given by

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$
$$j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{2.2}$$

We can now define a norm-preserving isomorphism of $\mathbb{R}^4$ with $\mathbb{H}$ by

$$(a,b,c,d) \mapsto a + bi + cj + dk.$$

In our realization unit quaternions constitute the symplectic group Sp(1), that is, we have the matrices

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad a,b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1, \tag{2.3}$$

and thus Sp(1) is the same as SU(2). Note that Sp(1) is isomorphic with S^3 and is the universal covering of the rotation group SO(3).

We now need the following basic lemma.

Lemma 2.1: Let q be a unit quaternion. There exists a unit quaternion p such that $p^{-1}qp = r + si$ with $r,s \in \mathbb{R}$, $s \geqslant 0$,

 0022-2488/87/091941-06$02.50

and $r^2+s^2=1$. Furthermore, r is the real part of q.

Proof: We use the representation of q by a unitary (2×2) matrix of the form

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},\quad a,b\in\mathbb{C},\quad |a|^2+|b|^2=1.$$

Hence there exists a unitary matrix $u=\left(\begin{smallmatrix} g & d \\ e & f \end{smallmatrix}\right)$ such that $u^{-1}qu$ is diagonal. We can normalize u and write it as $\exp(i\theta)p$, where θ is real and p is a unit quaternion, that is, a (2×2) unitary matrix with unit determinant. Hence $p^{-1}qp$ is diagonal and has the form $\left(\begin{smallmatrix} c & 0 \\ 0 & \bar{c} \end{smallmatrix}\right)$ with $c=r+si$. If $s\geqslant 0$ we are done, if $s<0$ then it is easy to see that with k as in (2.2) $k^{-1}p^{-1}qpk=\left(\begin{smallmatrix} \bar{c} & 0 \\ 0 & c \end{smallmatrix}\right)$, which corresponds to $\bar{c}$ by (2.1) and has the imaginary part positive, which is what we want.

To see that r is the real part of q, we write

$$r+si=p^{-1}qp \quad\text{or}\quad r=q-spip^{-1}.$$

Now

$$(pip^{-1})^2=pip^{-1}pip^{-1}=pi^2p^{-1}=-1.$$

Hence pip^{-1} and so also $spip^{-1}$ are pure imaginary. This proves that r is the real part of q. We are now finished with the proof.

Note that in quaternions we have three different linearly independent square roots of -1 and any linear combination of these with real coefficients is pure imaginary also. Note further that for any nonzero $p\in\mathbb{H}$, $q\mapsto p^{-1}qp$ is an automorphism of $\mathbb{H}$: it is not difficult to prove that every automorphism of $\mathbb{H}$ is of this form. Since $\mathbb{R}$ is the center of $\mathbb{H}$, $\mathbb{R}$ is invariant under every automorphism of $\mathbb{H}$. An automorphism of $\mathbb{H}$ is thus a rotation of $\mathbb{R}^4$ with the real axis invariant, that is, a rotation of $\mathbb{R}^3$ where $\mathbb{R}^3$ correponds to the purely imaginary part of $\mathbb{H}$.

If two quaternions q and t are related by $t=pqp^{-1}$ for some $p\in\mathbb{H}$, then we write $t\simeq q$. It is evident that $\simeq$ is an equivalence relation on $\mathbb{H}$.

III. HILBERT SPACE OVER $\mathbb{H}$

Let H be a vector space over $\mathbb{H}$. We can define a positive definite Hermitian form on H by

$$\langle\ ,\ \rangle: H\times H\to\mathbb{H},$$

$$\langle pu,qv\rangle=p\langle u,v\rangle q^*,\quad p,q\in\mathbb{H},\quad u,v\in H,$$

$$\langle u+v,w\rangle=\langle u,w\rangle+\langle v,w\rangle,$$

$$\langle u,v\rangle^*=\langle v,u\rangle,$$

$$\langle u,u\rangle=0 \quad\text{only if } u=0.$$

A vector space H over $\mathbb{H}$ which is complete in the metric topology induced by the positive definite Hermitian form on H is called a Hilbert space over $\mathbb{H}$.

Let H be a Hilbert space over $\mathbb{H}$. We say that a map

$$L: H\to H$$

is linear if and only if for all $u,v\in H$ and for all $p\in\mathbb{H}$

$$L(u+v)=Lu+Lv$$

and

$$L(pu)=pLu.$$

It is easy to prove a Riesz representation theorem for a complex Hilbert space and thus H is naturally isomorphic to its dual $\widetilde{H}$ as a complex space. It is then easy to define the adjoint L^* of a linear map L whose essential properties are

$$\langle u,Lv\rangle=\langle L^*u,v\rangle$$

and for its product LM with another linear map M

$$(LM)^*=M^*L^*.$$

A linear map N on H is said to be normal if $NN^*=N^*N$. A linear map A on H is said to be self-adjoint if $A^*=A$. A linear map U on H is said to be unitary if $U^{-1}=U^*$. A linear map from H to itself is called an operator on H.

IV. SOME BASIC PROPERTIES OF OPERATORS ON A QUATERNIONIC HILBERT SPACE

We prove a number of properties, some of which are true for operators on a Hilbert space on any field: when that is so we refer to a Hilbert space rather than a quaternionic Hilbert space in our propositions.

Proposition 4.1: Let L be a linear operator on a quaternionic Hilbert space.

(i) Let q be an eigenvalue of L, then so also is pqp^{-1} for each nonzero $p\in\mathbb{H}$.

(ii) Let ϕ be an eigenvector of L, then so also is $p\phi$, where p is as in (i) above.

Proof: $L\phi=q\phi$ implies that $Lp\phi=pL\phi=pq\phi=pqp^{-1}p\phi$.

Proposition 4.2: Let U be a unitary operator on a Hilbert space. Let q be an eigenvalue of U, then q^* is an eigenvalue of U^* and U and U^* have a common eigenvector with these eigenvalues.

Proof: Since $U\phi=q\phi$ we have $\phi=U^*U\phi=U^*q\phi=qU^*\phi$, that is, $U^*\phi=(1/q)\phi$. But $\langle\phi,\phi\rangle=\langle U\phi,U\phi\rangle=q\langle\phi,\phi\rangle q^*$, therefore $q^*q=1$ or $(1/q)=q^*$. Hence $U^*\phi=q^*\phi$.

Proposition 4.3: Let ϕ and ψ be two eigenvectors of the same unitary operator U on a quaternionic Hilbert space. Then either ϕ and ψ are orthogonal or ϕ and $p\psi$ have the same eigenvalue for some $p\in\mathbb{H}$.

Proof: Let $U\phi=q\phi$ and $U\psi=t\psi$. Then $\langle\phi,\psi\rangle=\langle U\phi,U\psi\rangle=q\langle\phi,\psi\rangle t^*$, which implies that either $\langle\phi,\psi\rangle=0$ or $q\langle\phi,\psi\rangle=\langle\phi,\psi\rangle t$ or $t=\langle\phi,\psi\rangle^{-1}q\langle\phi,\psi\rangle$. Take $p=\langle\phi,\psi\rangle$ then $Up\psi=pU\psi=pp^{-1}qp\psi=qp\psi$.

Proposition 4.4: Let ϕ and ψ be two eigenvectors of the same unitary operator U on a quaternionic Hilbert space belonging to the same eigenvalue q. Then $\langle\phi,\psi\rangle$ commutes with q.

Proof: $\langle\phi,\psi\rangle=\langle U\phi,U\psi\rangle=q\langle\phi,\psi\rangle q^*$ or $q\langle\phi,\psi\rangle=\langle\phi,\psi\rangle q$.

Proposition 4.5: Eigenvectors of a unitary operator U on a quaternionic Hilbert space belonging to the eigenvalue q and to the eigenvalue pqp^{-1} for any nonzero $p\in\mathbb{H}$ span the same subspace.

Proof: If ϕ is an eigenvector of U belonging to q, then $p\phi$ belongs to pqp^{-1}. Similarly a particular linear multiple of an eigenvector of U belonging to pqp^{-1} is an eigenvector of U belonging to q. Hence eigenvectors belonging to q and those belonging to pqp^{-1} span the same subspace.

Proposition 4.6: Let M be the subspace spanned by eigenvectors of a unitary operator U on a quaternionic Hilbert space belonging to the eigenvalue q. Then M has an orthonormal basis in which each member is an eigenvector of U belonging to the same eigenvalue.

Proof: Let $\phi_1,\phi_2,...,\phi_n$ be a basis of eigenvectors in M belonging to the same eigenvalue q. Construct an orthonormal set $\psi_1,\psi_2,...,\psi_n$ by the Gram–Schmidt process. Since all the coefficients used in the process are either real or real multiples of $\langle\phi_i,\phi_j\rangle$ which by Proposition 4.4 commutes with q, all vectors obtained by the process are eigenvectors of U belonging to the same eigenvalue q.

Remark: We have seen that $q = ptp^{-1}$ defines an equivalence relation $r\simeq t$. If we partition $\mathbb{H}$ into equivalence classes then we can use any member of a class to represent that class. By Lemma 2.1 the equivalence classes can be represented by points in the upper half of the complex plane. If q is an eigenvalue of a linear operator then linear multiples of the corresponding eigenvectors are also eigenvectors belonging to eigenvalues in the same equivalence class but not necessarily to the same eigenvalue and each member of an equivalence class is the eigenvalue for some linear multiple of an eigenvector belonging to some eigenvalue in the same class. Thus eigenspaces in a quaternionic Hilbert space do not correspond to a single eigenvalue but to an *eigenclass* of eigenvalues.

Proposition 4.7: Let U be a unitary operator on a Hilbert space H. Let M be the subspace of H spanned by eigenvectors of U belonging to the eigenvalues $+1$ and -1. Then a nonzero vector ψ in H satisfies $(U^* - U)\psi = 0$ if and only if $\psi\in M$.

Proof: $U\psi = U^*\psi$ implies that $U^2\psi = \psi$. Let $\phi = \psi + U\psi$ and $\chi = \psi - U\psi$.

Since $\psi\neq 0$, both ϕ and χ cannot be simultaneously zero. Hence there are three possibilities: (i) $\phi = 0$, (ii) $\chi = 0$, and (iii) $\phi\neq 0, \chi\neq 0$. In cases (i) and (ii) ψ is itself an eigenvector belonging to the eigenvalues -1 and $+1$, respectively, in either case $\psi\in M$. For case (iii)

$$U\phi = U\psi + U^2\psi = \psi + U\psi = \phi$$

and

$$U\chi = U\psi - U^2\psi = -(\psi - U\psi) = -\chi.$$

Hence ϕ and χ are eigenvectors of U belonging to the eigenvalues $+1$ and -1, respectively. But $\psi = \frac{1}{2}(\phi + \chi)$. Hence $\psi\in M$. Conversely if $\psi\in M$, $\psi = \Sigma_i q_i\phi_i + \Sigma_j t_j\chi_j$, where the ϕ_i's are eigenvectors of U belonging to $+1$ and χ_j's to -1. By Proposition 4.2, these are also eigenvectors of U^* belonging to the same eigenvalues which implies that the restriction to M of U and U^* are identical. Hence $U\psi = U^*\psi$.

Corollary 4.7.1: With U and M as in Proposition 4.7, $(U - U^*)\psi = 0$ for

$$\psi\in M^\perp \Leftrightarrow \psi = 0.$$

Proof: Follows immediately from Proposition 4.7.

V. THE SPECTRAL THEOREM FOR UNITARY OPERATORS ON A QUATERNIONIC HILBERT SPACE

A quaternionic Hilbert space H is also a Hilbert space H_c over $\mathbb{C}$. If H is regarded as a Hilbert space H_c over $\mathbb{C}$, then for any vector ϕ and $j\phi$ are no longer linearly dependent but mutally perpendicular (in this context k has the representation $k = ij = -ji$) and j, since it is neither a vector nor a member of the field $\mathbb{C}$, must now be regarded as an antilinear operator from H_c to H_c that takes every vector into a vector perpendicular to itself and is such that $j^* = -j$ and $j^2 = -I$, where I is the identity operator on H_c. With these remarks we prove a few propositions on a quaternionic Hilbert space regarded as a complex Hilbert space. However, we first need a definition.

Definition 5.1: Let $\phi\in H_c$. Here ϕ is said to be an ϵ-approximate eigenvector belonging to the approximate eigenvalue λ of the operator U if and only if

$$||(U - \lambda I)\phi|| < \epsilon.$$

Proposition 5.1: Let H_c be a quaternionic Hilbert space regarded as a complex Hilbert space. Let E be the projection on a subspace M. Then $-jEj$ is a projection on the subspace $N = \{j\phi : \phi\in M\}$.

Proof: Since $(-jEj)^2 = -jEj$ and $(-jEj)^* = -jEj$, $-jEj$ is a projection. Let $E\phi = \phi$. Then $-jEjj\phi = jE\phi = j\phi$. Conversely let $-jEj\psi = \psi$. Now let $\phi = -j\psi$, then $\psi = j\phi$ and $E\phi = -Ej\psi = jjEj\psi = -j\psi = \phi$. We are done with the proof.

Proposition 5.2: (a) Let E be the projection on the eigenspace of a unitary operator U on H_c belonging to the eigenvalue λ, then $-jEj$ is the projection on the eigenspace of U belonging to the eigenvalue $\bar{\lambda}$.

(b) Let E be the projection on the ϵ-approximate eigenspace of a unitary operator U on H_c belonging to the approximate eigenvalue λ, then $-jEj$ is the projection on the ϵ-approximate eigenspace belonging to the approximate eigenvalue $\bar{\lambda}$.

Proof: (a) Let $E\phi = \phi$. Then $U\phi = \lambda\phi$ and $Uj\phi = ju\phi = -j\lambda jj\phi = \bar{\lambda}\phi$. Hence $j\phi$ belongs to the eigenvalue $\bar{\lambda}$ of U.

(b) Let $E\phi = \phi$. Then $||u\phi - \lambda\phi|| < \epsilon$ and

$$||Uj\phi - \bar{\lambda}j\phi|| = ||Uj\Phi + j\lambda jj\phi||$$

$$= ||j(U\phi - \lambda\phi)|| = ||j||\,||U\phi - \lambda\phi|| < \epsilon.$$

We are finished with the proof.

Corollary 5.2.1: The spectrum $\Lambda(U)$ of U regarded as an operator on H_c is such that $\lambda\in\Lambda(U)\Leftrightarrow\bar{\lambda}\in\Lambda(U)$.

Proof: Since the spectrum of a normal operator on a complex Hilbert space is identical with its approximate point spectrum,[8] our corollary follows immediately from Proposition 5.2.

Let $\{\hat{E}_\lambda\}$ be the spectral family for U regarded as an operator on H_c. Then by the spectral theorem for unitary operators on a complex Hilbert space we have

$$U = \int_0^{2\pi} \exp(i\lambda)\, d\hat{E}_\lambda, \tag{5.1}$$

where 0 and 2π are regarded as the same point and counted only once in the integration. As a consequence of Corollary 5.2.1, $\hat{E}_{2\pi} - \hat{E}_{2\pi-\lambda} = -j\hat{E}_\lambda j$ for $\lambda < \pi$ and $d\hat{E}_{2\pi-\lambda} = jd\hat{E}_\lambda j$. When H is regarded as a quaternionic Hilbert space, we drop the hats from the projection operators and then image spaces of E_λ and $E_{2\pi} - E_{2\pi-\lambda}$ are the same and they must be regarded as the same projection operator and we have

$$dE_{2\pi-\lambda} = -dE_\lambda. \quad (5.2)$$

However, in formula (5.1) the multiplying factor for dE_λ and $dE_{2\pi-\lambda}$ are $\exp(i\lambda)$ and $\exp(-i\lambda)$, respectively, and account must be taken of this in finding the correct spectral representation of U as an operator on H rather than on H_c [note that λ throughout is real and $\exp(i\lambda)$ complex]. To this end we define an operator $(\exp(i\lambda)E_\lambda)$ on H by the following device: this operator is well defined on H_c and its definition is extended H by requiring that if $E_\lambda\phi = \phi$, then

$$(\exp(i\lambda)E_\lambda)q\phi = q\exp(i\lambda)E_\lambda\phi = q\exp(i\lambda)\phi$$

for any $q\in\mathbb{H}$: thus our definition gives us a linear operator. A consequence of this definition is that if $E_\lambda\phi = \phi$, then

$$(\exp(i\lambda)E_\lambda)q\phi = \exp(i\lambda)(E_\lambda q\phi) \quad \text{if } q\in\mathbb{C},$$

but

$$(\exp(i\lambda)E_\lambda)q\phi = q(\exp(i\lambda)E_\lambda)\phi \neq \exp(i\lambda)(E_\lambda q\phi) \quad \text{if } q\notin\mathbb{C},$$

and in particular

$$(\exp(i\lambda)E_\lambda)j\phi = j\exp(i\lambda)E_\lambda\phi = -j\exp(i\lambda)\phi = -j\exp(i\lambda)jj\phi = \exp(-i\lambda)j\phi = \exp(-i\lambda)(E_\lambda j\phi). \quad (5.3)$$

Proposition 5.3: Let U be a unitary operator on H. Then

$$U = \int_0^\pi (\exp(i\lambda)dE_\lambda), \quad (5.4)$$

where $\{E_\lambda\}$ is the spectral family for U regarded as an operator on H_c and $(\exp(i\lambda)dE_\lambda)$ is an operator on H in the sense described in the preceding paragraph. [In this representation for each λ between 0 and π, $\exp(i\lambda)$ is on the upper unit semicircle in the complex plane and represents a whole equivalence class of unit quaternions in the set $\{\gamma^{-1}\exp(i\lambda)\gamma;\ \gamma\in\mathbb{H}\}$.]

Proof: Any vector $\psi\in\mathbb{H}$ can be written in the form

$$\psi = \phi_1 + j\phi_2, \quad (5.5)$$

where $\phi_1,\phi_2\in E_\pi(H_c)$ and $j\phi_2\in((E_{2\pi} - E_\pi)(H_c))$.

By the spectral theorem for U regarded as an operator on H_c, we can write

$$U\psi = \int_0^{2\pi} \exp(i\lambda)dE_\lambda\,\psi = \int_0^\pi \exp(i\lambda)dE_\lambda\,\phi_1 + \int_\pi^0 \exp(i(2\pi-\lambda))dE_{2\pi-\lambda}\,j\phi_2$$
$$= \int_0^\pi \exp(i\lambda)dE_\lambda\,\phi_1 - \int_0^\pi \exp(-i\lambda)dE_{2\pi-\lambda}\,j\phi_2$$
$$= \int_0^\pi \exp(i\lambda)dE_\lambda\,\phi_1 + \int_0^\pi (\exp(i\lambda)dE_\lambda)j\phi_2 = \int_0^\pi (\exp(i\lambda)dE_\lambda)\psi,$$

where we have used (5.2) and (5.4). Since ψ is any vector (5.4) follows. We are finished with the proof.

Though the result looks neat, its actual application requires going back to H_c and writing every vector in the form (5.5). This is inevitable because linear multiples of eigenvectors of U in H do not belong to the same eigenvalue and $\exp(i\lambda)$ is only one member of a whole eigenclass; the spectrum of U in H is not just a subset of the unit semicircle in the upper half of the complex plane but is a subset of the whole unit sphere in $\mathbb{H}$ ($\simeq\mathbb{R}^4$). These features are also present in Ref. 3 but are obscured by lack of precise definitions and some lengthy but unnecessary formulation of an operator J. Thus in order to use the theorem, the spectral decomposition of U in H_c, that is, all the eigenvalues and the eigenvectors, must be completely known, but for more abstract applications it may be enough to know that such a decomposition always exists.

The next proposition shows us how simple the calculation of the relationship between a general eigenvector in an equivalence class and its representative belonging to an eigenvalue on the upper unit semicircle in the complex plane is.

Proposition 5.4: Let U be a unitary operator on H. Let ϕ be an eigenvector of U belonging to the eigenvalue $q = q_0 + q_1 i + q_2 j + q_3 k$, which is not pure complex, that is, either q_2 or q_3 or both are nonzero. Then the eigenvalue belonging to the same eigenclass and situated on the upper unit semicircle of the complex plane is given by $c = c_0 + c_1 i$ with $c_0 = q_0$ and $c_1 = (1 - q_0^2)^{1/2}$ and the corresponding eigenvector is $\gamma\phi$, where $\gamma\in\mathbb{H}$ and is given by

$$\gamma = \gamma_0 + \gamma_1 i + \gamma_2 j + \gamma_3 k$$

with

$$\gamma_0 = ((c_1 + q_1)/2c_1)^{1/2}, \quad \gamma_1 = 0,$$
$$\gamma_2 = \gamma_0 q_3(c_1 - q_1)/(q_2^2 + q_3^2),$$

and

$$\gamma_3 = -\gamma_2 q_2/q_3.$$

Let ψ be an eigenvector of U belonging to the purely complex eigenvalue c. Then each member of the set $\{\gamma c\gamma^{-1}:\ \gamma\in\mathbb{H}\}$ belongs to an equivalence class of eigenvalues whose representative is c. The corresponding equivalence class of eigenvectors of U is the set $\{\gamma\psi:\ \gamma\in\mathbb{H}\}$ with $\gamma\psi$ belonging to the eigenvalue $\gamma c\gamma^{-1}$.

Proof: From Proposition 4.5 we know that c must be of the form $c = \gamma q\gamma^{-1}$ and since real numbers commute with quaternions, γ can be taken to be a unit quaternion. Lemma 2.1 tells us that $c_0 = q_0$ and c being the eigenvalue of a unitary operator must have unit modulus. Hence $c_1 = (1 - q_0^2)^{1/2}$. Lemma 2.1 also tells us that γ^{-1} is represented by the matrix whose columns are the normalized eigenvectors of the matrix $\begin{pmatrix} a & b \\ -\bar{b} & \bar{a}\end{pmatrix}$ representing q and belonging to the eigenvalues c and $\bar{c}$, respectively. From (2.2) we know that $a = q_0 + q_1 i$ and $b = q_2 + q_3 i$. Let the eigenvector of q belonging to the eigenvalue c be $\binom{x}{z}$, where we have chosen the phase in such a way that x is positive real. Then

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = c \begin{pmatrix} x \\ z \end{pmatrix}$$

is equivalent to

$$bz = (c - a)x \tag{5.6}$$

and

$$(c - \bar{a})z = -\bar{b}x. \tag{5.7}$$

Multiplying together (5.6) and the complex conjugate of (5.7) gives us

$$x^2 = -[(\bar{c} - a)/(c - a)]|z|^2, \tag{5.8}$$

which together with the normalization condition

$$x^2 + |z|^2 = 1 \tag{5.9}$$

gives

$$x = ((c_1 + q_1)/2c_1)^{1/2}. \tag{5.10}$$

Then (5.6) immediately gives us

$$z = (c - a)x/b. \tag{5.11}$$

Since γ^{-1} is a unit quaternion and its first column is $\binom{x}{z}$ we know from (2.3) that γ^{-1} is represented by $\begin{pmatrix} x & -\bar{z} \\ z & x \end{pmatrix}$ and therefore γ is represented by $\begin{pmatrix} x & \bar{z} \\ -z & x \end{pmatrix}$. Using (2.2) and (2.3) we immediately get the values of $\gamma_0, \gamma_1, \gamma_2$, and γ_3 as stated above in the Proposition. [Note that neither γ nor $\gamma\phi$ is unique—they both depend on a choice of phase. We can replace γ by γ' given by $\gamma' = \gamma \exp(ir)$, where r is any real number. Now $\gamma'\phi$ and $\gamma\phi$ belong to the same eigenvalue but differ in phase.]

The second part is an immediate corollary of Proposition 4.5.

VI. CONCLUDING REMARKS

In this section we shall briefly discuss why we prefer to treat our quaternionic Hilbert space as a left module (multiplication by scalars on the left) while most other authors dealing with applications of such a space to quantum mechanics treat it as a right module (multiplication by scalars on the right), and what happens to the matrix representation in our formalism. We shall also briefly discuss the relationship between the Hilbert spaces H and H_c in order to remove any unease the reader might have felt when H_c was introduced in Sec. IV.

Whether we treat a module over a ring as a right or a left module is purely a matter of convention or fashion. The two structures are isomorphic with each other and neither can have any inherent advantages or disadvantages. Vector spaces are very special cases of modules over a ring and there the convention that scalars are multiplied on the left is very firmly established. Having two different conventions for two very similar structures (one a particular case of the other) can be confusing. At the time the work on Refs. 3 and 5 was done, treating a module as a right module was very much in fashion among mathematicians and the authors were merely following the fashion current then. However, they put a linear operator on the left of a module rather than on the right as a mathematician working with a right module would have done. This had the advantage that in the matrix representation there was no need of taking the scalar across the matrix; linearity merely became a kind of associativity. The disadvantage of this approach is that one starts using notations and conventions different from those used by mathematicians and it becomes difficult for users of the two conventions to understand each other and have a meaningful dialogue. The present convention among mathematicians is very much in favor of treating a module as a left module, which is fortunately consistent with the overwhelming majority of papers on quantum mechanics where a Hilbert space is treated as a left vector space. A linear map on a quaternionic Hilbert space is a particular case of an R-module (R for ring) homomorphism and our approach follows the convention popular among mathematicians at the present time [see the highly authoritative texts by Jacobson[9] (p. 162) and by MacLane and Birkoff[10] (p. 193)]. It may, therefore, come as a surprise to the reader that even for commutative rings, while dealing with matrix representations, it is not unusual for mathematicians to put the matrix on the right of the vector [see Eq. (19) on p. 167 of Ref. 9 and Eq. (1) on p. 256 of Ref. 10]. So in our treatment, if we had to deal with matrix representations, we would put the scalar on the left, then the vector as a row vector followed by the matrix representation of the linear map; this is merely a transpose of the representation used in Ref. 3. Having the abstract objects appearing in the same order as their matrix counterparts as in Ref. 3 is, no doubt, an advantage but there is a small price to pay in that some obstruction is created in one's contact with the flow of the mainstream of mathematical literature. In conclusion one must say that conventions and fashions are merely a matter of personal preferences, in using the particular conventions in our work we are with the majority of mathematicians and with all those who are familiar with quantum mechanics in complex Hilbert space but in the very specialized field of applications of quaternions to quantum theory we are in a minority.

In trying to cope with different notations and conventions, one is liable to lose track of what is the crux of the problem in developing a spectral theory in a quaternionic Hilbert space and this is that the linear operators on an R-module do not themselves form an R-module unless the ring R over which the module is defined is commutative; thus linear multiples of linear operators are not in general linear operators and the primary concern of the spectral theory is to represent an operator as a sum or integral of linear multiples of projection operators. This is precisely the reason why in order to get anything like a spectral theorem in the present case there is no alternative but to view H as a complex space where the lack of commutativity disappears. It ought to be pointed out once again that $(\exp(i\lambda)E_\lambda)$ of Sec. V was defined in a very particular way to make it linear and sums of linear operators are, of course, linear in R-modules even though scalar multiples of linear operators are, in general, not linear.

A little ingenuity went into our work to ensure that after the spectral decomposition of U had acted on j times an eigenvector belonging to the purely complex eigenvalue

$(\exp(i\lambda))$, the eigenvalue appears in the final expression between j and the eigenvector. Considerable ingenuity is used in Ref. 3 to ensure the same thing. It is interesting to note that the same effect is achieved with remarkable ease in the Dirac notation, where the spectral decomposition of U takes the form

$$U = \int_0^\pi |\phi_\lambda\rangle \exp(i\lambda) d\lambda \, \langle\phi_\lambda|, \tag{6.1}$$

where $\{|\phi_\lambda\rangle\}$ is an orthonormal basis in the Dirac sense (that is, even points in the continuous spectrum have eigenvectors with δ-function normalizations) in H with eigenvalues on the upper unit semicircle in the complex plane. Even mathematicians appreciate the magic in the Dirac formalism, but so far it has not been found possible to produce a rigorous justification of it particularly in the case of continuous spectrum.

We now turn to the relation between H and H_c. Here H_c was defined by requiring that ϕ and $j\phi$ are mutually perpendicular for every ϕ in H and H_c is regarded as a complex Hilbert space—all the remaining structure is inherited from H with these conditions. If the inner product of two vectors in H is q which is a quaternion, then q has a unique decomposition in the form $q = c_0 + c_1 j$. It is easy to see that as a result of perpendicularity between ϕ and $j\phi$ the inner product in H_c is simply c_0. Because of the uniqueness of the decomposition of a quaternion as a complex number plus another complex number times j, this definition of inner product is unique and as all real numbers remain unchanged by the transformation which takes the inner product in H to the inner product in H_c, vectors continue to have the same norm in H_c and hence the topology of H_c is the same as that of H; and remembering that the underlying set of vectors for both H and H_c is the same and norms are preserved, unitary operators on H remain unitary when viewed as operators on H_c. A more interesting question is as follows: given a complex Hilbert space H_c, can we define a quaternionic Hilbert space H with the help of it? Even though the answer to this question has no relevance to the present work, we will give a brief answer. The answer is "yes" with the proviso that in the finite-dimensional case H_c must have even complex dimension and in the infinite-dimensional case the dimension must be countable. We find an antiunitary transformation J on H_c satisfying $\langle\phi, J\phi\rangle = 0$ for every ϕ in H_c and satisfying $J^2 = -I$, where I is the identity operator on H_c. We define, for complex c_0 and c_1 and the quaternionic j of Sec. II,

$$c_0\phi + c_1 j\phi = c_0\phi + c_2 J\phi, \tag{6.2}$$

$$\langle\phi, j\phi\rangle = \langle\phi,\phi\rangle j^*; \tag{6.3}$$

all other structures are inherited from H_c. Note that ϕ and $j\phi$ are now linearly dependent. This construction of a quaternionic space is very similar to the construction of a complex space from a real one and it is customary to call in that case the new structure a complex structure[11]; we can similarly call the structure we have just defined a "symplectic structure" on a complex Hilbert space.

It should perhaps be pointed out that the operator J used in the construction is not unique—if the dimension of the space is even in the finite-dimensional case and countable in the infinite-dimensional case, J can be defined in infinitely many ways: it is merely a question of arranging an orthonormal set in pairs and thus the definition of J in each case is necessarily coordinate dependent. It is easy to prove that a unitary operator on H_c continues to be unitary in H if and only if it commutes with J. For a given unitary operator U on H_c an operator J of the kind defined above and commuting with U exists if and only if $\exp(-i\lambda)$ belongs to the spectrum of U whenever $\exp(i\lambda)$ does.

ACKNOWLEDGMENT

It is a pleasure to acknowledge some helpful suggestions from the referee.

[1]G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (Benjamin, New York, 1963).
[2]C. S. Sharma, Phys. Lett. A **80**, 135 (1980).
[3]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).
[4]J. B. Fraleigh, *A First Course in Linear Algebra* (Addison–Wesley, Reading, MA, 1968).
[5]D. Finkelstein, J. M. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **3**, 207 (1962).
[6]L. P. Horwitz and L. C. Biedenharn, Ann. Phys. (NY) **157**, 432 (1984).
[7]S. L. Adler, Commun. Math. Phys. **104**, 611 (1986).
[8]P. R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (Chelsea, New York, 1957).
[9]N. Jacobson, *Basic Algebra I* (Freeman, San Francisco, 1974).
[10]S. MacLane and G. Birkoff, *Algebra* (Macmillan, New York, 1967).
[11]R. O. Wells, *Differential Analysis on Complex Manifolds* (Springer, New York, 1980).

Generating relations for reducing matrices. III. Kronecker products

R. Dirl and P. Kasperkovitz
Institut für Theoretische Physik, TU Wien, A-1040 Wien, Karlsplatz 13, Austria
M. I. Aroyo and J. N. Kotzev
Faculty of Physics, University of Sofia, BG-1126 Sofia, Bulgaria
B. L. Davies
Department of Applied Mathematics and Computation, University of North Wales, Bangor, Gwynedd, LL57 2UW, United Kingdom

(Received 16 December 1986; accepted for publication 13 May 1987)

In the first two of this series of papers [R. Dirl *et al.*, J. Math. Phys. **27**, 37 (1986); M. I. Aroyo *et al.*, *ibid.* **27**, 2236 (1986)] a systematic method to calculate the elements of reducing matrices has been developed. In this paper this "auxiliary group approach" is adapted to multiple Kronecker products. Three examples are worked out to illustrate the efficiency of generating relations and the proposed reduction of the multiplicity problem.

I. INTRODUCTION

This is the third of a series of papers in which we discuss how to calculate matrices that transform reducible (co)representations into direct sums of irreducible constituents in a systematic and efficient way. In the first two papers,[1,2] henceforth referred to as I and II, our "auxiliary group approach" has been introduced and discussed for arbitrary reducible (co)representations. In this paper we adapt the method to multiple Kronecker products.

Our approach is described in detail in Sec. II where, for the reader's convenience, some basic material from papers I and II is briefly reviewed. Sections III–IV contain three applications of our scheme. In the first example, where twofold Kronecker products of a space group are considered, the drastic reduction of the number of Clebsch–Gordan coefficients that actually have to be computed is most noteworthy. The other two examples, where threefold Kronecker products of magnetic point groups are discussed, serve mainly to illustrate the peculiar features arising from the algebraic structure of corepresentations.

II. THE AUXILIARY GROUP APPROACH TO THE REDUCTION OF KRONECKER PRODUCTS

A. Preliminaries

Before specifying the auxiliary group approach to Kronecker products we briefly review our notation and some basic facts needed later on; for details the reader is referred to papers I and II. In these papers transformations of unitary matrix (co)representations, of a group G, were introduced. A typical transformation q of a representation (rep) is composed of three different kinds of operations: an association, i.e., multiplication with a one-dimensional rep; an automorphism of the group resulting in a permutation of the matrices representing different group elements; and—for one-half of the considered transformations—the complex conjugation of the matrix elements. Although these transformations, denoted by the symbols $q_1,q_2,\ldots$, are defined for all (co)reps it is sufficient to consider only irreducible ones in the following. The set of all these transformations endowed with the natural multiplication law for bijective mappings forms the group $Q^{\rm rep}$ (transformation for ordinary reps) or $Q^{\rm co}$ (transformation of coreps), respectively. In the following we shall denote both these groups by the single symbol Q (thereby deviating from the notation of papers I and II) to avoid tedious repetitions in describing our approach.

The ordinary one-dimensional reps are uniquely determined by the multiplication law of the group G. However, if $G=G(H)=\{H,Ha_0\}$ is a magnetic group, its one-dimensional coreps are unique only up to a phase factor common to all elements $g\in Ha_0$. By means of a unitary transformation this phase can be varied in such a way that

$$D^j(a_0)=1 \quad (\dim D^j=1), \tag{2.1}$$

where a_0 is a fixed element of $G\setminus H$. The group ASS defined by the one-dimensional reps satisfying convention (2.1) is then uniquely determined by the reps $D^j(h)$, $h\in H$. We furthermore assume that the same element a_0, which is used to fix the phases of the one-dimensional coreps, is the one that is used to construct irreducible corepresentations (coirreps) of type III in Wigner canonical form.[3] For coirreps of type I and II it does not matter which element $a\in G\setminus H$ is used to construct the Wigner canonical form. This can be seen from the following list where for all three types Γ^k is assumed to be a unitary irreducible representation (irrep) of H, group elements are denoted by $h\in H$ and $a\in G\setminus H$, and M^T is the transpose of the matrix M,

type I:

$$\begin{aligned} &D^k(h)=\Gamma^k(h),\quad D^k(a)=Z^k(a),\\ &Z^k(a^{-1})=Z^k(a)^T,\\ &Z^k(a_1)\Gamma^k(h)^*Z^k(a_2)^*=\Gamma^k(a_1ha_2). \end{aligned} \tag{2.2}$$

type II:

$$D^k(h)=\begin{pmatrix}\Gamma^k(h) & 0\\ 0 & \Gamma^k(h)\end{pmatrix},$$

$$D^k(a)=\begin{pmatrix}0 & Z^k(a)\\ -Z^k(a) & 0\end{pmatrix},$$

$$Z^k(a^{-1})=-Z^k(a)^T, \tag{2.3}$$

$$Z^k(a_1)\Gamma^k(h)^*Z^k(a_2)^* = -\Gamma^k(a_1ha_2).$$

type III:

$$D^k(h) = \begin{pmatrix} \Gamma^k(h) & 0 \\ 0 & \Gamma^k(a_0^{-1}ha_0)^* \end{pmatrix}, \quad D^k(a) = \begin{pmatrix} 0 & \Gamma^k(aa_0) \\ \Gamma^k(a_0^{-1}a)^* & 0 \end{pmatrix}. \tag{2.4}$$

It is evident from (2.4) that for a coirrep $D^k(g)$ of type III convention (2.1) entails that the submatrices occurring in the coirrep $D^j(g)D^k(g) = D^l(g)$, which is also of type III, are related in exactly the same way as in $D^k(g)$. These conventions, to which we adhere in the following, have already been used in the previous paper (cf. Sec. IV of II) but not stated explicitly.

Convention (2.1) also influences the automorphisms which are included in the auxiliary group. As has been stated in paper II we only consider automorphisms of magnetic groups $G(H)$ that leave the subgroup H invariant. To be consistent with (2.1) we furthermore only admit automorphisms β for which

$$D^j(\beta^{-1}(a_0)) = 1 \tag{2.5}$$

holds true for all one-dimensional coirreps.

Transformation of an irrep D^k with one of the mappings $q\in Q$ yields a second one which is in general different from the first one and may even be inequivalent to it. This fact can be used to combine the irreps into disjoint subsets, called Q-classes, whose members are related by equations of the form $D^l = qD^k$ $(l \not\sim k)$. The same can be done for coirreps where the above mentioned conventions together with the restriction (2.5) ensure that qD^k is in Wigner canonical form if D^k is in that form. In the following we shall always assume that such a "standard set" of (co)irreps has already been constructed (cf. Sec. II G of paper I).

If an operation q transforms a (co)rep D^k into an equivalent one then there exists a unitary matrix $U^k(q)$ such that

$$(qD^k)(g) = U^k(q)^\dagger D^k(g)U^k(q)^{(g)}. \tag{2.6}$$

The meaning of the superscript (g) is fixed by the following definitions:

$$\text{reps of } G\text{: } M^{(g)} = M, \qquad \text{coreps of } G(H)\text{: } M^{(g)} = \begin{cases} M & \text{for } g\in H, \\ M^* & \text{for } g\in G\setminus H. \end{cases} \tag{2.7}$$

The transformation $D^k \to qD^k$ $(\sim D^k)$ determines the matrix $U^k(q)$ only to a certain extent: If the unitary matrices M^k and M_q^k belong to the commuting algebras of D^k and qD^k, respectively, i.e.,

$$M^kD^k(g) = D^k(g)M^{k(g)}, \qquad M_q^k(qD^k)(g) = (qD^k)(g)M_q^{k(g)} \tag{2.8}$$

for all $g\in G$, then the matrix $U^k(q)$ in (2.6) can be replaced by $M^{k\dagger}U^k(q)M_q^k$. The whole freedom in the definition of $U^k(q)$ is, however, already contained in one of the two commuting algebras. This follows from the fact that, for a fixed matrix $U^k(q)$, each left factor $\overline{M}^k$ may be replaced by a right factor $\overline{M}_q^k$ related to $\overline{M}^k$ by

$$U^k(q)\overline{M}_q^k = \overline{M}^kU^k(q) \tag{2.9}$$

and vice versa. From Schur's lemma and its generalization[4] the commuting algebras of (co)irreps are known to be isomorphic to (skew)fields of characteristic 0 and their unitary elements form matrix groups isomorphic to well-known compact groups [cf. Eqs. (3.21) of paper II]. (See Table I.)

In the following we shall assume that sufficiently many matrices $U^k(q)$ have been fixed by some convention for the representative D^k of each Q class. Here the number of q's has to be chosen in such a way that these transformations generate the group Q^k consisting of all transformations satisfying (2.6) (cf. Sec. II G of paper I),

$$q\in Q^k(\subset Q) \quad \text{iff } qD^k \sim D^k. \tag{2.10}$$

From now on we shall reserve the letter k for the Q-classes and the corresponding representative (co)irreps. An arbitrary (co)irrep is denoted by D^l; it is generated from the representative D^k by a relation of the form

$$D^l = q_l^{(k)}D^k, \tag{2.11}$$

where

$$Q = q_{l_1}^{(k)}Q^k \cup q_{l_2}^{(k)}Q^k \cup \cdots, \qquad R^k = \{q_{l_1}^{(k)}, q_{l_2}^{(k)}, \ldots\} \tag{2.12}$$

= fixed set of coset representatives,

$q_{l_1}^{(k)} = q_0$ (identical transformation).

For $D^l \sim qD^k$ we write, in a shortened form, $l = qk$, and the Q-class with representative D^k is denoted by

$$[k] = \{l \mid l = q^k \text{ for some } q\in Q\}. \tag{2.13}$$

It follows from (2.12) that for each $q\in Q$ and $q_l^{(k)}\in R^k$ there exist elements $q_{l'}^{(k)}\in R^k$ and $q'\in Q^k$ such that the following relation holds true:

$$qq_l^{(k)} = q_{l'}^{(k)}q', \quad q\in Q, \quad q'\in Q^k. \tag{2.14}$$

Note that the transformations on the rhs are uniquely determined by those on the lhs,

$$l' = ql, \quad q' = [q_{l'}^{(k)}]^{-1}qq_l^{(k)}, \tag{2.15}$$

$$k = \text{representative of the } Q\text{-class containing } l. \tag{2.16}$$

Equation (2.14) allows us to express the transformation $D^l \to qD^l$ in terms of coset representatives and the matrices $U^k(q)$ appearing in (2.6),

$$(qD^l)(g) = U^{l',l}(q)^\dagger D^{l'}(g)U^{l',l}(q)^{(g)}, \tag{2.17}$$

$$U^{l',l}(q) = q_l^{(k)}U^k(q') \quad \text{[cf. (2.12)]}. \tag{2.18}$$

In Eq. (2.18) we used the definition

TABLE I. Commuting algebras for (co)irreps.

Representation D^k	irrep	coirrep of type I	III	II
Commuting algebra	$\mathbb{C}$	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{Q}$
Unitary elements	U(1)	S_2	SO(2)	SU(2)
Freedom in $U^k(q)$	one real parameter	one sign	one real parameter	three real parameters

$$q^M = \begin{cases} M^*, & \text{if } q \text{ contains the complex conjugation,} \\ M, & \text{otherwise,} \end{cases} \tag{2.19}$$

valid for all matrices that do not represent elements of G $[M \neq D(g)]$.

B. The auxiliary group for Kronecker products

In this paper the reducible (co)reps of papers I and II are specialized to Kronecker products of the form

$$D^{\mathbf{k}}(g) = D^{k_1}(g) \otimes \cdots \otimes D^{k_n}(g). \tag{2.20}$$

The tensor product on the rhs of this equation means that the rows (columns) of $D^{\mathbf{k}}$ are labeled by $\mathbf{s} = (s_1, s_2, \ldots)$ if the rows (columns) of the factors D^{k_1}, $D^{k_2}, \ldots$ are labeled by $s_1, s_2, \ldots$, respectively, and that

$$D^{\mathbf{k}}_{\mathbf{s},\mathbf{s}'}(g) = D^{k_1}_{s_1,s_1'}(g) \cdots D^{k_n}_{s_n,s_n'}(g) \tag{2.21}$$

[cf. Eq. (2.49) of paper I].

If the (co)rep (2.20) were considered just as an ordinary reducible (co)representation nothing would emerge that has not yet been described in the preceding papers. However, the structure of $D^{\mathbf{k}}$ allows us to proceed. To this end we introduce a new group with elements

$$\mathbf{q} = (q_1, \ldots, q_n), \tag{2.22}$$

the q_i's being transformations of the kind considered before. The action of $\mathbf{q}$ on $D^{\mathbf{k}}$ is defined by

$$\mathbf{q}D^{\mathbf{k}} = q_1 D^{k_1} \otimes \cdots \otimes q_n D^{k_n}. \tag{2.23}$$

These transformations form the direct product group

$$Q^{(n)} = Q \times \cdots \times Q. \tag{2.24}$$

In view of the applications we have in mind we need also a group $\mathbf{Q}$ that is, in part, generated by the following transformations $\mathbf{q} \in \mathbf{Q}$:

$$\mathbf{a} = (q_1, \ldots, a_n), \quad a_i \in \text{ASS}, \tag{2.25}$$

$$\mathbf{b} = (b, \ldots, b), \quad b \in \text{AUT}, \tag{2.26}$$

$$\mathbf{c} = (c, \ldots, c), \quad c \in \text{CON}. \tag{2.27}$$

Here the group ASS (associations), AUT (automorphisms), and CON (complex conjugation), are those defined in the preceding papers. Besides the transformations (2.25)–(2.27) we consider one more kind of generating transformations, namely transformations $D^{\mathbf{k}} \to D^{\mathbf{k}'}$ related to permutations,

$$\begin{aligned} &\mathbf{p}D^{\mathbf{k}} = D^{\mathbf{k}'}; \\ &\mathbf{k} = (k_1, \ldots, k_n), \quad \mathbf{k}' = (k_{1'}, \ldots, k_{n'}); \\ &(1', \ldots, n') = \text{permutation of } (1, \ldots, n), \end{aligned} \tag{2.28}$$

i.e.,

$$p = \begin{pmatrix} 1, \ldots, n \\ 1', \ldots, n' \end{pmatrix}.$$

It is easy to see that the structure of the group $\mathbf{Q}$ generated by the transformations (2.25)–(2.28) is

$$\mathbf{Q} \cong \text{ASS}^{(n)} \circledS (\text{AUT} \times \text{CON} \times S_n), \tag{2.29}$$

where

$$\text{ASS}^{(n)} = \text{ASS} \times \cdots \times \text{ASS} \quad (n\text{-fold direct product}), \tag{2.30}$$

and S_n is the symmetric group containing all permutations of n objects.

There exists a natural homomorphism $\mathscr{H}$ from $\mathbf{Q}$ onto Q defined by

$$\begin{aligned} &\mathbf{a} = (a_1, \ldots, a_n)\colon \ \mathscr{H}(\mathbf{a}) = a_1, \ldots, a_n \ (n\text{-fold product}) \in \text{ASS}; \\ &\mathbf{b} = (b, \ldots, b)\colon \ \mathscr{H}(\mathbf{b}) = b \in \text{AUT}; \\ &\mathbf{c} = (c, \ldots, c)\colon \ \mathscr{H}(\mathbf{c}) = c \in \text{CON}; \\ &\mathscr{H}(\mathbf{p}) = q_0 \quad (\text{identical transformation}). \end{aligned} \tag{2.31}$$

In the following the images of the elements $\mathbf{q}, \mathbf{q}', \ldots, \in \mathbf{Q}$ will be denoted by $q, q', \ldots, \in Q$.

The transform $\mathbf{q}D^{\mathbf{l}}$ of a Kronecker product of irreps is obviously also a Kronecker product but contains in general nonstandard factors even if $D^{\mathbf{l}}$ is composed of standard irreps only. The relation between $D^{\mathbf{l}}$ and $\mathbf{q}D^{\mathbf{l}}$ is given by (2.28) for $\mathbf{q} = \mathbf{p}$; for $\mathbf{q} \in Q^{(n)}$ this relation follows from Eqs. (2.23) and (2.17). Using the definition

$$U^{\mathbf{l}',\mathbf{l}}(q) = U^{l_1',l_1}(q_1) \otimes \cdots \otimes U^{l_n',l_n}(q_n), \tag{2.32}$$

we can write this relation in the concise form

$$(\mathbf{q}D^{\mathbf{l}})(g) = U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})^\dagger D^{\mathbf{l}'}(g) U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})^{(g)}. \tag{2.33}$$

Equation (2.33) is the natural generalization of relation (2.17) to n-fold Kronecker products. Both these relations are needed in the derivation of generating and symmetry relations of Clebsch–Gordan matrices. A relation similar to (2.33) exists also for $\mathbf{q} = \mathbf{p}$, namely,

$$(\mathbf{p}D^{\mathbf{l}})(g) = U^{\mathbf{l}',\mathbf{l}}(\mathbf{p}) D^{\mathbf{l}}(g) U^{\mathbf{l}',\mathbf{l}}(\mathbf{p})^{\dagger(g)} = D^{\mathbf{l}'}(g), \tag{2.34}$$

where $U^{\mathbf{l}',\mathbf{l}}(\mathbf{p})$ is a permutation matrix depending only on the number of factors D^{l_i} and their dimensions. However, one should note the difference between (2.34) and (2.33): Equation (2.34) shows that the new rep $D^{\mathbf{l}'}$ may be obtained from the original rep $D^{\mathbf{l}}$ *either* by performing the substitution (2.28) *or* by transforming $D^{\mathbf{l}}$ with the permutation matrix $U^{\mathbf{l},\mathbf{l}'}(\mathbf{p})$. Equation (2.33), on the other hand, implies that for $\mathbf{q} \in Q^{(n)}$ one obtains $D^{\mathbf{l}'}$ from $D^{\mathbf{l}}$ in two steps, namely, *first* substituting $D^{\mathbf{l}}$ by $qD^{\mathbf{l}}$ and *afterwards* transforming it with $U^{\mathbf{l},\mathbf{l}'}(\mathbf{q})$.

Our ultimate goal is the systematic decomposition of all Kronecker products $D^{\mathbf{k}}$, each factor D^{k_i} ranging over a full set of standard irreps. The set of all these Kronecker products can be decomposed into subsets whose members are related by the auxiliary group $\mathbf{Q}$. To define this partition we use an equivalence relation between n-fold Kronecker products of the form $D^{\mathbf{k}} \approx D^{\mathbf{l}}$, where

$$D^{\mathbf{k}} \approx D^{\mathbf{l}} \quad \text{means } D^{k_i} \sim D^{l_i} \quad \text{for all } i. \tag{2.35}$$

This equivalence relation is obviously more stringent than general unitary equivalence ($D^{\mathbf{k}} \sim D^{\mathbf{l}}$) but at the same time weaker than complete equality ($D^{\mathbf{k}} = D^{\mathbf{l}}$). It is obvious from Eqs. (2.32), (2.33), and definition (2.35), that the equivalence relations defined by this relation are mapped onto each other under the transformations $\mathbf{q} \in Q^{(n)} \cap \mathbf{Q}$ and $\mathbf{q} = \mathbf{p}$, i.e., $\mathbf{q}D^{\mathbf{k}} \approx \mathbf{q}D^{\mathbf{l}}$ iff $D^{\mathbf{k}} \approx D^{\mathbf{l}}$. A typical subset defined by the equivalence relation considered here is

$$[\mathbf{h}] = \{D^{\mathbf{l}} | D^{\mathbf{l}} \approx \mathbf{q} D^{\mathbf{h}} \quad \text{for some } \mathbf{q} \in \mathbf{Q}\}. \tag{2.36}$$

Note that the reps $D^{\mathbf{l}}$ appearing on the rhs are all products of standard irreps D^{l}. As we did before for single irreps, we reserve the symbol $\mathbf{h}$ for the class representatives. The letter $\mathbf{h}$ has been chosen to indicate that the components h_i cannot always be chosen to be representatives of Q classes; however, we shall set $h_i = k_i$ whenever this is possible.

For each of the representatives $\mathbf{h}$ we furthermore define a subgroup of $\mathbf{Q}$ by

$$\mathbf{q} \in \mathbf{Q}^{\mathbf{h}} \ (\subset \mathbf{Q}) \quad \text{iff } \mathbf{q} D^{\mathbf{h}} \approx D^{\mathbf{h}} \tag{2.37}$$

[cf. definition (2.10)]. Finally we assume, in analogy to (2.12), that for each group $\mathbf{Q}^{\mathbf{h}} \subset \mathbf{Q}$ a set of coset representatives has been fixed by some convention,

$$\begin{aligned} &\mathbf{Q} = \mathbf{q}_{\mathbf{l}_1}^{(\mathbf{h})} \mathbf{Q}^{\mathbf{h}} \cup \mathbf{q}_{\mathbf{l}_2}^{(\mathbf{h})} \mathbf{Q}^{\mathbf{h}} \cup \cdots, \\ &\mathbf{R}^{\mathbf{h}} = \{\mathbf{q}_{\mathbf{l}_1}^{(\mathbf{h})}, \mathbf{q}_{\mathbf{l}_2}^{(\mathbf{h})}, \ldots\}, \\ &\mathbf{q}_{\mathbf{l}_1}^{(\mathbf{h})} = \mathbf{q}_0 \quad \text{(identity transformation).} \end{aligned} \tag{2.38}$$

C. Transformation properties of Clebsch–Gordan matrices

The decomposition of a Kronecker product $D^{\mathbf{l}}$ into (co)irreps D^{l}, obtained by means of a unitary matrix, reads

$$D^{\mathbf{l}}(g) C^{\mathbf{l}(g)} = C^{\mathbf{l}} \Big[\bigoplus_{l} E(\mathbf{l}|l) \otimes D^{l}(g) \Big] \quad \text{for all } g \in G. \tag{2.39}$$

In this equation C has been chosen for the reducing matrix to indicate that its elements are usually denoted as (generalized) Clebsch–Gordan coefficients. The matrices E appearing on the rhs are unit matrices,

$$E(d) = \text{unit matrix of dimension } d; \tag{2.40}$$

their dimensions are given by

$$(\mathbf{l}|l) = \text{multiplicity of } D^{l} \text{ in } D^{\mathbf{l}}. \tag{2.41}$$

In accordance with the definition of the Kronecker product, Eq. (2.21), the rows of $C^{\mathbf{l}}$ are labeled by $\mathbf{s} = (s_1, \ldots, s_n)$, where s_i is the row index of D^{l_i},

$$\begin{aligned} &\text{(co)irrep label: } l, \\ &\text{row index: } s = 1, \ldots, n_l, \quad (2.42) \\ &n_l = \dim D^{l}. \quad (2.43) \end{aligned}$$

As is evident from Eq. (2.39), the definition of the tensor product [paper I, Eq. (2.49)], and that of a direct sum of matrices, the columns of $C^{\mathbf{l}}$ are labeled by the triple (l,m,s), the second index referring to multiple occurrence of D^{l},

$$\text{multiplicity index: } m = m(\mathbf{l},l) = 1, \ldots, (\mathbf{l}|l). \tag{2.44}$$

If corepresentations are chosen in Wigner canonical form it is often more convenient to replace the row index s by the pair σ, z where σ depends on the type of the coirrep and z labels the rows of the irreps Γ^{k} of H contained in the coirrep D^{k} of $G(H)$,

$$\begin{aligned} &\text{row index of coirreps: } \sigma, z (\hat{=} s), \\ &\text{type I: } \sigma = 1; z = 1, \ldots, n_k, \\ &\text{type II or III: } \sigma = 1,2; \quad z = 1, \ldots, \tfrac{1}{2} n_k. \end{aligned} \tag{2.45}$$

Just as the matrix $U^{k}(q)$ is not uniquely determined by relation (2.6) so the Clebsch–Gordan matrix $C^{\mathbf{l}}$ is not uniquely fixed by (2.39). Reasoning as for $U^{k}(q)$, one sees that any two Clebsch–Gordan matrices satisfying (2.39) are related by right multiplication with a unitary matrix $M^{\mathbf{l}}$ which satisfies

$$M^{\mathbf{l}} \Big[\bigoplus_{l} E(\mathbf{l}|l) \otimes D^{l}(g) \Big] = \Big[\bigoplus_{l} E(\mathbf{l}|l) \otimes D^{l}(g) \Big] M^{\mathbf{l}(g)} \tag{2.46}$$

for all $g \in G$. This matrix has the structure

$$M^{\mathbf{l}} = \bigoplus_{l} M^{\mathbf{l},l}, \tag{2.47}$$

where the submatrices $M^{\mathbf{l},l}$ depend on the multiplicity $(\mathbf{l}|l)$ and the type of the coirrep (cf. Sec. III of paper II),

irreps:

$$M^{\mathbf{l},l}_{ms,m's'} = L_{m,m'} \delta_{s,s'}, \qquad L \text{ unitary}, \quad \dim L = (\mathbf{l}|l). \tag{2.48}$$

coirreps:

$$M^{\mathbf{l},l}_{m\sigma z,m'\sigma' z'} = L_{m\sigma,m'\sigma'} \delta_{z,z'}, \qquad L \text{ unitary}, \quad \dim L = (\mathbf{l}|l)\tau_l. \tag{2.49}$$

$$\begin{aligned} &\text{type I: } \tau_{\mathrm{I}} = 1; \quad \sigma = \sigma' = 1, \\ &L_{m1,m'1} = L^{*}_{m1,m'1} \quad (L \text{ orthogonal}). \end{aligned} \tag{2.50}$$

$$\begin{aligned} &\text{type II: } \tau_{\mathrm{II}} = 2; \quad \sigma, \sigma' \in \{1,2\}, \\ &L_{m1,m'1} = L^{*}_{m2,m'2}, \quad L_{m1,m'2} = -L^{*}_{m2,m'1}. \end{aligned} \tag{2.51}$$

$$\begin{aligned} &\text{type III: } \tau_{\mathrm{III}} = 2; \quad \sigma, \sigma' \in \{1,2\}, \\ &L_{m1,m'1} = L^{*}_{m2,m'2}, \quad L_{m1,m'2} = L_{m2,m'1} = 0. \end{aligned} \tag{2.52}$$

As in the preceding papers we now transform Eq. (2.39) into equivalent ones by a sequence of steps consisting of the following operations: (i) Both sides of Eq. (2.39) are multiplied with $D^{\mathbf{j}}(g) = D^{j_1}(g) \cdots D^{j_n}(g)$, $\dim D^{j} = 1$, D^{j} in standard form [see Eq. (2.1)]. (ii) The argument of the functions $D^{\mathbf{l}}(g)$ and $D^{l}(g)$ is substituted by $\beta^{-1}(g)$, where $g \to \beta(g)$ is an automorphism of the group G. (iii) Both sides of (2.39) are conjugated, i.e., $M \to M^{*}$ for all matrices. Comparing these operations with the generators of $\mathbf{Q} \cap Q^{(n)}$, Eqs. (2.25)–(2.27), we see that for each $\mathbf{q} \in \mathbf{Q} \cap Q^{(n)}$ the set of Eqs. (2.39) is transformed into a similar set of equations that is, in general, different from the original one. Using the definitions (2.23) and (2.19) and the homomorphism $\mathbf{Q} \to Q$ defined by (2.31) the equation obtained from (2.39) by applying the operation corresponding to $\mathbf{q} \in \mathbf{Q} \cap Q^{(n)}$ reads

$$(\mathbf{q} D^{\mathbf{l}})(g) (q C^{\mathbf{l}})^{(g)} = (q C^{\mathbf{l}}) \Big[\bigoplus_{l} E(\mathbf{l}|l) \otimes (q D^{l})(g) \Big]. \tag{2.53}$$

By means of relations (2.33) and (2.17) this can be transformed into

$$\begin{aligned} &U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})^{\dagger} D^{\mathbf{l}'}(g) U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})^{(g)} (q C^{\mathbf{l}})^{(g)} \\ &\quad = (q C^{\mathbf{l}}) \Big[\bigoplus_{l} E(\mathbf{l}|l) \otimes U^{l',l}(q)^{\dagger} D^{l'}(g) U^{l',l}(q)^{(g)} \Big]. \end{aligned} \tag{2.54}$$

Multiplying with $U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})$ from the left and with the matrix $Z^{\mathbf{l}}(q)^{(g)\dagger}$ from the right, where

$$Z^{\mathbf{l}}(q) = \bigoplus_{l} E(\mathbf{l}|l) \otimes U^{l',l}(q) \tag{2.55}$$

and $l' = ql$ [cf. Eqs. (2.18) and (2.15)], we obtain finally

$$D^{\mathbf{l}'}(g)\{U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})(qC^{\mathbf{l}})Z^{\mathbf{l}}(q)^{\dagger}\}^{(g)} = \{U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})(qC^{\mathbf{l}})Z^{\mathbf{l}}(q)^{\dagger}\}\left[\bigoplus_{l} E(\mathbf{l}|l)\otimes D^{l'}(g)\right]. \qquad (2.56)$$

This equation is also valid for $\mathbf{q}=\mathbf{p}$ because

$$D^{\mathbf{l}'}(g) = U^{\mathbf{l}',\mathbf{l}}(\mathbf{p})D^{\mathbf{l}}(g)U^{\mathbf{l}',\mathbf{l}}(\mathbf{p})^{\dagger(g)}$$

[cf. Eqs. (2.28) and (2.34)] and $\mathbf{p}\to q_0$, $q_0C^{\mathbf{l}} = C^{\mathbf{l}}$, $Z^{\mathbf{l}}(q)$ = identity matrix.

Comparison of (2.56) and (2.39) shows that the matrix in the curly brackets decomposes the Kronecker product $D^{\mathbf{l}'}$. It also states implicitly that $D^{\mathbf{l}'}$ contains $D^{l'}$ as often as $D^{\mathbf{l}}$ contains D^{l}, i.e.,

$$(\mathbf{l}|l) = (\mathbf{ql}|ql) = (\mathbf{l}'|l'); \qquad (2.57)$$

this follows also directly from the relation between the characters of D^k and qD^k. Because of (2.57) the matrix $E(\mathbf{l}|l)$ appearing in (2.56) can be written as $E(\mathbf{l}'|l')$. However, it should be noted that the order in which the (co)irreps $D^{l'}$ appear in (2.56) is determined by the label l ($= q^{-1}l'$) and not by l'. If we want the (co)irreps to appear in the direct sum in a certain standard order, i.e., a lexicographical order fixed by some convention once and for all, then the matrix in the curly bracket has to be multiplied from the right by a permutation matrix $P(q)$ with elements

$$P_{lms,l'm's'}(q) = \delta_{l,ql'}\delta_{m,m'}\delta_{s,s'}. \qquad (2.58)$$

$$D^{\mathbf{l}'}(g)\{U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})(qC^{\mathbf{l}})Z^{\mathbf{l}}(q)^{\dagger}P(q)\}^{(q)} = \{U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})(qC^{\mathbf{l}})Z^{\mathbf{l}}(q)^{\dagger}P(q)\} \times\left[\bigoplus_{l} E(\mathbf{l}'|l')\otimes D^{l'}(g)\right]. \qquad (2.59)$$

Comparing (2.59) with (2.39) one might be tempted to set the matrix in the curly bracket equal to $C^{\mathbf{l}'}$. However, this would lead to contradictory results since two different transformations $\mathbf{q}_1$ and $\mathbf{q}_2$ with $\mathbf{q}_1\mathbf{l} = \mathbf{q}_2\mathbf{l} = \mathbf{l}'$ would in general give rise to two different definitions of $C^{\mathbf{l}'}$. Keeping in mind the freedom in the definition of the Clebsch–Gordan matrices we can only conclude that

$$C^{\mathbf{l}'}M^{\mathbf{l}'}(\mathbf{q}) = U^{\mathbf{l}',\mathbf{l}}(\mathbf{q})(qC^{\mathbf{l}})Z^{\mathbf{l}}(q)^{\dagger}P(q), \qquad (2.60)$$

where the matrix $M^{\mathbf{l}'}(\mathbf{q})$ has the structure given by Eqs. (2.47)–(2.52). In the following we shall use Eq. (2.60) to establish generating relations by fixing the matrix $M^{\mathbf{l}'}(\mathbf{q})$ or parts of it for a carefully selected set of transformations $\mathbf{q}$. We shall also use this equation to reduce the freedom in the definition of Clebsch–Gordan matrices by requiring $M^{\mathbf{l}'}(\mathbf{q})$ to have a certain structure as $\mathbf{q}$ varies over various subgroups of $\mathbf{Q}$ (symmetry relations).

The first kind of generating relations relate all matrices $C^{\mathbf{l}}$, where $\mathbf{l}$ belongs to the $\mathbf{Q}$ class $[\mathbf{h}]$, to one single Clebsch–Gordan matrix $C^{\mathbf{h}}$.

$$\begin{aligned}
&\mathbf{q} = \mathbf{q}_{\mathbf{l}}^{(\mathbf{h})}\in\mathbf{R}^{\mathbf{h}},\ \mathscr{H}(\mathbf{q}) = q = q_{\mathbf{l}}^{(\mathbf{h})}:\\
&\quad \mathbf{l} = \mathbf{qh},\\
&\quad M^{\mathbf{l},l}(\mathbf{q}) = E(\mathbf{l}|l)\otimes E(n_l)\quad \text{for all } l \text{ with } (\mathbf{l}|l)\neq 0,\\
&\quad C^{\mathbf{l}} = U^{\mathbf{l},\mathbf{h}}(\mathbf{q})(qC^{\mathbf{h}})Z^{\mathbf{h}}(q)^{\dagger}P(q).
\end{aligned} \qquad (2.61)$$

Having exploited all relations obtained for the coset representatives $\mathbf{q}_{\mathbf{l}}^{(\mathbf{h})}\in\mathbf{R}^{\mathbf{h}}$ we are left with a problem as already considered in the previous papers. The only difference is that in the present case the matrix $U^{\mathbf{h},\mathbf{h}}(\mathbf{q})$ relating $D^{\mathbf{h}}$ and $\mathbf{q}D^{\mathbf{h}}$ need not be computed in addition to the matrices $U^{l',l}(q)$, but may be composed of a finite number of these matrices.

For the mathematical formulation of the subsequent procedure we need a few definitions. First we introduce the image of $\mathbf{Q}^{\mathbf{h}}$ under the homomorphism $\mathbf{Q}\to Q$,

$$Q^{\mathbf{h}} = \mathscr{H}(\mathbf{Q}^{\mathbf{h}}) \qquad (2.62)$$

and use it to define a new partition of the (co)irreps,

$$/k_x/_{\mathbf{h}} = \{l\,|\,l = qk_x,\quad q\in Q^{\mathbf{h}}\},\qquad \bigcup_x /k_x/_{\mathbf{h}} = [k]. \qquad (2.63)$$

For each representative k_x of one of these classes there exists a subgroup $\mathbf{Q}^{\mathbf{h},k_x}\subset Q^{\mathbf{h}}$ defined by

$$q\in Q^{\mathbf{h},k}\quad \text{iff } qD^k\sim D^k \quad \text{and } q\in Q^{\mathbf{h}}. \qquad (2.64)$$

We assume that sets of coset representatives with respect to these subgroups have been fixed by some convention,

$$\begin{aligned}
&Q^{\mathbf{h}} = q_{l_1}^{(\mathbf{h},k)}Q^{\mathbf{h},k}\cup q_{l_2}^{(\mathbf{h},k)}Q^{\mathbf{h},k}\cup\cdots,\\
&R^{\mathbf{h},k} = \{q_{l_1}^{(\mathbf{h},k)}, q_{l_2}^{(\mathbf{h},k)},\ldots\},\\
&q_{l_1}^{(\mathbf{h},k)} = q_0.
\end{aligned} \qquad (2.65)$$

Finally we split the square matrix $C^{\mathbf{h}}$ into rectangular blocks that belong to the (co)irreps D^{l} contained in $D^{\mathbf{h}}$,

$C^{\mathbf{h}}_l$ = rectangular matrix consisting of all columns of $C^{\mathbf{h}}$ with fixed index l
$[m = 1,\ldots,(\mathbf{h}|l);\ s = 1,\ldots,n_l]$. (2.66)

These definitions allow us to formulate the second kind of generating relations, namely those that relate the blocks $C^{\mathbf{h}}_l$ with $l\in/k_x/_{\mathbf{h}}$ to the block $C^{\mathbf{h}}_k$, $k = k_x$.

$$\begin{aligned}
&\mathbf{q}\in\mathbf{Q}^{\mathbf{h}},\ (\mathbf{q}) = q = q_l^{(\mathbf{h},k)}\in R^{\mathbf{h},k},\ k = k_x:\\
&\quad \mathbf{l} = \mathbf{l}' = \mathbf{h},\quad l = qk,\\
&\quad M^{\mathbf{h},l}(q) = E(\mathbf{h}|l)\otimes E(n_l),\\
&\quad C^{\mathbf{h}}_l = U^{\mathbf{h},\mathbf{h}}(\mathbf{q})(qC^{\mathbf{h}}_k)Z^{\mathbf{h}}_k(q)^{\dagger},\\
&\quad Z^{\mathbf{h}}_k(q) = E(\mathbf{h}|k)\otimes U^{l,k}(q).
\end{aligned} \qquad (2.67)$$

In the last step we consider the transformations $\mathbf{q}\in\mathbf{Q}^{\mathbf{h}}$ with $\mathscr{H}(\mathbf{q})\in Q^{\mathbf{h},k}$, $k = k_x$. These transformations form a group, the inverse image of $Q^{\mathbf{h},k}$ with respect to the homomorphism $\mathscr{H}$: $\mathbf{Q}\to Q$:

$$\mathbf{Q}^{\mathbf{h},k} = \mathscr{H}^{-1}(Q^{\mathbf{h},k})\subset\mathbf{Q}. \qquad (2.68)$$

The effect of these transformations on the block $C^{\mathbf{h}}_k$ can be seen from Eq. (2.60),

$$C^{\mathbf{h}}_k M^{\mathbf{h},k}(\mathbf{q}) = U^{\mathbf{h},\mathbf{h}}(\mathbf{q})(qC^{\mathbf{h}}_k)[E(\mathbf{h}|k)\otimes U^k(q)^{\dagger}]. \qquad (2.69)$$

To explain our strategy let us assume for the moment that the block $C^{\mathbf{h}}_k$ is given and has been split into $(\mathbf{h}|k)$ subblocks $C^{\mathbf{h}}_{km}$:

$C^{\mathbf{h}}_{km}$ = rectangular matrix consisting of all columns of $C^{\mathbf{h}}$ with fixed indices km $[s = 1,\ldots,n_k]$. (2.70)

Each of these subblocks satisfies the equation

$$D^{\mathbf{h}}(g)C^{\mathbf{h}(g)}_{km} = C^{\mathbf{h}}_{km}D^k(g)\quad \text{for all } g\in G. \qquad (2.71)$$

Furthermore, if $C^{\mathbf{h}}$ is unitary,

$$C^{\mathbf{h}\dagger}_{km} C^{\mathbf{h}}_{km'} = \delta_{m,m'} E(n_k). \tag{2.72}$$

It is easily verified that any linear combination

$$C^{\mathbf{h}}_{ka} = \sum_m C^{\mathbf{h}}_{km} L^k_{(m,a)}, \tag{2.73}$$

where the matrices $L^k_{(m,a)}$ belong to the commuting algebra of D^k,

$$D^k(g) L^{k(g)} = L^k D^k(g) \quad \text{for all } g \in G, \tag{2.74}$$

is also a solution of Eq. (2.71). As can be seen from the form of the matrices L^k, Eqs. (2.48)–(2.52) with $m = m' = 1$, the commuting algebra is isomorphic to a (skew) field: $\{L^k\} \cong \mathbb{R}$ for coirreps of type I, $\{L^k\} \cong \mathbb{Q}$ for coirreps of type II, and $\{L^k\} \cong \mathbb{C}$ for coirreps of type III and for irreps. The set of solutions of (2.71) is therefore a linear space over the corresponding (skew) field, its dimension being always $(\mathbf{h}|k)$ because of (2.72). Accordingly the number of free real parameters appearing in the most general subblock of the form (2.73) is $(\mathbf{h}|k)\rho_k$, where ρ_k is the number of real parameters needed to specify an element of the (skew) field ($\rho = 1$ for $\mathbb{R}$, $\rho = 2$ for $\mathbb{C}$, and $\rho = 4$ for $\mathbb{Q}$). That the most general solution of (2.71) has to be of the form (2.73) follows from the fact that the projection matrix

$$P^{\mathbf{h}}_k = C^{\mathbf{h}} \Big[\oplus_l \delta_{l,k} E(\mathbf{h}|l) \otimes E(n_l) \Big] C^{\mathbf{h}\dagger}$$

$$= C^{\mathbf{h}}_k C^{\mathbf{h}\dagger}_k = \sum_m C^{\mathbf{h}}_{km} C^{\mathbf{h}\dagger}_{km} \tag{2.75}$$

is uniquely determined by $D^{\mathbf{h}}$ and D^k and independent of the special choice of subblocks $C^{\mathbf{h}}_{km}$,

$$P^{\mathbf{h}}_k = P^{\mathbf{h}\dagger}_k = (P^{\mathbf{h}}_k)^2. \tag{2.76}$$

If $P^{\mathbf{h}}_k$ is given in terms of known subblocks $C^{\mathbf{h}}_{km}$ this allows one to calculate the expansion coefficients $L^k_{(m,a)}$ for a general solution of (2.71),

$$C^{\mathbf{h}}_{ka} = P^{\mathbf{h}}_k C^{\mathbf{h}}_{ka} = \sum C^{\mathbf{h}}_{km} C^{\mathbf{h}\dagger}_{km} C^{\mathbf{h}}_{ka} = \sum C^{\mathbf{h}}_{km} L^k_{(m,a)},$$
$$L^k_{(m,a)} = C^{\mathbf{h}\dagger}_{km} C^{\mathbf{h}}_{ka}. \tag{2.77}$$

Equations (2.72) and (2.77) show that the relation

$$C^{\mathbf{h}\dagger}_{ka} C^{\mathbf{h}}_{kb} = L^k_{(a,b)} \tag{2.78}$$

may be interpreted as a scalar product that assigns to any two "vectors" $C^{\mathbf{h}}_{ka}$, $C^{\mathbf{h}}_{kb}$ a "number" $L^k_{(a,b)}$ of the corresponding (skew) field. Since this product has all the usual properties of a scalar product the set of solutions of (2.71) can be considered as a unitary space (for the properties of the less familiar quaternionic Hilbert space see, e.g., Ref. 5). Note especially that in all cases the square of the norm of $C^{\mathbf{h}}_{ka}$ has the form

$$L^k_{(a,a)} = \gamma_a E(n_k), \quad \gamma_a = \gamma_a^* \geqslant 0. \tag{2.79}$$

Let us now consider the operators $T(\mathbf{q})$, $\mathbf{q} \in \mathbf{Q}^{\mathbf{h},k}$, defined by

$$T(\mathbf{q}) C^{\mathbf{h}}_{ka} = U^{\mathbf{h},\mathbf{h}}(\mathbf{q}) (q C^{\mathbf{h}}_{ka}) U^k(q)^\dagger. \tag{2.80}$$

Because of (2.78) and the unitarity of the matrices $U^{\mathbf{h},\mathbf{h}}$ and U^k these operators are certainly norm preserving. Furthermore,

$$T(\mathbf{q}) [C^{\mathbf{h}}_{ka} L^k] = [T(\mathbf{q}) C^{\mathbf{h}}_{ka}] L^{k\{q\}}, \tag{2.81}$$

where

$$L^{k\{q\}} = U^k(q) (q L^k) U^k(q)^\dagger. \tag{2.82}$$

If we transfer the transformation $q \in Q^k$ to Eq. (2.74) in exactly the same way as the transformations $\mathbf{q} \in \mathbf{Q} \cap Q^{(n)}$ were transferred to relation (2.39) we obtain

$$[q D^k(q)] [q L^k]^{(g)} = [q L^k] [q D^k(g)] \quad \text{for all } g \in G, \tag{2.83}$$

or, using (2.6) and (2.82),

$$D^k(g) L^{k\{q\}(g)} = L^{k\{q\}} D^k(g) \quad \text{for all } g \in G. \tag{2.84}$$

Therefore the matrices $L^{k\{q\}}$ also belong to the commuting algebra of D^k and it is obvious from (2.82) that

$$(L^k_1 + L^k_2)^{\{q\}} = L^{k\{q\}}_1 + L^{k\{q\}}_2,$$
$$(L^k_1 L^k_2)^{\{q\}} = L^{k\{q\}}_1 L^{k\{q\}}_2. \tag{2.85}$$

This shows that each mapping $L^k \to L^{k\{q\}}$ may be considered as an automorphism of the corresponding (skew) field. The number and the properties of these automorphisms are well known from the literature and can also be deduced from the form of the matrices L^k. For ordinary irreps and coirreps of type I these matrices are multiples of the unit matrix,

$$\text{irreps: } L^{k\{q\}} = q L^k, \tag{2.86}$$

$$\text{coirreps of type I: } L^{k\{q\}} = L^k. \tag{2.87}$$

For coirreps of type II where $\{L^k\} \cong \mathbb{Q}$ and all automorphisms of $\mathbb{Q}$ are inner ones one can always find unitary matrices $L^k(q)$ such that[5]

$$L^{k\{q\}} = L^k(q)^\dagger L^k L^k(q). \tag{2.88}$$

Thus if we use the matrices

$$\check{U}^k(q) = L^k(q) U^k(q) \tag{2.89}$$

instead of the matrices $U^k(q)$, they also satisfy (2.6) but

$$\check{L}^{k\{q\}} = \check{U}^k(q) (q L^k) \check{U}^k(q)^\dagger = L^k. \tag{2.90}$$

Likewise we could use matrices $\hat{U}^k(q) = \hat{L}^k(q) U^k(q)$ such that

$$\hat{L}^{k\{q\}} = \hat{U}^k(q) (q L^k) \hat{U}^k(q)^\dagger = q L^k \tag{2.91}$$

since $L^k \to (L^k)^*$ is also an automorphism of $\{L^k\} \cong \mathbb{Q}$. For coirreps of type II we therefore always choose one of the following two conventions for the matrices $U^k(q)$,

$$\text{coirreps of type II: } L^{k\{q\}} = L^k, \tag{2.92a}$$

$$L^{k\{q\}} = q L^k. \tag{2.92b}$$

For coirreps of type III a redefinition of $U^k(q)$, i.e., a substitution $U^k(q) \to L^k(q) U^k(q)$, has no effect on $L^{k\{q\}}$ since all matrices L^k commute. The automorphism $L^k \to L^{k\{q\}}$ is therefore uniquely determined by the transformation q and the coirrep D^k. Since there exists only one nontrivial automorphism of $\{L^k\} \cong \mathbb{C}$, namely $L^k \to (L^k)^*$, we arrive at the following result:

$$\text{coirreps of type III: } L^{k\{q\}} = L^k \text{ or } L^{k*}. \tag{2.93}$$

Summing up we see that for ordinary irreps and coirreps of type III the operators $T(\mathbf{q})$ are either linear or antilinear and, since they are also norm preserving, either unitary or antiunitary. For coirreps of type I they have to be linear and unitary, and so they are for coirreps of type II if the matrices $U^k(q)$ are properly chosen [Eqs. (2.89) and (2.90)]. An alternative choice of these matrices leads to operators that are linear/antilinear and unitary/antiunitary (in a generalized sense).

It follows from (2.56) and (2.73) that

$$T(\mathbf{q})C^{\mathbf{h}}_{km} = \sum_{m'} C^{\mathbf{h}}_{km'} L^{k}_{(m',m)}(\mathbf{q}). \quad (2.94)$$

The matrices $L^{k}_{(m,m')}(\mathbf{q})$, $m,m' = 1,...,(\mathbf{h}|k)$, of dimension n_k may be combined into a matrix of dimension $(\mathbf{h}|k)n_k$,

$$M^{\mathbf{h},k}(\mathbf{q}) = L^{\mathbf{h},k}(\mathbf{q}) \otimes E(n_k \tau_k^{-1}), \quad \tau_k = 1 \text{ or } 2 \quad (2.95)$$

[cf. Eqs. (2.48)–(2.52)], which represents the action of $T(\mathbf{q})$ on the basis blocks $C^{\mathbf{h}}_{km}$. Because of (2.81) the product $T(\mathbf{q}_1)T(\mathbf{q}_2)$ is represented by

$$M^{\mathbf{h},k}(\mathbf{q}_1) * M^{\mathbf{h},k}(\mathbf{q}_2) = M^{\mathbf{h},k}(\mathbf{q}_1) M^{\mathbf{h},k}(\mathbf{q}_2)^{\{q_1\}}, \quad (2.96)$$

where the rhs is either the ordinary product or a coproduct of the two matrices (note the definition of conjugation for type III coirreps!). If neither convention (2.92) nor (2.92b) is adopted for coirreps of type II Eq. (2.96) has to be considered as a generalization of the usual coproduct.

If for a transformation $\mathbf{q} \in \mathbf{Q}^{\mathbf{h},k}$ both matrices $U^{\mathbf{h},\mathbf{h}}(\mathbf{q})$ and $U^k(q)$ are given then the operator $T(\mathbf{q})$ is uniquely defined by (2.80). However, usually these matrices are known only for a few generators of $\mathbf{Q}^{\mathbf{h},k}$ so that a variety of products of the corresponding operators is needed to close them into an operator group $\widetilde{Q}^{\mathbf{h},k}$. The form of the matrix (co)representation of this group, $T(\mathbf{q}) \to M^{\mathbf{h},k}(\mathbf{q})$, or equivalently $T(\mathbf{q}) \to L^{\mathbf{h},k}(\mathbf{q})$, depends on which Clebsch–Gordan matrix is used to fix the basis blocks $C^{\mathbf{h}}_{km}$. We exploit this fact in requiring the following property.

$\mathbf{q} \in \mathbf{Q}^{\mathbf{h},k}$, $\mathscr{H}(\mathbf{q})q \in Q^{\mathbf{h},k}$, $k = k_x$:
the matrices $L^{\mathbf{h},k}(\mathbf{q})$ are direct sums of (co)irreps of the operator group $\widetilde{Q}^{\mathbf{h},k} = \{T(\mathbf{q})\}$. (2.97)

In general this requirement will not be fulfilled for a given set of subblocks if $(\mathbf{h}|k) > 1$ but it can always be achieved by a suitable unitary transformation. If all (co)irreps occurring in $L^{\mathbf{h},k}(\mathbf{q})$, $k = k_x$, are inequivalent then the multiplicity $(\mathbf{h}|k_x)$ may be explained in terms of the auxiliary operator group $\widetilde{Q}^{\mathbf{h},k_x}$. If at least one of the (co)irreps occurs more than once the multiplicity problem is only reduced but not completely solved.

D. Summary of the scheme

Although it should be clear from the preceding discussion how to proceed in a concrete calculation of Clebsch–Gordan coefficients we list once more the essential steps for those who are mainly interested in possible applications of our method.

(1) *Prerequisites:* The multiplication law of the given (finite or compact) group determines the inner automorphisms and gives hints for the existence (or nonexistence) of outer automorphisms. From this knowledge the group AUT (or a subgroup of it) may be constructed; note the restriction (2.5) for magnetic groups $G(H)$.

Next determine the simple characters of G [or H if $G = G(H)$]. They include one-dimensional irreps and hence the definition of the group ASS. Take care of convention (2.1) for magnetic groups.

Define the group Q ($= \mathbf{Q}^{\mathrm{rep}}$ or Q^{co}) by means of generating elements and use them to find the qk table (see papers I and II and the examples in Secs. III–V) and the Q-classes [Eq. (2.13)]. Choose representatives D^k and determine for each of them the subgroup $Q^k \subset Q$ [Eq. (2.10)] and a set of R^k of coset representatives [Eq. (2.12)]. Moreover construct for the generators of Q^k unitary matrices $U^k(q)$ relating D^k and qD^k according to Eq. (2.6).

(2) *Generating relations of the first kind:* Define the group $\mathbf{Q}$ for the n-fold Kronecker products of interest and use it to find the $\mathbf{Q}$-classes $[\mathbf{h}]$. Choose representatives $\mathbf{h}$ and find the subgroups $\mathbf{Q}^{\mathbf{h}}$ and coset representatives $\mathbf{q}^{\mathbf{h}}_{\mathbf{l}}$ [Eqs. (2.37) and (2.38)]. Then determine for the components of $\mathbf{q}^{(\mathbf{h})}_{\mathbf{l}}$, $q_i = (\mathbf{q}^{(\mathbf{h})}_{\mathbf{l}})_i$, of each of these transformations, the matrices $U^{l_i,h_i}(q_i)$. In doing so use the matrices $U^k(q)$ and the decomposition of Q into cosets of Q^k [Eqs. (2.14)–(2.16), (2.18)]. The tensor product $U^{\mathbf{l},\mathbf{h}}(\mathbf{q})$, $\mathbf{q} = \mathbf{q}^{(\mathbf{h})}_{\mathbf{l}}$, composed of these matrices is one of the three matrices needed to calculate the Clebsch–Gordan matrices $C^{\mathbf{l}}$, $\mathbf{l} \in [\mathbf{h}]$, from the matrix $C^{\mathbf{h}}$ [Eq. (2.61)]. The other two matrices are $Z^{\mathbf{h}}(q)$ and $P(q)$, where $q = \mathscr{H}(\mathbf{q}^{(\mathbf{h})}_{\mathbf{l}})$ [Eqs. (2.31)]. The matrix $P(q)$, Eq. (2.58), is easily found from the qk table. For $Z^{\mathbf{h}}(q)$, Eq. (2.55), one has to calculate the matrices $U^{l',l}(q)$, $l' = ql$, from the matrices $U^k(q)$ and the coset representatives $q^{(k)}_l \in R^k$ [Eqs. (2.15) and (2.18)].

(3) *Generating relations of the second kind:* For each $\mathbf{h}$ construct the classes $/k_x/_{\mathbf{h}}$ [Eq. (2.63)] and choose representatives k_x for each of them. Then determine the group $Q^{\mathbf{h},k}$ and coset representatives $R^{\mathbf{h},k}$ for each $k = k_x$ [Eqs. (2.64) and (2.65)]. Find for each $q \in R^{\mathbf{h},k}$ a transformation $\mathbf{q} \in Q^{\mathbf{h}}$ such that $\mathscr{H}(\mathbf{q}) = q$ according to (2.31). Construct for these transformations $\mathbf{q}$ the matrix $U^{\mathbf{h},\mathbf{h}}(\mathbf{q})$ as a product of matrices $q_1 U^k(q_2)$ [Eqs. (2.32) and (2.18)]. Finally use the matrices $U^{l,k}(q)$, constructed like the factors of $U^{\mathbf{h},\mathbf{h}}(\mathbf{q})$ before, to find the matrix $Z^{\mathbf{h}}_k(q)$ [Eq. (2.67)]. This enables one to calculate the blocks $C^{\mathbf{h}}_l$, $l \in /k/$, from the block $C^{\mathbf{h}}_k$, $k = k_x$, according to (2.67).

(4) *The multiplicity problem:* First calculate from the defining equation

$$D^{\mathbf{h}}(q) C^{\mathbf{h}(g)}_{ka} = C^{\mathbf{h}}_{ka} D^k(g) \quad \text{for all } g \in G, \quad (2.98)$$

the general form of $C^{\mathbf{h}}_{ka}$ for each pair $\mathbf{h}$, $k = k_x$. This rectangular matrix must contain $2(\mathbf{h}|k)$ free real parameters, if D^k is an irrep. If D^k is a coirrep the number of free real parameters in $C^{\mathbf{h}}_{ka}$ has to be N for type I, $4N$ for type II, and $2N$ for type III, where $N = (\mathbf{h}|k)$. This number of real parameters is reduced by one if $C^{\mathbf{h}}_{ka}$ is normalized to $E(n_k)$ according to Eq. (2.79). Then define a basis $\{C^{\mathbf{h}}_{km}\}$ by fixing the free parameters. Next determine the groups $\mathbf{Q}^{\mathbf{h},k}$ and choose a set of generators for each of them. For these transformations $\mathbf{q}$ find the operators $T(\mathbf{q})$ [Eq. (2.80)] and the structure of the operator group generated by them. Finally pass to a new basis that transforms according to (co)irreps of $\widetilde{Q}^{\mathbf{h},k}$.

Before discussing several examples treated along these lines we would like to emphasize once more what freedom is left in following the proposed rules. For instance, while the transformation groups Q and $\mathbf{Q}$ and all their subgroups are uniquely defined, the selection of generating elements or coset representatives is to a large extent arbitrary. Furthermore all the matrices U relating (co)irreps to equivalent ones are only determined up to factors that lie in the commuting algebra. Finally, a third kind of arbitrariness enters through the choice of the class representatives k, $\mathbf{h}$, and k_x.

III. TWOFOLD KRONECKER PRODUCTS FOR THE DOUBLE SPACE GROUP *P*23

In this example we consider twofold Kronecker products (KP's) for the double space group $P23$ (Ref. 6). Here we adopt the same notations and conventions as in Ref. 6 as long as they do not conflict with our auxiliary group approach. The present example has been briefly sketched in Ref. 7. Here we give a much more detailed discussion which should enable the reader to gain more insight into our procedure.

The double space group $G = P23$ is, as every symmorphic space group, a semidirect product

$$G = T \circledS P = \bigcup_{R \in P} T(R|\vec{0}) \tag{3.1}$$

with the multiplication law

$$(R|\vec{t})(R'|\vec{t}') = (RR'|\vec{t} + R\vec{t}') = (R''|\vec{t}''), \tag{3.2}$$

where $\vec{t} \in T$ (translation group) and $R \in P \simeq G/T$ (point group). For the sake of simplicity we choose the lattice constant equal to 1 so that the basic translations $\vec{t}_j, j = 1,2,3$ are normalized. An arbitrary translation $\vec{t}$ is an integral linear combination of the basic translations $\vec{t}_j$. The point group P assigned to $P23$ is the tetrahedral double point group consisting of 24 elements,

$$\begin{aligned} P = \{&E,\ C_{2x},\ C_{2y},\ C_{2z},\ C_{31}^-,\ C_{32}^-,\ C_{33}^-,\ C_{34}^-,\\ &C_{31}^+,\ C_{32}^+,\ C_{33}^+,\ C_{34}^+,\\ &\overline{E},\ \overline{C}_{2x},\ \overline{C}_{2y},\ \overline{C}_{2z},\ \overline{C}_{31}^-,\ \overline{C}_{32}^-, \overline{C}_{33}^-,\ \overline{C}_{34}^-,\\ &\overline{C}_{31}^+,\ \overline{C}_{32}^+,\ \overline{C}_{33}^+,\ \overline{C}_{34}^+\}. \end{aligned} \tag{3.3}$$

Space group irreps are sometimes called "standard" irreps if they are determined by induction (Ref. 8) out of the one-dimensional irreps of the translation group T. These standard space group irreps of $G = P23$ have the form

$$D^K_{\underline{R},\underline{S}}((R|\vec{t})) = \Delta^{\vec{q}}(\underline{R},R\underline{S})e^{-i\underline{R}\vec{q}\cdot\vec{t}}D^k(\underline{R}^{-1}R\underline{S}), \tag{3.4}$$

where in part a matrix notation has been adopted. The symbols $D^k(R')$, $R' \in P(\vec{q})$ denote matrix irreps of the "little cogroup" $P(\vec{q})$, where

$$P(\vec{q}) = \{R \in P | R\vec{q} = \vec{q} + \vec{Q}\} \cong G(\vec{q})/T, \tag{3.5}$$

$G(\vec{q})$ is the corresponding "little group" and Q is a vector of the reciprocal lattice. In detail our notation has to be understood as follows:

$K = (\vec{q},k)\uparrow G = (\vec{q},k)$,	standard G-irrep label,	(3.6)
$\vec{q} \in \Delta\mathrm{BZ}(G)$,	representation domain of the Brillouin zone BZ(G),	
$k \in A(\vec{q})$,	set of $P(\vec{q})$-irrep labels,	
$\underline{R},\underline{S} \in \underline{P}(\vec{q})$,	fixed set of coset representatives (CR's) for the decomposition of P with respect to $P(\vec{q})$,	
$A(G) = \{(\vec{q},k) \mid \vec{q} \in \Delta\mathrm{BZ}(G),\ k \in A(\vec{q})\}$,	set of standard G-irrep labels.	

The meaning of the symbol $\Delta^{\vec{q}}$ is defined as

$$\Delta^{\vec{q}}(\underline{R},R\underline{S}) = \begin{cases} 1, & \text{if } \underline{R}^{-1}R\underline{S} \in P(\vec{q}),\\ 0, & \text{otherwise.}\end{cases} \tag{3.7}$$

This uniquely determines the permutational structure of standard space group irreps of $G = P23$ if the coset representatives (CR's) are fixed. Note that

$$G = \bigcup_{\underline{R} \in \underline{P}(\vec{q})} (\underline{R}|\vec{0})G(\vec{q}) \text{ and } P = \bigcup_{\underline{R} \in \underline{P}(\vec{q})} \underline{R}P(\vec{q}). \tag{3.8}$$

Finally we choose for the many-dimensional $P(\vec{q})$-irreps D^k, $k \in A(\vec{q})$, the corresponding Miller–Love matrices (ML matrices, Ref. 6).

Clearly when applying the auxiliary group approach we have to modify correspondingly what we have called up to now standard G-irreps. This comes from the definition of standard G-irreps within the framework of the auxiliary group approach. Therefore, to be consistent, we define only the irreps D^K of Q-class representatives K along the lines given above. The irreps of the remaining members of a Q-class $[K]$ are, on the other hand, determined by choosing the unitary matrices $U^{K',K}(q_{K'}^{(K)})$, $q_{K'}^{(K)} \in R^K$ to be unit matrices. This defines a specific modification of the usual standard space group irreps.

The first task in our approach is to determine the auxiliary group Q for $G = P23$. In this case it only requires to inspect the character tables for $P23$ (Ref. 6) to obtain the group ASS. The automorphism groups of space groups are well known in the literature,[9]

$$\mathrm{AUT}(P23) = \mathrm{AUT} = Im3m, \tag{3.9}$$

$$\mathrm{ASS}(P23) = \mathrm{ASS} = \{G1,\ G2,\ G3,\ R1,\ R2,\ R3\}. \tag{3.10}$$

It should be pointed out that $\mathrm{AUT}(P23) = Im3m$ must have the same lattice constant as $P23$. Moreover note that the $P23$-irrep labels $K = (\vec{q},k)$ for the special points $\vec{G}$ (gamma), $\vec{R}$ (R point), etc. are abbreviated by Gk, $k \in A(\vec{G})$, Rk, $k \in A(\vec{R})$, etc.

Note that the group AUT contains $P23$ as a normal subgroup. However, as inner automorphisms do not affect $P23$-irrep labels, we carry out a coset decomposition of $Im3m$ with respect to $P23$. One readily finds

$$\begin{aligned}\mathrm{AUT} = {}&(E|\vec{0})G \cup (C_{2b}|\vec{0})G \cup (I|\vec{0})G \cup (\sigma_{db}|\vec{0})G\\ &\cup (E|\vec{B})G \cup (C_{2b}|\vec{B})G \cup (I|\vec{B})G \cup (\sigma_{db}|\vec{B})G.\end{aligned} \tag{3.11}$$

Note, in particular, that the special translation

$$\vec{B} = \tfrac{1}{2}(\vec{t}_1 + \vec{t}_2 + \vec{t}_3) \tag{3.12}$$

does not belong to the translation group T of $G = P23$. This comes from the fact that the translation group of $Im3m$ contains the translation group T of $P23$ as a subgroup of index two and $\vec{B}$ can be chosen as the nontrivial CR. The factor group AUT/G is an Abelian group of order 8 and shows that only seven nontrivial outer automorphisms (modulus G) exist. For the set of CR's defining (3.11) we introduce a set of generators that create all outer automorphisms (modulus G). We choose

$$b_1 = (C_{2b}|\vec{0}), \quad b_2 = (I|\vec{0}), \quad b_3 = (E|\vec{B}) \tag{3.13}$$

and abbreviate them by

$$b_j = (S_j|\vec{B}_j), \quad j = 1,2,3, \tag{3.14}$$

where the entries S_j and $\vec{B}_j$ have to be identified correspondingly. According to our approach we assign to every b_j a mapping β_j of G onto G, where

$$\beta_j((R|\vec{t})) = (S_j|\vec{B}_j)(R|\vec{t})(S_j|\vec{B}_j)^{-1}. \tag{3.15}$$

These three mappings together with those that are associated with inner automorphisms are sufficient to assign to every element of AUT the corresponding mapping of G onto G. Conveniently only the automorphisms (3.13) are needed to determine the Q-classes of $A(G)$. According to our procedure we have

$$Q = \mathrm{ASS} \circledS (\mathrm{AUT} \times \mathrm{CON})$$
$$\simeq (C_3 \times C_2) \circledS (Im3m \times C_2), \tag{3.16}$$

which is a discrete group. Later we shall show that we cannot ignore the inner automorphisms as they arise automatically when applying our approach.

For the following we limit our considerations to a subset of $A(G)$. This set, denoted by SA(G), contains all $P23$-irrep labels of the special points $\vec{G}, \vec{R}, \vec{X}, \vec{M} \in \Delta BZ(P23)$:

$$\mathrm{SA}(G) = \{G1, G2, G3, G4, G5, G6, G7; \\ R1, R2, R3, R4, R5, R6, R7; \\ X1, X2, X3, X4, X5; \\ M1, M2, M3, M4, M5\}. \tag{3.17}$$

The choice of SA(G) is influenced by the auxiliary group approach because this set is invariant with respect to all operations of Q. In other words the set SA(G) decomposes into disjoint Q-classes without leading to other $P23$-irrep labels which are not contained in SA(G). Moreover let us recall[6]

$$P(\vec{G}) = P(\vec{R}) = P,$$
$$P(\vec{X}) = P(\vec{M}) = \{E, C_{2x}, C_{2y}, C_{2z}, \overline{E}, \overline{C}_{2x}, \overline{C}_{2y}, \overline{C}_{2z}\}. \tag{3.18}$$

Finally to fix the standard $P23$-irreps for the Q-class representatives K we have to fix CR-sets for $\vec{G}, \vec{R}, \vec{X}, \vec{M}$. We choose

$$\underline{P}(\vec{G}) = \underline{P}(\vec{R}) = \{E\},$$
$$\underline{P}(\vec{X}) = \underline{P}(\vec{M}) = \{E, C_{31}^-, C_{31}^+\}, \tag{3.19}$$

and take for the $P(\vec{q})$-irreps D^k, $k \in A(\vec{q})$ the corresponding ML matrices.[6]

To determine the Q-classes one has to inspect

$$\mathrm{Tr}((a_J D^K)(R|\vec{t})) = \chi^J(R|\vec{t})\chi^K(R|\vec{t}), \tag{3.20}$$
$$\mathrm{Tr}((b_l D^K)(R|\vec{t})) = \chi^K(\beta_l^{-1}(R|\vec{t})), \tag{3.21}$$
$$\mathrm{Tr}((cD^K)(R|\vec{t})) = \chi^K(R|\vec{t})^*, \tag{3.22}$$

where $a_J \in$ ASS and J abbreviates the irrep labels of the one-dimensional irreps of $P23$.

The mappings $a_J: K \to K'$ defined by (3.20) are readily obtained from the KP Tables of Ref. 10. We arrive at the following qK tables:

a_J Gk	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$
$G1$	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$
$G2$	$G2$	$G3$	$G1$	$G4$	$G6$	$G7$	$G5$
$G3$	$G3$	$G1$	$G2$	$G4$	$G7$	$G5$	$G6$,
$R1$	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
$R2$	$R2$	$R3$	$R1$	$R4$	$R6$	$R7$	$R5$
$R3$	$R3$	$R1$	$R2$	$R4$	$R7$	$R5$	$R6$

a_J Rk	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
$G1$	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
$G2$	$R2$	$R3$	$R1$	$R4$	$R6$	$R7$	$R5$
$G3$	$R3$	$R1$	$R2$	$R4$	$R7$	$R5$	$R6$,
$R1$	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$
$R2$	$G2$	$G3$	$G1$	$G4$	$G6$	$G7$	$G5$
$R3$	$G3$	$G1$	$G2$	$G4$	$G7$	$G5$	$G6$

(3.23)

a_J Xk	$X1$	$X2$	$X3$	$X4$	$X5$
$G1$	$X1$	$X2$	$X3$	$X4$	$X5$
$G2$	$X1$	$X2$	$X3$	$X4$	$X5$
$G3$	$X1$	$X2$	$X3$	$X4$	$X5$,
$R1$	$M1$	$M3$	$M4$	$M2$	$M5$
$R2$	$M1$	$M3$	$M4$	$M2$	$M5$
$R3$	$M1$	$M3$	$M4$	$M2$	$M5$

a_J Mk	$M1$	$M2$	$M3$	$M4$	$M5$
$G1$	$M1$	$M2$	$M3$	$M4$	$M5$
$G2$	$M1$	$M2$	$M3$	$M4$	$M5$
$G3$	$M1$	$M2$	$M3$	$M4$	$M5$.
$R1$	$X1$	$X4$	$X2$	$X3$	$X5$
$R2$	$X1$	$X4$	$X2$	$X3$	$X5$
$R3$	$X1$	$X4$	$X2$	$X3$	$X5$

The mappings $b_l: K \to K'$ defined by (3.21) require a careful analysis especially because $P23$ is assumed to be a double space group,

b_l Gk	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$
$(C_{2b}\|\vec{0})$	$G1$	$G3$	$G2$	$G4$	$G5$	$G7$	$G6$
$(I\|\vec{0})$	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$,
$(E\|\vec{B})$	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$

b_l Rk	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
$(C_{2b}\|\vec{0})$	$R1$	$R3$	$R2$	$R4$	$R5$	$R7$	$R6$
$(I\|\vec{0})$	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$,
$(E\|\vec{B})$	$R1$	$R3$	$R2$	$R4$	$R5$	$R7$	$R6$

b_l Xk	$X1$	$X2$	$X3$	$X4$	$X5$
$(C_{2b}\|\vec{0})$	$X1$	$X3$	$X2$	$X4$	$X5$
$(I\|\vec{0})$	$X1$	$X2$	$X3$	$X4$	$X5$,
$(E\|\vec{B})$	$X4$	$X3$	$X2$	$X1$	$X5$

(3.24)

b_l Mk	$M1$	$M2$	$M3$	$M4$	$M5$
$(C_{2b}\|\vec{0})$	$M1$	$M2$	$M4$	$M3$	$M5$
$(I\|\vec{0})$	$M1$	$M2$	$M3$	$M4$	$M5$.
$(E\|\vec{B})$	$M2$	$M1$	$M4$	$M3$	$M5$

Finally the mappings $c: K \rightarrow K'$ are readily obtained from (3.22) by comparing the complex conjugate characters with the character table of $P23$.

c Gk	$G1$	$G2$	$G3$	$G4$	$G5$	$G6$	$G7$
*	$G1$	$G3$	$G2$	$G4$	$G5$	$G7$	$G6$,

c Rk	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$
*	$R1$	$R3$	$R2$	$R4$	$R5$	$R7$	$R6$,

(3.25)

c Xk	$X1$	$X2$	$X3$	$X4$	$X5$
*	$X1$	$X2$	$X3$	$X4$	$X5$,

c Mk	$M1$	$M2$	$M3$	$M4$	$M5$
*	$M1$	$M2$	$M3$	$M4$	$M5$.

Inspecting the tables for the various mappings one easily deduces the following Q-classes $[K] \subset \mathrm{SA}(G)$:

$$\begin{aligned}
&[G1] = \{G1,\ G2,\ G3,\ R1,\ R2,\ R3\},\\
&[G4] = \{G4,\ R4\},\\
&[G5] = \{G5,\ G6,\ G7,\ R5,\ R6,\ R7\},\\
&[X1] = \{X1,\ X4,\ M1,\ M2\},\\
&[X2] = \{X2,\ X3,\ M3,\ M4\},\\
&[X5] = \{X5,\ M5\}.
\end{aligned} \tag{3.26}$$

Accordingly the set $\mathrm{SA}(G)$ consisting of 24 elements decomposes into six Q-classes, where $G1, G4, G5, X1, X2, X5$ are chosen as representatives. The next task is to determine the groups Q^K and to fix the corresponding sets R^K of CR's. Taking into account (3.9), (3.10), and (3.16) one readily finds from (3.26),

$$\begin{aligned}
&Q^{G1} = \mathrm{AUT} \times \mathrm{CON},\\
&Q^{G4} = \{G1,\ G2,\ G3\} \circledS (\mathrm{AUT} \times \mathrm{CON}),\\
&Q^{G5} = \mathrm{AUT} \times \mathrm{CON},\\
&Q^{X1} = \{G1,\ G2,\ G3\} \circledS (Pm3m \times \mathrm{CON}),\\
&Q^{X2} = \{G1,\ G2,\ G3\} \circledS (Im3 \times \mathrm{CON}),\\
&Q^{X5} = \{G1,\ G2,\ G3\} \circledS (\mathrm{AUT} \times \mathrm{CON}).
\end{aligned} \tag{3.27}$$

To derive the fourth and fifth equations of (3.27) two different coset decompositions of $\mathrm{AUT} = Im3m$ have to be utilized, namely,

$$\begin{aligned}
&Im3m = Pm3m \cup (E\,|\vec{B})Pm3m,\\
&Im3m = Im3 \cup (C_{2b}\,|\vec{0})Im3.
\end{aligned} \tag{3.28}$$

These decompositions can be easily deduced from Ref. 11. From (3.16), (3.27), and (3.28) we infer that

$$\begin{aligned}
&R^{G1} = R^{G5} = \mathrm{ASS},\\
&R^{G4} = R^{X5} = \{G1,\ R1\},\\
&R^{X1} = \{G1,\ R1\} \times \{(E\,|\vec{0}),\ (E\,|\vec{B})\},\\
&R^{X2} = \{G1,\ R1\} \times \{(E\,|\vec{0}),\ (C_{2b}\,|\vec{0})\},
\end{aligned} \tag{3.29}$$

can be chosen as CR's R^K. For the sake of clearness it is useful to list the various sets R^K in more detail. Due to the definition of the auxiliary group Q the CR's are ordered triplets.[1] In detail we have

$$\begin{aligned}
R^{G1} &= \{(a_0,b_0,c_0),\ (G2,b_0,c_0),\ (G3,b_0,c_0),\\
&\qquad (R1,b_0,c_0),\ (R2,b_0,c_0),\ (R3,b_0,c_0)\},\\
R^{G4} &= \{(a_0,b_0,c_0),\ (R1,b_0,c_0)\},\\
R^{X1} &= \{(a_0,b_0,c_0),\ (R1,b_0,c_0),\\
&\qquad (a_0,b_3,c_0),\ (R1,b_3,c_0)\},\\
R^{X2} &= \{(a_0,b_0,c_0),\ (R1,b_0,c_0),\\
&\qquad (a_0,b_1,c_0),\ (R1,b_1,c_0)\},
\end{aligned} \tag{3.30}$$

where the symbols $a_0 = G1$, $b_0 = (E\,|\vec{0})$, and c_0 are the trivial operations of the groups ASS, AUT, and CON, respectively. Very often, however, we omit the trivial constitutents of the triplets (a, b, c) in order to keep the notation as concise as possible.

At this point we recall briefly how we modify the usual standard $P23$ irreps. First we have to fix the $P23$ irreps for the Q-class representatives $G1, G4, G5, X1, X2, X5$. Then the remaining $P23$ irreps within a Q class are obtained by using

$$D^{K'}(R\,|\vec{t}) = (q_{K'}^{(K)} D^K)(R\,|\vec{t}), \tag{3.31}$$

where $q_{K'}^{(K)} \in R^K$. Note that our standard $P23$ irreps may differ from commonly used standard $P23$ irreps by nontrivial similarity transformations.

The next task is to compute the matrices $U^K(q)$, $q \in Q^K$ for the various Q-class representatives. For this purpose we have to use definition (2.6). Before summarizing our results it is useful to explain how to treat inner automorphisms. Let us start from our basic equation (2.6) where $q = b$ is an arbitrary automorphism belonging to Q^K,

$$\begin{aligned}
(bD^K)(R\,|\vec{t}) &= D^K(\beta^{-1}(R\,|\vec{t}))\\
&= U^K(b)^\dagger D^K(R\,|\vec{t}) D^K(b).
\end{aligned} \tag{3.32}$$

If $b = b'$ is an inner automorphism then we can identify b' with a specific group element $(R'|\vec{t}') \in P23$. Hence

$$D^K(\beta'^{-1}(R\,|\vec{t})) = D^K((R'|\vec{t}')^{-1}(R\,|\vec{t})(R'|\vec{t}'))$$

and therefore we may identify

$$U^K(b') = D^K(R'|\vec{t}'). \tag{3.33}$$

This shows how simple inner automorphisms can be treated. Although they are trivial one must not forget that they can occur because of (2.7) and (2.8). Moreover we only need to know the matrices $U^K(b_j)$ for the generating automorphisms b_j, $j = 1, 2, 3$ (provided that they belong to Q^K). Straightforward computations yield for the various Q-class representatives K the following similarity transformations:

$$K = G1: \; U^K(q) = 1, \quad q \in Q^{G1}. \tag{3.34}$$

$$K = G4: \; U^K(G2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \alpha^* & 0 \\ 0 & 0 & \alpha \end{vmatrix},$$

$$\alpha = \exp(i2\pi/3),$$

$$U^K(C_{2b}|\vec{0}) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}, \tag{3.35}$$

$$U^K(I|\vec{0}) = U^K(E|\vec{B}) = U^K(c) = E(3).$$

$$K = G5: \; U^K(C_{2b}|\vec{0}) = \begin{vmatrix} 0 & \gamma^* \\ \gamma & 0 \end{vmatrix}, \quad \gamma = \exp(i\pi/4),$$

$$U^K(I|\vec{0}) = U^K(E|\vec{B}) = E(2), \tag{3.36}$$

$$U^K(c) = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}.$$

$$K = X1: \; U^K(G2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^* \end{vmatrix},$$

$$\alpha = \exp(i2\pi/3),$$

$$U^K(C_{2b}|\vec{0}) = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{3.37}$$

$$U^K(I|\vec{0}) = U^K(c) = E(3).$$

$$K = X2: \; U^K(G2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^* \end{vmatrix},$$

$$\alpha = \exp(i2\pi/3),$$

$$U^K(C_{2b}|\vec{B}) = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{3.38}$$

$$U^K(I|\vec{0}) = U^K(c) = E(3).$$

$$K = X5: \; U^K(G2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^* \end{vmatrix} \otimes E(2),$$

$$\alpha = \exp(i2\pi/3),$$

$$U^K(C_{2b}|\vec{0}) = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} \otimes \frac{1}{\sqrt{2}} \begin{vmatrix} 1 & -i \\ i & -1 \end{vmatrix}, \tag{3.39}$$

$$U^K(E|\vec{B}) = E(3) \otimes \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix},$$

$$U^K(c) = E(3) \otimes \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}.$$

These similarity transformations are unique up to arbitrary phase factors which is in agreement with the general theory. We have taken the "simplest" choices. One should bear in mind that whenever the matrices $U^K(q)$ are involved they have to satisfy the comultiplication law.[1]

As already pointed out at the beginning we consider twofold KP's of $P23$-irreps. We discuss the set of KP's which is assigned to the elements (= ordered pairs) of the product set

$$\begin{aligned} \mathrm{SA}(G \times G) &= \mathrm{SA}(G) \times \mathrm{SA}(G) \\ &= \{((\vec{q}_1, k_1), (\vec{q}_1, k_2)) | \\ &\quad \times \vec{q}_j \in \{\vec{G}, \vec{R}, \vec{X}, \vec{M}\}, \; k_j \in A(\vec{q}_j)\}. \end{aligned} \tag{3.40}$$

Clearly this set is a subset of the set $A(G \times G) = A(G) \times A(G)$ of all ordered pairs of $P23$-irreps labels. The order of $\mathrm{SA}(G \times G)$ is given by

$$|\mathrm{SA}(G \times G)| = 24^2 = 576. \tag{3.41}$$

Now the crucial question is which simplifications and reductions in the calculation of the CG matrices for the set $\mathrm{SA}(G \times G)$ can be achieved by means of the auxiliary group approach. Or, in other words, how many CG matrices or parts of them actually have to be computed when generating relations of the first and second kind are exploited.

To answer these questions we first have to define the auxiliary group **Q**. Due to Sec. II we have

$$\mathbf{Q} \simeq (\mathrm{ASS} \times \mathrm{ASS}) \circledS (\mathrm{AUT} \times \mathrm{CON} \times \mathrm{PERM}). \tag{3.42}$$

Because of $\mathrm{ASS} = C_3 \times C_2$ we can write

$$\mathbf{Q} \cong (C_3 \times C_2)^{(2)} \circledS (Im3m \times C_2 \times S_2), \tag{3.43}$$

where the "diagonal" product groups AUT[×]AUT and CON[×]CON are simply written as AUT and CON, respectively.

The next task is to subdivide the set $\mathrm{SA}(G \times G)$ into disjoint **Q**-classes [**H**]. The corresponding manipulations are straightforward. One has to take into account Eqs. (3.23)–(3.25) and the group **Q** given by (3.42). However one must not forget that the groups **Q** and the direct product group $Q^{(2)} = Q \times Q$ are different and are not in a group–subgroup relation. This implies that in general it is not sufficient to take only the Q-class representatives to form ordered pairs in order to get a complete set of **Q**-classes [**H**]. This comes from the fact that AUT and CON are Kronecker product groups and therefore cannot act individually on the constituents of the KP's.

We summarize the **Q**-classes in the following equation. Here the sets $[K,K']$ listed at the intersections of two Q-classes (column $[K_1]$, row $[K_2]$) contain all KP's $K \otimes K'$, $K \in [K_1]$, $K' \in [K_2]$. Note that a full **Q**-class is always of the form $[K,K'] \cup [K',K]$, the second class being omitted here.

[G1]	[G4]	[G5]	[X1]	[X2]	[X5]	
[G1,G1]	[G1,G4]	[G1,G5]	[G1,X1]	[G1,X2]	[G1,X5]	[G1]
	[G4,G4]	[G4,G5]	[G4,X1]	[G4,X2]	[G4,X5]	[G4]
		[G5,G5]	[G5,X1]	[G5,X2]	[G5,X5]	[G5]
			[X1,X1] [X1,X4]	[X1,X2]	[X1,X5]	[X1]
				[X2,X2] [X2,X3]	[X2,X5]	[X2]
					[X5,X5]	[X5]

. (3.44)

Let us introduce

$$[K_1]\times[K_2] = [\vec{A}_1,k_1]\times[\vec{A}_2,k_2] \tag{3.45}$$

as the notation for product sets of ordered pairs where K_j are the Q_j-class representatives. Using this definition we conclude from (3.44) that except for two cases the sets $[K_1]\times[K_2]$ coincide with the **Q**-classes $[K_1,K_2]$. The two exceptions are

$$\begin{aligned}[X1]\times[X1] &= [X1,X1]\cup[X1,X4],\\ [X2]\times[X2] &= [X2,X2]\cup[X2,X3].\end{aligned} \tag{3.46}$$

What is the reason for this result? We may infer that if at least one CR set R^{K_j} consists only of elements of ASS then $[K_1]\times[K_2]\cup[K_2]\times[K_1] = [K_1,K_2]$ holds. This relation is also valid if elements of AUT or CON occur in R^{K_j}, provided that $R^{K_1}\cap R^{K_2} = \{q_0\}$. Therefore only in the cases given in (3.46) do we arrive at a splitting of the sets $[K_1]\times[K_2]\cup[K_2]\times[K_1]$ into disjoint subsets because there the same elements of AUT occur in both sets R^{K_1} and R^{K_2}.

We know from the preceding discussion that the sets R^K are sets of ordered triplets which also can be written as product sets [see (3.29)]. But in the following we keep the notation as concise as possible and write the sets $\mathbf{R}^{\mathbf{H}}$ (if possible) as product sets omitting the trivial group elements of the various subgroups of Q. For convenience we denote the trivial subgroup of ASS by C_1 and recall that $C_3=\{G1, G2, G3\}$ and $C_2=\{G1, R1\}$ [cf. (3.9)]. In the following we summarize the groups $\mathbf{Q}^{\mathbf{H}}$ together with their CR sets $\mathbf{R}^{\mathbf{H}}$:

$$\begin{aligned}\mathbf{Q}^{G1,G1} &= Im3m\times C_2\times S_2,\\ \mathbf{R}^{G1,G1} &= \mathrm{ASS}\times\mathrm{ASS},\end{aligned} \tag{3.47}$$

$$\begin{aligned}\mathbf{Q}^{G1,G4} &= (C_1\times C_3)\circledS(Im3m\times C_2),\\ \mathbf{R}^{G1,G4} &= (\mathrm{ASS}\times C_2)\times S_2,\end{aligned} \tag{3.48}$$

$$\begin{aligned}\mathbf{Q}^{G1,G5} &= Im3m\times C_2,\\ \mathbf{R}^{G1,G5} &= (\mathrm{ASS}\times\mathrm{ASS})\times S_2,\end{aligned} \tag{3.49}$$

$$\begin{aligned}\mathbf{Q}^{G1,X1} &= (C_1\times C_3)\circledS(Pm3m\times C_2),\\ \mathbf{R}^{G1,X1} &= (\mathrm{ASS}\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B})\}\times S_2,\end{aligned} \tag{3.50}$$

$$\begin{aligned}\mathbf{Q}^{G1,X2} &= (C_1\times C_3)\circledS(Im3\times C_2),\\ \mathbf{R}^{G1,X2} &= (\mathrm{ASS}\times C_2)\times\{(E\,|\vec{0}),(C_{2b}|\vec{0})\}\times S_2,\end{aligned} \tag{3.51}$$

$$\begin{aligned}\mathbf{Q}^{G1,X5} &= (C_1\times C_3)\circledS(Im3m\times C_2),\\ \mathbf{R}^{G1,X5} &= (\mathrm{ASS}\times C_2)\times S_2,\end{aligned} \tag{3.52}$$

$$\begin{aligned}\mathbf{Q}^{G4,G4} &= (C_3\times C_3)\circledS(Im3m\times C_2\times S_2),\\ \mathbf{R}^{G4,G4} &= C_2\times C_2,\end{aligned} \tag{3.53}$$

$$\begin{aligned}\mathbf{Q}^{G4,G5} &= (C_3\times C_1)\circledS(Im3m\times C_2),\\ \mathbf{R}^{G4,G5} &= (C_2\times\mathrm{ASS})\times S_2,\end{aligned} \tag{3.54}$$

$$\begin{aligned}\mathbf{Q}^{G4,X1} &= (C_3\times C_3)\circledS(Pm3m\times C_2),\\ \mathbf{R}^{G4,X1} &= (C_2\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B})\}\times S_2,\end{aligned} \tag{3.55}$$

$$\begin{aligned}\mathbf{Q}^{G4,X2} &= (C_3\times C_3)\circledS(Im3\times C_2),\\ \mathbf{R}^{G4,X2} &= (C_2\times C_2)\times\{(E\,|\vec{0}),(C_{2b}|\vec{0})\}\times S_2,\end{aligned} \tag{3.56}$$

$$\begin{aligned}\mathbf{Q}^{G4,X5} &= (C_3\times C_3)\circledS(Im3m\times C_2),\\ \mathbf{R}^{G4,X5} &= (C_2\times C_2)\times S_2,\end{aligned} \tag{3.57}$$

$$\begin{aligned}\mathbf{Q}^{G5,G5} &= Im3m\times C_2\times S_2,\\ \mathbf{R}^{G5,G5} &= \mathrm{ASS}\times\mathrm{ASS},\end{aligned} \tag{3.58}$$

$$\begin{aligned}\mathbf{Q}^{G5,X1} &= (C_1\times C_3)\circledS(Pm3m\times C_2),\\ \mathbf{R}^{G5,X1} &= (\mathrm{ASS}\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B})\}\times S_2,\end{aligned} \tag{3.59}$$

$$\begin{aligned}\mathbf{Q}^{G5,X2} &= (C_1\times C_3)\circledS(Im3\times C_2),\\ \mathbf{R}^{G5,X2} &= (\mathrm{ASS}\times C_2)\times\{(E\,|\vec{0}),(C_{2b}|\vec{0})\}\times S_2,\end{aligned} \tag{3.60}$$

$$\begin{aligned}\mathbf{Q}^{G5,X5} &= (C_1\times C_3)\circledS(Im3m\times C_2),\\ \mathbf{R}^{G5,X5} &= (\mathrm{ASS}\times C_2)\times S_2.\end{aligned} \tag{3.61}$$

At this point we arrive at the two exceptions which require a more careful analysis. To determine $\mathbf{Q}^{X1,X1}$ and $\mathbf{R}^{X1,X1}$ one can proceed in the same way as before,

$$\begin{aligned}\mathbf{Q}^{X1,X1} &= (C_3\times C_3)\circledS(Pm3m\times C_2\times S_2),\\ \mathbf{R}^{X1,X1} &= (C_2\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B})\}.\end{aligned} \tag{3.62}$$

It is less trivial to find $\mathbf{Q}^{X1,X4}$ and $\mathbf{R}^{X1,X4}$ because $X4$ is a member of the Q-class $[X1]$ but does not coincide with the Q-class representative $X1$. Therefore we have to establish a relationship between Q^{X1} and Q^{X4}. We expect them to be conjugate subgroups of Q. The proof can be given quite generally. Assume that k and l belong to the Q-class $[k]$, then there must exist a $q=q_l^{(k)}\in Q$ such that $k=q_l^{(k)}l$. Therefore if $q\in Q^k$ then $q_l^{(k)}q(q_l^{(k)})^{-1}\in Q^l$ or, in other words,

$$Q^l = q_l^{(k)}Q^k(q_l^{(k)})^{-1}. \tag{3.63}$$

This means for our special case

$$Q^{X4} = (E\,|\vec{B})Q^{X1}(E\,|\vec{B})^{-1} = Q^{X1}. \tag{3.64}$$

To prove the identity $Q^{X4}=Q^{X1}$ one needs the relation $(E\,|\vec{B})Pm3m(E\,|\vec{B})^{-1}=Pm3m$ and the fact that C_3 is a normal subgroup of ASS. However, we must not infer from

(3.64) that $\mathbf{Q}^{X1,X1}=\mathbf{Q}^{X1,X4}$ because $\mathbf{Q}^{X1,X1}$ contains S_2 whose nontrivial element p_{12} cannot belong to $\mathbf{Q}^{X1,X4}$. Hence the group $\mathbf{Q}^{X1,X4}$ has a more complicated structure. In fact the combined group element

$$\mathbf{q}'=((E\,|\vec{B}),p_{12}) \tag{3.65}$$

belongs to $\mathbf{Q}^{X1,X4}$ whereas neither $(E\,|\vec{B})$ nor p_{12} belongs to this group. Therefore we arrive at

$$\begin{aligned}&\mathbf{Q}^{X1,X4}=((C_3\times C_3)\circledS(Pm3m\times C_2))\circledS\{\mathbf{q}_0,\mathbf{q}'\},\\&\mathbf{R}^{X1,X4}=\mathbf{R}^{X1,X1},\end{aligned} \tag{3.66}$$

where the choice $\mathbf{R}^{X1,X4}=\mathbf{R}^{X1,X1}$ has been made for convenience. However, note that

$$(E\,|\vec{B})D^{X1,X4}=p_{12}D^{X1,X4}=D^{X4,X1} \tag{3.67}$$

holds, i.e., both group elements are equally well suited to serve as CR's. To determine $\mathbf{Q}^{X1,X2}$ and $\mathbf{R}^{X1,X2}$ we need some further subgroup relations for the space groups $Pm3m$ and $Im3$. We deduce from Ref. 11 that

$$\begin{aligned}&Pm3m=Pm3\cup(C_{2b}\,|\vec{0})Pm3,\\&Im3=Pm3\cup(E\,|\vec{B})Pm3,\end{aligned} \tag{3.68}$$

holds which leads us to $Pm3m\cap Im3=Pm3$. Therefore we have

$$\begin{aligned}&\mathbf{Q}^{X1,X2}=(C_3\times C_3)\circledS(Pm3\times C_2),\\&\mathbf{R}^{X1,X2}=(C_2\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B}),\\&\qquad(C_{2b}\,|\vec{0}),(C_{2b}\,|\vec{B})\}\times S_2.\end{aligned} \tag{3.69}$$

The determination of $\mathbf{Q}^{X1,X5}$ and $\mathbf{R}^{X1,X5}$ is simple and we obtain

$$\begin{aligned}&\mathbf{Q}^{X1,X5}=(C_3\times C_3)\circledS(Pm3\times C_2),\\&\mathbf{R}^{X1,X5}=(C_2\times C_2)\times\{(E\,|\vec{0}),(E\,|\vec{B})\}\times S_2.\end{aligned} \tag{3.70}$$

To find $\mathbf{Q}^{X2,X2}$ and $\mathbf{R}^{X2,X2}$ is again straightforward,

$$\begin{aligned}&\mathbf{Q}^{X2,X2}=(C_3\times C_3)\circledS(Im3\times C_2\times S_2),\\&\mathbf{R}^{X2,X2}=(C_2\times C_2)\times\{(E\,|\vec{0}),(C_{2b}\,|\vec{0})\}.\end{aligned} \tag{3.71}$$

However, the decomposition of $\mathbf{Q}$ with respect to $\mathbf{Q}^{X2,X3}$ requires a more careful analysis analogous to (3.66). We need

$$Q^{X3}=(C_{2b}\,|\vec{0})Q^{X2}(C_{2b}\,|\vec{0})^{-1}=Q^{X2}, \tag{3.72}$$

where the coincidence of Q^{X3} with Q^{X2} comes from the fact that $Im3$ and C_3 are normal subgroups of $Im3m$ and ASS, respectively. Again we must not infer from (3.72) that $\mathbf{Q}^{X2,X2}=\mathbf{Q}^{X2,X3}$ is valid. One readily proves that the special combined group element

$$\hat{\mathbf{q}}=((C_{2b}\,|\vec{0}),p_{12}) \tag{3.73}$$

belongs to $\mathbf{Q}^{X2,X3}$. Thus we arrive at

$$\begin{aligned}&\mathbf{Q}^{X2,X3}=((C_3\times C_3)\circledS(Im3\times C_2))\circledS\{\mathbf{q}_0,\hat{\mathbf{q}}\},\\&\mathbf{R}^{X2,X3}=\mathbf{R}^{X2,X2}.\end{aligned} \tag{3.74}$$

Again note that the group elements $(C_{2b}\,|\vec{0})$ and p_{12} have the same effect on $D^{X2,X3}$, namely, to permute the constituents of the KP. The remaining two cases are simple and we obtain the following groups and CR's:

$$\mathbf{Q}^{X2,X5}=(C_3\times C_3)\circledS(Im3\times C_2),$$

$$\mathbf{R}^{X2,X5}=(C_2\times C_2)\times\{(E\,|\vec{0}),(C_{2b}\,|\vec{0})\}\times S_2, \tag{3.75}$$

$$\begin{aligned}&\mathbf{Q}^{X5,X5}=(C_3\times C_3)\circledS(Im3m\times C_2\times S_2),\\&\mathbf{R}^{X5,X5}=C_2\times C_2.\end{aligned} \tag{3.76}$$

Before we start with the first step in our auxiliary group approach let us check the relation

$$|\mathrm{SA}(G\times G)|=\sum_{\mathbf{H}}|[\mathbf{H}]|=\sum_{\mathbf{H}}|\mathbf{R}^{H}|. \tag{3.77}$$

From Eq. (3.44) we know that $\mathrm{SA}(P23\times P23)$ decomposes into 23 disjoint $\mathbf{Q}$ classes. Summing the orders of the various CR sets $\mathbf{R}^{\mathbf{H}}$ one readily proves (3.77).

The first step in our auxiliary group approach is to utilize generating relations of the first kind. This requires the determination of the KP decompositions for the $\mathbf{Q}$-class representatives $\mathbf{H}$. To begin we merely need to know the G-irreps occurring in the KP decompositions. For the sake of simplicity the KP decompositions are written symbolically. The formulas we state emerge from Ref. 10,

$$\begin{aligned}&G1\otimes G1=G1,\quad G1\otimes G4=G4,\\&G1\otimes G5=G5,\quad G1\otimes X1=X1,\\&G1\otimes X2=X2,\quad G1\otimes X5=X5,\end{aligned} \tag{3.78}$$

$$\begin{aligned}&G4\otimes G4=G1\oplus G2\oplus G3\oplus 2\,G4,\\&G4\otimes G5=G5\oplus G6\oplus G7,\\&G4\otimes X1=X2\oplus X3\oplus X4,\\&G4\otimes X2=X1\oplus X3\oplus X4,\quad G4\otimes X5=3X5,\end{aligned} \tag{3.79}$$

$$\begin{aligned}&G5\otimes G5=G1\oplus G4,\quad G5\otimes X1=X5,\\&G5\otimes X2=X5,\quad G5\otimes X5=X1\oplus X2\oplus X3\oplus X4,\end{aligned} \tag{3.80}$$

$$\begin{aligned}&X1\otimes X1=G1\oplus G2\oplus G3\oplus 2\,M1,\\&X1\otimes X4=G4\oplus M3\oplus M4,\\&X1\otimes X2=G4\oplus M2\oplus M4,\\&X1\otimes X5=G5\oplus G6\oplus G7\oplus 2\,M5,\end{aligned} \tag{3.81}$$

$$\begin{aligned}&X2\otimes X2=G1\oplus G2\oplus G3\oplus 2\,M3,\\&X2\otimes X3=G4\oplus M1\oplus M2,\\&X2\otimes X5=G5\oplus G6\oplus G7\oplus 2\,M5,\end{aligned} \tag{3.82}$$

$$\begin{aligned}X5\otimes X5=\;&G1\oplus G2\oplus G3\oplus 3\,G4\\&\oplus 2(M1\oplus M2\oplus M3\oplus M4).\end{aligned} \tag{3.83}$$

To utilize generating relations of the first kind it is not necessary to know explicitly the generating CG matrices $C^{\mathbf{H}}$. We merely need the CR's $\mathbf{q}_{\mathrm{L}}^{(\mathbf{H})}\in\mathbf{R}^{\mathbf{H}}$, their homomorphic images $q_{\mathrm{L}}^{(\mathbf{H})}$, and the corresponding similarity matrices $U^{\mathbf{L},\mathbf{H}}(\mathbf{q}_{\mathrm{L}}^{(\mathbf{H})})$ and $U^{L',L}(q_{\mathrm{L}}^{(\mathbf{H})})$, where $L'=q_{\mathrm{L}}^{(\mathbf{H})}L$. The latter matrices compose the transformations $Z^{\mathbf{H}}(q_{\mathrm{L}}^{(\mathbf{H})})$,

$$Z^{\mathbf{H}}(q_{\mathrm{L}}^{(\mathbf{H})})=\bigoplus_{L}E(\mathbf{H}|L)\otimes U^{L',L}(q_{\mathrm{L}}^{(\mathbf{H})}). \tag{3.84}$$

Specifying formula (2.61) to the present examples, the generating relations of the first kind read

$$C^{\mathbf{L}}=U^{\mathbf{L},\mathbf{H}}(\mathbf{q}_{\mathrm{L}}^{(\mathbf{H})})(q_{\mathrm{L}}^{(\mathbf{H})}C^{\mathbf{H}})Z^{\mathbf{H}}(q_{\mathrm{L}}^{(\mathbf{H})})^{\dagger}P(q_{\mathrm{L}}^{(\mathbf{H})}), \tag{3.85}$$

where $P(q_{\mathrm{L}}^{(\mathbf{H})})$ is, in general, a nontrivial permutation matrix which ensures that the irreps D^{L}, $L\in A(G)$, occur in the decomposition in the lexicographical sequence given by (3.17).

Clearly for a given KP $D^{\mathbf{H}}$, where $\mathbf{H}$ presents the $\mathbf{Q}$-class representative of $[\mathbf{H}]$, the following must hold:

number of generating relations of the first kind

$$\text{for } D^{\mathbf{H}} = |[\mathbf{H}]| = |\mathbf{R}^{\mathbf{H}}|. \tag{3.86}$$

To get more insight into the structure of generating relations of the first kind we discuss three examples in full detail. The first example concerns the KP $D^{G4,G4}$. Because of

$$(G1,G1)D^{G4,G4} = D^{G4,G4}, \quad (G1,R1)D^{G4,G4} = D^{G4,R4},$$
$$(R1,G1)D^{G4,G4} = D^{R4,G4}, \quad (R1,R1)D^{G4,G4} = D^{R4,R4}, \tag{3.87}$$

which are identities due to the standardization procedure of G irreps, we choose the following CR's:

$$\begin{aligned} \mathbf{q}^{(G4,G4)}_{G4,G4} &= (G1,G1), \quad \mathbf{q}^{(G4,G4)}_{G4,R4} = (G1,R1), \\ \mathbf{q}^{(G4,G4)}_{R4,G4} &= (R1,G1), \quad \mathbf{q}^{(G4,G4)}_{R4,R4} = (R1,R1). \end{aligned} \tag{3.88}$$

Their homomorphic images are

$$q^{(G4,G4)}_{G4,G4} = q^{(G4,G4)}_{R4,R4} = G1, \quad q^{(G4,G4)}_{G4,R4} = q^{(G4,G4)}_{R4,G4} = R1. \tag{3.89}$$

Because of the standardization procedure of the G irreps we have

$$\begin{aligned} &U^{\mathbf{L};G4,G4}(\mathbf{q}^{(G4,G4)}_{\mathbf{L}}) = E(3) \otimes E(3), \\ &Z^{G4,G4}(\mathbf{q}^{(G4,G4)}_{\mathbf{L}}) = E(6) \quad \text{for all } \mathbf{L} \in [G4,G4]. \end{aligned} \tag{3.90}$$

The second relation holds by virtue of $U^{L',L}(q^{(G4,G4)}_{\mathbf{L}}) = E(n_L)$ for all $\mathbf{L} \in [G4,G4]$ and all $L \in A(G)$ which are contained in the KP decomposition of $D^{G4,G4}$. Hence in this example we arrive at generating relations of the first kind of the form

$$C^{\mathbf{L}} = C^{G4,G4} \quad \text{for all } \mathbf{L} \in [G4,G4]. \tag{3.91}$$

Note, in particular, that the simplicity of (3.91) is achieved because of the standardization procedure for the G irreps. This example does not represent the most general situation where nontrivial similarity transformations may occur.

To investigate if nontrivial similarity transformations may also appear in generating relations of the first kind we consider as a second example the KP $D^{X1,X1}$. Here we have eight generating relations of the first kind because of (3.86) and (3.62). In detail we have

$$\begin{aligned} &(G1,G1,(E|\vec{0}))D^{X1,X1} = D^{X1,X1}, \\ &(G1,R1,(E|\vec{0}))D^{X1,X1} = D^{X1,M1}, \\ &(R1,G1,(E|\vec{0}))D^{X1,X1} = D^{M1,X1}, \\ &(R1,R1,(E|\vec{0}))D^{X1,X1} = D^{M1,M1}, \\ &(G1,G1,(E|\vec{B}))D^{X1,X1} = D^{X4,X4}, \\ &(G1,R1,(E|\vec{B}))D^{X1,X1} = D^{X4,M2}, \\ &(R1,G1,(E|\vec{B}))D^{X1,X1} = D^{M2,X4}, \\ &(R1,R1,(E|\vec{B}))D^{X1,X1} = D^{M2,M2}, \end{aligned} \tag{3.92}$$

which are again identities due to the standardization procedure for G irreps. Therefore we choose the following CR's:

$$\begin{aligned} \mathbf{q}^{(X1,X1)}_{X1,X1} &= (G1,G1,(E|\vec{0})), \\ \mathbf{q}^{(X1,X1)}_{X1,M1} &= (G1,R1,(E|\vec{0})), \\ \mathbf{q}^{(X1,X1)}_{M1,X1} &= (R1,G1,(E|\vec{0})), \\ \mathbf{q}^{(X1,X1)}_{M1,M1} &= (R1,R1,(E|\vec{0})), \\ \mathbf{q}^{(X1,X1)}_{X4,X4} &= (G1,G1,(E|\vec{B})), \\ \mathbf{q}^{(X1,X1)}_{X4,M2} &= (G1,R1,(E|\vec{B})), \\ \mathbf{q}^{(X1,X1)}_{M2,X4} &= (R1,G1,(E|\vec{B})), \\ \mathbf{q}^{(X1,X1)}_{M2,M2} &= (R1,R1,(E|\vec{B})). \end{aligned} \tag{3.93}$$

Their homomorphic images are the following:

$$\begin{aligned} q^{(X1,X1)}_{X1,X1} &= q^{(X1,X1)}_{M1,M1} = (G1,(E|\vec{0})), \\ q^{(X1,X1)}_{X1,M1} &= q^{(X1,X1)}_{M1,X1} = (R1,(E|\vec{0})), \\ q^{(X1,X1)}_{X4,X4} &= q^{(X1,X1)}_{M2,M2} = (G1,(E|\vec{\mathrm{B}})), \\ q^{(X1,X1)}_{X4,M2} &= q^{(X1,X1)}_{M2,X4} = (R1,(E|\vec{B})). \end{aligned} \tag{3.94}$$

Now let us discuss the structure of the matrices $U^{\mathbf{L};X1,X1}(\mathbf{q}^{(X1,X1)}_{\mathbf{L}})$ since we want to demonstrate that they are in fact unit matrices. For instance let $\mathbf{q}^{(X1,X1)}_{\mathbf{L}} = \mathbf{q}^{(X1,X1)}_{X4,M2}$. We have to write this transformation as an element of $Q^{(2)}$,

$$\mathbf{q}^{(X1,X1)}_{X4,M2} = ((G1,(E|\vec{B}));(R1,(E|\vec{B}))) \in Q^{(2)}. \tag{3.95}$$

From this we conclude that the constituents $(G1,(E|\vec{B}))$ and $(R1,(E|\vec{B}))$ are elements of R^{X1} and thus generate standardized $P23$ irreps. This holds for every $\mathbf{q}^{(X1,X1)}_{\mathbf{L}} \in \mathbf{R}^{X1,X1}$. Therefore

$$U^{\mathbf{L};X1,X1}(\mathbf{q}^{(X1,X1)}_{\mathbf{L}}) = E(3) \otimes E(3). \tag{3.96}$$

Next we have to find out the structure of the matrices $U^{L',L}(q^{(X1,X1)}_{\mathbf{L}})$ which are the constituents of $Z^{X1,X1}(q^{(X1,X1)}_{\mathbf{L}})$. For that purpose we have to investigate

$$\mathbf{q}^{(X1,X1)}_{\mathbf{L}} D^{X1,X1} \sim q^{(X1,X1)}_{\mathbf{L}} (D^{G1} \oplus D^{G2} \oplus D^{G3} \oplus 2D^{M1}). \tag{3.97}$$

Let $\mathbf{q}^{(X1,X1)}_{\mathbf{L}} = \mathbf{q}^{(X1,X1)}_{X1,M1} = (G1,R1,(E|\vec{0}))$ then $q^{(X1,X1)}_{X1,M1} = (R1,(E|\vec{0}))$. Therefore on the rhs of (3.97) we have

$$\begin{aligned} &(R1)D^{G1} = D^{R1}, \quad (R1)D^{G2} = D^{R2}, \\ &(R1)D^{G3} = D^{R3}, \\ &(R1)D^{M1} = (R1)^2 D^{X1} = D^{X1}, \end{aligned} \tag{3.98}$$

where we have employed (2.11) and (2.15). Consequently,

$$\begin{aligned} &U^{R1,G1}(q) = U^{R2,G2}(q) = U^{R3,G3}(q) = E(1), \\ &U^{M1,X1}(q) = E(3), \end{aligned} \tag{3.99}$$

where $q = q^{(X1,X1)}_{X1,M1}$. This leads to the generating relation of the first kind $C^{X1,M1} = C^{X1,X1}$. Now let $\mathbf{q}^{(X1,X1)}_{\mathbf{L}} = \mathbf{q}^{(X1,X1)}_{X4,X4} = (G1,G2,(E|\vec{\mathrm{B}}))$ then $q^{(X1,X1)}_{X4,X4} = (G1,(E|\vec{B}))$. Accordingly the rhs of (3.97) undergoes the following transformations:

$$\begin{aligned} &(G1,(E|\vec{B}))D^{G1} = D^{G1}, \\ &(G1,(E|\vec{B}))D^{G2} = D^{G2}, \\ &(G1,(E|\vec{B}))D^{G3} = D^{G3}, \\ &(G1,(E|\vec{B}))D^{M1} = (G1,(E|\vec{B}))(R1,(E|\vec{0}))D^{X1} \\ &\qquad = (R1,(E|\vec{B}))D^{X1} = D^{M2}. \end{aligned} \tag{3.100}$$

Because D^{G1} is one dimensional we can choose

$$U^{G1,G1}(q) = U^{G2,G2}(q) = U^{G3,G3}(q) = E(1), \quad U^{M2,M1}(q) = E(3), \tag{3.101}$$

where $q = q^{(X1,X1)}_{X4,X4}$. One can verify quite generally that all generating relations of the first kind take the form

$$C^{\mathrm{L}} = C^{X1,X1} \quad \text{for all } \mathrm{L} \in [X1,X1]. \tag{3.102}$$

Again we arrive at generating relations of the first kind which possess the simplest form.

To demonstrate that nontrivial matrices may also occur in generating relations of the first kind we discuss as the third example the KP $D^{X1,X4}$. Here because of (3.86) and (3.66), we also have eight generating relations of the first kind. However, in contrast to $D^{X1,X1}$, the constituent D^{X4} of the KP $D^{X1,X4}$ is not a Q-class representative of $[X1]$. In detail we have the equivalence relations

$$\begin{aligned}
&(G1,G1,(E|\vec{0}))D^{X1,X4} = D^{X1,X4},\\
&(G1,R1,(E|\vec{0}))D^{X1,X4} \simeq D^{X1,M2},\\
&(R1,G1,(E|\vec{0}))D^{X1,X4} \simeq D^{M1,X4},\\
&(R1,R1,(E|\vec{0}))D^{X1,X4} \simeq D^{M1,M2},\\
&(G1,G1,(E|\vec{B}))D^{X1,X4} \simeq D^{X4,X1},\\
&(G1,R1,(E|\vec{B}))D^{X1,X4} \simeq D^{X4,M1},\\
&(R1,G1,(E|\vec{B}))D^{X1,X4} \simeq D^{M2,X1},\\
&(R1,R1,(E|\vec{B}))D^{X1,X4} \simeq D^{M2,M1},
\end{aligned} \tag{3.103}$$

which are in general not identities. Nevertheless we can choose the transformations occurring in (3.103) as CR's,

$$\begin{aligned}
&\mathbf{q}^{(X1,X4)}_{X1,X4} = (G1,G1,(E|\vec{0})),\\
&\mathbf{q}^{(X1,X4)}_{X1,M2} = (G1,R1,(E|\vec{0})),\\
&\mathbf{q}^{(X1,X4)}_{M1,X4} = (R1,G1,(E|\vec{0})),\\
&\mathbf{q}^{(X1,X4)}_{M1,M2} = (R1,R1,(E|\vec{0})),\\
&\mathbf{q}^{(X1,X4)}_{X4,X1} = (G1,G1,(E|\vec{B})),\\
&\mathbf{q}^{(X1,X4)}_{X4,M1} = (G1,R1,(E|\vec{B})),\\
&\mathbf{q}^{(X1,X4)}_{M2,X1} = (R1,G1,(E|\vec{B})),\\
&\mathbf{q}^{(X1,X4)}_{M2,M1} = (R1,R1,(E|\vec{B})).
\end{aligned} \tag{3.104}$$

The homomorphic images of these transformations $\mathbf{q}$ are the following elements $q \in Q$:

$$\begin{aligned}
&q^{(X1,X4)}_{X1,X4} = q^{(X1,X4)}_{M1,M2} = (G1,(E|\vec{0})),\\
&q^{(X1,X4)}_{X1,M2} = q^{(X1,X4)}_{M1,M4} = (R1,(E|\vec{0})),\\
&q^{(X1,X4)}_{X4,X1} = q^{(X1,X4)}_{M2,M1} = (G1,(E|\vec{B})),\\
&q^{(X1,X4)}_{X4,M1} = q^{(X1,X4)}_{M2,X1} = (R1,(E|\vec{B})).
\end{aligned} \tag{3.105}$$

Again we have to find out the structure of the matrices $U^{\mathrm{L};X1,X4}(\mathbf{q}^{(X1,X4)}_{\mathrm{L}})$. Let us take the CR $\mathbf{q}^{(X1,X4)}_{X4,X1} = (G1,G1,(E|\vec{B}))$. This group element reads in more detail

$$\mathbf{q}^{(X1,X4)}_{X4,X1} = ((G1,(E|\vec{B}));(G1,(E|\vec{B}))) \in Q^{(2)}. \tag{3.106}$$

Because of $(G1,(E|\vec{B}))D^{X1} = D^{X4}$, which is in an irrep in standard form, we take

$$U^{X4,X1}(G1,(E|\vec{B})) = E(3). \tag{3.107}$$

However, we see

$$\begin{aligned}
(G1,(E|\vec{B}))D^{X4} &= (G1,(E|\vec{B}))(G1,(E|\vec{B}))D^{X1}\\
&= (G1,(E|2\vec{B}))D^{X1},
\end{aligned} \tag{3.108}$$

that a nontrivial element of Q^{X1} appears, namely $(G1,(E|2\vec{B}),c_0) \in Q^{X1}$. Since this is an inner automorphism we can use (3.33) and arrive at

$$U^{X1,X4}(G1,(E|\vec{B})) = D^{X1}(E|2\vec{B}), \tag{3.109}$$

which yields

$$U^{X4,X1;X1,X4}(\mathbf{q}^{(X1,X4)}_{X4,X1}) = E(3) \otimes D^{X1}(E|2\vec{B}). \tag{3.110}$$

Accordingly we may conclude that in generating relations of the first kind some nontrivial matrices $U^{\mathrm{L,H}}(\mathbf{q}^{(\mathrm{H})}_{\mathrm{L}})$ may occur. To see whether the matrices $U^{L',L}(q^{(\mathrm{H})}_{\mathrm{L}})$ are also nontrivial we have to investigate the mapping

$$\mathbf{q}^{(X1,X4)}_{\mathrm{L}} D^{X1,X4} \sim q^{(X1,X4)}_{\mathrm{L}} (D^{G4} \oplus D^{M3} \oplus D^{M4}). \tag{3.111}$$

Let $q^{(X1,X4)}_{\mathrm{L}} = q^{(X1,X4)}_{X4,X1} = (G1,(E|\vec{B}))$. This group element acts on the rhs constituents as follows:

$$\begin{aligned}
&(G1,(E|B))D^{G4} \simeq D^{G4},\\
&(G1,(E|\vec{B}))D^{M3} = (G1,(E|\vec{B}))(R1,(E|\vec{0}))D^{X2}\\
&\qquad = (R1,(E|\vec{B}))D^{X2}\\
&\qquad = (R1,(C_{2b}|\vec{0}))(G1,(\overline{C}_{2b}|\vec{B}))D^{X2}\\
&\qquad \simeq D^{M4},\\
&(G1,(E|\vec{B}))D^{M4} = (G1,(E|\vec{B}))(R1,(C_{2b}|\vec{0}))D^{X2}\\
&\qquad = (R1,(C_{2b}|\vec{B}))D^{X2}\\
&\qquad = (R1,(E|\vec{0}))(G1,(C_{2b}|\vec{B}))D^{X2} \simeq D^{M3}.
\end{aligned} \tag{3.112}$$

To prove the preceding relations one has to utilize the fact that $(G1,(E|\vec{B})) \in Q^{G4}$, $(G1,(\overline{C}_{2b}|\vec{B})) \in Q^{X2}$, and $(G1,(C_{2b}|\vec{B})) \in Q^{X2}$. Employing the general relation (2.18) we arrive at the following matrices:

$$\begin{aligned}
&U^{G4,G4}(G1,(E|\vec{B})) = E(3),\\
&U^{M4,M3}(G1,(E|\vec{B})) = U^{X2}(\overline{C}_{2b}|\vec{B}),\\
&U^{M3,M4}(G1,(E|\vec{B})) = U^{X2}(C_{2b}|\vec{B}).
\end{aligned} \tag{3.113}$$

Note that $U^{X2}(\overline{C}_{2b}|\vec{B}) = U^{X2}(C_{2b}|\vec{B})$ since D^{X2} is a nonfaithful $P23$ irrep. Inserting these results into the general formula (2.67) we obtain

$$\begin{aligned}
&Z^{X4,X1;X1,X4}(G1,(E|\vec{B}))\\
&\quad = E(3) \oplus U^{X2}(C_{2b}|\vec{B}) \oplus U^{X2}(C_{2b}|\vec{B}).
\end{aligned} \tag{3.114}$$

If we ignore the lexicographical ordering of the constituents in the KP decomposition of $D^{X1,X4}$ the corresponding generating relations of the first kind take the form

$$\begin{aligned}
C^{X4,X1} &= E(3) \otimes D^{X2}(E|2\vec{B})C^{X1,X4}\\
&\quad \times Z^{X4,X1;X1,X4}(G1,(E|\vec{B}))^{\dagger}.
\end{aligned} \tag{3.115}$$

To establish this sort of generating relation one does not need to know explicitly the generating matrices C^{H}. Apart from this we realize that the generating relations of the first kind lead to a drastic reduction, namely from the original 576 KP's to the 23 KP's given in Eq. (3.44). Both aspects demonstrate the usefulness of generating relations of the first kind.

The second step in our auxiliary group approach is to look for generating relations of the second kind. If such relations exist we clearly gain a further reduction in the actual computation of the CG matrices because of (67). Whether generating relations of the second kind exist or not follows immediately from the coset decomposition of $Q^{\mathbf{H}}$ with respect to $Q^{\mathbf{H},K}$. (The definitions of $Q^{\mathbf{H}}$ and $Q^{\mathbf{H},K}$ are given in the general part [see (62), (64), and (65)]). Whenever $Q^{\mathbf{H},K}$ is a proper subgroup of $Q^{\mathbf{H}}$ then nontrivial generating relations of the second kind must exist. They read

$$C_L^{\mathbf{H}} = U^{\mathbf{H},\mathbf{H}}(\mathbf{q}_L^{(\mathbf{H},K)})(Q_L^{(\mathbf{H},K)}C_K^{\mathbf{H}})Z_K^{\mathbf{H}}(q_L^{(\mathbf{H},K)})^{\dagger}, \tag{3.116}$$

where CG blocks are denoted by $C_L^{\mathbf{H}}$ and where

$$Z_K^{\mathbf{H}}(q_L^{(\mathbf{H},K)}) = E(\mathbf{H}|K) \otimes U^{L,K}(q_L^{(\mathbf{H},K)}). \tag{3.117}$$

For the sake of brevity we do not give all generating relations of the second kind occurring for the 23 KP's of Eq. (3.44). Again we concentrate on the preceding three examples for which we discussed the generating relations of the first kind at full length.

We start with the KP $D^{G4,G4}$. From (3.53) we deduce the homomorphic image of $\mathbf{Q}^{G4,G4}$,

$$\mathscr{H}(\mathbf{Q}^{G4,G4}) = C_3 \circledS (Im3m \times C_2) = Q^{G4,G4}. \tag{3.118}$$

The kernel of the homomorphism restricted to the subgroup $\mathbf{Q}^{G4,G4}$ is

$$\ker \mathscr{H} \cap \mathbf{Q}^{G4,G4} = \{(G1,G1),(G2,G2),(G3,G3)\} \circledS S_2. \tag{3.119}$$

If one takes into account (3.27), (3.79), and (3.118), one obtains

$$/G1/_{G4,G4} = \{G1,G2,G3\},\quad R^{G4,G4;G1} = \{G1,G2,G3\} = \{q_L^{(G4,G4;G1)}\}, \tag{3.120}$$

$$/G4/_{G4,G4} = \{G4\},\quad R^{G4,G4;G4} = \{G1\}. \tag{3.121}$$

At this point we do not have to know the CG blocks $C_{G1}^{G4,G4}$ and $C_{G4}^{G4,G4}$ explicitly. Their actual computation is done in the third and last step of the auxiliary group approach. The next task is to fix by some convention inverse images of the CR's $q_L^{(\mathbf{H},K)} \in R^{\mathbf{H},K}$. We choose

$$\begin{aligned} (G1,G1) &= \mathbf{q}_1 \in \mathscr{H}^{-1}(G1),\\ (G2,G1) &= \mathbf{q}_2 \in \mathscr{H}^{-1}(G2),\\ (G3,G1) &= \mathbf{q}_3 \in \mathscr{H}^{-1}(G3). \end{aligned} \tag{3.122}$$

(Note that the inverse image of the element $q_L^{(\mathbf{H},K)}$ is a set of the order of $\ker \mathscr{H}$.) Employing (3.35) the matrices $U^{G4,G4;G4,G4}(\mathbf{q}_j), j = 1,2,3$ are found to be

$$\begin{aligned} U^{G4,G4;G4,G4}(\mathbf{q}_1) &= E(3) \otimes E(3),\\ U^{G4,G4;G4,G4}(\mathbf{q}_2) &= U^{G4}(G2) \otimes E(3),\\ U^{G4,G4;G4,G4}(\mathbf{q}_3) &= U^{G4}(G2)^2 \otimes E(3). \end{aligned} \tag{3.123}$$

Due to our standardization procedure for $P23$ irreps we have

$$U^{G1,G1}(G1) = U^{G2,G1}(G2) = U^{G3,G1}(G3) = E(1), \tag{3.124}$$

which implies

$$Z^{G4,G4}(Gj) = E(1), \quad j = 1,2,3. \tag{3.125}$$

Finally inserting into the generating relations (3.116) we obtain

$$\begin{aligned} C_{G2}^{G4,G4} &= U^{G4}(G2) \otimes E(3) C_{G1}^{G4,G4},\\ C_{G3}^{G4,G4} &= U^{G4}(G2)^2 \otimes E(3) C_{G1}^{G4,G4}. \end{aligned} \tag{3.126}$$

Again one realizes the merits of our approach. One merely has to compute the CG block $C_{G1}^{G4,G4}$ whereas the other two CG blocks are determined by (3.126).

To get a better insight into generating relations of the second kind we also discuss in detail the other two examples beginning with the KP $D^{X1,X1}$. From (3.62) one readily derives the homomorphic image of $\mathbf{Q}^{X1,X1}$ and its kernel,

$$\mathscr{H}(\mathbf{Q}^{X1,X1}) = C_3 \circledS (Pm3m \times C_2) = Q^{X1,X1},$$

$$\ker \mathscr{H} \cap \mathbf{Q}^{X1,X1} = \{(G1,G1),(G2,G2),(G3,G3)\} \circledS S_2. \tag{3.127}$$

From (3.81) and (3.27) we deduce

$$/G1/_{X1,X1} = \{G1,G2,G3\},\quad R^{X1,X1;G1} = \{G1,G2,G3\}, \tag{3.128}$$

$$/M1/_{X1,X1} = \{M1\},\quad R^{X1,X1;M1} = \{G1\}. \tag{3.129}$$

Note that we used the relation $(R1)Q^{X1}(R1)^{-1} = Q^{M1} = Q^{X1}$ to derive (3.129). Again we only need to determine the CG blocks $C_{G1}^{X1,X1}$ and $C_{M1}^{X1,X1}$ to generate the remaining ones. Analogous to the previous case we fix by conventions the inverse images of the CR's $q_L^{(X1,X1;K)} \in R^{X1,X1;K}$. We take the same choice as in (3.122) which allows us to determine the matrices $U^{X1,X1;X1,X1}(\mathbf{q}_j), j = 1,2,3$. Using (3.37) we arrive at

$$\begin{aligned} U^{X1,X1;X1,X1}(\mathbf{q}_1) &= E(3) \otimes E(3),\\ U^{X1,X1;X1,X1}(\mathbf{q}_2) &= U^{X1}(G2) \otimes E(3),\\ U^{X1,X1;X1,X1}(\mathbf{q}_3) &= U^{X1}(G2)^2 \otimes E(3). \end{aligned} \tag{3.130}$$

Because of (3.124) we also obtain

$$Z^{X1,X1}(Gj) = E(1), \quad j = 1,2,3, \tag{3.131}$$

which leads to the following generating relations of the second kind:

$$\begin{aligned} C_{G2}^{X1,X1} &= U^{X1}(G2) \otimes E(3)\ C_{G1}^{X1,X1},\\ C_{G3}^{X1,X1} &= U^{X1}(G2)^2 \otimes E(3)\ C_{G1}^{X1,X1}. \end{aligned} \tag{3.132}$$

As in the last example we investigate the KP $D^{X1,X4}$. From (3.66) we obtain as homomorphic image

$$\mathscr{H}(\mathbf{Q}^{X1,X4}) = C_3 \circledS (Tm3m \times C_2) = Q^{X1,X4}. \tag{3.133}$$

Note, in particular, that in contrast to the previous case the group $Q^{X1,X4}$ contains $Q^{X1,X1}$ as the proper subgroup. Accordingly the kernel of the homomorphism must be smaller,

$$\ker \mathscr{H} \cap \mathbf{Q}^{X1,X4} = \{(G1,G1),(G2,G2),(G3,G3)\}. \tag{3.134}$$

Consulting the KP decomposition of $D^{X1,X4}$ and (3.27) we find

$$/G4/_{X1,X4} = \{G4\},\quad R^{X1,X4;G4} = \{G1\}, \tag{3.135}$$

$$/M3/_{X1,X4} = \{M3,M4\},$$
$$R^{X1,X4;M3} = \{(E|\vec{0}),(C_{2b}|\vec{0})\}. \quad (3.136)$$

In order to verify (3.136) one has to take into account the relation $Q^{M3} = (R1)Q^{X2}(R1)^{-1} = Q^{X2}$. We choose as inverse images

$$(G1,G1,(E|\vec{0})) = \mathbf{q}_1 \in \mathscr{H}^{-1}((E|\vec{0})),$$
$$(G1,G1,(C_{2b}|\vec{0})) = \mathbf{q}_2 \in \mathscr{H}^{-1}((C_{2b}|\vec{0})), \quad (3.137)$$

which allow us to construct the matrices $U^{X1,X4;X1,X4}(\mathbf{q}_L^{(X1,X4;M3)}), L = M3, M4$. Moreover we need

$$(C_{2b}|\vec{0})D^{X4} = (C_{2b}|\vec{0})(E|\vec{B})D^{X1} = (C_{2b}|\vec{B})D^{X1}$$
$$= (E|\vec{B})(E|\vec{B})^{-1}(C_{2b}|\vec{B})D^{X1}$$
$$= (E|\vec{B})(C_{2b}|\vec{0})D^{X1}, \quad (3.138)$$

where $(C_{2b}|\vec{0}) \in Q^{X1}$. Employing (3.37) we obtain

$$U^{X1,X4;X1,X4}(\mathbf{q}_1) = E(3) \otimes E(3),$$
$$U^{X1,X4;X1,X4}(\mathbf{q}_2) = U^{X1}(C_{2b}|\vec{0}) \otimes U^{X1}(C_{2b}|\vec{0}). \quad (3.139)$$

Finally we have to determine $U^{M4,M3}(q_{M4}^{(X1,X4;M3)})$ to define $Z^{X1,X4}(q_{M4}^{(X1,X4;M3)})$. For this purpose we have to compute

$$(G1,(C_{2b}|\vec{0}))D^{M3} = (G1,(C_{2b}|\vec{0}))(R1,(E|\vec{0}))D^{X2}$$
$$= (R1,(C_{2b}|\vec{0}))D^{X2} = D^{M4}. \quad (3.140)$$

Because $q' \in Q^{X2}$ turns out to be the trivial element we have to take

$$U^{M4,M3}(G1,(C_{2b}|\vec{0})) = E(3). \quad (3.141)$$

But this does not represent the most general situation since the matrices $U^{L',L}(q_L^{(H,K)})$ can be different from the unit matrices. In the last example generating relations of the second kind therefore take the following form:

$$C_{M4}^{X1,X4} = U^{X1}(C_{2b}) \otimes U^{X1}(C_{2b}) C_{M3}^{X1,X4}. \quad (3.142)$$

The third and last step in our procedure is to reduce (resolve) the multiplicity problem by means of auxiliary operator groups.

To carry out this step the first task is to determine the inverse images of the groups $Q^{\mathbf{H},K}$,

$$\mathbf{Q}^{\mathbf{H},K} = \mathscr{H}^{-1}(Q^{\mathbf{H},K}). \quad (3.143)$$

Let K_x be the representative of the $\mathbf{Q}^{\mathbf{H}}$ class $/K_x/_{\mathbf{H}}$. Clearly if $(\mathbf{H}|K) = 1$ then it is superfluous to determine the groups $\mathbf{Q}^{\mathbf{H},K}$ since the corresponding CG blocks $C_{K_x}^{\mathbf{H}}$ are unique up to arbitrary phase factors. Therefore we determine the inverse images $\mathbf{Q}^{\mathbf{H},K}$ only if $(\mathbf{H}|K_x) > 1$. We discuss this task again only for our three examples.

For the first example which concerns the KP $D^{G4,G4}$ we conclude from its KP decomposition that only $Q^{G4,G4;G4}$ must be determined. Inspecting (3.118), (3.53), and (3.119), one derives

$$\mathscr{H}^{-1}(Q^{G4,G4;G4}) = \mathbf{Q}^{G4,G4}. \quad (3.144)$$

For the second example we infer from its KP decomposition that only $\mathscr{H}^{-1}(Q^{X1,X1;M1})$ must be determined because $(X1,X1|G1) = 1$. From (3.127) and (3.62) one deduces

$$\mathscr{H}^{-1}(Q^{X1,X1;M1}) = \mathbf{Q}^{X1,X1}. \quad (3.145)$$

For the third example there is no need to determine the inverse images $\mathscr{H}^{-1}(Q^{X1,X4;K_x})$, $K_x = G4$, $M3$ because $(X1,X4|G4) = (X1,X4|M3) = 1$.

Due to their definition the CG blocks $C_{K,m}^{\mathbf{H}}$ have to satisfy

$$D^{\mathbf{H}}(R|\vec{t})C_{K,m}^{\mathbf{H}} = C_{K,m}^{\mathbf{H}}D^{K}(R|\vec{t}). \quad (3.146)$$

To obtain the CG matrix $C^{G4,G4}$ it suffices to compute the CG blocks $C_{G1}^{G4,G4}$ and $C_{G4,m}^{G4,G4}$, $m = 1,2$. The remaining CG blocks are defined by (3.126). Simple manipulations yield

$$\{C_{G1}^{G4,G4}\}_{s_1,s_2} = (1/\sqrt{3})\delta_{s_1,s_2}, \quad (3.147)$$

$$C_{G4,m}^{G4,G4} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a_m \\ 0 & b_m & 0 \\ 0 & 0 & b_m \\ 0 & 0 & 0 \\ a_m & 0 & 0 \\ 0 & a_m & 0 \\ b_m & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{array}{l} m = 1,2, \\ a_1 = b_1 = 1/\sqrt{2}, \\ a_2 = -b_2 = 1/\sqrt{2}, \end{array} \quad (3.148)$$

where we have fixed the CG blocks $C_{G4,m}^{G4,G4}$ by conventions. What remains to be settled is the question of whether the multiplicity problem is resolved by the auxiliary group approach or not. To be able to do this we adapt the general formulas (2.80) and (2.94) to our example,

$$\mathbf{H} = G4,G4; \quad K = G4; \quad \mathbf{q} \in \mathbf{Q}^{\mathbf{H},K},$$
$$T(\mathbf{q})C_{K,m}^{\mathbf{H}} = U^{\mathbf{H},\mathbf{H}}(\mathbf{q})(qC_{K,m}^{\mathbf{H}})U^{K,K}(q)^{\dagger} \quad (3.149)$$
$$= \sum_{m'} C_{K,m'}^{\mathbf{H}} L_{m',m}(\mathbf{q}).$$

Straightforward matrix multiplications yield for the generating elements of $\mathbf{Q}^{G4,G4;G4}$ the following matrices $L(\mathbf{q})$:

$$L((G2,G1)) = \frac{1}{2}\begin{vmatrix} -1 & -i\sqrt{3} \\ -i\sqrt{3} & -1 \end{vmatrix},$$
$$L((G1,G2)) = \frac{1}{2}\begin{vmatrix} -1 & i\sqrt{3} \\ i\sqrt{3} & -1 \end{vmatrix},$$
$$L((C_{2b}|\vec{0})) = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = L(p_{12}), \quad (3.150)$$
$$L((I|\vec{0})) = L((E|\vec{B})) = L(c) = E(2).$$

To arrive at these matrices we have used (3.35). Note in particular that (3.150) defines an irreducible corepresentation of the operator group $\tilde{\mathbf{Q}}^{G4,G4;G4} = D_3 \times \Theta$. Accordingly the multiplicity problem is resolved by the auxiliary group approach. One may conclude from the specific form of $L(p_{12})$ that the CG blocks $C_{G4,m}^{G4,G4}$ are γ symmetrized with respect to the permutation group S_2. This was the main reason to choose the constants a_m, b_m as we did in (3.148). The matrix $L(p_{12})$ shows that the permutation group S_2 with its irreps would have been sufficient to resolve the multiplicity problem. Piecing all things together we arrive at the following CG matrix $C^{G4,G4}$:

$$C^{G4,G4}=\begin{vmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & -1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & -1/\sqrt{2} \\ 1/\sqrt{3} & \alpha^*/\sqrt{3} & \alpha/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1/\sqrt{2} & 0 & 0 & -1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{3} & \alpha/\sqrt{3} & \alpha^*/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}. \tag{3.151}$$

The generating relations of the first kind are given by (3.91) and read explicitly

$$C^{G4,G4}=C^{G4,R4}=C^{R4,G4}=C^{R4,R4}. \tag{3.152}$$

The second example concerns the KP $D^{X1,X1}$. To arrive at the CG matrix $C^{X1,X1}$ it suffices to determine the CG blocks $C^{X1,X1}_{G1}$ and $C^{X1,X1}_{M1,m}$, $m=1,2$. By means of computer generation we obtained

$$\{C^{X1,X1}_{G1}\}_{\mathbf{R}_1,\mathbf{R}_2}=(1/\sqrt{3})\delta_{\mathbf{R}_1,\mathbf{R}_2}, \quad \underline{R}_j\in\underline{P}(X), \tag{3.153}$$

$$C^{X1,X1}_{M1,1}=\begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad C^{X1,X1}_{M1,2}=\begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \tag{3.154}$$

where in (3.154) we have chosen a special solution of the multiplicity problem (Refs. 12 and 13). The remaining CG blocks are defined by (3.132). The next task is to compute (3.149) for $\mathbf{q}\in\mathbf{Q}^{X1,X1;M1}=\mathbf{Q}^{X1,X1}$, the transformation properties of the CG blocks $C^{X1,X1}_{M1,m}$. To derive them we need (3.37) and the matrices $U^{M1}(q)$, $q\in\mathscr{H}(\mathbf{Q}^{X1,X1;M1})$. For that purpose we remember that $(R\,1)D^{X1}=D^{M1}$ holds. Therefore

$$q(R\,1)D^{X1}=(R\,1)(R\,1)q(R\,1)D^{X1}, \quad q\in Q^{X1,X1;M1}, \tag{3.155}$$

where $(R\,1)q(R\,1)\in Q^{X1,X1;M1}$ must be valid. From this we conclude

$$U^{M1}(q)=qU^{X1}((R\,1)q(R\,1)), \quad q\in Q^{X1,X1;M1}. \tag{3.156}$$

In detail we have for the generating elements of $Q^{X1,X1;M1}$ $(R\,1)q(R\,1)=q$. Consequently we arrive at

$$U^{M1}(G2)=U^{X1}(G2)=\begin{vmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha^* \end{vmatrix},$$

$$\alpha=\exp(i2\pi/3),$$

$$U^{M1}(C_{2b}|\vec{0})=U^{X1}(C_{2b}|\vec{0})=\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \tag{3.157}$$

$$U^{M1}(I|\vec{0})=U^{M1}(c)=E(3).$$

Employing these matrices together with (3.37) the corresponding transformation matrices $L(q)$ turn out to be

$$\begin{aligned} L((G2,G1))&=\begin{vmatrix} \alpha & 0 \\ 0 & \alpha^2 \end{vmatrix}, \\ L((G1,G2))&=\begin{vmatrix} \alpha^2 & 0 \\ 0 & \alpha \end{vmatrix}, \\ L((C_{2b}|\vec{0}))&=\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}=L(p_{12}), \\ L((I|\vec{0}))&=L(c)=E(2). \end{aligned} \tag{3.158}$$

We conclude from the matrix $L(p_{12})$ that the CG blocks $C^{X1,X1}_{M1,m}$ are not symmetrized with respect to the permutation group S_2. On the other hand the irreps of S_2 would again be sufficient to resolve the multiplicity problem. Apart from this note that (3.158) defines an irreducible corepresentation of the operator group $\tilde{\mathbf{Q}}^{X1,X1;M1}=D_3\times\Theta$. We therefore finally arrive at the following CG matrix $C^{X1,X1}$:

$$C^{X1,X1}=\begin{vmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1/\sqrt{3} & \alpha/\sqrt{3} & \alpha^*/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1/\sqrt{3} & \alpha^*/\sqrt{3} & \alpha/\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}. \tag{3.159}$$

To complete this example we use the generating relations (3.102) to obtain all CG matrices which belong to the Q-class $[X1,X1]$:

$$\begin{aligned} C^{X1,X1} = C^{X1,M1} &= C^{M1,X1} = C^{M1,M1} \\ &= C^{X4,X4} = C^{X4,M2} \\ &= C^{M2,X4} = C^{M2,M2}. \end{aligned} \tag{3.160}$$

Finally, let us summarize the simplifications that arise from the auxiliary group approach when the CG matrices of all KP's contained in the set $SA(G\times G)$, Eq. (3.40), are calculated. As already pointed out the generating relations of the first kind reduce the 576 KP's to 23 KP'S which refer to the Q-classes $[\mathbf{H}]$. To display the role of generating relations of the second kind and the reduction (resolution) of the multiplicity problem we repeat the decompositions (3.78)–(3.83) in a condensed form where the trivial KP's defined in (3.78) are omitted,

$$\begin{aligned} G4\otimes G4 &= \underline{G1\oplus G2\oplus G3}\oplus \underline{\underline{2}}\,G4, \\ G4\otimes G5 &= \underline{G5\oplus G6\oplus G7}, \\ G4\otimes X1 &= \underline{X2\oplus X3\oplus X4}, \\ G4\otimes X2 &= \underline{X1\oplus X4}\oplus X3, \\ G4\otimes X5 &= (\underline{\underline{2}}\oplus \underline{\underline{1}})X5, \end{aligned} \tag{3.161}$$

$$\begin{aligned} G5\otimes G5 &= G1\oplus G4, \\ G5\otimes X1 &= X5, \\ G5\otimes X2 &= X5, \\ G5\otimes X5 &= \underline{X1\oplus X4}\oplus \underline{X2\oplus X3}, \end{aligned} \tag{3.162}$$

$$\begin{aligned} X1\otimes X1 &= \underline{G1\oplus G2\oplus G3}\oplus \underline{\underline{2}}\,M1, \\ X1\otimes X4 &= G4\oplus \underline{M3\oplus M4}, \\ X1\otimes X2 &= G4\oplus M2\oplus M4, \\ X1\otimes X5 &= \underline{G5\oplus G6\oplus G7}\oplus \underline{\underline{2}}\,M5, \end{aligned} \tag{3.163}$$

$$\begin{aligned} X2\otimes X2 &= \underline{G1\oplus G2\oplus G3}\oplus \underline{\underline{2}}\,M3, \\ X2\otimes X3 &= G4\oplus \underline{M1\oplus M2}, \\ X2\otimes X5 &= \underline{G5\oplus G6\oplus G7}\oplus \underline{\underline{2}}\,M5, \end{aligned} \tag{3.164}$$

$$\begin{aligned} X5\otimes X5 = {} & \underline{G1\oplus G2\oplus G3}\oplus (\underline{\underline{2}}\oplus \underline{\underline{1}})G4 \\ & \oplus (\underline{\underline{1}}\oplus \underline{\underline{1}})(\underline{M1\oplus M2}\oplus \underline{M3\oplus M4}). \end{aligned} \tag{3.165}$$

In these formulas the effect of generating relations of the second kind are shown by underlining certain direct sums of inequivalent G irreps. For instance the meaning of these lines is that CG blocks $C^{G4,G4}_{G2}$ and $C^{G4,G4}_{G3}$ are generated by $C^{G4,G4}_{G1}$, etc. The results of third step are indicated by double underbars. For instance, $(\underline{\underline{2}}\oplus\underline{\underline{1}})$ means that the three $P23$ irreps $X5$ occurring in $G4\otimes X5$ transform according to two- and one-dimensional irreducible corepresentations, respectively. In this example *all* multiplicities can be explained in terms of irreducible corepresentations of auxiliary operator groups. Clearly the most effective step is to exploit generating relations of the first kind because they reduce the set of 576 KP's to 23 KP's. Generating relations of the second kind lead to further reductions as only CG blocks $C^{\mathbf{H}}_{K}$ have to be computed explicitly whereas the CG blocks $C^{\mathbf{H}}_{L}$, $L\in/K/$ are defined by simple matrix multiplications. But we are aware that the validity of our approach relies upon the fact that "standardized" G irreps have to be used. Finally we would like to mention that in Ref. 7 our approach is presented in a slightly modified form because there the permutation group has been incorporated in a different way.

IV. THREEFOLD KRONECKER PRODUCTS FOR THE GREY GROUP $C_4^*\times\Theta$

In the second example we discuss threefold KP's of coirreps of the grey group

$$\begin{aligned} &G(H) = C_4^*\times\Theta, \\ &C_4^* = \{E,C_4,C_4^2,C_4^3,\overline{E},\overline{C}_4,\overline{C}_4^2,\overline{C}_4^3\}, \\ &\Theta = \{E,\theta\}. \end{aligned} \tag{4.1}$$

Its generating elements are C_4 and θ. Coirreps of this group are tabulated in Ref. 14 from where we take over the notation. From these tables we see that $G(H)$ has two one-dimensional and three two-dimensional coirreps. The one-dimensional coirreps, labeled by 1 and 2, are of type I, where 1 denotes the trivial one. The two-dimensional coirreps, labeled by 3, 5, and 8, are of type III. To be consistent with (2.1) we choose

$$D^2(\theta) = 1. \tag{4.2}$$

Next we determine the auxiliary group taking into account (2.5) when fixing the admissible automorphisms. The automorphism group AUT turns out to consist of four elements. The corresponding mappings are

$$\begin{aligned} &\beta_0(C_4) = C_4, \quad \beta_0(\theta) = \theta; \\ &\beta_1(C_4) = \overline{C}_4^3, \quad \beta_1(\theta) = \overline{C}_4^3\theta; \\ &\beta_2(C_4) = C_4^3, \quad \beta_2(\theta) = C_4^3\theta; \\ &\beta_3(C_4) = \overline{C}_4, \quad \beta_3(\theta) = \overline{C}_4\theta. \end{aligned} \tag{4.3}$$

Assigning to each mapping β_j the corresponding group element b_j, the automorphism group AUT reads

$$\mathrm{AUT} = \{b_0,b_1,b_2,b_3\}\cong D_2. \tag{4.4}$$

Thus we arrive at

$$\begin{aligned} &\mathrm{ASS} = \{1,2\} = \{a_0,a\}\cong C_2', \\ &\mathrm{AUT} = \{b_0,b_1,b_2,b_3\}\cong D_2, \\ &\mathrm{CON} = \{c_0,c\}\cong C_2'', \end{aligned} \tag{4.5}$$

where the nontrivial element of ASS is denoted by a. Because of the structure of ASS the auxiliary group forms a direct product group,

$$\begin{aligned} Q &= \{(a^m,b_1^n b_2^r,c^s)\,|\,m,n,r,s = 1,2\} \\ &\cong C_2'\times D_2\times C_2'' = D_{2h}\times\Theta, \end{aligned} \tag{4.6}$$

where for the sake of clearness the subgroups of order 2 are distinguished by different superscripts. We take as generators of this group the elements (a,b_0,c_0), (a_0,b_1,c_0), (a_0,b_2,c_0), and (a_0,b_0,c). Later on we sometimes omit the trivial transformations a_0, b_0, c_0 to make the notation more concise.

The next task is to determine the Q-classes. We denote by $A\,(G(H))$ the index set of all coirrep labels

$$A(G(H)) = \{1,2,3,5,8\}. \quad (4.7)$$

The Q-classes can be found by inspecting the qk table for the coirreps of $G(H)$:

qk	1	2	3	5	8
a	2	1	3	8	5
b_1	1	2	3	5	8
b_2	1	2	3	8	5
c	1	2	3	5	8

(4.8)

From this we readily obtain the following Q-classes:

$$[1] = \{1,2\}, \quad [3] = \{3\}, \quad [5] = \{5,8\}. \quad (4.9)$$

The corresponding groups Q^k and CR sets R^k are given in the following list:

$$Q^1 = \mathrm{AUT}\times\mathrm{CON} \cong D_2 \times C_2'', \quad R^1 = \mathrm{ASS}, \quad (4.10)$$

$$Q^3 = Q = D_{2h}\times\Theta, \quad R^3 = \{q_0\}, \quad q_0 = (a_0,b_0,c_0), \quad (4.11)$$

$$Q^5 = \{q_0,(a_0,b_1),(a,b_2),(a,b_3)\}\times\mathrm{CON}, \quad R^5 = \{q_0,(a,b_1,c_0)\}. \quad (4.12)$$

Note that the choice for R^5 is not the simplest one as (a_0,b_1,c_0) is a nontrivial element of Q^5. The reason for taking this peculiar element is to arrive at standard coirreps which coincide with those tabulated in Ref. 15.

Next we choose the coirreps 1, 3, 5 as Q-class representatives and their generating matrices in the following form:

$$k=1: D^1(C_4) = 1 = D^1(\theta), \quad (4.13)$$

$$k=3:\ D^3(C_4) = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}, \quad D^3(\theta) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad (4.14)$$

$$k=5:\ D^5(C_4) = \begin{vmatrix} \alpha & 0 \\ 0 & \alpha^* \end{vmatrix}, \quad \alpha = \exp(i\pi/4),$$

$$D^5(\theta) = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \quad (4.15)$$

For the sake of completeness we also state the nontrivial one-dimensional coirrep of $G(H)$,

$$D^2(C_4) = -1, \quad D^2(\theta) = 1. \quad (4.16)$$

The next task is to determine the matrices $U^k(q), q\in Q^k$, where k denotes the Q-class representatives. Simple manipulations lead us to the following matrices:

$$k=1:\ U^1(q) = 1 \quad \text{for all } q\in Q^1, \quad (4.17)$$

$$k=3:\ U^3(a) = U^3(b_1) = U^3(b_2) = U(c) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad (4.18)$$

$$k=5:\ U^5(b_1) = U^5((a,b_2)) = U^5(c) = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \quad (4.19)$$

The next step is to fix the KP's considered in the following as this determines the structure of the auxiliary group **Q**. We discuss threefold KP's. Accordingly,

$$\mathbf{Q} = (\mathrm{ASS}^{(3)}\times\mathrm{AUT}\times\mathrm{CON})\circledS\mathrm{PERM}(3) \cong (C_2'^{(3)}\times D_2\times C_2'')\circledS S_3, \quad (4.20)$$

where the generating elements of $\mathrm{ASS}^{(3)}$ are (a,a_0,a_0), (a_0,a,a_0), and (a_0,a_0,a). The generating elements of AUT and CON are the triplets (b_1,b_1,b_1), (b_2,b_2,b_2), and (c,c,c), respectively. Finally we choose $p_{123} = (123)$ and $p_{12} = (12)(3)$ as generating elements of the permutation group S_3. Because of the structure of **Q**, Eq. (4.20), the elements of this group will be labeled by three associations, an automorphism, a conjugation, and a permutation, part of which may be trivial transformations,

$$\mathbf{q} = (a',a'',a''',b',c',p). \quad (4.21)$$

Following our approach we now have to determine the **Q**-classes. Adopting a notation similar to that in the previous example for the irrep label sets we have

$$|A(G(H)^{(3)})| = |A(G(H))|^3 = 5^3 = 125. \quad (4.22)$$

The **Q**-classes are disjoint subsets of $A(G(H)^{(3)})$ and turn out to be the following sets:

$$\begin{aligned}
[111] &= \{111, 112, 122, 222, \text{and permutations}\},\\
[113] &= \{113, 123, 223, \text{and permutations}\},\\
[115] &= \{115, 125, 225, 118, 128, 228, \text{and permutations}\},\\
[135] &= \{135, 235, 138, 238, \text{and permutations}\},\\
[155] &= \{155, 158, 188, 255, 258, 288, \text{and permutations}\},\\
[133] &= \{133, 233, \text{and permutations}\},\\
[333] &= \{333\},\\
[335] &= \{335, 338, \text{and permutations}\},\\
[355] &= \{355, 358, 388, \text{and permutations}\},\\
[555] &= \{555, 558, 588, 888, \text{and permutations}\}.
\end{aligned} \quad (4.23)$$

In contrast to the previous example it suffices to consider products of Q-class representatives to obtain all **Q**-classes. Again it is straightforward to show that

$$|A(G(H)^{(3)})| = \sum_{\mathbf{h}} |[\mathbf{h}]|. \quad (4.24)$$

We now focus on the single **Q**-class [335] and demonstrate how to proceed along the lines of our approach. In doing so we have to determine the group $\mathbf{Q}^{335}$ and to fix a CR set $\mathbf{R}^{335}$. One readily finds

$$\begin{aligned}
\mathbf{Q}^{335} &= (\mathbf{D}_2\times\mathbf{D}_2'\times\mathbf{C}_2)\circledS\mathbf{S}_2,\\
\mathbf{D}_2 &= \{(a^m,a^n,a_0,b_0,c_0,p_0)\,|\,m,n=1,2\},\\
\mathbf{D}_2' &= \{(a_0,a_0,a^m,b_2^m b_1^n,c_0,p_0)\,|\,m,n=1,2\},\\
\mathbf{C}_2 &= \{(a_0,a_0,a_0,b_0,c^n,p_0)\,|\,n=1,2\},\\
\mathbf{S}_2 &= \{(a_0,a_0,a_0,b_0,c_0,p_{12}^n)\,|\,n=1,2\}.
\end{aligned} \quad (4.25)$$

We choose the CR's from the set

$$\mathbf{R}^{335} = \{\mathbf{q}_0,\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4,\mathbf{q}_5\}, \quad (4.26)$$

where

$$\begin{aligned}
\mathbf{q}_0 &= \mathbf{q}_{335}^{(335)} = (a_0,a_0,a_0,b_0,c_0,p_0),\\
\mathbf{q}_1 &= \mathbf{q}_{533}^{(335)} = (a_0,a_0,a_0,b_0,c_0,p_{123}),\\
\mathbf{q}_2 &= \mathbf{q}_{353}^{(335)} = (a_0,a_0,a_0,b_0,c_0,p_{23}),\\
\mathbf{q}_3 &= \mathbf{q}_{338}^{(335)} = (a,a,a,b_1,c_0,p_0),\\
\mathbf{q}_4 &= \mathbf{q}_{833}^{(335)} = (a,a,a,b_1,c_0,p_{123}),\\
\mathbf{q}_5 &= \mathbf{q}_{383}^{(335)} = (a,a,a,b_1,c_0,p_{23}).
\end{aligned} \tag{4.27}$$

To establish generating relations of the first kind we now have to determine the matrices $U^{l,\mathbf{h}}(q_l^{(\mathbf{h})})$ and $U^{l',l}(q_l^{(\mathbf{h})})$. For the latter we need the KP decomposition of D^{335},

$$D^{335} \cong 2\,D^5 \oplus 2\,D^8. \tag{4.28}$$

Both D^5 and D^8 are of type III. As we use standard $G(H)$ coirreps we have $U^{8,5}((a,b_1)) = E(2)$. Moreover it follows from (4.18) that $U^3((a,b_1)) = U^3(a)U^3(b_1) = E(2)$. Therefore

$$\begin{aligned}
U^{338,335}(\mathbf{q}_3) &= U^3((a,b_1)) \otimes U^3((a,b_1)) \otimes U^3((a,b_1))\\
&= E(8).
\end{aligned} \tag{4.29}$$

Apart from this we need the matrices $U^{338,335}(\mathbf{q})$ for $\mathbf{q} = \mathbf{q}_1$, $\mathbf{q}_2$ to generate all the matrices $U^{338,335}(\mathbf{q}_j), j = 0,1,2,...,5$,

$$U^{338,335}(\mathbf{q}_1) = \begin{vmatrix}
1&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0\\
0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&1&0\\
0&1&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0\\
0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&1
\end{vmatrix},$$

$$U^{338,335}(\mathbf{q}_2) = \begin{vmatrix}
1&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0\\
0&1&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0\\
0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&1&0\\
0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&1
\end{vmatrix}. \tag{4.30}$$

The rows and columns of these matrices are enumerated in lexicographical order (in accordance with the KP $D^{335} = D^3 \otimes D^3 \otimes D^5$). To obtain the matrices $Z^{\mathbf{h}}(q_l^{(\mathbf{h})})$ we need the homomorphic images of the elements $\mathbf{q}_j \in \mathbf{R}^{335}$. These are easily found to be

$$\begin{aligned}
&\mathscr{H}(\mathbf{q}_0) = q_0 = q_0, \quad \mathscr{H}(\mathbf{q}_1) = q_1 = q_0,\\
&\mathscr{H}(\mathbf{q}_2) = q_2 = q_0, \quad \mathscr{H}(\mathbf{q}_3) = q_3 = (a,b_1),\\
&\mathscr{H}(\mathbf{q}_4) = q_4 = (a,b_1), \quad \mathscr{H}(\mathbf{q}_5) = q_5 = (a,b_1).
\end{aligned} \tag{4.31}$$

We see from (2.55) and (4.28) that the matrix $Z^{335}(q_3)$ must have the form

$$\begin{aligned}
Z^{335}(q_3) &= U^{8,5}((a,b_1)) \oplus U^{8,5}((a,b_1))\\
&\quad \oplus U^{5,8}((a,b_1)) \oplus U^{5,8}((a,b_1)).
\end{aligned} \tag{4.32}$$

Due to our standardization the matrix $U^{8,5}((a,b_1))$ is assumed to be a unit matrix. To obtain $U^{5,8}((a,b_1))$ we need to inspect the relation between D^8 and D^5,

$$(a,b_1)D^8 = (a,b_1)^2 D^5 = D^5, \tag{4.33}$$

which turns out to be an identity. Accordingly, $U^{5,8}((a,b_1)) = E(2)$ and hence

$$Z^{335}((a,b_1)) = E(8). \tag{4.34}$$

By similar arguments we finally arrive at

$$Z^{335}(q_j) = E(8), \quad j = 0,1,...,5. \tag{4.35}$$

Inserting this result into the general formula (2.61) we obtain the following generating relations of the first kind:

$$C^{\mathbf{k}_j} = U^{\mathbf{k}_j,335}(\mathbf{q}_j)C^{335}, \quad j = 0,1,...,5, \tag{4.36}$$

where $\mathbf{k}_j = \mathbf{q}_j(335) \in [335]$.

The next step in our approach is to establish generating relations of the second kind. For that purpose one has to find the subgroups $Q^{\mathbf{h},k}$ of $Q^{\mathbf{h}} = \mathscr{H}(\mathbf{Q}^{\mathbf{h}})$. If some of these groups are proper subgroups of $Q^{\mathbf{h}}$ then there exist nontrivial generating relations of the second kind. To state them one merely has to compute the matrices $U^{\mathbf{h},\mathbf{h}}(\mathbf{q}_l^{(\mathbf{h},k)})$ and $Z^{\mathbf{h}}(q_l^{(\mathbf{h},k)})$. First we determine the homomorphic image Q^{335} of $\mathbf{Q}^{335}$. One easily finds

$$\mathscr{H}(\mathbf{Q}^{335}) = Q^{335} = Q, \tag{4.37}$$

$$\ker \mathscr{H} \cap \mathbf{Q}^{335} = \{(a^m,a^m,a_0,b_0,p_{12}^n) \mid m,n = 1,2\}. \tag{4.38}$$

In this example the partitions $/k/_{\mathbf{h}}$ and the CR's $R^{\mathbf{h},k}$ are

$$/3/_{335} = \{3\}, \quad R^{335,3} = \{q_0\}, \tag{4.39}$$

$$\begin{aligned}
&/5/_{335} = \{5,8\},\\
&R^{335,5} = \{q_0,q'\}, \quad q' = q_8^{(335,5)} = (a,b_1).
\end{aligned} \tag{4.40}$$

Now we have to fix by suitable conventions inverse images of the transformations q_0, $q' \in R^{335,5}$. We choose

$$\begin{aligned}
&\mathscr{H}^{-1}(q_0) = \mathbf{q}_0,\\
&\mathscr{H}^{-1}(q') = \mathbf{q}' = (a,a_0,a_0,b_1,c_0,p_0).
\end{aligned} \tag{4.41}$$

Due to formula (2.32) we have

$$\begin{aligned}
U^{335,335}(\mathbf{q}') &= U^3((a,b_1)) \otimes U^3(b_1) \otimes U^5(b_1)\\
&= E(2) \otimes \begin{vmatrix} 0 & 1\\ 1 & 0 \end{vmatrix} \otimes \begin{vmatrix} 0 & -1\\ 1 & 0 \end{vmatrix},
\end{aligned} \tag{4.42}$$

the last equation following from (4.18) and (4.19). From (4.33) we infer

$$Z^{335}(q') = E(2) \oplus E(2) = E(4). \tag{4.43}$$

Accordingly we arrive at the following generating relations of the second kind:

$$C_8^{335} = U^{335,335}(\mathbf{q}')C_5^{335}, \tag{4.44}$$

where the corresponding CG blocks are denoted by C_5^{335} and C_8^{335}, respectively. Note that until now it was not necessary to compute the generating CG blocks explicitly.

The last step in our approach is to reduce (resolve) the multiplicity problem by means of (co)irreps of an auxiliary operator group. To this end we need the inverse image of $Q^{335,5}$,

$$Q^{335,5} = \{(a^m, b_2^m b_1^r, c^s) \mid m,r,s = 1,2\}, \tag{4.45}$$

$$\mathscr{H}^{-1}(Q^{335,5})$$
$$= \mathbf{Q}^{335,5} = \{(a^m,a^m,a^n,b_2^n b_1^r,c^s,p_{12}^t)\,|\,m,n,r,s,t = 1,2\}. \quad (4.46)$$

Next we compute the general CG block $C^{335}_{5,a}$. A simple calculation yields

$$C^{335}_{5,a} = \begin{array}{c} \sigma=1 \quad \sigma=2 \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \alpha & 0 \\ 0 & \beta^* \\ \beta & 0 \\ 0 & \alpha^* \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \end{array}, \quad \alpha,\beta \in \mathbb{C}. \quad (4.47)$$

Considering now the action of the operators $T(\mathbf{q})$, $\mathbf{q} \in \mathbf{Q}^{335,5}$, on the CG block (4.47) we see from (4.18) and (4.19) that these operators coincide for some of the generators $\mathbf{q}$ or even reduce to the unit operator,

$$T((a,a,a_0,b_0,c_0,p_0)) = T((a_0,a_0,a_0,b_0,c_0,p_{12})), \quad (4.48)$$
$$T((a_0,a_0,a,b_2,c_0,p_0)) = T((a_0,a_0,a_0,b_1,c_0,p_0)), \quad (4.49)$$
$$T((a_0,a_0,a_0,b_0,c,p_0)) = T((a_0,a_0,a_0,b_0,c_0,p_0)). \quad (4.50)$$

Note that (4.49) is antilinear whereas (4.48) is a linear operator. Accordingly,

$$\tilde{\mathbf{Q}}^{335,5} = C_2 \times \Theta. \quad (4.51)$$

The two-dimensional space spanned by the blocks (4.47) contains two irreducible subspaces spanned by the blocks $C^{335}_{5,m}$, where

$$\alpha = \beta = 1/\sqrt{2} \quad \text{for } m = 1,$$
$$\alpha = -\beta = 1/\sqrt{2} \quad \text{for } m = 2. \quad (4.52)$$

The corresponding coirreps are

$$L^1((a_0,a_0,a_0,b_0,c_0,p_{12})) = L^1((a_0,a_0,a_0,b_1,c_0,p_0)) = 1,$$
$$-L^2((a_0,a_0,a_0,b_0,c_0,p_{12})) = L^2((a_0,a_0,a_0,b_1,c_0,p_0)) = 1. \quad (4.53)$$

Putting all pieces together we arrive at the CG matrix

$$\sqrt{2}C^{335} = \begin{array}{c} \begin{array}{cccccccc} 511 & 512 & 521 & 522 & 811 & 812 & 821 & 822 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \end{array}, \quad (4.54)$$

where for convenience the column indices $lm\sigma$ are written explicitly. From this we obtain all the other CG matrices for the **Q**-class [335] using the generating relations given in (4.36).

V. THREEFOLD KRONECKER PRODUCTS FOR THE GREY GROUP $C_3^* \times \Theta$

In the last example we investigate threefold KP's of coirreps of the grey group

$$G(H) = C_3^* \times \Theta,$$
$$C_3^* = \{E,C_3,C_3^{-1},\bar{E},\bar{C}_3,\bar{C}_3^{-1}\}, \quad (5.1)$$
$$\Theta = \{E,\theta\}.$$

We choose the elements C_3 and θ as generators of this group. Coirreps of this group can be found in Ref. 16. From these tables we see that $G(H)$ has one one-dimensional and three two-dimensional coirreps.

According to our approach we first have to determine the corresponding auxiliary group taking into account (2.5) when defining admissible automorphisms. One readily verifies that the automorphism group AUT consists of two elements. The nontrivial element, denoted by b, corresponds to the following mapping:

$$\beta(C_3) = C_3^5, \quad \beta(\theta) = \theta. \quad (5.2)$$

Denoting the trivial automorphism by β_0 we have

$$\text{AUT} = \{b_0,b\}; \quad (5.3)$$

in the present example the group ASS is trivial and is therefore omitted. Thus we arrive at

$$Q = \text{AUT} \times \text{CON} \cong C_2 \times C_2' = C_2 \times \Theta, \quad (5.4)$$

where $b \in \text{AUT}$ and $c \in \text{CON}$ are the generating elements of this group.

Next we have to determine the Q-classes of $A(G(H))$ where this symbol denotes the set of coirrep labels of $G(H)$. From the character table[16] one readily derives the following qk table:

qk	1	2	4	6
b	1	2	4	6.
c	1	2	4	6

(5.5)

Consequently the Q-classes consist only of single elements, i.e.,

$$[k] = \{k\} \quad \text{for } = 1,2,4,6. \quad (5.6)$$

Therefore

$$Q^k = Q \quad \text{and} \quad R^k = \{q_0\} \quad (5.7)$$

for all Q-class representatives.

Because of (5.6) we have to fix the explicit form of all coirreps of $G(H)$,

$k=1$ (type I):

$$D^1(C_3)=1,\quad D^1(\theta)=1. \tag{5.8}$$

$k=2$ (type III):

$$D^2(C_3)=\begin{vmatrix}\alpha^5 & 0\\ 0 & \alpha\end{vmatrix},\quad \alpha=\exp(i\pi/3),$$
$$D^2(\theta)=\begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix}. \tag{5.9}$$

$k=4$ (type III):

$$D^4(C_3)=\begin{vmatrix}\alpha & 0\\ 0 & \alpha^5\end{vmatrix},$$
$$D^4(\theta)=\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}. \tag{5.10}$$

$k=6$ (type II):

$$D^6(C_3)=\begin{vmatrix}-1 & 0\\ 0 & -1\end{vmatrix},$$
$$D^6(\theta)=\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}. \tag{5.11}$$

These equations allow us to determine the corresponding matrices $U^k(q)$, $q\in Q^k$,

$$k=1{:}\ U^1(q)=1,\quad q\in Q. \tag{5.12}$$

$$k=2{:}\ U^2(b)=U^2(c)=\begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix}. \tag{5.13}$$

$$k=4{:}\ U^4(b)=U^4(c)=\begin{vmatrix}0 & -1\\ 1 & 0\end{vmatrix}. \tag{5.14}$$

$$k=6{:}\ U^6(b)=U^6(c)=\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix}. \tag{5.15}$$

It is worth emphasizing that $U^6(c)$ has been chosen to be different from $D^6(\theta)$. Such a choice is possible since $U^6(c)$ is unique only up to a unitary element of the corresponding commuting algebra.

In order to determine the structure of $\mathbf{Q}$ we have to fix the type of KP's, i.e., the number of factors, which we want to consider. We shall investigate threefold KP's. The corresponding auxiliary group reads

$$\begin{aligned}\mathbf{Q}&=\{(b^m,c^n,p_{23}^r\,p_{123}^s)\,|\,m,n,r=1,2;\ s=1,2,3\}\\ &=\mathrm{AUT}\times\mathrm{CON}\times\mathrm{PERM}\\ &\cong C_2\times C_2'\times S_3.\end{aligned} \tag{5.16}$$

For the sake of brevity we discuss in the following only the KP D^{446}. One easily finds

$$\mathbf{Q}^{446}\cong C_2\times C_2'\times S_2, \tag{5.17}$$

$$\mathbf{R}^{446}=\{\mathbf{q}_0,\mathbf{q}_1,\mathbf{q}_2\}, \tag{5.18}$$

where the choice of the CR's is as follows:

$$\begin{aligned}\mathbf{q}_0&=q_{446}^{(446)}=(b_0,c_0,p_0),\\ \mathbf{q}_1&=q_{644}^{(446)}=(b_0,c_0,p_{123}),\\ \mathbf{q}_2&=q_{464}^{(446)}=(b_0,c_0,p_{23}).\end{aligned} \tag{5.19}$$

Equations (5.18) and (5.19) show that $[446]=\{446,644,464\}$.

The first step in the auxiliary group approach for KP's is to establish generating relations of the first kind. This only requires to determine the matrices $U^{\mathbf{l},\mathbf{h}}(\mathbf{q}_l^{(\mathbf{h})})$ and $U^{l',l}(q_l^{(\mathbf{h})})$ which compose the matrices $Z^{\mathbf{h}}(q_l^{(\mathbf{h})})$. For the latter we need the KP decomposition of D^{446},

$$D^{446}=D^4\otimes D^4\otimes D^6\sim 2\,D^4\oplus 2\,D^6. \tag{5.20}$$

Because of (5.19) the matrices $U^{\mathbf{l},\mathbf{h}}(\mathbf{q}_l^{(\mathbf{h})})$ must be permutation matrices that correspond to p_{123} and p_{23}. As $\dim D^4=\dim D^6=2$ we can take the same matrices as in Sec. IV, i.e.,

$$\begin{aligned}U^{644,446}(\mathbf{q}_1)&=U^{338,335}(\mathbf{q}_1)\quad\text{of (4.30)},\\ U^{464,446}(\mathbf{q}_2)&=U^{338,335}(\mathbf{q}_2)\quad\text{of (4.30)}.\end{aligned} \tag{5.21}$$

To determine the matrices $Z^{446}(q_j)$, $j=1,2$, we have to know the homomorphic images of the elements $\mathbf{q}_j\in\mathbf{R}^{446}$. Due to (2.31) they are

$$\mathscr{H}(\mathbf{q}_j)=q_0;\quad j=0,1,2, \tag{5.22}$$

and therefore

$$Z^{446}(q_j)=E(8),\quad j=0,1,2. \tag{5.23}$$

Inserting this into the general formula (2.61) we arrive at the following generating relations of the first kind:

$$C^{\mathbf{k}_j}=U^{\mathbf{k}_j,446}(\mathbf{q}_j)C^{446},\quad j=0,1,2, \tag{5.24}$$

$$\mathbf{k}_j=\mathbf{q}_j(446)\in[446]. \tag{5.25}$$

Note that there are no generating relations of the second kind because of (5.6).

To tackle the multiplicity problem we first need the groups $Q^{446,4}$ and $Q^{446,6}$ and their inverse images. We have

$$\mathscr{H}^{-1}(Q^{446,4})=\mathscr{H}^{-1}(Q^{446,6})=\mathbf{Q}^{446}, \tag{5.26}$$

$$\mathbf{Q}^{446}=\{(b^m,c^n,p_{12}^r)\,|\,m,n,r=1,2\}. \tag{5.27}$$

As the KP decomposition (5.20) contains two inequivalent constituents which are not related by generating relations of the second kind we have to treat each of them separately.

We start with the CG block $C_{4,a}^{446}$, and try to resolve the multiplicity problem by means of coirreps of an operator group $\widetilde{\mathbf{Q}}^{446,4}$. The general form of this CG block is

$$C_{4,a}^{446}=\begin{vmatrix}0 & -\beta^*\\ 0 & \alpha^*\\ 0 & 0\\ 0 & 0\\ 0 & 0\\ 0 & 0\\ \alpha & 0\\ \beta & 0\end{vmatrix},\quad \alpha,\beta\in\mathbb{C}. \tag{5.28}$$

The blocks of this configuration form a two-dimensional linear space over $\mathbb{C}$ which is in accordance with the general considerations of Sec. II [$(446|4)=2$, D^2 is of type III].

The action of the operators $T(\mathbf{q})$, $\mathbf{q}\in\mathbf{Q}^{446}$, on these blocks follows from their definition, Eq. (2.80), the special form of the matrices $U^k(q)$, $k=4$, 6, given in (5.14) and (5.15), and

$$U^{446,446}(b_0,c_0,p_{12}) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \otimes \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}. \tag{5.29}$$

It is evident from (5.28) and (5.29) that $T(b_0,c_0,p_{12})$ is the unit operator; hence we omit the label p for the operators $T(\mathbf{q})$. It is easily verified that $T(b_0,c)$ is a linear operator of order 4 and that $T(b^3,c)$ is the operator of complex conjugation commuting with all other operators $T(\mathbf{q})$. Therefore

$$\tilde{\mathbf{Q}}^{446,4} = C_4 \times \Theta. \tag{5.30}$$

Choosing

$$\begin{aligned} &\alpha = 1/\sqrt{2}, \quad \beta = i/\sqrt{2} \quad \text{for } m=1, \\ &\alpha = 1/\sqrt{2}, \quad \beta = -i/\sqrt{2} \quad \text{for } m=2, \end{aligned} \tag{5.31}$$

we obtain two orthonormalized blocks which transform according to the corep generated by

$$L^4(b_0,c) = \begin{vmatrix} I & 0 \\ 0 & I^* \end{vmatrix}, \quad L^4(b^3,c) = \begin{vmatrix} 0 & E \\ E & 0 \end{vmatrix}. \tag{5.32}$$

The matrices

$$I = \text{diag}(i,-i), \quad E = E(2) \tag{5.33}$$

are elements of the ground field of the linear space corresponding to the numbers i, $1\in\mathbb{C}$. The corepresentation L^4 is therefore a coirrep of $C_4\times\Theta$ in standard form[16] where the complex numbers have been replaced by elements of the commuting algebra of D^4. Note that in this case the multiplicity $(446|4)$ equals the dimension of L^4 over the field $\{\text{diag}(\alpha,\alpha^*)|\alpha\in\mathbb{C}\}\cong\mathbb{C}$ and that the two blocks $C^{446}_{4,m}$, $m=1,2$, are uniquely determined by the corresponding irrep of the unitary subgroup $\{T(b_0,c)^n|n=1,2,3,4\}\cong C_4$.

To obtain the full CG matric C^{446} we also have to consider the blocks belonging to D^6. Their general form is

$$C^{446}_{6,\alpha} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ \alpha & \delta^* \\ \beta & -\gamma^* \\ \gamma & \beta^* \\ \delta & -\alpha^* \\ 0 & 0 \\ 0 & 0 \end{vmatrix}, \quad \alpha,\beta,\gamma,\delta\in\mathbb{C}. \tag{5.34}$$

Here we have eight real parameters since $(446|6)=2$ and D^6 is of type II. The action of the operators $T(\mathbf{q})$, $\mathbf{q}\in\mathbf{Q}^{446}$, follows again from Eqs. (2.80), (5.14), (5.15), and (5.29). One finds that the operators $T(b,c_0,p_0)$, $T(b_0,c_0,p_{12})$, and $T(b_0,c,p_0)$ generating $\tilde{\mathbf{Q}}^{446,6}$ are all of order 2; the first two are linear whereas the last one is antilinear. Moreover $T(b,c_0,p_{12})$ is the negative unit operator and $T(b,c,p_0)$ is the operator of complex conjugation commuting with all operators $T(\mathbf{q})$. Accordingly

$$\tilde{\mathbf{Q}}^{446,6} \cong D_2 \times \Theta. \tag{5.35}$$

The grey group on the rhs has four one-dimensional coirreps of type I.[14] Substituting numbers $\rho\in\mathbb{R}$ ($\subset\mathbb{Q}$) by 2×2 matrices $\rho E(2)$ belonging to the commuting algebra of D^6 we can find two of them in the two-dimensional quaternionic linear space spanned by the blocks (5.34). They are carried by basis blocks with

$$\begin{aligned} &\alpha=\gamma=1/\sqrt{2}, \quad \beta=\delta=0, \quad \text{for } m=1, \\ &\alpha=\gamma=0, \quad \beta=-\delta=1/\sqrt{2}, \quad \text{for } m=2, \end{aligned} \tag{5.36}$$

and generated by the matrices

$$\begin{aligned} &L^{6,1}(b,c_0,p_0) = -L^{6,1}(b_0,c_0,p_{12}) = -E(2), \\ &L^{6,2}(b,c_0,p_0) = -L^{6,2}(b_0,c_0,p_{12}) = E(2), \\ &L^{6,1}(b,c,p_0) = L^{6,2}(b,c,p_0) = E(2). \end{aligned} \tag{5.37}$$

Here too the multiplicity problem is resolved since the two blocks transforming D^{466} into D^6 may be distinguished by their transformation properties under the operator group $\tilde{\mathbf{Q}}^{446,6}$.

Combining all these results we arrive at the following CG matrix:

$$\sqrt{2}C^{446} = \begin{array}{c} \begin{matrix} 411 & 412 & 421 & 422 & 611 & 612 & 621 & 622 \end{matrix} \\ \begin{vmatrix} 0 & i & 0 & -i & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ i & 0 & -i & 0 & 0 & 0 & 0 & 0 \end{vmatrix} \end{array}. \tag{5.38}$$

From this matrix the other CG matrices $C^{\mathbf{h}}$, $\mathbf{h}\in[446]$, can be generated according to (5.24).

ACKNOWLEDGMENTS

This work is part of the scientific technical cooperation between Austria and Bulgaria.

Financial support of research visits in Sofia (R.D. and P.K.) and in Vienna (M.I.A. and J.N.K.) is gratefully acknowledged.

[1] R. Dirl, P. Kasperkovitz, M. I. Aroyo, J. N. Kotzev, and M. N. Angelova-Turkedjieva, J. Math. Phys. 27, 37 (1986).
[2] M. I. Aroyo, J. N. Kotzev, M. N. Angelova-Turkedjieva, R. Dirl, and P.

Kasperkovitz, J. Math. Phys. **27**, 2236 (1986).
[3]L. Jansen and M. Boon, *Theory of Finite Groups. Applications in Physics* (North-Holland, Amsterdam, 1967), Chap. II.7.
[4]J. N. Kotzev and M. I. Aroyo, in "Group theoretical methods in physics," in *Proceedings of the International Seminar*, Zvenigorod, USSR, 1982, edited by M. A. Markov (Nauka, Moscow, 1983), Vol. I, p. 363; M. I. Aroyo and J. N. Kotzev, in *XIIIth International Colloquium on Group Theoretical Methods in Physics*, College Park, MD, 1984, edited by W. W. Zachary (World Scientific, Singapore, 1984), p. 360.
[5]P. Kasperkovitz and G. Kahl, J. Math. Phys. **22**, 2404 (1981).
[6]A. P. Cracknell, B. L. Davies, S. C. Miller, and W. F. Love, *Kronecker Product Tables* (Plenum, New York, 1979), Vol. 1.
[7]R. Dirl, P. Kasperkovitz, M. I. Aroyo, J. N. Kotzev, and B. L. Davies, J. Phys. A **19**, L303 (1986).
[8]C. J. Bradley and A. P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Oxford U. P., Oxford, 1972).
[9]H. Burzlaff and H. Zimmermann, Z. Kristallogr. Kristallgeom. Kristallphys. Kristallchem. **153**, 151 (1980).
[10]B. L. Davies and A. P. Cracknell, *Kronecker Product Tables* (Plenum, New York, 1979), Vol. 2.
[11]*International Tables for Crystallography*, edited by T. Hahn (Reidel, Dordrecht, 1983).
[12]R. Dirl, J. Math. Phys. **20**, 659, 664, 671, 679 (1979).
[13]B. L. Davies, J. Phys. A **19**, 831, 841, 847 (1986).
[14]J. N. Kotzev and M. I. Aroyo, J. Phys. A **15**, 711 (1982).
[15]G. F. Koster, J. O. Dimmock, R. G. Weeler, and H. Statz, *Properties of the Thirty-Two Point Groups* (MIT, Cambridge, MA, 1963).
[16]J. N. Kotzev and M. I. Aroyo, J. Phys. A **15**, 725 (1982).

Nonlinear representations of Poincaré group in three dimensions

J. Bertrand and G. Rideau
Laboratory de Physique Theorique et Mathematique, University Paris VII and E. R. Centre National de la Recherche Scientifique No. 177, Tour 33-43 ler etage 2, Place Jussieu, 75251 Paris Cedex 05, France

(Received 25 November 1986; accepted for publication 8 April 1987)

The method of formal series is applied to the construction of nonlinear representations of Poincaré's group in three dimensions. The first term of the series must be a linear massless representation. In the special case of discrete helicities, the cocycles of extension of these representations by their tensor product are determined. It turns out that a nonlinear representation with a nonzero quadratic term must have a helicity zero representation as leading term. It is proved by induction how to avoid the successive obstructions to the computation of each term in the series.

I. INTRODUCTION AND FUNDAMENTALS

Interacting field equations are nonlinear equations and contain the transformation properties of the fields under the Poincaré group $\mathscr{P}$. Insofar as the solutions of such equations are uniquely determined from initial data, these transformation properties induce a nonlinear representation of $\mathscr{P}$ on the space of initial conditions. Since $\mathscr{P}$ includes time translations, the field dynamics is completely known from such a representation.

Thus we are naturally led to determine directly the nonlinear representations of $\mathscr{P}$ and to study their classification. Indeed, this is an extensive program and only partial answers have been given so far. In this paper, we follow the method initiated in Ref. 1, restricting the study to the Poincaré group in three space-time dimensions $\mathscr{P}_3$. This rather particular choice originates in the relative simplicity of $\mathscr{P}_3$, which nevertheless does not exhibit the oversimplified features of $\mathscr{P}$ in two space-time dimensions.

First, we recall the basic principles of Ref. 1 and the corresponding fundamental equations. Generally speaking, a representation of a group G in a topological vector space E is a separately continuous mapping

$$S\colon\ E\times G\to E$$

such that

$$S(0,g)=0, \tag{1.1}$$

$$S(\varphi,gg')=S(S(\varphi,g'),g), \tag{1.2}$$

where $\varphi\in E$ and $g,g'\in G$.

If S is a linear mapping, this reduces to the usual definition of a linear representation. The method proposed in Ref. 1 amounts to considering linear representations as the first term of a formal expansion:

$$S(\varphi,g)=\sum_1^\infty S_n(g)\varphi^{\otimes n}, \tag{1.3}$$

where $\otimes$ denotes the projective tensorial product and $S_n(g)$ is a linear mapping from $\otimes^n E$ into E. Here S is thus defined by a formal series. It is invertible in the formal sense if S_1 is invertible.

In this scheme, two representations S and S' are said to be equivalent if there exists a formal invertible mapping t: $E\to E$ such that

$$S'(\varphi,g)=t^{-1}(S(t(\varphi),g)). \tag{1.4}$$

In particular, if S' is linear, we will say that the representation S is formally linearizable.

Defining the operators

$$Z^n(g)=S_n(g)S_1(g^{-1})^{\otimes n},\quad n\geqslant 1, \tag{1.5}$$

we can write the group law (1.2) applied to the series (1.3) as

$$Z^n(gg')-Z^n(g)-S_1(g)Z^n(g')S_1(g^{-1})^{\otimes n}$$
$$=\sum_{q=2}^{n-1}Z^q(g)S_1(g)^{\otimes q}\sum_{\substack{l_1+\cdots+l_q=n\\ l_i\geqslant 1}}Z^{l_1}(g')\otimes\cdots\otimes Z^{l_q}(g')\sigma_n S_1(g^{-1})^{\otimes n}, \tag{1.6}$$

where σ_n is the customary symmetrization operator. When compatible, these equations allow us to construct the formal series from the first term S_1 by induction on the order n. For $n=2$, we have, in particular,

$$Z^2(gg')-Z^2(g)-S_1(g)Z^2(g')S_1(g^{-1})\otimes S_1(g^{-1})=0. \tag{1.7}$$

This defines a one-cocycle of extension of $S_1(g)$ by $S_1(g)\otimes S_1(g)$. For $n\geqslant 3$, the right-hand side is a two-cocycle of extension of S_1 by $S_1^{\otimes n}$. Equation (1.6) makes the cocycle a coboundary.

In this paper, we want to build nonlinear representations of $\mathscr{P}_3$ when S_1 is a unitary irreducible representation restricted to a convenient subspace of the Hilbert space. It has been proved in Ref. 2 that we can expect a nontrivial construction only if S_1 is a massless representation. So, we first describe with some details these massless representations (Sec. II) and the nontrivial cocycle of extension of one of them by the tensor product of two (Sec. III). We show that Eq. (1.7) has nontrivial solutions only if S_1 is a helicity zero representation. Then we prove that all the successive equations (1.6) can be solved, essentially because the little group of $\mathscr{P}_3$ on the cone is a one-parameter Abelian group (Sec. IV).

II. POINCARÉ GROUP $\mathscr{P}_3$: DEFINITION AND REPRESENTATIONS

We denote the elements of the three-dimensional translation group $\mathscr{T}_3$ by $a\equiv(a^0,\mathbf{a})\equiv(a^0,a^1,a^2)$ and those of SL(2,$\mathbb{R}$) by

$$\Lambda=\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix},\quad \alpha,\beta,\gamma,\delta\in\mathbb{R},\quad \alpha\delta-\beta\gamma=1.$$

The Poincaré group in three space-time dimensions $\mathscr{P}_3$ is the semidirect product $\mathscr{T}_3\wedge \mathrm{SL}(2,\mathbb{R})$, where the action of SL(2,$\mathbb{R}$) on $\mathscr{T}_3$, $a\to\Lambda a$, is given by

$$\begin{vmatrix}a^0-a^2 & a^1\\ a^1 & a^0+a^2\end{vmatrix}\overset{\wedge}{\to}\Lambda\begin{vmatrix}a^0-a^2 & a^1\\ a^1 & a^0+a^2\end{vmatrix}\Lambda^+.$$

Let M^3 be the three-dimensional Minkowski space with elements $x=(x^0,\mathbf{x})$ and M_3 its dual; the Minskowskian scalar product is denoted by $\langle\ ,\ \rangle$.

We consider the mass zero positive energy representations of $\mathscr{P}_3$. The corresponding orbit in $\widehat{M}_3$ is the cone C^+ without the origin. To any point $k\in C^+$ we associate the matrix Λ_k:

$$\Lambda_k=\begin{vmatrix}e^t & 0\\ 0 & e^{-t}\end{vmatrix}\begin{vmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{vmatrix},$$

$$t\in\mathbb{R},\quad -\pi/2<\varphi<\pi/2, \tag{2.1}$$

such that

$$k=\Lambda_k^{-1}\omega,\quad \omega=(\tfrac{1}{2},0,-\tfrac{1}{2})\in C^+.$$

Thus we can use (t,w),$w=\tan\varphi$, as coordinates of k:

$$\begin{aligned}k^0&=\tfrac{1}{2}e^{-2t},\\ k^1&=\tfrac{1}{2}e^{-2t}\sin 2\varphi,\\ k^2&=-\tfrac{1}{2}e^{-2t}\cos 2\varphi.\end{aligned} \tag{2.2}$$

The stabilizer of the point ω is $\Gamma_\omega=\mathscr{T}_3\wedge H$, where H is the Abelian group of matrices

$$h\equiv(\epsilon,z)\equiv\epsilon\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix},\quad \epsilon=\pm1,\quad z\in\mathbb{R}.$$

Starting with the unitary one-dimensional representations of H given by

$$(\epsilon,z)\in H\to\epsilon^\eta,\quad \eta=0,1,$$

we obtain, by induction, unitary representations of $\mathscr{P}_3$ realized in $L^2(C^+)$ according to

$$U_\eta(a,\Lambda)\colon f(\mathbf{k})\to e^{i\langle a,k\rangle}\epsilon^\eta(\Lambda,k)f(\Lambda^{-1}\mathbf{k}), \tag{2.3}$$

where the coordinates (t',w') of $\Lambda^{-1}k$ are given by

$$e^{-2t'}=e^{-2t}[(\alpha w-\gamma)^2+(\beta w-\delta)^2](1+w^2)^{-1},$$

$$w'=(\alpha w-\gamma)(\delta-\beta w)^{-1},$$

and $\epsilon(\Lambda,k)$ by

$$\epsilon(\Lambda,k)=\mathrm{sgn}(-\beta\tan\varphi+\delta). \tag{2.4}$$

Furthermore, we have

$$\epsilon(\Lambda^{-1},\Lambda^{-1}k)=\epsilon(\Lambda,k), \tag{2.5a}$$

$$\epsilon(\Lambda\Lambda',k)=\epsilon(\Lambda,k)\epsilon(\Lambda',\Lambda^{-1}k), \tag{2.5b}$$

$$\epsilon(\Lambda_k,\omega)=1. \tag{2.5c}$$

Actually, the Hilbert space $L^2(C^+)$ is not well adapted to the formulation of our cohomological problems. Its use would lead to the occurrence of formal coboundaries, which are difficult to interpret either mathematically or physically, and would discard significant cocycles meaningful only on dense subspaces.

Consequently, we consider the restriction V_η of U_η to the space $\mathscr{D}_\eta$ (see Ref. 3) of functions $f(t,w)$ such that (1) $f(t,w)$ has compact support in t, (2) $f(t,w)$ is C^∞ in t and w, and (3) $\hat{f}(t,w)\equiv\mathrm{sgn}^\eta(w)f(t,-1/w)$ has the same properties. A sequence $f_n(t,w)$ in $\mathscr{D}_\eta$ converges to zero if $f_n(t,w)$ and $\hat{f}_n(t,w)$ together with all their derivatives converge uniformly to zero on a fixed compact in t and on every compact in w.

Here $\mathscr{D}_\eta$ is a space dense in $L^2(C^+)$ [for the norm topology of $L^2(C^+)$] and contains as a subspace the space $\mathscr{D}(\mathbb{R}^2)$ whose elements are the C^∞ functions $f(t,w)$ of compact support in t and w. Also $\mathscr{D}(\mathbb{R}^2)$ is not invariant under the whole Poincaré group but only under a subgroup of elements (a,Λ), where a is any translation and Λ has the general form

$$\begin{vmatrix}\lambda & 0\\ z & \lambda^{-1}\end{vmatrix},\quad \lambda,z\in\mathbb{R}.$$

III. EXTENSIONS OF V_η by $V_{\eta_1}\otimes V_{\eta_2}$

We want to find an operator $Z(g)$, $g\in\mathscr{P}_3$, from $\mathscr{D}_{\eta_1}\times\mathscr{D}_{\eta_2}$ into $\mathscr{D}_\eta$ satisfying the one-cocycle equation

$$Z(gg')=Z(g)+V_\eta(g)Z(g')V_{\eta_1\eta_2}^{-1}(g), \tag{3.1}$$

where the notation $V_{\eta_1,\eta_2}(g)$ stands for $V_{\eta_1}(g)\otimes V_{\eta_2}(g)$ and Z for Z^2. Such an operator is determined up to a coboundary and we can choose as a representative of each equivalence class a cocycle which is C^∞ on $\mathscr{P}_3$ and equal to zero on SU(1) (see Ref. 4).

We define a continuous linear functional $Z_k(a,\Lambda)$ on $\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$ by

$$\begin{aligned}&(Z_{\mathbf{k}}(a,\Lambda),f_1(\mathbf{k}_1)\otimes f_2(\mathbf{k}_2))\\ &\quad=(Z(a,\Lambda)f_1\otimes f_2)(\mathbf{k}),\quad f_i\in\mathscr{D}_{\eta_i}.\end{aligned} \tag{3.2}$$

When we choose as $f_i(\mathbf{k}_i)$ functions in $\mathscr{D}(\mathbb{R}^2)$, then the left-hand side of (3.2) defines a continuous functional on $\mathscr{D}^2(\mathbb{R}^2)\equiv\mathscr{D}(\mathbb{R}^2)\otimes\mathscr{D}(\mathbb{R}^2)$. Therefore there exists a distribution $Z_{\mathbf{k}}(a,\Lambda;\mathbf{k}_1,\mathbf{k}_2)$ such that

$$\begin{aligned}&\int Z_{\mathbf{k}}(a,\Lambda;\mathbf{k}_1,\mathbf{k}_2)f_1(\mathbf{k}_1)f_2(\mathbf{k}_2)d\mathbf{k}_1\,d\mathbf{k}_2\\ &\quad=(Z_{\mathbf{k}}(a,\Lambda),f_1(\mathbf{k}_1)\otimes f_2(\mathbf{k}_2))\end{aligned}$$

whenever $f_i(\mathbf{k}_i)$ are C^∞ functions with compact support in (t_i,w_i).

A. Determination of $Z_{\mathbf{k}}(a,I)$

Lemma 1: There exists a functional $A_{\mathbf{k}}\in(\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2})'$ such that

$$Z_{\mathbf{k}}(a,I)=A_{\mathbf{k}}-e^{i\langle a,k\rangle}A_{\mathbf{k}}V_{\eta_1\eta_2}^{-1}(a,I). \tag{3.3}$$

Proof: First, we compute $Z_\omega(a,I)$. Let us consider the associated distribution $Z_\omega(a,I;\mathbf{k}_1,\mathbf{k}_2)$. As $\mathscr{D}(\mathbb{R}^2)$ is invariant under the translations subgroup, we can apply the considerations of Ref. 2 (cf. also Ref. 4) to prove the existence of a distribution $t(\mathbf{k}_1,\mathbf{k}_2)$ on $\mathscr{D}^2(\mathbb{R}^2)$ such that

$$Z_\omega(a,I;\mathbf{k}_1,\mathbf{k}_2) = (1 - e^{i\langle a,\omega - k_1 - k_2\rangle})t(\mathbf{k}_1,\mathbf{k}_2). \quad (3.4)$$

This gives us $Z_\omega(a,I)$ only on a subspace $\mathscr{D}(\mathbb{R}^2)$ of $\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$. To get it on the whole space, we proceed in two steps.[3]

(1) Let $f_i\in\mathscr{D}_{\eta_i}$ be such that f_1,f_2 are simultaneously equal to zero when w belongs to a bounded open set $\Omega\subset\mathbb{R}$. We can find z such that $-z^{-1}\in\Omega$; this implies

$$V_{\eta_i}^{-1}(0,h)f_i(t,w)\in\mathscr{D}(\mathbb{R}^2), \quad i=1,2, \quad (3.5)$$

for

$$h=\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}.$$

Then, from the cohomological equation, we have

$$\begin{aligned}(Z_\omega(a,I),f_1\otimes f_2) &= (Z_\omega(h^{-1}a,I),V_{\eta_1}^{-1}(0,h)f_1\otimes V_{\eta_2}^{-1}(0,h)f_2)\\ &\quad+(Z_\omega(0,h),f_1\otimes f_2)-e^{i\langle a,\omega\rangle}\\ &\quad\times(Z_\omega(0,h),V_{\eta_1}^{-1}(a,I)f_1\otimes V_{\eta_2}^{-1}(a,I)f_2).\end{aligned}$$

Taking (3.5) into account, we can apply (3.4) and get

$$\begin{aligned}(Z_\omega(a,I),f_1\otimes f_2) &= (Z^\Omega,f_1\otimes f_2)-e^{i\langle a,\omega\rangle}\\ &\quad\times(Z^\Omega,V_{\eta_1}^{-1}(a,I)f_1\otimes V_{\eta_2}^{-1}(a,I)f_2),\end{aligned}$$

where the functional Z^Ω is given by

$$\begin{aligned}(Z^\Omega,f_1\otimes f_2) &= (Z_\omega(0,h),f_1\otimes f_2)\\ &\quad+(t(\mathbf{k}_1,\mathbf{k}_2),V_{\eta_1}^{-1}(0,h)f_1\otimes V_{\eta_2}^{-1}(0,h)f_2).\end{aligned} \quad (3.6)$$

(2) Next, let $\Omega_1,\Omega_2,\Omega_3$ be three open bounded sets on $\mathbb{R}$ such that

(i) $\Omega_1\cap\Omega_2\neq\varnothing,\quad \Omega_i\cap C\Omega_j\neq\varnothing,\quad i,j=1,2,$

(ii) $\overline{\Omega_1\cap\Omega_2}\subset\Omega_3,$

(iii) $\overline{\Omega}_3\subset\Omega_1\cup\Omega_2,$

where C indicates the complement of a set. Now, any function $f_i\in\mathscr{D}_{\eta_i}$ can be written as

$$f_i=f_i^1+f_i^2+f_i^3,$$

where $f_i^j=0$ on Ω_j, $j=1,2$ and f_i^3 has a compact support such that

$$\overline{\Omega_1\cap\Omega_2}\subset\operatorname{supp}f_i^3\subset\overline{\Omega}_3.$$

Then we have

$$(Z_\omega(a,I),f_1\otimes f_2)=\sum_{i,j=1}^{3}(Z_\omega(a,I),f_1^i\otimes f_2^j)$$

and we can apply the previous results. Indeed, either both functions f_1^i,f_2^j are zero on some common bounded open set or both are of compact support. We eventually get

$$\begin{aligned}(Z_\omega(a,I),f_1\otimes f_2) &= (T,f_1\otimes f_2)-e^{i\langle a,\omega\rangle}(T,V_{\eta_1}^{-1}(a,I)f_1\\ &\quad\otimes V_{\eta_2}^{-1}(a,I)f_2),\end{aligned} \quad (3.7)$$

where T is a continuous linear functional on $\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$, according to the continuity of the mappings $f_i\to f_i^j, j=1,2,3$.

We can now compute $Z_\mathbf{k}(a,I)$. Indeed, applying (3.1) to the identity

$$Z_\omega((0,\Lambda_k)(a,I))=Z_\omega((\Lambda_k a,I)(0,\Lambda_k)),$$

noticing that, from (2.5c),

$$Z_\mathbf{k}(a,I)=V_\eta(0,\Lambda_k)Z_\omega(a,I)$$

and, taking (3.7) into account, we get

$$Z_\mathbf{k}(a,I)=A_\mathbf{k}-e^{i\langle a,k\rangle}A_\mathbf{k}V_{\eta_1\eta_2}^{-1}(a,I)$$

with

$$A_\mathbf{k}=[T-Z_\omega(0,\Lambda_k)]V_{\eta_1\eta_2}(0,\Lambda_k).$$

Proposition 1: The restriction of $Z(a,\Lambda)$ to the translation subgroup is a coboundary.

Proof: We have to verify that $\mathscr{A}_\mathbf{k}\equiv(A_\mathbf{k},f_1\otimes f_2)$ belongs to $\mathscr{D}_\eta$ for $f_i\in\mathscr{D}_{\eta_i}$ and that $A_\mathbf{k}$ is continuous.

From the assumptions above, $Z(0,\Lambda_k)$ is actually a C^∞ function of t. Using then the explicit expression of $V_{\eta_i}(0,\Lambda_k)$:

$$\begin{aligned}V_{\eta_i}(0,\Lambda_k)f_i(t_i,w_i) &= \operatorname{sgn}^{\eta_i}(e^{-t}-e^t ww_i)\\ &\quad\times f_i\left(t_i-\frac{1}{2}\log\frac{w_i^2e^{2t}+e^{-2t}}{1+w_i^2},\frac{e^tw_i+e^{-t}w}{e^{-t}-e^tww_i}\right),\end{aligned}$$

we conclude from the very definition of $\mathscr{D}_{\eta_i}$ that $V_{\eta_1\eta_2}(0,\Lambda_k)f_1\otimes f_2$ is C^∞ in t,w. Therefore $\mathscr{A}_\mathbf{k}$ is C^∞ in (t,w).

Let us now discuss $\hat{\mathscr{A}}_\mathbf{k}$. We write

$$\hat{\mathscr{A}}_\mathbf{k}=(\operatorname{sgn}^\eta w[T-Z_\omega(0,\Lambda_{\hat{k}})]V_{\eta_1\eta_2}(0,\Lambda_{\hat{k}}),f_1\otimes f_2), \quad (3.8)$$

where

$$\hat{k}=(t,-w^{-1}),\quad \Lambda_{\hat{k}}=u\tilde{\Lambda}_k,$$

$$\tilde{\Lambda}_k=\begin{vmatrix}e^t & 0\\ 0 & e^{-t}\end{vmatrix}\begin{vmatrix}\sin\varphi & -\cos\varphi\\ \cos\varphi & \sin\varphi\end{vmatrix},$$

$$u=(w/|w|)I.$$

Because of (3.1), we have

$$Z_\omega(0,\Lambda_{\hat{k}})=\operatorname{sgn}^\eta wZ_\omega(0,\tilde{\Lambda}_k)V_{\eta_1\eta_2}^{-1}(0,u).$$

On the other hand, straightforward manipulations using (3.1) and the identity

$$(a,I)=(0,u)(u^{-1}a,I)(0,u^{-1})$$

lead to another expression of $Z_\omega(a,I)$:

$$\begin{aligned}Z_\omega(a,I) &= (\operatorname{sgn}^\eta w)TV_{\eta_1\eta_2}(0,u)\\ &\quad-e^{i\langle a,w\rangle}(\operatorname{sgn}^\eta w)TV_{\eta_1\eta_2}(0,u)V_{\eta_1\eta_2}(-a,I).\end{aligned}$$

This allows us to substitute T for $(\operatorname{sgn}^\eta w)TV_{\eta_1\eta_2}(0,u)$ in Eq. (3.8) which becomes

$$\hat{\mathscr{A}}_k=([T-Z_\omega(0,\tilde{\Lambda}_k)]V_{\eta_1\eta_2}(0,\tilde{\Lambda}_k),\quad f_1\otimes f_2)$$

and the C^∞ property is proved as above.

To study the support of $\mathscr{A}_k$, choose $f(\mathbf{k}_1,\mathbf{k}_2)\in\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$ with $S=\operatorname{supp}f$. Then the following properties hold.

(i) There exist real numbers ϵ and R, depending only on S, such that

$$(Z_\mathbf{k}(a,I),f)=0,$$

whenever $|\mathbf{k}|<\epsilon$ or $|\mathbf{k}|>R$.

(ii) The function $\tilde{f}$ defined by

$$\tilde{f}(\mathbf{k}_1,\mathbf{k}_2) = (1 - e^{i\langle a,k' - k_1 - k_2\rangle})^{-1} f(\mathbf{k}_1,\mathbf{k}_2)$$

belongs to $\mathscr{D}_{\eta_1} \otimes \mathscr{D}_{\eta_2}$ and has support S provided $|\mathbf{k}'| < \epsilon'$ or $|\mathbf{k}'| > R'$, where again the real numbers ϵ', R' depend only on S.

From this, we conclude that

$$\mathscr{A}_{\mathbf{k}} \equiv (A_{\mathbf{k}}, f) = (Z_{\mathbf{k}}(a,I), \tilde{f}) = 0$$

for $|\mathbf{k}| < \min \epsilon,\epsilon'$ or $|\mathbf{k}| > \max R,R'$.

It should be remarked that we have incidentally proved that the support of $\mathscr{A}_{\mathbf{k}}$ depends only on the support of $f(\mathbf{k}_1,\mathbf{k}_2)$. The continuity of $\mathscr{A}_{\mathbf{k}}$ is easily inferred from this property. This completes the proof.[5]

Therefore, in the following, we shall assume

$$Z_{\mathbf{k}}(a,I) = 0. \tag{3.9}$$

B. Determining $Z_\omega(a,h),(a,h)\in\Gamma_\omega$

Developing the relation

$$Z_\omega((a,I)(0,h)) = Z_\omega((0,h)(h^{-1}a,I))$$

and using assumption (3.9), we get

$$Z_\omega(0,h) = e^{i\langle a,\omega\rangle} Z_\omega(0,h) V^{-1}_{\eta_1\eta_2}(a,I). \tag{3.10}$$

Again, we restrict ourselves to functions $f_1(\mathbf{k}_1) \otimes f_2(\mathbf{k}_2)$, $f_i \in \mathscr{D}_{\eta_i}$, with compact support in W_i. Thus the f_i are functions of compact support in k_i^1, k_i^2 and equal to zero in a neighborhood of the half-line $k_i^1 = 0$, $k_i^2 \geq 0$. Equation (3.10) can then be written in terms of a distribution kernel according to

$$(1 - e^{i\langle a,\omega - k_1 - k_2\rangle}) Z_\omega((0,h);\mathbf{k}_1,\mathbf{k}_2) = 0.$$

Proposition 2: If $\eta + \eta_1 + \eta_2 \neq 0 \bmod 2$, we have

$$Z_\omega((a,h);\mathbf{k}_1,\mathbf{k}_2) = 0.$$

If $\eta + \eta_1 + \eta_2 = 0 \bmod 2$, then

$$Z_\omega((a,h);\mathbf{k}_1,\mathbf{k}_2) = \delta(\boldsymbol{\omega} - \mathbf{k}_1 - \mathbf{k}_2)\{\delta(k_1^1)(u'(\lambda)z^2 + v(\lambda)z) + \delta'(k_1^1)u(\lambda)z\}, \tag{3.11}$$

where

$$u(\lambda), v(\lambda) \in \mathscr{D}'(]0,1[).$$

Proof: We must have

$$(\boldsymbol{\omega} - \mathbf{k}_1 - \mathbf{k}_2) Z_\omega((0,h);\mathbf{k}_1,\mathbf{k}_2) = 0, \tag{3.12}$$

$$(\tfrac{1}{2} - |\mathbf{k}_1| - |\mathbf{k}_2|) Z_\omega((0,h);\mathbf{k}_1,\mathbf{k}_2) = 0. \tag{3.13}$$

From (3.12), we get

$$(Z_\omega(0,h), f_1 \otimes f_2) = \int S(h;\mathbf{k}) \psi(\mathbf{k}) d\mathbf{k}, \tag{3.14}$$

with

$$\psi(\mathbf{k}) \equiv f_1(\mathbf{k}) f_2(\boldsymbol{\omega} - \mathbf{k}).$$

Therefore $S(h;\mathbf{k})$ is a distribution on the space of functions with compact support in k^1, k^2, equal to zero in a neighborhood of the set $k^1 = 0$, $k^2 \geq 0$, or $k^2 \leq -\frac{1}{2}$.

Equation (3.13) then yields

$$(k^1)^2 S(h,\mathbf{k}) = 0.$$

Hence

$$S(h,\mathbf{k}) = \delta(k^1) s_0(h;\lambda) + \delta'(k^1) s_1(h;\lambda), \tag{3.15}$$

where $\lambda \equiv -2k^2$ and s_0, s_1 are elements of $\mathscr{D}'(]0,1[)$. These are to be determined through the use of the cohomological equation (3.1) for H which reads

$$\int [\delta(k^1) s_0(hh';\lambda) + \delta'(k^1) s_1(hh';\lambda)] \psi(\mathbf{k}) d\mathbf{k}$$

$$= \int [\delta(k^1) s_0(h;\lambda) + \delta'(k^1) s_1(h;\lambda)] \psi(\mathbf{k}) d\mathbf{k}$$

$$+ \epsilon^\eta \int \epsilon^{\eta_1}(h^{-1},k) \epsilon^{\eta_2}(h^{-1},\omega - k) \times [\delta(k^1) s_0(h';\lambda) + \delta'(k^1) s_1(h';\lambda)] \psi(\mathbf{hk}) d\mathbf{k}.$$

Identifying terms, we get

$$s_0(hh';\lambda) = s_0(h;\lambda) + \epsilon^{\eta + \eta_1 + \eta_2} s_0(h';\lambda) + 2\epsilon^{\eta + \eta_1 + \eta_2} z \frac{\partial s_1}{\partial \lambda}(h';\lambda), \tag{3.16}$$

$$s_1(hh';\lambda) = s_1(h;\lambda) + \epsilon^{\eta + \eta_1 + \eta_2} s_1(h';\lambda). \tag{3.17}$$

If $\eta + \eta_1 + \eta_2 \neq 0 \bmod 2$, the only solution of Eqs. (3.16), (3.17) is $s_0 = s_1 = 0$.

If $\eta + \eta_1 + \eta_2 = 0 \bmod 2$, the solution of (3.17) is

$$s_1(h;\lambda) = u(\lambda) z, \tag{3.18}$$

where $u \in \mathscr{D}'(]0,1[)$.

Solving (3.16), we get

$$s_0(h;\lambda) = u'(\lambda) z^2 + v(\lambda) z, \tag{3.19}$$

where

$$v \in \mathscr{D}'(]0,1[).$$

Collecting results from Eqs. (3.14), (3.15), (3.18), and (3.19), we can write finally

$$Z_\omega((a,h);\mathbf{k}_1,\mathbf{k}_2) = \delta(\boldsymbol{\omega} - \mathbf{k}_1 - \mathbf{k}_2)\{\delta(k_1^1)(u'(\lambda)z^2 + v(\lambda)z) + \delta'(k_1^1)u(\lambda)z\}.$$

Now, we extend $Z_\omega(a,h)$ to the whole $\mathscr{D}_{\eta_1} \otimes \mathscr{D}_{\eta_2}$.

Lemma 2: $(Z_\omega(0,h), f) = 0$, whenever $f(t_1,w_1,t_2,w_2) \in \mathscr{D}_{\eta_1} \otimes \mathscr{D}_{\eta_2}$ is such that there exists a neighborhood W of the origin on the real line with

$$f(t_1,w_1,t_2,w_2) = 0,$$

for w_1 or w_2 in W.

Proof: First, (3.10) implies that for any $F \in \mathscr{D}_{\eta_1} \otimes \mathscr{D}_{\eta_2}$, we have

$$(Z_\omega(0,h), (1 - e^{i\langle a,\omega - k_1 - k_2\rangle}) F) = 0.$$

Taking derivatives, we infer

$$(Z_\omega(0,h), (e^{-2t_1}[w_1^2/(1 + w_1^2)] + e^{-2t_2}[w_2^2/(1 + w_2^2)]) F) = 0. \tag{3.20}$$

If f has the properties stated in the lemma, then

$$F(t_1,w_1,t_2,w_2) = (e^{-2t_1}[w_1^2/(1 + w_1^2)] + e^{-2t_2}[w_2^2/(1 + w_2^2)])^{-1} f(t_1,w_1,t_2,w_2)$$

belongs to $\mathscr{D}_{\eta_1} \otimes \mathscr{D}_{\eta_2}$ since the factor in front of f is well defined on the support of f.

From (3.20), we get

$$(Z_\omega(0,h), f) = 0.$$

Let $\chi(w)$ be a C^∞ function with compact support equal to one on a neighborhood of the origin. For $F\in\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$ we have

$$F=F\chi(w_1)\chi(w_2)+F(1-\chi(w_1))\chi(w_2)$$
$$+F\chi(w_1)(1-\chi(w_2))+F(1-\chi(w_1))(1-\chi(w_2)).$$

By linearity and Lemma 2, we then get

$$(Z_\omega(0,h),F)=(Z_\omega(0,h),F(t_1,w_1,t_2,w_2)\chi(w_1)\chi(w_2)),$$

so that the operator $Z_\omega(0,h)$ is thoroughly determined by the distribution $Z_\omega((0,h);\mathbf{k}_1,\mathbf{k}_2)$.

Remark: In (3.11), the v-dependent term is a coboundary.

Proof: If $Z_\omega((a,h);\mathbf{k}_1,\mathbf{k}_2)$ is a coboundary, there exists a distribution $A_\omega(\mathbf{k}_1,\mathbf{k}_2)$ such that

$$Z_\omega((a,h);\mathbf{k}_1,\mathbf{k}_2)$$
$$=A_\omega(\mathbf{k}_1,\mathbf{k}_2)-e^{i(a,\omega-k_1-k_2)}\epsilon^\eta(h,\omega)$$
$$\times\prod_{i=1,2}\epsilon^{\eta_i}(h,k_i)|\mathbf{h}^{-1}\mathbf{k}_i|/|\mathbf{k}_i|A_\omega(\mathbf{h}^{-1}\mathbf{k}_1,\mathbf{h}^{-1}\mathbf{k}_2). \tag{3.21}$$

Since $Z_\omega((a,I);\mathbf{k}_1,\mathbf{k}_2)$ is equal to zero, we conclude that A_ω satisfies (3.12), (3.13) and can be written as

$$A_\omega(\mathbf{k}_1,\mathbf{k}_2)=\delta(\omega-\mathbf{k}_1-\mathbf{k}_2)$$
$$\times[\delta(k_1^1)\widetilde{A}_0(\lambda)+\delta'(k_1^1)\widetilde{A}_1(\lambda)]$$

with $\widetilde{A}_0,\widetilde{A}_1\in\mathscr{D}'(]0,1[)$. Identifying (3.11) and (3.21), we obtain the result with

$$\widetilde{A}_0(\lambda)=0,\quad v(\lambda)=-2\widetilde{A}_1'(\lambda).$$

C. Final determination of $Z_\mathbf{k}(a,\Lambda)$

Applying the cohomological equation to both sides of the identity

$$Z_\omega((0,\Lambda_k)(a,\Lambda))=Z_\omega((\Lambda_k a,h_k(\Lambda))(0,\Lambda_{\Lambda^{-1}k})),$$
$$k=\Lambda_k^{-1}\omega,\quad h_k(\Lambda)=\Lambda_k\Lambda\Lambda_{\Lambda^{-1}k}^{-1},$$

we get the relation

$$Z_\omega(0,\Lambda_k)+V_\eta(0,\Lambda_k)Z_\omega(a,\Lambda)V_{\eta_1\eta_2}^{-1}(0,\Lambda_k)$$
$$=Z_\omega(\Lambda_k a,h_k(\Lambda))+V_\eta(\Lambda_k a,h_k(\Lambda))$$
$$\times Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}^{-1}(\Lambda_k a,h_k(\Lambda)). \tag{3.22}$$

But we have from (2.5c),

$$V_\eta(0,\Lambda_k)Z_\omega(a,\Lambda)=Z_\mathbf{k}(a,\Lambda)$$

and, from the group law,

$$(\Lambda_k a,h_k(\Lambda))^{-1}(0,\Lambda_k)=(0,\Lambda_{\Lambda^{-1}k})(a,\Lambda)^{-1}.$$

Hence, Eq. (3.22) yields

$$Z_\mathbf{k}(a,\Lambda)=Z_\omega(\Lambda_k a,h_k(\Lambda))V_{\eta_1\eta_2}(0,\Lambda_k)$$
$$-Z_\omega(0,\Lambda_k)V_{\eta_1\eta_2}(0,\Lambda_k)+V_\eta(\Lambda_k a,h_k(\Lambda))$$
$$\times Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}^{-1}(a,\Lambda).$$

The explicit computation of the last term gives

$$V_\eta(\Lambda_k a,h_k(\Lambda))Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}(0,\Lambda_{\Lambda^{-1}k})$$
$$=e^{i(\Lambda_k a,\omega)}\epsilon^\eta(h_k(\Lambda),\omega)Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}(0,\Lambda_{\Lambda^{-1}k})$$

or, from Eq. (2.5),

$$V_\eta(\Lambda_k a,h_k)Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}(0,\Lambda_{\Lambda^{-1}k})$$
$$=e^{i(a,k)}\epsilon^\eta(\Lambda,k)Z_\omega(0,\Lambda_{\Lambda^{-1}k})V_{\eta_1\eta_2}(0,\Lambda_{\Lambda^{-1}k}). \tag{3.23}$$

Defining a function of k by

$$B(k)\equiv Z_\omega(0,\Lambda_k)V_{\eta_1\eta_2}(0,\Lambda_k),$$

we then recognize in the rhs of (3.23) the action of $V_\eta(a,\Lambda)$ on $B(k)$.

As shown above (Sec. III A), the operator B associated with the linear functional $B(k)$ [as in (3.2)] maps $\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$ into $\mathscr{D}_\eta$. Thus we have, up to a trivial cocycle to be neglected,

$$Z_\mathbf{k}(a,\Lambda)=Z_\omega(\Lambda_k a,h_k(\Lambda))V_{\eta_1\eta_2}(0,\Lambda_k). \tag{3.24}$$

A priori, the full expression (3.11) has to be substituted here. However, it is easily shown that the coboundary part in (3.11) leads to a coboundary contribution in the cocycle $Z_\mathbf{k}(a,\Lambda)$. Therefore only the $u(\lambda)$-dependent part must be used in (3.24). So we get the following.

Theorem 1: For $\eta+\eta_1+\eta_2=0 \bmod 2$, the classes of extension cocycles of V_η by $V_{\eta_1}\otimes V_{\eta_2}$ can be labeled by a distribution $u(\lambda)\in\mathscr{D}'(]0,1[)$. Each class contains a representative $Z_\mathbf{k}(a,\Lambda)$ with

$$(Z_\mathbf{k}(a,\Lambda),f)$$
$$=\int_0^1 d\lambda\left\{u'(\lambda)z^2(\Lambda,k)+u(\lambda)\frac{z(\Lambda,k)}{|k|}\right.$$
$$\left.\times\left[k^2\left(\frac{\partial}{\partial\mathbf{k}_1^1}-\frac{\partial}{\partial\mathbf{k}_2^1}\right)-k^1\left(\frac{\partial}{\partial\mathbf{k}_1^2}-\frac{\partial}{\partial\mathbf{k}_2^2}\right)\right]\right\}$$
$$\times f(\mathbf{k}_1,\mathbf{k}_2)\Big|_{\substack{\mathbf{k}_1=\lambda\mathbf{k}\\ \mathbf{k}_2=(1-\lambda)\mathbf{k}}}, \tag{3.25}$$

where $f\in\mathscr{D}_{\eta_1}\otimes\mathscr{D}_{\eta_2}$ and $z(\Lambda,k)$ is defined by

$$h_k(\Lambda)\equiv\Lambda_k\Lambda\Lambda_{\Lambda^{-1}k}^{-1}=\epsilon(\Lambda,k)\begin{vmatrix}1 & z(\Lambda,k)\\ 0 & 1\end{vmatrix},$$
$$k=(k^0,k^1,k^2). \tag{3.26}$$

IV. REMOVAL OF OBSTRUCTIONS TO THE CONSTRUCTION OF A NONLINEAR REPRESENTATION WITH A QUADRATIC TERM

We now proceed to the construction of a nonlinear representation (1.3) of $\mathscr{P}_3$ driven by V_η as the linear part and containing, at least, a quadratic term. We see at once from Proposition 2, that this will not be possible unless $\eta=0$. According to the general scheme (Sec. I), we first study the two-cocycle

$$C(g,g')=Z(g)(Y(gg')-Y(g)),$$

where

$$Y(g)=Z(g)\otimes I+I\otimes Z(g).$$

Proposition 3: $C(g,g')$ is a coboundary, i.e., there exists $Z^3(g)$ mapping $\mathscr{D}_0^3$ (the third symmetrical tensor power of $\mathscr{D}_0$) into $\mathscr{D}_0$ such that

$$C(g,g')=Z^3(gg')-Z^3(g)-V(g)Z^3(g')V^{(3)-1}(g), \tag{4.1}$$

where

$$V_0=V,\quad V_0\otimes V_0\otimes V_0=V^{(3)}.$$

Proof: As above, we introduce $C_{\mathbf{k}}(g,g')$ and $Z^3_{\mathbf{k}}(g)$ in $(\mathscr{D}^3_0)'$ by

$$(C_{\mathbf{k}}(g,g'),f) = (C(g,g')f)(\mathbf{k}), \quad f\in\mathscr{D}^3_0,$$

$$(Z^3_{\mathbf{k}}(g),f) = (Z^3(g)f)(\mathbf{k}).$$

Since, according to the definition of C,

$$C((0,u);g') = C(g;(0,u)) = 0, \quad u\in \mathrm{SU}(1),$$

Eq. (4.1) implies that the restriction of $Z^3(g)$ to SU(1) is a one-cocycle of extension over a compact group. Hence it can be assumed that

$$Z^3(0,u) = 0. \tag{4.2}$$

Furthermore, as

$$C((a,I);g') = C(g;(a,I)) = 0,$$

we get from (4.1) the equation

$$Z^3_\omega(0,\Lambda_k) + Z^3_{\mathbf{k}}(a,I)V^{(3)-1}(0,\Lambda_k) = Z^3_\omega(\Lambda_k a,I) + e^{i(a,k)}Z^3_\omega(0,\Lambda_k)V^{(3)-1}(\Lambda_k a,I),$$

which is consistent with the assumptions

$$Z^3_\omega(0,\Lambda_k) = Z^3_{\mathbf{k}}(a,I) = 0. \tag{4.3}$$

Now, using again the identity

$$(0,\Lambda_k)(a,\Lambda) = (\Lambda_k a, h_k(\Lambda))(0,\Lambda_{\Lambda^{-1}k}),$$

$$h_k(\Lambda) = \Lambda_k \Lambda \Lambda^{-1}_{\Lambda^{-1}k},$$

we obtain from (4.1)

$$Z^3_{\mathbf{k}}(a,\Lambda) = Z^3_\omega(0,h_k(\Lambda))V^{(3)}(0,\Lambda_k) + Z_\omega(0,h_k(\Lambda)) \times [Y(\Lambda_k a,\Lambda_k\Lambda) - Y(\Lambda_k a,h_k(\Lambda))]V^{(3)}(0,\Lambda_k). \tag{4.4}$$

Therefore we need to determine $Z^3_\omega(0,h_k(\Lambda))$. It satisfies the following cohomological equation:

$$Z^3_\omega(0,hh') - Z^3_\omega(0,h) - Z^3_\omega(0,h')V^{(3)-1}(0,h) = Z_\omega(0,h)[Y(0,hh') - Y(0,h)], \tag{4.5}$$

where

$$h = \epsilon\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}, \quad h' = \epsilon'\begin{vmatrix}1 & z'\\ 0 & 1\end{vmatrix}.$$

Then, it can be verified directly that

$$Z^3_\omega(0,h) = \int_0^z Z_\omega\left(0,\begin{vmatrix}1 & \sigma\\ 0 & 1\end{vmatrix}\right)\frac{dY}{d\sigma}\left(0,\begin{vmatrix}1 & \sigma\\ 0 & 1\end{vmatrix}\right)d\sigma \tag{4.6}$$

is the solution of (4.5) whose derivative is zero at the origin. Substituting (4.6) into (4.4), we end up with a solution of (4.1) satisfying (4.2) and (4.3). Thus we have removed the first obstruction to the general construction. We now prove by induction that the same holds for all other obstructions.

Theorem 2: Assume that $Z^i(a,\Lambda)$ is known for $i\leq n-1$ with the properties

$$Z^i_{\mathbf{k}}(a,I) = Z^i_\omega(0,\Lambda_k) = 0,$$

$$Z^i(0,u) = 0, \quad u\in\mathrm{SU}(1),$$

$$\frac{dZ^i_\omega}{dz}\left(0,\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}\right)\bigg|_{z=0} = 0, \quad i>2.$$

Then there exists $Z^n(a,\Lambda)$, mapping $\mathscr{D}^n_0$ (the nth symmetrized tensor power of $\mathscr{D}_0$) into $\mathscr{D}_0$ such that

$$Z^n(gg') - Z^n(g) - V(g)Z^n(g')V^{(n)}(g^{-1}) = \sum_{q=2}^{n-1} Z^q(g)V^{(q)}(g) \sum_{\substack{l_1+\cdots+l_q=n\\ l_i\geq 1}} Z^{l_1}(g') \otimes\cdots\otimes Z^{l_q}(g')\sigma_n V^{(n)}(g^{-1}),$$

with

$$V^q(g) = V_0(g)^{\otimes q}, \quad q \text{ integer}$$

and

$$Z^n_{\mathbf{k}}(a,I) = Z^n_\omega(0,\Lambda_{\mathbf{k}}) = 0,$$

$$Z^n(o,u) = 0, \quad u\in\mathrm{SU}(1),$$

$$\frac{d}{dz}Z^n_\omega\left(0, \begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}\right)\bigg|_{z=0} = 0.$$

Proof: The result is proved as in Proposition 3, the only difference being the expression of $Z^n_\omega(0,h)$ now given by

$$Z^n_\omega(0,h) = \sum_{q=2}^{n-1}\int_0^z d\sigma\, Z^q_\omega\left(0,\begin{vmatrix}1 & \sigma\\ 0 & 1\end{vmatrix}\right)V^{(q)}\left(0,\begin{vmatrix}1 & \sigma\\ 0 & 1\end{vmatrix}\right) \times\sum_{p=0}^{q-1} I^{\otimes p}\otimes\frac{dZ^{n-q+1}}{dz}(0) \otimes I^{\otimes q-p-1}\sigma_n V^{(n)}\left(0,\begin{vmatrix}1 & \sigma\\ 0 & 1\end{vmatrix}\right)^{-1}$$

with

$$\frac{dZ^r}{dz}(0) = \frac{d}{dz}Z^r\left(0,\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}\right)\bigg|_{z=0}, \quad 2<r\leq n-1,$$

$$h = \epsilon\begin{vmatrix}1 & z\\ 0 & 1\end{vmatrix}.$$

V. CONCLUDING REMARKS

In conclusion, we want to stress some features of the previous construction. We have been able to build each term of the formal series (1.3) once the existence of a nontrivial cocycle has been proved. This has been obtained here using essentially the Abelian nature of the C^+ stabilizer group. Since the latter is no longer Abelian for Poincaré groups of higher dimensionality, a similar result cannot be extrapolated without further examination.

It may look striking that the nonlinearities of the representation concern only pure Lorentz transformations. In fact, this property is strictly dependent on the choice made for the representation space; it ensures that the extension cocycle restricted to space-time translations is a coboundary instead of a pseudocoboundary. This would be false if we had chosen, for instance, the space of Fourier transforms of functions in $\mathscr{D}(\mathbb{R}^2)$. Such a situation would deserve a separate study.

[1]M. Flato, G. Pinczon, and J. Simon, Ann. Sci. Ec. Norm. Sup. **10**, 405 (1977).

[2]E. Taflin, J. Math. Phys. **25**, 765 (1984).

[3]I. M. Gel'fand, M. I. Graev, and N. Ya. Vilenkin, *Generalized Functions* (Academic, New York, 1966), Vol. 5.

[4]G. Rideau, Rep. Math. Phys. **16**, 251 (1979).

[5]G. Rideau, Lett. Math. Phys. **8**, 421 (1984).

Path integration over compact and noncompact rotation groups[a)]

Manfred Böhm and Georg Junker
Physikalisches Institut der Universität Würzburg, Am Hubland, 8700 Würzburg, Federal Republic of Germany

(Received 31 December 1986; accepted for publication 6 May 1987)

Applications of group theoretical methods in the path integral formalism of nonrelativistic quantum theory are considered. Analysis of the symmetry of the Lagrangian leads to the expansion of the short time propagator in matrix elements of unitary irreducible representations of the symmetry group. Identification of the coordinates with the group parameters transforms the path integral to integrals over the group manifold. The integration is performed using the orthogonality of the representations. Compact and noncompact rotation groups are considered, where the corresponding path integral is embedded in Euclidean and pseudo-Euclidean spaces, respectively. The unit sphere and unit hyperboloid may either be viewed as the group manifold itself or at least as a group quotient. In the first case Fourier analysis leads to an expansion in group characters. In the second case an expansion in zonal spherical functions is obtained. As examples the groups SO(n), SU(2), SO($n-1$,1), and SU(1,1) are explicitly discussed. The path integral on SO($n+m$) and SO(n,m) in bispherical coordinates is also treated.

I. INTRODUCTION

In the year 1948 Feynman[1,2] had established the path integration formalism of quantum theory. In field theories the functional integration has been successfully applied in the last two decades. In nonrelativistic quantum theory, however, not much progress in solving exactly particular problems has been made up to 1979. Only quadratic Lagrangians, including a $1/r^2$ potential, could be integrated due to their Gaussian nature. The breakthrough was made in 1979 by Duru and Kleinert,[3] who solved the path integral (in phase space) of the hydrogen atom for the first time. In configuration space this problem has been treated explicitly by Ho and Inomata[4] (a critique of this work was made in the paper by Kleinert[5]). For later calculations see Ref. 6. This success has become possible by employing new techniques such as local time rescaling and dimensional extension. With these tricks the list of exactly soluble problems has increased rapidly. Common to all these is the fact that the dimensional extension has been used for the realization of the dynamical symmetry of the Lagrangian. For example, the dynamical symmetry of the Coulomb[3,4] and dyonium problem[5] has been utilized by the Kustaanheimo–Stiefel transformation being a nonlinear map from $\mathbb{R}^3$ into $\mathbb{R}^4$. Various problems having SU(2) as dynamical symmetry have become solvable by using similar methods. Examples are the Pöschl–Teller,[7,8] Rosen–Morse,[9,10] Hartmann,[11] and Hulthén potentials.[12] For noncompact groups only the SU(1,1) symmetry of the modified Pöschl–Teller potential[13] and the Kepler problem in a uniformly curved space[14] have been realized. Therefore the path integration on symmetry groups, especially on compact and noncompact rotation groups, is of great importance.

In the Schrödinger theory the solution of symmetric problems is usually simplified by choosing proper coordinates, e.g., spherical polar coordinates for spherically symmetric potentials. In the path integral formalism this transition is not that simple since for non-Cartesian coordinates additional quantum corrections of order $O(\hbar^2)$ do appear in the Lagrangian.[15,16] Indeed, the Feynman integral in the usual sliced-time basis[2,17] is only valid in Cartesian coordinates. The aim of the present paper is to derive a general procedure for the path integral treatment on compact and noncompact rotation groups. For this we have to embed the group manifold in Euclidean or pseudo-Euclidean spaces, respectively. We will proceed as follows.

In the next section we start with the definition of our notation. For this we have to recall some properties of transformation groups and their representations. Section III is devoted to the extension of the Feynman ansatz in pseudo-Euclidean space, in order to include the noncompact groups. This makes necessary a modification of the usual regularization scheme.[16] In Sec. IV we introduce generalized polar coordinates and develop two equivalent methods for performing the angular integration. The first one is the character expansion. In lattice gauge theories this technique, called cluster expansion, is used extensively.[18] Actually, the character expansion of Dosch and Müller,[19] where the cluster expansion of a SU(2) Yang–Mills gauge theory on a two-dimensional lattice is done, looks very similar to the expansion formula of Junker and Inomata,[10] where the path integral on the SU(2) manifold is expanded in SU(2) group characters. However, it will be shown that SU(2) and SU(1,1) are the only simple Lie groups where this technique is applicable in ordinary quantum mechanics. In looking for a method having a wider application we develop an expansion in zonal spherical functions. This technique does indeed work on all homogeneous spaces, which may be viewed as a group quotient G/H. In the last part of Sec. IV the connection between both expansions is shown. Finally we discuss an example for both methods for compact and noncompact groups. As compact groups we choose SO(n) and SU(2).

[a)] Dedicated to Hans Joos on the occasion of his 60th birthday.

For SU(2) we briefly review the path integral treatment in the Pöschl–Teller problem in order to perform the path integral on $SO(n+m)$ in bispherical coordinates using the group chain $SO(n+m)\supset SO(m)\times SO(n)$. As noncompact groups we will take $SO(n-1,1)$ and $SU(1,1)$. An application of the SU(1,1) propagator is made for the modified Pöschl–Teller potential leading to a path integral treatment on $SO(n,m)$ by using $SO(n,m)\supset SO(m)\times SO(n)$. In the Appendix we give the calculation of the Fourier coefficient for the SU(1,1) expansion which has been omitted by us in Ref. 13.

II. TRANSFORMATION GROUPS AND THEIR REPRESENTATIONS

In order to define our notation we repeat some properties of transformation groups and their representations.[20,21] A group G is called a transformation group of a space $\mathscr{M}$, if one may associate with each element $g\in G$ a transformation $\mathbf{x}\to g\mathbf{x}$ on $\mathscr{M}$. If there exists for any $\mathbf{x},\mathbf{y}\in\mathscr{M}$ an element g such that $g\mathbf{x}=\mathbf{y}$, then G is called a transitive transformation group and $\mathscr{M}$ a homogeneous space.

Let G be a transitive transformation group of $\mathscr{M}$. Furthermore let $\mathscr{L}$ be a linear vector space of functions $f(\mathbf{x})$, $\mathbf{x}\in\mathscr{M}$ such that

$$f(\mathbf{x})\in\mathscr{L}\Leftrightarrow f(g\mathbf{x})\in\mathscr{L}, \tag{2.1}$$

for any $g\in G$. With $f\in\mathscr{L}$ and $g\in G$ a representation of the group G is given by

$$D(g)f(\mathbf{x})=f(g^{-1}\mathbf{x}). \tag{2.2}$$

Choose $\mathscr{L}$ to be the Hilbert space of square integrable functions with respect to a group invariant measure $d\mu(\mathbf{x})$ on $\mathscr{M}$. Then the above representation is unitary relative to the scalar product

$$(f_1,f_2)=\int_{\mathscr{M}} f_1^*(\mathbf{x})f_2(\mathbf{x})d\mu(\mathbf{x}). \tag{2.3}$$

Such a representation is called a regular representation. For compact groups the regular representation is decomposable into a direct sum of unitary irreducible representations D^l of this group on $\mathscr{M}$. (A generalization for noncompact groups may be found in Chap. 5 of Ref. 21.) They form a complete basis in the Hilbert space.

Take $D^l(g)$ to be a unitary irreducible representation of G in the Hilbert space $\mathscr{L}$. Furthermore, let H be a subgroup of G which leaves the nonzero vector $a\in\mathscr{L}$ invariant, i.e.,

$$D^l(h)a=a,\quad h\in H\subset G. \tag{2.4}$$

Then $D^l(g)$ is called representation of class 1 relative to H. With each vector $f\in\mathscr{L}$ we may associate a scalar function

$$f^l(g)=(D^l(g)f,a). \tag{2.5}$$

Here $f^l(g)$ is called spherical function of the representation $D^l(g)$. Choosing a basis $\{b_i\}$ in $\mathscr{L}$ such that $b_0=a$, the matrix elements of $D^l(g)$ are given by

$$d^l_{nm}(g)=(D^l(g)b_m,b_n). \tag{2.6}$$

The $d^l_{0m}(g)$ are called associate spherical functions and the $d^l_{00}(g)$ are the zonal spherical functions. Obviously,

$$d^l_{m0}(gh)=d^l_{m0}(g),\quad d^l_{00}(h^{-1}gh)=d^l_{00}(g). \tag{2.7}$$

The spherical functions are eigenfunctions of the Laplace–Beltrami operator on the homogeneous space $\mathscr{M}=G/H$. The Hilbert space is spanned by a complete set $\{l\}$ of associate spherical functions.

Finally we give the general Fourier analysis on compact and noncompact groups[22]:

$$\begin{aligned} f(g)&=\sum_l d_l\sum_{m,n}\hat{f}_{mn}(l)d^l_{nm}(g),\\ \hat{f}_{mn}(l)&=\int_G f(g)d^l_{mn}(g^{-1})dg. \end{aligned} \tag{2.8}$$

The sum Σ_l is to be taken over the complete set $\{l\}$. For compact groups d_l is the dimension of the representation. However, we will call d_l the dimension also in the case of infinite-dimensional unitary representations of noncompact groups. In this case we may take

$$\int_G d^l_{mn}(g)d^{l'*}_{m'n'}(g)dg=\frac{\delta(l,l')}{d_l}\delta_{mm'}\delta_{nn'} \tag{2.9}$$

as a definition for d_l. In (2.9) $\delta(l,l')$ stands for $\delta(l-l')$ in the continuous and for $\delta_{ll'}$ in the discrete case, as noncompact groups in general contain both series. For the continuous series Σ_l is replaced by an integral in (2.8).

III. THE FEYNMAN PROPAGATOR IN PSEUDO-EUCLIDEAN SPACE

According to Feynman[1,2] the nonrelativistic quantum propagator $K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a)$ is given by a functional integral over the action,

$$\begin{aligned} K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a)&=\int_{\mathbf{r}_a=\mathbf{r}(t_a)}^{\mathbf{r}_b=\mathbf{r}(t_b)}\exp\left\{\frac{i}{\hbar}S[\mathbf{r}]\right\}\mathscr{D}\mathbf{r}(t)\\ &=\int_{\mathbf{r}_a=\mathbf{r}(t_a)}^{\mathbf{r}_b=\mathbf{r}(t_b)}\exp\left\{\frac{i}{\hbar}\int_{t_a}^{t_b}L\,dt\right\}\mathscr{D}\mathbf{r}(t). \end{aligned} \tag{3.1}$$

On the sliced-time basis the path integral in n dimensions is usually written as

$$\begin{aligned} K(\mathbf{r}_b,\mathbf{r}_a;\mathbf{t}_b-\mathbf{t}_a)&=\lim_{N\to\infty}\int\prod_{j=1}^N\exp\left\{\frac{i}{\hbar}S_j\right\}\\ &\quad\times\prod_{j=1}^N\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{n/2}\prod_{j=1}^{N-1}d^n\mathbf{r}_j, \end{aligned} \tag{3.2}$$

with the short time action

$$S_j=(m/2\epsilon)[(\Delta x_j^1)^2+\cdots+(\Delta x_j^n)^2]-V(\mathbf{r}_j)\epsilon. \tag{3.3}$$

For convenience we have chosen an equidistant time slicing $N\epsilon=t_b-t_a$, x^μ $(\mu=1,\ldots,n)$ are the Cartesian coordinates of $\mathbf{r}$ and $\Delta x_j^\mu=x_j^\mu-x_{j-1}^\mu$.

In many physically interesting problems the Lagrangian corresponding to (3.3) has a symmetry, which means that it is invariant under group transformations of the symmetry group. Therefore the Hamiltonian of the system may be expressed by Casimir invariants of the dynamical symmetry group. The wave functions correspond to unitary irreducible representations in the Hilbert space. This is the well-known procedure used by the algebraic method.[23]

However, the symmetry of the action may be also very useful in the path integral treatment. Expanding the phase $\exp\{(i/\hbar)S_j\}$ via the Fourier decomposition (2.8) in a series of unitary irreducible representations, the path integral may be performed (at least partially) using the orthogonality (2.9) of the matrix elements. The use of this group property in path integration has already been suggested in 1970 by Dowker.[24] The above path integral (3.2) is defined on a Euclidean space with metric $g_{\mu\nu} = \delta_{\mu\nu}$. As Feynman[1] has already mentioned a generalization to an indefinite metric

$$g_{\mu\nu} = \text{diag}\{\underbrace{+1,...,+1}_{p\text{ times}},\underbrace{-1,...,-1}_{q\text{ times}}\} \tag{3.4}$$

is possible. The pseudo-Euclidean space will be denoted by $E_{p,q}$. With metric (3.4) the short time action is given by

$$S_j = (m/2\epsilon)[(\Delta x_j^1)^2 + \cdots + (\Delta x_j^p)^2 - (\Delta x_j^{p+1})^2 - \cdots - (\Delta x_j^{p+q})^2] - V(\mathbf{r}_j)\epsilon. \tag{3.5}$$

In order to match the boundary condition

$$\lim_{t_b \to t_a} K(\mathbf{r}_b,\mathbf{r}_a;t_b - t_a) = \delta(\mathbf{r}_b - \mathbf{r}_a), \tag{3.6}$$

the measure has to be chosen in the following way:

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b - t_a) = \lim_{N\to\infty} \int \prod_{j=1}^{N} \exp\left\{\frac{i}{\hbar} S_j\right\} \times \prod_{j=1}^{N} \left(\frac{m}{2\pi i\hbar\epsilon}\right)^{p/2} \left(\frac{mi}{2\pi\hbar\epsilon}\right)^{q/2} \prod_{j=1}^{N-1} d^{p+q}\mathbf{r}_j. \tag{3.7}$$

For $p = n$ and $q = 0$ we recover the Euclidean propagator (3.2). The short time action (3.5) still remains invariant under some group transformation depending on $V(\mathbf{r})$. Therefore the above arguments are valid in the pseudo-Euclidean space, too. Here the symmetry group will be in general noncompact.

However, the above extension of the Feynman ansatz to $E_{p,q}$ requires some modification of the usual path integral formalism. First we have to regularize the path integral in the following way: *Integration over compact coordinates $x^1,...,x^p$ is regularized, as usual, by a mass having a small positive imaginary part, $m \to m + i\eta$ ($\eta > 0$), that over the noncompact ($x^{p+1},...,x^{p+q}$), however, by a small negative imaginary part of the mass, $m \to m - i\eta$.*

For $q = 0$, i.e., the Euclidean case, this reduces to the prescription of Langguth and Inomata.[16] Second, due to the topology of $E_{p,q}$ the scalar product

$$(\mathbf{r},\mathbf{r}) = (x^1)^2 + \cdots + (x^p)^2 - (x^{p+1})^2 - \cdots - (x^{p+q})^2 \tag{3.8}$$

can be positive, negative, or zero. Consequently, the space $E_{p,q}$ may be divided into three different subspaces T_α:

$$\begin{aligned} T_{+1} &= \{\mathbf{r}|(\mathbf{r},\mathbf{r}) > 0\}, \quad \text{timelike},\\ T_{-1} &= \{\mathbf{r}|(\mathbf{r},\mathbf{r}) < 0\}, \quad \text{spacelike},\\ T_0 &= \{\mathbf{r}|(\mathbf{r},\mathbf{r}) = 0\}, \quad \text{lightlike}. \end{aligned} \tag{3.9}$$

Integration over Cartesian coordinates in $E_{p,q}$ is similar to the usual one in E_n. We still have Gaussian integrals. Up to signs the methods of path integration in Cartesian coordinates on E_n may be applied here similarly.[2,17,25] The propagator has contributions from space-, time- and lightlike paths. And also from paths intersecting different regions T_α.

Systems that may evolve only along one kind of path are also physically interesting. For example, quantum mechanics on a space of constant negative curvature may be discussed in one region T_α of $E_{n-1,1}$ (see Ref. 26).

In this paper we will explore the path integral on such subspaces $T_\alpha \in E_{p,q}$. For this we introduce generalized polar coordinates r and θ^μ, $\mu = 1,...,p+q-1$. In general we have

$$x^\nu = re^\nu(\theta^1,...,\theta^{p+q-1}), \quad \nu = 1,...,p+q. \tag{3.10}$$

The functions e^ν define a unit vector in T_α,

$$\mathbf{e} = (e^1,...,e^{p+q}). \tag{3.11}$$

The set of all such vectors forms a hyperboloid $\mathscr{H}_\alpha \in T_\alpha$. We will call $\mathscr{H}_\alpha$ the unit sphere of T_α,

$$\mathscr{H}_\alpha = \{\mathbf{e}|(\mathbf{e},\mathbf{e}) = \alpha\}, \quad \alpha = 1, -1, 0. \tag{3.12}$$

To be more explicit one should also distinguish the nonconnected regions of T_α.

The short time action of the free system on $T_{\pm 1}$ reads in polar coordinates ($0 \leqslant r_j < \infty$, $\Delta r_j = r_j - r_{j-1}$)

$$S_j = \pm(m/2\epsilon)\Delta r_j^2 \pm (m/\epsilon) r_j r_{j-1}[1 \mp (\mathbf{e}_j,\mathbf{e}_{j-1})]. \tag{3.13}$$

In T_0 we have $S_j = -(\mathbf{e}_j,\mathbf{e}_{j-1}) m r_j r_{j-1}/\epsilon$. The corresponding path integral separates into an angular and radial part.

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b - t_a) = \lim_{N\to\infty} \int \prod_{j=1}^{N} \exp\left\{\frac{i}{\hbar} S_j\right\} \prod_{j=1}^{N} \left(\frac{m}{2\pi i\hbar\epsilon}\right)^{p/2} \times \left(\frac{mi}{2\pi\hbar\epsilon}\right)^{q/2} \prod_{j=1}^{N-1} r_j^{p+q-1}\, dr_j\, d^{p+q-1}\Omega_j. \tag{3.14}$$

IV. PATH INTEGRATION IN GENERALIZED POLAR COORDINATES

In this section we derive a general procedure for the angular path integration on $E_{p,q}$ using group theoretical methods.

Let G be a transformation group of $\mathscr{H}_\alpha$, i.e.,

$$\mathbf{e} = g\mathbf{a}. \tag{4.1}$$

The $n \times n$ matrix representation $g \in G$ ($n = p + q$) maps the fixed vector $\mathbf{a}$ into the vector $\mathbf{e}$, both being unit vectors on $\mathscr{H}_\alpha$. In (4.1) the vectors $\mathbf{e}$ and $\mathbf{a}$ have to be in the same subspace T_α. The unit sphere $\mathscr{H}_\alpha$ is covered by all possible rotations (4.1). For example, we have

$$\begin{aligned} &\mathbf{e}^2 = +1\colon\ \mathbf{a} = (+1,0,...,0) \quad \text{for } \mathscr{H}_{+1} \text{ with } x^1 > 0,\\ &\mathbf{e}^2 = +1\colon\ \mathbf{a} = (-1,0,...,0) \quad \text{for } \mathscr{H}_{+1} \text{ with } x^1 < 0,\\ &\mathbf{e}^2 = -1\colon\ \mathbf{a} = (0,...,0,+1) \quad \text{for } \mathscr{H}_{-1}, \text{ etc.} \end{aligned} \tag{4.2}$$

Note that to $\mathbf{a} \in \mathscr{H}_\alpha$ corresponds $a \in \mathscr{L}$ of Sec. II.

A possible choice of the group G is one that contains SO(p,q), $G \subseteq$ SO(p,q). However, other groups like SU(u,v) may do as well. For example, the unit sphere S^3 in the four-dimensional space E_4 is isomorphic to the group manifold of SU(2). Therefore instead of SO(4) we may choose SU(2) as a transformation group of $S^3 = SO(4)/SO(3)$.

In this paper we restrict ourselves to the cases where $\mathscr{H}_\alpha$ is isomorphic to the group manifold of G, $\mathscr{H}_\alpha \simeq G$, or $\mathscr{H}_\alpha$ is given by a group quotient G/H, $\mathscr{H}_\alpha = G/H$. Here $G = SO(p,q)$ and H is the stationary subgroup of **a**.

A. Expansion in group characters, $\mathscr{H}_\alpha \simeq G$

First we consider the special case $\mathscr{H}_\alpha \simeq G$, where the unit sphere is isomorphic to G. In order to find all rotation groups having this property we use the necessary condition $\dim \mathscr{H}_\alpha = \dim G$. From Table I it follows that SO(2), SO(1,1), SU(2), and SU(1,1) are the only candidates.

For the one-parameter groups SO(2) and SO(1,1) the irreducible representations are one-dimensional [nonunitary for SO(1,1)]. Obviously their characters and zonal spherical functions are identical and therefore these groups will be included in the general theory of the next section. Actually the expansions reduce to the Fourier and Laplace expansions, respectively.

Therefore we are left with the groups SU(2) and SU(1,1). First we consider the group SU(2) which is isomorphic to S^3. The infinitesimal generators are given by Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.3}$$

Defining

$$s^\mu = (i\sigma, \mathbf{1}), \quad \bar{s}^\mu = (-i\sigma, \mathbf{1}), \tag{4.4}$$

the isomorphism between points on the unit sphere $x \in S^3$ corresponding to unit vectors e^μ_x ($e^\mu_x e_{x\mu} = 1$) and the group elements $g \in SU(2)$ may be established by

$$g_x = e^\mu_x s_\mu = \begin{pmatrix} e^4_x + ie^3_x & ie^1_x + e^2_x \\ ie^1_x - e^2_x & e^4_x - ie^3_x \end{pmatrix}, \tag{4.5}$$

$$e^\mu_x = \tfrac{1}{2}\,\mathrm{Tr}(g_x \bar{s}^\mu). \tag{4.6}$$

Note that indeed $\det g_x = 1$, $g_x^\dagger g_x = \mathbf{1}$ and therefore $g \in SU(2)$. From Eq. (4.5) follows

$$\mathrm{Tr}(g_a^{-1} g_b) = 2\mathbf{e}_a \cdot \mathbf{e}_b. \tag{4.7}$$

The explicit identification of the coordinates will be given later.

The group manifold of SU(1,1) is isomorphic to the hyperboloid[13]

$$\begin{aligned} (\mathbf{e},\mathbf{e}) &= e^\mu e_\mu = -(e^1)^2 - (e^2)^2 + (e^3)^2 + (e^4)^2, \\ e^\mu &= (e^1, e^2, e^3, e^4), \quad e_\mu = (-e^1, -e^2, e^3, e^4). \end{aligned} \tag{4.8}$$

The infinitesimal generators may be given by

TABLE I. Solutions for $\dim G = \dim H_\alpha$.

G	$\dim G$	$\dim H_\alpha$	$\dim G = \dim H_\alpha$
$SO(p,q)$	$(p+q)(p+q-1)/2$	$p+q-1$	$p+q=2$
$SU(u,v)$	$(u+v)^2-1$	$2(u+v)-1$	$u+v=2$

$$\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.9}$$

The isomorphism can be established through

$$s_\mu = (i\sigma, \mathbf{1}), \quad \bar{s}_\mu = (-i\sigma^\dagger, \mathbf{1}), \tag{4.10}$$

using the scalar product (4.8) on $E_{2,2}$:

$$g_x = e^\mu_x s_\mu = \begin{pmatrix} e^4_x + ie^3_x & e^1_x - ie^2_x \\ e^1_x + ie^2_x & e^4_x - ie^3_x \end{pmatrix}, \tag{4.11}$$

$$e^\mu_x = \tfrac{1}{2}\,\mathrm{Tr}(g_x \bar{s}^\mu). \tag{4.12}$$

For g_x being an element of SU(1,1) it has to fulfill the following conditions. The pseudounitarity $g^{-1} = \sigma_3 g^\dagger \sigma_3$ is obviously true. But in order to get $\det g = +1$ we must have $(\mathbf{e},\mathbf{e}) = +1$. This means the hyperboloid $\mathscr{H}_{+1}$ has to be chosen. Again we find that the scalar product on $\mathscr{H}_{+1}$ may be written as a trace:

$$\mathrm{Tr}(g_a^{-1} g_b) = 2(\mathbf{e}_a, \mathbf{e}_b). \tag{4.13}$$

We conclude that for the case with $G \simeq \mathscr{H}_\alpha$, the corresponding short time propagator depends only on $\mathrm{Tr}(\hat{g}_j)$, $\hat{g}_j = g_{j-1}^{-1} g_j$ and is therefore invariant under group transformations $f(\hat{g}) \to f(g\hat{g}g^{-1})$. Such functions are called central functions and may be expanded in group characters.[20,21] The Fourier decomposition (2.8) simplifies to

$$\begin{aligned} f(g) &= \sum_l d_l\, \chi^{(l)}(g) \hat{f}(l), \\ \hat{f}(l) &= \frac{1}{d_l} \int_G f(g) \chi^{(l)*}(g)\, dg. \end{aligned} \tag{4.14}$$

Applying (4.14) to the short time propagator,

$$K(\mathbf{r}_j, \mathbf{r}_{j-1}; \epsilon) = \left(\frac{m}{2\pi i\hbar\epsilon}\right)^{p/2} \left(\frac{mi}{2\pi\hbar\epsilon}\right)^{q/2} \exp\left\{\frac{i}{\hbar} S_j\right\}, \tag{4.15}$$

leads to

$$K(\hat{g}_j; \epsilon) = \sum_l K_l(r_j, r_{j-1}; \epsilon) d_l\, \chi^{(l)}(\hat{g}_j). \tag{4.16}$$

The radial short time propagator $K_l(r_j, r_{j-1}; \epsilon)$ is determined by the Fourier coefficient $\hat{f}(l)$.

Using the group properties

$$\begin{aligned} \int_G \chi^{(l)}(g_{j-1}^{-1} g_j) \chi^{(l')}(g_j^{-1} g_{j+1})\, dg_j &= \frac{\delta(l,l')}{d_l} \chi^{(l)}(g_{j-1}^{-1} g_{j+1}), \end{aligned} \tag{4.17}$$

$$\chi^{(l)}(g_a^{-1} g_b) = \sum_{m,n} d^{\,l}_{mn}(g_b) d^{\,l*}_{mn}(g_a),$$

the angular integration can be performed.The $d^{\,l}_{mn}(g)$ are the unitary irreducible representations of G in the Hilbert space $\mathscr{L}$ being infinite dimensional for noncompact groups. Note that $d\Omega$ in Eq. (3.14) is given by $d\Omega = |\mathscr{H}_\alpha| dg$, where $|\mathscr{H}_\alpha|$ denotes the volume of $\mathscr{H}_\alpha$. The resulting propagator reads

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_{l,m,n} K_l(r_b,r_a;t_b-t_a) Y_{lmn}(\mathbf{e}_b) Y^*_{lmn}(\mathbf{e}_a), \quad (4.18)$$

with

$$Y_{lmn}(\mathbf{e}) = \sqrt{d_l}\, d^{\,l}_{mn}(g), \quad (4.19)$$

which we may call generalized harmonics. As for $G = SU(2)$ they lead to the monopole harmonics of Wu and Yang.[27] Here $K_l(r_b,r_a;t_b-t_a)$ is given by

$$K_l(r_b,r_a;t_b-t_a) = \lim_{N\to\infty} \int \prod_{j=1}^{N} K_l(r_j,r_{j-1};\epsilon) \prod_{j=1}^{N-1} r_j^{p+q-1}\, dr_j. \quad (4.20)$$

The expansions for the compact groups $SO(2) \simeq U(1) \simeq S^1$ and $SU(2) \simeq S^3$ have been discussed in detail by Junker and Inomata.[10] The expansion of the SU(1,1) propagator in $E_{2,2}$ with metric $(+1,+1,-1,-1)$ has been given by the authors.[13] A detailed discussion for SU(1,1) in $E_{2,2}$ with metric (4.8) follows in Sec. VI. The method of character expansion has also been used in the high temperature expansion of field theories on the lattice.[18]

Since as a homogeneous space $\mathscr{H}_\alpha$ usually may be viewed as a quotient G/H, we do need a general scheme for performing the path integration. Such a method may be found by using the expansion of the short time propagator in zonal spherical functions.

B. Expansion in zonal spherical functions, $\mathscr{H}_\alpha = G/H$

In this subsection we consider the case $\mathscr{H}_\alpha = G/H$, where the unit sphere is given by a group quotient. The subgroup $H \subset G$ is the little group of $\mathbf{a}$, i.e., $h\mathbf{a} = \mathbf{a}$, $h \in H$.

With (4.1) the scalar product in the short time action (3.13) may be written as

$$(\mathbf{e}_j,\mathbf{e}_{j-1}) = (g_j\mathbf{a}, g_{j-1}\mathbf{a}) = (g_{j-1}^{-1} g_j \mathbf{a}, \mathbf{a}). \quad (4.21)$$

The short time propagator (4.15) again depends on the group element

$$\hat{g}_j = g_{j-1}^{-1} g_j \quad (4.22)$$

and is invariant with respect to left and right transformations of the subgroup H:

$$K(h\hat{g}_j h^{-1};\epsilon) = K(\hat{g}_j;\epsilon), \quad h \in H. \quad (4.23)$$

Functions having this property may be expanded in zonal spherical functions of the representation of class 1 relative to H (see Ref. 20). The angles θ^μ can be identified with the group parameters of G which do not belong to the subgroup H. As by construction $\mathscr{H}_\alpha = G/H$, $\dim \mathscr{H}_\alpha = \dim G - \dim H$, this identification is always possible.

For functions having the property $f(hgh^{-1}) = f(g)$ the Fourier decomposition (2.8) simplifies to

$$f(g) = \sum_l d_l\, d^{\,l}_{00}(g) \hat{f}(l), \qquad \hat{f}(l) = \int_{\mathscr{H}_\alpha} f(g) d^{\,l*}_{00}(g) d\Gamma. \quad (4.24)$$

In (4.24) the integration over the subgroup H has already been performed using $dg = d\Gamma\, dh$. Here $d\Gamma$ and dh are the normalized measures of $\mathscr{H}_\alpha$ and H, respectively. Note that here $d\Omega$ is given by $|\mathscr{H}_\alpha| d\Gamma$.

Since the short time propagator (4.23) belongs to this class of functions the expansion yields

$$K(\hat{g}_j;\epsilon) = \sum_l K_l(r_j,r_{j-1};\epsilon) d_l\, d^{\,l}_{00}(\hat{g}_j), \quad (4.25)$$

where the radial short time propagator $K_l(r_j,r_{j-1};\epsilon)$ is again determined by the Fourier coefficient $\hat{f}(l)$.

Using the group properties

$$d^{\,l}_{00}(\hat{g}_j) = \sum_m d^{\,l}_{m0}(g_j) d^{\,l*}_{m0}(g_{j-1}), \qquad \int_{\mathscr{H}_\alpha} d^{\,l}_{m0}(g) d^{\,l'*}_{m'0}(g) d\Gamma = \frac{\delta_{mm'}}{d_l} \delta(l,l'), \quad (4.26)$$

the angular path integration can be performed. The result is

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_{l,m} K_l(r_b,r_a;t_b-t_a) Y_{lm}(\mathbf{e}_b) Y^*_{lm}(\mathbf{e}_a), \quad (4.27)$$

where

$$Y_{lm}(\mathbf{e}) = \sqrt{d_l}\, d^{\,l}_{m0}(g) \quad (4.28)$$

are the hyperspherical harmonics on $\mathscr{H}_\alpha$. Note $K_l(r_b,r_a;t_b-t_a)$ is the remaining radial path integral (4.20).

Since any unit sphere $\mathscr{H}_\alpha$ in $E_{p,q}$ can be viewed as a quotient G/H, the expansion in zonal spherical functions is a general method for performing the path integral on $\mathscr{H}_\alpha$. As examples we will discuss the cases $G = SO(n)$ and $SO(n-1,1)$ with $H = SO(n-1)$.

C. Equivalence of both methods

Above we have discussed two different methods for the path integration on homogeneous spaces. However, as the expansion in zonal spherical functions will always work by construction there arises the question of whether both methods are equivalent or not. In the following we will show that they are indeed identical in the cases where the character expansion does work.

According to Maurin (Ref. 21, p. 237ff), a character of a compact group H can be considered as a zonal spherical function on the group $G = H \times H$. The homogeneous space G/H may be identified with H. For Abelian groups G the characters are also zonal spherical functions of G. Here G need not be compact.

Restricting ourselves to simple Lie groups, the only isomorphism having the above structure is $D_2 \simeq A_1 \times A_1$ (Ref. 28). The following isomorphisms are obtained[23]:

$$SO(4) \simeq SU(2) \times SU(2)/Z_2, \qquad SO(2,2) \simeq SU(1,1) \times SU(1,1)/Z_2. \quad (4.29)$$

With $SU(2)/Z_2 \simeq SO(3)$ and $SU(1,1)/Z_2 \simeq SO(2,1)$ we identify the group manifolds of SU(2) and SU(1,1) with the quotients SO(4)/SO(3) and SO(2,2)/SO(2,1), respectively. Therefore the discussion of SU(2) and SU(1,1) in Sec. IV A contains all simple Lie groups where the expansion in group characters works. By Marinov and Terentyev[29] the fact that the path integral over the SU(n) manifold with

$n>2$ cannot be embedded into a flat space has already been noticed.

V. EXAMPLES FOR COMPACT GROUPS

A. Path integration on SO(n), $\mathscr{H}_\alpha = \mathrm{SO}(n)/\mathrm{SO}(n-1)$

The path integral over $S^{n-1} = \mathrm{SO}(n)/\mathrm{SO}(n-1)$ has been discussed by Marinov and Terentyev[30] for the first time, see also Refs. 10 and 31. However, up to now no explicit path integral treatment has been given.

Introducing spherical polar coordinates

$$\begin{aligned} x^1 &= r\sin\phi^{(n-1)}\cdots\sin\phi^{(1)}, & 0\leqslant r<\infty,\\ x^2 &= r\sin\phi^{(n-1)}\cdots\cos\phi^{(1)}, & 0\leqslant\phi^{(1)}<2\pi,\\ &\vdots & 0\leqslant\phi^{(k)}<\pi \quad (k\neq 1),\\ x^n &= r\cos\phi^{(n-1)}, & \end{aligned} \tag{5.1}$$

the Feynman ansatz on E_n reads

$$\begin{aligned} K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) &= \lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\}\\ &\times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{n/2}\prod_{j=1}^{N-1}r_j^{n-1}\,dr_j\,d^{\,n-1}\Omega_j, \end{aligned}\tag{5.2}$$

with

$$S_j = (m/2\epsilon)\Delta r_j^2 + (m/\epsilon)r_jr_{j-1}[1-\mathbf{e}_j\cdot\mathbf{e}_{j-1}], \tag{5.3}$$

$$\begin{aligned} d^{\,n-1}\Omega_j &= \sin^{n-2}\phi_j^{(n-1)}\cdots\sin^2\phi_j^{(3)}\\ &\times\sin\phi_j^{(2)}\,d\phi_j^{(n-1)}\cdots d\phi_j^{(1)}. \end{aligned}\tag{5.4}$$

In order to perform the expansion of the short time propagator in zonal spherical functions we have to recall some properties of the SO(n) representations.[20]

A $n\times n$ matrix representation may be given by a product of rotation matrices

$$\begin{aligned} g &= g^{n-1}\cdots g^k\cdots g^1,\\ g^k &= g_1(\theta_1^k)\cdots g_i(\theta_i^k)\cdots g_k(\theta_k^k), \end{aligned}\tag{5.5}$$

where $g_i(\theta_i^k)$ represents a rotation in the $(i,i+1)$ plane by an angle θ_i^k:

$$\begin{pmatrix} x'^i\\ x'^{i+1}\end{pmatrix} = \begin{pmatrix}\cos\theta_i^k & \sin\theta_i^k\\ -\sin\theta_i^k & \cos\theta_i^k\end{pmatrix}\begin{pmatrix} x^i\\ x^{i+1}\end{pmatrix}. \tag{5.6}$$

The $n(n-1)/2$ parameters θ_i^k are called Eulerian angles of the rotation g,

$$\begin{aligned} &0\leqslant\theta_i^k<\pi, \quad i=2,3,\ldots,k,\\ &0\leqslant\theta_1^k<2\pi, \quad k=1,2,\ldots,n-1 \end{aligned}\tag{5.7}$$

The associate invariant measure is

$$dg = \prod_{k=1}^{n-1}\left\{\frac{\Gamma((k+1)/2)}{2\pi^{[(k+1)/2]}}\prod_{i=1}^{k}\sin^{i-1}\theta_i^k\,d\theta_i^k\right\}. \tag{5.8}$$

Choosing $\mathbf{a} = (0,\ldots,0,1)$ as the stationary vector, each point $\mathbf{e}$ on S^{n-1} may be obtained by a rotation $\mathbf{e} = g\mathbf{a}$. The parameters $\theta_1^{n-1},\ldots,\theta_{n-1}^{n-1}$ of g are identical with the polar coordinates $\phi^{(1)},\ldots,\phi^{(n-1)}$ of $\mathbf{e}$. The stationary subgroup $H=\mathrm{SO}(n-1)$ of $\mathbf{a}$ is given by the elements $h=g^k$ $(k\neq n-1)$. Integrating (5.8) over all parameters of H yields the normalized volume element on S^{n-1}:

$$d\Gamma = [\Gamma(n/2)/2\pi^{n/2}]\,d\Omega. \tag{5.9}$$

The dimension of the unitary irreducible representation $D^l(g)$ in the Hilbert space is

$$\begin{aligned} d_l &= (2l+n-2)[(l+n-3)!/l!(n-2)!],\\ l &= 0,1,2,\ldots. \end{aligned}\tag{5.10}$$

The zonal spherical functions depend only on the parameter θ_{n-1}^{n-1} and are given by Gegenbauer polynomials,

$$\begin{aligned} d_{00}^{\,l}(g) &= [(n-3)!l!/(l+n-3)!]\\ &\times C_l^{(n-2)/2}(\cos\theta_{n-1}^{n-1}). \end{aligned}\tag{5.11}$$

Note, that $\Theta\equiv\theta_{n-1}^{n-1}$ is the angle between the vectors $\mathbf{a}$ and $\mathbf{e}$, i.e., $\mathbf{a}\cdot\mathbf{e} = \cos\Theta$.

The associate zonal spherical functions are denoted by $d_{M0}^{\,l}(g)$, where M stands for the $(n-2)$-tuple

$$\begin{aligned} M &= (m_1,m_2,\ldots,m_{n-2}),\\ l &\equiv m_0\geqslant m_1\geqslant\cdots\geqslant m_{n-3}\geqslant|m_{n-2}|. \end{aligned}\tag{5.12}$$

An explicit expression is given by Vilenkin,[20] see also Eq. (5.25).

Now we are well prepared for the expansion of the short time propagator. According to the general theory of Sec. IV, the action may be written as

$$S_j = (m/2\epsilon)\Delta r_j^2 + (m/\epsilon)r_jr_{j-1}[1-\hat{g}_j\mathbf{a}\cdot\mathbf{a}], \tag{5.13}$$

and depends only on the parameter $\Theta\equiv\theta_{n-1}^{n-1}$ of $\hat{g}_j$. Actually we have $\hat{g}_j\mathbf{a}\cdot\mathbf{a} = \cos\Theta$. For the Fourier analysis only the factor $\exp(iz\cos\Theta)$, where $z = -mr_jr_{j-1}/\epsilon\hbar$, has to be considered. We have

$$\exp\{iz\cos\Theta\} = \sum_{l=0}^{\infty} d_l\,d_{00}^{\,l}(\hat{g}_j)\hat{f}(l), \tag{5.14}$$

$$\hat{f}(l) = \int_{S^{n-1}} e^{iz\cos\Theta}d_{00}^{\,l*}(\hat{g}_j)\,d\Gamma. \tag{5.15}$$

The integral can be simplified to

$$\begin{aligned} \hat{f}(l) &= \frac{\Gamma(n/2)\Gamma(n-2)l!}{2\sqrt{\pi}\,\Gamma((n-1)/2)\Gamma(n+l-2)}\\ &\times\int_0^{\pi} e^{iz\cos\Theta}C_l^{(n-2)/2}(\cos\Theta)\sin^{n-2}\Theta\,d\Theta \end{aligned}\tag{5.16}$$

and yields (p. 221 in Ref. 32),

$$\hat{f}(l) = \Gamma(n/2)(2/z)^{(n-2)/2}i^lJ_{l+(n-2)/2}(z). \tag{5.17}$$

Replacing the Bessel function $J_\nu(z)$ by the modified one $I_\nu(iz)$ leads to the well-known Gegenbauer formula $[\nu=(n-2)/2]$:

$$e^{iz\cos\Theta} = \left(\frac{2}{iz}\right)^{\nu}\Gamma(\nu)\sum_{l=0}^{\infty}(l+\nu)C_l^{\nu}(\cos\Theta)I_{l+\nu}(iz). \tag{5.18}$$

This formula has been used earlier for the path integration in polar coordinates.[10,30,31]

The short time propagator now reads

$$K(\hat{g}_j,\epsilon) = \frac{\Gamma(n/2)}{2\pi^{n/2}}\sum_{l=0}^{\infty} d_l\,d_{00}^{\,l}(\hat{g}_j)K_l(r_j,r_{j-1};\epsilon). \tag{5.19}$$

The angular path integral can be performed and we find

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \frac{\Gamma(n/2)}{2\pi^{n/2}}\sum_{l=0}^{\infty} d_l\, d^{\,l}_{00}(g_a^{-1}g_b)K_l(r_a,r_b;t_b-t_a). \tag{5.20}$$

The radial propagator is given by

$$K_l(r_a,r_b;t_b-t_a) = (r_b r_a)^{(1-n)/2}\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j^l\right\} \times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}dr_j,$$

$$S_j^l = \frac{m}{2\epsilon}\Delta r_j^2 - \left[\left(l+\frac{n-2}{2}\right)^2-\frac{1}{4}\right]\frac{\hbar^2\epsilon}{2mr_jr_{j-1}}, \tag{5.21}$$

where we have made use of the asymptotic formula[16]

$$I_\nu(iz) = (2\pi iz)^{-1/2}\exp\left\{iz+i\frac{\nu^2-\frac{1}{4}}{2z}+O\left(\frac{1}{z^2}\right)\right\},$$

$$\mathrm{Im}\, z<0 \Leftrightarrow \mathrm{Im}\, m>0. \tag{5.22}$$

The propagator (5.20) may be also expressed in terms of hyperspherical harmonics [see Eq. (4.27)]

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_{l=0}^{\infty}K_l(r_a,r_b;t_b-t_a)\frac{\Gamma(n/2)}{2\pi^{n/2}} \times\sum_M Y_{lM}(\mathbf{e}_b)Y^*_{lM}(\mathbf{e}_a), \tag{5.23}$$

where

$$\sum_M \equiv \sum_{m_1=0}^{m_0}\sum_{m_2=0}^{m_1}\cdots\sum_{m_{n-3}=0}^{m_{n-4}}\sum_{m_{n-2}=-m_{n-3}}^{m_{n-3}}. \tag{5.24}$$

The $Y_{lM}(\mathbf{e})$ are given explicitly by[20]

$$Y_{lM}(\mathbf{e}) = A^l_M\prod_{k=0}^{n-3}\{C^{m_{k+1}+(n-k-2)/2}_{m_k-m_{k+1}}(\cos\phi^{(n-k-1)})\sin^{m_{k+1}}\phi^{(n-k-1)}\}\exp(im_{n-2}\phi^{(1)}),$$

$$(A^l_M)^2 = \frac{1}{\Gamma(n/2)}\prod_{k=0}^{n-3}\left\{\frac{2^{2m_{k+1}+n-k-4}(m_k-m_{k+1})!}{\sqrt{\pi}\,\Gamma(m_{k+1}+m_k+n-k-2)}(n-k-2+2m_k) \times[\Gamma(m_{k+1}+(n-k-2)/2)]^2\right\}. \tag{5.25}$$

They form a complete set on S^{n-1}:

$$\int_{S^{n-1}} Y_{lM}(\mathbf{e})Y^*_{l'M'}(\mathbf{e})d\Gamma = \delta_{ll'}\delta_{MM'}. \tag{5.26}$$

The result (5.23) is identical with that obtained earlier.[10,15,31,33] The SO(n) propagator has already been proved useful in the path integration of the n-dimensional harmonic oscillator and the singular potential $V(\mathbf{r}) = -\alpha/r$ (see Refs. 8 and 31) which is sometimes erroneously called the n-dimensional Coulomb problem.

B. Path integration over the SU(2) manifold, $\mathscr{H}_\alpha = S^3$

The path integration over the SU(2) manifold has recently attracted much attention in the Feynman quantization of various problems having SU(2) as dynamical symmetry. Examples are the nonsymmetric Rosen–Morse,[10] Pöschl–Teller,[7] Hartmann,[11] and Hulthén potentials.[12] Even for the dyonium problem[6] the expansion of the Feynman ansatz in SU(2) matrix elements has been proved useful. In this section we would like to show how this SU(2) expansion, derived by Junker and Inomata,[10] can be incorporated into the general scheme of Sec. IV.

The spinor representation of SU(2) is usually parametrized in Eulerian angles[20]:

$$g(\phi,\theta,\psi) = \begin{pmatrix} e^{i\varphi/2} & 0\\ 0 & e^{-i\varphi/2}\end{pmatrix} \times\begin{pmatrix}\cos\theta/2 & \sin\theta/2\\ -\sin\theta/2 & \cos\theta/2\end{pmatrix}\begin{pmatrix} e^{i\psi/2} & 0\\ 0 & e^{-i\psi/2}\end{pmatrix},$$

$$0\leqslant\varphi<2\pi,\quad 0<\theta<\pi,\quad 0\leqslant\psi<4\pi. \tag{5.27}$$

The matrix elements of the $(2J+1)$-dimensional unitary irreducible representation in the Hilbert space are the well-known Wigner functions,

$$d^{\,J}_{mn}(g) = e^{-im\varphi}d^{\,J}_{mn}(\theta)e^{-in\psi},$$

$$J=0,\ \tfrac{1}{2},1,\tfrac{3}{2},\ldots,\quad -J\leqslant m,n\leqslant J. \tag{5.28}$$

The characters are

$$\chi^{(J)}(g) = \sum_{m=-J}^{J}d^{\,J}_{mm}(g) = \frac{\sin(2J+1)\Theta/2}{\sin\Theta/2}, \tag{5.29}$$

where

$$\cos\left(\frac{\Theta}{2}\right) = \cos\left(\frac{\theta}{2}\right)\cos\left(\frac{(\varphi+\psi)}{2}\right). \tag{5.30}$$

The invariant volume element follows to be

$$dg = (1/16\pi^2)\sin\theta\, d\theta\, d\varphi\, d\psi. \tag{5.31}$$

Comparing (5.27) with the spinor representation (4.5) of Sec. IV suggests the following parametrization of E_4:

$$\begin{aligned} x^1 &= r\sin(\theta/2)\sin((\varphi-\psi)/2),\\ x^2 &= r\sin(\theta/2)\cos((\varphi-\psi)/2),\\ x^3 &= r\cos(\theta/2)\sin((\varphi+\psi)/2),\\ x^4 &= r\cos(\theta/2)\cos((\varphi+\psi)/2), \end{aligned}\qquad \begin{aligned} &0\leqslant\varphi<2\pi,\\ &0\leqslant\theta<\pi,\\ &0\leqslant\psi<4\pi. \end{aligned} \tag{5.32}$$

The corresponding Feynman ansatz reads

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\} \times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^2\prod_{j=1}^{N-1}r_j^3\,dr_j\,2\pi^2\,dg_j,$$

$$S_j = \frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_jr_{j-1}\left[1-\frac{1}{2}\mathrm{Tr}(\hat{g}_j)\right], \tag{5.33}$$

where we have made use of (4.7).

The expansion of the factor $\exp\{z\,\mathrm{Tr}(g)\}$ in SU(2) characters has already been investigated in lattice gauge theories[19]:

$$\exp\{z\,\mathrm{Tr}(g)\} = \sum_J (2J+1)\frac{1}{z} I_{2J+1}(2z)\chi^{(J)}(g). \tag{5.34}$$

For $2z = mr_j r_{j-1}/i\epsilon\hbar$ we find using the asymptotic formula (5.22) for small ϵ

$$\exp\left\{\frac{imr_j r_{j-1}}{\epsilon\hbar}\left[1-\frac{1}{2}\mathrm{Tr}(\hat{g}_j)\right]\right\} \approx \frac{1}{2\pi^2}\left(\frac{2\pi i\hbar\epsilon}{mr_j r_{j-1}}\right)^{3/2}\sum_J (2J+1) \times\exp\left\{-\frac{i}{\hbar}\left[J(J+1)+\frac{3}{16}\right]\frac{2\hbar^2\epsilon}{mr_j r_{j-1}}\right\}\chi^{(J)}(\hat{g}_j). \tag{5.35}$$

This is the expansion derived by Junker and Inomata.[10] A similar formula has been given by Duru.[7] It contains only integer angular momenta J and therefore does not yield the complete SU(2) propagator.

Performing the angular integration leads to

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_J \frac{2J+1}{2\pi^2} K_J(r_b,r_a;t_b-t_a)\chi^{(J)}(g_a^{-1}g_b), \tag{5.36}$$

where

$$K_J(r_b,r_a;t_b-t_a) = (r_b r_a)^{-3/2}\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j^J\right\} \times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}dr_j, \tag{5.37}$$

$$S_j^J = \frac{m}{2\epsilon}\Delta r_j^2 - \left[J(J+1)+\frac{3}{16}\right]\frac{2\hbar\epsilon}{mr_j r_{j-1}}.$$

The result is identical with (5.20) for $n=4$, as expected.

As already mentioned, the above expansion has been used for various problems having SU(2) symmetry. As an instructive example we may take the one-dimensional Pöschl–Teller potential[34]

$$V(x) = \frac{\hbar^2 a^2}{2m}\left(\frac{\kappa^2-\frac{1}{4}}{\sin^2 ax}+\frac{\lambda^2-\frac{1}{4}}{\cos^2 ax}\right),\quad 0<x<\frac{\pi}{2a}. \tag{5.38}$$

A detailed discussion may be found in Refs. 8, 10, and 35. Here we just state that for $\theta = 2ax$, the Feynman ansatz reads

$$K(x_b,x_a;t_b-t_a) = a\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\} \times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar a^2\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}\frac{1}{2}d\theta_j, \tag{5.39}$$

with

$$S_j = \frac{m}{a^2\epsilon}\left(1-\cos\frac{\Delta\theta_j}{2}\right) - \left[\frac{\kappa^2-\frac{1}{4}}{\sin(\theta_j/2)\sin(\theta_{j-1}/2)} + \frac{\lambda^2-\frac{1}{4}}{\cos(\theta_j/2)\cos(\theta_{j-1}/2)}+\frac{1}{4}\right]\frac{a^2\hbar^2}{2m}\epsilon. \tag{5.40}$$

For $\kappa, \lambda\in N$ the one-dimensional path integral can be transformed into that of the SU(2) propagator[9,10]:

$$K(x_b,x_a;t_b-t_a) = \frac{a}{4}(\sin\theta_b\sin\theta_a)^{1/2}\exp\left\{-\frac{i\hbar a^2}{8m}(t_b-t_a)\right\} \times\int_0^{2\pi}\int_0^{4\pi} Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a) \times\exp\left\{i\left(\frac{\lambda+\kappa}{2}\varphi_b+\frac{\lambda-\kappa}{2}\psi_b\right)\right\}d\psi_b\,d\varphi_b, \tag{5.41}$$

where $Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a)$ is indeed a path integral over SU(2),

$$Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a) = \lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}\tilde{S}_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar a^2\epsilon}\right)^{3/2} \times\prod_{j=1}^{N-1}\frac{1}{8}\sin\theta_j\,d\theta_j\,d\varphi_j\,\psi_j, \tag{5.42}$$

$$\tilde{S}_j = (m/a^2\epsilon)\left[1-\tfrac{1}{2}\mathrm{Tr}(\hat{g}_j)\right]. \tag{5.43}$$

The integration can now be performed and yields

$$K(x_b,x_a;t_b-t_a) = a(\sin\theta_b\sin\theta_a)^{1/2}\sum_{J=(\kappa+\lambda)/2}^{\infty}(2J+1) \times d^J_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_b)\,d^{J*}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_a) \times\exp\left\{-\frac{i}{\hbar}(2J+1)^2\frac{\hbar^2a^2}{2m}(t_b-t_a)\right\}. \tag{5.44}$$

Here J is either an integer or a half-integer depending on $(\lambda+\kappa)/2$. Shifting the summation index yields the standard form [in terms of Jacobi polynomials $P_n^{(\alpha,\beta)}(z)$]

$$K(x_b,x_a;t_b-t_a) = \sum_{n=0}^{\infty}\exp\left\{-\frac{i}{\hbar}E_n(t_b-t_a)\right\}\Psi_n(x_b)\Psi_n^*(x_a), \tag{5.45}$$

where

$$E_n = (2n+\kappa+\lambda+1)^2(\hbar^2a^2/2m), \tag{5.46}$$

$$\Psi_n(x) = \left[\frac{2a(2n+\kappa+\lambda+1)n!(n+\kappa+\lambda)!}{(n+\kappa)!(n+\lambda)!}\right]^{1/2} \times\sin^{\kappa+1/2}ax\,\cos^{\lambda+1/2}ax\,P_n^{(\kappa,\lambda)}(1-2\sin^2 ax). \tag{5.47}$$

Path integration in bispherical coordinates: The above solution of the Pöschl–Teller problem now enables us to perform the path integral in bispherical coordinates

$$
\begin{aligned}
&x^1 = r\sin(\theta/2)\sin\alpha^{(n-1)}\cdots\sin\alpha^{(1)}, \\
&\vdots \\
&x^n = r\sin(\theta/2)\cos\alpha^{(n-1)}, \\
&x^{n+1} = r\cos(\theta/2)\sin\beta^{(m-1)}\cdots\sin\beta^{(1)}, \\
&\vdots \\
&x^{n+m} = r\cos(\theta/2)\cos\beta^{(m-1)},
\end{aligned}
\qquad
\begin{aligned}
&0\leqslant r<\infty, \\
&0\leqslant\alpha^{(1)},\beta^{(1)}<2\pi, \\
&0\leqslant\alpha^{(i)},\beta^{(i)},\theta<\pi(i\neq 1).
\end{aligned}
\tag{5.48}
$$

The Jacobian is

$$
d^{m+n}\mathbf{r} = r^{m+n-1}\,dr\,\frac{1}{2}\sin^{n-1}\frac{\theta}{2}\cos^{m-1}\frac{\theta}{2}\,d\theta\, d^{n-1}\Omega(\alpha)d^{m-1}\Omega(\beta), \tag{5.49}
$$

where $d\Omega$ is given similar to Eq. (5.4).

The propagator of the free system in E_{m+n} is

$$
K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{(m+n)/2}\prod_{j=1}^{N-1}d^{m+n}\mathbf{r}_j, \tag{5.50}
$$

$$
\begin{aligned}
S_j = {}& \frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_j r_{j-1}\left(1-\cos\frac{\Delta\theta_j}{2}\right) \\
&+\frac{m}{\epsilon}r_j r_{j-1}\sin\frac{\theta_j}{2}\sin\frac{\theta_{j-1}}{2}(1-\mathbf{e}^{\alpha}_{j-1}\cdot\mathbf{e}^{\alpha}_j) + \frac{m}{\epsilon}r_j r_{j-1}\cos\frac{\theta_j}{2}\cos\frac{\theta_{j-1}}{2}(1-\mathbf{e}^{\beta}_{j-1}\cdot\mathbf{e}^{\beta}_j),
\end{aligned}
\tag{5.51}
$$

where $\mathbf{e}^{\alpha}$ and $\mathbf{e}^{\beta}$ are the unit vectors in the subspaces E_n and E_m, respectively.

Guided by the group chain $SO(m+n)\supset SO(m)\times SO(n)$ the integration over the angles $\alpha^{(i)}$ and $\beta^{(i)}$ can be performed analogously to Sec. V A. We find

$$
K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_{l,\lambda=0}^{\infty}K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a)\frac{\Gamma(n/2)}{2\pi^{n/2}}\sum_N Y_{lN}\mathbf{e}^{\alpha}_b\,Y^*_{lN}(\mathbf{e}^{\alpha}_a)\frac{\Gamma(m/2)}{2\pi^{m/2}}\sum_M Y_{\lambda M}(\mathbf{e}^{\beta}_b)Y^*_{\lambda M}(\mathbf{e}^{\beta}_a), \tag{5.52}
$$

where

$$
\begin{aligned}
K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a) = {}& \left(r_b r_a\sin\frac{\theta_b}{2}\sin\frac{\theta_a}{2}\right)^{(1-n)/2}\left(r_b r_a\cos\frac{\theta_b}{2}\cos\frac{\theta_a}{2}\right)^{(1-m)/2} \\
&\times\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}\tilde{S}_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)\prod_{j=1}^{N-1}r_j\,dr_j\,\frac{1}{2}\,d\theta_j,
\end{aligned}
\tag{5.53}
$$

$$
\tilde{S}_j = \frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_j r_{j-1}\left(1-\cos\frac{\Delta\theta_j}{2}\right) - \left[\frac{\mu^2-\frac{1}{4}}{\sin(\theta_j/2)\sin(\theta_{j-1}/2)} + \frac{\nu^2-\frac{1}{4}}{\cos(\theta_j/2)\cos(\theta_{j-1}/2)}\right]\frac{\hbar^2\epsilon}{2mr_j r_{j-1}}, \tag{5.54}
$$

with $\mu = l+(n-2)/2$ and $\nu=\lambda+(m-2)/2$. The θ integral is now formal identical with the Pöschl–Teller problem leading to

$$
\begin{aligned}
&K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a) \\
&\quad = 2\left(r_b r_a\sin\frac{\theta_b}{2}\sin\frac{\theta_a}{2}\right)^{(2-n)/2} \\
&\quad\times\left(r_b r_a\cos\frac{\theta_b}{2}\cos\frac{\theta_a}{2}\right)^{(2-m)/2} \\
&\quad\times\sum_{J=(\mu+\nu)/2}(2J+1) \\
&\quad\times d^J_{(\mu+\nu)/2,(\nu-\mu)/2}(\theta_b)d^{J*}_{(\mu+\nu)/2,(\nu-\mu)/2}(\theta_b) \\
&\quad\times K_J(r_b,r_a;t_b-t_a),
\end{aligned}
\tag{5.55}
$$

where $K_J(r_b,r_a;t_b-t_a)$ is given by Eq. (5.37). Note that (5.53) can only be transformed into a SU(2) integral for even m and n.

VI. EXAMPLES FOR NONCOMPACT GROUPS

Up to now we have dealt only with compact groups, where the final results were already known by other methods. However, the general theory of Secs. III and IV was formulated in such a way that noncompact groups can also be treated. Here we will choose as examples the n-dimensional Lorentz group $SO(n-1,1)$ and $SU(1,1)$. These noncompact groups are often used for scattering problems in quantum theory. Both can be viewed as analytical continuations of $SO(n)$ and $SU(2)$, respectively. Therefore we will keep close to the calculation of the previous section.

A. Path integration on $G=SO(n-1,1)$, $\mathscr{H}_\alpha \subset SO(n-1,1)/SO(n)$

As already mentioned, the space $E_{n-1,1}$ having the metric

$$(\mathbf{r},\mathbf{r}) = (x^1)^2 + \cdots + (x^{n-1})^2 - (x^n)^2 \tag{6.1}$$

consists of topologically different subspaces. Here we will perform the path integral on $\mathscr{H}_{-1} \in T_{-1} = \{\mathbf{r} | (\mathbf{r},\mathbf{r}) < 0\}$, the spacelike subspace. The polar coordinates on T_{-1} may be introduced via

$$\begin{aligned} x^1 &= r\sinh\phi^{(n-1)}\sin\phi^{(n-2)}\cdots\sin\phi^{(1)}, \\ x^2 &= r\sinh\phi^{(n-1)}\sin\phi^{(n-2)}\cdots\cos\phi^{(1)}, \\ &\vdots \\ x^n &= r\cosh\phi^{(n-1)}, \end{aligned} \qquad \begin{aligned} &0\leqslant r, \quad \phi^{(n-1)} < \infty, \\ &0\leqslant \phi^{(1)} < 2\pi, \\ &0\leqslant \phi^{(k)} < \pi \quad (k \neq 1, n-1). \end{aligned} \tag{6.2}$$

The Feynman ansatz reads

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \lim_{N\to\infty}\int \prod_{j=1}^{N} \exp\left\{\frac{i}{\hbar}S_j\right\} \times \prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{(n-1)/2} \times \left(\frac{mi}{2\pi\hbar\epsilon}\right)^{1/2} \prod_{j=1}^{N-1} d^n\mathbf{r}_j, \tag{6.3}$$

with

$$S_j = -(m/2\epsilon)\Delta r_j^2 - (m/\epsilon) r_j r_{j-1}[1 + (\mathbf{e}_j,\mathbf{e}_{j-1})], \tag{6.4}$$

$$d^{n-1}\Omega = \sinh^{n-2}\phi^{(n-1)}\sin^{n-3}\phi^{(n-2)}\cdots\sin\phi^{(2)} \times d\phi^{(n-1)}\cdots d\phi^{(1)}. \tag{6.5}$$

Before proceeding we have to recall some properties of the $SO(n-1,1)$ representations.[20]

The $n\times n$ matrix representation may be given by products of hyperbolic and ordinary rotations:

$$g = g^{(n-1)}h, \tag{6.6}$$

where h is a $n\times n$ representation of the maximal compact subgroup $H = SO(n-1)$ given by Eq. (5.5), and

$$g^{(n-1)} = g_1(\theta_1^{n-1})\cdots g_k(\theta_k^{n-1})\cdots g_{n-1}(\theta_{n-1}^{n-1}), \tag{6.7}$$

where $g_k(\theta_k^{n-1})$ $(k\neq n-1)$, is a rotation in the $(k,k+1)$ plane [see Eq. (5.6)]. Here $g_{n-1}(\theta_{n-1}^{n-1})$ is the Lorentz transformation

$$\begin{pmatrix} x'^{n-1} \\ x'^n \end{pmatrix} = \begin{pmatrix} \cosh\theta_{n-1}^{n-1} & \sinh\theta_{n-1}^{n-1} \\ \sinh\theta_{n-1}^{n-1} & \cosh\theta_{n-1}^{n-1} \end{pmatrix} \begin{pmatrix} x^{n-1} \\ x^n \end{pmatrix}. \tag{6.8}$$

The parameter θ_{n-1}^{n-1} is in the interval $0\leqslant\theta_{n-1}^{n-1} < \infty$ and all others are limited analogously to Eq. (5.7). The invariant volume element may be obtained by analytical continuation of (5.8):

$$dg = \frac{\Gamma(n/2)}{2\pi^{n/2}}\sinh^{n-2}\theta_{n-1}^{n-1}\sin^{n-3}\theta_{n-2}^{n-1}\cdots\sin\theta_2^{n-1} \times d\theta_{n-1}^{n-1}\cdots d\theta_1^{n-1}\,dh, \tag{6.9}$$

dh is the corresponding measure of $h\in SO(n-1)$.

Taking the northpole $\mathbf{a} = (0,\ldots,0,+1)$ as stationary vector, each $\mathbf{e}$ on the spacelike hyperboloid $\mathscr{H}_{-1}$ may be obtained by the transformation $\mathbf{e} = g\mathbf{a}$. The polar coordinates $\phi^{(1)},\ldots,\phi^{(n-1)}$ of $\mathbf{e}$ are given by the parameters $\theta_1^{n-1},\ldots,\theta_{n-1}^{n-1}$ of $g^{(n-1)}$. The little group is $SO(n-1)$.

Here $\mathscr{H}_{-1}$ is also called the $(n-1)$-dimensional Lobachevsky space, denoted by Λ^{n-1} (see Ref. 20). Group theoretically we have

$$\Lambda^{n-1} \subset SO(n-1,1)/SO(n-1), \tag{6.10}$$

where Λ^{n-1} is a model of a space of constant negative curvature, similar to the way S^{n-1} represents a space of constant positive curvature. Quantum mechanics on spaces with negative curvature is of interest.[26] For example quantum chaos is recently studied on such topologies.[36]

The normalized volume element on $\mathscr{H}_{-1}$, in the sense of $\int_{\mathscr{H}_{-1}} f(\mathbf{r})\delta(\mathbf{r})d\Gamma = f(0)$, is

$$d\Gamma = [\Gamma(n/2)/2\pi^{n/2}]d^{n-1}\Omega. \tag{6.11}$$

The unitary irreducible representations D^l in the Hilbert space are continuous,[20]

$$\text{fundamental series: } l = -(n-2)/2 + i\rho, \quad -\infty < \rho < +\infty, \tag{6.12}$$

complementary series: $-n+2 < l < 0$.

The zonal spherical functions depend only on the parameter $\Theta \equiv \theta_{n-1}^{n-1}$, $(\mathbf{e},\mathbf{a}) = -\cosh\Theta$ (see Ref. 20):

$$d_{00}^{l}(g) = 2^{(n-3)/2}\frac{\Gamma((n-1)/2)}{\sinh^{(n-3)/2}\Theta}P_{l+(n-3)/2}^{(3-n)/2}(\cosh\Theta). \tag{6.13}$$

Expressing the Legendre function $P_l^m(z)$ in terms of Gegenbauer functions shows the analytical continuation of (5.11) explicitly:

$$d_{00}^{l}(g) = \frac{(n-3)!\Gamma(l+1)}{\Gamma(l+n-2)}C_l^{(n-2)/2}(\cosh\Theta) \tag{6.14}$$

The associate spherical functions $d_{K0}^{l}(g)$ may be written as a product due to Eq. (6.6):

$$d_{K0}^{l} = d_{K'0}^{l}(g^{(n-1)})d_{M0}^{k}(h), \tag{6.15}$$

with

$$\begin{aligned} K &= (k,m_1,\ldots,m_{n-3}), \\ M &= (m_1,\ldots,m_{n-3}), \\ K' &= (k,0,\ldots,0), \end{aligned} \qquad \begin{aligned} &k\equiv m_0 \geqslant \cdots \geqslant |m_{n-3}|, \\ &k = 0,1,2,\ldots. \end{aligned} \tag{6.16}$$

Actually, $d_{M0}^{k}(h)$ is the associate spherical function of the subgroup $SO(n-1)$ [see Eq. (5.25)]. Here $d_{K'0}^{l}(g^{(n-1)})$ again depends only on Θ and is given by[20]

$$d^{\,l}_{K'0}(g^{(n-1)})$$
$$=(-1)^k 2^{(n-5)/2}\frac{\Gamma((n-3)/2)\Gamma(l+1)}{\Gamma(n-3)\Gamma(l-k+1)}$$
$$\times\left[(n+2k-3)\frac{\Gamma(n-2)\Gamma(n+k-3)}{k!}\right]^{1/2}$$
$$\times\sinh^{(3-n)/2}\Theta\, P^{(3-n)/2+k}_{l+(n-3)/2}(\cosh\Theta). \qquad (6.17)$$

For the expansion in zonal spherical functions only the fundamental series in (6.12) has to be considered. Vilenkin[20] distinguishes between even and odd dimension [$\mathbf{e}=g\mathbf{a}, f(\mathbf{e})$ being invariant under SO($n-1$) transformations]:

$n=2m+2$:

$$f(\mathbf{e})=[(-1)^m 2^{2m+1}\pi^{m+1/2}\Gamma(m+1/2)]^{-1}\times\int_{-\infty}^{+\infty}\frac{\Gamma(l+2m)}{\Gamma(l)}\hat f(l)d^{\,l}_{00}(g)d\rho,$$

$n=2m+1$:

$$f(\mathbf{e})=[(-1)^{m+1}2^{2m}\pi^m\Gamma(m)]^{-1}\times\int_{-\infty}^{+\infty}\frac{\Gamma(l+2m-1)}{\Gamma(l)}\times\cot(\pi l)\hat f(l)d^{\,l}_{00}(g)d\rho, \qquad (6.18)$$

with $l=-(n-2)/2+i\rho$ and

$$\hat f(l)=\int_{\mathscr{H}_a} f(\mathbf{e})d^{\,l}_{00}(g^{-1})d^{\,n-1}\Omega. \qquad (6.19)$$

Using some group properties one finds the shorter formulation

$$f(g)=\int_{-\infty}^{+\infty}\frac{|\Gamma((n-2)/2+i\rho)|^2}{|\Gamma(i\rho)|^2\Gamma(n-1)}\times\left(1\pm i\frac{n-2}{2}\rho\right)\hat f(l)d^{\,l}_{00}(g)d\rho, \qquad (6.20)$$

$$\hat f(l)=\int_{\mathrm{SO}(n-1,1)} f(g)d^{\,l}_{00}(g^{-1})dg, \qquad (6.21)$$

where the upper sign has to be taken for even dimensions and the lower one for odd n, respectively. However, as the Legendre function $P^{\alpha}_{-1/2+i\rho}(z)$ is symmetric in the index ρ (see Ref. 36), i.e., $P^{\alpha}_{-1/2+i\rho}(z)=P^{\alpha}_{-1/2-i\rho}(z)$, the separation between even and odd n is obsolete. The integration in (6.20) is reducible to one along the positive ρ axis and the substitution

$$\tilde f(\cosh\Theta)=2^{(n-3)/2}\frac{\Gamma(n/2)}{\sqrt\pi}\sinh^{(n-3)/2}\Theta f(g) \qquad (6.22)$$

leads to the generalized Mehler transformation[37]:

$$\tilde f(t)=\frac{|\Gamma((n-2)/2+i\rho)|^2}{|\Gamma(i\rho)|^2}\int_0^\infty c(\rho)P^{(3-n)/2}_{-1/2+i\rho}(t)d\rho, \qquad (6.23)$$

$$c(\rho)=\int_1^\infty \tilde f(t)P^{(3-n)/2}_{-1/2+i\rho}(t)dt. \qquad (6.24)$$

Here n may be an arbitrary complex number.

In the following we consider the Fourier analysis on SO($n-1,1$) in the reduced form

$$f(g)=\int_0^\infty 2\frac{|\Gamma((n-2)/2+i\rho)|^2}{|\Gamma(i\rho)|^2\Gamma(n-1)}\hat f(l)d^{\,l}_{00}(g)d\rho, \qquad (6.25)$$

with $\hat f(l)$ given by (6.21). Comparison with Eq. (2.8) leads to the definition of the dimension

$$d_l=2[|\Gamma((n-2)/2+i\rho)|^2/|\Gamma(i\rho)|^2\Gamma(n-1)]. \qquad (6.26)$$

Indeed analytical continuation of the dimension $d_l^{\mathrm{SO}(n)}$ for SO(n) in $l\to-(n-2)/2+i\rho$ gives

$$d_l^{\mathrm{SO}(n)}\to d_l(-1)^{(n-2)/2}\begin{cases}1, & \text{for even } n,\\ \coth\pi\rho, & \text{for odd } n.\end{cases} \qquad (6.27)$$

The above definition for d_l is confirmed by the orthogonality

$$\int_{\mathrm{SO}(n-1,1)} d^{\,l}_{K0}(g)d^{\,l'*}_{K0}(g)dg=\frac{\delta(\rho-\rho')}{d_l}\delta_{KL}. \qquad (6.28)$$

For the expansion of the short time propagator we rewrite the action (6.4) using $(\mathbf{e}_j,\mathbf{e}_{j-1})=(\hat g_j\mathbf{a},\mathbf{a})=-\cosh\Theta$. Note, that now Θ is the parameter $\theta^{\,n-1}_{\,n-1}$ of the group element $\hat g_j=g_{j-1}^{-1}g_j$. Again only the factor $\exp\{z\cosh\Theta\}$ with $z=imr_jr_{j-1}/\hbar\epsilon$ has to be considered. For the Fourier coefficient we have ($\nu=(3-n)/2$)

$$\hat f(l)=\frac{\Gamma(n/2)}{2^\nu\sqrt\pi}\int_1^\infty e^{zt}(t^2-1)^{-\nu/2}P^{\nu}_{-1/2+i\rho}(t)dt. \qquad (6.29)$$

With Re $\nu<1$ ($\Rightarrow n>1$) and Re $z<0$ ($\Rightarrow$ Im $m>0$) the integral can be performed (p. 194 in Ref. 32):

$$\hat f(l)=\frac{\Gamma(n/2)}{\sqrt\pi}2^{(n-2)/2}(-z)^{\nu-1/2}K_{i\rho}(-z), \qquad (6.30)$$

where $K_{i\rho}(-z)$ is the modified Bessel function of the third kind. Using the asymptotic form for $|z|\to\infty$ and $|\arg z|<3\pi/2$,

$$K_{i\rho}(-z)=\sqrt{\frac{\pi}{-2z}}\exp\left\{z+\frac{\rho^2+1/4}{2z}+O\left(\frac{1}{z^2}\right)\right\}, \qquad (6.31)$$

the path integration results in

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a)=\int_0^\infty\frac{\Gamma(n/2)}{2\pi^{n/2}}d_l\,d^{\,l}_{00}(g_a^{-1}g_b)K_\rho(r_b,r_a,t_b-t_a)d\rho, \qquad (6.32)$$

with

$$K_\rho(r_b,r_a,t_b-t_a)=(r_br_a)^{(1-n)/2}\lim_{N\to\infty}\int\prod_{j=1}^N\exp\left\{\frac{i}{\hbar}S^\rho_j\right\}\times\prod_{j=1}^N\left(\frac{mi}{2\pi\hbar\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}dr_j, \qquad (6.33)$$

$$S^\rho_j=-(m/2\epsilon)\Delta r_j^2-[(\rho^2+\tfrac14)/2mr_jr_{j-1}]\hbar^2\epsilon.$$

The propagator on a space of constant negative curvature: For $r\equiv1$ we set[10]

$$\sqrt{\frac{mi}{2\pi\hbar\epsilon}}\exp\left\{\frac{-im\Delta r_j^2}{2\hbar\epsilon}\right\}=\delta(r_j-r_{j-1}). \qquad (6.34)$$

The radial path integral can be performed immediately. The final integration over the end position r_b yields the propagator on a space of constant negative curvature:

$$K(\mathbf{e}_b,\mathbf{e}_a;t_b-t_a) = \int_0^\infty d\rho \exp\left\{-\frac{i}{\hbar}E_\rho(t_b-t_a)\right\} \times \sum_{k=0}^{\infty} Z_{\rho k}(\phi_b^{(n-1)})Z^*_{\rho k}(\phi_a^{(n-1)}) \times \frac{\Gamma((n-1)/2)}{2\pi^{(n-1)/2}}\sum_M Y_{kM}(\mathbf{e}_b)Y^*_{kM}(\mathbf{e}_a), \quad (6.35)$$

where

$$Z_{\rho k}(\phi) = \frac{\Gamma((n-2)/2+k+i\rho)}{\Gamma(i\rho)} \times \sinh^{(3-n)/2}\phi\, P^{(3-n)/2-k}_{-1/2+i\rho}(\cosh\phi). \quad (6.36)$$

The $Y_{kM}(\mathbf{e})$ are the hyperspherical harmonics of the $(n-1)$-dimensional compact subspace [see Eq. (5.25)]. The $Z_{\rho k}(\phi)$, already discussed by Bander and Itzykson,[38] obey the orthogonality relation

$$\int_0^\infty Z_{\rho k}(\phi)Z^*_{\rho' k}(\phi)d\phi = \delta(\rho-\rho'). \quad (6.37)$$

Finally we remark that the energy spectrum is continuous,

$$E_\rho = (\rho^2+\tfrac{1}{4})(\hbar^2/2m). \quad (6.38)$$

It is, up to the additive constant, identical with that of a free particle having the momentum $p=\hbar\rho$. Therefore the above treatment may be a useful tool for solving scattering problems via path integration. The constant energy shift $\hbar^2/8m$ has also recently been obtained by Balazs and Voros.[26]

The path integral on the timelike hyperboloid $\mathscr{H}_{+1}$ can be performed similarly and has been done in Ref. 8.

B. Path integration over the SU(1,1) manifold

As a last application we consider the Feynman propagator on the group manifold of SU(1,1). The unitary irreducible representations of SU(1,1) have been constructed by Bargmann[39] for the first time. In recent years the group SU(1,1) has attracted much attention in the group theoretical approach to scattering theory.[40] In path integral formalism there exists much interest on SU(1,1) symmetries.[35] A first explicit path integral has been performed by the authors,[8,13] where the SU(1,1) manifold has been realized on the upper sheet ($x^1>0$) of a timelike hyperboloid $\mathscr{H}_{+1}$. In this section we consider quantum mechanics in $E_{2,2}$ with metric (4.8).

The spinor representation is in analogy to SU(2) parametrized in the following way:

$$g(\phi,\theta,\psi) = \begin{pmatrix} e^{i\varphi/2} & 0 \\ 0 & e^{-i\varphi/2}\end{pmatrix} \times \begin{pmatrix}\cosh\theta/2 & \sinh\theta/2 \\ \sinh\theta/2 & \cosh\theta/2\end{pmatrix}\begin{pmatrix} e^{i\psi/2} & 0 \\ 0 & e^{-i\psi/2}\end{pmatrix},$$
$$0\leqslant\varphi<2\pi,\quad 0\leqslant\theta<\infty,\quad 0\leqslant\psi<4\pi. \quad (6.39)$$

The associate invariant measure is

$$dg = (1/16\pi^2)\sinh\theta\, d\theta\, d\varphi\, d\psi. \quad (6.40)$$

Comparison of Eq. (6.39) with (4.11) yields the following explicit identification of the parameters with coordinates in $E_{2,2}$:

$$\begin{aligned} x^1 &= r\sinh(\theta/2)\sin((\psi-\varphi)/2), \\ x^2 &= r\sinh(\theta/2)\cos((\psi-\varphi)/2), \\ x^3 &= r\cosh(\theta/2)\sin((\psi+\varphi)/2), \\ x^4 &= r\cosh(\theta/2)\cos((\psi+\varphi)/2), \end{aligned} \quad \begin{aligned} &0\leqslant\varphi<2\pi, \\ &0\leqslant\theta<\infty, \\ &0\leqslant\psi<4\pi. \end{aligned} \quad (6.41)$$

The Feynman ansatz is then

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \lim_{N\to\infty}\int\prod_{j=1}^N \exp\left\{\frac{i}{\hbar}S_j\right\}\prod_{j=1}^N\left(\frac{m}{2\pi i\hbar\epsilon}\right)\left(\frac{im}{2\pi\hbar\epsilon}\right) \times \prod_{j=1}^{N-1} r_j^3\, dr_j\, 2\pi^2\, dg_j, \quad (6.42)$$

$$S_j = \frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_j r_{j-1}\left[1-\frac{1}{2}\mathrm{Tr}(\hat{g}_j)\right],$$

where we have made use of the results of Sec. IV A.

As is well known, the unitary irreducible representations $D^{l,\sigma}(g)$ in the Hilbert space may be divided into two fundamental and one supplementary series. The fundamental ones are (Bargmann's notation is $k=l+1$)

$$\text{continuous series: } l = -\frac{1}{2}+i\rho\begin{cases}\rho\geqslant 0, & m=0,\pm 1,\pm 2,\ldots, \quad \text{for } \sigma=0,\\ \rho>0, & m=\pm\frac{1}{2},\pm\frac{3}{2},\ldots, \quad \text{for } \sigma=\frac{1}{2},\end{cases} \quad (6.43)$$

$$\text{discrete series: } l = -\frac{1}{2},0,\frac{1}{2},1,\ldots\begin{cases} m=l+1,l+2,\ldots, & \text{for } \sigma=+,\\ m=-l-1,-l-2,\ldots, & \text{for } \sigma=-.\end{cases} \quad (6.44)$$

The matrix elements are given by the multiplier representation

$$d^{l,\sigma}_{mn}(g) = e^{-im\varphi}d^{l,\sigma}_{mn}(\theta)e^{-in\psi}. \quad (6.45)$$

The functions $d^{l,\sigma}_{mn}(\theta)$ are called Bargmann functions and may be viewed as an analytical continuation of the Wigner polynomials $d^J_{mn}(\theta)\in\mathrm{SU}(2)$. Explicitly they are given by hypergeometric functions for $m\geqslant n$:

$$d^{l,+}_{mn}(\theta) = \frac{1}{(m-n)!}\left[\frac{\Gamma(1+m+l)\Gamma(m-l)}{\Gamma(1+n+l)\Gamma(n-l)}\right]^{1/2}\cosh^{-m-n}\frac{\theta}{2}\sinh^{m-n}\frac{\theta}{2} \times {}_2F_1(1-n+l,-n-l;1+m-n;-\sinh^2(\theta/2)), \quad (6.46)$$

$$d^{l,-}_{mn}(\theta) = \frac{1}{(m-n)!}\left[\frac{\Gamma(1-n+l)\Gamma(-n-l)}{\Gamma(1-m+l)\Gamma(-m-l)}\right]^{1/2} \cosh^{m+n}\frac{\theta}{2}\sinh^{m-n}\frac{\theta}{2}$$
$$\times {}_2F_1(1+m+l, m-l; 1+m-n; -\sinh^2(\theta/2)). \quad (6.47)$$

The functions with $n<m$ may be obtained via the relation $d^{l,\sigma}_{mn}(\theta) = (-1)^{m-n} d^{l,\sigma}_{nm}(\theta)$. For the continuous series one finds $d^{-1/2+ip,\sigma}_{mn}(\theta)$ by analytical continuation of Eq. (6.46) or (6.47) in $l \to -\frac{1}{2}+ip$. Note that for $m=n$ we have $d^{l,+}_{mm}(\theta) = d^{l,-}_{mm}(\theta)$.

According to a theorem of Bargmann,[39] the Hilbert space of square integrable functions on SU(1,1) is spanned by the fundamental continuous series and the discrete series with $l \geqslant 0$. The representations $D^{-1/2,\pm}(g)$ are excluded.

From the orthogonality[41]

$$\int_{SU(1,1)} d^{l',\sigma}_{m'n'}(g) d^{l,\sigma *}_{mn}(g) dg = \begin{cases} \dfrac{\delta_{ll'}}{2l+1}\delta_{mm'}\delta_{nn'}, & \text{for } \sigma = (+,-), \\ \dfrac{\delta(\rho-\rho')}{2\rho\tanh\pi(\rho+i\sigma)}\delta_{mm'}\delta_{nn'}, & \text{for } \sigma = (0,\frac{1}{2}), \end{cases} \quad (6.48)$$

follows the explicit Fourier decomposition

$$f(g) = \sum_\sigma \left\{\left[\sum_{2l=0}^{\infty}(2l+1) + \int_0^\infty d\rho\, 2\rho\tanh\pi(\rho+i\sigma)\right] \times \sum_{mn}\hat f_{mn}(l) d^{l,\sigma}_{nm}(g)\right\}, \quad (6.49)$$

$$\hat f_{mn}(l) = \int_{SU(1,1)} f(g) d^{l,\sigma *}_{nm}(g) dg. \quad (6.50)$$

Let $\hat\varphi_j, \hat\theta_j, \hat\psi_j$ be the parameters of $\hat g_j = g_{j-1}^{-1} g_j$, then the trace in Eq. (6.42) is given by $\frac{1}{2}\mathrm{Tr}(\hat g_j) = \cosh(\hat\theta_j/2)\cos((\hat\varphi_j+\hat\psi_j)/2)$. Therefore we have to consider the expansion of the term $\exp\{-iz\cosh(\theta/2)\cos((\varphi+\psi)/2)\}$ with $z = m r_j r_{j-1}/\hbar\epsilon$. The calculation, given in the Appendix, leads to

$$\hat f_{mn}(l) = (2/\pi z)\{K_{2l+1}(ze^{i\pi/2}) + (-1)^{2m}K_{2l+1}(ze^{-i\pi/2})\}\delta_{mn}. \quad (6.51)$$

The complete expansion reads

$$\exp\left\{-\frac{iz}{2}\mathrm{Tr}(g)\right\} = \sum_\sigma \left[\sum_{2l=0}^{\infty}(2l+1) + \int_0^\infty d\rho\, 2\rho\tanh\pi(\rho+i\sigma)\right] \times \frac{2}{\pi z}[K_{2l+1}(iz) + (-1)^{2m}K_{2l+1}(-iz)]\chi^{l,\sigma}(g). \quad (6.52)$$

For the path integration we do need only the asymptotic form for large $|z|$ of the expression

$$F^\sigma_l(z) = (2/\pi z)[K_{2l+1}(iz) + (-1)^{2m}K_{2l+1}(-iz)]. \quad (6.53)$$

For this we have to distinguish between the discrete and continuous case.

As the continuous series is a consequence of the noncompact nature of $E_{2,2}$ we associate this series with the integration over the noncompact coordinates (x^1,x^2), where the mass has to be regularized by a negative imaginary part ($\Rightarrow \mathrm{Im}\, z<0$). If we look at the asymptotic behavior of $K_\nu(iz)$,

$$K_\nu(iz) = \sqrt{\frac{\pi}{2iz}}\exp\left\{-iz + \frac{\nu^2-\frac{1}{4}}{2iz} + O\left(\frac{1}{z^2}\right)\right\}, \quad (6.54)$$

we realize, that in Eq. (6.53) the first term is increasing exponentially for $|z|\to\infty$ with $\mathrm{Im}\, z<0$ and the second one is damping out. Therefore we may drop the last term for continuous l [a similar argument has been used for the asymptotic form (5.22) in Ref. 16.]:

$$F^{(0,1/2)}_l(z) \approx (2/\pi z)K_{ip}(iz), \quad |z| \text{ large}. \quad (6.55)$$

The discrete series, however, may be associated with the compact subspace (x^3,x^4) and the regularization requires a positive imaginary part of the mass ($\Rightarrow \mathrm{Im}\, z>0$). Using the identity[32]

$$K_\nu(iz) = e^{-i\pi\nu}K_\nu(-iz) - i\pi I_\nu(-iz) \quad (6.56)$$

we find

$$F^\sigma_l(z) = (2/\pi z)\{(e^{-2\pi i m} - e^{-2\pi i l}) \times K_{2l+1}(-iz) - i\pi I_{2l+1}(-iz)\}. \quad (6.57)$$

In the discrete case m and l are both integer or half-integer and therefore

$$F^{(+,-)}_l(z) = (2/iz)I_{2l+1}(-iz). \quad (6.58)$$

For $\mathrm{Im}\, z>0$ the asymptotic formula (5.22) is applicable.

Explicitly we have in both cases

$$F^\sigma_l(z) = \frac{1}{2\pi^2}\frac{2\pi}{iz}\left(\frac{2\pi i}{z}\right)^{1/2} \times \exp\left\{-iz - i\frac{(2l+1)^2-\frac{1}{4}}{2z} + O\left(\frac{1}{z^2}\right)\right\}, \quad (6.59)$$

where $\mathrm{Im}\, z>0$ for $\sigma=(+,-)$ and $\mathrm{Im}\, z<0$ for $\sigma=(0,\frac{1}{2})$.

The expansion (6.52) may be now applied to the short time propagator. Using the orthogonality (6.48) the angular integration can be performed and we find

$$K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \sum_\sigma \left\{\left[\sum_{2l=0}^{\infty}\frac{(2l+1)}{2\pi^2}\chi^{l,\sigma}(g_a^{-1}g_b) + \int_0^\infty d\rho\frac{2\rho\tanh\pi(\rho+i\sigma)}{2\pi^2}\chi^{-1/2+ip,\sigma}(g_a^{-1}g_b)\right] \times K_\sigma(r_b,r_a;t_b-t_a)\right\}, \quad (6.60)$$

with

$$
\begin{aligned}
K_\sigma(r_b,r_a;t_b-t_a)&=(r_br_a)^{-3/2}\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j^\sigma\right\}\\
&\quad\times\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}dr_j,
\end{aligned}\tag{6.61}
$$

$$
S_j^\sigma=\frac{m}{2\epsilon}\Delta r_j^2-\left[(2l+1)^2-\frac{1}{4}\right]\frac{\hbar^2\epsilon}{2mr_jr_{j-1}}.
$$

The above SU(1,1) propagator has been recently applied to the one-dimensional modified Pöschl–Teller potential

$$
V(x)=\frac{\hbar^2a^2}{2m}\left(\frac{\kappa^2-\frac14}{\sinh^2ax}-\frac{\lambda^2-\frac14}{\cosh^2ax}\right),\quad 0<x<\infty. \tag{6.62}
$$

This case can be treated similarly to the ordinary Pöschl–Teller problem of Sec. V B. Namely, for $\theta=2ax$ the Feynman ansatz

$$
K(x_b,x_a;t_a-t_b)=a\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\}\prod_{j=1}^{N}\left(\frac{im}{2\pi\hbar a^2\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}\frac12 d\theta_j, \tag{6.63}
$$

$$
S_j=\frac{m}{a^2\epsilon}\left(1-\cosh\frac{\Delta\theta_j}{2}\right)+\left[\frac{\kappa^2-\frac14}{\sinh(\theta_j/2)\sinh(\theta_{j-1}/2)}-\frac{\lambda^2-\frac14}{\cosh(\theta_j/2)\cosh(\theta_{j-1}/2)}-\frac14\right]\frac{a^2\hbar^2}{2m}\epsilon, \tag{6.64}
$$

may be converted into a SU(1,1) path integral for $\kappa,\lambda\in N$. Note that we have used the time reversal trick of Ref. 13:

$$
\begin{aligned}
K(x_b,x_a;t_a-t_b)&=(a/4)(\sinh\theta_b\sinh\theta_a)^{1/2}\exp\{-(i\hbar a^2/8m)(t_b-t_a)\}\\
&\quad\times\int_0^{2\pi}\int_0^{4\pi}Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a)\exp\left\{i\left(\frac{\lambda+\kappa}{2}\varphi_b+\frac{\lambda-\kappa}{2}\psi_b\right)\right\}d\psi_b\,d\varphi_b.
\end{aligned}\tag{6.65}
$$

With $Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a)$ the SU(1,1) symmetry of (6.62) is realized:

$$
Q(\theta_b,\varphi_b,\psi_b;\theta_a,0,0;t_b-t_a)=\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}\tilde S_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar a^2\epsilon}\right)^{1/2}\left(\frac{im}{2\pi\hbar a^2\epsilon}\right)\prod_{j=1}^{N-1}2\pi^2dg_j, \tag{6.66}
$$

$$
\tilde S_j=(m/a^2\epsilon)[1-\tfrac12\operatorname{Tr}(\hat g_j)]. \tag{6.67}
$$

The integration gives

$$
\begin{aligned}
K(x_b,x_a;t_b-t_a)&=a(\sinh\theta_b\sinh\theta_a)^{1/2}\\
&\quad\times\Bigg\{\sum_{l=\sigma}^{(\lambda-\kappa)/2-1}(2l+1)d^{\,l,\sigma}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_b)d^{\,l,\sigma*}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_a)\exp\left\{\frac{i}{\hbar}(2l+1)^2\frac{\hbar^2a^2}{2m}(t_b-t_a)\right\}\\
&\quad+\int_0^\infty dp\,2p\tanh\pi(p+i\sigma)d^{\,-1/2+ip,\sigma}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_b)d^{\,-1/2+ip,\sigma*}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(\theta_a)\exp\left\{-\frac{i}{\hbar}\frac{(2ap)^2\hbar^2}{2m}(t_b-t_a)\right\}\Bigg\},
\end{aligned}\tag{6.68}
$$

where $\sigma=0$ ($\frac12$) for $\kappa+\lambda$ even (odd). With $k=2ap$ we find the standard form

$$
K(x_b,x_a;t_b-t_a)=\sum_{l=\sigma}^{(\lambda-\kappa)/2-1}e^{-i/\hbar E_l(t_b-t_a)}\Psi_l(x_b)\Psi_l^*(x_a)+\int_0^\infty dk\,e^{-i/\hbar E_k(t_b-t_a)}\Phi_k(x_b)\Phi_k^*(x_a), \tag{6.69}
$$

where the bound and scattering states are found simultaneously via path integration:

$$
\begin{aligned}
E_l&=-(2l+1)^2\frac{\hbar^2a^2}{2m},\quad \Psi_l(x)=[a(2l+1)\sinh 2ax]^{1/2}d^{\,l,\sigma}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(2ax),\\
E_k&=\frac{\hbar^2k^2}{2m},\quad \Phi_k(x)=\left[\frac{k}{2a}\sinh 2ax\tanh\pi\left(\frac{k}{2a}+i\sigma\right)\right]^{1/2}d^{\,-1/2+i(k/2a),\sigma}_{(\lambda+\kappa)/2,(\lambda-\kappa)/2}(2ax).
\end{aligned}\tag{6.70}
$$

The energy eigenvalues and eigenfunctions are identical with that obtained by the algebraic method.[42]

The above technique is also applicable to the Coulomb problem in a space of constant positive curvature.[14]

Path integration on SO(n,m) *in bispherical coordinates:* The path integral solution of the modified Pöschl–Teller problem may be used for the calculation of the Feynman propagator in $E_{n,m}$. Choosing the subspace T_{+1} with the parametrization

$$
\begin{aligned}
&x^1 = r\cosh\frac{\sigma}{2}\sin\alpha^{(n-1)}\cdots\sin\alpha^{(1)},\\
&\vdots\\
&x^n = r\cosh\frac{\theta}{2}\cos\alpha^{(n-1)}, && 0\leqslant r,\theta<\infty,\\
& && 0\leqslant\alpha^{(1)},\beta^{(1)}<2\pi,\\
&x^{n+1} = r\sinh\frac{\theta}{2}\sin\beta^{(m-1)}\cdots\sin\beta^{(1)}, && 0\leqslant\alpha^{(i)},\beta^{(i)}<\pi\quad(i\neq 1),\\
&\vdots\\
&x^{n+m} = r\sinh\frac{\theta}{2}\cos\beta^{(m-1)},
\end{aligned}
\tag{6.71}
$$

the Jacobian is

$$
d^{m+n}\mathbf{r} = r^{m+n-1}\,dr\,\frac{1}{2}\sinh^{m-1}\frac{\theta}{2}\cosh^{n-1}\frac{\theta}{2}\,d\theta\,d^{n-1}\Omega(\alpha)d^{m-1}\Omega(\beta),
\tag{6.72}
$$

and the propagator reads

$$
K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = \lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}S_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{n/2}\left(\frac{mi}{2\pi\hbar\epsilon}\right)^{m/2}\prod_{j=1}^{N-1}d^{m+n}\mathbf{r}_j,
\tag{6.73}
$$

$$
\begin{aligned}
S_j = {}&\frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_jr_{j-1}\left(1-\cosh\frac{\Delta\theta_j}{2}\right)\\
&+\frac{m}{\epsilon}r_jr_{j-1}\cosh\frac{\theta_j}{2}\cosh\frac{\theta_{j-1}}{2}\left[1-\mathbf{e}_{j-1}^{\alpha}\cdot\mathbf{e}_j^{\alpha}\right]-\frac{m}{\epsilon}r_jr_{j-1}\sinh\frac{\theta_j}{2}\sinh\frac{\theta_{j-1}}{2}\left[1-\mathbf{e}_{j-1}^{\beta}\cdot\mathbf{e}_j^{\beta}\right].
\end{aligned}
\tag{6.74}
$$

After integration over the $\alpha^{(i)}$'s and $\beta^{(i)}$'s using $SO(n,m)\supset SO(n)\times SO(m)$ we have

$$
\begin{aligned}
K(\mathbf{r}_b,\mathbf{r}_a;t_b-t_a) = {}&\sum_{l,\lambda=0}^{\infty}K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a)\\
&\times\frac{\Gamma(n/2)}{2\pi^{n/2}}\sum_N Y_{lN}(\mathbf{e}_b^{\alpha})Y_{lN}^*(\mathbf{e}_a^{\alpha})\frac{\Gamma(m/2)}{2\pi^{m/2}}\sum_M Y_{\lambda M}(\mathbf{e}_b^{\beta})Y_{\lambda M}^*(\mathbf{e}_a^{\beta}),
\end{aligned}
\tag{6.75}
$$

with

$$
\begin{aligned}
K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a) = {}&\left(r_br_a\sinh\frac{\theta_b}{2}\sinh\frac{\theta_a}{2}\right)^{(1-m)/2}\left(r_br_a\cosh\frac{\theta_b}{2}\cosh\frac{\theta_a}{2}\right)^{(1-n)/2}\\
&\times\lim_{N\to\infty}\int\prod_{j=1}^{N}\exp\left\{\frac{i}{\hbar}\widetilde{S}_j\right\}\prod_{j=1}^{N}\left(\frac{m}{2\pi i\hbar\epsilon}\right)^{1/2}\left(\frac{mi}{2\pi\hbar\epsilon}\right)^{1/2}\prod_{j=1}^{N-1}r_j\,dr_j\,\frac{1}{2}\,d\theta_j,
\end{aligned}
\tag{6.76}
$$

$$
\widetilde{S}_j = \frac{m}{2\epsilon}\Delta r_j^2 + \frac{m}{\epsilon}r_jr_{j-1}\left(1-\cosh\frac{\Delta\theta_j}{2}\right)+\left[\frac{\mu^2-\frac{1}{4}}{\sinh(\theta_j/2)\sinh(\theta_{j-1}/2)}-\frac{\nu^2-\frac{1}{4}}{\cosh(\theta_j/2)\cosh(\theta_{j-1}/2)}\right]\frac{\hbar^2\epsilon}{2mr_jr_{j-1}},
\tag{6.77}
$$

where we have defined $\mu=\lambda+(m-2)/2$ and $\nu=l+(n-2)/2$. The remaining θ integration may be transformed into an $SU(1,1)$ path integral. Indeed,it is formally identical with the modified Pöschl–Teller problem. Using this result leads to

$$
\begin{aligned}
K_{l,\lambda}(r_b,\theta_b;r_a,\theta_a;t_b-t_a) = 2&\left(r_br_a\sinh\frac{\theta_b}{2}\sinh\frac{\theta_a}{2}\right)^{(2-m)/2}\left(r_br_a\cosh\frac{\theta_b}{2}\cosh\frac{\theta_a}{2}\right)^{(2-n)/2}\\
&\times\left[\sum_{l=\sigma}^{(\nu-\mu)/2-1}(2J+1)+\int_0^{\infty}d\rho\,2\rho\tanh\pi(\rho+i\sigma)\right]K_l(r_b,r_a;t_b-t_a)\\
&\times d^{\,l,\sigma}_{(\nu+\mu)/2,(\nu-\mu)/2}(\theta_b)d^{\,l,\sigma *}_{(\nu+\mu)/2,(\nu-\mu)/2}(\theta_b),
\end{aligned}
\tag{6.78}
$$

where $K_l(r_b,r_a;t_b-t_a)$ is given by Eq. (5.37) for $J=l$ and $\sigma=0$ or $\frac{1}{2}$ for $(\nu-\mu)$ even or odd, respectively. For $r\equiv 1$ we have the spectrum $K_l(1,1;t_b-t_a)=\exp\{-(i/\hbar)E_l(t_b-t_a)\}$ with

$$
E_l = \begin{cases}[(2l+1)^2-\frac{1}{4}](\hbar^2/2m), & \text{for } l \text{ discrete},\\ -(\rho^2+\frac{1}{4})(\hbar^2/2m), & \text{for } l=-\frac{1}{2}+i\rho.\end{cases}
\tag{6.79}
$$

VII. DISCUSSION AND OUTLOOK

In the present paper we have discussed the path integral on compact and noncompact rotation groups. The group manifold has been embedded into Euclidean and pseudo-Euclidean spaces, respectively. For this we had to generalize the usual path integral formalism, where the construction is very much similar to that of Feynman. Especially the regularizing scheme had to be modified in order to get well-defined Feynman integrals. Restricting the discussion to a connected subspace of the pseudo-Euclidean space the introduction of polar coordinates leads to a separation into a radial and angular part.

Application of group theory enables us to perform the angular integration, where group theory is introduced through identification of coordinates with group parameters. Writing the short time propagator as a function of group elements, the Fourier analysis on the group leads to an expansion of the propagator in unitary irreducible representations. We have found two methods. For dim $G = \dim \mathscr{H}_\alpha$ the short time action may be written in terms of the character of the fundamental representation, $\chi(g^{(f)}) = \mathrm{Tr}\, g^{(f)}$. Note that the short time action is formally identical with the Wilson action in lattice gauge theories. The character expansion, which has been already used extensively in lattice gauge theories, is applicable. The angular integration reduces to an application of the orthogonality relation of group characters. The only simple Lie groups that may be treated in this way are SU(2) and SU(1,1). In the general case $\dim G \geqslant \dim \mathscr{H}_\alpha$ and the short time action is, by construction, invariant under transformations of the subgroup H as $\mathscr{H}_\alpha = G/H$. Here the expansion in zonal spherical functions is a proper treatment and the application of their orthogonality relation enables us to perform the angular integrals. In both cases the remaining radial path integral is expressed in terms of modified Bessel functions of the first and third kind for compact and noncompact groups, respectively.

The formalism has been applied to the physically most important groups. For the compact groups SO(n) and SU(2) we have recovered known expansion formulas, which have found many applications in path integration in the recent years. For noncompact groups such an expansion has been applied in path integration only by the authors.[13] Here we have chosen the n-dimensional Lorentz group SO($n-1$,1) and the group SU(1,1). The SO($n-1$,1) propagator is found to have the continuous spectrum of a free particle and therefore may become an important tool in scattering theory via path integration [e.g., Rutherford scattering has a SO(3,1) symmetry]. The spectrum generating property of the SU(1,1) algebra, which has numerous applications in group theory, has been used in path integration, too. With the SU(1,1) propagator the bound and scattering states of various problems (here the modified Pöschl–Teller potential has been taken) may be found simultaneously.

Besides nonrelativistic quantum theory the proposed expansion methods may also be very useful in quantum field theories. For pure Yang–Mills lattice gauge theories the character expansion has already been used for a long time. It would be interesting to know whether the expansion in zonal spherical functions still works in theories with matter fields like the symmetry breaking Higgs field. Another area of applications is the path integral formalism of statistical physics. Here the partition function is given as a functional integral over the Boltzmann factor $\exp\{-\beta H\}$. Especially for scalar theories the expansion in zonal spherical functions seems to be successful. How far both techniques may be applied in field theories of elementary particle and solid state physics is under present investigation by the authors.[43]

ACKNOWLEDGMENTS

The authors would like to thank A. Inomata and G. Reents for useful discussions.

APPENDIX: DERIVATION OF THE SU(1,1) FOURIER COEFFICIENT

The formula (6.50) of the Fourier coefficient reads, for $f(g) = \exp\{-(iz/2)\mathrm{Tr}(g)\}$,

$$\hat{f}_{mn}(l) = \frac{1}{16\pi^2}\int_0^{4\pi}\int_0^{2\pi}\int_0^{\infty} \exp\left\{-iz\cosh\left(\frac{\theta}{2}\right)\cos\frac{\varphi+\psi}{2}\right\} \times e^{-im\varphi}e^{-in\psi}\, d^{\,l,\sigma}_{mn}(\theta)\sinh\theta\, d\theta\, d\varphi\, d\psi. \tag{A1}$$

Using the generating function of Bessel functions,

$$\exp\left\{-iz\cosh\left(\frac{\theta}{2}\right)\cos\frac{\varphi+\psi}{2}\right\} = \sum_{2p=-\infty}^{\infty} e^{ip(\varphi+\psi)}e^{-i\pi p}J_{2p}\left(z\cosh\left(\frac{\theta}{2}\right)\right), \tag{A2}$$

the integrals over φ and ψ may be performed and yield

$$\hat{f}_{mn}(l) = \frac{\delta_{mn}}{2}e^{-i\pi m}\int_0^{\infty} J_{2m}\left(z\cosh\left(\frac{\theta}{2}\right)\right)\cosh^{2m}(\theta/2) \times {}_2F_1(l+m+1, m-l; 1; -\sinh^2(\theta/2)) \times \sinh\theta\, d\theta, \tag{A3}$$

where we have used the explicit form (6.47) of the Bargmann functions. Writing the Bessel function in terms of the Meijer G function we find with $x = \cosh^2(\theta/2)$:

$$\hat{f}_{mn}(l) = \delta_{mn}e^{-i\pi m}\int_1^{\infty} x^m G^{10}_{02}\left(\frac{xz^2}{4}\middle| m, -m\right) \times {}_2F_1(l+m+1, m-l; 1; 1-x^2)\, dx. \tag{A4}$$

This integral is a special case of the integral formula #7.831 in Ref. 44. The second set of integrability conditions gives for $\mathrm{Re}\, l \geqslant -\frac{1}{2}$,

$$\hat{f}_{mn}(l) = \delta_{mn}e^{-i\pi m}G^{30}_{24}\left(\frac{z^2}{4}\middle| \begin{matrix} m, -m \\ l, -l-1, m, -m \end{matrix}\right). \tag{A5}$$

Using some properties[45] of the G function the order may be reduced in three steps and the final G functions can be identified with modified Bessel functions of the third kind:

$$\begin{aligned}\hat{f}_{mn}(l) &= \delta_{mn}e^{-i\pi m}G^{20}_{13}\left(\frac{z^2}{4}\middle| \begin{matrix} -m \\ l, -l-1, -m \end{matrix}\right) \\ &= \frac{\delta_{mn}}{2\pi i}e^{-i\pi m}\left\{e^{-i\pi m}G^{21}_{13}\left(\frac{z^2}{4}e^{-i\pi}\middle| \begin{matrix} -m \\ l, -l-1, -m \end{matrix}\right) - e^{i\pi m}G^{21}_{13}\left(\frac{z^2}{4}e^{i\pi}\middle| \begin{matrix} -m \\ l, -l-1, -m \end{matrix}\right)\right\} \\ &= \frac{\delta_{mn}}{2\pi i}\left\{e^{-2\pi i m}G^{20}_{02}\left(\frac{z^2}{4}e^{-i\pi}\middle| l, -l-1\right) - G^{20}_{02}\left(\frac{z^2}{4}e^{i\pi}\middle| l, -l-1\right)\right\} \\ &= \frac{2}{\pi z}\{K_{2l+1}(ze^{i\pi/2}) + e^{-2\pi i m}K_{2l+1}(ze^{-i\pi/2})\}\delta_{mn}.\end{aligned} \tag{A6}$$

[1]R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
[2]R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[3]I. H. Duru and H. Kleinert, Phys. Lett. B **84**, 185 (1979); Fortschr. Phys. **30**, 401 (1982).
[4]R. Ho and A. Inomata, Phys. Rev. Lett. **48**, 185 (1982); R. Ho, Ph.D. thesis, State University of New York at Albany, 1982.
[5]G. Junker and A. Inomata, University of New York at Albany preprint, 1986; H. Kleinert, Phys. Lett. A **116**, 201 (1986).
[6]A. Inomata, Phys. Lett. A **101**, 253 (1984); F. Steiner, *ibid.* **106**, 363 (1984).
[7]I. H. Duru, Phys. Rev. D **30**, 2121 (1984).
[8]G. Junker, Diploma thesis, Universität Würzburg, 1986.
[9]A. Inomata and M. A. Kayed, J. Phys. A **18**, 235 (1985).
[10]G. Junker and A. Inomata, in *Path Integrals From meV to MeV*, edited by M. C. Gutzwiller, A. Inomata, J. R. Klauder, and L. Streit (World Scientific, Singapore, 1986), p. 315.
[11]M. V. Carpio and A. Inomata in Ref. 10, p. 261.
[12]J. M. Cai, P. Y. Cai, and A. Inomata, Phys. Rev. A **34**, 4621 (1986).
[13]G. Junker and M. Böhm, Phys. Lett. A **117**, 375 (1986).
[14]A. O. Barut, A. Inomata, and G. Junker, "Path integral treatment of the hydrogen atom in curved space of constant curvature," Universität Würzburg preprint, 1987.
[15]D. Peak and A. Inomata, J. Math. Phys. **10**, 1422 (1969).
[16]W. Langguth and A. Inomata, J. Math. Phys. **20**, 499 (1979).
[17]L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981).
[18]P. Becher, M. Böhm, and H. Joos, *Eichtheorien der Starken und Elektroschwachen Wechselwirkung* (Teubner, Stuttgart, 1983), 2nd ed.
[19]H. G. Dosch and V. F. Müller, Fortschr. Phys. **27**, 547 (1979).
[20]N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (Am. Math. Soc., Providence, RI, 1968).
[21]K. Maurin, *General Eigenfunction Expansion and Unitary Representations of Topological Groups* (Polish Scientific, Warsaw, 1968).
[22]R. Hermann, *Fourier Analysis on Groups and Partial Wave Analysis* (Benjamin, New York, 1969), p. 191.
[23]B. G. Wybourne, *Classical Groups for Physicists* (Wiley, New York, 1974).
[24]J. S. Dowker, J. Phys. A **3**, 451 (1970).
[25]*Path Integrals from meV to MeV*, edited by M. C. Gutzwiller, A. Inomata, J. R.Klauder, and L. Streit (World Scientific, Singapore, 1986).
[26]N. L. Balazs and A. Voros, Phys. Rep. **143**, 109 (1986).
[27]T. T. Wu and C. N. Yang, Nucl. Phys. B **107**, 365 (1976); Phys. Rev. D **16**, 1018 (1977).
[28]S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces* (Academic, New York, 1978), p. 465.
[29]M. S.Marinov and M. V. Terentyev, Sov. J. Nucl. Phys. **28**, 729 (1978).
[30]M. S. Marinov and M. V. Terentyev, Fortschr. Phys. **27**, 511 (1979).
[31]F. Steiner, in Ref. 25, p. 335.
[32]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin, 1966), 3rd ed.
[33]S. F. Edwards and Y. V. Gulyaev, Proc R. Soc. London Ser. A **279**, 229 (1964).
[34]G. Pöschl and E. Teller, Z. Phys. **83**, 143 (1933).
[35]A. Inomata and R. Wilson, "Path integral realization of a dynamical group," in *Conformal Groups and Related Structures, Lecture Notes in Physics,* Vol. 261, edited by A. O. Barut and H. D. Doebner (Springer, Berlin, 1986), p. 42; A. Inomata and R. Wilson, "Factorization - Algebraization - Path integration and dynamical groups," in *Symmetries in Science II*, edited by B. Gruber and R. Lenczewski (Plenum, New York, 1986), p. 255.
[36]M. C. Gutzwiller, Phys. Rev. Lett. **45**, 150 (1980); Phys. Scr. T **9**, 184 (1985).
[37]F. Oberhettinger and T. P. Higgins, *Tables of Lebedev, Mehler and Generalized Mehler Transforms* (Boeing Scientific Research Lab, Seattle, WA 1961).
[38]M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330, 346 (1966).
[39]V. Bargmann, Ann. Math. **48**, 568 (1947).
[40]Y. Alhassid, F. Gürsey, and F. Iachello, Ann. Phys. (NY) **148**, 346 (1983); **167**, 181 (1986); Phys. Rev. Lett. **50**, 873 (1983); Y. Alhassid, J. Engel, and J. Wu, Phys. Rev. Lett. **53**, 17 (1984); A. Frank and K. B. Wolf, Phys. Rev. Lett. **52**, 1737 (1984).
[41]W. J. Holmann and L. C. Biedenharn, Ann. Phys. (NY) **39**, 1 (1966).
[42]D. Basu and K. B. Wolf, J. Math. Phys. **24**, 478 (1983); A. Frank and K. B. Wolf, *ibid.* **26**, 973 (1985).
[43]G. Junker, Ph.D. thesis, Universität Würzburg, in preparation.
[44]I. S. Gradsteyn and I. M. Rhyzhik, *Tables of Integrals Series and Products* (Academic, New York, 1969), p. 903.
[45]A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions Vol. 1,* Bateman Manuscript Project (McGraw-Hill, New York, 1953), p. 209ff.

The polynomial-type analysis of SU(3) group-theoretical quantities

J. A. Castilho Alcarás
Instituto de Física Teórica, Rua Pamplona 145, CEP: 01405, São Paulo, Brazil
V. Vanagas
Institute of Physics, Academy of Sciences of Lithuanian SSR, Vilnius, USSR

(Received 30 July 1985; accepted for publication 11 February 1987)

Formulas for the transformation matrix between the canonical and noncanonical SU(3) bases in both $\lambda \geqslant \mu$ and $\lambda < \mu$ cases are derived. An expression for the isoscalar factor of the SU(3) coupling coefficient, in the noncanonical basis, with one symmetric representation is also derived. Both expressions are put in a polynomial-type form suitable for coding computer programs to obtain explicit and exact algebraic expressions.

I. INTRODUCTION

Despite the considerable progress during the last two decades in the field of developing the algebraic technique of compact groups, many problems connected with both analytical and numerical calculations still remain unsolved. For example, the general theory of SU(N) coupling coefficients is far behind the stage reached by angular momentum theory. The main problems connected with the generalization of the SU(2) group-theoretical quantities to higher rank groups can be explained by the example of the SU(3) group, largely used in both nuclear and elementary particle physics.

The Kronecker product $(\lambda\mu) \times (\lambda'\mu')$ of SU(3) irreducible representations is not simply reducible, thus the external multiplicity problem is involved. There are at least two types of bases considered, the canonical one, suitable for elementary particle physics applications and the noncanonical one, labeled by the chain SU(3) $\supset$ SO(3) and suitable for applications in nuclear physics. In the last basis an internal multiplicity label is present; thus the isoscalar factor of the general SU(3) coupling coefficient depends on nine parameters, five labeling the irreducible representations and four labeling the multiplicity.

The last ones are not uniquely defined, which makes it essentially more difficult to find analytical expressions as well as to perform numerical calculations.

Similar problems exist for the group-theoretical quantities of higher rank groups, such as coupling and recoupling coefficients, transformation matrices between bases labeled by different chains of subgroups, and so on. Due to the great complexity, even in the cases when the explicit expressions are known, it is difficult to use them in practical applications. One meets an additional problem in the case of large values of the parameters needed, in particular, in nuclear physics theory, namely the uncontrollable growing of rounding errors when calculations are done in floating point numbers.

One way of handling these problems is to derive expressions for the Wigner–Racah algebra quantities in a polynomial-type form. For example, it is well known that the SU(2) coupling coefficients for $j_1 \times j_2 \rightarrow j$, in the case of numerical $j_2 = a$, $m_2 = \alpha$, and $j = j_1 + b$, with numerical b can be presented as a polynomial in j_1 and m_1, with integer coefficients, multiplied by the square root of the ratio of two factorized polynomials also in j_1 and m_1 and with integer coefficients (see, for instance, Ref. 1). We will refer to expressions with such structure as "polynomial-type expressions." For the SU(2) group, polynomial-type expressions in three variables are known for $6j$ and $9j$ coefficients (see, for example, tables in Refs. 2 and 3). Polynomial-type expressions are useful in several aspects: they can help to disclose the analytical dependence of the mean value of operators representing physical quantities on the quantum numbers; they are useful in the study of asymptotic properties of physical quantities; by their use one can obtain *exact* values of the quantities that they represent in much shorter computing time as compared with the one spent by the standard crude expressions.

The question of what polynominal-type expressions exist for higher rank groups, and what methods must be used to obtain them, has never been studied properly. At the present stage of our knowledge it seems impossible to give a definite answer in a general form.

Intuition in this direction can be obtained by considering particular typical examples. In this paper we shall start a project aimed towards the development of the algebraic machinery of the compact groups, mainly unitary and orthogonal groups, based on the polynomial-type decomposition of Wigner-Racah algebraic quantities. Our first goal is to convert the available expressions into polynomial-type expressions, simple to inspect and easy to use. We shall do it via composing the explicit expressions and developing both the special algorithms and the computer programs. Sometimes the known expressions of group-theoretical quantities cannot be put into a polynomial-type expression. In such a case, new expressions for those quantities must be derived. As the next step we plan to apply them in physics, in particular, in collective phenomena of strong interacting particle systems. We shall start with the SU(3) group.

Some details of this project are given in the next section, where the publications on the SU(3) group-theoretical quantities are summarized briefly, and also some general relationships between them are presented. In Sec. III the polynomial-type analysis of the transformation matrix between the canonical and the noncanonical bases is carried out. In Sec. IV the explicit expressions of the isoscalar factor of the SU(3) coupling coefficients with one symmetric representation are derived and presented in a polynomial-type form. In the last section some properties of the quantities considered,

 0022-2488/87/091995-12$02.50

as well as the orthogonalization procedure, are described. In the Appendices the explicit expressions and the notation used are presented.

II. REVIEW AND SOME RELATIONSHIPS BETWEEN SU(3) GROUP-THEORETICAL QUANTITIES

The noncanonical basis of the SU(3) irreducible representation $(\lambda\mu)$ labeled by the angular momentum L and its projection M, as well as by the inner multiplicity index K, having the meaning of the L projection in the "intrinsic" (body-fixed) z axis has been introduced by Elliott.[4] Another possibility of classification of the multiple appearance of L in $(\lambda\mu)$ was proposed by Bargmann and Moshinsky,[5] using the eigenvalues of a Hermitian SO(3) scalar operator $\hat{\Omega}$. More details on the internal labeling of the SU(3) $\supset$ SO(3) basis and the transformations connecting it with the canonical one SU(3) $\supset$ SU(2) $\supset$ U(1) can be found in the review articles (Refs. 6 and 7).

As far as we were able to trace, the first publications on the systematical studies of the SU(3) coupling coefficients appeared in 1962. In Ref. 8 the explicit expression of the SU(3) coupling coefficients for the simply reducible Kronecker product $[h'_1 h'_2]\times[h''_1]\rightarrow[h_1 h_2 h_3]$ in the canonical basis has been obtained. (We will use labels inside parentheses ( ) to indicate noncanonical basis and labels inside brackets [ ] for the canonical.) The explicit expression of the matrix element $\langle\sigma t\tau(\lambda\mu)qLM\rangle$ of the matrix $A^{(\lambda\mu)}$ in the case of $M=\tau/2=L$, connecting the canonical basis SU(3) $\supset$ SU(2) $\supset$ U(1) labeled by $\sigma t\tau$, and the SU(3) $\supset$ SO(3) basis labeled by qLM, where q denotes a multiplicity index taking some definite value, was also obtained there. In Ref. 9 a recurrence procedure has been proposed and some numerical tables were presented for the SU(3) coupling coefficient in the basis SU(3) $\supset$ SO(3) for the case of the Kronecker product $(\epsilon 0)\times(\epsilon' 0)\rightarrow(\lambda\mu)$.

Further developments in this field went on in several stages and directions. Some publications[10–19] related to SU(3) or to more general context have been aimed towards the external multiplicity problem. Many papers[20–39] have been devoted to the SU(3) coupling coefficients in the canonical basis. In some of them[26,27,30,31,35] more general SU(N) coupling coefficients have been considered. For the simply reducible case $[h'_1 h'_2]\times[h''_1]\rightarrow[h_1 h_2 h_3]$, an expression of the isoscalar factor containing only a double sum has been obtained by Ališauskas[29] and then generalized to SU(N) (in Refs. 20 and 31) and presented in few different forms, convenient for specific applications (for details see also the review paper, Ref. 40).

In the general case $[h'_1 h'_2]\times[h''_1 h''_2]\rightarrow\alpha[h_1 h_2 h_3]$ a closed expression for the isoscalar factor, in canonical basis, containing a sum over six indices has been obtained[32] and considered[34–39] from the point of view of the external multiplicity problem. Some classification schemes for the external multiplicity problem have been used. One of them, giving orthogonal states, proposed and developed in Refs. 11, 13, 14, 17, and 18, rests on the structure of the unit tensor operators. Following this line, in Ref. 33 a polynomial-type expression was obtained for the denominator function of the SU(3) isoscalar factors in a canonical basis. The biorthogonal basis has been introduced[34] (see also Ref. 40) and used for a classification leading to simple expressions, especially convenient for applications in the Wigner–Eckart theorem. In numerical calculations the external multiplicity index classified by means of the eigenvalue of SO(3) isoscalar Hermitian operators have been used.[10,12,24,36,39] This question was also investigated from a more general point of view in Ref. 16, and in Ref. 19 the matrix representation of most general classification operators has been obtained. Convenient classification operators have been introduced[15] leading to simple symmetry properties similar to those for the SU(2) coupling coefficients. Summarizing, we conclude that at present the expressions for the SU(3) coupling coefficients in the *canonical* basis are known in a form suitable to adapt the classification of the multiplicity index to both analytical consideration or numerical calculations.

The isoscalar factors of SU(3) in the noncanonical basis SU(3) $\supset$ SO(3) have been considered[9,41–53] for more than twenty years, but a closed form has not yet been derived. In some particular cases polynomial-type expressions have been obtained, some properties including phase convention analyzed,[41–47] and some special direct product cases considered. The general SU(N) isoscalar factors in the noncanonical basis SU(N) $\supset$ SO(N) have also been investigated.[49–52]

A few algorithms and programs for numerical calculations have been developed based on the formulas presented in Refs. 37 and 39 for the canonical basis and in Refs. 54–56 for both canonical and noncanonical bases. Due to the irrational eigenvalues of the classification operators, some of them, for example, those mentioned in Refs. 39, 54, and 55 give the answers only in floating point numbers. As a result, accumulation of rounding errors occurs, which makes it difficult to carry out precise calculations with large values of the parameters, as well as in cases when the quantities considered are internal blocks of more complex expressions.

We will avoid these problems by employing the polynomial algebra technique with only integers involved.

We will use two expressions relating known SU(3) quantities with unknown ones. The first gives the isoscalar factor for the SU(3) coupling coefficients in the noncanonical basis in terms of the same quantities in the canonical basis and the transformation matrices between the bases. The second relates recoupling coefficients and isoscalar factors in the canonical basis. Before presenting the results let us set up the notation.

Let γ, f', and f denote the U(3) irreducible representations $[\gamma_1\gamma_2\gamma_3]$, $[f'_1 f'_2 f'_3]$, and $[f_1 f_2 f_3]$, respectively, and $K_0 L_0, K'L'$, and KL denote their bases in the noncanonical chain $U(3)\supset$ SO(3) with "intrinsic" projections K_0, K', and K, and finally β denotes a multiplicity index in the Kronecker product $\gamma\times f'\rightarrow\beta f$ in some arbitrary classification. Let $m=[m_2 m_3]$ be the U(2) irreducible representation labeling the canonical basis of γ, and similarly $h'=[h'_2 h'_3]$ for f' and $h=[h_2 h_3]$ for f. Then the general isoscalar factor of the U(3) coupling coefficient in the noncanonical chain can be presented in the following form:

$$C^{[\gamma_1\gamma_2\gamma_3]\,[f_1'f_2'f_3']\,\beta\,[f_1f_2f_3]}_{K_0L_0\quad K'L'\qquad KL} = \left(A^{(f_1-f_2,f_2-f_3)K}_{KL(h_2-f_2,h_3-f_3)}\right)^{-1} \sum_{M_1M_2m_2m_3h_2'h_3'} A^{(\gamma_1-\gamma_2,\gamma_2-\gamma_3)M_1}_{K_0L_0(m_2-\gamma_2,m_3-\gamma_3)} A^{(f_1'-f_2',f_2'-f_3')M_2}_{K'L'(h_2'-f_2',h_3'-f_3')}$$

$$\times C^{L_0\,L'\,L}_{M_1\,M_2K}\, C^{(1/2)(m_2-m_3)\,(1/2)(h_2'-h_3')\,(1/2)(h_2-h_3)}_{(1/2)M_1\,(1/2)M_2\,(1/2)K}\, C^{[\gamma_1\gamma_2\gamma_3]\,[f_1'f_2'f_3']\,\beta\,[f_1f_2f_3]}_{[m_2m_3]\,[h_2'h_3']\,[h_2h_3]} . \tag{2.1}$$

In this result $[h_2\,h_3]$ denotes $[f_2f_3]$ in cases $f_1-f_2 \geqslant f_2-f_3$, and $[f_1f_2]$ in case $f_1-f_2<f_2-f_3$. The C's depending on L projections M_1, M_2, K and $M_1/2$, $M_2/2$, $K/2$ are the usual SU(2) coupling coefficient often called Clebsch–Gordan coefficients. The other C's are the isoscalar factors of SU(3) coupling coefficients. The A's are matrix elements of the matrix $A^{(\lambda\mu)M}$ which transforms the SU(3) canonical basis into the noncanonical one:

$$\phi((\lambda\mu)KLM) = \sum_{p=0}^{\lambda}\sum_{q=0}^{\mu} \phi\left((\lambda\mu)pq\frac{M}{2}\right) A^{(\lambda\mu)M}_{KL(pq)} , \tag{2.2}$$

where M and K denote the usual projection and the "intrinsic" projection of L. In (2.1) use was made of the special correlation between the U(1) and SO(2) bases proposed by Elliott.[4] Due to this correlation, M is related to the Gel'fand labels $[m_2\,m_3]n$ for the chain $U(2) \supset U(1)$ by $M=m_2+m_3-2n$. This choice simplifies the expression for the isoscalar factor in the nonncanonical basis and, as a result, the transformation (2.2) appears only twice under the sums of (2.1).

The scalar product of functions (2.2) with different values of K defines an overlap given by

$$B^{(\lambda\mu)L}_{KK'} \equiv \langle(\lambda\mu)KLM\,|\,(\lambda\mu)K'LM\rangle = \sum_{p,q}\left(A^{(\lambda\mu)}_{KL(pq)}\right)^{*} A^{(\lambda\mu)M}_{K'L(pq)} , \tag{2.3}$$

which is symmetric in K and K'.

The nonorthogonality of Elliott's basis $\phi((\lambda\mu)KLM)$ with respect to K makes it possible to introduce the so-called dual Elliott's basis, orthogonal to Elliott's basis, by the relation

$$\phi^{D}((\lambda\mu)KLM) = \sum_{K'} \phi((\lambda\mu)K'LM) D^{(\lambda\mu)L}_{K'K} , \tag{2.4}$$

where the D's are the inverse of the B's, namely

$$\sum_{K} B^{(\lambda\mu)L}_{K'K} D^{(\lambda\mu)L}_{KK''} = \delta_{K'K''} . \tag{2.5}$$

If one follows the Engeland[42] procedure for coupling two Elliott's bases to a resultant Elliott's basis one arrives to an equation similar to (2.1) to which the A's inside the summations are replaced by their inverse. Equation (2.1) is obtained when a resultant Elliott's basis is obtained by coupling two *dual* Elliott's bases, namely,

$$\phi((\gamma f')\beta fKLM) = \sum_{\substack{K_0L_0M_0\\K'L'M'}} \phi^{D}(\gamma K_0L_0M_0)\phi^{D}(f'K'L'M')\, C^{\gamma\quad f'\quad \beta f}_{K_0L_0\,K'L'\,KL}\, C^{L_0L'L}_{M_0M'M} . \tag{2.6}$$

Not all the coefficients involved in (2.6) are orthogonal with respect to the "intrinsic" projections K of L. Nevertheless we will refer to the lhs of (2.1) as the isoscalar factor of SU(3) coupling coefficients in the noncanonical chain $SU(3) \supset SO(3)$. More details on this question will be given in Sec. V.

The second expression that we will use in the development of the SU(3) algebra technique involves both recoupling coefficients and isoscalar factors of coupling coefficients in the canonical basis. We will use it in the form

$$\sum_{\beta}\begin{bmatrix} [\omega_1\,\omega_2\,\omega_3] & [\epsilon+\gamma_0\,\gamma_0\,\gamma_0] & \underset{\beta}{[f_1+\gamma_0\,f_2+\gamma_0 f_3+\gamma_0]} \\ [\omega_1\,\omega_2\,\omega_3] & [\epsilon'\,0\,0] & [f_1'\,f_2'\,f_3'] \\ [0\,0\,0] & [\gamma_1\,\gamma_2\,\gamma_3] & [\gamma_1\,\gamma_2\,\gamma_3] \end{bmatrix} C^{[\gamma_1\gamma_2\gamma_3]\,[f_1'f_2'f_3']\,\beta\,[f_1+\gamma_0 f_2+\gamma_0 f_3+\gamma_0]}_{[m_2\,m_3]\,[h_2'h_3']\,[h_2+\gamma_0,h_3+\gamma_0]}$$

$$= \sum_{n_2n_3qq'} \delta(n_2+n_3+q,h_2+h_3)\delta(n_2+n_3+q',h_2'+h_3')(-)^{(1/2)(q_2+h_2'-h_3'+n_2-n_3+m_2-m_3)}$$

$$\times C^{[\omega_1\omega_2\omega_3]\,[\epsilon+\gamma_0\gamma_0\gamma_0]\,[f_1+\gamma_0 f_2+\gamma_0 f_3+\gamma_0]}_{[n_2n_3]\,[q+\gamma_0\gamma_0]\,[h_2+\gamma_0 h_3+\gamma_0]}\, C^{[\omega_1\omega_2\omega_3]\,[\epsilon'00]\,[f_1'f_2'f_3']}_{[n_2n_3]\,[q'0]\,[h_2'h_3']}\, C^{[\gamma_1\gamma_2\gamma_3]\,[\epsilon'00]\,[\epsilon+\gamma_0\gamma_0\gamma_0]}_{[m_2m_3]\,[q'0]\,[q+\gamma_0,\gamma_0]}$$

$$\times\left[(h_2'-h_3'+1)(q+1)\right]^{1/2}\begin{Bmatrix} q/2 & (h_2-h_3)/2 & (n_2-n_3)/2 \\ (h_2'-h_3')/2 & q'/2 & (m_2-m_3)/2 \end{Bmatrix}, \tag{2.7}$$

which is adapted (see Refs. 56–58) to the density matrix technique in nuclear theory.

The first factor in the lhs of (2.7) denotes the matrix element of recoupled SU(3) basis functions,

$$\langle(([0],[\omega])[\omega],([\gamma],[\epsilon'])[\epsilon])[f]\,|\,(([0],[\gamma])[\gamma],([\omega],[\epsilon'])[f'])\beta\,[f]\rangle , \tag{2.8}$$

where $[\omega] = [\omega_1\, \omega_2\, \omega_3]$, etc. The last factor in the rhs is the standard $6j$ coefficient of SU(2).

We will see later that all the quantities in the rhs of (2.7) are available. Thus if for each value of β, one of the factors in the lhs is known, Eq. (2.7) gives a system of linear equations which allows us to obtain the other factor. Suppose, the SU(3) isoscalar factors in the canonical basis, for some definite classification scheme β, are known. Then, from (2.7) we can find the recoupling coefficients (2.8) and from (2.1) one obtains the isoscalar factors in the noncanonical basis. Thus Eqs. (2.1) and (2.7) provide explicit relations for the most important quantities of the SU(3) algebra: the isoscalar factors in the noncanonical basis and the recoupling coefficients.

We will follow this line of research. As the first step we examine the building blocks in (2.1) and (2.7), in particular their polynomial structure in order to develop the algorithms and computer programs for both numerical calculations and polynomial-type analysis of the SU(3) quantities mentioned above. We intend to carry out this project in the following sequence.

(1) Polynomial-type analysis and corresponding computer programs for the usual coupling coefficient, $6j$ and $9j$ symbols of the group SU(2). This part of the project is already finished and reported in Ref. 59.

(2) Study of the matrix element $A^{(\lambda\mu)M}$ in (2.2). We will find out that the same quantity can be presented in various forms showing different features when treated from the point of view of polynomial-type analysis. This property of the A's requires special efforts in order to adapt them to our purposes.

(3) Due to the wide field of applications of (2.1) with $f_2' = f_3' = 0$ in nuclear theory, we will carry out the analysis of the isoscalar factors in noncanonical basis in this external multiplicity-free case. Some particular isoscalar factors in polynomial-type form have been derived from (2.1) in Ref. 48.

(4) The explicit polynomial-type expressions of the SU(3) isoscalar factors will be used in a special density matrix representation. We also intend to give some applications along this line in the many-body problem in microscopic nuclear theory.

(5) As the next step we plan to develop in the polynomial-like form the general isoscalar factors of the SU(3) coupling coefficients in noncanonical basis using some definite classification β as well as the recoupling coefficients adapted for this classification.

(6) We intend to continue the studies of the polynomial-type analysis of the group-theoretical quantities of higher rank compact groups.

In this paper we will cover the second and third items of the plan above.

III. POLYNOMIAL-TYPE EXPRESSIONS FOR THE BASIS TRANSFORMATION COEFFICIENTS

Besides the two SU(2) coupling coefficient, the general SU(3) isoscalar factor in the noncanonical basis, defined by the rhs of (2.1), depends on the general SU(3) isoscalar factor in the canonical basis and on the matrix elements of $A^{(\lambda\mu)M}$. We shall start with the last ones. First we will discuss their features and then we will derive a polynomial-type expression for them. Explicit expressions for the matrix elements of $A^{(\lambda\mu)M}$ have been obtained using the Hill–Wheeler integral method[54] as well as the projection technique in an infinitesimal form.[60]

In the last approach an operator $\hat{P}$, which projects from the maximum weight basis state of the SU(3) irreducible representation $(\lambda\mu)$ a general basis state of the noncanonical basis, has been constructed. For the case $\lambda \geqslant \mu$ the matrix elements of $A^{(\lambda\mu)M}$ and the overlap (2.3), respectively, are written, in terms of the matrix elements of this operator, as

$$A^{(\lambda\mu)M}_{KL(pq)} = P^{(\lambda\mu)\,M}_{KL\,(pq)} / [P^{(\lambda\mu)\,K}_{KL\,(00)}]^{1/2}, \tag{3.1}$$

$$B^{(\lambda\mu)\,L}_{KK'} = P^{(\lambda\mu)\,K'}_{KL\,(00)} / [P^{(\lambda\mu)\,K}_{KL\,(00)}\, P^{(\lambda\mu)\,K'}_{K'L\,(00)}]^{1/2}, \tag{3.2}$$

where, according to Elliott[4]

$$\begin{aligned} &K = \mu, \mu-2, \ldots, 0 \text{ or } 1, \\ &L = K, K+1, \ldots, K+\lambda, \text{ for } K \neq 0, \\ &L = \lambda, \lambda-2, \ldots, 0 \text{ or } 1, \text{ for } K = 0. \end{aligned} \tag{3.3}$$

The notations used in (2.1) and (3.1) are related by

$$\begin{aligned} &\lambda = \gamma_1 - \gamma_2, \quad \mu = \gamma_2 - \gamma_3, \\ &p = m_2 - \gamma_2, \quad q = m_3 - \gamma_3. \end{aligned} \tag{3.4}$$

The branching rules for the labels of the canonical chain require

$$\begin{aligned} &p = 0,1,\ldots,\lambda, \quad q = 0,1,\ldots,\mu, \\ &(\mu + p - q \pm M)/2 \geqslant 0 \text{ and integer}. \end{aligned} \tag{3.5}$$

An expression for the matrix elements of the projection operator, valid for $\lambda \geqslant \mu$, is given by Eq. (27) of Ref. 60. (There is a misprint in this expression. The fifth factor in the numerator under square root has no factorial symbol.) The particular case $\mu = 0$, however, can be written in the closed form[6]

$$P^{(\lambda 0)\,M}_{KL\,(pq)} = \delta_{K0}\delta_{q0}(-)^{(p-M)/2} \frac{(2L+1)}{2^{M/2}(\lambda+L+1)!!} [\lambda!((p+M)/2)!((p-M)/2)!(L+M)!(L-M)!(\lambda-p)!]^{1/2}$$
$$\times \sum_x \frac{(-)^x}{2^x x!(x+M)!(L-M-2x)!((p-M)/2-x)!(\lambda-L-p+M+2x)!!}. \tag{3.6}$$

Expressions for the matrix elements of the $A^{(\lambda\mu)M}$ and the overlap, valid for the case $\lambda < \mu$ have been found in Ref. 61, by projecting the basis states of the noncanonical Elliott's basis from the lowest weight basis state, namely,

$$A^{(\lambda\mu)\,M}_{KL\,(pq)} = \frac{(-)^{(\lambda+\mu+p-q-K-M)/2}\, P^{(\mu\lambda)\,M}_{KL\,(\mu-q,\lambda-p)}}{[P^{(\mu\lambda)\,K}_{KL\,(00)}]^{1/2}}, \quad \lambda < \mu, \quad B^{(\lambda\mu)\,L}_{KK'} = (-)^{(K-K')/2} B^{(\mu\lambda)\,L}_{KK'}, \quad \lambda < \mu. \tag{3.7}$$

(The phase factor recommended in p. 90 of Ref. 62 is not correct.)

It follows from (3.1), (3.2), and (3.7) that the matrix elements of the projection operator $\hat{P}$ are the only quantities needed in order to perform the basis transformation (2.2), as well as to orthogonalize the Elliott's basis states.

The expression for the matrix elements of the operator $\hat{P}$ obtained in Ref. 60 is not convenient for a polynomial-type analysis for the following reasons. Expressions under sums can be converted into *explicit* polynomials only when the ranges of indices that control the sums can be put in terms of definite numerical quantities. In the applications of the SU(3) noncanonical basis in nuclear theory it is convenient to treat L as a variable. From the analysis of Eq. (27) of Ref. 60 one concludes that there is no possibility of keeping both λ and μ as variables. Since (27) is valid for $\lambda \geqslant \mu$, we will keep λ as a variable and take μ as a numerical parameter. Ascribing numerical values to K, M, p, and q, three sums can be taken, leaving only one with algebraic bounds depending on λ and L. For each value of μ this sum can be rearranged using recursion relations of hypergeometric functions. These relations however become more and more complicated for increasing μ and it is impossible to implement then in a systematic way in computer programs.

From this analysis we decided to search for more convenient expressions for the matrix elements of $\hat{P}$. Ališauskas suggested that we use an alternative formula composed from some building blocks following from Eqs. (4.1) and (4.10) of Ref. 7 and (4.23b) of Ref. 6. Taking also into account the relationship between different analytical expressions for the basis states of the noncanonical basis,[7] the following expression can be derived:

$$P^{(\lambda\mu)\,M}_{KL\,(pq)} = \sum_{\bar{l}_2} B_{\bar{L}_1\bar{l}_2}((\lambda\mu)\,KL) \sum_{\substack{m_1 m_2\\ p_1 p_2}} \delta(p_1+p_2,\mu+p+q)\, C^{\bar{L}_1\,\bar{l}_2\,L}_{m_1\,m_2\,M}$$
$$\times C^{(\mu+p-q)/2\;(\lambda-p_1-p_2+2\mu)/2\;\lambda/2}_{(p_1-p_2)/2\;(\lambda-p_1+p_2)/2\;\lambda/2}\, C^{p_1/2\;p_2/2\;(\mu+p-q)/2}_{m_1/2\;m_2/2\;M/2}\, A^{(\lambda+\mu,0)m_1}_{0\bar{L}_1\,(p_1 0)}\, A^{(\mu 0)m_2}_{0\bar{l}_2\,(p_2 0)}\,. \tag{3.8}$$

In (3.8) the C's denote the usual SU(2) coupling coefficients, $\bar{L}_1 = \bar{l}_2 + L - \Delta$, and the B coefficient is defined by

$$B_{\bar{L}_1\bar{l}_2}((\lambda\mu)KL) = (-)^{(\mu-K)/2+\delta}\,\frac{[1+(-)^{\bar{l}_2-\Delta-\delta}]}{2}\times\frac{[1+\operatorname{sgn}(\bar{l}_2-\Delta-\delta)]}{2}$$
$$\times\Bigg[\frac{\lambda!(\mu-\bar{l}_2)!!(L+K)!(L-K)!(\lambda+\mu+1)!(2L+\delta+1)!(2\bar{L}_1+1)!!}{(L+1)^\delta(2L)!(\lambda+1)!(2L-\Delta)!(\lambda+\mu-\bar{L}_1)!!(\lambda+\mu+\bar{L}_1+1)!!(\mu+K)!!}$$
$$\times\frac{(\mu+\bar{l}_2+1)!!\,\bar{l}_2^{\Delta}\,2^{\bar{L}_1-\delta}}{(\mu-K)!!(\bar{l}_2-\Delta)!(2\bar{l}_2+1)!}\Bigg]^{1/2}\frac{K^\delta(K-\delta-\Delta)!!}{(K-\bar{l}_2)!!}$$
$$\times\left\{\delta_{K0}+(1-\delta_{K0})\,\frac{(K+\bar{l}_2-2)!!}{(K+\delta+\Delta-2)!!}\right\}, \tag{3.9}$$

with $\Delta = \pi(\lambda - L)$, $\delta = \pi(\lambda - L + \mu)$. Here π is a parity balancing function defined as

$$\pi(n) = [1-(-)^n]/2\begin{cases} =0, & \text{for } n \text{ even},\\ =1, & \text{for } n \text{ odd}.\end{cases} \tag{3.10}$$

Expression (3.9) has been obtained from Eqs. (4.1) and (4.10) of Ref. 7 with the help of the summation formula

$$\sum_{l_{20}}\frac{(-)^{(l_{20}-K)/2+\delta}(l_{20}-\delta-\Delta)!!(\bar{l}_2+l_{20}-2)!!}{(l_{20}+K)!!(l_{20}-K)!!(\mu-l_{20})!!(\bar{l}_2+l_{20}-\mu-\Delta-\delta)!!}$$
$$=\frac{(-)^{(\mu-K)/2+\delta}(K-\delta-\Delta)!!(\mu-\bar{l}_2)!!(\mu+\Delta+\delta-2)!!}{(\mu+K)!!(\mu-K)!!(\bar{l}_2-\Delta-\delta)!!(K-\bar{l}_2)!!}\left[\delta_{K0}+(1-\delta_{K0})\,\frac{(K+\bar{l}_2-2)!!}{(K+\delta+\Delta-2)!!}\right]. \tag{3.11}$$

Although Eq. (27) of Ref. 60 and our Eq. (3.8) represent the same quantity, from the point of view of the polynomial-type analysis, the latter is substantially better since all bounds on the sums are *numerical* when we ascribe numerical values to μ, K, M, p, and q and keep λ, L, and Δ as variables.

The structured form of (3.8) allows one to find the range of summations by simply taking into account the branching rules involved in its explicit building blocks. Using the symmetries of these blocks one finds the existence of the following symmetry relation

$$P^{(\lambda\mu)\,-M}_{KL\,(pq)} = (-)^{\Delta+q}\,P^{(\lambda\mu)\,M}_{KL\,(pq)}\,, \tag{3.12}$$

which, in particular, explains the appearance of some unexpected zeros in the matrix elements of $\hat{P}$.

In the analysis of (3.8) it is convenient to use the following generalization of Pochhammer's symbol (from now on refered to as POCH for short)

$$P(x;k;d) = x(x+d)(x+2d)\cdots(x+(k-1)d),\ \text{for } k=1,2,\dots,$$
$$=1 \quad \text{for} \quad k=0, \tag{3.13}$$

introduced in Ref. 59. The $P(x;k;d)$ is a product of k factors of an arithmetic series with starting term x and increment d. Obviously, when x is a variable, $P(x;k;d)$ is a polynomial of degree k in x. From its definition it follows that

$$P(x/a;k;d) = P(x;k;ad)/a^k, \quad a\neq 0 . \tag{3.14}$$

Many commonly used quantities are particular examples of POCH's. For example,

$$\begin{aligned} &n! = P(1;n;1), \quad x^n = P(x;n;0), \quad (x+n)!/x! = P(x+1;n;1), \\ &(x+2n)!!/x!! = P(x+2;n;2), \quad \Gamma(x/a+n)/\Gamma(x/a) = P(x;n;a)/a^n, \quad a\neq 0 . \end{aligned} \tag{3.15}$$

The POCH (3.13) is defined only for non-negative k. However, in order to present some formulas below in a more compact form we will use the following *formal* extension of POCH for negative second argument:

$$P(x;k;d) = [P(x+kd;-k;d)]^{-1}, \quad k=-1,-2,\dots, \tag{3.16}$$

in analogy to the extension of powers to negative exponent. Returning to Eq. (3.8), substituting the explicit expressions for its internal blocks, canceling factors and rearranging in terms of POCH's, one gets, for $M\geqslant 0$ the following expression which is suitable for algebraic calculations via the computer:

$$\begin{aligned} P^{(\lambda\mu)M}_{KL(pq)} &= \frac{(-)^{(p-q-K+M)/2}((K-\delta-\Delta)/2)!(2L+1)(\lambda-p)!}{2^{\mu+2\Delta+\delta+(p+q)/2}(K-\delta)!(2K)![\mu!(\mu+p)!]^2(L+M)^\Delta} \\ &\times B(K,\Delta)\, R^{(\lambda\mu)M}_{KL(pq)}\, [(\lambda+L+\mu+K+1-\Delta)!!(\lambda-L+\mu+K+\Delta)!!P(\lambda+L+\mu-K+2-\Delta;k;2)]^{-1} \\ &\times\left[\frac{q!(\mu+p-q+1)P(L-M+1;M-K;1)P(L+K+1;M-K;1)P(\lambda-p+1;p;1)P(\lambda+\mu-q+2;q;1)}{((\mu+K)/2)!((\mu-K)/2)!p!P(\mu-q+1;p+q+1;1)P((\mu+p-q-M)/2+1;M;1)}\right]^{1/2}, \end{aligned} \tag{3.17}$$

where

$$B(K,\Delta) = \begin{cases} (K+(2-K)\Delta)/2, & \text{for } \mu \text{ even and } K>2, \\ 1, & \text{otherwise}, \end{cases} \tag{3.18}$$

and R is a polynomial in λ, L, and Δ. Its explicit expression is too long and inconvenient to print. For this reason we give below only a schematic expression of it that shows its polynomial structure:

$$\begin{aligned} R^{(\lambda\mu)M}_{KL(pq)} &= (-)^\mu \sum_{\bar{l}_2=\pi(\mu)}^{K}{}' N_1P_1(\lambda+L-\Delta)P_2(L)P_3(\lambda-L+\Delta)P_4(\lambda+L-\Delta) \\ &\times \sum_{p_1=p+q}^{p+\mu} N_2P_5(\lambda) \sum_{m_2=-\bar{l}_2}^{\bar{l}_2}{}' N_3[1+\Delta(2Lm_2+2\bar{l}_2M-1)]P_6(L-\Delta)X(\lambda,L,\Delta)YW, \end{aligned} \tag{3.19}$$

$$\begin{aligned} X(\lambda,L,\Delta) &= \sum_{n=0}^{(p_1-m_1)/2}(-2)^n\binom{(p_1-m_1)/2}{n}P((p_1-m_1)/2-n+1;n+(p_1-m_1)/2;1) \\ &\times P_7(L-\Delta)P_8(\lambda-L+\Delta) . \end{aligned} \tag{3.20}$$

The P_i's are polynomial POCH's $P_i(x+n_i;k_i;d_i)$ of their arguments x, while the N_i's, Y, and W are numbers. The reader interested in the explicit expression for R will find its pieces listed in the Appendices. This same printing strategy will be used for other quantities in this paper. We will use also the convention that sums with a prime have their indexes running in steps of 2.

Expressions (3.17)–(3.20) are adapted for $\lambda\geqslant\mu$ and $M\geqslant 0$. For negative M one simply uses symmetry relation (3.12). When $\lambda<\mu$ we must treat μ as a variable $\bar{\lambda}$, while λ will be a number $\bar{\mu}$. Using then (3.7) for $(\bar{\mu}\bar{\lambda})$, we will see that we need an analogous expression for $P^{(\lambda\mu)M}_{KL(pq)}$ in which p will be no longer a number but will have the form $p=\lambda-\alpha$, with $\alpha=0,1,2,\dots$. By rearranging terms differently in (3.8), one obtains another expression for (3.8) which is adapted again to the case $M\geqslant 0$:

$$\begin{aligned} P^{(\lambda\mu)M}_{KL(\lambda-\alpha,q)} &= \frac{(-)^{(K+\mu)/2}K^\delta(-)^{(\lambda-L+\Delta)/2}(2L+1)(\lambda+\mu-\alpha-q-M)!!}{2^a((\alpha-q+\mu-\nu-\sigma+2)/2)!(\mu+K)!(\mu!)^2(\lambda-L+\mu-\delta)!!} \\ &\times(L+M-\Delta+\beta-1)!![(L+K-\Delta-M+\mu-\nu-\sigma+2)!!(\lambda+L-\Delta+\mu+K+1)!! \\ &\times P(L+M-\Delta+2-\beta;(K+\mu-\nu-\sigma+2+\beta)/2;2)]^{-1}\frac{1}{P(\lambda+L+\mu+2+\delta;(K-\Delta-\delta)/2;2)} \\ &\times\left[\frac{q!(\lambda+\mu-\alpha-q+1)P(\lambda+\mu-q+2;q;1)\lambda!}{2^\mu(\mu-q)!\alpha!((\mu+K)/2)!((\mu-K)/2)!P(L-K+1;2K;1)}\right. \\ &\times\left.\frac{P(\lambda+\mu-\alpha-q-M+2;M;2)}{P(L-M+1;2M;1)P(\lambda-\alpha+1;\mu+1;1)}\right]^{1/2}\tilde{R}^{(\lambda\mu)M}_{KL(\lambda-\alpha,q)}, \end{aligned} \tag{3.21}$$

where

$$\begin{aligned} &\nu=\pi(q-\mu), \quad \sigma=\pi(\alpha), \quad \beta=\pi(\nu+\sigma), \\ &a=\Delta-q+2\nu+3\mu+(\Delta+\delta+\alpha-q+\nu+\sigma)/2+(\mu-\Delta-\delta)/2 . \end{aligned}$$

Again, $\tilde{R}$ is a polynomial in λ, L, and Δ with the following structure:

$$\tilde{R}^{(\lambda\mu)M}_{KL(\lambda-\alpha,q)} = \sum_{\bar{l}_2=\pi(\mu)}^{K}{}' N_4 P_9(\lambda+L-\Delta)P_{10}(L)P_{11}(\lambda-L+\Delta)P_{12}(L-\Delta)P_{13}(L-\Delta)P_{14}(\lambda+L-\Delta)$$
$$\times\sum_{p_2=q}^{\mu} N_5 \begin{Bmatrix} P_{15}(L-\Delta)P_{16}(L-\Delta) \\ P_{17}(L-\Delta)P_{18}(L-\Delta)\end{Bmatrix} \sum_{m_1 m_2=M}{}' N_6 P_{19}(L-\Delta)P_{20}(L-\Delta)P_{21}(L-\Delta)P'(L,\Delta)$$
$$\times\tilde{X}(\lambda,L,\Delta)\tilde{Y}\tilde{W}(\lambda)\,, \tag{3.22}$$

where

$$P'(L,\Delta) = \begin{cases} -2\bar{l}_2, \text{ for } m_2=-\bar{l}_2 \text{ and } \Delta=1\,, \\ P(L-M+1;\bar{l}_2+m_2-\Delta;1)\,[1+\Delta(2\bar{l}_2M+2Lm_2-1)], \text{ otherwise}, \end{cases}$$
$$\tilde{X}(\lambda,L,\Delta) = \sum_{r=-p_2}^{\alpha-q-\eta} N_7 P_{22}(L-\Delta)P_{23}(\lambda-L+\Delta)P_{24}(L-\Delta)\,, \tag{3.23}$$
$$\tilde{W}(\lambda) = \sum_{k=0}^{p_2-q}(-)^k\binom{p_2-q}{k}P_{25}(\lambda)P_{26}(\lambda), \quad \eta=\pi(\alpha-q+p_2)\,.$$

As before, the P_i's are polynomial POCH's in their arguments and are listed in the Appendices together with the N_i's and $\tilde{Y}$. The quantity between braces in (3.22) means that one should pick the upper row when $\pi(p_2-q)=0$ and the lower row when $\pi(p_2-q)=1$.

Equations (3.17)–(3.20) and (3.23), together with their pieces listed in the Appendices, provide explicit expressions of the matrix elements needed to evaluate the $A^{(\lambda\mu)M}$'s and $B^{(\lambda\mu)L}$'s in a polynomial-type form.

IV. NONCANONICAL SU(3) ISOSCALAR FACTORS WITH ONE SYMMETRIC REPRESENTATION

We will now cover the third item listed in Sec. II. The only quantity in the rhs of (2.1) not yet discussed is the SU(3) isoscalar factor in the canonical basis with the special values of $[h_2h_3]$, tied to the left $[h_2h_3]=[f_1f_2]$ or tied to the right $[h_2h_3]=[f_2f_3]$. When $[f'_1f'_2f'_3]$ is symmetric, i.e., equal to $[\bar{p}\,0\,0]$, and $[h_2h_3]$ is tied, one can obtain from the general expressions given in Ref. 30, the following results with no sums at all:

$$C^{[\omega_1\omega_2\omega_3]\,[\bar{p}\,0\,0]\,[f_1f_2f_3]}_{[n_2n_3]\,[\bar{q}\,0]\,[f_1f_2]}$$
$$=(-)^{\omega_1-\omega_2-n_2-n_3}\{(n_2-n_3+1)(f_1-n_2)!(f_1-n_3+1)![(\bar{p}-\bar{q})!(\omega_1-n_2)!(\omega_1-n_3+1)!]^{-1}$$
$$\times(f_2-n_3)!(\omega_1-f_2)!(n_3-\omega_3)!(f_2-f_3+1)!(f_1-f_3+2)!$$
$$\times(f_3-\omega_3)!(n_2-\omega_2)!(n_2-\omega_3+1)![(\omega_2-f_3)!(\omega_1-f_3+1)!(f_1-\omega_1)!(f_1-\omega_2+1)!(f_2-\omega_2)!$$
$$\times(f_1-\omega_3+2)!(f_2-\omega_3+1)!(n_2-f_2)!(\omega_2-n_3)!]^{-1}\}^{1/2}\,, \tag{4.1}$$

$$C^{[\omega_1\omega_2\omega_3]\,[\bar{p}\,0\,0]\,[f_1f_2f_3]}_{[n_2n_3]\,[\bar{q}\,0]\,[f_2f_3]}$$
$$=(-)^{f_2+f_3-n_2-n_3}\{(n_2-n_3+1)(\bar{p}-\bar{q})!(f_1-f_2+1)![(f_1-\omega_1)!(f_1-\omega_2+1)!(f_1-\omega_3+2)!]^{-1}$$
$$\times(f_1-f_3+2)!(\omega_1-n_2)!(\omega_1-n_3+1)!(\omega_2-n_3)!(n_2-f_3)!$$
$$\times(f_2-\omega_2)!(f_2-\omega_3+1)!(f_3-\omega_3)![(n_2-\omega_2)!(n_2-\omega_3+1)!(n_3-\omega_3)!(f_2-n_2)!(f_2-n_3+1)!$$
$$\times(f_3-n_3)!(\omega_1-f_2)!(\omega_1-f_3+1)!(\omega_2-f_3)!]^{-1}\}^{1/2}\,. \tag{4.2}$$

In both (4.1) and (4.2) the rules of the Kronecker product require $\bar{p}=(f_1+f_2+f_3)-(\omega_1+\omega_2+\omega_3)$ and $\bar{q}=(h_2+h_3)-(n_2+n_3)$.

Now all the building blocks are known and (2.1) with $f'_2=f'_3=0$ gives a closed expression for the noncanonical SU(3) isoscalar factor with one symmetric representation.

The first question to be discussed is what kind of polynomial-type expressions can be developed from (2.1). When $f'_2=f'_3=0$, the SU(3) isoscalar factors for tied $[h_2h_3]$ have no sums. The sums over indices M_1,M_2,m_2,m_3 and h'_2 $(h'_3=0)$ are restricted by relations $M_1+M_2=K$ and $m_2+m_3+h'_2=h_2+h_3$. Thus only three summation indices are free. Their values must be determined by numerical bounds, consequently either $\lambda_1\equiv\gamma_1-\gamma_2$, $\mu_1\equiv\gamma_2-\gamma_3$ are numbers and λ_2 is variable or λ_2 and one of λ_1, μ_1 are numbers and the other is variable. For convenience of future applications in nuclear theory, we choose the first possibility, keeping λ_2 as a variable.

Let us examine, with this choice, the polynomial-type structure of (2.1) when $f'_2=f'_3=0$. From the last section it is already known that the transformation brackets A can be presented as polynomial-type expressions in λ and L for $\lambda\geqslant\mu$ or in μ and L for $\lambda<\mu$, with numerical entries M,K,p, and q. In order to restrict M_1, we take L_0 as a number. With this, all the parameters of the first representation are numbers and consequently the first A inside the summation is purely numerical. For the second A, since the second representation is symmetric, only the case $\lambda\geqslant\mu$ is needed. To compute this A and the one outside the sums we need to have both h_2 and h_3 as numbers. The sum in the SU(2) coupling coefficient depending on the L's can be performed taking L_0 and M_1 as numbers. The other SU(2) coupling coefficient is purely numerical.

From this analysis one concludes that our formulas provide polynomial-type expressions for the SU(3) isoscalar factors in the variables $\lambda_2 \equiv f_1'$, L, and L_2. In case $\lambda \equiv f_1 - f_2 \geqslant \mu \equiv f_2 - f_3$ one has $\lambda = \lambda_2 + (\lambda_1 + 2\mu_1 - 2h_2 - h_3)$ and $\mu = h_2 - h_3$. In case $\lambda < \mu$ one has $\lambda = h_1 - h_2$ and $\mu = -\lambda_2 + (h_2 + 2h_3 - \lambda_1 - 2\mu_1)$. The polynomial-type expression that one obtains from (2.1) after replacing all the terms in the rhs, simplifying, and rearranging factors in POCH's is

$$C^{(\lambda_1\mu_1)(\lambda_2 0)(\lambda\mu)}_{K_1L_1\,0L_2\,KL} = (-)^{L_1 - L_2 + L} N_8(\lambda_2 - h_2 - h_3 + \mu_1)! \left[\frac{N_9(\lambda+1)(\mu+1)(f_1 - f_3 - h_2 + h_3)!}{(\lambda_1 + \mu_1 - f_3 + 1)!(f_1 - \lambda_1 - \mu_1)!}\right.$$
$$\times \frac{(\lambda+\mu+2)!(2L_2+1)(2L+1)}{(\lambda_2 + L_2 + 1)!!(\lambda_2 - L_2)!!P_{27}(L)P_{28}(L_2+L)(f_1 - \mu_1 + 1)!(f_1+2)!(\mu_1 - f_3)!(L_2 - L + L_1)!}$$
$$\left.\times \frac{1}{(L_1 - L_2 + L)!\,P^{(\bar\lambda_1\bar\mu_1)K_1}_{K_1L_1(00)}\,P^{(\bar\lambda\bar\mu)K}_{KL(00)}}\right]^{1/2} \left\{\frac{(\mu_1 - f_3)!\theta(\lambda-\mu)}{[(\lambda_1+\mu_1 - f_2)!(f_2 - \mu_1)!f_3!(f_2+1)!]^{1/2}}\right.$$
$$\left. + \left[\frac{(\lambda_1 + \mu_1 - f_2)!}{(f_2 - \mu_1)!(f_2+1)!f_3!}\right]^{1/2} \frac{\bar\theta(\lambda-\mu)}{(\lambda_1+\mu_1)!(\lambda_1+\mu_1+1)!}\right\} R\begin{pmatrix}\lambda_1\mu_1 & \lambda_2 & \lambda\mu\\ K_1L_1 & L_2 & KL\end{pmatrix}, \tag{4.3}$$

where

$$\theta(x)\begin{cases} = 1, & \text{for } x \geqslant 0,\\ = 0, & \text{for } x < 0,\end{cases} \quad \bar\theta(x) = 1 - \theta(x), \tag{4.4}$$

and R is a polynomial in λ_2, L_2, L with the following structure:

$$R\begin{pmatrix}\lambda_1\mu_1 & \lambda_2 & \lambda\mu\\ K_1L_1 & L_2 & KL\end{pmatrix} = \frac{1}{P(L_2 + \gamma_1 + 1; 2\gamma_1; 1)} \sum_{m_2=\mu_1}^{\lambda_1+\mu_1} \sum_{m_3=0}^{\mu_1} N_{10}P_{29}(L_2)P_{30}(L_2)$$
$$\times [N_{11}Z_<(\lambda_2, L_2, L) + (\lambda_1+\mu_1)!Z_>(\lambda_2, L_2, L)]P_{31}(\lambda_2)[N_{12}\theta(\lambda-\mu) + N_{13}P_{32}(\lambda_2)\bar\theta(\lambda-\mu)]. \tag{4.5}$$

The polynomials Z are defined as

$$Z \equiv \sum_{M_1} [Z_>\theta(M_1) + Z_<\bar\theta(M_1)]$$
$$= \sum_{M_1 = -(m_2 - m_3)}^{m_2 - m_3}{}' N_{14}P_{33}(L_2)P_{34}(L_2)$$
$$\times \mathscr{P}_{\text{cleb}}\begin{pmatrix} L_2 & L & L_1\\ -M_2 & K & M_1\end{pmatrix} \mathscr{P}_{\text{cleb}}\begin{pmatrix} (h_2 - h_3)/2 & (m_2 - m_3)/2 & p_2/2\\ K/2 & -M_1/2 & M_2/2\end{pmatrix} R^{(\bar\lambda_1\bar\mu_1)|M_1|}_{K_1L_1(\bar p_1\bar q_1)} \bar R^{(\lambda_2 0)M_2}_{0L_2(p_2 0)}, \tag{4.6}$$

where $p_2 = h_2 + h_3 - m_2 - m_3$, $M_1 + M_2 = K$, $\mathscr{P}_{\text{Cleb}}$ is the polynomial part of the SU(2) coupling coefficient defined as

$$\mathscr{P}_{\text{cleb}}\begin{pmatrix} j_1 & j_2 & j\\ m_1 & m_2 & m\end{pmatrix} = (-)^{2j}\sum_{z=0}^{j+m} (-)^{j+m+z}\binom{j+m}{z} P(j - j_1 + j_2 - z + 1; z; 1)$$
$$\times P(j_1 - m_1 + 1; z; 1)P(j_1 - j_2 - m + z; j + m - z; 1)P(j_2 - m_2 + 1; j + m - z; 1) \tag{4.7}$$

and

$$\bar R^{(\lambda_2 0)M_2}_{0L_2(p_2 0)} = \sum_{k=0}^{(p_2 - M_2)/2} (-2)^k \binom{(p_2 - M_2)/2}{k} P((p_2 + M_2)/2 - k + 1; k + (p_2 - M_2)/2; 1)$$
$$\times P(L_2 - p_2 + 2k + 1; p_2 - M_2 - 2k; 1)P(\lambda_2 - L_2 - 2k + 2; k; 2). \tag{4.8}$$

The constants N_i's and the parameters of the polynomial POCH's $P_i(x_i + n_i; k_i; d_i)$ are given in the Appendices. Despite the presence of a polynomial in the denominator, the R given by (4.5) is really a polynomial since the numerator is exactly divisible by the denominator.

The isoscalar factor given by (4.3)–(4.8) depends on two extra numerical parameters h_2, h_3, whose ranges are determined by the rules of the Kronecker product in U(3); for $\lambda \geqslant \mu$, one has $h_2 = f_2$, $h_3 = f_3$, and for $\lambda < \mu$, $h_2 = f_1$, $h_3 = f_2$. To make a one-to-one correspondence between the U(3) parameters used in (2.1) and those of SU(3) used in (4.3), one lets $\gamma_3 = 0$ and, for symmetric $[f_1' f_2' f_3']$ obtains

$$\gamma_1 = \lambda_1 + \mu_1, \quad \gamma_2 = \mu_1, \quad \gamma_3 = 0,$$
$$f_1' = \lambda_2, \quad f_2' = f_3' = 0, \tag{4.9}$$

$$f_1 = (\lambda_1 + 2\mu_1 + \lambda_2 + 2\lambda + \mu)/3,$$
$$f_2 = (\lambda_1 + 2\mu_1 + \lambda_2 - \lambda + \mu)/3, \tag{4.10}$$
$$f_3 = (\lambda_1 + 2\mu_1 + \lambda_2 - \lambda - 2\mu).$$

The rules of the Kronecker product, when applied to $[f_1' f_2' f_3'] = [\lambda_2\, 0\, 0]$ give us

$[\lambda_1 + \mu_1, \mu_1, 0] \times [\lambda_2\, 0\, 0]$

$$= \sum_{\substack{k=0 \\ k+s\leqslant\lambda_2}}^{\lambda_1} \sum_{s=0}^{\mu_1} [\lambda_2+\lambda_1+\mu_1-s-k;\mu_1+k;s] . \quad (4.11)$$

By fixing h_2 and h_3 one has, in the case when $\lambda \geqslant \mu$,

$$\lambda = \lambda_2 + (\lambda_1 + 2\mu_1 - 2h_2 - h_3), \quad \mu = h_2 - h_3,$$
$$f_1 = \lambda_2 + (\lambda_1 + 2\mu_1 - h_2 - h_3), \quad f_2 = h_2, \quad f_3 = h_3 \quad (4.12)$$

and, in the case when $\lambda < \mu$,

$$\lambda = h_2 - h_3, \quad \mu = -\lambda_2 + (h_2 + 2h_3 - \lambda_1 - 2\mu_1),$$
$$f_1 = h_2, \quad f_2 = h_3, \quad f_3 = \lambda_2 + (\lambda_1 + 2\mu_1 - h_2 - h_3). \quad (4.13)$$

In the case when $\lambda \geqslant \mu$, Eqs. (4.12) fix μ and the difference $\lambda - \lambda_2$. The variable λ_2 is restricted from below by $\lambda_2 \geqslant k + s$ and remains entirely free from above.

In case $\lambda < \mu$, Eqs. (4.13) show that λ and the sum $\mu + \lambda_2$ are fixed. Here, however, λ_2 besides being restricted from below by $\lambda \geqslant k + s$ is also restricted from above by $\lambda_2 \leqslant h_2 + 2h_3 - \lambda_1 - 2\mu_1$, since one must have $\mu \geqslant 0$. This restricts severely the allowed values of λ_2 making comparatively rare the occurrence of cases $\lambda < \mu$. For low-dimensional multiplets $(\lambda_1\mu_1)$ this implies that all entry parameters in (4.13) by numerically fixed. For multiplets $(\lambda_1\mu_1)$ with $\mu_1 = 0$, Eqs. (4.13) fix the value of λ_2. Only for high-dimensional multiplets $(\lambda_1\mu_1)$ with $\mu_1 > 0$ does the formula (4.3) allow several possibilities for λ_2 and L.

V. OVERLAP FOR THE COUPLED FUNCTIONS AND THE ORTHONORMALIZATION PROCEDURE

Assuming that the classification scheme is such that the composite functions are orothogonal in β, the overlap for functions (2.6) is related to the overlap (2.3) by

$$B^{((\lambda_1\mu_1),(\lambda_2\mu_2))\beta(\lambda\mu)L}_{KK'} = \sum_{\substack{K_1K'_1L_1 \\ K_2K'_2L_2}} D^{(\lambda_1\mu_1)L_1}_{K_1K'_1} D^{(\lambda_2\mu_2)L_2}_{K_2K'_2} \times C^{(\lambda_1\mu_1)(\lambda_2\mu_2)\beta(\lambda\mu)}_{K_1L_1\ K_2L_2\ KL} C^{(\lambda_1\mu_1)(\lambda_2\mu_2)\beta(\lambda\mu)}_{K'_1L_1\ K'_2L_2\ K'L} . \quad (5.1)$$

Formally (5.1) depends on $(\lambda_1\mu_1)$, $(\lambda_2\mu_2)$, and β. However, if the isoscalar factors possess some well defined symmetry properties, one can expect that (5.1) is independent of $(\lambda_1\mu_1)$, $(\lambda_2\mu_2)$, and is identical to the usual overlap $B^{(\lambda\mu)L}_{KK'}$. It is very difficult to prove this proposition analytically. In order to confirm it for the isoscalar factors that we obtained in the previous section, we derived a polynomial-type expression for the rhs of (5.1) in the special multiplicity-free case $\mu_2 = 0$ and coded a computer program to evaluate it algebraically. For all cases that we have examined, the proposition was confirmed. This is a powerful test of the consistency of the transformation basis coefficients and the isoscalar factors here obtained. The polynomial-type analysis allowed us to make this test algebraically.

Isoscalar factors with this property have many advantages. One of them is manifested in the orthonormalization procedure. Let $M^{(\lambda\mu)L}$ be the matrix that transforms the noncanonical basis into an orthonormal one, namely,

$$\phi((\lambda\mu)\delta LM) = \sum_K M^{(\lambda\mu)L}_{\gamma K} \phi((\lambda\mu)KLM) . \quad (5.2)$$

The orthonormality property requires then that

$$(M^{(\lambda\mu)L})^\dagger B^{(\lambda\mu)L} M^{(\lambda\mu)L} = I . \quad (5.3)$$

Consider now the linear combination of composite states in the nonorthogonal basis

$$\phi(((\lambda_1\mu_1),(\lambda_2\mu_2))\beta(\lambda\mu)\delta KL) = \sum_K M^{(\lambda\mu)L}_{\gamma K} \phi(((\lambda_1\mu_1),(\lambda_2\mu_2))\beta(\lambda\mu)KLM) . \quad (5.4)$$

The scalar product of such states with labels γ and γ' gives

$$[(M^{(\lambda\mu)L})^\dagger B^{((\lambda_1\mu_1),(\lambda_2\mu_2))\beta(\lambda\mu)L} M^{(\lambda\mu)L}]_{\gamma\gamma'} . \quad (5.5)$$

Then, due to the identity of the overlap matrices, those linear combinations are also orthonormal, that is, the matrix that orthonormalizes the single states also orthonormalizes the composite states.

The orthonormalization procedure that will procedure the matrices $M^{(\lambda\mu)L}$ can be carried out by using the Gram–Schmidt orthonormalization procedure and presents no additional difficulties.

Based on Eqs. (3.17)–(3.23) and (4.3)–(4.13) computer programs were coded to procedure explicit algebraic formulas for the matrix elements of the projection operator $\hat{P}$ and for the isoscalar factors (2.1) in the Elliott's basis. A tabulation of these formulas will be published elsewhere.

ACKNOWLEDGMENTS

We are indebted to S. J. Ališauskas for many helpful discussions, in particular, for pointing out that in the rhs of (5.1) appear the matrix elements of $D^{(\lambda\mu)L}$ instead of $B^{(\lambda\mu)L}$.

One of us (JACA) was supported in part by FINEP, CNPq, and UNESP-IFT contracts. We acknowledge the financial support of FAPESP (Brazilian agency Fundação de Amparo á Pesquisa do Estado de São Paulo) for a visit of one of us (JACA) to the Institute of Physics of the Academy of Sciences of Lithuanian SSR where this work was initiated, and also the financial support of the Scientific and Technological Agreement between Brazil and the Soviet Union for a visit of the other (V V) to the Instituto de Física Teórica.

APPENDIX A: INTERNAL COMPONENTS OF $R^{(\lambda\mu)M}_{KL(pq)}$, EQ. (3.19)

We have

$$N_1 = 2^{\bar{l}_2} P((K+\Delta+\delta-2)/2+1;(\bar{l}_2-\rho)/2;1)P((K-\bar{l}_2)/2+1;(K+\bar{l}_2)/2;1)P(2\bar{l}_2+1;2(K-\bar{l}_2);1), \quad (A1)$$

$$N_2 = (p_1-q)![P(p_1+1;p+\mu-p_1;1)P(\mu+p+q-p_1+1;p_1-p-q;1)]^2, \quad \rho = \pi(\mu), \quad (A2)$$

$$N_3 = (-)^{(p_1+m_1)/2} P((p_1-m_1)/2+1;(p_1+m_1)/2;1)P((p_2-m_2)/2+1;(p_2+m_2)/2;1) \times P(\bar{l}_2-m_2+1;\bar{l}_2+m_2;1), \quad (A3)$$

TABLE I. Parameters $n_i k_i d_i$ of $P_i(x+n_i;k_i;d_i)$.

i	1	2	3	4	5	6	7	8
n_i	$\mu+\bar{l}_2+2$	$\bar{l}_2+1$	$\mu+2-\bar{l}_2$	$\mu-K+2$	$1-p$	$M+1$	$\bar{l}_2-p_1+2n+1$	$\mu-\bar{l}_2-2n+2$
k_i	$K-\bar{l}_2$	$K-\bar{l}_2$	$(K+\bar{l}_2)/2$	$(K+\bar{l}_2)/2$	$p+\mu-p_1$	$\bar{l}_2-m_2$	p_1-m_1-2n	n
d_i	1	1	2	2	1	1	1	2

$$Y=\sum_{j=0}^{(p_2-m_2)/2}(-2)^j\binom{(p_2-m_2)/2}{j}P((p_2+m_2)/2-j+1;j+(p_2-m_2)/2;1)$$
$$\times P(\bar{l}_2-p_2+2j+1;p_2-m_2-2j;1)P(\mu-\bar{l}_2-2j+2;j;2)\,, \tag{A4}$$
$$W=\sum_{\sigma=0}^{(\mu+p-q+M)/2}(-)^\sigma\binom{(\mu+p-q+M)/2}{\sigma}P(\mu-p_1+p-\sigma+1;(p_1-m_1)/2;1)$$
$$\times P(p_1-(\mu+p+q+M)/2+\sigma+1;(p_2+m_2)/2;1)\,. \tag{A5}$$

See Table I.

APPENDIX B: INTERNAL COMPONENTS OF $R_{KL\,(\lambda-\alpha,q)}^{(\lambda\mu)\,M}$, EQ. (3.22)

We have

$$N_4=(-)^{(\mu+\bar{l}_2)/2}\,2^{\bar{l}_2}P((K+\delta+\Delta-2)/2+1;(\bar{l}_2-\delta-\Delta)/2;1)P((K-\bar{l}_2)/2+1;(\bar{l}_2-\Delta-\delta)/2;1)$$
$$\times P(\bar{l}_2+\mu+1;K-\bar{l}_2;1)\,, \tag{B1}$$

$$N_5=\begin{Bmatrix}2^{5(\mu+\nu-p_2)/2}P((\alpha-q+\sigma+p_2)/2+1;(\mu-\nu-p_2)/2+(1-\sigma);1)\\ 2^{5(\mu+\nu-1-p_2)/2+\sigma+2}P((\alpha-q-\sigma+1+p_2)/2+1;(\mu-\nu+1-p_2)/2;1)\end{Bmatrix}P(p+1;\mu-p_2;1)$$
$$\times P(p_2-q+1;\mu-p_2;1)\;[P(p_2+1;\mu-p_2;1)]^2\,P(\bar{l}_2+p_2+1;\mu-p_2;1)\,. \tag{B2}$$

In the braces one should use the first row for $\pi(p_2-q)=0$ and the second for $\pi(p_2-q)=1$,

$$N_6=P((p_2-m_2)/2+1;(p_2+m_2)/2;1)P(\bar{l}_2+m_2+1;p_2-m_2;1)P((p_2+m_2)/2+1;(p_2-m_2)/2;1)\,, \tag{B3}$$
$$N_7=(-)^{(r-p_2)/2}\,2^{(\alpha-q+\eta-r)/2}P(\alpha-q-r+1;r+p_2;1)P((r+p_2)/2+1;(\alpha-q+\eta-r)/2;1)\,, \tag{B4}$$
$$\tilde{Y}=\sum_{y=0}^{(p_2-m_2)/2}(-)^y2^{p_2-m_2-2y}\binom{(p_2-m_2)/2}{y}P(\bar{l}_2-m_2+2y+1;2(y+m_2);1)$$
$$\times P(y+m_2+1;(p_2-m_2)/2-y;1)P((\mu-\bar{l}_2-p_2+m_2)/2+y+1;(p_2-m_2)/2-y;1)\,. \tag{B5}$$

See Tables II–IV.

TABLE II. Parameters n_i,k_i,d_i of $P_i(x+n_i;k_i;d_i)$.

i	9	10	11	12	13	14	15
n_i	$\mu+\bar{l}_2+2$	$\bar{l}_2+1$	$\mu-\bar{l}_2+2$	$\bar{l}_2+M+\mu-\nu-\sigma+4$	$\bar{l}_2-M+\mu-\nu-\sigma+4$	$\mu+\Delta+\delta+2$	$\bar{l}_2-M+p_2+\sigma+2$
k_i	$K-\bar{l}_2$	$K-\bar{l}_2$	$(\bar{l}_2-\Delta-\delta)/2$	$(K-\bar{l}_2)/2$	$(K-\bar{l}_2)/2$	$(\bar{l}_2-\Delta-\delta)/2$	$(\mu-\nu-p_2)/2+1-\sigma$
d_i	1	1	2	2	2	2	2

TABLE III. Parameters n_i, k_i, d_i of $P_i(x+n_i;k_i;d_i)$.

i	16	17	18	19	20	21
n_i	$\bar{l}_2+M+p_2+\sigma+2$	$\bar{l}_2-M+p_2-\sigma+3$	$\bar{l}_2+M+p_2-\sigma+3$	$\bar{l}_2-m_1+\eta+2$	$\bar{l}_2+m_1+\eta+2$	$M+1$
k_i	$(\mu-\nu-p_2)/2+1-\sigma$	$(\mu-\nu+1-p_2)/2$	$(\mu-\nu+1-p_2)/2$	$(p_2-m_2)/2$	$(p_2+m_2)/2$	$\bar{l}_2-m_2$
d_i	2	2	2	2	2	1

TABLE IV. Parameters n_i, R_i, d_i of $P_i(x+n_i;R_i;d_i)$.

i	22	23	24	25	26
n_i	$r-\alpha+\bar{l}_2-m_1+q+2$	$\mu-r-\bar{l}_2-p_2+2$	$r-\alpha+\bar{l}_2+m_1+q+2$	$\mu-\alpha-q+M-2k+2$	$\mu-\alpha+q-M-2p_2+2k+2$
k_i	$(\alpha-q+\eta-r)/2$	$(r+p_2)/2$	$(\alpha-q+\eta-r)/2$	$(p_2-m_2)/2$	$(p_2+m_2)/2$
d_i	2	2	2	2	2

APPENDIX C: INTERNAL COMPONENTS FOR EQS. (4.3)–(4.6)

We have

$$N_8=\frac{(2L_1+1)((K_1-\delta_1-\bar{\Delta}_1)/2)!((\lambda_1+\mu_1-K_1+L_1-\bar{\Delta}_1)/2)!}{(\lambda_1+\mu_1+L_1+k_1-\bar{\Delta}_1+1)!((\lambda_1+\mu_1+K_1-L_1+\bar{\Delta}_1)/2)!(K_1-\delta_1)!(2K_1)!(\bar{\mu}_1!)^2}$$
$$\times\frac{1}{2^{\bar{\mu}_1+3\bar{\Delta}_1-L_1+K_1+\delta_1}[(\lambda_1+\mu_1)!]^3(\lambda_1+\mu_1+1)!(2L_1)!(h_2+h_3-\mu_1)!},\tag{C1}$$

$$N_9=\frac{(L_1-K_1)!(\lambda_1+\mu_1+1)!((h_2-h_3-K)/2)!\,[2^{\lambda_1+3\mu_1}\theta(\lambda_1-\mu_1)+\bar{\theta}(\lambda_1-\mu_1)]}{\bar{\lambda}_1!(L_1+K_1)!((h_2-h_3+K)/2)!((\bar{\mu}_1+K_1)/2)!((\bar{\mu}_1-K_1)/2)!2^{h_2+h_3+\lambda_1+2\mu_1}},\tag{C2}$$

$$N_{10}=P(1;\mu_1+\lambda_1-m_2;1)P(m_2+2;\lambda_1+\mu_1-m_2;1)P(p_2+1;m_2+m_3-\mu_1;1)(m_2-m_3+1)$$
$$\times\begin{cases}P(m_2-\mu_1+1;\lambda_1+\mu_1-m_2;1)\,[P(m_2+1;\lambda_1+\mu_1-m_2;1)]^2, & \lambda_1\geqslant\mu_1,\\ 2^{m_2+m_3}P(\mu_1-m_3+1;m_3;1)\,[P(\lambda_1+\mu_1-m_3+1;m_3;1)]^2, & \lambda_1<\mu_1,\end{cases}\tag{C3}$$

$$N_{11}=P(m_2-m_3+1;\lambda_1+\mu_1-m_2+m_3;1),\tag{C4}$$

$$N_{12}=(-)^{p_2}P(\mu_1-h_3+1;m_2-\mu_1;1)(\mu_1-m_3)!P(h_2-m_2+1;m_2-\mu_1;1)$$
$$\times P(h_3-m_3+1;m_3;1)P(h_2-m_3+2;m_3;1),\tag{C5}$$

$$N_{13}=(-)^{\lambda_1-m_2-m_3}m_3!(\mu_1+1)!P(\lambda_1+\mu_1-m_2+1;m_2;1)P(\lambda_1+\mu_1-m_3+2;m_3;1)$$
$$\times(m_2-\mu_1)!P(\mu_1+2;m_2-\mu_1;1),\tag{C6}$$

$$N_{14}=(-)^{(\bar{p}_1-\bar{q}_1-K_1+|M_1|)/2}(L_1+|M_1|-\bar{\Delta}_1)!P(L_1+M_1+1;L_1-M_1;1)\,[\theta(\lambda_1-\mu_1)$$
$$+(-)^{(\lambda_1+\mu_1+p_1-q_1-K_1-M_1)/2}\bar{\theta}(\lambda_1-\mu_1)]\,[\theta(M_1)+(-)^{\bar{\Delta}_1+\bar{q}_1}((m_2-m_3+M_1)/2)!$$
$$\times P((m_2-m_3-M_1)/2+1;(m_2-m_3+M_1)/2;1)\bar{\theta}(M_1)]\,(-)^{(M_1-K+p_2)/2},\tag{C7}$$

$$p_1=m_2-\mu_1;\quad p_2=h_2+h_3-m_2-m_3;\quad q_1=m_3;\quad \Delta_1=\pi(\lambda_1-L_1);$$
$$\delta_1=\pi(\lambda_1+\mu_1-L_1),\quad \bar{\Delta}_1=\pi(\bar{\lambda}_1-L_1),\tag{C8}$$

$$(\bar{\lambda}_1,\bar{\mu}_1,\bar{p}_1,\bar{q}_1)=(\lambda_1,\mu_1,p_1,q_1),\quad\text{for }\lambda_1\geqslant\mu_1$$
$$=(\mu_1,\lambda_1,\mu_1-q_1,\lambda_1-p_1),\quad\text{for }\lambda_1<\mu_1.\tag{C9}$$

See Table V.

TABLE V. Parameters n_i, k_i, d_i of $P_i(x+n_i;k_i;d_i)$.

i	27	28	29	30	31	32	33	34
n_i	$1-K$	$1-L_1$	$1-\lambda_1-\mu_1$	m_2-m_3+1	$\mu_1+1-h_2-h_3$	$1-p_2$	$1-m_2+m_3$	$1+M_1-K$
k_i	$2K$	$2L_1+1$	$\lambda_1+\mu_1-m_2+m_3$	$\lambda_1+\mu_1-m_2+m_3$	$m_2+m_3-\mu_1$	$\lambda_1+2\mu_1-m_2-m_3$	$K+m_2-m_3-M_1$	$K+m_2-m_3-M_1$
d_i	1	1	1	1	1	1	1	1

[1]E. V. Condon and G. H. Shortley, in *The Theory of Atomic Spectra* (Cambridge U. P., London, 1935).
[2]A. R. Edmonds, *Angular momentum in Quantum Mechanics* (Princeton U. P., Princeton, 1974).
[3]D. A. Varchalovitch, A. N. Moscalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975) (in Russian).
[4]J. P. Elliott, Proc. R. Soc. London Ser. A **245**, 128, 562 (1958).
[5]V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 697 (1960); **23**, 177 (1961).
[6]M. Moshinsky, J. Patera, R. T. Sharp, and P. Winternitz, Ann. Phys. (NY) **95**, 139 (1975).
[7]S. J. Ališauskas, Liet. Fiz. Rinkinys **18**(5), 567 (1978) (in Russian).
[8]M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).
[9]V. Vanagas and A. Jucys, Liet. Fiz. Rink. **2**(3 and 4), 199 (1962).
[10]M. Moshinksy, J. Math. Phys. **4**, 1128 (1963).
[11]B. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[12]C. K. Chew and R. T. Sharp, Can. J. Phys. **44**, 2789 (1966).
[13]J. A. Castilho Alcarás, L. C. Biedenharn, K. T. Hecht, and G. Neely, Ann. Phys. (NY) **60**, 85 (1970).
[14]L. C. Biedenharn and J. D. Louck, J. Math. Phys. **13**, 1985 (1972).
[15]Z. Plukhař, Preprint Karlov University, Praga, 1973.
[16]T. Sharp, J. Math. Phys. **16**, 2050 (1975).
[17]M. A. Lohe, L. C. Biedenharn, and J. D. Louck, J. Math. Phys. **18**, 1883 (1977).
[18]L. C. Biedenharn and D. E. Flath, Commun. Math. Phys. **93**, 143 (1984).
[19]S. J. Ališauskas and P. P. Kulish, Sov. J. Math. **18**, 232 (1985).
[20]H. Pei-Yu, Sci. Sin. **14**, 369 (1965).
[21]K. T. Hecht, Nucl. Phys. **62**, 1 (1965).
[22]G. Ponzano, Nuovo Cimento A **41**, 142 (1966).
[23]R. T. Sharp and H. C. Baeyer, J. Math. Phys. **7**, 1105 (1966).
[24]M. Resnikoff, J. Math. Phys. **8**, 63 (1967).
[25]C. K. Chew and R. T. Sharp, Nucl. Phys. B **2**, 697 (1967).
[26]L. A. Shelepin and V. P. Karasev, Yad. Fiz. **5**, 221 (1967) [English translation: Sov. J. Nucl. Phys. **5**, 156 (1967)].
[27]V. P. Karasev and L. A. Shelepin, Yad. Fiz. **7**, 1131 (1968) [Sov. J. Nucl. Phys. **7**, 678 (1968)].
[28]R. M. Asherova and Yu. F. Smirnov, Nucl. Phys. B **4**, 399 (1968).
[29]S. J. Ališauskas, Liet. Fiz. Rink. **9**(4), 641 (1969) (in Russian).
[30]S. J. Ališauskas, A. A. Jucys, and A. P. Jucys, J. Math. Phys. **13**, 1329 (1972).
[31]E. Chacon, M. Ciftan, and L. C. Biedenharn, J. Math. Phys. **13**, 577 (1972).
[32]S. J. Ališauskas, Liet. Fiz. Rink. **12**(5), 721 (1972) (in Russian).
[33]L. C. Biedenharn, M. A. Lohe, and J. D. Louck, J. Math. Phys. **26**, 1458 (1985).
[34]S. J. Ališauskas, Liet. Fiz. Rink. **18**(6), 701 (1978) (in Russian).
[35]A. U. Klimyk, *Matrix Elements and Clebsch–Gordan Coefficients of Group Representations* (Naukova Dumka, Kiev, 1979) (in Russian).
[36]Z. Plukhař, Yu. F. Smirnov, and V. N. Tolstoy, Preprint Karlov University, Praga, 1981 (in Russian).
[37]I. S. Guseva, Yu. F. Smirnov, and V. N. Tolstoy, Preprint 678 Leningrad Institute of Nuclear Physics, Leningrad, 1981 (in Russian).
[38]S. J. Ališauskas, Liet. Fiz. Rink. **22**(2), 13 (1982) (in Russian).
[39]I. S. Guseva, Yu. F. Smirnov, V. N. Tolstoy, and Yu. I. Kharitonov, Preprint 837, Leningrad Institute of Nuclear Physics, 1983 (in Russian).
[40]S. J. Ališauskas, Fiz. Elem. Chastits At. Yadra **14**, 1336 (1983) [Sov. J. Part. Nucl. **14**, 563 (1983)].
[41]H. Horie, J. Phys. Soc. Jpn. **19**, 1783 (1964).
[42]T. Engeland, Nucl. Phys. **72**, 67 (1965).
[43]M. Josifescu and E. Stancu, Nucl. Phys. B **1**, 471 (1967); A **142**, 481 (1970).
[44]J. D. Vergados, Nucl. Phys. A **111**, 681 (1968).
[45]T. Sebe, Nucl. Phys. A **109**, 65 (1968).
[46]R. T. Sharp, H. C. Baeyer, and S. C. Pieper, Nucl. Phys. A **127**, 573 (1969).
[47]H. C. Baeyer and R. T. Sharp, Nucl. Phys. A **140**, 118 (1970).
[48]M. Taurinskas and V. Vanagas, Liet. Fiz. Rink. **17**(6), 709 (1977) (in Russian).
[49]S. J. Ališauskas and E. Z. Norvaišas, Liet. Fiz. Rink. **20**(2), 3 (1980) (in Russian).
[50]S. J. Ališauskas, Liet. Fiz. Rink. **22**(3), 3 (1982) (in Russian).
[51]S. J. Ališauskas, Liet. Fiz. Rink. **23**(3), 3 (1983) (in Russian).
[52]S. J. Ališauskas, J. Phys. A: Math. Gen. **17**, 2899 (1984); **18**, 737(E) (1985).
[53]K. T. Hecht and Y. Suzuki, J. Math. Phys. **24**, 785 (1983).
[54]J. P. Draayer and S. A. Williams, Nucl. Phys. A **129**, 647 (1969).
[55]J. P. Draayer and Y. Akiyama, J. Math. Phys. **14**, 1904 (1973).
[56]V. Vanagas and M. Taurinskas, Liet. Fiz. Rink. **6**, 699 (1977) (in Russian).
[57]V. Vanagas and M. B. Taurinskas, Liet. Fiz. Rink. **17**(6), 699, 709 (1977) (in Russian).
[58]V. Vanagas and M. B. Taurinskas, Liet. Fiz. Rink. **17**(6), 718 (1977) (in Russian).
[59]J. A. Castilho Alcarás and V. Vanagas, Rev. Bras. Fis. **14**, 121 (1984).
[60]R. M. Asherova and Yu. F. Smirnov, Nucl. Phys. A **144**, 116 (1970).
[61]V. Vanagas, O. Katkevičius, Liet. Fiz. Rink. **23**(2), 3 (1983) (in Russian).
[62]R. M. Asherova and Y. F. Smirnov, Rep. Math. Phys. **4**, 83 (1973).

Classification and interpretation of the supertableaux of the orthosymplectic groups OSP($m|4$)

Michel Gourdin[a)]
Laboratoire de Physique Théorique et Hautes Energies,[b)] *Paris, France*

(Received 2 April 1986; accepted for publication 22 April 1987)

The supertableaux of the orthosymplectic group OSP $(m|4)$ are analyzed and interpreted as representations of the corresponding superalgebras. An extension of the results to the general case OSP $(m|2p)$ is proposed.

I. INTRODUCTION

The orthosymplectic group OSP$(m|2p)$ leaves invariant an even bilinear form G in a Z_2-graded vector space V of dimension $m+2p$. The associated superalgebra is basic and simple[1] and its Bose sector is the ordinary Lie algebra of the direct product SO$(m) \otimes$ Sp$(2p)$. When $m=2$ this even subalgebra contains a U(1) factor and we have a class I superalgebra $C(p+1)$. When $m \neq 2$ the even subalgebra is semisimple, or simple for $m=1$ and we have a class II superalgebra.

The starting point of the present investigation is a paper by Farmer and Jarvis[2] where orthosymplectic supertableaux have been discussed. A full analysis is given here by using the notion of generalized atypical supertableaux already introduced for the superunitary case by Delduc and the author.[3] The problem is the following. To a supertableau T of the orthosymplectic group OSP$(m|2p)$ we associate a supertensor $\mathscr{T}$ of the general graded linear group PL$(m|2p)$ whose indices have the supersymmetry properties as indicated by the supertableau T. This supertensor is an irreducible representation of PL$(m|2p)$. Because of the existence of G for orthosymplectic transformations $\mathscr{T}$ is not, in general, an irreducible representation of OSP$(m|2p)$. In order to isolate the irreducible parts we must subtract traces in all the possible ways and we now write symbolically

$$\mathscr{T} = \mathscr{T}_{\rm IRR} + \sum_{1\ \rm Trace} G\mathscr{T}^{(1)}_{\rm IRR} + \sum_{2\ \rm Traces} GG\mathscr{T}^{(2)}_{\rm IRR} + \cdots .$$

If r is the rank of the tensor $\mathscr{T}$, the irreducible components $\mathscr{T}_{\rm IRR}$, $\mathscr{T}_{\rm IRR}(1)$, $\mathscr{T}_{\rm IRR}(2)$,..., are, respectively, of rank r, $r-2$, $r-4$,... . When all the traces can be separated $\mathscr{T}$ is a fully reducible tensor representation of OSP$(m|2p)$. The supertableau T is called an irreducible supertableau and it corresponds to $\mathscr{T}_{\rm IRR}$. We shall say that the supertableau T describes a reducible—or a fully reducible—representation of OSP$(m|2p)$. This situation occurs, in particular, for typical supertableaux.

When some trace terms cannot be separated from $\mathscr{T}_{\rm IRR}$, then $\mathscr{T}$ is a nonfully reducible supertensor. The supertableau T is called nonirreducible and, with the other supertableaux corresponding to the nonseparated parts of the trace, the supertableau T belongs to what we define as a generalized atypical supertableau (GAST). Such an object describes a nonfully reducible representation of OSP$(m|2p)$.

In order to recognize if an atypical supertableau T is irreducible or nonirreducible we can use the following empirical rule.

(a) Compute the dimension $d_{\rm ST}$ of the supertableau T as given by the determinant of Balentekin and Bars.[4] For a legal supertableau[2] $d_{\rm ST}$ is a positive integer.

(b) Compute the Kac–Dynk in parameters of the highest weight $\Lambda_{\rm ST}$ of the supertableau T.[2] This highest weight is also the highest weight of an atypical irreducible representation R of OSP$(m|2p)$.

(c) By analyzing the SO$(m) \otimes$ Sp$(2p)$ content of the representation R compute the dimension d_R of R (Ref. 5) and introduce the integer q_R with $q_R=1$ if R is a self-contragradient representation and $q_R=2$ if it is not.

Then

$$d_{\rm sr} \geqslant q_R d_R .$$

The equality occurs only when T is irreducible.

Notice that, whereas the leading supertableau of a GAST is always nonirreducible by definition, the other atypical supertableaux participating to the GAST are either nonirreducible or irreducible. Such a feature plays a central role in the counting of the atypical components forming the associated non-fully-reducible representation.

Let us keep in mind an important result of the theory of Young tableaux for the orthogonal group O(m) (Ref. 6). In the restriction to unimodular orthogonal transformations O$(m) \Rightarrow$ SO(m) each irreducible representation of O(m) remains irreducible unless the corresponding Young tableau t is a self-associate tableau. The representations corresponding to different associate Young tableaux t and t' become equivalent and no other equivalence appears. When t is a self-associate Young tableau—this case occurring only when $m=2\nu$—the representation splits into two inequivalent representations of SO(2ν) of same dimension. We shall refer to this phenomena as the O$(2\nu) \Rightarrow$ SO(2ν) reduction and it will affect the supertableaux of OSP$(2\nu|2p)$. In particular this fact is at the origin of the factor q_R previously introduced.

The case $m=1$ is trivial and the Young supertableaux of the orthosymplectic group OSP$(1|2p)$ are all typical and irreducible.[1] The case $m=2$ has been considered in a previous paper[7] and we restrict ourselves now to orthosymplec-

[a)] Postal address; Université Pierre et Marie Curie, Tour 16–1er étage 4, place Jussieu, 75252 Paris Cedex 05, France.
[b)] Laboratoire associé au CNRS UA 280.

tic tableaux of OSP$(m|2p)$ with $m\geqslant 3$. With a particular emphasis to the case $p=2$, which might be relevant for supersymmetry and supergravity. The paper is made self-consistent by briefly discussing in Sec. II the main features of the orthosymplectic superalgebras and of the orthosymplectic supertableaux. Then come two sections with a detailed study of the orthosymplectic supertableaux of OSP$(m|4)$: a discussion of the irreductibility in Sec. III, a description of the general atypical supertableaux and of the atypical components of the corresponding non-fully-reducible representations in Sec. IV. A relation between the size and the degeneracy of atypically of supertableaux is proposed in Sec. V in the form of three theorems analogous to those previously obtained for the supertableaux of the superunitarily groups SU$(m|n)$ (Ref. 3).

II. GENERAL RESULTS ON ORTHOSYMPLECTIC SUPERTABLEAUX

A. Definitions

Let us call as ν the integer part of $m/2$. In Kac's notation[1] the superalgebras of the orthosymplectic groups are

$$m=2\nu+1,\quad B(\nu,p),\quad \nu\geqslant 0,\quad p\geqslant 1\,,$$
$$m=2\nu,\quad D(\nu,p),\quad \nu\geqslant 2,\quad p\geqslant 1\,,$$
$$m=2,\quad C(p+1),\quad \nu=1,\quad p\geqslant 1\,.$$

The supertableaux of the orthosymplectic group OSP$(2|2p)$ have been studied in a previous publication[7] and they will not be considered here. General results concerning the superalgebras $B(\nu,p)$ and $D(\nu,p)$ are brought in mind in Appendix A.

B. Legal orthosymplectic supertableaux

The supertableaux of the orthosymplectic group OSP$(m|2p)$ have p columns and ν rows of arbitrary length[2] as shown in Fig. 1. The row and column parameters are constrained by the usual inequalities of positivity

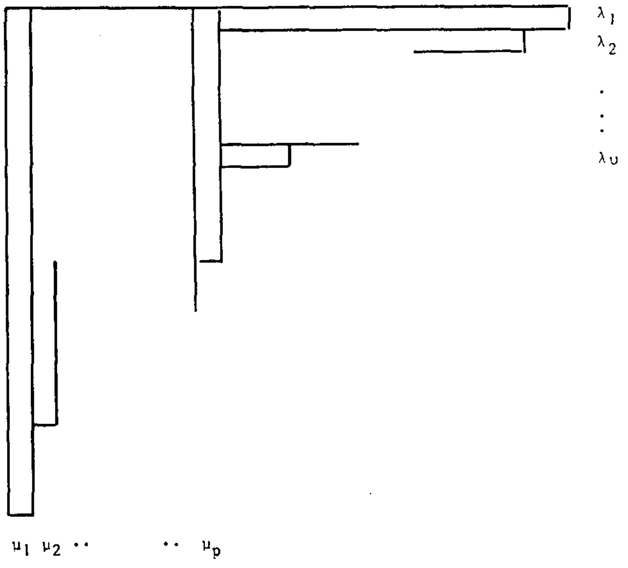


FIG. 1. General supertableau of OSP$(2\nu|2p)$ and OSP$(2\nu+1|2p)$.

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_2\geqslant 0,\quad \mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_p\geqslant 0\,.$$

As a consequence we have the legality conditions in one of the two equivalent forms: (i) the number of boxes of the $(\nu+1)$th row is less or equal to p, (ii) the number of boxes of the $(p+1)$th column is less or equal to ν.

C. Highest weight of an orthosymplectic supertableau ($m\neq 2$)

The highest weight $\Lambda_{\rm ST}$ of an orthosymplectic supertableau is defined by a set of Kac–Dynkin parameters[2]

$$a_j=\mu_j-\mu_{j+1},\quad j=1,2,\ldots,p-1\,,$$
$$a_p=\mu_p+\lambda_1\,,$$
$$a_{p+\gamma}=\lambda_\gamma-\lambda_{\gamma+1},\quad \gamma=1,2,\ldots,\nu-1\,,$$

and

$$a_{p+\nu}=\lambda_{\nu-1}+\lambda_\nu,\quad \text{for } D(\nu,p)\,,$$
$$a_{p+\nu}=2\lambda_\nu,\quad \text{for } B(\nu,p)\,.$$

The eigenvalues b of the hidden Sp$(2p)$ Cartan generator k is given by

$$b=\mu_p\,.$$

In this language the necessary consistency conditions of Kac[1] are trivial if

$$\mu_p=j\leqslant\nu,\quad \text{then } \lambda_{j+1}=\cdots=\lambda_\nu=0\,.$$

Let us point out that the correspondence between the $(p+\nu)$ supertableau parameters (λ,μ) and the $(p+\nu)$ Kac–Dynkin parameters of its highest weight $\Lambda_{\rm ST}$ is bijective.

D. Atypical supertableaux

The Kac–Dynkin parameter a_p of an irreducible representation R of OSP$(m|2p)$ can take $2\nu p$ possible atypical values which are linear functions of the other $p+\nu-1$ Kac–Dynkin parameters (see Appendix A),

$$a_p=A_{j\alpha}\,,$$
$$j=0,1,\ldots,p-1,\quad \alpha=0,1,\ldots,\nu-1\,.$$
$$a_p=B_{j\alpha}\,,$$

The highest weight $\Lambda_{\rm ST}$ of a supertableau is atypical when one or several relations between the row and the column parameters are satisfied:

$$\text{atypicity}\quad a_p=A_{j\alpha},\quad \mu_{p-j}+\lambda_{1+\alpha}=\alpha-j\,, \tag{1}$$
$$\text{atypicity}\quad a_p=B_{j\alpha},\quad \mu_{p-j}-\lambda_{1+\alpha}=m-2-j-\alpha\,, \tag{2}$$

with

$$0\leqslant\alpha\leqslant\nu-1,\quad 0\leqslant j\leqslant p-1\,.$$

Let us notice that in the case $m=2\nu$ when $\lambda_\nu=0$ the atypicities $A_{j\nu-1}$ and $B_{j\nu-1}$ become equal for every j as a consequence of the equality $a_{p-1+\nu}=a_{p+\nu}$.

The degeneracy of atypicity of a supertableau is the number of relations (1) and (2) which are fulfilled by the row and column parameters of the supertableau. Using the inequalities of positivity of the row and column parameters it is straightforward to derive an upper bound on the degeneracy of atypicity δ,

$$0 \leqslant \delta \leqslant L \tag{3}$$

with

$$L = \min[m-1, p] . \tag{4}$$

E. Comment on the atypical supertableau A_j

Using again positivity arguments it is straightforward to obtain the following results.

(a) The atypicities $A_{j\alpha}$ with $j > \alpha$ cannot be realized for the highest weight of a supertableau.

(b) The atypicities $A_{j\alpha}$ with $j \leqslant \alpha$ have only one possible realization for a supertableau

$$\lambda_{1+\alpha} = 0, \quad \mu_{p-j} = \alpha - j .$$

For instance the atypicity $A_{00} = 0$ corresponds to supertableaux with only $(p-1)$ columns and no row ($\mu_p = 0$). Of course such supertableaux may have, in addition, other atypicities of type A or B.

III. IRREDUCTIBILITY OF THE SUPERTABLEAUX OF OSP($m|4$)

(1) The consideration of the supertableaux of OSP($m|4$) with only two columns gives valuable information on the behavior of more complicated supertableaux. Let us first study the atypicity of the two column supertableaux. Using Eqs. (1) and (2) we obtain $2(m-1)$ relations of atypicities presented on Table I. Of course by positivity the atypicity A_{10} cannot be realized for a supertableau. We then obtain a first classification of the two-column supertableaux of OSP($m|4$) accordingly to the degeneracy of atypicity δ of their highest weight

(a) $\delta = 0$ typical, $\mu_1 \geqslant \mu_2 \geqslant m-1$,

(b) $\delta = 1$ atypical, $\mu_1 \geqslant m-2 \geqslant \mu_2$,

(c) $\delta = 2$ atypical, $m-3 \geqslant \mu_1 \geqslant \mu_2$.

The next problem of interest is that of the irreducibility of these supertableaux and we shall use for that purpose the language of supertensors introduced in Sec. I. The even bilinear form G being supersymmetric the removing of one trace, 2 traces, etc., for a two-column supertableau is realized in the following way:

one box in each column for one trace,

two boxes in each column for two traces,

⋮

n boxes in each column for n traces.

Of course such a truncation is possible if and only if $0 \leqslant n \leqslant \mu_2$.

As a direct consequence the one-column supertableaux of the orthosymplectic groups being associated to fully superantisymmetric tensors are irreducible.

TABLE I. Atypicity for the two-column supertableaux of OSP($m|4$).

A_{00} $\mu_2 = 0$	A_{01} $\mu_2 = 1$	A_{02} $\mu_2 = 2$	B_{01} $\mu_2 = m-3$	B_{00} $\mu_2 = m-2$
A_{10} $\mu_1 = -1$	A_{11} $\mu_1 = 0$	A_{12} $\mu_1 = 1$	B_{11} $\mu_1 = m-4$	B_{10} $\mu_1 = m-3$

An algebraic study of the reduction of a supertensor $\mathscr{T}$ of PL($m|2p$) under orthosymplectic transformations shows that the relevant parameter for the irreductibility is the superdimension $N = m - 2p$. In our case $p = 2$ we have $m = N - 4$. The result of our investigation of the two-column supertableaux $T(\mu_1,\mu_2)$ is the following. In the tensor space $\mathscr{T}(\mu_1,\mu_2)$ the term with n traces cannot be separated from the irreducible part when m takes the value

$$m = \mu_1 + \mu_2 + 3 - n . \tag{5}$$

The number of traces n being upper bounded by μ_2 we obtain, from Eq. (5), the characterization of the subset of nonirreducible two-column supertableaux of OSP($m|4$) as

$$\mu_1 + \mu_2 \geqslant m-2, \quad \mu_1 \leqslant m-3 . \tag{6}$$

From Table I these supertableaux have a degeneracy of atypicity $\delta = 2$. Let us call such a supertableau T_1. The truncation of n traces gives a supertableau T_2 simply related to T_1:

$$T_1(\mu_1,\mu_2), \quad T_2(\mu_1 - n,\ \mu_2 - n) .$$

Let us point out that, for a given supertableau of the subset (6) and a fixed value of m, Eq. (5) produces only one solution for n and the pair (T_1,T_2) called a two-generalized atypical supertableau is then uniquely defined by T_1. We observe that T_2 has also a degeneracy of atypicity $\delta = 2$ but contrary to T_1, T_2 does not belong to the subset (6).

The two-GAST (T_1,T_2) describes a non-fully-reducible representation of OSP($m|4$) with three atypical components given by

$$\{\mu_1 - \mu_2|\mu_2|0,...,0\} + 2\{\mu_1 - \mu_2|\mu_2 - n|0,...,0\}.$$

The results concerning the degeneracy of atypicity and the irreductibility of the two-column supertableaux of OSP($m|4$) are summarized in Table II, where the empty squares are trivial consequences of the positivity constraint $\mu_1 \geqslant \mu_2$.

(2) As pointed out in Sec. I a supertableau is either irreducible or nonirreducible. In the later case it is the leading supertableau of a set of ρ atypical supertableaux of same degeneracy of atypicity δ forming a ρ-GAST. It is trivial to check, by elementary counting, that ρ is non-negative integer power of 2. At fixed value of δ by positivity the maximal value of ρ is $\rho_M = 2^\delta$. We can distinguish two cases: (a) the normal case where $\rho = \rho_M$, and (b) the abnormal case

TABLE II. Irreductibility and nonirreductibility of the two-column supertableaux of OSP($m|4$).

μ_1 \ μ_2	$A_{0\alpha}$	$B_{0\alpha}$	Typical
$A_{1\beta}, \beta > \alpha$	$\delta = 2$ IRR		
$B_{1\beta}$ $\quad \beta \geqslant \alpha$	$\delta = 2$ IRR		
$B_{1\beta}$ $\quad \beta < \alpha$	$\delta = 2$ two-GAST	$\delta = 2$ two-GAST	
Typical	$\delta = 1$ IRR	$\delta = 1$ IRR	$\delta = 0$ IRR

where $\rho < \rho_M$. Of course in this language a one-GAST is an irreducible supertableau. The typical supertableaux, $\delta = 0$, are irreducible, $\rho = 1$, and normal. The zero-box and the one-box supertableaux are atypical, irreducible, and abnormal.

In order to define the size of a supertableau we make a partition of the full set S_0 of legal supertableau of OSP$(m|2p)$ in $L + 1$ classes

$$S_0 = \bigoplus_{l=0}^{l=L} \Delta_l .$$

The typical and the atypical supertableaux of the normal case belong to the class Δ_0. An atypical supertableau of the abnormal case with the two associated parameters δ and ρ belongs to the class Δ_l with the positive integer l defined by

$$2^l = \rho_M / \rho$$

or, equivalently

$$\rho = 2^{(\delta - l)} . \tag{7}$$

(3) Coming back to OSP$(m|4)$ we easily see that the set of atypical two-column supertableaux considered in the first paragraph of this section belongs to the abnormal case, the typical two-column supertableaux being of course normal. In the normal case we have three possibilities:

typical $\delta = 0$ irreducible $\rho = 1$,

atypical $\delta = 1$ two-GAST $\rho = 2$,

atypical $\delta = 2$ four-GAST $\rho = 4$.

For the abnormal case we have again three possibilities:

atypical $\delta = 1$ irreducible $\rho = 1$,

atypical $\delta = 2$ irreducible $\rho = 1$,

atypical $\delta = 2$ two-GAST $\rho = 2$.

The study of the properties of a general supertableau of OSP$(m|4)$ with two columns and at most ν rows is conveniently made starting from the associated two-column supertableau of same μ_1 and μ_2. The detailed discussion is given in Appendix B and the results are summarized in Table III.

For the orthosymplectic groups OSP$(m|4)$ with $m \geqslant 3$ we have three classes of supertableaux Δ_0, Δ_1, Δ_2. Using the results of Appendix B we obtain a simple characterization of

TABLE III. Irreductibility and nonirreductibility of the supertableaux of OSP$(m|4)$.

μ_1 \ μ_2	$A_{0\alpha}$	$B_{0\alpha}$ $\lambda_{1+\alpha} = 0$	$B_{0\alpha}$ $\lambda_{1+\alpha} \neq 0$	Typical
$A_{0\beta}, \beta > \alpha$	$\delta = 2$ IRR			
$\beta \geqslant \alpha$	$\delta = 2$ IRR			
$B_{1\beta}$ $\lambda_{1+\beta} = 0$ $\beta < \alpha$	$\delta = 2$ two-GAST	$\delta = 2$ two-GAST		
$B_{1\beta}$ $\lambda_{1+\beta} \neq 0$	$\delta = 2$ two-GAST	$\delta = 2$ two-GAST	$\delta = 2$ four-GAST	$\delta = 1$ two-GAST
Typical	$\delta = 1$ IRR	$\delta = 1$ IRR	$\delta = 1$ two-GAST	$\delta = 0$ IRR

these classes with the three column parameters $c_1 = \mu_1$, $c_2 = \mu_2$, and c_3, which is the number of nonvanishing rows. The result is

$$\begin{aligned} &\text{class } \Delta_0\text{: } c_2 + c_3 \geqslant m - 1 , \\ &\text{class } \Delta_1\text{: } c_1 + c_2 \geqslant m - 2 \geqslant c_2 + c_2 , \\ &\text{class } \Delta_2\text{: } c_1 + c_2 \leqslant m - 3 . \end{aligned} \tag{8}$$

A technical remark is now in order. Because of the positivity we have

$$c_1 \geqslant c_2, \quad c_3 \leqslant \min[c_2, \nu] ,$$

and in the $X = c_1 + c_2$, $Y = c_2 + c_3$ plane the allowed domain for legal supertableaux is limited by three straight lines and given by

$$Y \geqslant 0, \quad Y \leqslant X, \quad Y \leqslant \nu + X/2 .$$

We have represented on Figs. 2 and 3 the diagram of classes for the two cases $m = 2\nu$ and $m = 2\nu + 1$.

IV. GENERALIZED ATYPICAL SUPERTABLEAUX OF OSP$(m|4)$

When the atypical supertableau T_1 is not irreducible it is the leading supertableau of a generalized atypical supertableau describing a non-fully-reducible representation of the orthosymplectic group. The partners of T_1 in the GAST and the atypical components of the associated non-fully-reducible representation are completely determined by T_1.

We now discuss in details these problems for the simplest case of two-GAST of OSP$(m|4)$. The analysis of four-GAST is only more complicated in the details but non-fundamentally-different. The various situations are those described in Table III.

(1) The supertableau T_1 belongs to the class Δ_1 and its highest weight has a degeneracy of atypicity $\delta = 2$. The μ_2 atypicity is $A_{0\alpha}$ $(1 \leqslant \mu_2 = \alpha \leqslant \nu - 1)$ or $B_{0\alpha}$ $(\nu \leqslant \mu_2 = m - 2 - \alpha \leqslant m - 2)$ with $\lambda_{1+\alpha} = 0$. The μ_1 atypicity is $B_{1\beta}$ $(0 \leqslant \beta < \alpha)$. We then have

$$\mu_1 = m - 3 - \beta + \lambda_{1+\beta} .$$

It is convenient to study separately the two cases $\lambda_{1+\beta} \geqslant 1$ and $\lambda_{1+\beta} = 0$,

$$a - \lambda_{1+\beta} \geqslant 1.$$

Let us define as σ the largest non-negative integer such that

$$\lambda_{1+\beta} = \cdots = \lambda_{1+\beta+\sigma}, \quad 0 \leqslant \sigma \leqslant \alpha - 1 - \beta .$$

The supertableau T_2 is deduced from T_1 by suppressing $2 + 2\sigma$ boxes as follows:

$$T_2 \begin{cases} \tilde{\mu}_1 = \mu_1 - 1 - \sigma, \quad \tilde{\mu}_2 = \mu_2 , \\ \tilde{\lambda}_j = \lambda_j, \quad \text{for } 1 \leqslant j \leqslant \beta, \quad 2 + \beta + \sigma \leqslant j \leqslant \nu , \\ \tilde{\lambda}_j = \lambda_j - 1, \quad \text{for } 1 + \beta \leqslant j \leqslant 1 + \beta + \sigma . \end{cases}$$

The new column parameter $\tilde{\mu}_1$ can be rewritten as

$$\begin{aligned} \tilde{\mu}_1 &= m - 3 - \beta + \lambda_{1+\beta} - 1 - \sigma \\ &= m - 3 - (\beta + \sigma) + \tilde{\lambda}_{1+\beta+\sigma} , \end{aligned}$$

and the supertableau T_2 has the $\tilde{\mu}_1$ atypicity $B_{1\beta+\sigma}$ $(\beta + \sigma < \alpha)$ and the same μ_2 atypicity as T_1. Therefore the supertableau T_2 has a degeneracy of atypicity $\delta = 2$ and it belongs, like T_1, to the class Δ_1,

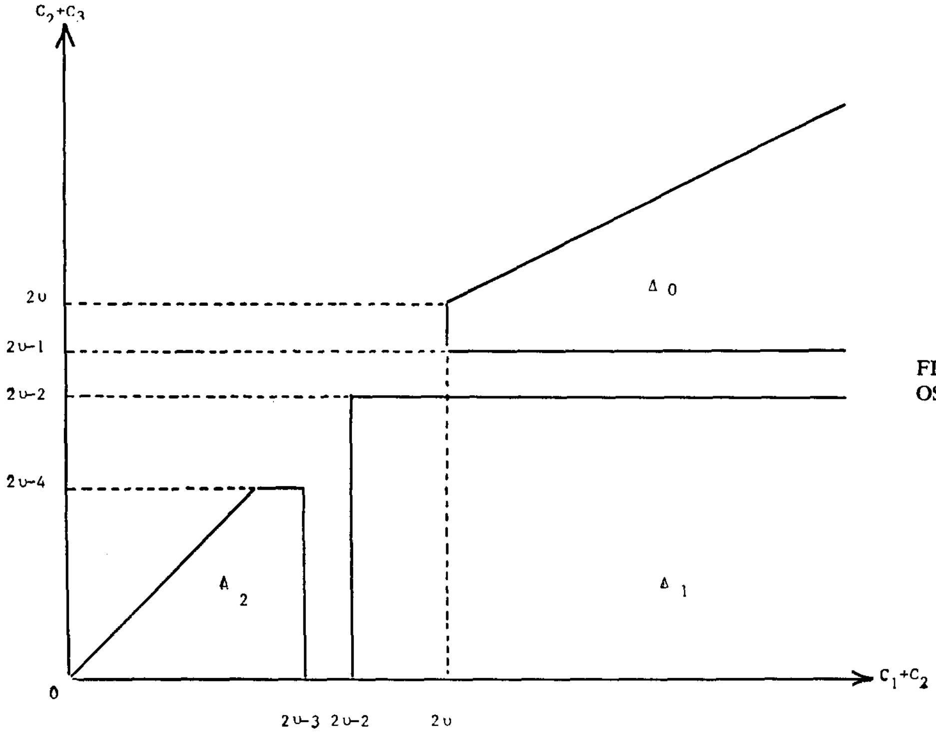


FIG. 2. Partition in classes of the supertableaux of OSP($2\nu|4$).

$$b-\lambda_{1+\beta}=0.$$

The supertableau T_1 has at most β nonvanishing rows that do not play any role in the μ_1 and μ_2 atypicities. We then introduce the two-column supertableau T_1^0 obtained from T_1 by suppressing the possibly nonvanishing β rows. Like T_1 the supertableau T_1^0 has $\delta=2$ and it belongs to the class Δ_1. We have a two-GAST (T_1^0, T_2^0) and the supertableau T_2^0 has been determined in Sec. III,

$$\tilde{\mu}_1=\mu_1-n=m-3-\mu_2,\quad \tilde{\mu}_2=\mu_2-n=m-3-\mu_1.$$

The supertableau T_2 forming with T_1 a two-GAST is now obtained from T_2^0 by reintroducing the possibly nonvanishing β rows of T_1. We get

$$T_2\begin{cases}\tilde{\mu}_1=m-3-\mu_2, & \tilde{\mu}_2=m-3-\mu_1=\beta,\\ \tilde{\lambda}_j=\lambda_j, & \text{for } 1\leqslant j\leqslant\beta.\end{cases}$$

Like T_2^0 the supertableau T_2 has $\delta=2$ and belongs to the class Δ_2. It is irreducible.

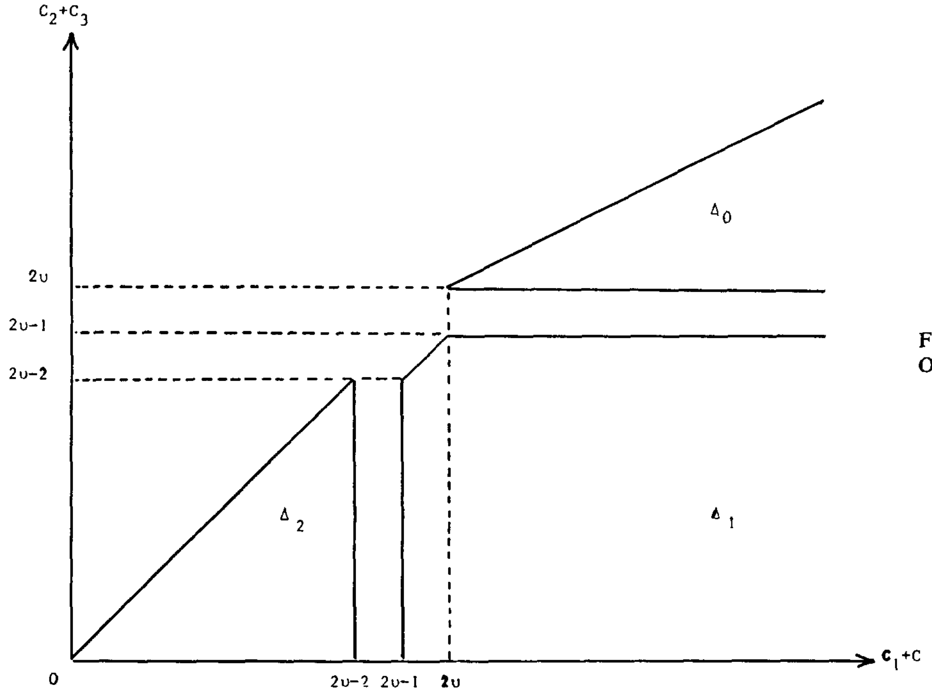


FIG. 3. Partition in classes of the supertableaux of OSP($2\nu+1|4$).

(2) The supertableau T_1 belongs to the class Δ_0 and its highest weight has a μ_2 atypicity $\delta = 1$. This μ_2 atypicity is $B_{0\alpha}$ $(0 \leqslant \alpha \leqslant \nu - 1)$ with $\lambda_{1+\alpha} \geqslant 1$ and we have

$$\mu_2 = m - 2 - \alpha + \lambda_{1+\alpha} .$$

As previously we define as σ the largest non-negative integer such that

$$\lambda_{1+\alpha} = \cdots = \lambda_{1+\alpha+\sigma}, \quad 0 \leqslant \sigma \leqslant \nu - 1 - \alpha .$$

The supertableau T_2 is deduced from T_1 by suppressing $2 + 2\sigma$ boxes as follows:

$$T_2 \begin{cases} \tilde{\mu}_1 = \mu_1, \quad \tilde{\mu}_2 = \mu_2 - 1 - \sigma , \\ \tilde{\lambda}_j = \lambda_j, \quad \text{for } 1 \leqslant j \leqslant \alpha, \quad 2 + \alpha + \sigma \leqslant j \leqslant \nu , \\ \tilde{\lambda}_j = \lambda_j - 1, \quad \text{for } 1 + \alpha \leqslant j \leqslant 1 + \alpha + \sigma . \end{cases}$$

The new column parameters $\tilde{\mu}_2$ can be rewritten as

$$\begin{aligned} \tilde{\mu}_2 &= \mu_2 - 1 - \sigma = m - 2 - \alpha + \lambda_{1+\alpha} - 1 - \sigma \\ &= m - 2 - (\alpha + \sigma) + \tilde{\lambda}_{1+\alpha+\sigma} , \end{aligned}$$

and the supertableau T_2 has the $\tilde{\mu}_2$ atypicity $B_{0\alpha+\sigma}$ and no μ_1 atypicity like T_1. When $\tilde{\lambda}_{1+\alpha+\sigma} \geqslant 1$ or $\lambda_{1+\alpha} \geqslant 2$ the supertableau T_2 is irreducible and it belongs to the class Δ_1.

Remark: When $m = 2\nu$, $\sigma = \nu - 1 - \alpha$, and $\lambda_{1+\alpha} = 1$ the atypicity $B_{0\nu-1}, \tilde{\lambda}_\nu = 0$ becomes $A_{0\nu-1}$.

(3) The supertableau T_1 belongs to the class Δ_0 and its highest weight has a μ_1-atypicity $\delta = 1$. The μ_1-atypicity is $B_{1\alpha}$ with $\lambda_{1+\alpha} \geqslant 1$ and we have

$$\mu_1 = m - 3 - \alpha + \lambda_{1+\alpha} .$$

The column parameter $\mu_2 \leqslant \mu_1$ being typical, at fixed μ_1 and λ's, it can only take $\lambda_{1+\alpha}$ typical values located in the range $\nu \leqslant \mu_2 \leqslant \mu_1$. Let us now define as σ the largest non-negative integer such that

$$\lambda_{1+\alpha} = \cdots = \lambda_{1+\alpha+\sigma}, \quad 0 \leqslant \sigma \leqslant \nu - 1 - \alpha .$$

The largest allowed typical value of μ_2 is then

$$(\mu_2)_{\max} = m - 3 - \alpha - \sigma + \lambda_{1+\alpha}$$

and it corresponds to the minimal value of the difference $\tau = \mu_1 - \mu_2$,

$$(\tau)_{\min} = \sigma .$$

When $\lambda_{1+\alpha} \geqslant 2$ we have other typical values of μ_2 for which

$$\tau \geqslant 1 + \sigma .$$

It is now convenient to distinguish the two cases $\tau \geqslant 1 + \sigma$ and $\tau = \sigma$,

$$a - \tau \geqslant 1 + \sigma \quad \text{implies } \lambda_{1+\alpha} \geqslant 2 .$$

The supertableau T_2 is obtained from T_1 by suppressing $2 + 2\sigma$ boxes as follows:

$$T_2 \begin{cases} \tilde{\mu}_1 = \mu_1 - 1 - \sigma, \quad \tilde{\mu}_2 = \mu_2 , \\ \tilde{\lambda}_j = \lambda_j, \quad \text{for } 1 \leqslant j \leqslant \alpha, \quad 2 + \alpha + \sigma \leqslant j \leqslant \nu , \\ \tilde{\lambda}_j = \lambda_j - 1, \quad \text{for } 1 + \alpha \leqslant j \leqslant 1 + \alpha + \sigma . \end{cases}$$

The new column parameter $\tilde{\mu}_1$ can be rewritten as

$$\begin{aligned} \tilde{\mu}_1 &= m - 3 - \alpha + \lambda_{1+\alpha} - 1 - \sigma \\ &= m - 3 - (\alpha + \sigma) + \tilde{\lambda}_{1+\alpha+\sigma} \end{aligned}$$

and the supertableau T_2 has the μ_1-atypicity $B_{1\alpha+\sigma}$ with $\tilde{\lambda}_{1+\alpha+\sigma} \geqslant 1$ and no μ_2-atypicity. Like T_1 the supertableau T_2 has $\delta = 1$ and it belongs to the class Δ_0,

$$b - \tau = \sigma .$$

The two-column parameters μ_1 and μ_2 of the supertableau T_1 are given by

$$\mu_1 = m - 3 - \alpha + \lambda_{1+\alpha}, \quad \text{atypical} ,$$

$$\mu_2 = m - 3 - \alpha - \sigma + \lambda_{1+\alpha+\sigma}, \quad \text{typical} ,$$

and we have, in addition at least $1 + \alpha + \sigma$ nonvanishing rows. It is convenient to study separately the two cases $\lambda_{1+\alpha} \geqslant 2$ and $\lambda_{1+\alpha} = 1$,

$$b - 1, \quad \tau = \sigma, \quad \lambda_{1+\alpha} \geqslant 2 .$$

We define as σ_1 the smallest non-negative integer such that

$$\lambda_{1+\alpha} - \lambda_{2+\alpha+\sigma+\sigma_1} \geqslant 2, \quad 0 \leqslant \sigma + \sigma_1 \leqslant \nu - 1 - \alpha .$$

The supertableau T_2 is obtained from T_1 by suppressing $2 + 2\sigma + 2\sigma_1$ boxes as follows:

$$T_2 \begin{cases} \tilde{\mu}_1 = \mu_1 - 1 - \sigma , \\ \tilde{\mu}_2 = \mu_2 - 1 - \sigma_1 , \\ \tilde{\lambda}_j = \lambda_j, \quad \text{for } 1 \leqslant j \leqslant \alpha, \quad 2 + \alpha + \sigma + \sigma_1 \leqslant j \leqslant \nu , \\ \tilde{\lambda}_j = \lambda_j - 1, \quad \text{for } 1 + \alpha \leqslant j \leqslant \alpha + \sigma, \\ \qquad\qquad 2 + \alpha + \sigma \leqslant j \leqslant 1 + \alpha + \sigma + \sigma_1, \\ \tilde{\lambda}_j = \lambda_j - 2, \quad \text{for } j = 1 + \alpha + \sigma . \end{cases}$$

The two-column parameters of the supertableau T_2 are now rewritten in the form

$$\begin{aligned} \tilde{\mu}_1 &= m - 3 - \alpha + \lambda_{1+\alpha} - 1 - \sigma \\ &= m - 3 - (\alpha + \sigma) + \tilde{\lambda}_{1+\alpha+\sigma} + 1 , \\ \mu_2 &= m - 3 - \alpha - \sigma + \lambda_{1+\alpha+\sigma} - 1 - \sigma_1 \\ &= m - 2 - (\alpha + \sigma + \sigma_1) + \tilde{\lambda}_{1+\alpha+\sigma+\sigma_2} . \end{aligned}$$

The value of $\tilde{\mu}_1$ is typical and the value of $\tilde{\mu}_2$ is atypical $B_{0\alpha+\sigma+\sigma_1}$. The supertableau T_2 has again $\delta = 1$. When $\tilde{\lambda}_{1+\alpha+\sigma+\sigma_1} \geqslant 1$ or $\lambda_{1+\alpha} \geqslant 3$ the supertableau T_2 like T_1 belongs to the class Δ_0. When $\tilde{\lambda}_{1+\alpha+\sigma+\sigma_1} = 0$ or $\lambda_{1+\alpha} = 2$ the supertableau T_2 belongs to the class Δ_1 and it is irreducible.

Remark: When $m = 2\nu$, $\sigma + \sigma_1 + \alpha = \nu - 1$, and $\lambda_{1+\alpha} = 2$ the atypicity $B_{0\nu-1}, \tilde{\lambda}_\nu = 0$ reduces to $A_{0\nu-1}$,

$$b - 2, \quad \tau = \sigma, \quad \lambda_{1+\alpha} = 1 .$$

The two-column parameters μ_1 and μ_2 of the supertableau T_1 are given by

$$\mu_1 = m - 2 - \alpha, \quad \text{atypical } B_{1\alpha} ,$$

$$\mu_2 = m - 2 - (\alpha + \sigma), \quad \text{typical} ,$$

and we have, in addition $1 + \alpha + \sigma$ rows, $1 + \sigma$ of which at least heaving only one box. Using Kac's consistency condition $\mu_2 \geqslant 1 + \alpha + \sigma$ we derive an upper bound for the sum $\alpha + \sigma$,

$$\alpha + \sigma \leqslant m - \nu - 2 .$$

As a consequence for $m = 2\nu$ the maximal value of nonvanishing rows is $\nu - 1$ and therefore $\lambda_\nu = 0$.

The supertableau T_2 is obtained from T_1 by suppressing $2 + 2\sigma$ boxes as follows:

$$T_2\begin{cases}\tilde{\mu}_1=\mu_1, & \tilde{\mu}_2=\mu_2-1-\sigma,\\ \tilde{\lambda}_j=\lambda_j, & 1\leqslant j\leqslant\alpha,\\ \tilde{\lambda}_j=0, & 1+\alpha\leqslant j\leqslant\nu.\end{cases}$$

The value of $\tilde{\mu}_1$ becomes typical and that of $\tilde{\mu}_2$ atypical. Rewriting $\tilde{\mu}_2$ in a more transparent way we get

$$\tilde{\mu}_2=m-2-(\alpha+\sigma)-1-\sigma$$
$$=m-(3+\alpha+2\sigma)\geqslant m-3-\alpha.$$

The parameter $\tilde{\mu}_2$ is atypical either of the $A_{0\beta}$ type or the $B_{0\beta}$ type with $\beta\geqslant\alpha+1$. Of course the supertableau T_2 is irreducible and it belongs to the class Δ_1.

(4) We now study the atypical components of the nonfully-reducible representations associated to the generalized atypical supertableaux.

Let us call R_j the irreducible representation of the orthosymplectic group whose highest weight is the same as the highest weight of the supertableau T_j. We have two possibilities for R_j: $\alpha-R_j$ is self-contragradient; or β-R_j is not self-contragradient and its contragradient is $\overline{R}_j$. The tensor representations of the orthosymplectic group $OSP(2\nu+1|2p)$ are self-contragradient. For the orthosymplectic group $OSP(2\nu|2p)$ only the representations with $a_{p+\nu-1}=a_{p+\nu}$ or, in the supertableau language, $\lambda_\nu=0$ are self-contragradient. If $a_{p+\nu-1}\neq a_{p+\nu}$ the representation $\overline{R}_j$ is obtained from R_j by exchanging only the two Kac–Dynkin parameters $a_{p+\nu-1}$ and $a_{p+\nu}$.

As an illustration consider a typical irreducible supertableau T. If the typical irreducible representation R is self-contragradient it is uniquely associated to T and we have

$$\dim T=\dim R.$$

If R is not self-contragradient then T describes the direct sum of two contragradient typical irreducible representations $R\oplus\overline{R}$ of same dimension and we have

$$\dim T=2\dim R.$$

Let us come back now to the case where T is atypical and nonirreducible. The discussion of the atypical components of the nonfully reducible representation of the orthosymplectic group described by a GAST can be formulated in a rather general way. Consider an atypical supertableau T_1 belonging to the class Δ_l with a degeneracy of atypicity $\delta=l+1$. We have a two-GAST (T_1,T_2) and the following possible cases may occur for T_2:

(α) T_2 belongs to the class Δ_l,

(β) T_2 belongs to the class Δ_{l+1}.

In the first case T_2 can be considered as the leading supertableau of a second two-GAST (T_2,T_3). In the second case T_2 is irreducible.

We are now in position to give the results concerning the atypical components of the non-fully-reducible representation described by the two-GAST (T_1,T_2).

(a) $T_2\in\Delta_l$ the supertableau T_3 exists.

(a-1) $R_1\neq\overline{R}_1$ then $R_2\neq\overline{R}_2$,

$$(T_1,T_2)\Rightarrow(R_1+2R_2+R_3)\oplus(\overline{R}_1+2\overline{R}_2+\overline{R}_3),$$

the representation R_3 may or may not be self-contragradient.

(a-2) $R_1=\overline{R}_1$ then $R_2=\overline{R}_2$ and $R_3=\overline{R}_3$,

$$(T_1,T_2)\Rightarrow(R_1+2R_2+R_3).$$

(b) $T_2\in\Delta_{l+1}$ T_2 is irreducible.

(b-1) $R_1\neq\overline{R}_1$ with $R_2=\overline{R}_2$,

$$(T_1,T_2)\Rightarrow(R_1+2R_2+\overline{R}_1).$$

(b-2) $R_1=\overline{R}_1$ with $R_2=\overline{R}_2$,

$$(T_1,T_2)\Rightarrow R_1+2R_2.$$

The number of atypical components of the non-fully-reducible representation described by the two-GAST (T_1,T_2) varies between 3 and 2×4.

V. SIZE AND ATYPICITY FOR ORTHOSYMPLECTIC SUPERTABLEAUX

(1) In this section we intend to make *conjectures* for an extension to $OSP(m|2p)$ of the results obtained in Sec. III for $OSP(m|4)$. The size of a legal supertableau of $OSP(m|2p)$ is determined by $p+1$ quantities, the column parameters $c_j=\mu_j$, $j=1,2,\ldots,p$, and c_{p+1} measuring the number of nonvanishing rows, $c_{p+1}\leqslant\nu$. The natural extension of Eqs. (8) is the following:

$$\begin{aligned}&\text{class }\Delta_0:\quad c_p+c_{p+1}\geqslant m-1,\\ &\text{class }\Delta_l:\quad c_{p-1}+c_{p-l+1}\geqslant m-1-l\\ &\qquad\qquad\geqslant c_{p-l+1}+c_{p-l+2},\\ &\qquad\qquad 1\leqslant l\leqslant L-1,\\ &\text{class }\Delta_L:\quad c_1+c_2\leqslant m-1-p\quad\text{if } p<m-1,\\ &\qquad\qquad c_{p+2-m}=0,\quad\text{if } p\geqslant m-1.\end{aligned}\tag{9}$$

(2) By analogy with the case of superunitary supertableaux[3] we formulate the relation between the size and the atypicity in the form of three theorems.

Theorem I: When a legal supertableau T belongs to the class Δ_l, $0<l\leqslant L$, then the degeneracy of atypicity δ of its highest weight Λ_{ST} is lower bounded by l and we get

$$l\leqslant\delta\leqslant L.$$

Two trivial consequences of this theorem are (i) the typical supertableaux all belong to the class Δ_0, (ii) the atypical supertableaux of the class Δ_L have $\delta=L$.

Theorem II: When a legal supertableau T belongs to the class Δ_l, $0<l\leqslant L$ and has a degeneracy of atypicity $\delta=l$ then T is irreducible. Calling as R the irreducible representation of highest weight Λ_{ST} we have the correspondence (i) for R self-contragradient $T\Rightarrow R$, and (ii) for R not self-contragradient $T\Rightarrow R\oplus\overline{R}$.

Theorem III: When a legal supertableau T belongs to the class Δ_l, $0<l<L$, and has a degeneracy of atypicity δ, $l<\delta\leqslant L$, then T is nonirreducible and it is the leading supertableau of a ρ-GAST given by Eq. (7).

The ρ-GAST describes a non-fully-reducible representation of $OSP(m|2p)$.

(3) Comment on the theorems: Theorem I has been obtained in Appendix C for orthosymplectic supertableaux of $OSP(m|2p)$ with p columns only. Its extension to more complete supertableaux with at most ν nonvanishing rows can be made as it has been done in Sec. II for $p=2$. The general proof is straightforward although a little tedious.

On the other hand for Theorems II and III we only have plausibility arguments and examples of particular situations where these theorems are satisfied,
$m=1,\ \ p\geqslant 1,\ \ L=0,\ \ \mathrm{OSP}(1|2p)$.

One class Δ_0 of irreducible supertableaux
$m=2,\ \ p\geqslant 1,\ \ L=1,\ \ \mathrm{OSP}(2|2p)$.

Two classes Δ_0, Δ_1 of supertableaux already studied in (Ref. 7):

class Δ_1: $C_p=0$, irreducible atypical supertableaux,

class Δ_0: $C_p\geqslant 1$, irreducible typical supertableaux and two-generalized atypical supertableaux,

$m\geqslant 3,\ p=2,\ L=2,\ \mathrm{OSP}(m|4)$.

Three classes of supertableaux studied in Sec. III.

VI. CONCLUDING REMARKS

The exhaustive study of the supertableaux of the orthosymplectic groups $\mathrm{OSP}(2|2p)$ (Ref. 6) and $\mathrm{OSP}(m|4)$ make plausible the conjectures proposed in Sec. V concerning the classification and the interpretation of any legal orthosymplectic supertableau. We notice the similarity of the results obtained in the superunitary case[3] and in the orthosymplectic one. In fact the partition in classes of the set of legal orthosymplectic supertableaux made in Sec. III was designed for that purpose and in this respect the key equations are Eqs. (8) and (9).

Let us now end this paper with a remark concerning the topology of the set S_0 of legal supertableaux of the orthosymplectic group $\mathrm{OSP}(m|2p)$. Let us call $\mathscr{N}$ the total number of boxes of a supertableau $T\in S_0$:

$$\mathscr{N}=\sum_{1}^{p}\mu_j+\sum_{1}^{\nu}\lambda_\alpha .$$

Such a quantity is obviously conserved mod 2 by tensor product. As a consequence the set S_0 has a Z_2 structure with two classes S_0^+ and S_0^- whose direct sum is S_0:

$$S_0=S_0^+\oplus S_0^-$$

and

$$\mathscr{N}\equiv 0(2)\quad \text{if } T\in S_0^+ ,$$
$$\mathscr{N}\equiv 1(2)\quad \text{if } T\in S_0^- .$$

Only the subset S_0^+ is closed under the tensor product operation and it is generated by the superantisymmetric two-box supertableau describing the adjoint representation. Of course S_0^+ contains the zero-box supertableau of the scalar representation. The subset S_0^- is not a subgroup and it contains, in particular, the one-box supertableau of the fundamental representation.

The Z_2 structure obviously extends to the generalized atypical supertableaux where two atypical supertableaux entering in the GAST always differ by an even number of boxes.

Such a topology has already been pointed out for the supertableaux of $\mathrm{OSP}(2|2p)$.

ACKNOWLEDGMENT

The author is happy to thank Jean Thierry-Mieg for communicating his tables prior to publication.

APPENDIX A: SUPERALGEBRAS $B(\nu,p)$ and $D(\nu,p)$

Definitions: The Z_2 graded algebras $B(\nu,p)$ and $D(\nu,p)$ have the decomposition $L=L_0\oplus L_1$. The Bose sector L_0 of the even generators is simple for $B(0,p)$ and semisimple in all the other cases

$$L_0=B_\nu\oplus C_p\quad \text{for } B(\nu,p),\quad \nu\geqslant 1,\quad p\geqslant 1 ,$$
$$L_0=D_\nu\oplus C_p\quad \text{for } D(\nu,p),\quad \nu\geqslant 2,\quad p\geqslant 1 ,$$

where B_ν, D_ν, and C_p are the familiar Cartan simple Lie algebras. The dimension of L_0 is

$$\dim L_0=p(2p+1)+\tfrac{1}{2}m(m-1) .$$

The Fermi sector L_1 of the odd generators is irreducible and it has the dimension

$$\dim L_1=2mp .$$

The orthosymplectic superalgebras $B(\nu,p)$ and $D(\nu,p)$ are basic, simple graded Lie algebras and the Cartan subalgebra H of mutually commuting generators is the maximal Abelian subalgebra of L_0 and it has the dimension $p+\nu$.

Canonical basis: It is possible to extract a system of $p+\nu$ simple positive generators α_i^+ and their conjugate α_i^- involving only one odd generator $\alpha_p^+=\beta^+$ and its conjugate $\alpha_p^-=\beta^-$ and such that

(1) $[\alpha_i^\pm,\alpha_j^\pm]=\pm\delta_{ij}h_i,\quad h_i\in H$

(of course the commutator becomes an anticommutator when $i=j=p$, $\{\beta^+,\beta^-\}=h_\beta,\quad h_\beta\in H$);

(2) $[h_i,\alpha_j^\pm]=\pm a_{ij}\alpha_j^\pm$,

where a_{ij} is the Cartan matrix. Let us notice that the diagonal elements of this matrix but a_{pp} are $a_{ii}=2$ and $a_{pp}=0$ excepted for $B(0,p)$.

Irreducible representations: The highest weight Λ of a finite-dimensional irreducible representation is annihilated by any simple positive generator

$$\alpha_i^+|\Lambda\rangle=0,\quad i=1,2,\ldots,p+\nu .$$

The Kac–Dynkin parameters are the eigenvalues of the Cartan generators h_i for the highest weight Λ,

$$h_i|\Lambda\rangle=a_i|\Lambda\rangle,\quad i=1,2,\ldots,p+\nu ,$$

and we shall use the following notation of an irreducible representation of the superalgebra:

$$\Lambda\Rightarrow\{a_1,\ldots,a_{p-1}|a_p|a_{p+1},\ldots,a_{p+\nu}\} .$$

The parameters $a_1,\ldots,a_{p-1}$ are the Dynkin parameters of a $\mathrm{SU}(p)$ subgroup of $\mathrm{Sp}(2p)$ and consequently they are $p-1$ non-negative integers. The parameters $a_{p+1},\ldots,a_{p+\nu}$ are the Dynkin parameters of the $\mathrm{SO}(m)$ group and they are ν non-negative integers.

The Cartan generator h_β corresponding to the simple odd generators $\beta^\pm$ combines orthogonal and symplectic Cartan generators. The hidden $\mathrm{Sp}(2p)$ Cartan generator k is related to the h_i's by

$$k=h_p-\sum_1^{\nu-1} h_{p+\gamma}-\frac{1}{2}h_{p+\nu} \quad \text{for } B(\nu,p)\,,$$

$$k=h_p-\sum_1^{\nu-2} h_{p+\gamma}-\frac{1}{2}(h_{p+\nu-1}+h_{p+\nu})$$

for $D(\nu,p)$.

The eigenvalue b of k for the highest weight Λ is also a non-negative integer and we get

$$b=a_p-\sum_1^{\nu-1} a_{p+\gamma}-\frac{1}{2}a_{p+2} \quad \text{for } B(\nu,p)\,,$$

$$b=a_p-\sum_1^{\nu-2} a_{p+\gamma}-\frac{1}{2}(a_{p+\nu-1}+a_{p+\nu})$$

for $D(\nu,p)$.

The parameters a_i—but a_p—and b being non-negative integers it follows that the Kac–Dynkin parameter a_p is either a non-negative integer or a positive half-integer. We then define two types of representations:

(1) the tensor representations where a_p is integer

$a_{p+\nu}$ is even for $B(\nu,p)$,

$a_{p+\nu-1}+a_{p+\nu}$ is even for $D(\nu,p)$;

(2) the spinor representations where a_p is half-integer

$a_{p+\nu}$ is odd for $B(\nu,p)$,

$a_{p+\nu-1}+a_{p+\nu}$ is odd for $D(\nu,p)$.

Of course only the subset of tensor representations is closed under the operation of tensor product and we have a Z_2 structure in the set of representations. Notice that the tensor representations are the only one realized in supertableaux.

Atypical representations: For specific values of the Kac–Dynkin parameter a_p the irreducible representations become atypical. We have $2\nu p$ such values[1]

$$a_p=A_{j\alpha}\,,\quad A_p=B_{j\alpha}\,,$$

with $j=0,1,\ldots,p-1$ and $\alpha=0,1\ldots,\nu-1$. The expression $A_{j\alpha}$ and $B_{j\alpha}$ are the following:

$$A_{j\alpha}=\sum_1^{\alpha}(1+a_{p+\gamma})-\sum_1^{j}(1+a_{p-k})\,, \tag{A1}$$

$$B_{j\alpha}=\sum_1^{m-\nu-2}(1+a_{p+\gamma})+\sum_{\alpha+1}^{\nu}(1+a_{p+\gamma}) - \sum_1^{j}(1+a_{p-k})\,. \tag{A2}$$

The atypical values $A_{j\alpha}$ and $B_{j\alpha}$ are, in order,

$$A_{j+1\,\alpha}<A_{j\alpha}<A_{j\,\alpha+1}\,,\quad B_{j+1\,\alpha}<B_{j\alpha}<B_{j\,\alpha-1}$$

and the difference $B_{j\alpha}-A_{j\alpha}$ is strictly positive for any $j=0,1,\ldots,p-1$ and any α such that $2\alpha\leqslant m-3$. When $m=2\nu$ and $\alpha=\nu-1$ we have

$$B_{j\,\nu-1}-A_{j\,\nu-1}=A_{p+\nu}-A_{p+\nu-1}$$

and the difference is either positive, or negative, or zero for any $j=0,1,\ldots,p-1$.

The degeneracy of atypicity δ of an irreducible representation

$$\{a_1,\ldots,a_{p-1}|a_p|a_{p+1},\ldots,a_{p+\nu}\}$$

is the number of time the value of a_p enters in the set of the atypical values $\{A_{j\alpha},B_{j\alpha}\}$ constructed from the other Kac–Dynkin parameters by using Eqs. (A1) and (A2). Because of the positivity of these last Dynkin parameters we obtain the two relations

$$0\leqslant\sigma\leqslant L,\quad L=\min[m-1,p]\,.$$

Kac's consistency conditions[1] are written for $0\leqslant b\leqslant\nu-1$,

$$a_{p+b+1}=\cdots=a_{p+\nu}=0$$

excepted on the case $m=2\nu$, $b=\nu-1$ where we have

$$a_{p+\nu-1}=a_{p+\nu}\,.$$

The corresponding value of a_p is atypical A_{0b}.

Typical representations: When a_p does not take one atypical value the irreducible representation is typical. Notice that from Kac's consistency condition a necessary condition for a representation to be typical is $b\geqslant\nu$. For an irreducible typical representation the dimension is independent of a_p and it is given by

$$2^{2\nu p}\times N_{\rm SYMP}(a_1,\ldots,a_{p-1},\bar b)\times N_{\rm ORTH}(a_{p+1},\ldots,a_{p+\nu})\,,$$

where $\bar b=b-\nu$. The symplectic and orthogonal factors are as follows.

(1) $m=2\nu+1$:

$$N_{\rm SYMP}(a_1,\ldots,a_{p-1},\bar b) =\prod_{1\leqslant i\leqslant p}\left(1+\frac{2\,\Sigma_i^{p-1}a_k+2\bar b}{2p+1-2j}\right)\prod_{1\leqslant i<j\leqslant p}\left(1+\frac{\Sigma_i^{p-1}a_k+\Sigma_j^{p-1}a_k+2b}{2p+1-i-j}\right)\left[1+\frac{\Sigma_i^{j-1}a_k}{j-i}\right],$$

$$N_{\rm ORTH}(a_{p+1},\ldots,a_{p+\nu}) =\prod_{1\leqslant\alpha\leqslant\nu}\left(1+\frac{2\,\Sigma_\alpha^{\nu-1}a_{p+\gamma}+a_{p+\nu}}{2\nu+l-2\alpha}\right)\prod_{1\leqslant\alpha<\beta\leqslant\nu}\left(1+\frac{\Sigma_\alpha^{\nu-1}a_{p+\gamma}+\Sigma_\beta^{\nu-1}a_{p+\gamma}+a_{p+\nu}}{2\nu+1-\alpha-\beta}\right)\left(1+\frac{\Sigma_\alpha^{\beta-1}a_{p+\gamma}}{\beta-\alpha}\right).$$

(2) $m=2\nu$:

$$N_{\rm SYMP}(a_1,\ldots,a_{p-1},\bar b) =\prod_{1\leqslant i\leqslant p}\left(1+\frac{\Sigma_i^{p-1}a_k+\bar b}{p+1-i}\right)\prod_{1\leqslant i<j\leqslant p}\left(1+\frac{\Sigma_i^{p-1}a_k+\Sigma_j^{p-1}a_k+\bar b}{2p+2-i-j}\right)\left(1+\frac{\Sigma_i^{j-1}a_k}{j-i}\right),$$

$$N_{\text{ORTH}}(a_{p+1},\dots,a_{p+\nu})$$
$$= \prod_{1<\alpha<\nu}\left(1+\frac{\Sigma_\alpha^{\nu-2}a_{p+\gamma}+a_{p+\nu-1}}{\nu-\alpha}\right)\left(1+\frac{\Sigma_\alpha^{\nu-2}a_{p+\gamma}+a_{p+\nu}}{\nu-\alpha}\right)$$
$$\times\prod_{1<\alpha<\beta<\nu}\left(1+\frac{\Sigma_\alpha^{\nu-2}a_{p+\gamma}+\Sigma_\beta^{\nu-2}a_{p+\gamma}+a_{p+\nu-1}+a_{p+2}}{2\nu-\alpha-\beta}\right)\left(1+\frac{\Sigma_\alpha^{\beta-1}a_{p+\gamma}}{\beta-\alpha}\right).$$

APPENDIX B: SUPERTABLEAUX OF OSP(m|4)

Starting from the results of Table II for the two-column supertableaux of OSP(m|4) we construct more general supertableaux keeping μ_1 and μ_2 fixed and adding rows.

Case I: $0\leqslant\mu_2\leqslant\nu-1$, $\mu_1+\mu_2\leqslant m-3$. Here μ_2 is atypical $A_{0\alpha}$, $\mu_2=\alpha$, $\lambda_{1+\alpha}=0$, μ_1 is atypical $A_{1\beta}$ $(\alpha<\beta\leqslant\nu-1)$ or $B_{1\beta}$ $(\alpha\leqslant\beta\leqslant m-\nu-2)$. The two-column supertableau has $\delta=2$ and is irreducible $\rho=1=2^{\delta-2}$.

To this two-column supertableau we can add up to α rows of arbitrary length without changing δ and ρ.

Case II: $0\leqslant\mu_2\leqslant\nu-1$, $\mu_1+\mu_2\geqslant m-2$. Here μ_2 is atypical $A_{0\alpha}$, $\mu_2=\alpha$, $\lambda_{1+\alpha}=0$, μ_1 is either atypical $B_{1\beta}$ $(0\leqslant\beta<\alpha)$ or typical. When μ_1 is atypical the two-column supertableau has $\delta=2$ and we have a two-GAST; when μ_1 is typical the two-column supertableau has $\delta=1$ and is irreducible. In both cases the relation between δ and ρ is $\rho=2^{\delta-1}$.

To this two-column supertableau we can add up to α rows of arbitrary length without changing the μ_2 atypicity. However, the μ_1 atypicity can be modified, μ_1 atypical $B_{1\beta}$ may stay atypical $B_{1\gamma}$ $(\beta\leqslant\gamma<\alpha)$ or become typical, μ_1 typical may stay typical or become atypical $B_{1\gamma}$ $(0\leqslant\gamma<\alpha)$. In both situations the relation between ρ and δ is not modified.

Case III: $\nu\leqslant\mu_2\leqslant m-2$. Here μ_2 is atypical $B_{0\alpha}$, $\alpha=m-2-\mu_2$, $\lambda_{1+\alpha}=0$, μ_1 is typical or atypical $B_{1\beta}$ $(0\leqslant\beta<\alpha)$. The properties of these two-column supertableaux are as in Case II and, in particular, we have $\rho=2^{\delta-1}$.

Let us add to this two-column supertableau r rows of arbitrary lengths $(r\leqslant\mu_2)$. When $r\leqslant\alpha$ $(\lambda_{1+\alpha}=0)$ the discussion proceeds as in the case II and the relation $\rho=2^{\delta-1}$ is preserved. When $r\geqslant\alpha+1$ $(\lambda_{1+\alpha}\neq0)$ we change both atypicities in μ_2 and μ_1. For μ_2 we obtain either an atypicity $B_{0\gamma}$ $(\alpha<\gamma\leqslant m-\nu-2)$ with $\lambda_{1+\gamma}\neq0$ or no atypicity; for μ_1 we obtain either an atypicity $B_{1\delta}$ $(\beta<\delta\leqslant m-\nu-2)$ with $\lambda_{1+\delta}\neq0$ or no atypicity. Of course $\beta\leqslant\delta<\gamma\leqslant m-\nu-2$. Now the supertableau becomes of the normal type and we have the relation $\rho=2^\delta$.

Case IV: The two column supertableau is typical $\delta=0$ and irreducible $\rho=1$. Adding at most ν rows of arbitrary length it may stay typical or become atypical $\delta=1$ or $\delta=2$ of the normal type $\rho=2^\delta$.

APPENDIX C: SUPERTABLEAUX WITH p COLUMNS OF OSP(m|2p)

We study the supertableaux of OSP($m|2p$) with at most p columns $(\lambda_1=0)$. Using Eqs. (1) and (2) of Sec. II we write the relations of atypicity in the form

$$\mu_j=j+k-(p+1)$$
$$\text{with } j=1,2,\dots,p,\quad k=1,2,\dots,m-1\,.$$

These relations are presented in the $(m-1)xp$ Table IV. Taking into account the positivity properties of the column parameters we easily see that for a supertableau, only few values of k are allowed:

$$p+1-j\leqslant k\leqslant m-1\,.$$

By inspection of the Table IV it is straightforward to obtain the degeneracy of atypicity for the p-column supertableau:

(1) $\delta=0$ typical,

$$\mu_p\geqslant m-1\,;$$

(2) $1<\delta<L$,

$$\mu_{p-\delta}\geqslant m-1-\delta\geqslant\mu_{p-\delta+1}\,;\qquad\text{(C1)}$$

(3) $\delta=L$ maximal degeneracy of atypicity,

$$\mu_1\leqslant m-1-p\quad\text{for } p<m-1\,,$$
$$\mu_{p+Q-m}=0\quad\text{for } p\geqslant m-1\,.$$

Let us now consider the partition in classes defined by Eq. (9). The typical p-column supertableaux obviously belong to the class Δ_0. Consider now a p-column supertableau with a degeneracy of atypicity δ, T_δ, and $\delta<L$. From Eq. (C1) we have

$$C_{p+\delta}+C_{p-\delta+1}\geqslant C_{p-\delta}\geqslant m-1-\delta\,.$$

With the definitions of classes given in Eqs. (9) this inequality implies that the supertableau T_δ belongs to one of the

TABLE IV. Atypicity for the p-column supertableaux of OSP($m|2p$).

$\mu_p=0$	$\mu_p=1$	$\mu_p=\alpha$	$\mu_p=m-3$	$\mu_p=m-2$
$\mu_{p-1}=-1$ $\mu_j=j-p$ $\mu_1=1-p$	$\mu_{p-1}=0$ $\mu_j=j-p+1$ $\mu_1=2-p$	$\mu_{p-1}=\alpha-1$ $\mu_j=j-p+\alpha$ $\mu_1=1+\alpha-p$	$\mu_{p-1}=m-4$ $\mu_j=j-p+m-3$ $\mu_1=m-2-p$	$\mu_{p-1}=m-3$ $\mu_j=j-p+m-2$ $\mu_1=m-1-p$

classes Δ_l with $0 \leqslant l \leqslant \delta$,

$$T_\delta \in \Delta_0 \oplus \Delta_1 \oplus \cdots \oplus \Delta_\delta \ . \tag{C2}$$

The relation (C2) is just the Theorem I for p column supertableaux. The case $\delta = L$ is trivial and the supertableau T_L belongs to any class Δ_l.

[1]V. C. Kac, Adv. Math. B **26**, (1977); Commun. Math. Phys. **53**, 31 (1977); *Lecture Notes in Mathematics*, Vol. 675 (Springer, Berlin, 1978).
[2]R. J. Farmer and P. D. Jarvis, J. Phys. A: Math. Gen. **17**, 2365 (1984).
[3]F. Delduc and M. Gourdin, J. Math. Phys. **25**, 1651 (1984); **26**, 1865 (1985); M. Gourdin, Paris Preprints PAR/LPTHE 84/26, 84/31, 1984.
[4]A. B. Balentekin and I. Bars, J. Math. Phys. **22**, 1149, 1810 (1981).
[5]J. Thierry-Mieg, Berkeley Preprint, UCB PTH 85/14 1BL 19353, 1985.
[6]M. Hammermesh, *Group Theory and Its Application to Physical Problems* (New York, 1962); M. Gourdin, *Basic of Lie Groups* (Editions Frontiéres, Gif sur Yvette, 1982).
[7]M. Gourdin, J. Math. Phys. **27**, 2832 (1986).

Integrable equations in (2+1) dimensions associated with symmetric and homogeneous spaces

Chris Athorne and Allan Fordy
Department of Applied Mathematical Studies and Centre for Nonlinear Studies, Leeds University, Leeds LS2 9JT, United Kingdom

(Received 3 February 1987; accepted for publication 29 April 1987)

Generalizations of the N-wave, Davey–Stewartson, and Kadomtsev–Petviashvili equations associated with homogeneous and symmetric spaces are presented. These equations are $(2+1)$-dimensional generalizations of those presented by Fordy and Kulish [Commun. Math. Phys. **89**, 427 (1983)] and Athorne and Fordy [J. Phys. A **20**, 1377 (1987)]. Examples are explicitly presented that are associated with the simplest spaces. In particular, a single component, $(2+1)$-dimensional generalization of the KdV equation is presented.

I. INTRODUCTION

In a recent series of papers we have presented a number of systems of equations associated with Hermitian symmetric spaces and reductive homogeneous spaces. The list of equations includes generalizations of the NLS and N-wave equations,[1] the DNLS equation,[2] and the KdV and MKdV equations.[3] All of these equations are isospectral flows of the linear spectral problem:

$$\hat{\psi}_x = (\lambda \mathsf{A} + \mathsf{Q})\hat{\psi}, \tag{1.1}$$

where constant matrix A and potential $\mathsf{Q}(x,t_N)$ are related to any of the Hermitian symmetric spaces (sometimes extended to reductive homogeneous spaces) in a way described in Sec. II. The matrix A is diagonal, but usually highly degenerate: ad A has only three distinct eigenvalues in the symmetric space case. This is in contrast to most other discussions of linear problems of the form (1.1), which usually assume A to be regular (see, for example, Ref. 4).

The above equations are all in $(1+1)$ space-time dimensions. In the present paper we generalize these equations to the case of $(2+1)$ space-time dimensions. This is achieved by the usual method[5–7] of replacing the spectral parameter λ by a new "spatial" derivative $\partial/\partial y$ in (1.1):

$$\psi_x = \mathsf{A}\psi_y + \mathsf{Q}\psi, \tag{1.2}$$

where A is the same constant matrix and the potential $\mathsf{Q}(x,y,t_N)$ now also depends upon y. Solutions $\psi(x,y,t_N)$ of (1.2) can be simply related to those of (1.1) by Fourier transform. The t_N dependance (for each $N=1,2,\ldots$) of ψ is defined by the linear evolution

$$\psi_{t_N} = \sum_{l=0}^{N} \mathsf{S}^{(N-l)}\,\partial_y^l \psi, \tag{1.3}$$

where $\partial_y^l \equiv (\partial/\partial y)^l$. The calculation of the coefficients $\mathsf{S}^{(N-l)}$ is much more complicated here than in the $(1+1)$-dimensional case corresponding to (1.1). In the latter case the coefficients would lie in the ring of polynomials of Q and its x derivatives whereas in the present case we must introduce some potentials with nonlocal definitions. This problem arises because ∂_y does not commute with the functions Q and $\mathsf{S}^{(N-l)}$, whereas λ does. Furthermore, the calculation here is not purely within the context of Lie algebras, since the associative matrix product is explicitly used. This causes problems in the case of class BDI symmetric spaces, which are thus omitted from our discussion.

In this paper we specifically discuss the cases $N=1$, 2, and 3. These correspond respectively to N-wave, Davey–Stewartson, and (in reduced form) KP-like equations. The case $N=1$ is only nonlinear in the homogeneous space case. The case $N=3$ is rather complicated, so we only present results in the simplest symmetric space case.

In the Appendix we derive generalizations of the Calogero–Degasperis boomeron equation[8] and further generalizations of the KP equation.

II. SYMMETRIC AND HOMOGENEOUS SPACES AND THE SPECTRAL PROBLEM

First we review some of the basic facts concerning Hermitian symmetric and reductive homogeneous spaces. More details can be found in Refs. 1 and 9.

A homogeneous space of a Lie group G is any differentiable manifold M on which G acts transitively ($\forall p_1,p_2,\in M$, $\exists g\in G$: $g\cdot p_1 = p_2$). The subgroup of G that leaves a given point $p_0 \in M$ fixed is called the isotropy group at p_0 and is defined by

$$K \equiv K_{p_0} = \{g\in G:\ g\cdot p_0 = p_0\}.$$

It is a theorem that each such M can be identified with a coset space G/K for some subgroup K and that this K plays the role of isotropy group of some point. There are many topological and differential geometric subtleties, but we have no need of them in this paper. We are only interested in the decompositions of the corresponding Lie algebras.

Let $\mathbf{g}$ and $\mathbf{k}$ be the Lie algebras of G and K, respectively, and let $\mathbf{m}$ be the vector space complement of $\mathbf{k}$ in $\mathbf{g}$. Then

$$\mathbf{g} = \mathbf{k} \oplus \mathbf{m}, \quad [\mathbf{k},\mathbf{k}] \subset \mathbf{k},$$

and $\mathbf{m}$ is identified with the tangent space $T_{p_0}M$ of $M = G/K$ at point p_0. At the moment we have $[\mathbf{k},\mathbf{k}] \subset \mathbf{k}$, but know nothing of $[\mathbf{k},\mathbf{m}]$ and $[\mathbf{m},\mathbf{m}]$.

When $\mathbf{g}$ satisfies the more stringent conditions

$$\mathbf{g} = \mathbf{k} \oplus \mathbf{m}, \quad [\mathbf{k},\mathbf{k}] \subset \mathbf{k}, \quad [\mathbf{k},\mathbf{m}] \subset \mathbf{m},$$

then G/K is called a reductive homogeneous space. These spaces possess canonically defined connections with curvature and torsion. Evaluated at fixed point p_0, the curvature

 0022-2488/87/092018-07$02.50

and torsion tensors are given purely in terms of the Lie bracket operation on $\mathbf{m}$:

$$(R(X,Y)Z)_{p_0} = -[[X,Y]_{\mathbf{k}},Z], \quad X,Y,Z\in\mathbf{m},$$

$$T(X,Y)_{p_0} = -[X,Y]_{\mathbf{m}}, \quad X,Y\in\mathbf{m},$$

where the subscripts $\mathbf{k}$ and $\mathbf{m}$ refer to the components of $[X,Y]$ in those vector subspaces.

When $\mathbf{g}$ satisfies the conditions

$$\mathbf{g}=\mathbf{k}\oplus\mathbf{m}, \quad [\mathbf{k},\mathbf{k}]\subset\mathbf{k},$$

$$[\mathbf{k},\mathbf{m}]\subset\mathbf{m}, \quad [\mathbf{m},\mathbf{m}]\subset\mathbf{k},$$

then $\mathbf{g}$ is called a symmetric algebra and G/K is a symmetric space. For these spaces the above mentioned canonical connection is derived from a metric, which is itself given by the restriction of the Killing form to $\mathbf{m}$. This connection is torsion free. Evaluated at fixed point p_0, the curvature is given by

$$(R(X,Y)Z)_{p_0} = -[[X,Y],Z], \quad X,Y,Z\in\mathbf{m},$$

where we now automatically have $[X,Y]\in\mathbf{k}$.

A special feature of Hermitian symmetric spaces is the existence of an element $A\in\mathbf{k}$ such that $\mathbf{k}=C_{\mathbf{g}}(A) = \{B\in\mathbf{g}: [A,B]=0\}$. The element A can (and therefore will) be chosen to be diagonal: $A\in\mathbf{h}\subset\mathbf{k}\subset\mathbf{g}$, where $\mathbf{h}$ is the Cartan subalgebra of $\mathbf{g}$. This element is highly degenerate; indeed, ad A (which is a dim $\mathbf{g}\times$dim $\mathbf{g}$ matrix) has only three distinct eigenvalues: 0, $\pm a$. Specifically, we have $\mathbf{m}=\mathbf{m}^+\oplus\mathbf{m}^-$ and $[A,\mathbf{k}]=0$, $[A,X^\pm]=\pm aX^\pm$, with a being the same value for all $X^\pm\in\mathbf{m}^\pm$. For any $X\in\mathbf{g}$, $X=X^0+X^++X^-$ and $X^+=\Sigma_\alpha X_\alpha e_\alpha$, $X^- = \Sigma_\alpha X_{-\alpha} e_{-\alpha}$, where α is summed over a special subset θ^+ of the positive root system Φ^+. In particular, Q of (1.2) is given by $Q=Q^++Q^-$ with $Q^+=\Sigma_\alpha q^\alpha e_\alpha$, $Q^-=\Sigma_\alpha r^{-\alpha}e_{-\alpha}$.

In the case of reductive homogeneous spaces we still have $\mathbf{m}=\mathbf{m}^+\oplus\mathbf{m}^-$, but each of $\mathbf{m}^\pm$ is further split into blocks, each of which is an eigenspace of ad A. The subset θ^+ of positive roots thus splits into a number of subsets θ_j^+ on each of which $\alpha(A)$ takes a constant value: $\alpha(A)=a_j$, $\alpha\in\theta_j^+$. We write $Q^\pm = \Sigma_j Q_j^\pm$ with $[A,Q_j^\pm] = \pm a_j Q_j^\pm$.

In Hermitian symmetric spaces we have the convenient property that $[X^+,Y^+]=0$ for all pairs of elements of $\mathbf{m}^+$, and similarly for $\mathbf{m}^-$. This is not true for the reductive homogeneous spaces so that the calculations of Sec. III are much simpler than those of Sec. IV. In previous papers the nonlinear terms of the resulting differential equations have been written in terms of the curvature and (for homogeneous spaces) torsion tensors. Although the equations of this paper could be similarly written we do not do this here. The $(2+1)$ calculations of this paper involve the associative as well as the Lie product. For all the symmetric spaces other than class BDI we have the property $X^+Y^+=0$ for all pairs of elements of $\mathbf{m}^+$ (similarly for $\mathbf{m}^-$). Since we use this property, the results of this paper do not apply to class BDI spaces.

With A defined as above and $Q(x,y,t_N)\in\mathbf{m}$ consider the linear equation

$$\psi_x = A\psi_y + Q\psi \tag{2.1}$$

together with the time evolution

$$\psi_{t_N} = \sum_{l=0}^{N} S^{(N-l)}\partial_y^l\psi. \tag{2.2}$$

Equating coefficients of ∂_y^m, $m=0,\ldots,N+1$ in the commutator,

$$\left[\partial_x - A\partial_y - Q,\ \partial_{t_N} - \sum_{l=0}^{N} S^{(N-l)}\partial_y^l\right] = 0, \tag{2.3}$$

leads to a system of equations for the $S^{(N-l)}$. The first two equations are

$$[A,S^{(0)}]=0, \tag{2.4a}$$

$$S_x^{(0)} - AS_y^{(0)} - [Q,S^{(0)}] = [A,S^{(1)}], \tag{2.4b}$$

and correspond to the coefficients of ∂_y^{N+1} and ∂_y^N, respectively. The situation is immediately more complicated and less algebraically precise than the $(1+1)$ case. Equation (2.4b) includes the associative product $AS_y^{(0)}$, which takes us out of the purely Lie algebraic context.

In the $(1+1)$-dimensional case the equations corresponding to (2.3) can be solved recursively for $S^{(l)}$, the solutions being purely in terms of Q and is x derivatives. In the present case we have to distinguish between the components $S_\pm^{(l)}$ and $S^{(l)}$ lying, respectively, in the spaces $\mathbf{m}$ and $\mathbf{k}$. All but two of the $\mathbf{k}$ components are defined nonlocally as solutions of differential equations involving the differential operator $\partial_x - A\partial_y$. When all functions are independent of y these equations can be integrated with respect to (w.r.t.) x to give local expressions. In the $(2+1)$-dimensional case these components are best considered as additional potentials, which is a familiar feature of the Davey–Stewartson equations.[7]

For our usual choice of $S^{(0)}$ there are $(N-1)$ such additional potentials for the Nth-order flow. The $\mathbf{m}$ components $S_\pm^{(l)}$ are still recursively defined in terms of Q and its x derivatives and *previously defined* $\mathbf{k}$ potentials. The final equation of (2.3) is

$$Q_{t_N} = \sum_{l=1}^{N} S^{(N-l)}Q_{ly} + (\partial_x - A\partial_y)S^{(N)} - [Q,S^{(N)}], \tag{2.5}$$

where $Q_{ly}\equiv\partial_y^l Q$. In the case of symmetric spaces this equation is easily decoupled into $\mathbf{m}$ and $\mathbf{k}$ components

$$Q_{t_N}^\pm = (\partial_x - A\partial_y)S_\pm^{(N)} - [Q^\pm,S_0^{(N)}] + \sum_{l=1}^{N} S_0^{(N-l)}Q_{ly}^\pm, \tag{2.6a}$$

$$(\partial_x - A\partial_y)S_0^{(N)} = [Q^+,S_-^{(N)}] + [Q^-,S_+^{(N)}] - \sum_{l=1}^{N}(S_+^{(N-l)}Q_{ly}^- + S_-^{(N-l)}Q_{ly}^+). \tag{2.6b}$$

We now explicitly construct some of the lower-order flows included in the above system of equations.

III. SYMMETRIC SPACE CASE: SECOND- AND THIRD-ORDER FLOWS

The first-order equation (n wave) is linear in the symmetric space case, so we start with $N=2$.

Generalized Davey–Stewartson equations: $N=2$: The simplest solution of (2.4) is

$$S^{(0)}=A, \quad S^{(1)}=Q. \tag{3.1}$$

The remaining equations are then

$$[\mathsf{A},\mathsf{S}^{(2)}] = (\partial_x + \mathsf{A}\,\partial_y)\mathsf{Q}, \tag{3.2a}$$

$$\mathsf{Q}_{t_2} = (\partial_x - \mathsf{A}\,\partial_y)\mathsf{S}^{(2)} + \mathsf{A}\mathsf{Q}_{yy} + \mathsf{Q}\mathsf{Q}_y - [\mathsf{Q},\mathsf{S}^{(2)}]. \tag{3.2b}$$

Equation (3.2a) determines the **m** component of $\mathsf{S}^{(2)}$,

$$\mathsf{S}^{(2)} = \mathsf{S}_0^{(2)} + (1/a)(\partial_x + \mathsf{A}\,\partial_y)(\mathsf{Q}^+ - \mathsf{Q}^-),$$

and the **k** component of (3.2b) gives an equation for $\mathsf{S}_0^{(2)}$,

$$(\partial_x - \mathsf{A}\,\partial_y)\mathsf{S}_0^{(2)} = (1/a)(\partial_x + \mathsf{A}\,\partial_y)[\mathsf{Q}^-,\mathsf{Q}^+]. \tag{3.3a}$$

If we write $\mathsf{S}_0^{(2)} = (1/a)[\mathsf{Q}^-,\mathsf{Q}^+] + (1/a)\mathsf{V}_y$, we can integrate, w.r.t. y, the resulting equation for V to obtain

$$(\partial_x - \mathsf{A}\,\partial_y)\mathsf{V} = 2\mathsf{A}[\mathsf{Q}^-,\mathsf{Q}^+]. \tag{3.3b}$$

The **m** components of (3.2b) are

$$a\mathsf{Q}^+_{t_2} = \mathsf{Q}^+_{xx} - (\mathsf{A}^2 - a\mathsf{A})\mathsf{Q}^+_{yy} + [\mathsf{Q}^+,[\mathsf{Q}^+,\mathsf{Q}^-]] - [\mathsf{Q}^+,\mathsf{V}_y], \tag{3.4a}$$

$$-a\mathsf{Q}^-_{t_2} = \mathsf{Q}^-_{xx} - (\mathsf{A}^2 + a\mathsf{A})\mathsf{Q}^-_{yy} + [\mathsf{Q}^-,[\mathsf{Q}^-,\mathsf{Q}^+]] - [\mathsf{V}_y,\mathsf{Q}^-]. \tag{3.4b}$$

In this paper (since we exclude class BDI symmetric spaces) all linear problems are block diagonal in structure

$$\mathsf{A} = \begin{pmatrix} \alpha \mathsf{I}_m & 0 \\ 0 & \beta \mathsf{I}_n \end{pmatrix}, \quad \mathsf{Q} = \begin{pmatrix} 0 & \mathsf{q} \\ \mathsf{r} & 0 \end{pmatrix}, \quad \mathsf{V} = \begin{pmatrix} \mathsf{S} & 0 \\ 0 & \mathsf{T} \end{pmatrix},$$

where $\alpha = na/(m+n)$, $\beta = -ma/(m+n)$, q, r, S, and T are, respectively, $m\times n$, $n\times m$, $m\times m$, and $n\times n$ matrices and I_m, I_n are identity matrices. Equations (3.4) and (3.3b) then take the form

$$a\mathsf{q}_{t_2} = \mathsf{q}_{xx} + \frac{mna^2}{(m+n)^2}\mathsf{q}_{yy} - 2\mathsf{qrq} + \mathsf{S}_y\mathsf{q} - \mathsf{q}\mathsf{T}_y, \tag{3.5a}$$

$$-a\mathsf{r}_{t_2} = \mathsf{r}_{xx} + \frac{mna^2}{(m+n)^2}\mathsf{r}_{yy} - 2\mathsf{rqr} + \mathsf{r}\mathsf{S}_y - \mathsf{T}_y\mathsf{r}, \tag{3.5b}$$

$$(\partial_x - \alpha\,\partial_y)\mathsf{S} = -2\alpha\mathsf{qr}, \tag{3.5c}$$

$$(\partial_x - \beta\,\partial_y)\mathsf{T} = 2\beta\mathsf{rq}. \tag{3.5d}$$

In this paper we consider two reductions of the above general system, being, respectively, valid in hyperbolic and elliptic[5] linear problems.

Hyperbolic: Here the matrix A is real so that $a^* = a$. If we set $\partial/\partial t_2 \to i(\partial/\partial t_2)$ (corresponding to the choice $\mathsf{S}^{(0)} = -i\mathsf{A}$ instead of A) then we can make the reduction

$$\mathsf{Q}^\dagger = \pm\mathsf{Q}, \quad \mathsf{V}^\dagger = \mathsf{V}$$

corresponding to

$$\mathsf{r} = \pm\mathsf{q}^\dagger, \quad \mathsf{S}^\dagger = \mathsf{S}, \quad \mathsf{T}^\dagger = \mathsf{T}.$$

Elliptic: Here the matrix A has imaginary eigenvalues, so that $a^* = -a$. When $m = n$, so that q and r are both square matrices, we can make the reduction

$$\mathsf{Q}^* = \pm\epsilon\mathsf{Q}\epsilon, \quad \mathsf{V}^* = \epsilon\mathsf{V}\epsilon,$$

where

$$\epsilon = \begin{pmatrix} 0 & \mathsf{I}_n \\ -\mathsf{I}_n & 0 \end{pmatrix}.$$

This corresponds to

$$\mathsf{r} = \pm\mathsf{q}^*, \quad \mathsf{T} = -\mathsf{S}^*.$$

Third-order flow: 2D KdV equation: The solution of (2.4) can still be taken as

$$\mathsf{S}^{(0)} = \mathsf{A}, \quad \mathsf{S}^{(1)} = \mathsf{Q}.$$

The remaining equations are

$$[\mathsf{A},\mathsf{S}^{(2)}] = \mathsf{D}_2\mathsf{Q}, \tag{3.6a}$$

$$[\mathsf{A},\mathsf{S}^{(3)}] = \mathsf{D}_{-1}\mathsf{S}^{(2)} + 3\mathsf{A}\mathsf{Q}_{yy} + 2\mathsf{Q}\mathsf{Q}_y - [\mathsf{Q},\mathsf{S}^{(2)}], \tag{3.6b}$$

$$\mathsf{Q}_{t_3} = \mathsf{D}_{-1}\mathsf{S}^{(3)} + \mathsf{A}\mathsf{Q}_{yyy} + \mathsf{Q}\mathsf{Q}_{yy} + \mathsf{S}^{(2)}\mathsf{Q}_y - [\mathsf{Q},\mathsf{S}^{(3)}], \tag{3.6c}$$

where $\mathsf{D}_l = \partial_x + l\mathsf{A}\,\partial_y$. The first equation determines the **m** component of $\mathsf{S}^{(2)}$:

$$\mathsf{S}^{(2)} = \mathsf{S}_0^{(2)} + (1/a)\mathsf{D}_2(\mathsf{Q}^+ - \mathsf{Q}^-) \tag{3.7a}$$

and the **k** component of (3.6b) gives an equation for $\mathsf{S}_0^{(2)}$:

$$\mathsf{D}_{-1}\mathsf{S}_0^{(2)} = (1/a)\mathsf{D}_2[\mathsf{Q}^-,\mathsf{Q}^+]. \tag{3.7b}$$

Similarly,the **m** component of (3.6b) gives

$$\mathsf{S}^{(3)} = \mathsf{S}_0^{(3)} + (1/a^2)\mathsf{D}_2\mathsf{D}_{-1}\mathsf{Q} + (3/a)\mathsf{A}(\mathsf{Q}^+ - \mathsf{Q}^-)_{yy} - (1/a)[\mathsf{Q}^+ - \mathsf{Q}^-,\mathsf{S}_0^{(2)}] \tag{3.8a}$$

with $\mathsf{S}_0^{(3)}$ given by the **k** component of (3.6c):

$$\begin{aligned}\mathsf{D}_{-1}\mathsf{S}_0^{(3)} &= (1/a)\{[\mathsf{S}_0^{(2)},[\mathsf{Q}^-,\mathsf{Q}^+]] + (\mathsf{Q}^-\mathsf{Q}^+_y)_x - (\mathsf{Q}^+\mathsf{Q}^-_y)_x + 4\mathsf{A}(\mathsf{Q}^+\mathsf{Q}^-_{yy} - \mathsf{Q}^-\mathsf{Q}^+_{yy}) \\ &\quad + 3\mathsf{A}[\mathsf{Q}^-,\mathsf{Q}^+]_{yy} - 4\mathsf{A}[\mathsf{Q}^-_y,\mathsf{Q}^+_y]\} \\ &\quad + (1/a^2)\{[\mathsf{Q},\mathsf{Q}_{xx}] + \mathsf{A}[\mathsf{Q},\mathsf{Q}_{yx}] - 2\mathsf{A}^2[\mathsf{Q},\mathsf{Q}_{yy}]\}.\end{aligned} \tag{3.8b}$$

The time evolution of Q is then given by

$$\mathsf{Q}_{t_3} = \mathsf{D}_{-1}\mathsf{S}_m^{(3)} + \mathsf{A}\mathsf{Q}_{yyy} + \mathsf{S}_0^{(2)}\mathsf{Q}_y - [\mathsf{Q},\mathsf{S}_0^{(3)}], \tag{3.9a}$$

where the **k** components $\mathsf{S}_0^{(2)}$ and $\mathsf{S}_0^{(3)}$ are given nonlocally by (3.7) and (3.8), while the **m** component $\mathsf{S}_m^{(3)}$ is given by (3.8a) as

$$\begin{aligned}\mathsf{S}_m^{(3)} &= (1/a^2)\{\partial_x^2 + \mathsf{A}\,\partial_x\,\partial_y + (3a\mathsf{A} - 2\mathsf{A}^2)\partial_y^2\}\mathsf{Q}^+ \\ &\quad + (1/a^2)\{\partial_x^2 + \mathsf{A}\,\partial_x\,\partial_y - (3a\mathsf{A} + 2\mathsf{A}^2)\partial_y^2\}\mathsf{Q}^- \\ &\quad - (1/a)[\mathsf{Q}^+,\mathsf{S}_0^{(2)}] + (1/a)[\mathsf{Q}^-,\mathsf{S}_0^{(2)}].\end{aligned} \tag{3.9b}$$

After some manipulation the equations of motion take the form

$$\begin{aligned}\mathsf{Q}^+_{t_3} &= (1/a^2)\{\partial_x^3 + (3a\mathsf{A} - 3\mathsf{A}^2)\partial_x\,\partial_y^2 + (a^2\mathsf{A} - 3a\mathsf{A}^2 + 2\mathsf{A}^3)\partial_y^3\}\mathsf{Q}^+ \\ &\quad - (1/a)\mathsf{D}_{-1}[\mathsf{Q}^+,\mathsf{S}_0^{(2)}] + \mathsf{S}_0^{(2)}\mathsf{Q}^+_y - [\mathsf{Q}^+,\mathsf{S}_0^{(3)}],\end{aligned} \tag{3.10a}$$

$$\begin{aligned}\mathsf{Q}^-_{t_3} &= (1/a^2)\{\partial_x^3 - (3a\mathsf{A} + 3\mathsf{A}^2)\partial_x\,\partial_y^2 + (a^2\mathsf{A} + 3a\mathsf{A}^2 + 2\mathsf{A}^3)\partial_y^3\}\mathsf{Q}^- \\ &\quad + (1/a)\mathsf{D}_{-1}[\mathsf{Q}^-,\mathsf{S}_0^{(2)}] + \mathsf{S}_0^{(2)}\mathsf{Q}^-_y - [\mathsf{Q}^-,\mathsf{S}_0^{(3)}].\end{aligned} \tag{3.10b}$$

Remark: If we take all functions to be independent of y then

$$\mathsf{S}_0^{(2)} = (1/a)[\mathsf{Q}^-,\mathsf{Q}^+], \quad \mathsf{S}_0^{(3)} = (1/a^2)[\mathsf{Q},\mathsf{Q}_x],$$

and Q^+ satisfies the simple equation

$$a^2 Q^+_{t_3} = Q^+_{xxx} - 3(Q^+_x Q^- Q^+ + Q^+ Q^- Q^+_x). \qquad (3.10c)$$

This is identical to Eq. (3.1) of Ref. 3.

IV. HOMOGENEOUS SPACE CASE: FIRST- AND SECOND-ORDER FLOWS

In this case the commutation relations are more complicated. We still usually take (3.1) as the solution of (2.4). However, for the first-order equation this gives trivial results, so we choose $S^{(0)}$ to be any constant, diagonal matrix.

Generalized n-wave equations: $N=1$: First, we introduce an element $P \in \mathbf{m}$ such that $Q=[A,P]$. Then the solution of (2.4) is taken as

$$S^{(0)} = B, \quad S^{(1)} = [B,P], \qquad (4.1)$$

where B is a constant diagonal matrix. The equations of motion are then

$$Q_{t_1} = (\partial_x - A\,\partial_y)[B,P] + BQ_y - [Q,[B,P]]_{\mathbf{m}}, \qquad (4.2)$$

where the **k** and **m** suffices signify components in the respective subspaces. Here $S_0^{(1)}$ has been chosen to be zero, as is always possible since $[[A,P],\ [B,P]]_{\mathbf{k}} = 0$, in general.

Generalized Davey–Stewartson equations: $N=2$: Once again we define P by $Q=[A,P]$ and take the simplest solution of (2.4):

$$S^{(0)} = A, \quad S^{(1)} = Q. \qquad (4.3)$$

The remaining equations are

$$[A,S^{(2)}] = (\partial_x + A\,\partial_y)Q, \qquad (4.4a)$$

$$(Q_{t_2} = (\partial_x - A\,\partial_y)S^{(2)} + AQ_{yy} + QQ_y - [Q,S^{(2)}]). \qquad (4.4b)$$

Equation (4.4a) determines the **m** component of $S^{(2)}$:

$$S^{(2)} = S_0^{(2)} + (\partial_x + A\,\partial_y)P.$$

Now recall that $Q^\pm = \Sigma_j Q_j^\pm$, with $[A,Q_j^\pm] = \pm a_j Q_j^\pm$. The **k** component of Eq. (4.4b) is

$$(\partial_x - A\,\partial_y)S_0^{(2)} = \sum_j \{[Q_j^+,(\partial_x + A\,\partial_y)P_j^-] + [Q_j^-,(\partial_x + A\,\partial_y)P_j^+] - Q_j^+Q_{jy}^- - Q_j^-Q_{jy}^+\}. \qquad (4.5a)$$

Here we have used such properties as $[A,Q_j^+Q_{iy}^-] = (a_j - a_i)Q_j^+Q_{iy}^-$ so that $Q_j^+Q_{iy}^- \in \mathbf{k}$ iff $a_j - a_i = 0$, so that $j=i$. Since we have $Q_j^\pm = \pm a_j P_j^\pm$, Eq. (4.5a) simplifies to

$$(\partial_x - A\,\partial_y)S_0^{(2)} = \sum_j (\partial_x + A\,\partial_y)[Q_j^+,P_j^-]. \qquad (4.5b)$$

Defining $V(x,y)$ by

$$S_0^{(2)} = \sum_j [Q_j^+,P_j^-] + V_y, \qquad (4.6a)$$

we find

$$(\partial_x - A\,\partial_y)V = \sum 2A[Q_j^+,P_j^-]. \qquad (4.6b)$$

Our normalization of V is slightly different here than in (3.3b). This gives rise to a factor $a_i V_y$ in Eq. (4.7) below, so that it *seems* not to reduce correctly to (3.4).

The $\mathbf{m}^\pm$ components of (4.4b) are then

$$a_i Q^+_{it_2} = \{\partial_x^2 + (a_i A - A^2)\partial_y^2\}Q_i^+ + \sum_j \frac{a_i}{a_j}[Q_i^+,[Q_j^+,Q_j^-]] - a_i[Q_i^+,V_y] + a_i\{(\vec{D}_1 P)Q - Q(P\overleftarrow{D}_1)\}_{+i}, \qquad (4.7a)$$

$$-a_i Q^-_{it_2} = \{\partial_x^2 - (a_i A + A^2)\partial_y^2\}Q_i^- + \sum_j \frac{a_i}{a_j}[Q_i^-,[Q_j^-,Q_j^+]] + a_i[Q_i^-,V_y] - a_i\{(\vec{D}_1 P)Q - Q(P\overleftarrow{D}_1)\}_{-i}, \qquad (4.7b)$$

where $\vec{D}_1 P = P_x + AP_y$ and $P\overleftarrow{D}_1 = P_x + P_y A$. We write $\{\ \}_{\pm i}$ to describe the components of these terms in the corresponding subspaces. These terms are easily written down in terms of weight spaces but rather messy in general, so we content ourselves with our explicit example in the next section. The third-order flow contains even more complicated terms so we do not present it here.

V. EXAMPLES

The equations we derived in Sec. III can be classified according to the three symmetric space classes AIII, CI, and DIII (p. 518 of Ref. 9). As previously mentioned, our discussion does not include class BDI symmetric spaces. In this section we only present the hyperbolic case with reduction $Q^\dagger = -Q$, corresponding to $r = -q^\dagger$ or $r^{-\alpha} = -(q^\alpha)^*$, which is the compact real form. We have

AIII $\quad SU(m+n)/S(U(m)\times U(n))$.

We consider two special cases.

Vector equations: $m=1$:

$$\begin{pmatrix}\psi_0\\ \vdots\\ \psi_n\end{pmatrix}_x = \left(\begin{array}{c|cccc} [n/(n+1)]\partial_y & q_1 & \cdots & & q_n \\ \hline -q_1^* & [-1/(n+1)]\partial_y & & & \\ \vdots & & & & \\ -q_n^* & & & & [-1/(n+1)]\partial_y \end{array}\right)\begin{pmatrix}\psi_0\\ \vdots\\ \psi_n\end{pmatrix}. \qquad (5.1)$$

Here our potential V takes the form

$$V = \left(\begin{array}{c|c} S & \\ \hline & T_{ij}\end{array}\right),$$

where S is just a scalar. We form the potential $U_{ij} = S\delta_{ij} - T_{ij}$. Then

$$iq_{jt_2} = q_{jxx} + \frac{n}{(n+1)^2}q_{jyy} - 2|\mathbf{q}|^2 q_j + q_k U_{kjy}, \qquad (5.2a)$$

$$\left(\partial_x^2 + \left(\frac{1-n}{1+n}\right)\partial_x \partial_y - \frac{n}{(n+1)^2}\partial_y^2 \right) U_{jk}$$
$$= -\frac{2}{n+1}\partial_x (q_j^* q_k - n|\mathbf{q}|^2\delta_{jk})$$
$$+ \frac{2n}{(n+1)^2}\partial_y (q_j^* q_k + |\mathbf{q}|^2 \delta_{jk}), \quad (5.2b)$$

which is a vector generalization of the hyperbolic Davey–Stewartson equation.

We have calculated the full third-order vector equations, but they are too complicated to present here. We content ourselves with the much simpler case of $n = 1$, starting with two real scalar potentials q and r and later reducing by taking $r = 1$. First, we define some new potentials, $S^{(2)}_{+-}$, etc. by the following:

$$S_0^{(2)} = \begin{pmatrix} S^{(2)}_{+-} & 0 \\ 0 & S^{(2)}_{-+} \end{pmatrix}, \quad S_0^{(3)} = \begin{pmatrix} S^{(3)}_{+-} & 0 \\ 0 & S^{(3)}_{-+} \end{pmatrix}.$$

Furthermore, the potential $S_0^{(3)}$ only appears in the evolution equations in the combination $S^{(3)}_{+-} - S^{(3)}_{-+}$. The resulting three potentials can be defined in terms of just two functions Φ and Ψ:

$$S^{(2)}_{+-} = -(\partial_x^2 + \tfrac{3}{2}\partial_x \partial_y + \tfrac{1}{2}\partial_y^2)\Phi,$$
$$S^{(2)}_{-+} = (\partial_x^2 - \tfrac{3}{2}\partial_x \partial_y + \tfrac{1}{2}\partial_y^2)\Phi,$$
$$S^{(3)}_{+-} - S^{(3)}_{-+} = \Psi.$$

These new potentials are defined nonlocally in terms of the functions q and r by

$$(\partial_x^2 - \tfrac{1}{4}\partial_y^2)\Phi = qr, \quad (5.3a)$$
$$(\partial_x^2 - \tfrac{1}{4}\partial_y^2)\Psi$$
$$= \{-(qr)_y + 2(qr_x - rq_x)\}_{xx} + \tfrac{3}{4}\{qr_x - rq_x\}_{xy}$$
$$+ \tfrac{1}{2}\{-(qr)_y + \tfrac{1}{2}(qr_x - rq_x)\}_{yy}. \quad (5.3b)$$

The time evolutions of q and r are then given by

$$q_{t_3} = (\partial_x^3 + \tfrac{3}{4}\partial_x\partial_y^2)q - (\partial_x - \tfrac{1}{2}\partial_y)\{q(2\partial_x^2 + \partial_y^2)\Phi\}$$
$$- q_y(\partial_x^2 + \tfrac{3}{2}\partial_x\partial_y + \tfrac{1}{2}\partial_y^2)\Phi + q\Psi, \quad (5.3c)$$
$$r_{t_3} = (\partial_x^3 + \tfrac{3}{4}\partial_x\partial_y^2)r - (\partial_x + \tfrac{1}{2}\partial_y)\{r(2\partial_x^2 + \partial_y^2)\Phi\}$$
$$+ r_y(\partial_x^2 - \tfrac{3}{2}\partial_x \partial_y + \tfrac{1}{2}\partial_y^2)\Phi - r\Psi. \quad (5.3d)$$

It is possible to make a 2-D KdV reduction of the system (5.3) by setting $r = 1$:

$$(\partial_x^2 - \tfrac{1}{4}\partial_y^2)\Phi = q, \quad (5.4a)$$
$$-(\partial_x + \tfrac{1}{2}\partial_y)(2\partial_x^2 + \partial_y^2)\Phi = \Psi, \quad (5.4b)$$
$$q_{t_3} = (\partial_x^3 + \tfrac{3}{4}\partial_x \partial_y^2)q - 2q_x\Phi_{xx} - q_x\Phi_{yy}$$
$$- 4q\Phi_{xxx} - 2q\Phi_{xyy} - \tfrac{3}{2}q_y\Phi_{xy}. \quad (5.4c)$$

Equations (5.4a) and (5.4c) constitute our 2-D generalization of the KdV equation, and are considerably more complicated than the KP equation. When all functions are independent of y then (5.4c) reduces to the KdV equation and (5.4a) is just Hirota's substitution if we let $\Phi = -2\ln\tau$. It is still possible to write the $(2+1)$-dimensional equations in Hirota form,

$$D^{(2)}\tau\cdot[(\bar{D}D_t - \bar{D}D^3)\tau\cdot\tau]$$
$$+ \bar{D}^{(2)}\tau\cdot[(DD_t - D\bar{D}^3)\tau\cdot\tau] = 0, \quad (5.5)$$

where $D = D_x + (a/2)D_y$; $\bar{D} = D_x - (a/2)D_y$; $D\rho\cdot\sigma = \rho'\sigma - \sigma'\rho$, and $D^{(2)}\rho\cdot\sigma = 2\rho'\sigma - \rho\sigma'$, and q is given by (5.4a). This is similar to (but crucially different form) the Hirota form of the KP equation. This aspect was reported in Ref. 10.

Remark: In this case the linear problem can be written as a scalar, second-order equation

$$(\partial_x^2 - \tfrac{1}{4}\partial_y^2)\Psi = q\Psi. \quad (5.6)$$

This scalar linear problem has been considered in Ref. 11, where an equation of the form similar to (5.4c) was derived.

Square matrix equations: $m = n$:

$$\begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{2n} \end{pmatrix}_x = \left(\begin{array}{c|c} \tfrac{1}{2}\mathrm{I}_n \partial_y & \mathrm{q} \\ \hline -\mathrm{q}^\dagger & -\tfrac{1}{2}\mathrm{I}_n \partial_y \end{array}\right) \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{2n} \end{pmatrix}, \quad (5.7)$$

where q is an $n \times n$ matrix. The matrices S and T are similarly $n \times n$. The resulting equations are (3.5) with $m = n$, $\mathrm{r} = -\mathrm{q}^\dagger$:

$$i\mathrm{q}_{t_2} = \mathrm{q}_{xx} + \tfrac{1}{4}\mathrm{q}_{yy} + 2\mathrm{q}\mathrm{q}^\dagger\mathrm{q} + \mathrm{S}_y\mathrm{q} - \mathrm{q}\mathrm{T}_y, \quad (5.8a)$$
$$(\partial_x - \tfrac{1}{2}\partial_y)\mathrm{S} = \mathrm{q}\mathrm{q}^\dagger, \quad (5.8b)$$
$$(\partial_x + \tfrac{1}{2}\partial_y)\mathrm{T} = \mathrm{q}^\dagger\mathrm{q}. \quad (5.8c)$$

Example: $m = n = 2$:

$$iq_{1t_2} = q_{1xx} + \frac{1}{4}q_{1yy} + 2q_1\sum_1^4 |q_j|^2 + 2q_3^*(q_2q_4 - q_1q_3)$$
$$+ (S_{11y} - T_{11y})q_1 + S_{12y}q_4 - q_2T_{21y}, \quad (5.9a)$$
$$iq_{2t_2} = q_{2xx} + \frac{1}{4}q_{2yy} + 2q_2\sum_1^4 |q_j|^2 + 2q_4^*(q_1q_3 - q_2q_4)$$
$$+ (S_{11y} - T_{22y})q_2 + S_{12y}q_3 - q_1T_{12y}, \quad (5.9b)$$
$$iq_{3t_2} = q_{3xx} + \frac{1}{4}q_{3yy} + 2q_3\sum_1^4 |q_j|^2 + 2q_1^*(q_2q_4 - q_1q_3)$$
$$+ S_{21y}q_2 + (S_{22y} - T_{22y})q_3 - q_4T_{12y}, \quad (5.9c)$$
$$iq_{4t_2} = q_{4xx} + \frac{1}{4}q_{4yy} + 2q_4\sum_1^4 |q_j|^2 + 2q_2^*(q_1q_3 - q_2q_4)$$
$$+ S_{21y}q_1 + (S_{22y} - T_{11y})q_4 - q_3T_{21y}, \quad (5.9d)$$
$$\left(\partial_x - \frac{1}{2}\partial_y\right)\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} |q_1|^2 + |q_2|^2 & q_1q_4^* + q_2q_3^* \\ q_4q_1^* + q_3^*q_3 & |q_4|^2 + |q_3|^2 \end{pmatrix}, \quad (5.9e)$$
$$\left(\partial_x + \frac{1}{2}\partial_y\right)\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} = \begin{pmatrix} |q_1|^2 + |q_4|^2 & q_1^*q_2 + q_4^*q_3 \\ q_1q_2^* + q_4q_3^* & |q_2|^2 + |q_3|^2 \end{pmatrix}. \quad (5.9f)$$

CI *and* DIII: The compact real forms of these symmetric spaces are, respectively, $\mathrm{Sp}(n)/\mathrm{U}(n)$ and $\mathrm{SO}(2n)/\mathrm{U}(n)$. In the representations we use[12] they correspond to the linear problem (5.7) with the reductions $\mathrm{q}^T = \mathrm{q}$ and $\mathrm{q}^T = -\mathrm{q}$, respectively. The corresponding differential equations are similarly reduced.

Since $\mathrm{q}^\dagger = \pm\mathrm{q}^*$ and $\mathrm{q}\mathrm{q}^* = (\mathrm{q}^*\mathrm{q})^*$ we have

$$(\partial_x - \tfrac{1}{2}\partial_y)\mathrm{S}^* = (\partial_y + \tfrac{1}{2}\partial_x)\mathrm{T}. \quad (5.10a)$$

Thus there exists a functon Ω such that

$$\mathrm{S} = (\partial_x + \tfrac{1}{2}\partial_y)\Omega, \quad \mathrm{T} = (\partial_x - \tfrac{1}{2}\partial_y)\Omega^*, \quad (5.10b)$$

where, from (5.8b), $\Omega = \Omega^\dagger$. Comparing (5.8a) with its transpose, we find that Ω must satisfy the further constraint

$$\mathsf{q}\Omega^*_{xy} = \Omega_{xy}\mathsf{q}. \tag{5.10c}$$

The CI case with $N=2$ is easily obtained from the system (5.9) by setting $q_4 = q_2$ and defining S and T in terms of Ω as above. Here Ω must satisfy equations (5.10c) together with (5.9e):

$$\left(\partial_x^2 - \frac{1}{4}\partial_y^2\right)\Omega = \begin{pmatrix} |q_1|^2 + |q_2|^2 & q_1 q_4^* + q_2 q_3^* \\ q_4 q_1^* + q_2^* q_3 & |q_4|^2 + |q_3|^2 \end{pmatrix}. \tag{5.10d}$$

Homogeneous spaces: The simplest nontrivial example is associatd with SU(3):

SU(3)/S(U(1)×U(1)×U(1)).

The first-order flow in this case is just the standard three-wave interaction given in Ref. 7, so we only present the generalized Davey–Stewartson equation here. Referring to (4.3)–(4.7) we have

$$\mathsf{A} = \begin{pmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{pmatrix}, \quad \mathsf{Q} = \begin{pmatrix} 0 & q_1 & q_2 \\ -q_1^* & 0 & q_3 \\ -q_2^* & -q_3^* & 0 \end{pmatrix},$$

$$\mathsf{V} = \begin{pmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{pmatrix}, \tag{5.11}$$

with $\alpha_1 + \alpha_2 + \alpha_3 = 0$ and V_i being real functions. Equations (4.6b) and (4.7a) are then

$$(\partial_x - \alpha_1\partial_y)V_1 = 2\alpha_1[|q_1|^2/\alpha_{12} + |q_2|^2/\alpha_{13}], \tag{5.12a}$$

$$(\partial_x - \alpha_2\partial_y)V_2 = 2\alpha_2[|q_3|^2/\alpha_{23} + |q_1|^2/\alpha_{21}], \tag{5.12b}$$

$$(\partial_x - \alpha_3\partial_y)V_3 = 2\alpha_3[|q_2|^2/\alpha_{31} + |q_3|^2/\alpha_{32}], \tag{5.12c}$$

and

$$iq_{1t_2} = \frac{1}{\alpha_{12}}(\partial_x^2 - \alpha_1\alpha_2\partial_y^2)q_1 + \left[\frac{2|q_1|^2}{\alpha_{12}} + \frac{|q_2|^2}{\alpha_{13}} + \frac{|q_3|^2}{\alpha_{32}}q_1\right] - (q_{2x} + \alpha_1 q_{2y})\frac{q_3^*}{\alpha_{13}} - (q_{3x}^* + \alpha_2 q_{3y}^*)\frac{q_2}{\alpha_{23}} - q_1 V_{21y}, \tag{5.13a}$$

$$iq_{2t_2} = \frac{1}{\alpha_{13}}(\partial_x^2 - \alpha_1\alpha_3\partial_y^2)q_2 + \left(\frac{|q_1|^2}{\alpha_{12}} + \frac{2|q_2|^2}{\alpha_{13}} + \frac{|q_3|^2}{\alpha_{23}}\right)q_2 + (q_{1x} + \alpha_1 q_{1y})\frac{q_3}{\alpha_{12}} - (q_{3x} + \alpha_3 q_{3y})\frac{q_1}{\alpha_{23}} - q_2 V_{31y}, \tag{5.13b}$$

$$iq_{3t_2} = \frac{1}{\alpha_{23}}(\partial_x^2 - \alpha_2\alpha_3\partial_y^2)q_3 + \left(\frac{|q_1|^2}{\alpha_{21}} + \frac{|q_2|^2}{\alpha_{13}} + \frac{2|q_3|^2}{\alpha_{23}}\right)q_3 + (q_{1x} + \alpha_2 q_{1y}^*)\frac{q_2}{\alpha_{12}} + (q_{2x} + \alpha_3 q_{2y})\frac{q_1^*}{\alpha_{13}} - q_3 V_{32y}, \tag{5.13c}$$

where $\alpha_{ij} = \alpha_i - \alpha_j$ and $V_{ij} = V_i - V_j$.

It can be seen from the definitions of A and Q that by setting $\alpha_2 = \alpha_3$ and $q_3 = 0$ the linear problem reduces to that of (5.1) with $n=2$. This reduction corresponds to the subspace relation

$$\frac{\mathrm{SU}(3)}{S(\mathrm{U}(1)\times\mathrm{U}(2))} \subset \frac{\mathrm{SU}(3)}{\mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(1)\times\mathrm{U}(1))}.$$

However, in terms of Eqs. (5.13) and of the definition (5.11)of V, the situation is a little more tricky. To achieve our goal we set $\alpha_1 = 1$, $\alpha_2 = -\frac{1}{2} - \delta$, $\alpha_3 = -\frac{1}{2} + \delta$, and $q_3 = 2\delta U$ and let $\delta \to 0$. Equation (5.13c) gives us

$$(\partial_x - \tfrac{1}{2}\partial_y)U = \tfrac{2}{3}q_2 q_1^*, \tag{5.14a}$$

while (5.13a) and (5.13b) reduce to

$$\frac{3}{2}i\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}_{t_2} = \left(\partial_x^2 + \frac{1}{2}\partial_y^2 + 2(|q_1|^2 + |q_2|^2)\right)\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} + \frac{3}{2}\begin{pmatrix} V_2 - V_1 & U^* \\ U & V_3 - V_1 \end{pmatrix}_y \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}. \tag{5.14b}$$

Thus the third component of Q has come to take the role of the off-diagonal part of the potential V.

SU(4)/S(U(1)×U(1)×U(2)).

Referring to (4.1) and (4.2) we have

$$\mathsf{A} = \left(\begin{array}{cc|cc} a_1 & & & \\ & a_2 & & \\ \hline & & a_3 & \\ & & & a_3 \end{array}\right), \quad \mathsf{B} = \left(\begin{array}{cc|cc} b_1 & & & \\ & b_2 & & \\ \hline & & b_3 & \\ & & & b_3 \end{array}\right),$$

$$\mathsf{P} = \left(\begin{array}{c|c|cc} 0 & p_5 & p_1 & p_2 \\ \hline -p_5^* & 0 & p_4 & p_3 \\ \hline -p_1^* & -p_4^* & 0 & 0 \\ -p_2^* & -p_3^* & 0 & 0 \end{array}\right), \tag{5.15}$$

with $a_1 + a_2 + 2a_3 = b_1 + b_2 + 2b_3 = 0$. Equation (4.2a) takes the form

$$\alpha p_{1t_1} = \beta p_{1x} + \gamma p_{1y} + 4\gamma p_5 p_4, \tag{5.16a}$$

$$\alpha p_{2t_1} = \beta p_{2x} + \gamma p_{2y} + 4\gamma p_5 p_3, \tag{5.16b}$$

$$\alpha' p_{3t_1} = \beta' p_{3x} + \gamma p_{3y} - 4\gamma p_5^* p_1, \tag{5.16c}$$

$$\alpha' p_{4t_1} = \beta' p_{4x} + \gamma p_{4y} - 4\gamma p_5^* p_2, \tag{5.16d}$$

$$[(\alpha - \alpha')/4]p_{5t_1} = [(\beta - \beta')/4]p_{5x} - 2\gamma p_{5y} - 4\gamma(p_1 p_4^* + p_2 p_3^*), \tag{5.16e}$$

where $\alpha = a_1 - a_3$, $\alpha' = a_3 - a_2$, $\beta = b_1 - b_3$, $\beta' = b_3 - b_2$, and $\gamma = \frac{1}{4}(\alpha'\beta - \alpha\beta')$. Corresponding to the inclusion

$$\frac{\mathrm{SU}(4)}{\mathrm{S}(\mathrm{U}(2)\times\mathrm{U}(2))} \subset \frac{\mathrm{SU}(4)}{\mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(1)\times\mathrm{U}(2))},$$

we have the reduction $a_1 = a_2$, $b_1 = b_2$, and $p_5 = 0$ and our nonlinear five-wave equation reduces to a linear four-wave equation associated with a symmetric space.

Note added in proof: Equation (5.4c) was first studied by Novika and Veselov.[13]

APPENDIX: MATRIX KP AND BOOMERON EQUATIONS IN (2+1) DIMENSIONS

In this appendix we remark upon the (2 + 1)-dimensional flows associated with the matrix Schrödinger equation:

$$\psi_y + \psi_{xx} - U\psi = 0, \tag{A1}$$

where U is an $n\times n$ matrix. However, U is not an arbitrary matrix but taken to be the product $U = rq$, where r and q are as in (3.5). This is a generalization of the matrix Schrödinger equation considered in Ref. 3. However, in the present case there is not a direct connection between (2.1) and (A1). The t_3 equation is given by the integrability conditions of (A1) and

$$\psi_{t_3} = 4\psi_{xxx} - 6U\psi_x - 3(U_x - V)\psi \tag{A2}$$

so that

$$U_{t_3} - U_{xxx} + 3(UU_x + U_xU) = 3V_y + 3[V,U], \tag{A3a}$$

$$V_x = U_y. \tag{A3b}$$

Note: We can use (A3b) to define a potential W such that $U = W_x$, $V = W_y$ so that (A3a) can be written as an equation in one dependent matrix variable which is a simple generalization of the potential KP equation

$$(W_{t_3} - W_{xxx} + 3W_x^2 - \tfrac{3}{2}[W_y,W])_x = (3W_y + \tfrac{3}{2}[W,W_x])_y. \tag{A4}$$

In terms of the same potential W, we can write down a $(2+1)$-dimensional "boomeron" equation[8] as the integrability conditions of (A1) and

$$\psi_{t_1} - B\psi_x - C\psi = 0 \tag{A5}$$

so that

$$2W_{xt_1} = \{W_{xx},B\} + [W_x,C] + [W_y,B] - [W_x,[W,B]], \tag{A6}$$

where $\{f,g\}$ denotes the anticommutator $fg + gf$.

It is a simple exercise to substitute the form of U given in Ref. 3, to obtain explicit forms for Eqs. (A3) and (A6).

[1]A. P. Fordy and P. P. Kulish, Commun. Math. Phys. **89**, 427 (1983).
[2]A. P. Fordy, J. Phys. A **17**, 1235 (1984).
[3]C. Athorne and A. P. Fordy, J. Phys. A **20**, 1377 (1987).
[4]A. C. Newell, Proc. R. Soc. London Ser. A **365**, 283 (1979).
[5]A. S. Fokas and M. J. Ablowitz, J. Math. Phys. **25**, 2494 (1984).
[6]B. G. Konopelchenko, Commun. Math. Phys. **88**, 531 (1983).
[7]S. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov, *Theory of Solitons* (Plenum, New York, 1984).
[8]F. Calogero and A. Degasperis, Nuovo Cimento B **39**, **1** (1977).
[9]S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces* (Academic, New York, 1978), 2nd ed.
[10]C. Athorne and A. P. Fordy, "A hierarchy of integrable KdV-like equations in 2 + 1 dimensions," preprint, 1987.
[11]M. Boiti, J. J-P Leon, M. Manna, and F. Pempinelli, Inverse Problems **2**, 271 (1986).
[12]J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, Berlin, 1972).
[13]S. P. Novikov and A. P. Veselov, Physica D **18**, 267 (1986).

Periodic fixed points of Bäcklund transformations

John Weiss[a)]

Institute for Pure and Applied Physical Sciences, University of California, San Diego, La Jolla, California 92093

(Received 10 October 1986; accepted for publication 13 May 1987)

The discussion of the periodic fixed points of Bäcklund transformations for the Korteweg–de Vries equation is completed. It will be shown that the systems of equations defined by the KdV periodic fixed points are *equivalent* to the periodic Kac–Van Moerbeke systems. As a consequence, for even order fixed points, the KdV systems are equivalent to the periodic Toda lattice. The periodic fixed points of the Bäcklund transformation for the Boussinesq equation are found to have a Hamiltonian structure. The integrals of these systems are found.

I. INTRODUCTION

The (Schwarzian) KdV equation[1]

$$\phi_t/\phi_x + \{\phi;x\} = \lambda \tag{1.1}$$

has the Bäcklund transformation[1]

$$\phi = (a\psi + b)/(c\psi + d)\,, \tag{1.2}$$

$$ad - bc = 1, \quad \phi_x = \psi_x^{-1}\,, \tag{1.3}$$

where

$$\psi_t/\psi_x + \{\psi;x\} = \lambda\,. \tag{1.4}$$

The expression

$$\{\phi;x\} = \frac{\partial}{\partial x}\left(\frac{\phi_{xx}}{\phi_x}\right) - \frac{1}{2}\left(\frac{\phi_{xx}}{\phi_x}\right)^2 \tag{1.5}$$

is the Schwarzian derivative, which is invariant under the Möbius group (1.2).[2,3]

The effective Bäcklund transformation (BT) for (1.1) is the composition of (1.2) and (1.3). We find that[4]

$$\phi_{n+1,x} = \frac{\phi_n^2}{\phi_{n,x}}\,, \tag{1.6}$$

$$\frac{\phi_{n+1,t}}{\phi_{n+1,x}} + \frac{\phi_{n,t}}{\phi_{n,x}} = \left(\frac{\phi_{n,xx}}{\phi_{n,x}}\right)^2 - 4\frac{\partial^2}{\partial x^2}\ln\phi_n + 2\lambda \tag{1.7}$$

is a BT for (1.1). The periodic fixed points of the BT are defined by Eqs. (1.6) and (1.7) with

$$n = 1,2,3,4,\ldots \ (\mathrm{mod}\ N)\,. \tag{1.8}$$

The periodic fixed points continue to define a strong BT for (1.1). That is, the integrability conditions

$$\phi_{n+1,xt} = \phi_{n+1,tx} \tag{1.9}$$

continue to imply that ϕ_n satisfy (1.1), and, by the periodicity mod N, the set

$$\{\phi_n,\ n = 1,2,\ldots \ (\mathrm{mod}\ N)\} \tag{1.10}$$

are solutions of (1.1).

In a previous work[4] we have found that if

$$\xi_j = \phi_{j,x}/\phi_j \tag{1.11}$$

then

$$\xi_{j+1,x}/\xi_{j+1} + \xi_{j,x}/\xi_j = \xi_j - \xi_{j+1}\,. \tag{1.12}$$

Define the $N\times N$ circulant matrices[5]

$$A = \begin{pmatrix} 1 & 1 & 0 & & & \\ 0 & 1 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 1 & & \\ & & & \ddots & \ddots & \\ 0 & & & & 1 & 1 \\ 1 & & & & & 1 \end{pmatrix}, \tag{1.13}$$

$$B = \begin{pmatrix} 1 & -1 & 0 & & & \\ 0 & 1 & -1 & 0 & & 0 \\ 0 & 0 & 1 & -1 & & \\ & & & \ddots & \ddots & \\ 0 & & & & 1 & -1 \\ -1 & & & & & 1 \end{pmatrix}. \tag{1.14}$$

Then with

$$\hat{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \tag{1.15}$$

$$\beta_j = \ln\xi_j\,, \tag{1.16}$$

$$\hat{\xi} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \tag{1.17}$$

Eqs. (1.12) are

$$A\hat{\beta}_{,x} = B\hat{\xi}\,. \tag{1.18}$$

For all N the one-dimensional null space of B is spanned by the N vector

[a)] Current address: 6 Lockeland Ave., Arlington, Massachusetts, 02174.

0022-2488/87/092025-15$02.50

$$\hat{b}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{1.19}$$

While for $N = 2k$

$$|A| = 0 \tag{1.20}$$

and A has a one-dimensional null space spanned by

$$\hat{a}_0 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}. \tag{1.21}$$

For $N = 2k+1$, A is invertible and

$$A^{-1} = \tfrac{1}{2}(I + \Omega), \tag{1.22}$$

where

$$\Omega = \begin{pmatrix} 0 & -1 & 1 & -1 & 1 & \cdots & -1 & 1 \\ 1 & 0 & -1 & 1 & -1 & \cdots & 1 & -1 \\ -1 & 1 & 0 & -1 & 1 & \cdots & -1 & 1 \\ 1 & -1 & 1 & 0 & -1 & \cdots & 1 & -1 \\ -1 & 1 & -1 & 1 & 0 & & & \\ \vdots & & & & \ddots & \ddots & \ddots & \\ 1 & & & & & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 & -1 & \cdots & 1 & 0 \end{pmatrix} \tag{1.23}$$

is a $(2k+1, 2k+1)$ antisymmetric matrix

$$\Omega^t = -\Omega \tag{1.24}$$

with

$$|\Omega| = 0. \tag{1.25}$$

The one-dimensional null space of (1.23) is spanned by (1.19) and it can be shown that

$$\Omega = A^{-1}B. \tag{1.26}$$

In the notation for circulant matrices[5]

$$A = \mathrm{circ}[1,1,0,0,\ldots,0],$$
$$B = \mathrm{circ}[1,-1,0,0,\ldots,0], \tag{1.27}$$
$$\Omega = \mathrm{circ}[0,-1,1,-1,1,\ldots,-1,1].$$

When N is odd Eqs. (1.18) can be written as Hamiltonian systems

$$\hat{\xi}_x = \begin{pmatrix} \xi_1 & & & 0 \\ & \xi_2 & & \\ & & \ddots & \\ 0 & & & \xi_N \end{pmatrix} \Omega \begin{pmatrix} \xi_1 & & & 0 \\ & \xi_2 & & \\ & & \ddots & \\ 0 & & & \xi_N \end{pmatrix} \nabla_{\xi} H_1, \tag{1.28}$$

where

$$H_1 = \sum_{j=1}^{N} \xi_j. \tag{1.29}$$

In Ref. 4 we find that (1.28) is a completely integrable, k-dimensional, Hamiltonian system with one Casimir

$$H_N = \prod_{j=1}^{N} \xi_j \tag{1.30}$$

and k independent integrals

$$H_{N-2m} = L^{m} \circ H_N, \tag{1.31}$$

where

$$N = 2k+1 \tag{1.32}$$

and

$$L = \sum_{j=1}^{N} \frac{\partial^2}{\partial \xi_j \, \partial \xi_{j+1}}. \tag{1.33}$$

The above integrals (and Casimir) are in involution with regard to the Poisson bracket

$$\{G,H\} = (\nabla_{\xi} G)^t M_{\xi} \nabla_{\xi} H, \tag{1.34}$$

where the cosymplectic form

$$M_{\xi} = \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \Omega \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}. \tag{1.35}$$

For all N the system (1.18), by contraction with (1.19), has the Casimir integral (1.29). For even N,

$$N = 2k+2, \tag{1.36}$$

contraction of (1.18) with the null vector of A, (1.21), obtains the *constraint* condition

$$C_1 = \sum_{j=1}^{N} (-1)^j \xi_j \equiv 0. \tag{1.37}$$

In Sec. II we find that, when $N = 2k+2$, the system (1.18) is a k-dimensional completely integrable Hamiltonian system with Casimir (1.29) and k independent integrals (1.31) in involution. The integrals are in involution with the constraint (1.37). That is, the constraint is preserved by the flows. Furthermore, we find that systems (1.18) are *equivalent* to the periodic Kac–Van Moerbeke (KM) equations.[6] In effect, the KM flow commutes with (1.18). This implies, by a known result,[7] that (1.18) is *equivalent* to the periodic Toda lattice when N is even.

In Sec. III we find that the periodic fixed points of the BT for the Boussinesq equation[8] is of the form (1.18) and (1.28) for appropriate A, B, Ω. Again, the system is shown to have a Hamiltonian structure and the integrals are found by a method similar to that developed for the KdV systems. However, the Boussinesq systems are not equivalent to the KdV systems.

In Sec. IV we define, by a generalization of the KdV and Boussinesq systems, a hierarchy of Hamiltonian systems of the form (1.18). Certain integrals are found.

II. THE KORTEWEG–DE VRIES SYSTEM

A. Even-order fixed points

With reference to Ref. 4 and Sec. I, the KdV periodic fixed points are solutions of the system

$$\xi_{j,x}/\xi_j + \xi_{j+1,x}/\xi_{j+1} = \xi_j - \xi_{j+1}, \tag{2.1}$$

where $j = 1,2,3,4,...$ (mod N). For any N there is a Casimir invariant

$$H_N = \prod_{j=1}^{N} \xi_j \tag{2.2}$$

and for any N

$$H_{N-2m} = L^{m} \circ H_N \tag{2.3}$$

are integrals of the system (2.1), where

$$L = \sum_{j=1}^{N} D_j D_{j+1}, \tag{2.4}$$

$$D_j = \frac{\partial}{\partial \xi_j}. \tag{2.5}$$

Proof: The basic identities are, for each j,

$$\xi_j D_j L^{m} \circ H_N = L^{m} \circ H_N - m D_j \times (D_{j-1} + D_{j+1}) L^{m-1} \circ H_N, \tag{2.6}$$

$$\xi_j D_j (D_{j-1} - D_{j+1}) L^{m} \circ H_N = (D_{j-1} - D_{j+1}) L^{m} \circ H_N. \tag{2.7}$$

Then

$$\frac{\partial}{\partial x}(L^{m} \circ H_N) = \sum_{j=1}^{N} \xi_{j,x} (D_j \circ L^{m} \circ H_N) = \sum \frac{\xi_{j,x}}{\xi_j} \circ \xi_j D_j L^{m} \circ H_N. \tag{2.8}$$

By (2.6) and $(\partial/\partial x) H_N = 0$,

$$= -m \sum \frac{\xi_{j,x}}{\xi_j} \circ D_j (D_{j-1} + D_{j+1}) L^{m-1} \circ H_N$$

$$= -m \sum \left(\frac{\xi_{j,x}}{\xi_j} + \frac{\xi_{j+1,x}}{\xi_{j+1}} \right) D_j D_{j+1} L^{m-1} \circ H_N.$$

By (2.1)

$$= -m \sum (\xi_j - \xi_{j+1}) D_j D_{j+1} L^{m-1} \circ H_N$$

$$= -m \sum \xi_j D_j (D_{j+1} - D_{j-1}) L^{m-1} \circ H_N.$$

By (2.7)

$$= -m \sum_{j=1}^{N} (D_{j+1} - D_{j-1}) L^{m-1} \circ H_N$$

and by periodicity mod N,

$$\equiv 0.$$

Let

$$N = 2k + 2. \tag{2.9}$$

Then for (2.1) we have one Casimir, H_N; k integrals $\{H_{2k+2-2m}; m = 1,2,3,...,k\}$ and one constraint

$$C_1 = \sum_{j=1}^{N} (-1)^j \xi_j = 0. \tag{2.10}$$

We claim that (2.1) is a k-dimensional Hamiltonian system

$$A \begin{pmatrix} \xi_{j,x}/\xi_j \\ \vdots \\ \xi_{N,x}/\xi_N \end{pmatrix} = B \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \nabla_{\xi} H_1, \tag{2.11}$$

where

$$H_1 = \sum_{j=1}^{N} \xi_j \tag{2.12}$$

and A, B are defined in Sec. I.

The higher-order flows

$$A \begin{pmatrix} \xi_{1,x}/\xi_1 \\ \vdots \\ \xi_{N,x}/\xi_N \end{pmatrix} = B \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \nabla_{\xi} L^{m} \circ H_N \tag{2.13}$$

for $m = 1,2,3,...,k$ are in involution and consistent with the constraint (2.10).

Observation 1: Since A and B are circulant matrices they commute

$$AB = BA. \tag{2.14}$$

Also from (1.19) and (1.21)

$$A\hat{a}_0 = 0, \quad B\hat{a}_0 = 2\hat{a}_0. \tag{2.15}$$

Then, contraction of (2.13) with

$$\nabla_{\xi} C_1 = -\hat{a}_0 \tag{2.16}$$

obtains the *constraints*

$$\hat{a}_0^t \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \nabla_{\xi} \circ H_N = 0 . \quad (2.17)$$

However, for even N, (2.17) vanishes identically [by (2.6)]. The constraint (2.10) is consistent [preserved by (2.13)].

Now, assume (for fixed m) (2.13) and consider

$$\frac{\partial}{\partial x} L^{\,p} \circ H_N = \sum_{j=1}^{N} \xi_{j,x} D_j L^{\,p} \circ H_N . \quad (2.18)$$

Then, using (2.13) and (2.6), (2.7) to raise and lower indices (p,m), it is evident that (2.12) vanishes. That is, the flows commute. For instance,

$$\begin{aligned} \frac{\partial}{\partial x} L^{\,p} \circ H_N &= -p \sum_{j=1}^{N} \{D_j L^{\,m} \circ H_N\}\{(D_{j+1} - D_{j-1}) L^{\,p-1} \circ H_N\} \\ &= p \sum \{(D_{j+1} - D_{j-1}) L^{\,m} \circ H_N\}\{D_j L^{\,p-1} \circ H_N\} \\ &= p \sum \{D_j (D_{j+1} - D_{j-1}) L^{\,m} \circ H_N\}\{\xi_j D_j L^{\,p-1} \circ H_N\} \\ &= -p(p-1) \sum \{D_j (D_{j+1} - D_{j-1}) L^{\,m} \circ H_N\} \circ \{D_j (D_{j+1} + D_{j-1}) L^{\,p-2} \circ H_N\} \\ &= \frac{p(p-1)}{m+1} \sum \{(D_{j+1} - D_{j-1}) L^{\,p-2} \circ H_N\}\{D_j L^{\,m+1} \circ H_N\} . \end{aligned} \quad (2.19)$$

It is useful to have a somewhat more explicit form of Eqs. (2.11) and (2.13). We define the circulant projection P onto the null space of A as

$$P = \text{circ}[1, -1, 1, -1, \ldots, 1, -1] . \quad (2.20)$$

Then

$$P^2 = P, \quad P\hat{a}_0 = \hat{a}_0, \quad PA = AP \equiv 0 . \quad (2.21)$$

The conditional inverse G of A satisfies

$$AG = GA = I - P . \quad (2.22)$$

A nonunique solution to (2.22) is the antisymmetric matrix

$$G' = (1/2N)\, \text{circ}[N, -N+2, N-4, -N+6, \ldots, -2+N] . \quad (2.23)$$

We require that $GB = BG$ be antisymmetric and find that

$$G = G' - \tfrac{1}{2} P \quad (2.24)$$

satisfies (2.22) and

$$GB = BG = \Omega_{k+1} , \quad (2.25)$$

where

$$\Omega_{k+1}^t = -\Omega_{k+1} . \quad (2.26)$$

By evaluation

$$\Omega_{k+1} = \begin{pmatrix} J_{k+1} & M_{k+1} \\ M_{k+1} & J_{k+1} \end{pmatrix}, \quad (2.27)$$

$$J_{k+1} = \frac{1}{k+1} \begin{pmatrix} 0 & -k & k-1 & -k+2 & k-3 & \cdots & (-)^k \\ k & 0 & -k & k-1 & -k+2 & \cdots & (-1)^k 2 \\ -k+1 & k & 0 & -k & k-1 & \cdots & \\ k-2 & -k+1 & k & 0 & -k & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & -k+1 & k & 0 & -k & k-1 \\ & & & -k+1 & k & 0 & -k \\ & & & & -k+1 & k & 0 \end{pmatrix}, \quad (2.28)$$

$$M_{k+1} = \frac{1}{k+1}\begin{pmatrix} 0 & (-1)^{k+1} & (-1)^k 2 & (-1)^{k+1}3 & \cdots & k \\ (-1)^k & 0 & (-1)^{k+1} & (-1)^k 2 & \cdots & -k+1 \\ (-1)^{k+1}2 & (-1)^k & 0 & (-1)^{k+1} & & \\ & \ddots & \ddots & \ddots & & \end{pmatrix}. \tag{2.29}$$

Applying G to (2.11) defines the equation to a scalar multiple of $\hat{a}_0$. This is uniquely determined by requiring that the constraint (2.10) be preserved. This obtains

$$\hat{\xi}_x = M_{\hat{\xi}} \nabla_{\hat{\xi}} H_1 - \frac{1}{H_1}\{C_1;H_1\}\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}\nabla_{\hat{\xi}} C_1, \tag{2.30}$$

where

$$M_{\hat{\xi}} = \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}\Omega_{k+1}\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}, \tag{2.31}$$

$$H_1 = \sum_{j=1}^N \xi_j,$$

$$C_1 = -\sum_{j=1}^N (-1)^j \xi_j,$$

$$\{C;H\} = (\nabla_{\hat{\xi}} C)^t M_{\hat{\xi}} \nabla_{\hat{\xi}} H. \tag{2.32}$$

Equations (2.30) imply that

$$\frac{\partial}{\partial x} C_1 = 0, \tag{2.33}$$

$$\frac{\partial}{\partial x} H_1 = -\frac{C_1}{H_1}\{C_1;H_1\}, \tag{2.34}$$

and $C_1 \equiv 0$, $(\partial/\partial x)H_1 = 0$ if $C_1 = 0$ when $x = 0$.

The corresponding equations for the higher-order flows are

$$\hat{\xi}_x = M_{\hat{\xi}} \nabla_{\hat{\xi}} L^{m} \circ H_N - \frac{1}{H_1}\{C_1;L^{m}\circ H_N\}\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}\nabla_{\hat{\xi}} C_1, \tag{2.35}$$

using (2.31) and (2.32). By direct calculation, using most of the previous results, it can be shown that

$$-mBA^t\nabla_{\hat{\xi}} L^{m-1}\circ H_N = M_{\hat{\xi}} \nabla_{\hat{\xi}} L^{m}\circ H_N - \frac{1}{H_1}\{C_1;L^{m}\circ H_N\}\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}\nabla_{\hat{\xi}} C_1, \tag{2.36}$$

where

$$BA^t = A^tB = \text{circ}[0,-1,0,\ldots,0,1]. \tag{2.37}$$

In effect, (2.36) is the (mth flow) dual-Hamiltonian formulation for the periodic fixed points when N is even. The dual-Hamiltonian structure for odd N is found in Ref. 4. The mth flow of Eq. (2.35) is dual to

$$\hat{\xi}_x = -m\Omega' \nabla_{\hat{\xi}} L^{m-1}\circ H_N, \tag{2.38}$$

where $\Omega' = BA^t$.

Observation 2: Both H_1 and C_1 are Casimir invariants of (2.37) and (2.38). Also, (2.36) does not determine a dual-Hamiltonian structure for (2.30) since $H_1 \neq L^{m}\circ H_N$ for any m. Here H_1 is an invariant of (2.30) and (2.35) if and only if C_1 vanishes.

The periodic fixed points of the Bäcklund transformation for the KdV equation are completely integrable (generalized) Hamiltonian systems. In the Appendix we show the transformation to canonical coordinates for these systems.

B. The Kac–Van Moerbeke system

The completely integrable, periodic Kac–Van Moerbeke (KM) system is[6,7]

$$\hat{\theta}_x = \Omega_{km} \nabla_{\hat{\theta}} H(e^{\theta}), \tag{2.39}$$

where

$$H = \sum_{j=1}^N e^{\theta_j}, \tag{2.40}$$

$$\Omega_{km} = BA^t. \tag{2.41}$$

[See (2.37).] For any N let

$$\hat{\theta} = -A\hat{\beta}, \tag{2.42}$$

i.e., $\theta_j = -\beta_j - \beta_{j+1}$. Then, under this change of variable,

$$\nabla_{\hat{\beta}} = -A^t\nabla_{\hat{\theta}} \tag{2.43}$$

and

$$A\hat{\beta}_x = B\nabla_{\hat{\beta}} H(e^{-\beta_j-\beta_{j+1}}), \tag{2.44}$$

where

$$H = \sum_{j=1}^N e^{-\beta_j-\beta_{j+1}}. \tag{2.45}$$

Let

$$\beta_j = \ln \xi_j. \tag{2.46}$$

Then

$$A\begin{pmatrix} \xi_{1,x}/\xi_1 \\ \vdots \\ \xi_{N,x}/\xi_N \end{pmatrix} = B\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix}\nabla_{\hat{\xi}} H, \tag{2.47}$$

where

$$H = \sum_{j=1}^N \frac{1}{\xi_j \xi_{j+1}} = \frac{L\circ H_N}{H_N}. \tag{2.48}$$

Comparing (2.47) and (1.18), the periodic KM system is, in a sense to be examined further, equivalent to the KdV periodic fixed point system. That is, the KM system is a higher-order flow of the KdV system [with Hamiltonian (2.48) an integral of these flows].

When N is *even*, the singularity of A requires that

$$\sum_{j=1}^{N} (-1)^j \theta_j = 0 . \tag{2.49}$$

Therefore, for odd N the KM and KdV systems are completely equivalent, and for even N the KdV system (2.47) is equivalent to KM systems satisfying this condition. Note that

$$H' = \sum_{j=1}^{N} \theta_j , \tag{2.50}$$

$$H'' = \sum_{j=1}^{N} (-1)^j \theta_j \tag{2.51}$$

are Casimir invariants of the KM system and $H'' = 0$ is the condition (2.49).

On the other hand, it is well known that for even N, the KM system is equivalent to the Toda lattice.[7] That is, for even N, let

$$\hat{\theta} = B\hat{p} \tag{2.52}$$

and find that

$$p_{j,x} = e^{\theta_{j-1}} + e^{\theta_j} - \alpha , \tag{2.53}$$

where α is a constant. Then, using

$$A'\hat{\theta} = A'B\hat{p} \tag{2.54}$$

and (2.39) find that

$$p_{j,xx} = e^{-p_j + p_{j-2}} - e^{-p_{j+2} + p_j} , \tag{2.55}$$

where $j = 1,2,3,...$ (mod N). Since N is even, (2.55) shows that the even and odd components of $\hat{p}$ decouple. Therefore, if

$$\hat{Q} = \begin{pmatrix} p_1 \\ p_3 \\ \vdots \\ p_{N-1} \end{pmatrix}, \quad \hat{Q}' = \begin{pmatrix} p_2 \\ p_4 \\ \vdots \\ p_N \end{pmatrix}, \tag{2.56}$$

the $N/2$ vectors $(\hat{Q},\hat{Q}')$ each satisfy the Toda lattice equations

$$Q_{j,xx} = e^{-Q_j + Q_{j-1}} - e^{-Q_{j+1} + Q_j} , \tag{2.57}$$

where $j = 1,2,3,...$ [mod $(N/2)$]. Note that the period of (2.57) is one-half of N, where N is the period of the KM-KdV systems.

From the singularity of the transformation (2.52),

$$H' = \sum_{j=1}^{N} \theta_j = 0 . \tag{2.58}$$

Previously, for (2.42), $H'' = 0$. Therefore, when both Casimirs vanish there is an equivalence between the KdV system and the Toda lattice. Condition (2.58) is trivial since it can be verified for an arbitrary solution of (2.39) by a suitable scaling in x and translation in $\hat{\theta}$. The transformation between (2.48) (KdV) and (2.55) (Toda) is

$$-A \begin{pmatrix} \ln \xi_1 \\ \vdots \\ \ln \xi_N \end{pmatrix} = B\hat{p} . \tag{2.59}$$

This requires

$$H_N = \prod_{j=1}^{N} \xi_j = 1 , \tag{2.60}$$

$$\sum_{j=1}^{N} (-1)^j p_j = 0 . \tag{2.61}$$

Again, by a suitable scaling (2.60) is trivial while the nontrivial (2.61) is equivalent to (2.58). In terms of (2.56), (2.61) is

$$\sum_{j=1}^{N/2} Q_j = \sum_{j=1}^{N/2} Q'_j . \tag{2.62}$$

Therefore (2.47) is equivalent to (two solutions of) (2.57) [which satisfy (2.62)].

To sum up: (1) For N odd the KM and KdV systems are completely equivalent; (2) for N even, the Toda lattice is completely equivalent to the system (2.47); (3) For even N, the KM system is, subject to the consistent condition (2.49), equivalent to the system (2.47); and (4) for N even, system (2.47) is, subject to the consistent constraint (2.10), a higher-order flow of the KdV system.

To put things somewhat differently, every solution of the KdV system is also a solution of the KM and/or Toda system.

III. THE BOUSSINESQ SYSTEMS

A. Existence of integrals

The (completely integrable) Boussinesq equation

$$U_{tt} = -\frac{\partial^2}{\partial x^2}\left(\frac{U_{xx}}{3} + U^2\right) \tag{3.1}$$

has the Bäcklund transformation[8]

$$U = 2\frac{\partial^2}{\partial x^2} \ln \phi + U_2 , \tag{3.2}$$

where ϕ satisfies the Schwarzian Boussinesq equation

$$\frac{\partial}{\partial t}\left(\frac{\phi_t}{\phi_x}\right) + \frac{1}{3}\frac{\partial}{\partial x}\left(\{\phi;x\} + \frac{3}{2}\left(\frac{\phi_t}{\phi_x}\right)^2\right) = 0 . \tag{3.3}$$

For the definition of the Schwarzian derivative see (1.5).

Equation (3.3) is invariant under the Möbius group

$$\phi = (a\psi + b)/(c\psi + d) , \quad ad - bc = 1 \tag{3.4}$$

and has the Bäcklund transformation[8]

$$\frac{\phi_{xx}}{\phi_x} = -\frac{1}{2}\frac{\psi_{xx}}{\psi_x} \mp \frac{3}{2}\frac{\psi_t}{\psi_x} , \tag{3.5}$$

$$\frac{\phi_t}{\phi_x} = \pm\frac{1}{2}\frac{\psi_{xx}}{\psi_x} - \frac{1}{2}\frac{\psi_t}{\psi_x} . \tag{3.6}$$

Equations (3.5) and (3.6) are a strong BT for (3.3) in that the compatibility condition

$$\phi_{xxt} = \phi_{txx} \tag{3.7}$$

is satisfied if and only if ψ satisfies (3.3). Here $\psi_{xxt} = \psi_{txx}$ requires ϕ also must satisfy (3.3).

The effective BT is the composition of (3.4)–(3.6). Say, in (3.5) and (3.6)

$$\phi = \phi_{j+1}, \quad \psi = -1/\phi_j . \tag{3.8}$$

Then, with

$$v_j = \phi_{j,xx}/\phi_{j,x}, \quad z_j = \phi_{j,t}/\phi_{j,x}, \tag{3.9}$$

we have [with the upper signs in (3.5) and (3.6)]

$$v_{j+1} + \tfrac{1}{2} v_j + \tfrac{3}{2} z_j = \phi_{j,x}/\phi_j, \tag{3.10}$$

$$z_{j+1} + \tfrac{1}{2} z_j - \tfrac{1}{2} v_j = -\phi_{j,x}/\phi_j. \tag{3.11}$$

As was the case for the KdV systems, let

$$\xi_j = \phi_{j,x}/\phi_j \tag{3.12}$$

and define

$$A = \text{circ}[1,1,0,\dots,0,1], \tag{3.13}$$

$$B = \text{circ}[0,-1,0,\dots,0,1]. \tag{3.14}$$

Then, the BT (3.10), (3.11) at the periodic fixed points of order N,

$$j = 1,2,3,\dots \ (\text{mod } N), \quad \xi_{j+N} = \xi_j \tag{3.15}$$

are solutions of the system

$$A \begin{pmatrix} \xi_{1,x}/\xi_1 \\ \vdots \\ \xi_{N,x}/\xi_N \end{pmatrix} = B\hat{\xi}, \tag{3.16}$$

where

$$\hat{\xi} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_N \end{pmatrix}. \tag{3.17}$$

In terms of the preceding

$$\hat{z} = -A\hat{\xi} \tag{3.18}$$

and

$$v_{j+1} + z_{j+1} + 2z_j = 0. \tag{3.19}$$

Observation 3: Except for the definition of the circulant matrices (A,B) the Boussinesq systems (3.18) have the same form as the KdV systems.

In component notation Eqs. (3.16) are

$$\frac{\xi_{j+1,x}}{\xi_{j+1}} + \frac{\xi_{j,x}}{\xi_j} + \frac{\xi_{j-1,x}}{\xi_{j-1}} = -\xi_{j+1} + \xi_{j-1}, \tag{3.20}$$

where $j = 1,2,3,\dots$ (mod N).

Since all circulant matrices commute[5] A and B have a set of simultaneous eigenvectors

$$\hat{\beta}_k = \begin{pmatrix} 1 \\ r^k \\ r^{2k} \\ \vdots \\ r^{(N-1)k} \end{pmatrix}, \tag{3.21}$$

for $k = 0,1,2,\dots,N-1$, where

$$r = \exp^{(2\pi i/N)}. \tag{3.22}$$

The spectra of A and B are, respectively,

$$\lambda_A = 1 + r^k + 1/r^k, \tag{3.23}$$

$$\lambda_B = -r^k + 1/r^k, \tag{3.24}$$

$k = 0,1,2,\dots,n-1$, with eigenvector (3.21). From (3.23) A is singular when

$$r^k = \exp^{\pm 2\pi i/3} \tag{3.25}$$

which, by (3.22) can occur iff

$$N = 3k. \tag{3.26}$$

On the other hand, B, for any N, has the null vector

$$\hat{\beta}_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{3.27}$$

and, when $N(=2k)$ is even, B has the additional null vector

$$\hat{\beta}_k = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \vdots \\ 1 \\ -1 \end{pmatrix}. \tag{3.28}$$

As was true for the KdV systems, the null vectors of A induce constraints and the null vectors of B induce Casimir invariants for the system (3.16). When N is not a multiple of 3, A is invertible.

For

$$N = 3k+1, \tag{3.29}$$

$$A^{-1} = -\text{circ}[0,0,1,0,0,1,0,0,\dots,0,1,0] + \tfrac{1}{3}\text{circ}[1,1,1,\dots,1,1], \tag{3.30}$$

and

$$\Omega = A^{-1}B = \text{circ}[0,-1,0,1,-1,0,1,\dots,-1,0,1]. \tag{3.31}$$

For

$$N = 3k+2, \tag{3.32}$$

$$A^{-1} = \text{circ}[0,1,0,0,1,0,0,1,\dots,0,0,1] - \tfrac{1}{3}\text{circ}[1,1,1,\dots,1,1], \tag{3.33}$$

and

$$\Omega = A^{-1}B = \text{circ}[0,0,-1,1,0,-1,1,0,\dots,-1,1,0]. \tag{3.34}$$

In both instances

$$\Omega^t = -\Omega \tag{3.35}$$

and the null vector of Ω are the null vectors of B.

For any N, associated with the null vector (3.27), the system (3.16) has the Casimir integral

$$H_N = \prod_{j=1}^{N} \xi_j. \tag{3.36}$$

That is,

$$\frac{\partial}{\partial x} H_N = \sum_{j=1}^{N} \xi_{j,x} \frac{\partial}{\partial \xi_j} H_N = H_N \sum \frac{\xi_{j,x}}{\xi_j} = \frac{H_N}{3} \sum \left(\frac{\xi_{j+1,x}}{\xi_{j+1}} + \frac{\xi_{j,x}}{\xi_j} + \frac{\xi_{j-1,x}}{\xi_{j-1}} \right), \tag{3.37}$$

by (3.20)

$$=\frac{H_N}{3}\sum_{j=1}^{N}(-\xi_{j+1}+\xi_{j-1})\equiv 0. \qquad (3.38)$$

We define the operator

$$L=\sum_{j=1}^{N}D_jD_{j+1}D_{j+2}, \qquad (3.39)$$

where $D_j=\partial/\partial\xi_j$, and note the following identities for $m=0,1,2,\ldots$ and each j

$$\xi_jD_jL^{m}\circ H_N=L^{m}\circ H_N - mD_j(D_{j-2}D_{j-1}+D_{j-1}D_{j+1} + D_{j+1}D_{j+2})L^{m-1}\circ H_N, \qquad (3.40)$$

$$\xi_jD_j(D_{j-2}D_{j-1}-D_{j+1}D_{j+2})L^{m}\circ H_N = (D_{j-2}D_{j-1}-D_{j+1}D_{j+2})L^{m}\circ H_N. \qquad (3.41)$$

With these we show that

$$H_{N-3m}=L^{m}\circ H_N \qquad (3.42)$$

for $m=0,1,2,\ldots,(N/3)$ are integrals of (3.16). That is,

$$\frac{\partial}{\partial x}L^{m}\circ H_N=\sum_{j=1}^{N}\xi_{j,x}\frac{\partial}{\partial\xi_j}L^{m}\circ H_N = \sum\frac{\xi_{j,x}}{\xi_j}\xi_jD_jL^{m}\circ H_N, \qquad (3.43)$$

by (3.36) and (3.40)

$$=-m\sum\left(\frac{\xi_{j+1,x}}{\xi_{j+1}}+\frac{\xi_{j,x}}{\xi_j}+\frac{\xi_{j-1,x}}{\xi_{j-1}}\right)\times D_{j-1}D_jD_{j+1}L^{m-1}\circ H_N,$$

by (3.20)

$$=m\sum(\xi_{j+1}-\xi_j)D_{j-1}D_jD_{j+1}L^{m-1}\circ H_N$$

$$=m\sum\xi_j(D_{j-2}D_{j-1}D_j-D_jD_{j+1}D_{j+2})\times L^{m-1}\circ H_N,$$

by (3.41)

$$=\sum_{j=1}^{N}(D_{j-2}D_{j-1}-D_{j+1}D_{j+2})L^{m-1}\circ H_N,$$

and by periodicity

$$\frac{\partial}{\partial x}L^{m}\circ H_N=0.$$

Therefore, when

$$N=3k+1,\ 3k+2,\ 3k+3, \qquad (3.44)$$

we have, from (3.42), $k+1$ integrals.

Furthermore, when N is even,

$$N=2k, \qquad (3.45)$$

there is associated with the null vector (3.28) of B the Casimir integral

$$H_0=\prod_{j=1}^{k}\frac{\xi_{2j-1}}{\xi_{2j}}. \qquad (3.46)$$

By multiplication of the Casimirs (3.36) and (3.46) there are produced the equivalent Casimirs

$$H_k=\prod_{j=1}^{k}\xi_{2j-1}, \qquad (3.47)$$

$$H'_k=\prod_{j=1}^{k}\xi_{2j}. \qquad (3.48)$$

The index refers to the degree of the integral under a homogeneous scaling

$$\xi_j\to\lambda\xi_j. \qquad (3.49)$$

The result of applying the operator (3.39) to (3.47), (3.48) is null and a calculation reveals that applying (3.39)–(3.46) does not produce integrals.

Therefore for a system of size $N\in\{3k+1,\ 3k+2,\ 3k+3\}$ we have one or two (N even) Casimir integrals and k integrals (3.42). In what follows we shall find that system (3.16) has a Hamiltonian structure. Unlike the KdV systems, we do not find, by the above procedure, a sufficient number of integrals to show that (3.16) is completely integrable. By finding the consistent reductions [which preserve the form of (3.16)] we also will find the *missing* integrals.

First consider the systems (3.16) without constraints. That is, when N is not a multiple of 3,

$$N=3k+1 \qquad (3.50)$$

or

$$N=3k+2. \qquad (3.51)$$

From (3.31) or (3.34) system (3.16) is

$$\hat{\xi}_{,x}=\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\Omega\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\circ\nabla_{\hat{\xi}}H_1, \qquad (3.52)$$

where

$$\Omega=A^{-1}B,\quad \hat{\xi}=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_N\end{pmatrix},\quad H_1=\sum_{j=1}^{N}\xi_j, \qquad (3.53)$$

and $N=3k+1$ or $N=3k+2$.

The Hamiltonian form of Eqs. (3.52) with Hamiltonian H_1 and cosymplectic form

$$M_{\hat{\xi}}=\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\Omega\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}, \qquad (3.54)$$

is evident. Recall that Ω is antisymmetric and the Poisson bracket is

$$\{G,H\}=(\nabla_{\hat{\xi}}G)^tM_{\hat{\xi}}\nabla_{\hat{\xi}}H. \qquad (3.55)$$

That (3.55) verifies the Jacobi identity is a simple consequence of the change of variables $\xi_j=e^{\theta_j}$ (Miura transformation between Hamiltonian systems).[4] It is perhaps worth noting at this time that

$$M_{\hat{\xi}}\nabla_{\hat{\xi}}L^{m}\circ H_N=-m\Omega_{\hat{D}}\nabla_{\hat{\xi}}L^{m-1}\circ H_N, \qquad (3.56)$$

where

$$\Omega_{\hat{D}} = \begin{pmatrix} 0 & 0 & -D_2 & 0 & \cdots & D_N & 0 \\ 0 & 0 & 0 & -D_3 & \cdots & 0 & D_1 \\ D_2 & 0 & 0 & 0 & -D_4 & \cdots & 0 \\ \vdots & \ddots & & & & \ddots & \\ & & & & & & -D_{N-1} \\ -D_N & & & & & & 0 \\ 0 & -D_1 & & & D_{N-1} & 0 & 0 \end{pmatrix}, \tag{3.57}$$

when $N = 3k + 1, 3k + 2$. The higher-order equations associated with the integrals (3.42) are

$$\hat{\xi}_{,x} = M_{\hat{\xi}} \nabla_{\hat{\xi}} L^{m} \circ H_N . \tag{3.58}$$

From the form of Eqs. (3.52) and (3.57) and inspection of (3.31) and (3.34) a consistent reduction of (3.52) is to set any three consecutive terms equal to zero. That is,

$$\xi_j = \xi_{j+1} = \xi_{j+2} = 0 \tag{3.59}$$

for fixed j. This preserves the form of the equations with

$$\begin{aligned} N &= 3k + 1 \rightarrow 3(k-1) + 1 , \\ N &= 3k + 2 \rightarrow 3(k-1) + 2 , \end{aligned} \tag{3.60}$$

and

$$L \circ H_N \rightarrow H_{N-3} . \tag{3.61}$$

When N $(=2k)$ is even, another consistent reduction of (3.52) and (3.58) is to set the even (or the odd) components of $\hat{\xi}$ equal to zero. That is,

$$\xi_{2j} = 0, \quad \xi_{2j-1} = \xi'_j \tag{3.62}$$

or

$$\xi_{2j-1} = 0, \quad \xi_{2j} = \xi'_j , \tag{3.63}$$

where $j = 1,2,...,k$ $N = 2k$. Note that (1) when $N = 2k = 3l + 1$, then $N/2 = k = 3p + 2$, where $l = 2p + 1$, and (2) when $N = 2k = 3l + 2$, then $N/2 = k = 3p + 1$, where $l = 2p$. Therefore, this reduction, while halving the period of the fixed point, exchanges the systems defined by (3.31) and (3.34). That the reduction is consistent can be seen immediately from (3.31) and (3.34).

In Table I we present a list of systems with $N \neq 3k$ and $N \leqslant 22$. The degree of freedom d as a Hamiltonian system (after subtracting the Casimirs) and the degree [under scaling (3.49)] of the complete (with regard to d) set of homogeneous, independent integrals are represented. The primary integrals are (3.42). We will show the existence and form of the secondary integrals. When N is even the highest weight secondary integral is the Casimir of the form (3.47) and (3.48).

For any N the degrees of the primary integrals are determined by (3.42). That is, in this case,

$$1,4,7,...,3k + 1 \tag{3.64}$$

or

$$2,5,8,...,3k + 2 . \tag{3.65}$$

When $N = 2l = 3k + 1$, the degrees of the secondary integrals are $(l = 3p + 2, k = 2p + 1)$

$$2,5,8,...,3p + 2 , \tag{3.66}$$

and when $N = 3k + 1 = 2l + 1$, $k = 2p$, $l = 3p$, the degrees are

$$2,5,8,...,3p - 1 . \tag{3.67}$$

When $N = 3k + 2 = 2l$, $k = 2p$, $l = 3p + 1$, the degrees are

$$1,4,7,...,3p + 1 , \tag{3.68}$$

and when $N = 3k + 2 = 2l + 1$, $k = 2p + 1$, $l = 3p + 2$ the degrees are

$$1,4,7,...,3p + 1 . \tag{3.69}$$

Now, for any N the primary integrals are known from (3.42). We claim that the secondary integrals at $N = N'$ are the primary integrals at $N = 2N'$ that survive reduction (3.62). From the preceding it is immediate that any primary integral that does survive the reduction will be a secondary integral of the reduced system. Note that reduction does not

TABLE I. Degree of primary/secondary integrals.

N	d								
4	1	1	4						
		2							
5	2	2	5						
		1							
7	3	1	4	7					
		2							
8	3	2	5	8					
		1	4						
10	4	1	4	7	10				
		2	5						
11	5	2	5	8	11				
		1	4						
13	6	1	4	7	10	13			
		2	5						
14	6	2	5	8	11	14			
		1	4	7					
16	7	1	4	7	10	13	16		
		2	5	8					
17	8	2	5	8	11	14	17		
		1	4	7					
19	9	1	4	7	10	13	16	19	
		2	5	8					
20	9	2	5	8	11	14	17	20	
		1	4	7	10				
22	10	1	4	7	10	13	16	19	22
		2	5	8	11				

change the (weight) degree of a nonvanishing term.

It is convenient to let N be even and greater than 4. Then N is of the form (for some k)

$$N=4k \tag{3.70}$$

or

$$N=4k+10\,. \tag{3.71}$$

Without loss of generality let

$$\xi_{2j}=0 \tag{3.72}$$

for $j=1,2,\ldots$ in (3.42). Any application of L that does not leave embedded even-order terms must remove terms in groups that are odd multiples of 3. Say, in

$$H_N=\xi_1\xi_2\xi_3\xi_4\xi_5\xi_6\xi_7\xi_8\xi_9\xi_{10}\xi_{11}\cdots, \tag{3.73}$$

keep ξ_1 and remove $\xi_2\xi_3\xi_4$, keep ξ_5 and remove $\xi_6\xi_7\xi_8$, etc. Or keep ξ_1 and remove $\xi_2\xi_3\xi_4\xi_5\xi_6\xi_7\xi_8\xi_9\xi_{10}$, keep ξ_{11}, etc.

Each excision will account for i terms,

$$i=4,10,16,22,28,\ldots, \tag{3.74}$$

where one odd indexed term is retained and $i-1$ terms are removed from (3.73). It is immediate that when $N=4k$ the highest weight terms that survive reduction (3.72) will have degree k (by k excisions with index $i=4$). Also, since $N=4k=16+4(k-4)$, there also result the next highest weight terms of degree $k-3$ ($k-4$ excisions with index 4 and one excision of index 16). In this way, it is evident that when $N=4k$ the terms (secondary integrals) in (3.42) of degree

$$k,k-3,k-6,\ldots \tag{3.75}$$

are uniquely the terms that do not vanish under reduction.

In the same manner, when $N=4k+10$, the degree of the surviving terms are

$$k+1,k-2,k-5,\ldots\,. \tag{3.76}$$

Therefore, we have established the existence of the secondary integrals (when N is not a multiple of 3). The number of primary and secondary integrals equals the degree of freedom of the Hamiltonian system.

For N a multiple of 3 it is not yet established that (3.62) is a consistent reduction. This will be examined later.

The form of the secondary integrals is readily obtained. For instance, let $N=22=4\circ 3+10$. Then by the preceding the highest term surviving reduction has degree 4 and

$$\begin{aligned} H_4(\xi_2=\xi_4=&\cdots\xi_N=0)\\ &=L^{6}\circ H_N(\xi_2=\cdots=0)\\ &=\sum_{j+1}^{11}\xi_{2j-1}\,\xi_{2j+3}\,\xi_{2j+7}\,\xi_{2j+11}\,. \end{aligned} \tag{3.77}$$

Therefore, when $N=11$ and relabeling terms $\xi_{2j-1}\to\xi_j$ for $j=1,2,\ldots,11$ we obtain the integral

$$H_4=\sum_{j=1}^{11}\xi_j\,\xi_{j+2}\,\xi_{j+4}\,\xi_{j+6}\,. \tag{3.78}$$

In the same way, when $N=13$

$$H_5=\sum_{j=1}^{13}\xi_j\,\xi_{j+2}\,\xi_{j+4}\,\xi_{j+6}\,\xi_{j+8}\,. \tag{3.79}$$

An explicit formula for the secondary integrals can be obtained through combinatorial considerations. We defer from further discussion of this point except to note that for any N (not a multiple of 3)

$$H_2=\tfrac{1}{2}\hat{\xi}^{\,t}A^{-1}\hat{\xi}\,. \tag{3.80}$$

See (3.30) and (3.33).

Observation 4: The results of Yoshida[9] may be used to find the hypothetical degree of integrals for the scale-invariant system (3.20). An application of this method to (3.20) is particularly interesting in that for the most obvious form of singularity the *leading orders* and *resonances* numerically coincide, thereby greatly reducing the computational effort.

B. The involution of integrals

For the KdV systems of Secs. I and II and Ref. 4 it is, in general (for any N), quite simple to demonstrate the involution of the integrals since the systems have the dual-Hamiltonian structure (2.36) for even N and

$$\hat{\xi}_{,x}=M_{\hat{\xi}}\nabla_{\hat{\xi}}L^{m}\circ H_N=-mBA^{t}\nabla_{\hat{\xi}}L^{m-1}\circ H_n \tag{3.81}$$

for odd N.[4] By the usual argument with (2.38), (3.81) for raising and lowering indices[4] it is readily seen that

$$\{L^{p}\circ H_N,L^{m}\circ H_N\}=(\nabla_{\hat{\xi}}L^{p}\circ H_N)^{t}M_{\hat{\xi}}\nabla_{\hat{\xi}}L^{m}\circ H_N=0\,. \tag{3.82}$$

That is, the Poisson bracket vanishes.

On the other hand, the corresponding formulation (3.56) for the Boussinesq systems,

$$\hat{\xi}_{,x}=M_{\hat{\xi}}\nabla_{\hat{\xi}}L^{m}\circ H_N=-m\Omega_{\hat{D}}\nabla_{\hat{\xi}}L^{m-1}\circ H_N\,, \tag{3.83}$$

does not obtain a transparent procedure for demonstrating the involution of the (primary) integrals. That is, it is not evident that

$$\begin{aligned}\{L^{p}\circ H_N,&L^{m}\circ H_N\}\\ &=m\sum_{j=1}^{N}D_jL^{p}\circ H_N\\ &\quad\times(-D_{j+1}D_{j+2}+D_{j-2}D_{j-1})\circ L^{m-1}\circ H_N\end{aligned} \tag{3.84}$$

must vanish, since $\Omega_{\hat{D}}$ is a differential operator and the usual raising/lowering arguments do not apply.

Various equivalent forms of this Poisson bracket are

$$\{L^{p}\circ H_N,L^{m}\circ H_N\}=-mp\widehat{W}_p^{t}A^{t}B\widehat{W}_m\,, \tag{3.85}$$

where (A,B) are as defined in Sec. III A,

$$A^{t}B=BA^{t}=\mathrm{circ}[0,-1,-1,0,\ldots,0,1,1]\,, \tag{3.86}$$

$$\widehat{W}_m=\begin{pmatrix}c_1\\ c_2\\ \vdots\\ c_N\end{pmatrix}L^{m-1}\circ H_N\,, \tag{3.87}$$

and

$$c_j=D_{j-1}D_jD_{j+1},\quad D_j=\frac{\partial}{\partial x}\,. \tag{3.88}$$

In divergence form

$$\{L^{p}\circ N_N,L^{m}\circ H_N\}=m\nabla_{\hat{\xi}}\circ(L^{p}\circ H_N\widehat{V}_m)\,, \tag{3.89}$$

where

$$\hat{V}_m = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix} \circ L^{m-1} \circ H_N , \qquad (3.90)$$

$$d_j = D_{j+1} D_{j+2} - D_{j-2} D_{j-1} , \qquad (3.91)$$

$$\nabla_{\hat{\xi}} \circ \hat{V}_m = 0 . \qquad (3.92)$$

We consider the Poisson brackets of the primary and secondary integrals defined in Table I and thereafter. Of course, for any N and G a generic integral

$$\{H_1, G\} = 0 , \qquad (3.93)$$

$$\{H_N, G\} = 0 , \qquad (3.94)$$

and, when $N = 2k$, the Casimir H_k commutes with any G

$$\{H_k, G\} = 0 . \qquad (3.95)$$

Let H_2 be defined as in (3.80). Then for any m

$$\{H_2, L^m \circ H_N\} = 0 . \qquad (3.96)$$

Since A is symmetric,

$$\nabla_{\hat{\xi}} H_2 = A^{-1} \hat{\xi} , \qquad (3.97)$$

and by (3.84)

$$\begin{aligned} \{H_2, L^m \circ H_N\} &= m \sum_{j=1}^{N} D_j H_2 \\ &\quad \circ (D_{j+1} D_{j+2} - D_{j-2} D_{j-1}) L^{m-1} \circ H_N \\ &= m \hat{\xi}\, {}^t A^{-1} \begin{pmatrix} D_{N-1} D_N - D_2 D_3 \\ D_N D_2 - D_3 D_4 \\ \vdots \\ D_{N-2} D_{N-1} - D_1 D_2 \end{pmatrix} \\ &\quad \times L^{m-1} \circ H_N \\ &= m \hat{\xi}\, {}^t \begin{pmatrix} D_N D_1 - D_1 D_2 \\ \vdots \\ D_{N-1} D_N - D_N D_1 \end{pmatrix} \\ &\quad \times L^{m-1} \circ H_N , \end{aligned} \qquad (3.98)$$

and by (3.40)

$$\{H_2, L^m \circ H_N\} = 0 .$$

For any N

$$\{H_{N-3}, H_{N-6}\} = 0 . \qquad (3.99)$$

By (3.85)

$$\begin{aligned} \{H_{N-3}, H_{N-6}\} &= -2 \hat{W}_1^t A\, {}^t B \hat{W}_2 \\ &= -2 H_N^2 \circ \begin{pmatrix} 1/\xi_N \xi_1 \xi_2 \\ \vdots \\ 1/\xi_{N-1} \xi_N \xi_1 \end{pmatrix} A\, {}^t B \begin{pmatrix} 1/\xi_N \xi_1 \xi_2 & & \\ & \ddots & \\ & & 1/\xi_{N-1} \xi_N \xi_1 \end{pmatrix} \\ &\quad \circ \operatorname{circ}[0,0,0,1,1,\ldots,1,1,0,0] \begin{pmatrix} 1/\xi_N \xi_1 \xi_2 \\ \vdots \\ 1/\xi_{N-1} \xi_N \xi_1 \end{pmatrix} \end{aligned} \qquad (3.100)$$

and by direct calculation,

$$\{H_{N-3}, H_{N-6}\} = 0 .$$

We claim that when $N = 3k+1$ or $3k+2$, $\{H_p, H_q\} = 0$ and $p + q < N + 3$, then $\{H_p, H_q\} = 0$ when $N = 3k+4$ or $3k+5$. The reduction (3.59) and N not a multiple of 3 implies that when $N = (3k+4, 3k+5)$ $\{H_p, H_q\}$ must contain the factor H_N. Since degree of $\{H_p, H_q\} = p + q < N$, $\{H_p, H_q\}$ must vanish.

For instance, in Table I the above argument shows that, say $\{H_5, H_7\} = 0$ when $N = 13$, but does not show that $\{H_7, H_{11}\}$ must vanish when $N = 17$.

C. Constraints: $N = 3k$

When $N = 3k$ the matrix A is singular and contraction of (3.16) with the null vectors, $\hat{\beta}_k$ and $\hat{\beta}_{2k}$, defined by (3.21) obtains the constraints

$$\hat{\beta}_k \circ \hat{\xi} = 0 , \qquad (3.101)$$

$$\hat{\beta}_{2k} \circ \hat{\xi} = 0 . \qquad (3.102)$$

Equations (3.101) and (3.102) are equivalent to the system of real constraints

$$C_1 = \hat{c}_1 \circ \hat{\xi} = 0 , \qquad (3.103)$$

$$C_2 = \hat{c}_2 \circ \hat{\xi} = 0 , \qquad (3.104)$$

where

$$\hat{c}_1 = (2, -1, -1, 2, -1, -1, \ldots, 2, -1, -1) , \qquad (3.105)$$

$$\hat{c}_2 = (0, 1, -1, 0, 1, -1, \ldots, 0, 1, -1) . \qquad (3.106)$$

We note that $\hat{c}_1 \circ \hat{c}_2 = 0$.

The symmetric circulant matrix

$$P = (1/N) \operatorname{circ}[2, -1, -1, 2, -1, -1, \ldots, 2, -1, -1] \qquad (3.107)$$

satisfies the conditions

$$P^2 = P, \quad P\hat{c}_1 = \hat{c}_1, \quad P\hat{c}_2 = \hat{c}_2 , \qquad (3.108)$$

and is the projection onto the null space of A.

We require a pseudoinverse, G of A to verify the condition

$$GA = AG = I - P. \tag{3.109}$$

Let

$$G = (1/N)\,\mathrm{circ}[N-2,0,3-N,\ N-2,-3,6-N,N-2,-6,9-N,\ldots,\ N-2,-3(j-1),3j-N,\ldots,N-2,-N+3,0]. \tag{3.110}$$

Then G is symmetric and satisfies (3.109).

System (3.16) is a Hamiltonian system

$$A\begin{pmatrix}\xi_{1,x}/\xi_1\\ \vdots\\ \xi_{N,x}/\xi_N\end{pmatrix} = B\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\nabla_{\hat\xi}H_1, \tag{3.111}$$

where $H_1 = \Sigma_{j=1}^N \xi_j$. From the primary integrals

$$H_{N-3m} = L^{m}\circ H_N,\qquad H_N = \prod_{j=1}^N \xi_j, \tag{3.112}$$

the higher-order systems are

$$A\begin{pmatrix}\xi_{1,x}/\xi_1\\ \vdots\\ \xi_{N,x}/\xi_N\end{pmatrix} = B\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\nabla_{\hat\xi}H_{N-3m}. \tag{3.113}$$

From identity (3.40), it is generally true that

$$\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\nabla_{\hat\xi}L^{m}\circ H_N = \begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix}L^{m}\circ H_N - mA\begin{pmatrix}D_ND_1D_2\\ \vdots\\ D_{N-1}D_ND_1\end{pmatrix}\circ H_N. \tag{3.114}$$

Using (3.112) it is found, by contraction of (3.113) with $(\hat c_1,\hat c_2)$, that the constraints are trivial for the systems (3.113). That is, both sides vanish identically.

Now apply (3.110) to (3.111) and find that

$$\hat\xi_{,x} = M_{\hat\xi}\nabla_{\hat\xi}H_1 - \frac{1}{h}\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\times(\tfrac{1}{3}\{C_1,H_1\}\nabla_{\hat\xi}C_1 + \{C_2,H_1\}\nabla_{\hat\xi}C_2), \tag{3.115}$$

where

$$M_{\hat\xi} = \begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\Omega_{3k}\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}, \tag{3.116}$$

$$\Omega_{3k} = GB, \tag{3.117}$$

$$\{F,H\} = (\nabla_{\hat\xi}F)^t M_{\hat\xi}\nabla_{\hat\xi}H, \tag{3.118}$$

$$C_1 = \hat c_1\circ\hat\xi, \tag{3.119}$$

$$C_2 = \hat c_2\circ\hat\xi, \tag{3.120}$$

$$h = \hat c_2^t\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\hat c_2 = \hat c_1^t\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\hat c_1 - \frac{2}{3}C_1. \tag{3.121}$$

We note that

$$\hat c_1^t\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\hat c_2 = -C_2. \tag{3.122}$$

By evaluation

$$GB = (1/N)\,\mathrm{circ}[0,2N-5,2-N,6-N,\ 2N-8,-1-N,12-N,\ldots,\ 2N-3j-2,5-3j-N,6j-N,\ldots,\ N+1,8-2N,N-6,N-2,5-2N]. \tag{3.123}$$

The higher-order equations are

$$\hat\xi_{,x} = M_{\hat\xi}\nabla_{\hat\xi}H_{N-3m} - \frac{1}{h}\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\times(\tfrac{1}{3}\{C_1,H_{N-3m}\}\nabla_{\hat\xi}C_1 + \{C_2,H_{N-3m}\}\nabla_{\hat\xi}C_2). \tag{3.124}$$

After a calculation it is found that

$$-m\Omega_{\hat D}\nabla_{\hat\xi}H_{N+3-3m} = M_{\hat\xi}\nabla_{\hat\xi}H_{N-3m} - \frac{1}{h}\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\circ(\tfrac{1}{3}\{C_1,H_{N-3m}\}\nabla_{\hat\xi}C_1 + \{C_2,H_{N-3m}\}\nabla_{\hat\xi}C_2), \tag{3.125}$$

where $\Omega_{\hat D}$ is defined by (3.56).

By the above the higher-order equations (3.124) are

$$\hat\xi_{,x} = -m\Omega_{\hat D}\nabla_{\hat\xi}H_{N+3-3m}. \tag{3.126}$$

Compare with (3.56) and (3.58).

In general, we find that for systems (3.115) and (3.124) the reductions (3.59) or (3.62) are not consistent (form preserving). Therefore the secondary integrals do not seem to have the same structure when $N = 3k$ and when $N \neq 3k$.

IV. THE GENERIC SYSTEM

The KdV and Boussinesq systems are instances of the general system in component form

$$\frac{\xi_{j,x}}{\xi_j} + \frac{\xi_{j+1,x}}{\xi_{j+1}} + \cdots + \frac{\xi_{j+p,x}}{\xi_{j+p}} = \xi_j - \xi_{j+p}, \tag{4.1}$$

where $j = 1,2,\ldots \pmod N$. The KdV systems correspond to $p = 1$ and the Boussinesq to $p = 2$. Let the circulant forward shift matrix be

$$C = \mathrm{circ}[0,1,0,0,\ldots,0]. \tag{4.2}$$

In the N-vector form Eqs. (4.1) are

$$A\begin{pmatrix}\xi_{1,x}/\xi_1\\ \vdots\\ \xi_{N,x}/\xi_N\end{pmatrix}=B\hat{\xi} \tag{4.3}$$

with

$$A=I+C+\cdots+C^p, \tag{4.4}$$

$$B=I-C^p. \tag{4.5}$$

Let

$$r=\exp^{2\pi i/N}. \tag{4.6}$$

Then the eigenvectors of (4.4) and (4.5) are

$$\hat{\beta}_k=\begin{pmatrix}1\\ r^k\\ r^{2k}\\ \vdots\\ r^{(N-1)k}\end{pmatrix}. \tag{4.7}$$

The spectra of (A,B) are

$$\lambda_A=\frac{1-z^{p+1}}{1-z}=1+z+\cdots+z^p, \tag{4.8}$$

$$\lambda_B=1-z^p, \tag{4.9}$$

where for $k=0,1,2,\ldots,N-1$,

$$z=r^k=\exp^{2\pi ik/N}. \tag{4.10}$$

Here λ_B is null for some k, $0\leqslant k\leqslant N-1$, if for integer m, $0\leqslant m\leqslant p-1$,

$$kp=mN, \tag{4.11}$$

λ_A is null for some k if

$$k(p+1)=mN. \tag{4.12}$$

The Casimir integrals of (4.1) correspond to the null vectors of B. The null vectors of A produce the constraints. We note that for fixed N, the Casimir vectors for the systems with $p=l+1$ are the principal Casimir vector $\hat{\beta}_0$, and the set of constraint vectors for the systems with $p=l$.

Associated with the principal Casimir, for any N

$$H_N=\prod_{j=1}^{N}\xi_j \tag{4.13}$$

we find the principal integrals of (4.1)

$$H_{N-pm}=L^m\circ H_N, \tag{4.14}$$

where $m=0,1,2,\ldots$, and

$$L=\sum_{j=1}^{N}D_jD_{j+1}\cdots D_{j+p}. \tag{4.15}$$

The identities, for each j,

$$\xi_jD_jL^m\circ H_N=L^m\circ H_N-m(D_{j-p}D_{j+1-p}\cdots D_j+D_{j+1-p}\cdots D_jD_{j+1}+\cdots+D_jD_{j+1}\cdots D_{j+p})L^{m-1}\circ H_N, \tag{4.16}$$

$$\xi_j(D_{j-p}D_{j+1-p}\cdots D_j-D_jD_{j+1}\cdots D_{j+p})L^m\circ H_N=(D_{j-p}D_{j+1-p}\cdots D_{j-1}-D_{j+1}D_{j+2}\cdots D_{j+p})\times L^m\circ H_N \tag{4.17}$$

imply that

$$\frac{\partial}{\partial x}L^m\circ H_N=0. \tag{4.18}$$

That is,

$$\frac{\partial}{\partial x}L^m\circ H_N=\sum_{j=1}^{N}\xi_{j,x}D_jL^m\circ H_N=\sum\frac{\xi_{j,x}}{\xi_j}\xi_jD_jL^m\circ H_N,$$

by (4.16)

$$=-m\sum\frac{\xi_{j,x}}{\xi_j}(D_{j-p}D_{j+1-p}\cdots D_j+\cdots+D_jD_{j+1}\cdots D_{j+p})\circ L^{m-1}\circ H_N$$

$$=-m\sum\left(\frac{\xi_{j,x}}{\xi_j}+\cdots+\frac{\xi_{j+p,x}}{\xi_{j+p}}\right)\times D_jD_{j+1}\cdots D_{j+p}L^{m-1}\circ H_N$$

by (4.1)

$$=-m\sum(\xi_j-\xi_{j+p})D_jD_{j+1}\cdots D_{j+p}L^{m-1}\circ H_N$$

$$=m\sum\xi_j(D_{j-p}D_{j+1-p}\cdots D_j-D_jD_{j+1}\cdots D_{j+p})\times L^{m-1}\circ H_N,$$

by (4.17)

$$=m\sum_{j=1}^{N}(D_{j-p}\cdots D_{j-1}-D_{j+1}\cdots D_{j+p})\times L^{m-1}\circ H_N,$$

by periodicity

$$=0.$$

The systems (4.3) have a Hamiltonian structure

$$A\begin{pmatrix}\xi_{1,x}/\xi_1\\ \vdots\\ \xi_{N,x}/\xi_N\end{pmatrix}=B\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\nabla_{\hat{\xi}}H_1, \tag{4.19}$$

where $H_1=\Sigma_{j=1}^{N}\xi_j$.

The higher-order equations associated with the integrals (4.14) are

$$A\begin{pmatrix}\xi_{1,x}/\xi_1\\ \vdots\\ \xi_{N,x}/\xi_N\end{pmatrix}=B\begin{pmatrix}\xi_1 & & \\ & \ddots & \\ & & \xi_N\end{pmatrix}\nabla_{\hat{\xi}}H_{N-pm}. \tag{4.20}$$

When A is invertible, then

$$\Omega=A^{-1}B \tag{4.21}$$

is an antisymmetric circulant matrix. Note that

$$\Omega=(I+C+\cdots+C^p)^{-1}(I-C^p). \tag{4.22}$$

Then

$$\Omega_t=(I+C^t+\cdots+C^{tp})^{-1}(I-C^{tp})$$

and

$$CC^t=I$$

implies that

$$\Omega^t=C^pC^{-p}\Omega^t=-\Omega.$$

We have the systems

$$\hat{\xi}_{,x} = M_{\hat{\xi}} \nabla_{\hat{\xi}} H_1 \tag{4.23}$$

and

$$\hat{\xi}_{,x} = M_{\hat{\xi}} \nabla_{\hat{\xi}} H_{N-pm}, \tag{4.24}$$

where

$$M_{\hat{\xi}} = \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \Omega \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \tag{4.25}$$

is the cosymplectic form.

Furthermore, for any N, by (4.16)

$$\begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \nabla_{\hat{\xi}} H_{N-pm}$$

$$= -mA \begin{pmatrix} D_{N+1-p} \cdots D_1 \\ \vdots \\ D_{N-p} \cdots D_N \end{pmatrix} H_{N+p-pm}. \tag{4.26}$$

This demonstrates that the constraints for (4.19) are trivial for systems (4.20) and that systems (4.20) are equivalent to

$$\hat{\xi}_{,x} = -m \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} B \begin{pmatrix} D_{N+1-p} \cdots D_1 \\ \vdots \\ D_{N-p} \cdots D_N \end{pmatrix}$$

$$\times H_{N+p-pm}. \tag{4.27}$$

By (4.17)

$$\hat{\xi}_{,x} = -m \begin{pmatrix} D_{N+1-p} \cdots D_N - D_2 \cdots D_{p+1} \\ \vdots \\ D_{N-p} \cdots D_{N-1} - D_1 \cdots D_p \end{pmatrix}$$

$$\times H_{N+p-pm}, \tag{4.28}$$

or

$$\hat{\xi}_{,x} = -m \Omega_{\hat{D}} \nabla_{\hat{\xi}} H_{N+p-pm}, \tag{4.29}$$

where for $p \geqslant 2$,

$$\Omega_{\hat{D}} = \text{diff circ } [\overbrace{0,0,0,\ldots,0, -d}^{p+1}{}_{p+1}\,0,0,0,\ldots,\overbrace{d_1,0,0,0,\ldots,0}^{p}], \tag{4.30}$$

where $d_j = D_{j+1-p} \cdots D_{j-1}$ and the differential circulant (diff circ) matrix has the terms $-d_{p+1}$, $-d_{p+2}$, $-d_{p+3}$,... (mod N) along the diagonal beginning at the $p+1$ place, etc. We note that $\Omega^t = -\Omega$.

From (4.24) when A is invertible

$$M_{\hat{\xi}} \nabla_{\hat{\xi}} H_{N-pm} = -m \Omega_{\hat{D}} \nabla_{\hat{\xi}} H_{N+p-pm}. \tag{4.31}$$

When $p = 1$ Eqs. (4.31) are the dual-Hamiltonian formulation of the KdV systems. When $p \geqslant 2$ (4.31) is a differential form of the dual-Hamiltonian structure and as was true for the Boussinesq systems ($p = 2$), does not directly imply the involution of (4.14).

ACKNOWLEDGMENTS

We thank S. R. Sreenivasan and R. Chaterjee for several helpful conversations during the course of the authors' visits to the University of Calgary.

This work was supported by the Air Force Office of Scientific Research Grant No. AFOSR 86-0068 and a Killam Visiting Scholar grant from the University of Calgary, Canada.

APPENDIX: CANONICAL FORM

Following Lax[10] it is shown how to transform a system of the form

$$\hat{\xi}_{,x} = \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \Omega \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_N \end{pmatrix} \nabla_{\hat{\xi}} H, \tag{A1}$$

where $\Omega = A^{-1}B$ is an antisymmetric circulant matrix, to standard Hamiltonian form.

To be specific, let $H = \Sigma_{j=1}^{N} \xi_j$, $N = 2m + 1$ and

$$\xi_j = e^{\beta_j}. \tag{A2}$$

Then

$$\hat{\theta}_{,x} = \Omega \nabla_{\hat{\theta}} H, \tag{A3}$$

where

$$H = \sum_{j=1}^{N} e^{\theta_j}. \tag{A4}$$

Again, the eigenvectors of (A,B,Ω) are

$$\hat{\beta}_k = (1, r^k, \ldots, r^{k(N-1)})^t, \tag{A5}$$

where $r = \exp^{2\pi i/N}$. Let

$$\hat{\beta}_k = \hat{s}_k + i\hat{t}_k. \tag{A6}$$

Then

$$\Omega \hat{\beta}_k = \lambda_k \hat{\beta}_k,$$

with $\lambda_k = -i\sigma_k$

$$\Omega \hat{s}_k = \sigma_k \hat{t}_k, \quad \Omega \hat{t}_k = -\sigma_k \hat{s}_k. \tag{A7}$$

Note that

$$\hat{s}_k = (1, c_k, c_{2k}, \ldots, c_{k(N-1)})^t,$$
$$\hat{t}_k = (0, s_k, s_{2k}, \ldots, s_{k(N-1)})^t, \tag{A8}$$

where

$$c_k = \cos((2\pi/N)k), \quad s_k = \sin((2\pi/N)k). \tag{A9}$$

Let

$$\Phi = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 1 & c_1 & c_2 & \cdots & c_m & s_1 & s_2 & \cdots & s_m \\ \vdots & \vdots & & & \vdots & \vdots & & & \vdots \\ 1 & c_{N-1} & c_{2(N-1)} & \cdots & c_{m(N-1)} & s_{(N-1)} & s_{2(N-1)} & \cdots & s_{m(N-1)} \end{pmatrix}. \tag{A10}$$

Then, by (A7)

$$\Phi^t\Omega\Phi = N\begin{pmatrix} 0 & & & & & & \\ & -\sigma_1 & & & & & \\ & & \ddots & & & & \\ & & & -\sigma_m & & & \\ & & & & \sigma_1 & & \\ & & & & & \ddots & \\ & & & & & & \sigma_m \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \tag{A11}$$

Let

$$J_m = \begin{pmatrix} \sigma_1^{1/2} & & \\ & \ddots & \\ & & \sigma_m^{1/2} \end{pmatrix}, \tag{A12}$$

$$F = \frac{1}{N^{1/2}}\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & -J_m & \\ 0 & & J_m \end{pmatrix}\Phi^t, \tag{A13}$$

and

$$\begin{pmatrix} 0 \\ \hat{p} \\ \hat{q} \end{pmatrix} = F\hat{\theta}.$$

Then

$$\begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix}_{,x} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\nabla_{(\hat{p},\hat{q})}H(\hat{p},\hat{q}), \tag{A14}$$

where for $j = 1,2,\ldots,m$

$$p_j = \sigma_j^{1/2}\left(\sum_{k=0}^{2m} c_{jk}\theta_k\right), \quad q_j = \sigma_j^{1/2}\left(\sum_{k=1}^{2m} s_{jk}\theta_k\right), \tag{A15}$$

and for $j = 0,1,2,\ldots,2m$,

$$\theta_{j+1} = \sum_{k=1}^{m} \sigma_k^{-1/2}(s_{jk}q_k - c_{jk}p_k). \tag{A16}$$

[1]J. Weiss, J. Math. Phys. **25**, 13 (1984).
[2]E. Hille, *Ordinary Differential Equations in the Complex Plane* (Wiley, New York, 1976).
[3]J. Weiss, J. Math. Phys. **24**, 1405 (1983).
[4]J. Weiss, J. Math. Phys. **27**, 2647 (1986).
[5]P. Davis, *Circulant Matrices* (Wiley, New York, 1979).
[6]M. Kac and P. Van Moerbeke, Adv. Math. **15**, 160 (1975).
[7]M. Toda, *Theory of Nonlinear Lattices* (Springer, New York, 1981).
[8]J. Weiss, J. Math. Phys. **26**, 258 (1985).
[9]H. Yoshida, Celestial Mech. **81**, 363, 381 (1983).
[10]P. Lax, in *Nonlinear Evolution Equations*, edited by M. G. Crandall (Academic, New York, 1978), pp. 207.

Remarks on some hypergeometric orthogonal polynomials of mathematical physics

E. Montaldi
Dipartimento di Fisica, Universitá di Milano, Istituto Nazionale di Fisica Nucleare, Sezione di Milano, via Celoria 16, 20133 Milano, Italy

G. Zucchelli
Centro CNR, Dipartimento di Biologia, Universitá di Milano, via Celoria 26, 20133 Milano, Italy

(Received 25 September 1986; accepted for publication 29 April 1987)

The main result is a simple evaluation of an integral related to the orthogonality property of hypergeometric polynomials occurring in mathematical physics. Some related formulas for generalized hypergeometric functions are also briefly discussed.

I. INTRODUCTION

In recent years interesting investigations[1] appeared concerning various hypergeometric orthogonal polynomials including (or generalizing) the 6*j* symbols of angular momentum, the classical polynomials, and related polynomials with discrete orthogonalities.

In this context, the polynomials (we adopt here the same notation as in Wilson, 1980[1])

$$p_n(-x^2;a,b,c,d)=(a+b)_n(a+c)_n(a+d)_n\,{}_4F_3\left[\begin{matrix}-n,\ a+b+c+d+n-1,\ a+ix,\ a-ix;\\ a+b,\ a+c,\ a+d\end{matrix}\right] \tag{1.1}$$

are particularly important. They satisfy the symmetry property

$$p_n(-x^2;a,b,c,d)=p_n(-x^2;b,a,c,d) \tag{1.2}$$

(this is a consequence of a well-known transformation formula[2] for a balanced ${}_4F_3$) and the orthogonality relation

$$\int_{-\infty}^{\infty} dx\,\Gamma(x^2;a,b,c,d)p_m(-x^2;a,b,c,d)p_n(-x^2;a,b,c,d)=n!(a+b+c+d+n-1)_nK_n(a,b,c,d)\delta_{mn}, \tag{1.3}$$

where

$$\Gamma(x^2;a_1,a_2,\ldots,a_n)\equiv\frac{\Pi_{r=1}^{n}|\Gamma(a_r+ix)|^2}{|\Gamma(ix)\Gamma(\tfrac{1}{2}+ix)|^2} \tag{1.4}$$

and

$$K_n=\frac{\Gamma(a+b+n)\Gamma(a+c+n)\Gamma(a+d+n)\Gamma(b+c+n)\Gamma(b+d+n)\Gamma(c+d+n)}{\Gamma(a+b+c+d+2n)}. \tag{1.5}$$

For the sake of simplicity, we start with real positive values of *a*, *b*, *c*, and *d*. This restriction may be removed by also allowing for complex values of the parameters, by performing the appropriate analytic continuation at each step of the procedure described in Sec. II.

The proof of Eq. (1.3) is quite easy, involving a term by term integration and a rearrangement of the resulting double series, by means of Saalschütz's theorem.[3] The key of this procedure is the formula

$$\int_{-\infty}^{\infty} dx\,\Gamma(x^2;a,b,c,d)=\frac{\Gamma(a+b)\Gamma(a+c)\Gamma(a+d)\Gamma(b+c)\Gamma(b+d)\Gamma(c+d)}{\Gamma(a+b+c+d)}, \tag{1.6}$$

which is an integral analog of a ${}_5F_4$ summation theorem.[4] According to Wilson, the derivation of Eq. (1.6) by contour integration requires a tedious trigonometric computation. The main purpose of the present note is to give a simpler proof of (1.6), based on a formal trick familiar to a physicist. This is done in Sec. II. Some related results, pertaining to the hypergeometric series ${}_2F_1$, are given in Sec. III. In Sec. IV, examples borrowed from the theory of generalized hypergeometric series are also discussed.

II. PROOF OF EQ. (1.6)

We begin by quoting the formula

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,\frac{\Gamma(a+ix)\Gamma(a-ix)}{\Gamma(b+ix)\Gamma(b-ix)}e^{i\xi x}=\frac{\Gamma(2a)}{\Gamma(b-a)\Gamma(b+a)}\left(2\cosh\frac{\xi}{2}\right)^{-2a}\times{}_2F_1\left(a,a+\frac{1}{2};\ b+a;\operatorname{sech}^2\frac{\xi}{2}\right). \tag{2.1}$$

To establish this result, we write ($0<\operatorname{Re} a<\operatorname{Re} b$)

$$\frac{\Gamma(a+ix)}{\Gamma(b+ix)}=\frac{1}{\Gamma(b-a)}\int_0^1 dt\, t^{a+ix-1}(1-t)^{b-a-1},$$

$$\frac{\Gamma(a-ix)}{\Gamma(b-ix)}=\frac{1}{\Gamma(b-a)}\int_0^1 du\, u^{a-ix-1}(1-u)^{b-a-1},$$

so that the lhs of Eq. (2.1) becomes

$$\frac{1}{2\pi[\Gamma(b-a)]^2}\int_0^1\int_0^1 dt\,du(tu)^{a-1}\times[(1-t)(1-u)]^{b-a-1}\int_{-\infty}^{\infty} dx\Big(\frac{t}{u}e^{\xi}\Big)^{ix}.$$

The x integration gives $2\pi uw\delta(t-uw)$, $w=e^{-\xi}$. Thus we are left with a single integral and, by using Euler's integral representation of ${}_2F_1$ together with a quadratic transformation,[5] Eq. (2.1) follows at once.

In particular, we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,\frac{\Gamma(a+ix)\Gamma(a-ix)}{\Gamma(ix)\Gamma(-ix)}e^{i\xi x}=\frac{\Gamma(2a)}{\Gamma(-a)\Gamma(a)}2^{-2a}\cosh\frac{\xi}{2}\left|\sinh\frac{\xi}{2}\right|^{-(2a+1)},\tag{2.2}$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,\frac{\Gamma(a+ix)\Gamma(a-ix)}{\Gamma(\frac12+ix)\Gamma(\frac12-ix)}e^{i\xi x}=\frac{\Gamma(2a)}{\Gamma(\frac12-a)\Gamma(\frac12+a)}2^{-2a}\left|\sinh\frac{\xi}{2}\right|^{-2a}.\tag{2.3}$$

Let us now consider the integral

$$I=\frac{1}{2\pi}\int_{-\infty}^{\infty}dx\times\frac{\Gamma(a+ix)\Gamma(a-ix)\Gamma(b+ix)\Gamma(b-ix)}{\Gamma(ix)\Gamma(-ix)}e^{i\xi x}=\frac{\Gamma(2b)}{2\pi}\int_0^1 dt\,[t(1-t)]^{b-1}\int_{-\infty}^{\infty}dx\times\frac{\Gamma(a+ix)\Gamma(a-ix)}{\Gamma(ix)\Gamma(-ix)}e^{i\mu x}\tag{2.4}$$

where $\mu\equiv\xi+\ln(t/(1-t))$. By using Eq. (2.2), we get

$$I=\frac{\Gamma(2a)\Gamma(2b)}{\Gamma(-a)\Gamma(a)}(w^{-1/2}A)^{2a+1}(I_1+I_2),\tag{2.5}$$

with $w=e^{-\xi}$, $A=w/(w+1)$, and [$\operatorname{Re} a<0$, $\operatorname{Re}(a+b)>0$]

$$I_1=\int_0^A dt\,[t(1-t)]^{a+b-1}[tw^{-1/2}+(1-t)w^{1/2}]\times(A-t)^{-(2a+1)},\tag{2.6}$$

$$I_2=\int_A^1 dt\,[t(1-t)]^{a+b-1}[tw^{-1/2}+(1-t)w^{1/2}]\times(t-A)^{-(2a+1)}.\tag{2.7}$$

The splitting into I_1+I_2 is necessary, since $|\sinh(\mu/2)|=\cosh(\xi/2)/[t(1-t)]^{1/2}|t-A|$. By putting $t=Au$ in I_1, and $t=1-(1-A)u$ in I_2, these integrals are readily evaluated in terms of hypergeometric (or Legendre) functions. Precisely, since

$$P_\nu^\mu(x)=\frac{1}{\Gamma(1-\mu)}\Big(\frac{1+x}{1-x}\Big)^{\mu/2}\times{}_2F_1(-\nu,\nu+1;1-\mu;(1-x)/2)$$

we first get

$$I_{\frac12}=2b\Gamma(a+b)\Gamma(-2a)(w^{-1/2}A)^{-(2a+1)}\times\Big(\frac12\operatorname{sech}\frac{\xi}{2}\Big)^{a+b}P^{a-b}_{a+b}\Big(\pm\tanh\frac{\xi}{2}\Big),\tag{2.8}$$

whence, according to a standard formula[6]

$$\int_{-\infty}^{\infty}dx\left|\frac{\Gamma(a+ix)\Gamma(b+ix)}{\Gamma(ix)}\right|^2 e^{i\xi x}=\Gamma\Big(a+\frac12\Big)\Gamma\Big(b+\frac12\Big)\Gamma(a+b)\cosh\frac{\xi}{2}\times{}_2F_1\Big(a+\frac12,b+\frac12;\frac12;-\sinh^2\frac{\xi}{2}\Big).\tag{2.9}$$

Similarly, one finds

$$\int_{-\infty}^{\infty}dx\left|\frac{\Gamma(c+ix)\Gamma(d+ix)}{\Gamma(\frac12+ix)}\right|^2 e^{i\xi x}=\Gamma(c)\Gamma(d)\Gamma(c+d)\,{}_2F_1\Big(c,d;\frac12;-\sinh^2\frac{\xi}{2}\Big).\tag{2.10}$$

We are now going to derive (1.6) from (2.9) and (2.10), by means of Parseval's formula[7]

$$\int_{-\infty}^{\infty}dxF(x)G(x)=\int_{-\infty}^{\infty}d\xi f(\xi)g(-\xi),$$

where $F(x)$ and $G(x)$ are the Fourier transforms of $f(\xi)$ and $g(\xi)$. Then, with

$$F(x)=\left|\frac{\Gamma(a+ix)\Gamma(b+ix)}{\Gamma(ix)}\right|^2,\quad G(x)=\left|\frac{\Gamma(c+ix)\Gamma(d+ix)}{\Gamma(\frac12+ix)}\right|^2,$$

and $\sinh^2(\xi/2)=t$, we are led to consider the integral

$$\int_0^\infty dt\,t^{-1/2}\,{}_2F_1\Big(a+\frac12,b+\frac12;\frac12;-t\Big)\times{}_2F_1\Big(c,d;\frac12;-t\Big)=\pi\frac{\Gamma(a+c)\Gamma(a+d)\Gamma(b+c)\Gamma(b+d)}{\Gamma(a+\frac12)\Gamma(b+\frac12)\Gamma(c)\Gamma(d)\Gamma(a+b+c+d)},$$
$$\operatorname{Re}(a+c),\operatorname{Re}(a+d),\operatorname{Re}(b+c),\operatorname{Re}(b+d)>0.\tag{2.11}$$

This formula is easily established as the particular case $u=1$ of the more general result

$$\int_0^\infty dt\,t^{-1/2}\,{}_2F_1\Big(a+\frac12,b+\frac12;\frac12;-ut\Big)\times{}_2F_1\Big(c,d;\frac12;-t\Big)=\pi\frac{\Gamma(a+c)\Gamma(a+d)\Gamma(b+c)\Gamma(b+d)}{\Gamma(a+\frac12)\Gamma(b+\frac12)\Gamma(c)\Gamma(d)\Gamma(a+b+c+d)}\times u^{-(a+1/2)}\,{}_2F_1(a+c,a+d;a+b+c+d;1-1/u)\tag{2.12}$$

whose validity can be checked by evaluating the Mellin transform (with respect to u) of both sides. This concludes our proof.

In the limiting case $d\to\infty$, Eq. (1.3) takes the form

$$\int_{-\infty}^{\infty} dx\,\Gamma(x^2;a,b,c)\bar{p}_m(-x^2;a,b,c)\bar{p}_n(-x^2;a,b,c) = n!\Gamma(a+b+n)\Gamma(a+c+n)\Gamma(b+c+n)\delta_{mn}, \tag{2.13}$$

where

$$\bar{p}_n(-x^2;a,b,c) = (a+b)_n(a+c)_n \times {}_3F_2\left[\begin{matrix} -n,\ a+ix,\ a-ix; \\ a+b,\ a+c \end{matrix}\right]. \tag{2.14}$$

It is instructive to give an alternative derivation of the orthogonality relation (2.13), based on some elementary properties of the coefficients $g_{s,l}$ defined by

$$x^{2s} = \sum_{l=0}^{s} g_{s,l}(a)(a+ix)_l(a-ix)_l. \tag{2.15}$$

Let us write

$$J_{r,s}(a,b,c) = \int_{-\infty}^{\infty} dx\,\Gamma(x^2;a,b,c)(a+ix)_r(a-ix)_r x^{2s}. \tag{2.16}$$

Then, we have [from (1.6), by letting $d\to\infty$],

$$J_{0,0}(a,b,c) = \Gamma(a+b)\Gamma(a+c)\Gamma(b+c), \tag{2.17}$$

$$J_{r,0}(a+s,b,c) = (a+b)_{r+s}(a+c)_{r+s}J_{0,0}(a,b,c), \tag{2.18}$$

and

$$J_{r,s}(a,b,c) = J_{0,0}(a,b,c)\sum_{l=0}^{s} g_{s,l} \times (a+r)(a+b)_{r+l}(a+c)_{r+l} \tag{2.19}$$

so that

$$\int_{-\infty}^{\infty} dx\,\Gamma(x^2;a,b,c)\,{}_3F_2\left[\begin{matrix} -n,\ a+ix,\ a-ix; \\ a+b,\ a+c \end{matrix}\right]x^{2s} = J_{0,0}(a,b,c)\sum_{r=0}^{n}(-1)^r\binom{n}{r}\sum_{l=0}^{s} g_{s,l} \times (a+r)(a+r+b)_l(a+r+c)_l. \tag{2.20}$$

Now, by mathematical induction, one easily proves that $\Sigma_{l=0}^{s} g_{s,l}(a+r)(a+r+b)_l(a+r+c)_l$ is a polynomial of degree s in r, the coefficient of r^s being $(b+c)_s$. By noting that

$$\sum_{r=0}^{n}(-1)^r\binom{n}{r}r^j = \left(x\frac{d}{dx}\right)^j(1-x)^n\Big|_{x=1} = \begin{cases} 0, & j<n, \\ (-1)^n n!, & j=n, \end{cases} \tag{2.21}$$

Eq. (2.13) follows at once.

As a further application of the method used in deriving Eq. (1.6), we consider the integral

$$J = \int_{-\infty}^{\infty} dx\,\Gamma(a+ix)\Gamma(b-ix)\Gamma(c+ix)\Gamma(d-ix) \times u_n(x;a,b,c,d)e^{i\xi x}, \tag{2.22}$$

where

$$u_n(x;a,b,c,d) = {}_3F_2\left[\begin{matrix} -n,\ a+b+c+d+n-1,\ a+ix; \\ a+b,\ a+d \end{matrix}\right]. \tag{2.23}$$

These polynomials, which have been discussed by Atakishiyev and Suslov,[1] are connected to Jacobi polynomials by Euler's transformation

$$B(a+ix,d-ix)u_n(x;a,b,c,d) = \int_0^1 dt\,t^{a+ix-1}(1-t)^{d-ix+1} \times {}_2F_1(-n,a+b+c+d+n-1;a+b;t). \tag{2.24}$$

By inserting (2.24) in (2.22), and by using the formula[8]

$$\int_{-\infty}^{\infty} dx\,\Gamma(b-ix)\Gamma(c+ix)e^{i\mu x} = 2\pi\Gamma(b+c)\left(\frac{1}{2}\operatorname{sech}\frac{\mu}{2}\right)^{b+c}e^{(b-c)\mu/2}, \tag{2.25}$$

we easily obtain

$$J = 2\pi\Gamma(a+d)\Gamma(b+c)e^{[(b-c)/2]\xi}2^{1-(a+b+c+d)} \times\frac{n!}{(a+b)_n}\int_{-1}^{1}dt(1-t)^{a+b-1}(1+t)^{c+d-1} \times\left(\cosh\frac{\xi}{2}-t\sinh\frac{\xi}{2}\right)^{-(b+c)} \times P_n^{(a+b-1,c+d-1)}(t). \tag{2.26}$$

By writing

$$\left(\cosh\frac{\xi}{2}-t\sinh\frac{\xi}{2}\right)^{-(b+c)} = \left(\cosh\frac{\xi}{2}\right)^{-(b+c)}\sum_{r=0}^{\infty}\frac{(b+c)_r}{r!}\left(t\tanh\frac{\xi}{2}\right)^r, \tag{2.27}$$

a straightforward calculation shows that

$$\int_{-\infty}^{\infty} dx\,\Gamma(a+ix)\Gamma(b-ix)\Gamma(c+ix)\Gamma(d-ix) \times u_m(x;a,b,c,d)u_n(x;a,b,c,d) = (-1)^n 2\pi n!\Gamma(a+b)\Gamma(a+d)\Gamma(b+c) \times\Gamma(c+d)\frac{(b+c)_n(c+d)_n}{(a+b)_n(a+d)_n} \times\frac{(a+b+c+d+n-1)_n}{\Gamma(a+b+c+d+2n)}\delta_{mn}. \tag{2.28}$$

Thus, the orthogonality relation for the polynomials u_n is a simple consequence of the orthogonality of Jacobi's polynomials[9]; this is, in fact, the only tool required in the above sketched proof.

We end this section by quoting the formula

$$\pi\int_{-\infty}^{\infty} dx\left|\left(\frac{1}{2}+ix\right)_r\left(\frac{1}{2}+ix\right)_s\right|^2 x\sinh(\pi x)\operatorname{sech}^2(\pi x)$$
$$= r!s!(r+s)!, \qquad (2.29)$$

which follows from Eq. (2.9) with $\xi=0$ and $a=r+\frac{1}{2}$, $b=s+\frac{1}{2}$ $(r,s=0,1,2,...)$. This result may be helpful in evaluating Mehler–Fock transforms; for instance, it is almost immediate to obtain Mehler's generalization of Heine's formula,[10]

$$(y-x)^{-1}=\pi\int_0^{\infty} dt\, P_{-1/2+it}(y)P_{-1/2+it}(-x)$$
$$\times t\sinh(\pi t)\operatorname{sech}^2(\pi t). \qquad (2.30)$$

III. SOME FORMULAS FOR $_2F_1$

An appropriate use of Fourier transforms like those considered in Sec. II leads to new integral representations of special functions. As an interesting example, we start here from Clausen's formula[11]

$$[{}_2F_1(a,b;a+b+\tfrac{1}{2};x)]^2={}_3F_2\left[\begin{matrix}2a,\ a+b,\ 2b;\ x\\ a+b+\frac{1}{2},\ 2a+2b\end{matrix}\right], \qquad (3.1)$$

which, with $a\to a+it$ and $b\to b-it$, can be written as

$$\Gamma(a+b+\tfrac{1}{2})\int_0^1\int_0^1 du\,dv\,u^{b-1}v^{a-1}(1-u)^{a-1/2}(1-v)^{b-1/2}(1-xu)^{-a}(1-xv)^{-b}\left[\frac{1-u}{u(1-xu)}\cdot\frac{v(1-xv)}{1-v}\right]^{it}$$
$$=\pi\cdot 4^{1-a-b}\sum_{n=0}^{\infty}\frac{\Gamma(2a+2it+n)\Gamma(2b-2it+n)(a+b)_n}{\Gamma(a+b+\frac{1}{2}n)(2a+2b)_n n!}x^n. \qquad (3.2)$$

By taking the Fourier transform of both sides, we get [recall Eq. (2.25)]

$$\int_0^1\int_0^1 du\,dv\,u^{-1}(1-u)^{a-1/2}(1-xu)^{-a}v^{a-1}(1-v)^{-1/2}\delta\left(\frac{1-u}{u(1-xu)}\cdot\frac{v(1-xv)}{1-v}\tau^2-1\right)$$
$$=B(a,\tfrac{1}{2})(1+\tau)^{-2a}\,{}_2F_1(a,a;2a;y), \qquad (3.3)$$

where $\tau=e^{\xi/2}$, $y=4x\tau/(1+\tau)^2$, and $a+b$ has been replaced by a. It is easy to check Eq. (3.3) in the special cases $x=0$ and $x=1$. For $\xi=0$, Eq. (3.3) reduces to

$$\int_0^1 dv\,v^{a-1}(1-v)^a(1-xv)^{1-a}(1-2xv+xv^2)^{-1}$$
$$=2^{-2a}B(a,\tfrac{1}{2})\,{}_2F_1(a,a;2a;x). \qquad (3.4)$$

A direct proof of (3.4) runs as follows. By partial integration, we have

$$\int_0^1 dx(1-x)^{2a-1}[(a+n)(1-x^2)^n$$
$$-n(1-x)(1-x^2)^{n-1}]=\tfrac{1}{2}, \qquad (3.5)$$

whence

$$\int_0^1 dx(1-x)^{2a-1}[1-t(1-x)][1-t(1-x^2)]^{-(a+1)}$$
$$=(1/2a)\,(1-t)^{-a}, \qquad (3.6)$$

which, in turn, implies

$$\int_0^1 dx(1-x)^{2a+n-1}\,{}_2F_1(-n,a+1;a;-x)$$
$$=\frac{1}{2a+n}\,{}_3F_2\left[\begin{matrix}-n,\ a+1,\ 1;\ -1\\ a,\ 2a+n+1\end{matrix}\right]=\frac{1}{2a}, \qquad (3.7)$$

because[12]

$$[1-t(1-x)][1-t(1-x^2)]^{-(a+1)}$$
$$=\sum_{n=0}^{\infty}\frac{(a)_n}{n!}[t(1-x)]^n\,{}_2F_1(-n,a+1;a;-x). \qquad (3.8)$$

Next, we rewrite (3.7) as

$$\int_0^1 dv\,v^{n+a-1}(1-v)^a\,{}_2F_1(-n,1;a;v-1)$$
$$=2^{-2a}B\left(a,\frac{1}{2}\right)\frac{(a)_n}{(2a)_n} \qquad (3.9)$$

and the desired result follows immediately.

In the general case, we have

$$\delta\left(\frac{1-u}{u(1-xu)}\cdot\frac{v(1-xv)}{1-v}\tau^2-1\right)$$
$$=\frac{\bar{u}^2(1-x\bar{u})^2}{1-2x\bar{u}+x\bar{u}^2}\cdot\frac{1-v}{v(1-xv)}\tau^{-2}\delta(u-\bar{u}), \qquad (3.10)$$

where $\bar{u}$ is the root of the quadratic equation

$$xu^2-\left(1+\frac{v(1-xv)}{1-v}\tau^2\right)u+\frac{v(1-xv)}{1-v}\tau^2=0, \qquad (3.11)$$

which, for $x=0$, reduces to $u_0=\tau^2 v/[1-(1-\tau^2)v]$. We can expand $\bar{u}$ as

$$\bar{u}=u_0+\sum_{r=1}^{\infty}u_r x^r, \qquad (3.12)$$

where $[A=1-(1-\tau^2)v]$

$$u_r=-\frac{1-v}{A^{r+1}}(v\tau)^{2r}+(r-1)!v^{r+1}\tau^2$$
$$\times\sum_{n+j=r-1}\frac{(\frac{1}{2})_{n+1}}{(n+2)!}\cdot\frac{[4(1-v)\tau^2]^{n+1}}{A^{2n+3}}\frac{1}{n!j!}$$
$$\times{}_2F_1(-j,2n+3;n+1;(1-v)/A). \qquad (3.13)$$

Equation (3.3) takes now the form

$$\int_0^1 dv\,\bar{u}(1-\bar{u})^{a-1/2}(1-x\bar{u})^{-(a-2)}v^{a-2}(1-v)^{1/2}$$
$$\times(1-xv)^{-1}(1-2x\bar{u}+x\bar{u}^2)^{-1}$$
$$=B(a,\tfrac{1}{2})[\tau^2/(1+\tau)^{2a}]\,{}_2F_1(a,a;2a;y). \quad (3.14)$$

If we expand both sides of (3.14) in power series of x, we get a sequence of identities involving ${}_2F_1$; for instance, at order 1, one has

$$(a+(1-\tau^2)/2)B(a+2,a+2)$$
$$\times{}_2F_1(a+\tfrac{5}{2},a+2;2a+4;1-\tau^2)+(a+1)$$
$$\times B(a+1,a+2)\,{}_2F_1(a+\tfrac{3}{2},a+1;2a+3;1-\tau^2)$$
$$+a\tau^2B(a+1,a+3)\,{}_2F_1(a+\tfrac{5}{2},a+3;2a+4;1-\tau^2)$$
$$=2aB(a,\tfrac{1}{2})\tau^{-1}(1+\tau)^{-2(a+1)}. \quad (3.15)$$

IV. SOME REMARKS ON ${}_3F_2$

In this section, we give two further examples illustrating the usefulness of Fourier transforms in the theory of generalized hypergeometric series. We first observe that ($w=e^{-\xi}$)

$$\int_{-\infty}^{\infty} dx\,\Gamma(\alpha+ix)\Gamma(\beta-ix)$$
$$\times{}_3F_2\left[\begin{matrix}\alpha+ix,\ \beta-ix,\ (\alpha+\beta-1)/2;\\ \alpha+\beta,\ (\alpha+\beta+1)/2\end{matrix}\right]e^{i\xi x}$$
$$=2\pi[\Gamma(\alpha+\beta)/(1+w^{-1})]$$
$$\times[w^{\alpha-1}\theta(\xi)+w^{-\beta}\theta(-\xi)]. \quad (4.1)$$

This result is a simple consequence of Eq. (2.25), combined with the formula[13]

$${}_2F_1(a-\tfrac{1}{2},a;2a;x)=((1+\sqrt{1-x})/2)^{1-2a}. \quad (4.2)$$

On the other hand, a standard integral representation[14] for the logarithmic derivative of the gamma function gives

$$\psi(\lambda+ix)-\psi(\mu+ix)=\int_0^1 dt\,t^{ix}\frac{t^{\mu-1}-t^{\lambda-1}}{1-t}, \quad (4.3)$$

whence one easily obtains

$$\int_{-\infty}^{\infty} dx[\psi(\lambda+ix)-\psi(\mu+ix)]e^{i\xi x}$$
$$=2\pi\frac{w^\mu-w^\lambda}{1-w}\theta(\xi). \quad (4.4)$$

By using (4.4), it is immediately seen that the rhs of (4.1) is the Fourier transform of

$$\frac{1}{2}\Gamma(\alpha+\beta)\left[\psi\frac{\alpha+1+ix}{2}\right]+\psi\left(\frac{\beta+1-ix}{2}\right)$$
$$-\psi\left(\frac{\alpha+ix}{2}\right)-\psi\left(\frac{\beta-ix}{2}\right);$$

hence we get Watson's formula[15]

$${}_3F_2\left[\begin{matrix}\alpha,\beta,(\alpha+\beta-1)/2;\\ \alpha+\beta,(\alpha+\beta+1)/2\end{matrix}\right]$$
$$=\frac{1}{2B(\alpha,\beta)}\left[\psi\left(\frac{\alpha+1}{2}\right)+\psi\left(\frac{\beta+1}{2}\right)\right.$$
$$\left.-\psi\left(\frac{\alpha}{2}\right)-\psi\left(\frac{\beta}{2}\right)\right]. \quad (4.5)$$

Next, let us consider Dixon's theorem

$${}_3F_2\left[\begin{matrix}a,\ b,\ c;\\ 1+a-b,\ 1+a-c\end{matrix}\right]=\frac{\Gamma(1+a/2)\Gamma(1+a-b)\Gamma(1+a-c)\Gamma(1+a/2-b-c)}{\Gamma(1+a)\Gamma(1+a/2-b)\Gamma(1+a/2-c)\Gamma(1+a-b-c)} \quad (4.6)$$

summing any convergent well-poised ${}_3F_2$ with unit argument. After the replacements $b\to b+ix$ and $c\to c-ix$, we take the Fourier transform of both sides. By using the formula ($w=e^{-\xi}$)

$$\int_{-\infty}^{\infty} dx\frac{\Gamma(\alpha+ix)\Gamma(\beta-ix)}{\Gamma(\lambda+ix)\Gamma(\mu-ix)}e^{i\xi x}$$
$$=2\pi\frac{\Gamma(\alpha+\beta)}{\Gamma(\lambda-\alpha)\Gamma(\alpha+\mu)}w^\alpha$$
$$\times{}_2F_1(\alpha-\lambda+1,\alpha+\beta;\alpha+\mu;w), \quad (4.7)$$

which generalizes Eq. (2.1), a simple calculation shows (by writing simply b instead of $b+c$) that

$$\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(1+a)_{2n}n!}w^n\,{}_2F_1(b-a,b+2n;1+a+2n;w)$$
$$={}_2F_1(b-a/2,b;1+a/2;w). \quad (4.8)$$

If we expand ${}_2F_1(b-a,b+2n;1+a+2n;w)$ and rearrange the double series, the lhs becomes a single power series in w. Then, by comparing the coefficients of w^r in both sides, we have (with the replacement $b\to b-n$)

$${}_3F_2\left[\begin{matrix}-n,a,b;\\ 1+a+n,1+a-b\end{matrix}\right]=\frac{(1+a)_n(1+a/2-b)_n}{(1+a/2)_n(1+a-b)_n} \quad (4.9)$$

and no new result is obtained, because (4.9) is nothing but (4.6) with $c=-n$. However, if we perform the above sketched transformations in the reverse order, a rather unexpected way to derive the general Dixon's theorem from the terminating one emerges. Now, the proof of Eq. (4.9) is very simple, involving the use of a suitable generating function. To this aim, we start from the formula

$${}_3F_2\left[\begin{matrix}-m,a,1+a/2;\\ a/2,a+m+1\end{matrix}\right]$$
$$=\frac{1}{B(a,m+1)}\int_0^1 dx\,x^{a-1}(1-x)^m$$
$$\times{}_2F_1\left(-m,1+\frac{a}{2};\frac{a}{2};x\right)$$
$$=\frac{1}{B(a,m+1)}\int_0^1 dx\,x^{a-1}(1-x)^{2m}$$
$$\times{}_2F_1\left(-m,-1;\frac{a}{2};-\frac{x}{1-x}\right)=\delta_{m0}, \quad (4.10)$$

which can be rewritten as

$$B(b,a-b+1)\delta_{m0} = \frac{(a/2)_m}{m!}\int_0^1 du\, u^{b+m-1}(1-u)^{a-b} \times {}_3F_2\left[\begin{matrix} -m,a,1+a/2;1-u \\ 1+a-b,a/2 \end{matrix}\right]. \tag{4.11}$$

By recalling that

$$\sum_{m=0}^{\infty} \frac{t^m}{m!}(\lambda)_m\, {}_3F_2\left[\begin{matrix} -m,\alpha,\beta;x \\ \lambda,\mu \end{matrix}\right] = (1-t)^{-\lambda}\, {}_2F_1\left(\alpha,\beta;\mu; -\frac{xt}{1-t}\right) \tag{4.12}$$

(this follows, after expansion of ${}_3F_2$, by rearrangement of the double series), Eq. (4.11) implies

$$\int_0^1 du\, u^{b-1}(1-u)^{a-b}(1-su)^{-a/2} \times {}_2F_1\left(1+\frac{a}{2},a;1+a-b; -\frac{su(1-u)}{1-su}\right) = B(b,1+a-b) \tag{4.13}$$

or, by setting $u=(1-x)/(1-sx)$

$$\int_0^1 dx\, x^{a-b}(1-x)^{b-1}(1-sx)^{-(1+a/2)} \times {}_2F_1\left(1+\frac{a}{2},a;1+a-b; -\frac{sx(1-x)}{1-sx}\right) = B(b,1+a-b)(1-s)^{-(1+a/2-b)}. \tag{4.14}$$

Now, as a particular case of Eq. (4.12), we have

$$\sum_{n=0}^{\infty} \frac{t^n}{n!}(a)_n\, {}_2F_1(-n,b;c;x) = (1-t)^{-a}\, {}_2F_1\left(a,b;c; -\frac{xt}{1-t}\right). \tag{4.15}$$

Thus Eq. (4.14) becomes

$$\int_0^1 dx\, x^{a-b+n}(1-x)^{b-1}\, {}_2F_1(-n,a;1+a-b;1-x) = B(b,1+a-b)\,\frac{(1+a/2-b)_n}{(1+a/2)_n} \tag{4.16}$$

and this is equivalent to the desired result (4.9).

In conclusion, we recall for the reader's convenience the formulas and the derivations that, as far as we know, are new and possibly useful. The key formulas of Sec. II, namely the Fourier transforms (2.9) and (2.10), together with the tools used to derive them, seem to be new. The same statement holds for Eq. (2.12), and for the alternative derivation of the orthogonality relation (2.13) based on the properties of the coefficients $g_{s,l}$ defined by (2.15). Likewise, Eqs. (3.4), (3.14), (4.8), and (4.14) do not seem to be known in the mathematical literature.

[1]R. Askey and J. A. Wilson, SIAM J. Math. Anal. **10**, 1008 (1979); **13**, 651 (1982); J. A. Wilson, SIAM J. Math. Anal. **11**, 690 (1980); N. M. Atakishiyev and S. K. Suslov, J. Phys. A: Math. Gen. **18**, 1583 (1985); and references therein.

[2]W. N. Bailey, *Generalized Hypergeometric Series* (Cambridge U. P., Cambridge, 1935), p. 56. As remarked by Wilson, this formula, when iterated, contains the symmetries of the 6*j* symbols.

[3]*Bateman Manuscript Project: Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1, p. 66.

[4]See Ref. 2, pp. 27 and 47.

[5]See Ref. 3, pp. 59 and 110.

[6]See Ref. 3, p. 126, formula (22).

[7]E. C. Titchmarsh, *Theory of Fourier Integrals* (Oxford U. P., New York, 1948), p. 50.

[8]By Fourier inversion, Eq. (2.25) is nothing but a standard integral representation of Euler's beta function.

[9]*Bateman Manuscript Project: Tables of Integral Transforms* (McGraw-Hill, New York, 1953), Vol. 2, p. 285.

[10]See Ref. 9, p. 174.

[11]See Ref. 3, p.185, formula (1); note (see Errata) that in the rhs one should read $2a+2b$ instead of $a+2b$. This formula can be obtained, in a quite simple manner, from the corresponding formula for q-hypergeometric functions, by letting $q\to 1$; see, for instance, F. H. Jackson, Q. J. Math. **11**, 1 (1940).

[12]See Ref. 3, p. 82.

[13]See Ref. 3, p. 101.

[14]See Ref. 3, p. 16.

[15]See Ref. 2, p. 98.

[16]See Ref. 2, p. 13.

On some classes of unbounded commutants of unbounded operator families

J. Shabani[a)]

International Centre for Theoretical Physics, Trieste, Italy and Institut de Physique Théorique, Université Catholique de Louvain, Louvain la Neuve, Belgium

(Received 15 October 1985; accepted for publication 8 April 1987)

Some classes of unbounded commutants and bicommutants and their behavior with respect to the quasiweak*-topology which seems to play here the role of the weak topology for bounded operators, are investigated. In particular, some sufficient conditions are given in order that the bicommutants be the quasiweak*-closure of the original set of operators.

I. INTRODUCTION

This paper continues the study of unbounded commutants and bicommutants of unbounded operator families, started in Refs. 1 and 2. In Refs. 1 and 2 we discussed, among other things, all possible definitions of unbounded commutants and we studied their behavior with respect to the weak topology which seems to be quite natural for the commutant $\mathscr{R}'_0$ (see Sec. II for definition).

As pointed out in Refs. 1–3 the weak topology is not natural for the weak unbounded commutant $\mathscr{R}'_\sigma$ or the strong unbounded commutant $\mathscr{R}'_c$, for in general these commutants need not be weakly closed.

In this paper, we will be concerned with the commutants $\mathscr{R}'_\sigma$ and $\mathscr{R}'_c$, as well as the bicommutants $\mathscr{R}''_{0c}$ and $\mathscr{R}''_{\sigma\sigma}$. In particular, we will investigate their behavior with respect to the quasiweak*-topology which plays here the role of the weak topology for bounded operators.

In Sec. II, after recalling some basic facts on partial inner product spaces,[4–8] we define the different topologies, as well as the different commutants and bicommutants we will discuss in this paper.

In Sec. III we study topological properties of our commutants and bicommutants and in Sec. IV we give some sufficient conditions in order that the bicommutants $\mathscr{R}''_{0c}$ and $\mathscr{R}''_{\sigma\sigma}$ be the closure of the original set of operators with respect to the quasiweak*-topology.

II. OPERATORS IN A PARTIAL INNER PRODUCT (PIP) SPACE

A. Definitions and basic properties

A PIP space[4–8] 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 requirements.

(i) The family $\mathscr{I} = \{V_r,\ r\in I\}$ covers V and is an involutive lattice with respect to set intersection, vector sum, and involution $\#$: $V_r \leftrightarrow V_{\bar r}$.

The lattice structure may be transferred to the index set I by writing: $V_{p\wedge q} \equiv V_p \cap V_q, V_{p\vee q} \equiv V_p + V_q, V_{\bar r} = (V_r)^\#$. Besides elements of $\mathscr{I}$, we consider also the extreme spaces $V^\# \equiv \cap_{r\in I} V_r$ and $V \equiv \cup_{r\in I} V_r$.

(ii) The Hermitian form $\langle\cdot|\cdot\rangle$, called the partial inner product, is defined on $\cup_{r\in I} V_r \times V_{\bar r}$.

Moreover, we assume that V possesses a central Hilbert space, i.e., there exists an element $0 = \bar 0$ in I such that $V_0 = V_{\bar 0} \equiv \mathfrak{H}$ is a Hilbert space with respect to $\langle\cdot|\cdot\rangle$.

It follows from the assumption of nondegeneracy $(V^\#)^\perp = \{0\}$ that every pair $\langle V_r, V_{\bar r}\rangle$ as well as $\langle V^\#, V\rangle$ is a dual pair with respect to $\langle\cdot|\cdot\rangle$. Consequently 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. If for every $r\in I$, V_r is a Hilbert space, then the PIP space V is called a nested Hilbert space.[9]

An operator[5] 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. The domain $\mathscr{D}(A)$ is a dense vector subspace of V containing $V^\#$, and A is uniquely determined by its restriction to $V^\#$. Such operators may be extremely singular, since the range of $A|_{V^\#}$ may be much larger than the central Hilbert space $\mathfrak{H}$. 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-x-algebra[10,11]); two operators A and B may always be added (as sesquilinear forms over $V^\#$), but their product AB is defined only if there is a continuous factorization through some $V_q \in \mathscr{I}$;

$$V^\# \xrightarrow{B} V_q \xrightarrow{A} V.$$

An operator $A\in$Op V is called regular[12] if $D(A) = D(A^\times) = V$; 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,[13] i.e., a *-subalgebra with unit of the algebra $L^+(V^\#)$ of all closable operators on $\mathfrak{H}$, which together with their (Hilbertian) adjoint have $V^\#$ for invariant domain. The space Op V contains another remarkable subset, namely,

$$C(V^\#, \mathfrak{H}) = \{A \text{ closable in } \mathfrak{H} | V^\# \subset D(A) \cap D(A^*)\}$$
$$= \{A\in \text{Op } V | A^{(*)}: V^\# \to \mathfrak{H}\}.$$

We have Reg $V \subseteq C(V^\#, \mathfrak{H}) \subseteq$ Op V.

From now on we will assume that V is quasicomplete in

[a)] On leave of absence from the Department of Mathematics, University of Burundi, BP 2700 Bujumbura, Burundi, Central Africa.

its Mackey topology. This implies in particular that we can simply identify Reg V with $L^+(V^\#)$; see Ref. 12, Proposition 2.5. The condition of Mackey quasicompleteness of V is actually satisfied in almost all examples; the only known exceptions are quite pathological.[14,15] It is of course automatic if $\langle V^\#,V\rangle$ is a reflexive dual pair in the sense of Köthe,[16] i.e., if the dual of $V^\#$ in the strong topology coincides with V and vice versa.

B. Topologies

Several topologies may be defined on $C(V^\#,\mathfrak{H})$. Here we will consider the following ones.

(a) The quasiweak-(qw-) topology[17] defined by the following family of seminorms:

$$A\in C(V^\#,\mathfrak{H})\mapsto|\langle h,Af\rangle|, \quad f\in V^\#, \quad h\in\mathfrak{H}.$$

(b) The quasiweak*-(qw*-) topology defined by the seminorms:

$$A\in C(V^\#,\mathfrak{H})\mapsto\max\{|\langle h,Af\rangle|,|\langle h,A^*f\rangle|\}$$

or

$$A\mapsto|\langle h,Af\rangle|+|\langle h,A^*f\rangle|, \quad f\in V^\#, \quad h\in\mathfrak{H}.$$

(c) The strong*-(s*-) topology[18] defined by the seminorms:

$$A\in C(V^\#,\mathfrak{H})\mapsto\max\{\|Af\|,\|A^*f\|\}$$

or

$$A\mapsto\|Af\|+\|A^*f\|, \quad f\in V^\#.$$

The qw*-topology is coarser than the s*-topology, finer than the qw-topology but not comparable to the strong topology. Actually the qw*-topology generalizes the so-called commutant topology introduced in Ref. 19.

On Reg $V\simeq L^+(V^\#)$ we will consider the qw-, qw*-, and s*-topologies inherited from $C(V^\#,\mathfrak{H})$.

C. Commutants

Since we are dealing with unbounded operators, several concepts of commutants may be introduced. Let $A,B\in$Op V. Following Ref. 11 we say that A is the left multiplier of B (resp. B is the right multiplier of A), and we note $A\in L(B)$ [resp. $B\in R(A)$], if the product AB is defined. Similarly, if $\mathscr{R}$ is a subset of Op V, we may define

$$L\mathscr{R}\equiv\bigcap_{B\in\mathscr{R}}L(B)=\{C\in\text{Op } V\,|\,CB \text{ exists } \forall B\in\mathscr{R}\}.$$

Definition 2.1: A subset $\mathscr{R}$ of Op V is called a partial-x-algebra, if the following conditions are satisfied: (i) $\mathscr{R}$ contains the identity; (ii) $\mathscr{R}$ is x invariant, i.e., $A\in\mathscr{R}$ implies $A^x\in\mathscr{R}$; (iii) if $A,B\in\mathscr{R}$ and $A\in L(B)$, then $AB\in\mathscr{R}$. It is easily verified that Op V is a partial-x-algebra.

Let us now recall the different notions of commutant we need in the following.

(a) If $\mathscr{R}$ is an x-invariant subset of Op V, we may define the following commutants.

(i) $\mathscr{R}'=\{X\in\text{Op } V\,|\,X\in L\mathscr{R}\cap R\mathscr{R},\ XA=AX,\ \forall A\in\mathscr{R}\}$. As already pointed out in Ref. 1, $\mathscr{R}'$ is a vector subspace of $L\mathscr{R}\cap R\mathscr{R}$. Furthermore, it is x-invariant and contains the identity, but in general it is not a partial-x-algebra.

(ii) $$\mathscr{R}'_c\equiv\mathscr{R}'\cap L^+(V^\#)=\{X\in L^+(V^\#)\,|\,AX=XA,\ \forall A\in\mathscr{R}\}.$$

Here $\mathscr{R}'_c$ will be called the strong unbounded commutant. It is an Op*-algebra on $V^\#$.

(b) Let now $\mathscr{R}$ be an x-invariant subset of $C(V^\#,\mathfrak{H})$. Then we may define the following commutants.

(i) $\mathscr{R}'_\sigma=\{X\in C(V^\#,\mathfrak{H})\,|\,\langle X^*f,Ag\rangle=\langle A^*f,Xg\rangle;\forall f,g\in V^\#,A\in\mathscr{R}\}$. Here $\mathscr{R}'_\sigma$ is called the weak unbounded commutant.[3] It is an *-invariant linear subset of $C(V^\#,\mathfrak{H})$. Its bounded part $\mathscr{R}'_w\equiv\mathscr{R}'_\sigma\cap B(\mathfrak{H})$ is the weak (bounded) commutant introduced in Ref. 20.

(ii) $$\mathscr{R}'_c\equiv\mathscr{R}'_\sigma\cap L^+(V^\#)=\{X\in L^+(V^\#)\,|\,\langle A^*X^*f,g\rangle=\langle f,AXg\rangle;\ \forall f,g\in V^\#,\ \forall A\in\mathscr{R}\},$$

where $\mathscr{R}'_c$ is an Op*-algebra on $V^\#$.

(c) Finally, if $\mathscr{R}$ is an Op*-algebra, then we may define the following commutants:

(i) $\mathscr{R}'_0=\{X\in\text{Op } V\,|\,AX=XA;\ \forall A\in\mathscr{R}\}$, where $\mathscr{R}'_0$ is an x-invariant subset of Op V;

(ii) $\mathscr{R}'_\sigma\equiv\mathscr{R}'_0\cap C(V^\#,\mathfrak{H})$;

(iii) $\mathscr{R}'_c\equiv\mathscr{R}'_\sigma\cap L^+(V^\#)$ coincides with Inoue's commutant.[21]

D. Bicommutants

In this paper we will be concerned with the following bicommutants (for the other unbounded bicommutants, we refer to Refs. 1–3 and 19).

(a) Let $\mathscr{R}$ be an Op*-algebra. We define

$$\mathscr{R}''_{0c}=\{Y\in L^+(V^\#)\,|\,YX=XY;\ \forall X\in\mathscr{R}'_0\},$$

where $\mathscr{R}''_{0c}$ is an Op*-algebra on $V^\#$.

(b) Now let $\mathscr{R}$ be an x-invariant subset of $C(V^\#,\mathfrak{H})$. Then we may define the following bicommutant:

$$\mathscr{R}''_{\sigma\sigma}=\{Y\in C(V^\#,\mathfrak{H})\,|\,\langle Y^*f,Xg\rangle=\langle X^*f,Yg\rangle,\ \forall f,g\in V^\#,\ X\in\mathscr{R}'_\sigma\}.$$

Here $\mathscr{R}''_{\sigma\sigma}$ is an x-invariant linear subset of $C(V^\#,\mathfrak{H})$. These two bicommutants are related in the following way:

$$\mathscr{R}''_{0c}\subseteq\mathscr{R}''_{\sigma\sigma}\subseteq C(V^\#,\mathfrak{H}).$$

III. TOPOLOGICAL PROPERTIES OF THE COMMUTANTS

It is well known that for the algebra $B(\mathfrak{H})$ of bounded operators the usual commutants and bicommutant are closed in the weak (and *a fortiori* in the strong) topology.[22] For unbounded operators and unbounded commutants, this result is no longer true and the question is as follows: Under which topology is each commutant and bicommutant closed?

A. The strong unbounded commutant

Let $\mathscr{R}$ be an x-invariant subset of Op V. As mentioned in Ref. 1, $\mathscr{R}'_c$ is not in general weakly or quasiweakly closed in $L^+(V^\#)$. Similarly, if $\mathscr{R}$ is an Op*-algebra on $V^\#$, then

$\mathscr{R}''_{0c}$ need not be closed with respect to the weak- or quasiweak-topologies.

Proposition 3.1: $\mathscr{R}'_c$ is closed in $L^+(V^\#)$ with respect to the quasiweak*-topology.

Proof: Let $A \in L^+(V^\#)$ be the limit of a qw*-converging net $A_\alpha \in \mathscr{R}'_c$. This means, in particular, that for every $f \in V^\#$ and $h \in V$, we have

$$\langle h, A_\alpha f\rangle \to \langle h, Af\rangle \quad \text{and} \quad \langle h, A^*_\alpha f\rangle \to \langle h, A^* f\rangle.$$

Let $B \in \mathscr{R}$, i.e., $BA_\alpha = A_\alpha B$. Then for every $f,g \in V^\#$, we have

$$\langle Bg, A^* f\rangle = \lim_\alpha \langle Bg, A^*_\alpha f\rangle = \lim_\alpha \langle A_\alpha g, B^* f\rangle = \langle Ag, B^* f\rangle,$$

i.e., $A \in \mathscr{R}'_\sigma$. Since $A \in L^+(V^\#)$, we finally get that $A \in \mathscr{R}'_c$. ■

Corollary 3.2: If $\mathscr{R}$ is an Op*-algebra on $V^\#$, then $\mathscr{R}''_{0c}$ is closed in $L^+(V^\#)$ with respect to the qw*-topology.

Proposition 3.3: If $\mathscr{R}$ is an Op*-algebra on $V^\#$, then the commutant of $\mathscr{R}$ is equal to the commutant of its qw*-closure, i.e.,

$$\mathscr{R}'_0 = (\overline{\mathscr{R}}^{\mathrm{qw}^*})'_0.$$

Proof: The inclusion $(\overline{\mathscr{R}}^{\mathrm{qw}^*})'_0 \subset \mathscr{R}'_0$ follows from the fact that $\mathscr{R} \subset \overline{\mathscr{R}}^{\mathrm{qw}^*}$.

Let us prove the opposite inclusion. Let $B \in \overline{\mathscr{R}}^{\mathrm{qw}^*}$, i.e., there exists a net $\{B_\alpha\} \subset \mathscr{R}$ such that $B_\alpha \to B$, i.e., for every $f \in V^\#$ and $h \in V$, we have

$$\lim_\alpha \langle h, B_\alpha f\rangle = \langle h, Bf\rangle$$

and

$$\lim_\alpha \langle h, B^*_\alpha f\rangle = \langle h, B^* f\rangle.$$

Let $X \in \mathscr{R}'_0$, i.e., $XB_\alpha = B_\alpha X$. Then, for every $f,g \in V^\#$, we have

$$\langle B^* X^* f, g\rangle = \lim_\alpha \langle B^*_\alpha X^* f, g\rangle = \lim_\alpha \langle f, XB_\alpha g\rangle$$

$$= \lim_\alpha \langle f, B_\alpha X g\rangle = \langle f, BXg\rangle = \langle X^* B^* f, g\rangle$$

hence $X \in (\overline{\mathscr{R}}^{\mathrm{qw}^*})'_0$.

Remark 3.4: In the proof of Proposition 3.3 we used only the fact that

$$\lim_\alpha \langle h, B_\alpha f\rangle = \langle h, Bf\rangle, \quad f \in V^\#, \quad h \in V.$$

Therefore the commutant of $\mathscr{R}$ is also equal to the commutant of its quasiweak closure, so that finally we have

$$\mathscr{R}'_0 = (\overline{\mathscr{R}}^{\mathrm{qw}})'_0 = (\overline{\mathscr{R}}^{\mathrm{qw}^*})'_0.$$

Corollary 3.5: Let $\mathscr{R}$ and $\mathscr{R}_1$ be two Op*-algebras such that $\mathscr{R}_1 \subset \mathscr{R}$ and $\mathscr{R}_1$ is dense in $\mathscr{R}$ with respect to the qw*-topology. Then we have

$$(\mathscr{R}_1)'_0 = \mathscr{R}'_0.$$

Proof: The inclusion $\mathscr{R}'_0 \subset (\mathscr{R}_1)'_0$ follows from the inclusion $\mathscr{R}_1 \subset \mathscr{R}$. Next, $\mathscr{R}_1$ is qw*-dense in $\mathscr{R}$, i.e., $\overline{\mathscr{R}}^{\mathrm{qw}^*} \supset \mathscr{R}$ and hence $(\overline{\mathscr{R}}_1^{\mathrm{qw}^*})'_0 \subset \mathscr{R}'_0$. From Proposition 3.3, we have $(\overline{\mathscr{R}}_1^{\mathrm{qw}^*})'_0 = (\mathscr{R}_1)'_0$ which implies that $(\mathscr{R}_1)'_0 \subset \mathscr{R}'_0$.

Remark 3.6: This corollary could be applied to symmetric Op*-algebras [i.e., such that for every $A \in \mathscr{R}$ $(1 + A^*A)^{-1}$ exists and lies in the bounded part $\mathscr{R}_b \equiv \mathscr{R} \cap B(\mathfrak{H})$], because in this case $\mathscr{R}_b$ is dense in $\mathscr{R}$ with respect to the qw*-topology.

B. The weak unbounded commutant

Let $\mathscr{R}$ be a *-invariant subset of $C(V^\#, \mathfrak{H})$. As pointed out in Ref. 3, in general, $\mathscr{R}'_\sigma$ is not weakly or strongly closed in $C(V^\#, \mathfrak{H})$. It has been shown in Ref. 19 that $\mathscr{R}'_\sigma$ is closed in $C(V^\#, \mathfrak{H})$ with respect to the strong*-topology. Here we consider the quasiweak*-topology which is weaker than the s^*-topology.

Proposition 3.7: $\mathscr{R}'_\sigma$ is closed in $(CV^\#, \mathfrak{H})$ with respect to the quasiweak*-topology.

Proof: Follows from Ref. 19 (Proposition 1 and the remark after Corollary 2). ■

Corollary 3.8: $\mathscr{R}''_{\sigma\sigma}$ is closed in $C(V^\#, \mathfrak{H})$ with respect to the quasiweak*-topology.

C. The role of the quasiweak*-topology

The implication of the last two paragraphs is that the commutants $\mathscr{R}'_c$ and $\mathscr{R}'_\sigma$ as well as the bicommutants $\mathscr{R}''_{0c}$ and $\mathscr{R}''_{\sigma\sigma}$ are closed in the quasiweak*-topology, respectively, in $L^+(V^\#)$ and $C(V^\#, \mathfrak{H})$. Therefore this topology seems to be quite natural for the above mentioned commutants and bicommutants, and it is clear that it will play the role of the weak topology for bounded operators.

IV. BICOMMUTANTS AND THE QUASIWEAK*-CLOSURE OF $\mathscr{R}$

Since the bicommutants $\mathscr{R}''_{0c}$ and $\mathscr{R}''_{\sigma\sigma}$ are closed with respect to the quasiweak*-topology, the natural question to ask is whether they coincide with the quasiweak*-closure of $\mathscr{R}$.

A. The strong unbounded bicommutant

Proposition 4.1: Let $\mathscr{M}$ be an Op*-algebra of bounded operators. Then, $\mathscr{M}''_{0c}$ is the closure of $\mathscr{M}$ in $L^+(V^\#)$ with respect to the quasiweak*-topology, i.e.,

$$\mathscr{M}''_{0c} = \overline{\mathscr{M}}^{\mathrm{qw}^*} \cap L^+(V^\#).$$

Proof: The inclusion $\overline{\mathscr{M}}^{\mathrm{qw}^*} \subseteq \mathscr{M}''_{0c}$ follows from Corollary 3.2. Let us prove that $\mathscr{M}''_{0c} \subseteq \overline{\mathscr{M}}^{\mathrm{qw}^*} \cap L^+(V^\#)$.

(a) Let $f \in V^\#$ and consider the norm closed subspace $\overline{\mathscr{M}f}$ of $\mathfrak{H}$. Let P be the projection on $\overline{\mathscr{M}f}$. Since every $M \in \mathscr{M}$ is bounded and $\mathscr{M}$ is an algebra, $\mathscr{M}$ leaves $\overline{\mu f}$ invariant, i.e., $P \in \mathscr{M}'_0$. Now take $Y \in \mathscr{M}''_{0c}$ and any $g \in V^\#$. Then

$$\langle (1-P)Yf, g\rangle = \langle Yf, (1-P)g\rangle = \langle (1-P)f, Y^* g\rangle = 0,$$

i.e., $Yf = PYf$ and therefore $Yf \in \overline{\mathscr{M}f}$. Consequently if $Y \in \mathscr{M}''_{0c}$, it follows that for every $\epsilon > 0$, $f \in V^\#$, there exists $M \in \mathscr{M}$ such that $\|(Y-M)f\| < \epsilon$. In particular, $|\langle h, (Y-M)f\rangle| < \epsilon; \forall f \in V^\#, h \in V$.

(b) Next we show that $\mathscr{M}''_{0c} \subseteq \overline{\mathscr{M}}^{\mathrm{qw}} \cap L^+(V^\#)$.

We recall that zero neighborhoods in the quasiweak-topology are of the form

$$\mathscr{V}_{f_1,\ldots,f_n;h_1,\ldots,h_n;\epsilon}(0) = \{A \in L^+(V^\#) \,|\, |\langle h_1, Af_1\rangle| < \epsilon, \ldots, |\langle h_n, Af_n\rangle| < \epsilon\},$$

for any finite sequences $f_1,\dots,f_n$ of elements of $V^\#$, $h_1,\dots,h_n$ of elements of V and any $\epsilon>0$. It is sufficient to prove that if $Y\in\mathscr{M}''_{0c}$; $f_1,f_2\in V^\#$; $h_1,h_2\in V$ and $\epsilon>0$, then there exists $M\in\mathscr{M}$ such that Y-M belongs to $\mathscr{V}_{f_1,f_2;h_1,h_2;\epsilon}(0)$.

Consider the PIP space $V\oplus V$, with central Hilbert space $\mathfrak{H}\oplus\mathfrak{H}$, and the subalgebra $\mathscr{M}\oplus\mathscr{M}$ of $L^+(V^\#\oplus V^\#)$. Every $M\in\mathscr{M}$ gives rise to a bounded operator

$$\tilde{M}=\begin{pmatrix} M & 0\\ 0 & M\end{pmatrix}\quad\text{in } \mathfrak{H}\oplus\mathfrak{H}.$$

We denote by $\tilde{\mathscr{M}}$ the set of such operators, i.e.,

$$\tilde{\mathscr{M}}=\left\{\tilde{M}=\begin{pmatrix} M & 0\\ 0 & M\end{pmatrix}\middle| M\in\mathscr{M}\right\}.$$

The unbounded commutant and bicommutant of $\tilde{\mathscr{M}}$ may be computed explicitly and we obtain

$$\tilde{\mathscr{M}}'_0=\left\{X=\begin{pmatrix} X_{11} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\middle| X_{ij}\in\mathscr{M}'_0,\ i,j=1,2\right\}$$

and

$$\tilde{\mathscr{M}}''_{0c}=\left\{\tilde{Y}=\begin{pmatrix} Y & 0\\ 0 & Y\end{pmatrix}\middle| Y\in\mathscr{M}''_{0c}\right\}=\widetilde{M''_{0c}}.$$

Applying the results of part (a) of this proof to $\tilde{\mathscr{M}}''_{0c}$ we get that $\forall\tilde{Y}\in\tilde{\mathscr{M}}''_{0c},\tilde{f}\in V^\#\oplus V^\#$, $\tilde{h}\in V\oplus V$, and $\epsilon>0$, there exists $\tilde{M}\in\tilde{\mathscr{M}}$ such that

$$|\langle\tilde{h},(\tilde{Y}\text{-}\tilde{M})\tilde{f}\rangle|<\epsilon,$$

i.e., if $\tilde{f}=(f_1,f_2)$ and $\tilde{h}=(h_1,h_2)$, then we have

$$|\langle h_1,(Y\text{-}M)f_1\rangle|<\epsilon/2\quad\text{and}\quad|\langle h_2,(M-Y)f_2\rangle|<\epsilon/2.$$

Thus $(Y-M)\in V_{f_1,f_2;h_1,h_2;\epsilon/2}(0)$. Since this is true for any neighborhood, we have that $Y\in\overline{\mathscr{M}}^{\text{qw}}$. Since $\mathscr{M}''_{0c}$ lies in $L^+(V^\#)$ we finally get that

$$\mathscr{M}''_{0c}\subseteq\overline{\mathscr{M}}^{\text{qw}}\cap L^+(V^\#).$$

(c) In order to show that $\mathscr{M}''_{0c}\subseteq\overline{\mathscr{M}}^{\text{qw*}}\cap L^+(V^\#)$, we have to prove that if $Y-M$ lies in some zero neighborhood of the form above, then Y^*-M^* also belongs to that neighborhood. We can follow step by step the proof of part (b), but in this case we consider direct sums $V\oplus\bar{V}$, $\mathfrak{H}\oplus\bar{\mathfrak{H}}$, and $V^\#\oplus\overline{V^\#}$; where $\bar{V}$ (resp. $\bar{\mathfrak{H}}$ and $\bar{V}^\#$), denotes the space conjugate to V (resp. $\mathfrak{H}$, $V^\#$), i.e., the same set considered with the conjugate scalar multiplication $\lambda\circ f\equiv\bar{\lambda}f$ and equipped with the complex conjugate scalar product. Next we consider the set

$$\hat{\mathscr{M}}=\left\{\hat{M}=\begin{pmatrix} M & 0\\ 0 & M^*\end{pmatrix}\middle| M\in\mathscr{M}\right\}$$

equipped with the product

$$\begin{pmatrix} M & 0\\ 0 & M^*\end{pmatrix}\begin{pmatrix} N & 0\\ 0 & N^*\end{pmatrix}=\begin{pmatrix} MN & 0\\ 0 & N^*M^*\end{pmatrix}$$

and with the scalar conjugate multiplication

$$\lambda\circ\begin{pmatrix} M & 0\\ 0 & M^*\end{pmatrix}=\begin{pmatrix} \lambda M & 0\\ 0 & \bar{\lambda}M^*\end{pmatrix}.$$

Then $\mathscr{M}$ is an Op*-algebra on $V^\#\oplus\overline{V^\#}$.

Now we compute the commutant and bicommutant. We obtain

$$\hat{\mathscr{M}}'_0=\left\{\begin{pmatrix} X_{11} & 0\\ 0 & X_{22}\end{pmatrix}\middle| X_{11},X_{22}\in\mathscr{M}'_0\right\},$$

$$\hat{\mathscr{M}}''_{0c}=\left\{\begin{pmatrix} Y_{11} & 0\\ 0 & Y_{22}\end{pmatrix}\middle| Y_{11},Y_{22}\in\mathscr{M}''_{0c}\right\}.$$

In particular, for every $Y\in\mathscr{M}'_{0c}$ the element $\hat{Y}=\left(\begin{smallmatrix} Y & 0\\ 0 & Y^*\end{smallmatrix}\right)$ belongs to $\hat{\mathscr{M}}''_{0c}$. The application of part (a) of this proof to this particular element shows that for every $\hat{f}\in V^\#\oplus\overline{V^\#}$, $\hat{h}\in V\oplus\bar{V}$, and $\epsilon>0$, there exists $\hat{M}\in\hat{\mathscr{M}}$ such that

$$|\langle\hat{h},(\hat{Y}-\hat{M})\hat{f}\rangle|<\epsilon.$$

If we take $\hat{f}=(f,f)$ and $\hat{h}=(h,h)$, for any $f\in V^\#$, $h\in V$, we get

$$|\langle h,(Y-M)f\rangle|<\epsilon/2\quad\text{and}\quad|\langle h,(Y^*-M^*)f\rangle|<\epsilon/2,$$

i.e., $Y\in\overline{\mathscr{M}}^{\text{qw*}}$. Since $\mathscr{M}''_{0c}$ is contained in $L^+(V^\#)$ we finally get that

$$\mathscr{M}''_{0c}\subseteq\overline{\mathscr{M}}^{\text{qw*}}\cap L^+(V^\#).\qquad\blacksquare$$

Proposition 4.2: Let $\mathscr{R}$ be an Op*-algebra on $V^\#$, and assume that there exists an Op*-algebra $\mathscr{M}$ in the bounded part $\mathscr{R}_b$ such that $\mathscr{M}'_0=\mathscr{R}'_0$. Then we have

$$\mathscr{R}''_{0c}=\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#).$$

Proof: By Proposition 4.1, $\mathscr{R}'_0=\mathscr{M}'_0$ implies that

$$\mathscr{R}''_{0c}=\mathscr{M}''_{0c}=\overline{\mathscr{M}}^{\text{qw*}}\cap L^+(V^\#)\subseteq\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#).$$

On the other hand, $\mathscr{R}''_{0c}$ is closed in $L^+(V^\#)$ with respect to the qw*-topology, so that we have

$$\mathscr{R}''_{0c}\supset\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#).\qquad\blacksquare$$

Corollary 4.3: Let $\mathscr{R}$ be an Op*-algebra. If there exists an Op*-algebra $\mathscr{M}\subseteq\mathscr{R}_b$ dense in $\mathscr{R}$ for the qw*-topology, then

$$\mathscr{R}''_{0c}=\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#).$$

Proof: We use Corollary 3.5 to obtain $\mathscr{M}'_0=\mathscr{R}'_0$ and then we conclude with Proposition 4.2. $\blacksquare$

Remark 4.4: The assumptions of Corollary 4.3 are automatically satisfied if $\mathscr{R}$ is a symmetric Op*-algebra and $\mathscr{M}=\mathscr{R}_b$.

Following Ref. 6 we 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 4.5: Let $\mathscr{R}$ be an Op*-algebra on $V^\#$ and assume that for every $f\in V^\#$, the $\sigma(V,V^\#)$ closure $\overline{\mathscr{R}f}^\sigma f$ of $\mathscr{R}f$ is orthocomplemented in V. Then we have

$$\mathscr{R}''_{0c}=\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#).$$

Proof: The inclusion $\overline{\mathscr{R}}^{\text{qw*}}\cap L^+(V^\#)\subset\mathscr{R}''_{0c}$ follows from the fact that $\mathscr{R}''_{0c}$ is closed in $L^+(V^\#)$ with respect to the qw*-topology. Now let P be the projection on $\overline{\mathscr{R}f}^\sigma$. Since every $A\in\mathscr{R}$ is $\sigma(V,V^\#)$ continuous, it leaves $\widehat{\mathscr{R}f}^\sigma$ invariant, which means that $PA=AP$, i.e., $P\in\mathscr{R}'_0$. On the other hand, by definition $P\in L^+(V^\#)$, so that finally $P\in\mathscr{R}'_c$. Take $Y\in\mathscr{R}''_{0c}$ and $g\in V^\#$. Then we have

$$\langle(1-P)Yf,g\rangle=\langle Yf,(1-P)g\rangle=\langle(1-P)f,Y^*g\rangle=0$$

hence

$Yf = PYf$, i.e., $Yf \in \overline{\mathscr{R} f}^{\sigma}$. Thus, given $Y \in \mathscr{R}''_{0c}$, for every $\epsilon > 0, f \in V^{\#}$, $h \in V$, there exists $M \in \mathscr{R}$ such that

$$|\langle h,(Y-M)f\rangle| < \epsilon.$$

In order to show that $\mathscr{R}''_{0c} \subseteq \overline{\mathscr{R}}^{\text{qw*}} \cap L^{+}(V^{\#})$, we follow step by step the proof of Proposition 4.1.

Remark 4.6: In this Proposition it suffices to require that there exists an Op*-algebra $\mathscr{R}_0$ dense in $\mathscr{R}$ with respect to the qw*-topology, such that for every $f \in V^{\#}$, $\overline{\mathscr{R}_0 f}^{\sigma}$ is orthocomplemented in V. The results of Proposition 4.5 hold for $\mathscr{R}$ because $(\mathscr{R}_0)'_0 = \mathscr{R}'_0$.

Corollary 4.7: Let $\mathscr{R}$ be a symmetric Op*-algebra on $V^{\#}$. If for every $f \in V^{\#}$, $\overline{\mathscr{R}_b f}^{\sigma}$ is orthocomplemented in V, then

$$\mathscr{R}''_{0c} = \overline{R}^{\text{qw*}} \cap L^{+}(V^{\#}).$$

B. The weak unbounded bicommutant

Proposition 4.8: Let $\mathscr{M} \subset C(V^{\#},\mathfrak{H})$ be a *-algebra of bounded operators containing the identity. Then $\mathscr{M}''_{\sigma\sigma}$ is the closure of $\mathscr{M}$ in $C(V^{\#},\mathfrak{H})$ with respect to the qw*-topology, i.e., $\mathscr{M}''_{\sigma\sigma} = \overline{\mathscr{M}}^{\text{qw*}}$.

Proof: The inclusion $\overline{\mathscr{M}}^{\text{qw*}} \subseteq \mathscr{M}''_{\sigma\sigma}$ follows from the fact that $\mathscr{M}''_{\sigma\sigma}$ is closed in $C(V^{\#},\mathfrak{H})$ with respect to the qw*-topology. On the other hand, from Ref. 19, Proposition 9 we have that

$$M''_{\sigma\sigma} \subseteq \overline{\mathscr{M}}^{s*} \subseteq \overline{\mathscr{M}}^{\text{qw*}}. \qquad \blacksquare$$

Remark 4.9: In Ref. 19 it was shown that $\mathscr{M}''_{\sigma\sigma}$ is also the closure of $\mathscr{M}$ with respect to the s*-topology, so that finally we have

$$\mathscr{M}''_{\sigma\sigma} = \overline{\mathscr{M}}^{\text{qw*}} = \overline{\mathscr{M}}^{s*}.$$

Proposition 4.10: Let $\mathscr{R}$ be a *-invariant subset of $C(V^{\#},\mathfrak{H})$ and assume there exists a *-algebra $\mathscr{M}$ with unit in the bounded part $\mathscr{R}_b$ such that $\mathscr{M}'_{\sigma} = \mathscr{R}'_{\sigma}$. Then we have

$$\mathscr{R}''_{\sigma\sigma} = \overline{\mathscr{R}}^{\text{qw*}} = \overline{\mathscr{R}}^{s*}.$$

Proof: The equality $\mathscr{R}'_{\sigma} = \mathscr{M}'_{\sigma}$ implies that $\mathscr{R}''_{\sigma\sigma} = \mathscr{M}''_{\sigma\sigma} = \overline{\mathscr{M}}^{\text{qw*}}$. Now, since $\overline{\mathscr{M}}^{\text{qw*}} \subset \overline{\mathscr{R}}^{\text{qw*}}$, it follows that $\mathscr{R}''_{\sigma\sigma} \subset \overline{\mathscr{R}}^{\text{qw*}}$. The inclusion $\overline{\mathscr{R}}^{\text{qw*}} \subset \mathscr{R}''_{\sigma\sigma}$ follows from the fact that $\mathscr{R}''_{\sigma\sigma}$ is closed in $C(V^{\#},\mathfrak{H})$ with respect to the qw*-topology. The equality $\mathscr{R}''_{\sigma\sigma} = \overline{\mathscr{R}}^{s*}$ has been proved in Ref. 19, Proposition 11.

ACKNOWLEDGMENTS

The author would like to thank Professor J. P. Antoine, Dr. F. Mathot, and Dr. C. Trapani for stimulating discussions. He would also like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste, where this work was completed.

[1]J. Shabani, J. Math. Phys. **25**, 3204 (1984).
[2]F. Mathot and J. Shabani, "Generalized unbounded commutants," 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).
[3]G. Epifanio and C. Trapani, J. Math. Phys. **25**, 2633 (1984).
[4]J. P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 369 (1976).
[5]J. P. Antoine and A. Grossmann, J. Funct. Anal. **23**, 379 (1976).
[6]J. P. Antoine and A. Grossmann, J. Math. Phys. **19**, 329 (1978).
[7]J. P. Antoine, J. Math. Phys. **21**, 268 (1980).
[8]J. P. Antoine, J. Math. Phys. **21**, 2067 (1980).
[9]A. Grossmann, Commun. Math. Phys. **2**, 1 (1966).
[10]H. J. Borchers, in RCP 25 (Strasbourg) **22**, 26 (1975); *Quantum Dynamics: Models and Mathematics*, edited by L. Streit, Acta Phys. Austriaca Suppl. **16**, 15 (1976).
[11]J. P. Antoine and W. Karwowski, Publ. RIMS, Kyoto Univ. **21**, 205 (1985).
[12]J. P. Antoine and F. Mathot, Ann. Inst. H. Poincaré **37**, 29 (1982).
[13]G. Lassner, Rep. Math. Phys. **3**, 279 (1972).
[14]M. Friedrich and G. Lassner, Wiss. Z. Karl-Marx-Univ. Leipzig Math.-Natur. **R27**, 245 (1978).
[15]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).
[16]G. Köthe, *Topological Vector Spaces I* (Springer, Berlin, 1969).
[17]A. Inoue, K. Kuriyama, and S. Ôta, Memoirs Fac. Sci. Kyushu Univ. Ser. A: Math. Vol. 33, No. 2 (1979).
[18]O. Bratteli and D. Robinson, *Operator Algebras and Quantum Statistical Mechanics I* (Springer, Berlin, 1979).
[19]F. Mathot, J. Math. Phys. **26**, 1118 (1985).
[20]D. Ruelle, Helv. Phys. Acta **35**, 162 (1962).
[21]A. Inoue, Proc. Am. Math. Soc. **69**, 97 (1978).
[22]J. Dixmier, *Les Algèbres d'Operateurs dans l'Espace Hilbertien* (Gauthier–Villars, Paris, 1962).

Affine collineations in Robertson–Walker space-time

R. Maartens
Centre for Nonlinear Studies and Department of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg 2050, South Africa

(Received 31 December 1986; accepted for publication 6 May 1987)

In a recent paper [M. L. Bedran and B. Lesche, J. Math. Phys. **27**, 2360 (1986)], an attempt was made to find an affine collineation in Robertson–Walker space-time. However, only a homothetic affine collineation was found, in the case of a linear scale factor, $R(t)\sim t$. It is pointed out that another homothetic affine collineation exists when $R(t)\sim t^b$, and a proper (nonhomothetic) affine collineation in the Einstein static space-times has been found.

I. INTRODUCTION

An affine collineation[1] in a space-time is generated by a vector field ξ which leaves invariant the connection, and therefore the geodesics, of space-time, and which is characterized by

$$\mathscr{L}_\xi \Gamma^\alpha_{\beta\gamma} = \xi^\alpha_{\ ;\beta\gamma} + R^\alpha_{\ \beta\gamma\sigma}\xi^\sigma = 0 \Leftrightarrow \xi_{(\alpha;\beta);\gamma} = 0. \tag{1}$$

Special (improper) affine collineations are generated by homothetic Killing vectors

$$\xi_{(\alpha;\beta)} = \psi g_{\alpha\beta}, \quad \psi_{;\alpha} = 0, \tag{2}$$

which include Killing vectors ($\psi = 0$) as a further special case.

For the Robertson–Walker metric

$$ds^2 = dt^2 - R^2(t)[(1-kr^2)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)], \tag{3}$$

where $k = 0, \pm 1$, Bedran and Lesche[1] look for an isotropic and homogeneous ξ, and find that (1) implies

$$\xi = R(t)\partial_t, \tag{4}$$

and

$$R(t) = at + b, \tag{5}$$

where a,b are constants. As they point out, (3)–(5) imply that ξ is in fact homothetic. Furthermore, the field equations with (5) lead[1] to the rather unrealistic equation of state, $\rho + 3p = 0$.

Two questions arise from their results. First, are there other homothetic affine collineations of (3)? Second, can we find proper (i.e., nonhomothetic) affine collineations of (3)? These questions are answered in the following two sections.

II. HOMOTHETIC AFFINE COLLINEATIONS

A study of the conformal Killing vectors of Robertson–Walker space-time[2] shows that apart from (4), the only other homothetic Killing vector is

$$\xi = \tau\,\partial_\tau + r\,\partial_r, \quad \tau = \int \frac{dt}{R}, \tag{6}$$

with

$$R(t) = at^b \quad \text{and} \quad k = 0, \quad b \neq 1, \tag{7}$$

which was found by Eardley.[3] For suitable values of b, the field equations with (7) give the standard equation of state, $p = (\gamma - 1)\rho$, $1 \leqslant \gamma \leqslant 2$.

Both of the homothetic affine collineations, given by (4) and (5) and by (6) and (7), may be found directly from the conformal Killing algebra,[2] without reference to Eq. (1). The vector field (4) is the orthogonal conformal Killing vector, which becomes homothetic iff (5) holds. The vector field (6) is one of the eight nonorthogonal conformal Killing vectors,[2]

$$\xi = (1-kr^2)^{1/2}[h(\tau)\partial_\tau + h'(\tau)r\,\partial_r],$$

where

$$h(\tau) = (\tau, \cos\tau, \cosh\tau) \quad \text{for} \quad k = (0, 1, -1),$$

which becomes homothetic iff (7) holds.

III. PROPER AFFINE COLLINEATION

The complexity of (1) for the metric (3) forces us to make some simplifying assumptions in the search for affine collineations. Bedran and Lesche[1] assume that ξ is isotropic and homogeneous, and then find only the homothetic solution (4) and (5). This seems to imply that (3) does not admit an isotropic and homogeneous proper affine collineation. However, this is not the case, since Bedran and Lesche omitted the special solution in the static case $\dot{R} = 0$. (The first of the equations (2.4) in Ref. 1 need not hold in this special solution.) It is readily verified that

$$\xi = t\,\partial_t \tag{8}$$

satisfies (1) provided

$$R(t) = \text{const}, \tag{9}$$

but does not satisfy (2), since

$$\xi_{(\alpha;\beta)} = -t^{-2}\xi_\alpha\xi_\beta. \tag{10}$$

Thus (8) and (9) give the unique isotropic and homogeneous proper affine collineation in Robertson–Walker space-time.

For $k = 0$, (9) reduces (3) to Minkowski space-time, and (8) is one of the nine proper affine collineations in Minkowski space-time. This follows from the generation solution

$$\xi_a = A_{ab}x^b + B_a,$$

of (1), where A_{ab} and B_a are constant, and $x^a = (t,x,y,z)$

 0022-2488/87/092051-02$02.50

are orthonormal coordinates. A basis for the affine collineation Lie algebra G_{20} is thus given by $\mathbf{P}_a = \partial_a$ ($A_{ab} = 0$), $\mathbf{M}_{ab} = x_{[a}\partial_{b]}$ ($A_{(ab)} = 0 = B_a$), $\mathbf{X}_{ab} = x_{(a}\partial_{b)}$ ($A_{[ab]} = 0 = B_a$), where $x_a = g_{ab}x^b = (-t,x,y,z)$. The G_{10} Killing subalgebra is spanned by $\{\mathbf{P}_a, \mathbf{M}_{ab}\}$, while $\{\mathbf{X}_{ab}\}$ spans the homothetic G_1 subalgebra [$g^{ab}\mathbf{X}_{ab} = x^a\partial_a$, which is just (6) with $b = 0$ in (7)], and gives nine proper affine collineation vectors, including (9) ($\xi = -\mathbf{X}_{00}$).

For $k = \pm 1$, (9) gives the Einstein static space-times, for which the general solution of (1) is not known. The proper affine collineation vector (9), together with the static Killing vector ∂_t [which is just (4) with $a = 0$ in (5)], and the six Killing vectors[2] on $t = \text{const}$, form a G_8 of affine collineations, since they close under the Lie bracket.

IV. CONCLUDING REMARKS

We have seen that the problem of finding affine collineations in Robertson–Walker space-time may be completely solved in the homothetic case, but only partially solved in the general case. Any further affine collineations would necessarily be nonhomothetic, and also inhomogeneous or anisotropic (or both).

Apart from the information that they can provide about the space-time geometry, affine collineations are important because they generate first integrals of geodesic motion. This follows since $\xi_{(\alpha;\beta)}$ is a Killing tensor [see Eq. (1)], which implies that[4]

$$K = \xi_\alpha p^\alpha - v\xi_{\alpha;\beta}p^\alpha p^\beta, \quad p^\alpha = \frac{dx^\alpha}{dv}, \tag{11}$$

is constant along any geodesic $x^\alpha(v)$, where v is an affine parameter. Bedran and Lesche[1] use (11) with (4) and (5) in order to regain the timelike geodesics for $R(t) \sim t$. A similar analysis could be performed for the affine collineations (6) and (7) and (8) and (9); however, in these cases, (11) would not provide any advantage over the standard method of using the Killing vectors to determine the geodesics. The real usefulness of (11) is in space-times without many Killing vectors which admit an affine collineation.

Finally, we give an alternative derivation to that in Sec. III for the proper affine collineation (8) and (9), which may be applicable in other space-times. Suppose $\mathbf{X}$ is a Killing vector and F a nonconstant scalar, and let $\xi = F\mathbf{X}$. Then

$$\xi_{(\alpha;\beta);\gamma} = X_{(\alpha}F_{;\beta);\gamma},$$

and it follows that ξ will satisfy (1) if $X_\alpha = aF_{;\alpha}$ ($a = \text{const}$), which is (locally) equivalent to $X_{\alpha;\beta} = 0$ since $\mathbf{X}$ is Killing. Furthermore, ξ will generate a proper affine collineation, since

$$\xi_{(\alpha;\beta)} = a^{-1}X_\alpha X_\beta = a^{-1}F^{-2}\xi_\alpha\xi_\beta,$$

of which (10) is a particular case.

We conclude that if a space-time admits a Killing vector $\mathbf{X}$ which is a gradient, or equivalently a covariantly constant vector $\mathbf{X}$, then $F\mathbf{X}$ is a proper affine collineation vector, where $aF_{;\alpha} = X_\alpha$ for some constant a. For example, in the plane-fronted wave solutions with parallel rays,[5] the ray vector $\mathbf{k}$ satisfies $k_{\alpha;\beta} = 0$, so that $\xi = \phi\mathbf{k}$ is a proper affine collineation vector, where ϕ is the phase function ($k_\alpha = \phi_{;\alpha}$).

[1]M. L. Bedran and B. Lesche, J. Math. Phys. 27, 2360 (1986). We follow the notation and conventions of their paper.

[2]R. Maartens and S. D. Maharaj, Class. Quant. Gravit. 3, 1005 (1986).

[3]D. M. Eardley, Commun. Math. Phys. 37, 287 (1974).

[4]S. Hojman, L. Nuñez, A. Patiño, and H. Rago, J. Math. Phys. 27, 281 (1986).

[5]H. Stephani, *General Relativity* (Cambridge U. P., Cambridge, 1982).

Graded spinors as an underlying geometry for extended supersymmetries

C. P. Luehr and M. Rosenbaum
Centro de Estudios Nucleares, Universidad Nacional Autonoma de México, Circuito Exterior, C. U., 04510, México, D. F., Mexico

(Received 4 December 1986; accepted for publication 1 April 1987)

Spinors associated with an eight-dimensional Euclidean space are used for constructing a graded vector representation space for OSp$(4/N,\mathbb{R})$. The underlying geometry fixes both the nature and highest dimensionality of the family of internal symmetry groups of the bosonic subspace, as well as those of the fermionic subspace through an appropriately defined inner product. The basic ingredients of the theory are fermionic variables while the bosonic fields occur as composite quantities made up from an even number of products of the fermionic entities.

I. INTRODUCTION

Extended supersymmetries are obtained essentially by constructing graded vector spaces given by the orthogonal direct sum of two subspaces for which the ring of operators consist of even and odd elements of a Grassman algebra. Defining a scalar product for these graded vector spaces which decompose (because of orthogonality) into a sum of two scalar products for each of the subspaces, allows one to obtain the transformations which leave each of the scalars invariant as well as those that interchange the two vector subspaces but preserve the inner product for the total graded space. These isometries determine the characteristic supergroup for the graded vector space under consideration.

Elements in the subspace with even (odd) Grassmann coefficients are identified as "semiclassical" limits of bosonic (fermionic) field operators. The structure and dimensionality of these subspaces are, however, totally unrelated and imposed in an *ad hoc* fashion when constructing each of the possible supersymmetric theories.[1] Such an approach was followed by the authors[2] in developing a theory for supergravity based on supertwistor fiber bundles which, although it served to clarify some aspects related to the gauging of translations, provided no additional insight for a geometrical interpretation of the internal degrees of freedom.

Since supersymmetry appears to be a fairly well-established ingredient in the construction of theories of grand unification of gravitation with the other fundamental interactions in nature, investigations which might lead to a deeper understanding of the underlying geometry for the bosonic and fermionic components of these theories and the Grassmann nature of their semiclassical limit are worthwhile undertaking.

Efforts in this direction have been reported recently in the literature. Some of them[3] are based on attempts to link the division algebras with the existence and properties of supersymmetric theories in various space-time dimensions; another interesting approach makes use of the Kähler–Atiyah algebra to suggest a certain fusion between fermions and bosons by writing both types of fields in terms of inhomogeneous differential forms.[4]

The program that will be presented here, although it also makes use of inhomogeneous exterior forms to represent spinors, is basically different from the one just cited. Our idea consists essentially of using the theory of spinors of Cartan,[5] worded in the language of modern differential geometry with an added gradation from the Grassmann algebra, to construct both the fermionic and bosonic elements of a graded vector space. Thus the forms representing spinor spaces in our formalism are constructed from isotropic subspaces of an underlying Euclidean geometry, and have dimensions $2^{[n/2]}$ (where $[n/2]$ denotes the integer part of $n/2$), while the forms used in Ref. 4 have dimension 2^n. This supporting geometry fixes both the nature and highest dimensionality of the family of internal symmetry groups of the bosonic subspace, as well as those of the fermionic subspace through an appropriately defined inner product which breaks up the spinor representation space into a direct sum of isomorphic orthogonal fermionic subspaces.

The bosonic part of our graded vector space is derived by making use of a general theorem by Cartan whereby the tensor product of two spinors is shown to be completely reducible with respect to the group of rotations and reversals into a sum of $n+1$ (n = dimension of the Euclidean space with which the spinors are associated) irreducible tensors or pseudotensors. By this procedure the even Grassmann gradation of the bosonic subspace appears as a consequence of the tensor product of the two odd graded spinors from which it is constructed, and the basic ingredients of the theory are fermionic variables while the bosonic fields occur as composite quantities made up from an even number of products of the fermionic entities.

The resulting graded vector space, with an inner product induced by those for the odd and even Grassmann-graded subspaces, serves as a representation space to generate the characteristic supergroup of the theory.

More specifically, in the present paper we shall restrict consideration to the construction of representation spaces for OSp$(4/N,\mathbb{R})$. One reason for this choice is that it serves to exhibit in a natural and rather direct manner some of the ideas stated above. Another reason is the intrinsic physical interest of OSp$(4/N,\mathbb{R})$ supersymmetry, since by means of a Wigner–Inönü[6] contraction one can go from this graded group to extended supersymmetry with SU(2,2) supersymmetry for the fermionic subspace and an O(N) internal symmetry group for the bosonic part. These extended supersym-

 0022-2488/87/092053-13$02.50

metries have been widely used in the treatment of supergravity.

Finally we also remark that although the specific construction of the local gauge theory for $OSp(4/N,\mathbb{R})$ will not be undertaken in this paper, such a construction can be achieved by a procedure similar to the fiber bundle approach previously developed by the authors.[2]

The presentation is organized as follows: In Sec. II we give a brief review of the essential features of Cartan's spinor theory formulated in terms of exterior and Clifford algebras. We consider in particular the spinors associated with Euclidean spaces in even n dimensions. An additional gradation is then introduced by taking the ring of operators which belong to the module of spinors to be elements of an odd Grassmann subalgebra.

Section III is dedicated to the construction of a graded vector space where the elements of the fermionic subspace are obtained from self-conjugate (in the Cartan sense) Grassmannized spinors associated with a Euclidean space in eight dimensions. These spinors are then projected on a four-dimensional subspace which allows for a fermionic structure with a symplectic and real inner product, as well as for the construction of supersymmetric transformations that map spinors into real bosonic vectors. The nature of these bosonic elements, which appear as composite spinorial entities with an added gradation in the even Grassman algebra, is discussed for the different cases of internal symmetries in $SO(N)$, and an explicit procedure for obtaining the superalgebras for $OSp(4/N,\mathbb{R})$ is presented using our graded vector space as a representation space for such a construction. In Sec. IV we give some concluding remarks and suggestions for further work.

Finally, in the Appendix A we present some additional formulas, related to material used in the text, for the spinor operators which generate the Clifford algebra $G(8,0)$ and its complex extension $G^c = G(8,0) \otimes \mathbb{C}$, and give representations for the infinitesimal generators of Spin(8) in terms of the elements in the Lie algebra of sp(4). In Appendix B we exhibit the relation between the exterior and matrix formulation of spinor algebra for the benefit of those readers interested in applications and in comparing our approach with the work of others.

II. SPINORS AND GRASSMANN GRADED SPINORS

A. Spinors in $\mathscr{E}_n$

The systematic development of the theory of spinors in n dimensions, from a purely geometrical point of view, is due to Cartan.[5] Here we review only those aspects of the theory which are needed in order to make the discussion in the following sections as self-contained as possible. Although for the present paper we are specifically interested in spinors associated with $\mathscr{E}_8$, in the discussion that follows the basic concepts and definitions will be formulated for the general case of n-dimensional (n = even) Euclidean spaces, making use of the modern language of exterior algebra, where quantities are defined in an intrinsic fashion. Such a formulation is particularly convenient for larger-dimensional spinor spaces, for which the matrix notation used by Cartan makes calculations rather cumbersome. A detailed presentation of this approach to n-dimensional spinor theory, both for even and odd general pseudo-Euclidean spaces, will be the subject of a review paper to be published by the authors elsewhere. However, for the benefit of those readers who may be interested in comparing and applying our formalism to other work which is expressed in the more familiar language of gamma matrices and spinors as one-column matrices, a short outline which provides such a translation is given in Appendix B.

Let $\mathscr{E}_n$ $(n = 2\nu)$, denote an n-dimensional Euclidean space for which, as it is well known, the fundamental quadric can be always reduced to a sum of n-positive squares. Moreover, as shown by Cartan, in a Euclidean (or pseudo-Euclidean) space of dimension $n = 2\nu$, any isotropic subspace (i.e., a subspace in which all vectors have zero norm) has dimension $\leqslant \nu$. Making use of this last fact, we can choose as a basis for $\mathscr{E}_n$ the basis elements of the two ν-dimensional isotropic subspaces $\{\mathbf{e}_1,...,\mathbf{e}_\nu\} \in \mathscr{N}_\nu$ and $\{\mathbf{e}'_1,...,\mathbf{e}'_\nu\} \in \mathscr{N}'_\nu$, where $\mathbf{e}_i$ and $\mathbf{e}'_i$ are complex conjugate to each other, and $\mathscr{E}_n = \mathscr{N}_\nu \oplus \mathscr{N}'_\nu$. We also have the orthogonality conditions

$$\mathbf{e}'_k \cdot \mathbf{e}_l = \tfrac{1}{2}\delta_{kl}, \quad \mathbf{e}_k \cdot \mathbf{e}_l = \mathbf{e}'_k \cdot \mathbf{e}'_l = 0. \tag{2.1}$$

In terms of this basis, a vector $\mathbf{x} \in \mathscr{E}_n$ may be written as

$$\mathbf{x} = \mathbf{r} + \mathbf{r}' = x^i \mathbf{e}_i + x'^i \mathbf{e}'_i, \quad i = 1,...,\nu, \tag{2.2}$$

and the fundamental quadric is given by the expression

$$F = x^1 x'^1 + \cdots + x^\nu x'^\nu, \tag{2.3}$$

where $(x^i)^* = x'^i$.

For notational convenience, we introduce the basis $\{\boldsymbol{\epsilon}^i\} \in \mathscr{M}'_\nu$, dual to $\{\mathbf{e}_i\} \in \mathscr{N}_\nu$, by means of the $\mathbb{R}$-valued nondegenerate bilinear function

$$g(\boldsymbol{\epsilon}^k, \mathbf{e}_l) = \delta^k{}_l. \tag{2.4}$$

We can now set up the isomorphism $\mathscr{M}'_\nu \approx \mathscr{N}'_\nu$, by $\boldsymbol{\epsilon}^i \to \mathbf{e}'_i$, such that

$$g(\boldsymbol{\epsilon}^k, \mathbf{e}_l) = 2\mathbf{e}'_k \cdot \mathbf{e}_l = \delta_{kl}. \tag{2.5}$$

In a similar fashion we can introduce a basis $\{\boldsymbol{\epsilon}'^i\} \in \mathscr{M}_\nu$, dual to $\{\mathbf{e}'_i\} \in \mathscr{N}'_\nu$, and the isomorphism $\mathscr{M}_\nu \approx \mathscr{N}_\nu$, by $\boldsymbol{\epsilon}'^i \to \mathbf{e}_i$, such that

$$g(\boldsymbol{\epsilon}'^k, \mathbf{e}'_l) = 2\mathbf{e}_k \cdot \mathbf{e}'_l = \delta_{kl}. \tag{2.6}$$

In terms of the dual spaces thus introduced, a spinor space $\mathscr{S}'$, associated with $\mathscr{E}_n$, is defined as the direct sum

$$\mathscr{S}' \equiv \wedge \mathscr{M}'_\nu = \bigoplus_{p=0}^{\nu} \wedge^p \mathscr{M}'_\nu, \tag{2.7a}$$

i.e., $\mathscr{S}'$ is the module of complex p-vectors formed from $\mathscr{M}'_\nu$, with degrees p ranging from zero to ν.

Identifying each $\wedge^p \mathscr{M}'_\nu$ with its image under the canonical injection i_p: $\wedge^p \mathscr{M}'_\nu \to \wedge \mathscr{M}'_\nu$, we can write

$$\mathscr{S}' = \sum_{p=0}^{\nu} \wedge^p \mathscr{M}'_\nu. \tag{2.7b}$$

A spinor $\xi \in \mathscr{S}'$ can be written, therefore, as

$$\xi = \sum_{p=0}^{\nu} \sum_{<} \xi_{i_1,...,i_p} \boldsymbol{\epsilon}^{i_1} \wedge \cdots \wedge \boldsymbol{\epsilon}^{i_p}, \tag{2.8}$$

where the symbol $<$ indicates that the indices $(i_1,...,i_p)$ are subject to the condition $i_1 < i_2 < \cdots < i_p$. The set of 2^ν complex coefficients $\xi_{i_1,\ldots,i_p}$ (which have the property of changing sign or being unaltered under odd or even permutations of the indices) are the components of the spinor ξ.

Similarly, we can define a dual spinor space

$$\mathscr{S} = \sum_{p=0}^{\nu} \wedge^p \mathscr{M}_\nu, \tag{2.9}$$

in terms of exterior products of elements in $\mathscr{M}_\nu$. From now on, however, we will make use only of spinors in $\mathscr{S}'$.

B. The spinor transformation associated with a vector

For a vector $\mathbf{x}\in\mathscr{E}_n$, we can construct a linear operator $H(\mathbf{x})$ whose action on spinors in $\mathscr{S}'$ is defined by the Clifford product:

$$H(\mathbf{x})\xi = \xi \wedge \sum_{k=1}^{\nu} (x'^k \boldsymbol{\epsilon}^k) + \bar{g}(\xi,\mathbf{r}), \tag{2.10}$$

where[7]

$$\bar{g}(\xi\wedge\boldsymbol{\tau},\mathbf{r}) = \xi\wedge\bar{g}(\boldsymbol{\tau},\mathbf{r}) + (-1)^p\bar{g}(\xi,\mathbf{r})\wedge\boldsymbol{\tau},$$
$$\text{for } \xi\in\mathscr{S}', \quad \boldsymbol{\tau}\in\wedge^p\mathscr{M}'_\nu, \quad \mathbf{r}\in\mathscr{N}_\nu, \tag{2.11}$$

and, in particular, $\bar{g}(\boldsymbol{\tau},\mathbf{r}) = g(\boldsymbol{\tau},\mathbf{r})$, when $\boldsymbol{\tau}\in\mathscr{M}'_\nu$. Thus $\bar{g}(\cdot,\mathbf{r})$ is an antiderivation operator which maps $\wedge^p\mathscr{M}'_\nu \to \wedge^{p-1}\mathscr{M}'_\nu$.

Making use of (2.10), (2.11), and (2.5), it can be readily shown that

$$H(\mathbf{x})H(\mathbf{y}) + H(\mathbf{y})H(\mathbf{x}) = 2\mathbf{x}\cdot\mathbf{y}E, \quad \text{for } \mathbf{x},\mathbf{y}\in\mathscr{E}_n \tag{2.12}$$

where E is the identity operator on $\mathscr{S}'$.

Writing $H_i \equiv H(\mathbf{e}_i)$, $H_{i'} \equiv H(\mathbf{e}'_i)$, we get from (2.12) the following basic anticommutation relations:

$$\begin{aligned} &H_i H_k + H_k H_i = 0,\\ &H_{i'} H_{k'} + H_{k'} H_{i'} = 0,\\ &H_i H_{k'} + H_{k'} H_i = \delta_{ik} E. \end{aligned} \tag{2.13}$$

Hence the elements $(H_1 + H_{1'})$, $-i(H_1 - H_{1'})$, $(H_2 + H_{2'})$, $-i(H_2 - H_{2'}),\ldots,(H_\nu + H_{\nu'})$, $-i(H_\nu - H_{\nu'})$, are the generators of the Clifford algebra $G(n,0)$ and its complex extension $G^c = G(n,0)\otimes\mathbb{C}$.

C. Bilinear fundamental form on $\mathscr{S}'$

In the space of spinors, we can define a fundamental bilinear form by means of

$$(\xi|\lambda) := \sum_{p=0}^{\nu} \sum_{i,k<} (-1)^{(1/2)p(p+1)} \xi_{i_1\ldots i_p} \times \epsilon^{i_1\cdots i_p k_1\ldots k_{\nu-p}} \lambda_{k_1\cdots k_{\nu-p}}, \tag{2.14}$$

where $\epsilon^{i_1\cdots i_p k_1\cdots k_{\nu-p}}$ is the Levi-Civita symbol.

It is easy to show that

$$(\xi|\lambda) = (-1)^{(1/2)\nu(\nu+1)}(\lambda|\xi), \tag{2.15}$$

and

$$(H(\mathbf{x})\xi|\lambda) = (-1)^\nu(\xi|H(\mathbf{x})\lambda). \tag{2.16}$$

D. Conjugate spinors

According to Cartan, for a Euclidean space $\mathscr{E}_{2\nu}$, the conjugate ξ^c of a spinor ξ associated with such a space is an antilinear map $\xi^c \mapsto \xi$ defined by

$$\xi^c = i^\nu C\xi^*, \tag{2.17}$$

where

$$C = (H_1 - H_{1'})\cdots(H_\nu - H_{\nu'}), \tag{2.18}$$

and ξ^* denotes the spinor that results from taking the ordinary complex conjugate of the components of ξ in (2.8). It is clear from this definition and the properties (2.13) that

$$(\xi^c)^c = (-1)^{(1/2)\nu(\nu+1)}\xi. \tag{2.19}$$

Thus the map $\xi \mapsto \xi^c$ defines an involution if $\nu\equiv 0$ or -1 (mod 4), and an anti-involution if $\nu\equiv 1$ or 2 (mod 4).

The definition (2.19) is based on the application of the identity

$$CH(\mathbf{x}^*) = (-1)^\nu H(x)C, \tag{2.20}$$

where

$$x^* = x'^1\mathbf{e}_1 + \cdots + x'^\nu\mathbf{e}_\nu + x^1\mathbf{e}'_1 + x^\nu\mathbf{e}'_\nu. \tag{2.21}$$

The proof of (2.20) follows rather directly from (2.13) and (2.18).

It is important to remark here that because of (2.20), $(H(\mathbf{x})\xi)^c = (-1)^\nu H(\mathbf{x})\xi^c$. Thus, for ν even, spinors and their conjugates transform under (2.10) in the same way, while the ordinary complex conjugate of a spinor transforms differently, in general.

We also note, for the purpose of avoiding any possible confusion, that the Cartan operation of conjugation is not the same as charge conjugation defined for quantum mechanics. In fact, for a hyperbolic $\mathscr{E}_{3,1'}$ space where charge conjugation is naturally defined, it is not difficult to show that the relation between these two operations is given by $\xi^Q \sim (H_1 - H_{1'})(H_1 + H_{1'})(H_2 - H_{2'})(H_2 + H_{2'})\xi^c$, ξ^Q is the spinor which is charge conjugate to ξ.

E. Spinors as graded modules over a Grassmann algebra

We are now ready to extend the spinor formalism summarized above to the case where the ring of operators belonging to the module of spinors is made up of elements of the odd subset $\mathscr{G}_0$ of an infinite-dimensional Grassmann algebra. Thus the spinor components in (2.8) will be complex anticommuting quantities, and spinor space will be defined, therefore, by the Krönecker product $\mathscr{S}'_g = \mathscr{G}_0\otimes(\wedge\mathscr{M}'_\nu)$, i.e., $\mathscr{S}'_g$ contains a double gradation: the one due to the direct sum of p-vectors and one generated by $\mathscr{G}_0$. The assumption that spinor components are contained in $\mathscr{G}_0$ seems reasonable when one considers the odd elements of a Grassmann algebra as a semiclassical limit of fermionic operators in quantum field theory.

Since in the remainder of this paper we will be dealing with such extended spinors, it is important to determine which other of the definitions and relations introduced so far need to be modified. We will retain, however, the same notation that we have used for ordinary spinors, as there is no risk of confusion. It turns out, in fact, that the only relations that need to be changed are those involving ordering of the spinor components. Thus instead of (2.15) we will now have

$$(\xi|\lambda) = -(-1)^{(1/2)(\nu+1)\nu}(\lambda|\xi), \tag{2.15'}$$

which follows immediately from the definition of the fundamental bilinear (2.14) and the anticommutativity of the spinor components.

Also, recalling that for Grassmann quantities $(\xi_\alpha \lambda_\beta)^* = \lambda_\beta{}^* \xi_\alpha{}^*$ (where α, β are composite spinor indices), it is easy to show that

$$(\xi|\lambda)^* = -(\xi^*|\lambda^*). \tag{2.22}$$

The rest of the formulas for ordinary spinors presented in this section remain the same for Grassmannized spinors.

Finally we note that (2.17), (2.18), and (2.22) together, imply

$$(\xi^c|\lambda^c) = (\xi^*|\lambda^*) = -(\xi|\lambda)^*. \tag{2.23}$$

III. SPINOR GENERATED REPRESENTATION SPACES FOR OSp(4/*N*,ℝ)

As pointed out in the Introduction (cf. also Ref. 1), the main families of simple superalgebras that are relevant to supergravity are the orthosymplectic superalgebras osp(M/N) and the superunitary algebras su(M/N). Of the latter family the most interesting ones are the superconformal algebras su$(2,2/N)$. Moreover, since by means of a Wigner–Inönü contraction of osp(4/1) one can pass to the simple super-Poincaré algebra, while extended supersymmetries can be derived by a similar contraction of osp$(4/N)$, thus leading to gauge theories of supergravity with SO(N) as an interior symmetry group, consideration of the possible underlying geometries of OSp$(4/N)$ appears to be an important issue for a deeper understanding of these theories.[8,9]

We will now proceed to develop a formalism for obtaining representation spaces of OSp$(4/N)$, where both the fermionic and bosonic subspaces have structures induced by the geometry of the double graded spinor space $\mathscr{S}'_g := \mathscr{G}_0 \otimes \mathscr{S}'$, with a fundamental bilinear given by (2.14) which also satisfies (2.15').

One interesting feature of this approach is that the basic ingredients of the theory are anticommuting fermionic variables. The bosonic elements occur as composite quantities made up, via the spinorial construction, from an even number of fermionic entities.

A. The fermionic subspace

Since OSp$(4/N)$ contains as bosonic part of the ordinary Lie groups Sp(4) and SO(N), the fermionic subspace for a representation of the orthosymplectic supergroup will have to be a four-dimensional vector space with a symplectic inner product. If we want to take the graded spinors $\mathscr{S}'_g$ as elements for the construction of this subspace, as well as for the different bosonic representations of SO(N) such that the supersymmetric transformations of OSp$(4/N)$ map spinors into real vectors and vice versa, we need to satisfy the following requirements: (1) the fundamental symplectic spinor bilinear has to be real valued; and (2) the dimensions of the original Euclidean space, with which the spinors are associated, should accommodate the largest possible values of N for the internal symmetry groups that are physically interesting.

The second of the above requirements clearly suggests choosing $N = 8$, which in turn implies taking $\mathscr{E}_8$ as the underlying geometry for our spinorial constructions. In order to satisfy the condition of real-valuedness for the symplectic spinor bilinear, we note first that (2.17)–(2.19) and (2.23) together imply

$$[i(\xi + \xi^c|\lambda + \lambda^c)]^* = i(\xi^c + (-1)^{(1/2)\nu(\nu+1)}\xi|\lambda^c + (-1)^{(1/2)\nu(\nu+1)}\lambda). \tag{3.1}$$

Thus from odd Grassmann graded spinors associated with $\mathscr{E}_8$ $(\nu = 4)$ we can indeed obtain real scalar products. Observe, however, that this spinor space has dimension 2^4 and also that its metric, according to (2.15'), is symmetric. We must, therefore, investigate the possibility of constructing projection operators on our 16-dimensional self-conjugate spinors to get a four-dimensional subspace and simultaneously find a different inner product for the resulting spinors in this subspace with has a symplectic metric. There are several different ways to meet these two conditions. One of them is to introduce the projection operators

$$\mathscr{P}^{(\pm)} := \tfrac{1}{2}\,[E \pm (H_2 H_{2'} - H_{2'} H_2)(H_{3'} H_3 - H_3 H_{3'})] \tag{3.2}$$

and

$$\mathscr{Q}^{(\pm)} := \tfrac{1}{2}[E \mp S], \tag{3.3}$$

where, for a homogeneous p-vector $\xi^{(p)}$,

$$S\xi^{(p)} = \begin{cases} \xi^{(p)} & (p \text{ even}), \\ -\xi^{(p)} & (p \text{ odd}). \end{cases} \tag{3.4}$$

By means of a straightforward calculation it can be shown that for a self-conjugate spinor

$$\begin{aligned} \xi + \xi^c = {} & (\xi_0 + \xi^*_{1234}) + (\xi_1 - \xi^*_{234})\epsilon^1 + (\xi_2 + \xi^*_{134})\epsilon^2 + (\xi_3 - \xi^*_{124})\epsilon^3 + (\xi_4 + \xi^*_{123})\epsilon^4 + (\xi_{12} - \xi^*_{34})\epsilon^1 \wedge \epsilon^2 \\ & + (\xi_{13} + \xi^*_{24})\epsilon^1 \wedge \epsilon^3 + (\xi_{14} - \xi^*_{23})\epsilon^1 \wedge \epsilon^4 + (\xi_{23} - \xi^*_{14})\epsilon^2 \wedge \epsilon^3 + (\xi_{24} + \xi^*_{13})\epsilon^2 \wedge \epsilon^4 \\ & + (\xi_{34} - \xi^*_{12})\epsilon^3 \wedge \epsilon^4 + (\xi_{123} + \xi^*_4)\epsilon^1 \wedge \epsilon^2 \wedge \epsilon^3 + (\xi_{124} - \xi^*_3)\epsilon^1 \wedge \epsilon^2 \wedge \epsilon^4 \\ & + (\xi_{234} - \xi^*_1)\epsilon^2 \wedge \epsilon^3 \wedge \epsilon^4 + (\xi_{134} + \xi^*_2)\epsilon^1 \wedge \epsilon^3 \wedge \epsilon^4 + (\xi_{1234} + \xi^*_0)\epsilon^1 \wedge \epsilon^2 \wedge \epsilon^3 \wedge \epsilon^4, \end{aligned} \tag{3.5}$$

the projection operators (3.2) and (3.3) yield

$$\xi^{(++)} := \mathscr{Q}^{(+)}\mathscr{P}^{(+)}(\xi+\xi^c) = (\xi_2+\xi^*_{134})\epsilon^2 + (\xi_3-\xi^*_{124})\epsilon^3 + (\xi_{124}-\xi^*_3)\epsilon^1\wedge\epsilon^2\wedge\epsilon^4 + (\xi_{134}+\xi^*_2)\epsilon^1\wedge\epsilon^3\wedge\epsilon^4, \tag{3.6a}$$

$$\xi^{(-+)} := \mathscr{Q}^{(-)}\mathscr{P}^{(+)}(\xi+\xi^c) = (\xi_{12}+\xi^*_{34})\epsilon^1\wedge\epsilon^2 + (\xi_{13}+\xi^*_{24})\epsilon^1\wedge\epsilon^3 + (\xi_{24}+\xi^*_{13})\epsilon^2\wedge\epsilon^4 + (\xi_{34}-\xi^*_{12})\epsilon^3\wedge\epsilon^4, \tag{3.6b}$$

$$\xi^{(+-)} := \mathscr{Q}^{(+)}\mathscr{P}^{(-)}(\xi+\xi^c) = (\xi_1-\xi^*_{234})\epsilon^1 + (\xi_4+\xi^*_{123})\epsilon^4 + (\xi_{123}+\xi^*_4)\epsilon^1\wedge\epsilon^2\wedge\epsilon^3 + (\xi_{234}-\xi^*_1)\epsilon^2\wedge\epsilon^3\wedge\epsilon^4, \tag{3.6c}$$

$$\xi^{(--)} := \mathscr{Q}^{(-)}\mathscr{P}^{(-)}(\xi+\xi^c) = (\xi_0+\xi^*_{1234}) + (\xi_{14}-\xi^*_{23})\epsilon^1\wedge\epsilon^4 + (\xi_{23}-\xi^*_{14})\epsilon^2\wedge\epsilon^3 + (\xi_{1234}+\xi^*_0)\epsilon^1\wedge\epsilon^2\wedge\epsilon^3\wedge\epsilon^4, \tag{3.6d}$$

and

$$\xi+\xi^c = \xi^{(++)}\oplus\xi^{(+-)}\oplus\xi^{(-+)}\oplus\xi^{(--)}. \tag{3.7}$$

Moreover, defining a new spinor bilinear product by

$$\langle\xi+\xi^c|\lambda+\lambda^c\rangle := (\xi+\xi^c|(H_4H_{4'}-H_{4'}H_4)(\lambda+\lambda^c)) \tag{3.8}$$

we find

$$\langle\xi^{(++)}|\lambda^{(++)}\rangle = [(\xi_2+\xi^*_{134})(\lambda_{134}+\lambda^*_2) - (\xi_3-\xi^*_{124})(\lambda_{124}-\lambda^*_3) + (\xi_{124}-\xi^*_3)(\lambda_3-\lambda^*_{124}) - (\xi_{134}+\xi^*_2)(\lambda_2+\lambda^*_{134})], \tag{3.9a}$$

$$\langle\xi^{(-+)}|\lambda^{(-+)}\rangle = [(\xi_{12}-\xi^*_{34})(\lambda_{34}-\lambda^*_{12}) - (\xi_{13}+\xi^*_{24})(\lambda_{24}+\lambda^*_{13}) - (\xi_{34}-\xi^*_{12})(\lambda_{12}-\lambda^*_{34}) + (\xi_{24}+\xi^*_{13})(\lambda_{13}+\lambda^*_{24})], \tag{3.9b}$$

$$\langle\xi^{(+-)}|\lambda^{(+-)}\rangle = [(\xi_{123}+\xi^*_4)(\lambda_4+\lambda^*_{123}) - (\xi_1-\xi^*_{234})(\lambda_{234}-\lambda^*_1) - (\xi_4+\xi^*_{123})(\lambda_{123}+\lambda^*_4) + (\xi_{234}-\xi^*_1)(\lambda_1-\lambda^*_{234})], \tag{3.9c}$$

$$\langle\xi^{(--)}|\lambda^{(--)}\rangle = [(\xi_{23}-\xi^*_{14})(\lambda_{14}-\lambda^*_{23}) - (\xi_0+\xi^*_{1234})(\lambda_{1234}+\lambda^*_0) - (\xi_{14}-\xi^*_{23})(\lambda_{23}-\lambda^*_{14}) + (\xi_{1234}+\xi^*_0)(\lambda_0+\lambda^*_{1234})], \tag{3.9d}$$

and

$$\langle\xi^{(++)}|\lambda^{(+-)}\rangle = \langle\xi^{(++)}|\lambda^{(-+)}\rangle = \langle\xi^{(++)}|\lambda^{(--)}\rangle = \langle\xi^{(+-)}|\lambda^{(-+)}\rangle = \langle\xi^{(+-)}|\lambda^{(--)}\rangle = \langle\xi^{(-+)}|\lambda^{(--)}\rangle = 0. \tag{3.10}$$

Clearly (3.9a)–(3.9d) are real and have a symplectic metric. Note also that since

$$C\mathscr{P}^{(\pm)} = \mathscr{P}^{(\pm)}C, \quad C\mathscr{Q}^{(\pm)} = \mathscr{Q}^{(\pm)}C, \tag{3.11}$$

it follows that

$$[\mathscr{Q}^{(\pm)}\mathscr{P}^{(\pm)}\xi]^c = \mathscr{Q}^{(\pm)}\mathscr{P}^{(\pm)}\xi^c, \tag{3.12}$$

i.e., if ξ is self-conjugate (real, in the sense of Cartan) then $\xi^{(\pm\pm)}$ is also self-conjugate (real).

We can therefore choose either one of the four-dimensional subspaces of spinors $\xi^{(++)}$, $\xi^{(+-)}$, $\xi^{(-+)}$, $\xi^{(--)}$ associated with $\mathscr{E}_4\subset\mathscr{E}_8$, with inner products defined according to (3.8), as our fermionic subspace. We shall select as our basis for Sp(4) the set of real spinors

$$\mathscr{T}'_g = \{\xi^{(++)}\in\mathscr{S}'_g|(\xi^{(++)})^c = \xi^{(++)}, \mathscr{Q}^{(+)}\mathscr{P}^{(+)}\xi^{(++)} = \xi^{(++)}\}.$$

Furthermore, for simplicity we rewrite the product (3.9a) in terms of a metric by choosing a standard basis $\{\mathbf{p}^\alpha\}$ via the isomorphism

$$\epsilon^2\approx\mathbf{p}^1, \quad \epsilon^1\wedge\epsilon^3\wedge\epsilon^4\approx\mathbf{p}^2, \quad \epsilon^1\wedge\epsilon^2\wedge\epsilon^4\approx\mathbf{p}^3, \quad \epsilon^3\approx\mathbf{p}^4.$$

In terms of this new basis we have

$$\xi^{(++)} = (\xi_2+\xi^*_{134})\mathbf{p}^1 + (\xi_{134}+\xi^*_2)\mathbf{p}^2 + (\xi_{124}-\xi^*_3)\mathbf{p}^3 + (\xi_3-\xi^*_{124})\mathbf{p}^4. \tag{3.13}$$

We now let $\{\mathbf{q}_\alpha\}$ be the basis dual to $\{\mathbf{p}^\alpha\}$, i.e., $q_\alpha\circ\mathbf{p}^\beta = \delta^\beta{}_\alpha$, and define a symplectic metric tensor

$$\Gamma_{(1)} = \mathbf{q}_1\wedge\mathbf{q}_2 + \mathbf{q}_3\wedge\mathbf{q}_4 \tag{3.14}$$

such that

$$\langle\xi^{(++)}|\lambda^{(++)}\rangle = \xi^{(++)}\circ\Gamma_{(1)}\circ\lambda^{(++)} = \xi_\alpha^{(++)}\Gamma^{\alpha\beta}_{(1)}\lambda_\beta^{(++)}, \tag{3.15}$$

where

$$\xi_\alpha^{(++)} = (\xi_2+\xi^*_{134},\ \xi_{134}+\xi^*_2,\ \xi_{124}-\xi^*_3,\ \xi_3-\xi^*_{124}),$$

and

$$\Gamma^{\alpha\beta}_{(1)} := \mathbf{p}^\alpha\circ\Gamma_{(1)}\circ\mathbf{p}^\beta = \left[\begin{array}{cc|cc} 0 & 1 & & \\ -1 & 0 & \multicolumn{2}{c}{0} \\ \hline & & 0 & 1 \\ \multicolumn{2}{c|}{0} & -1 & 0 \end{array}\right]. \tag{3.16}$$

The inverse of $\Gamma_{(1)}$ is given by the tensor

$$\mathbf{G}_{(1)} = -\mathbf{p}^1 \wedge \mathbf{p}^2 - \mathbf{p}^3 \wedge \mathbf{p}^4. \tag{3.17}$$

Evidently,

$$\begin{aligned} \Gamma_{(1)} \circ \mathbf{G}_{(1)} &= \mathbf{q}_\alpha \otimes \mathbf{p}^\alpha \quad (\text{right identity on } \mathscr{T}'_q), \\ \mathbf{G}_{(1)} \circ \Gamma_{(1)} &= \mathbf{p}^\alpha \otimes \mathbf{q}_\alpha \quad (\text{left identity on } \mathscr{T}'_q). \end{aligned} \tag{3.18}$$

We next derive an explicit representation for the generators of the Lie algebra sp(4) in terms of our basis $\{\mathbf{p}^\alpha\}$ and metric $\Gamma_{(1)}$. To this end note that the transformations $\mathbf{U}$, with real coefficients in the odd Grassmann subalgebra that leave invariant the bilinear form (3.15) and satisfy the restriction

$$\widetilde{\mathbf{U}} \circ \Gamma_{(1)} \circ \mathbf{U} = \widetilde{U}^\beta{}_\alpha \Gamma^{\alpha\rho}_{(1)} U^\tau_\rho = \Gamma^{\rho\tau}_{(1)} \tag{3.19}$$

are elements of $\mathrm{Sp}(4,\mathbb{R})$.

The generators of sp(4) follow directly by first letting

$$\mathbf{U} = \mathbf{G}_{(1)} \circ \Gamma_{(1)} + \epsilon \mathbf{B}, \tag{3.20}$$

where ϵ is a real infinitesimal parameter in the odd Grassmann subalgebra. Substituting this expression into (3.19), results in the condition

$$\mathbf{D} := \Gamma_{(1)} \circ \mathbf{B} = (\Gamma_{(1)} \circ \mathbf{B})^\sim = \widetilde{\mathbf{D}} \tag{3.21}$$

[the symbol $^\sim$ in (3.21) denotes the usual operation of transposition on tensors]. It is now an easy matter to obtain all possible forms for the generators $\mathbf{B}$ of sp(4). In fact, by noting that $\widetilde{\mathbf{D}} = \mathbf{D}$ and [by (3.18)] $\mathbf{B} = \mathbf{G}_{(1)} \circ \mathbf{D} = (\mathbf{G}_{(1)} \circ \mathbf{D} \circ \mathbf{G}_{(1)}) \circ \Gamma_{(1)}$, we immediately get

$$\mathbf{B}^{(\alpha\beta)} = (\mathbf{p}^\alpha \otimes \mathbf{p}^\beta + \mathbf{p}^\beta \otimes \mathbf{p}^\alpha) \circ \Gamma_{(1)}, \tag{3.22a}$$

which satisfy the Lie algebra

$$\begin{aligned} [\mathbf{B}^{(\alpha\beta)}, \mathbf{B}^{(\gamma\delta)}] &:= \mathbf{B}^{(\alpha\beta)} \circ \mathbf{B}^{(\gamma\delta)} - \mathbf{B}^{(\gamma\delta)} \circ \mathbf{B}^{(\alpha\beta)} \\ &= \Gamma^{\beta\gamma}_{(1)} \mathbf{B}^{\alpha\delta} + \Gamma^{\beta\delta}_{(1)} \mathbf{B}^{(\alpha\gamma)} \\ &\quad + \Gamma^{\alpha\gamma}_{(1)} \mathbf{B}^{(\beta\delta)} + \Gamma^{\alpha\delta}_{(1)} \mathbf{B}^{(\beta\gamma)}. \end{aligned} \tag{3.22b}$$

B. The bosonic subspaces

These are representation spaces for SO(N), which we will construct from bilinear forms containing the real odd-Grassmann graded spinors in $\mathscr{S}'_g$ or their projected subspaces defined above. In order to obtain the form of the representation spaces, we resort to a general theorem due to Cartan (see Ref. 5, Sec. 131) by means of which the product of two spinors can be shown to be completely reducible (with respect to the group of proper and improper orthogonal transformations) into $2\nu + 1$ irreducible tensor spaces.

We shall consider first the most interesting case $N = 8$. The generators for the irreducible tensor spaces into which a product of two real spinors $\theta, \rho \in \mathscr{S}'_g$ decompose are the spinor bilinears $(\theta | \underset{(p)}{H} \rho)$ ($p = 0,1,\ldots,8$), where $\underset{(p)}{H}$ is the p-vector operator given by

$$\underset{(p)}{H} := \frac{1}{p!} \sum_\sigma (-1)^\sigma H(\mathbf{x}_{\sigma(1)}) \cdots H(\mathbf{x}_{\sigma(p)}). \tag{3.23}$$

Here the sum is over the set of all permutations σ of $\{1,\ldots,p\}$, $(-1)^\sigma = \pm$ is the sign of σ, and $H(\mathbf{x}_{\sigma(i)})$ is the linear spinor operator defined in (2.10).

Note now that under a reflection $H(\mathbf{a}_i)$ in the hyperplane normal to the unit vector $\mathbf{a}_i$ real spinors in $\mathscr{S}'_g$ and p-vector operators transform as

$$\begin{aligned} \theta &\mapsto \theta' = H(\mathbf{a}_i)\theta, \\ \underset{(p)}{H} &\mapsto \underset{(p)}{H}' = (-1)^p H(\mathbf{a}_i) \underset{(p)}{H} H(\mathbf{a}_i), \end{aligned} \tag{3.24a}$$

where the spinor operators $H(\mathbf{a}_i)$ are the generators of the Clifford algebra $G(8,0)$:

$$\begin{aligned} &H(\mathbf{a}_1) = H_1 + H_{1'}, \quad H(\mathbf{a}_2) = -i(H_1 - H_{1'}), \\ &H(\mathbf{a}_3) = H_2 + H_{2'}, \quad H(\mathbf{a}_4) = -i(H_2 - H_{2'}), \\ &H(\mathbf{a}_5) = (H_3 + H_{3'}), \quad H(\mathbf{a}_6) = -i(H_3 - H_{3'}), \\ &H(\mathbf{a}_7) = H_4 + H_{4'}, \quad H(\mathbf{a}_8) = -i(H_4 - H_{4'}). \end{aligned} \tag{3.24b}$$

We thus have, after making use of (2.16),

$$\left(\theta' \middle| \underset{(p)}{H}' \rho'\right) = (-1)^p \left(\theta \middle| \underset{(p)}{H} \rho\right). \tag{3.25}$$

Consequently $(\theta | \underset{(p)}{H} \rho)$ transforms under reflections as a scalar or pseudoscalar according to whether p is even or odd, and we can write in general

$$\left(\theta \middle| \underset{(p)}{H} \rho\right) = x^{i_1 \cdots i_p} y_{i_1 \ldots i_p}, \quad \text{for } p \text{ even}, \tag{3.26a}$$

$$\left(\theta \middle| \underset{(p)}{H} \rho\right) = x^{i_1 \ldots i_p} \epsilon_{i_1 \ldots i_p i_{p+1} \cdots i_n} y^{i_{p+1} \cdots i_n}, \quad \text{for } p \text{ odd}, \tag{3.26b}$$

where $x^{i_1 \cdots i_p}$ are the contravariant components of the p-vector $\underset{(p)}{H}$. The tensors $y_{i_1 \ldots i_p}$ and pseudotensors $\epsilon_{i_1 \ldots i_{p+1} \cdots i_n} y^{i_{p+1} \cdots i_n}$, which are bilinear with respect to the components of θ and ρ, are elements of nine irreducible representation spaces for O(8) of dimensions $\binom{8}{p} \equiv 8!/(8-p)!p!$. Futhermore, for the unimodular transformations in which we are interested, the tensorial and pseudotensorial representations transform in the same way and we can therefore, use either one indistinctly. This observation allows us to choose the pseudovectors $y_j = i(\theta | H_j \rho)$, $j = 1,1',\ldots,4,4'$, as a natural eight-dimensional representation space for SO(8). (The inclusion of the pure imaginary factor i in the definition of the pseudovectors y_j is needed in order to have $y_j^* = y_j$ so that $y_j \in \mathscr{G}_e \otimes \mathscr{C}_8$.) Explicitly, we have

$$\begin{aligned} y_1 = i[&(\theta_{1234} + \theta_0^*)(\rho_1 - \rho_{234}^*) - (\theta_{14} - \theta_{23}^*)(\rho_{123} + \rho_4^*) + (\theta_{13} + \theta_{24}^*)(\rho_{124} - \rho_3^*) \\ &- (\theta_{12} - \theta_{34}^*)(\rho_{134} + \rho_2^*) + (\theta_1 - \theta_{234}^*)(\rho_{1234} + \rho_0^*) - (\theta_{134} + \theta_2^*)(\rho_{12} - \rho_{34}^*) \\ &+ (\theta_{124} - \theta_3^*)(\rho_{13} + \rho_{24}^*) - (\theta_{123} + \theta_4^*)(\rho_{14} - \rho_{23}^*)], \end{aligned} \tag{3.27a}$$

$$y_2 = i[(\theta_{23} - \theta^*_{14})(\rho_{124} - \rho^*_3) - (\theta_{24} + \theta^*_{13})(\rho_{123} + \rho^*_4) - (\theta_{12} - \theta^*_{34})(\rho_{234} - \rho^*_1)$$
$$- (\theta_{234} - \theta^*_1)(\rho_{12} - \rho^*_{34}) + (\theta_2 + \theta^*_{134})(\rho_{1234} + \rho^*_0) + (\theta_{124} - \theta^*_3)(\rho_{23} - \rho^*_{14})$$
$$- (\theta_{123} + \theta^*_4)(\rho_{24} + \rho^*_{13}) + (\theta_{1234} + \theta^*_0)(\rho_2 + \rho^*_{134})], \qquad (3.27b)$$
$$y_3 = i[(\theta_{1234} + \theta^*_0)(\rho_3 - \rho^*_{124}) + (\theta_{23} - \theta^*_{14})(\rho_{134} + \rho^*_2) - (\theta_{13} + \theta^*_{24})(\rho_{234} - \rho^*_1)$$
$$- (\theta_{34} - \theta^*_{12})(\rho_{123} + \rho^*_4) - (\theta_{234} - \theta^*_1)(\rho_{13} + \rho^*_{24}) + (\theta_{134} + \theta^*_2)(\rho_{23} - \rho^*_{14})$$
$$+ (\theta_3 - \theta^*_{124})(\rho_{1234} + \rho^*_0) - (\theta_{123} + \theta^*_4)(\rho_{34} - \rho^*_{12})], \qquad (3.27c)$$
$$y_4 = i[(\theta_{1234} + \theta^*_0)(\rho_4 + \rho^*_{123}) - (\theta_{14} - \theta^*_{23})(\rho_{234} - \rho^*_1) + (\theta_{24} + \theta^*_{13})(\rho_{134} + \rho^*_2)$$
$$- (\theta_{34} - \theta^*_{12})(\rho_{124} - \rho^*_3) - (\theta_{234} - \theta^*_1)(\rho_{14} - \rho^*_{23}) + (\theta_{134} + \theta^*_2)(\rho_{24} + \rho^*_{13})$$
$$- (\theta_{124} - \theta^*_3)(\rho_{34} - \rho^*_{12}) + (\theta_4 + \theta^*_{123})(\rho_{1234} + \rho^*_0)], \qquad (3.27d)$$

and

$$y'_1 = y^*_1, \quad y'_2 = y^*_2, \quad y'_3 = y^*_3, \quad y'_4 = y^*_4. \qquad (3.27e)$$

Extending now the canonical isomorphism introduced previously for the subspace $\mathscr{T}'_g$ [see discussion preceding (3.13)] to all of the real-spinor space $\mathscr{S}'_g$, by

$$\epsilon^2 \approx \mathbf{p}^1, \quad \epsilon^1\wedge\epsilon^3\wedge\epsilon^4 \approx \mathbf{p}^2, \quad \epsilon^1\wedge\epsilon^2\wedge\epsilon^4 \approx \mathbf{p}^3, \quad \epsilon^3 \approx \mathbf{p}^4,$$
$$\epsilon^1\wedge\epsilon^2\wedge\epsilon^3 \approx \mathbf{p}^5, \quad \epsilon^4 \approx \mathbf{p}^6, \quad \epsilon^2\wedge\epsilon^3\wedge\epsilon^4 \approx \mathbf{p}^7, \quad \epsilon^1 \approx \mathbf{p}^8,$$
$$\epsilon^1\wedge\epsilon^2 \approx \mathbf{p}^9, \quad \epsilon^3\wedge\epsilon^4 \approx \mathbf{p}^{10}, \quad \epsilon^2\wedge\epsilon^4 \approx \mathbf{p}^{11},$$
$$\epsilon^1\wedge\epsilon^3 \approx \mathbf{p}^{12}, \quad \epsilon^1\wedge\epsilon^2\wedge\epsilon^3\wedge\epsilon^4 \approx \mathbf{p}^{13},$$
$$\mathbf{1} \approx \mathbf{p}^{14}, \quad \epsilon^2\wedge\epsilon^3 \approx \mathbf{p}^{15}, \quad \epsilon^1\wedge\epsilon^4 \approx \mathbf{p}^{16},$$

we can write

$$y_j = (i\boldsymbol{\theta}\otimes\boldsymbol{\rho})\,\mathring{}\,\mathbf{E}_{(j)} = i\boldsymbol{\theta}\circ\mathbf{E}_{(j)}\circ\boldsymbol{\rho}$$
$$= (i\theta_\alpha\rho_\beta)E^{\alpha\beta}_{(j)}, \quad \alpha,\beta = 1,\ldots,16; \qquad (3.28)$$

where

$$\mathbf{E}_{(j)} = \mathbf{T}\circ H_j, \quad E^{\alpha\beta}_{(j)} = \mathbf{p}^\alpha\circ\mathbf{E}_{(j)}\circ\mathbf{p}^\beta, \quad j = 1,1',\ldots,4,4', \qquad (3.29)$$

$$\mathbf{T} := \sum_{i=1}^{8} (-1)^i(\mathbf{q}_{2i-1}\otimes\mathbf{q}_{2i} + \mathbf{q}_{2i}\otimes\mathbf{q}_{2i-1}). \qquad (3.30)$$

We thus see that the tensors $\mathbf{E}_{(j)}$ act as an isotropic basis for $\mathscr{G}_e\otimes\mathscr{E}_8$, whereby vectors are expressed as tensor products of real spinors with anticommuting components.

Equations (3.29) and (3.30) also serve to display the relationship between the generators of the isometries of our fermionic and bosonic subspaces. For this purpose note that if, in analogy to the metric tensor (3.14) for the symplectic product of real spinors in $\mathscr{T}'_g$, we introduce similar metrics for the other projected subspaces $\{\xi^{(+-)}\}$, $\{\xi^{(-+)}\}$, $\{\xi^{(--)}\}$, into which a real spinor splits up as a direct sum [cf. Eq. (3.7)], we can define a symplectic metric for $\mathscr{S}'_g$ by

$$\boldsymbol{\Gamma} := \boldsymbol{\Gamma}_{(1)}\oplus\boldsymbol{\Gamma}_{(2)}\oplus\boldsymbol{\Gamma}_{(3)}\oplus\boldsymbol{\Gamma}_{(4)}, \qquad (3.31)$$

where

$$\boldsymbol{\Gamma}_{(2)} = \mathbf{q}_5\wedge\mathbf{q}_6 + \mathbf{q}_7\wedge\mathbf{q}_8, \quad \boldsymbol{\Gamma}_{(3)} = \mathbf{q}_9\wedge\mathbf{q}_{10} + \mathbf{q}_{11}\wedge\mathbf{q}_{12},$$
$$\boldsymbol{\Gamma}_{(4)} = \mathbf{q}_{13}\wedge\mathbf{q}_{14} + \mathbf{q}_{15}\wedge\mathbf{q}_{16}. \qquad (3.32)$$

It is then easy to show that (3.30) may be written as

$$\mathbf{T} = \boldsymbol{\Gamma}\circ\left[\sum_{i=1}^{8}(-1)^i(\mathbf{p}^{2i-1}\otimes\mathbf{p}^{2i} + \mathbf{p}^{2i}\otimes\mathbf{p}^{2i-1})\right]\circ\boldsymbol{\Gamma}$$
$$= \sum_{i=1}^{8}(-1)^i\boldsymbol{\Gamma}\circ\mathbf{B}^{(2i-1,2i)}, \qquad (3.33)$$

where the $\mathbf{B}^{(2i-1,2i)}$ are the generators of the Lie algebra of sp(4) which we derived in (3.22).

In terms of (3.33), the expression (3.29) for the isotropic basis tensors for $\mathscr{G}_e\otimes\mathscr{E}_8$ now becomes

$$\mathbf{E}_{(j)} = \left(\sum_{i=1}^{8}(-1)^i\boldsymbol{\Gamma}\circ\mathbf{B}^{(2i-1,2i)}\right)\circ H_j$$
$$= [\boldsymbol{\Gamma}_{(1)}(B^{(34)} - B^{(12)})\oplus\boldsymbol{\Gamma}_{(2)}(B^{(34)} - B^{(12)})\oplus\cdots$$
$$\oplus\boldsymbol{\Gamma}_{(4)}(B^{(34)} - B^{(12)})]H_j. \qquad (3.34)$$

Here the quantity in square brackets is a 16×16 block diagonal matrix, where each of the block members in the direct sum is the same and has the form

$$\boldsymbol{\Gamma}_{(i)}(B^{(34)} - B^{(12)})$$
$$= \tfrac{1}{2}(B^{(11)} + B^{(22)} + B^{(33)} + B^{(44)})(B^{(34)} - B^{(12)})$$
$$= -\sigma_3\otimes\sigma_1$$

(the σ's are ordinary right-handed Pauli matrices), while the matrices H_j are also expressible in terms of elements of the Lie algebra of sp(4) as shown in Eqs. (A2) in Appendix A.

Consequently the H_j, whose even degree Clifford algebra serves to generate the elements of Spin(8), which is the 2–1 covering group of SO(8), and the representation vectors $\mathbf{E}_j$ [as given by (3.34)] for this latter group, are both determined by the Lie algebra of the generators of isometries of our fermionic subspace.

Our next step is to use the basis $\{\mathbf{E}_{(j)}\}$ to construct a metric tensor on $\mathscr{G}_e\otimes\mathscr{E}_8$. This leads us to the need to introduce the basis $\{\mathbf{E}^{(j)}\}$, dual to $\{\mathbf{E}_{(j)}\}$, and defined by

$$\mathbf{E}^{(j)} = \tfrac{1}{8}\,\mathbf{G}\circ\mathbf{T}\circ\mathbf{G}\circ\widetilde{H}_j. \qquad (3.35)$$

Clearly

$$\mathbf{E}_{(j)}\cdot\mathbf{E}^{(k)} = \mathbf{E}^{(k)}\cdot\mathbf{E}_{(j)} := \mathbf{E}^{(k)}\,\mathring{}\,\mathbf{E}_{(j)} = \delta^k{}_j. \qquad (3.36)$$

An appropriate metric tensor $\boldsymbol{\Delta}$ on $\mathscr{G}_e \otimes \mathscr{E}_8$ is then

$$\boldsymbol{\Delta} = \frac{1}{2} \sum_{j=i}^{4} (\mathbf{E}_{(j)} \otimes \mathbf{E}_{(j')} + \mathbf{E}_{(j')} \otimes \mathbf{E}_{(j)}). \tag{3.37}$$

In terms of this metric, the scalar product of any two vectors in $\mathscr{G}_e \otimes \mathscr{E}_8$,

$$\mathbf{w} = \sum_j [(i\theta_\alpha \rho_\beta) E^{\alpha\beta}_{(j)}] \mathbf{E}^{(j)} \quad (j = 1, 1', \ldots 4, 4'),$$
$$\mathbf{z} = \sum_j [(i\mu_\alpha \nu_\beta) E^{\alpha\beta}_{(j)}] \mathbf{E}^{(j)}, \tag{3.38}$$

is given by

$$\mathbf{w}\cdot\boldsymbol{\Delta}\cdot\mathbf{z} = \frac{1}{2} \sum_j [(i\theta_\alpha \rho_\beta) E^{\alpha\beta}_{(j)}][(i\mu_\gamma \nu_\delta) E^{\gamma\delta}_{(j')}]. \tag{3.39}$$

Moreover, if in analogy with (3.15), we now set

$$\langle \mathbf{E}^{(i)} | \mathbf{E}^{(j)} \rangle := \mathbf{E}^{(i)}\cdot\boldsymbol{\Delta}\cdot\mathbf{E}^{(j)}, \tag{3.40}$$

we then have the following equivalent expressions for the scalar product in $\mathscr{G}_e \otimes \mathscr{E}_8$:

$$\langle \mathbf{w} | \mathbf{z} \rangle := \mathbf{w}\cdot\boldsymbol{\Delta}\cdot\mathbf{z} = \langle (i\boldsymbol{\theta} \otimes \boldsymbol{\rho}) | (i\boldsymbol{\mu} \otimes \boldsymbol{\nu}) \rangle$$
$$= \langle (i\theta_\alpha \rho_\beta E^{\alpha\beta}_{(j)}) \mathbf{E}^{(j)} | (i\mu_\gamma \nu_\delta E^{\gamma\delta}_{(k)}) \mathbf{E}^{(k)} \rangle. \tag{3.41}$$

By construction [cf. Eqs. (3.27) and (3.28)], the components of any vector $\mathbf{w}$, defined as in (3.38), satisfy the property $(w_1)^* = w_1'$, $(w_2)^* = w_2'$, $(w_3)^* = w_3'$, $(w_4)^* = w_4'$. Hence the scalar (3.41) is real. In fact (3.37) can be rewritten in the form

$$\boldsymbol{\Delta} = \frac{1}{4} \sum_{a=1}^{8} \mathbf{A}_{(a)} \otimes \mathbf{A}_{(a)}, \tag{3.37$'$}$$

where

$$\begin{aligned}
\mathbf{A}_{(1)} &= \mathbf{E}_{(1)} + \mathbf{E}_{(1')}, & \mathbf{A}_{(2)} &= -i(\mathbf{E}_{(1)} - \mathbf{E}_{(1')}), \\
\mathbf{A}_{(3)} &= \mathbf{E}_{(2)} + \mathbf{E}_{(2')}, & \mathbf{A}_{(4)} &= -i(\mathbf{E}_{(2)} - \mathbf{E}_{(2')}), \\
\mathbf{A}_{(5)} &= \mathbf{E}_{(3)} + \mathbf{E}_{(3')}, & \mathbf{A}_{(6)} &= -i(\mathbf{E}_{(3)} - \mathbf{E}_{(3')}), \\
\mathbf{A}_{(7)} &= \mathbf{E}_{(4)} + \mathbf{E}_{(4')}, & \mathbf{A}_{(8)} &= -i(\mathbf{E}_{(4)} - \mathbf{E}_{(4')}).
\end{aligned} \tag{3.37$''$}$$

When (3.37′) is substituted into (3.41), it yields real quantities for each vector component in the scalar product.

It is easy to verify that the generators of the infinitesimal transformations which leave (3.41) invariant are of the form

$$\mathbf{M}^{(ab)} = (\mathbf{A}^{(a)} \wedge \mathbf{A}^{(b)})\cdot\boldsymbol{\Delta}, \tag{3.42}$$

and satisfy the commutation on relations

$$\begin{aligned}
[\mathbf{M}^{(ab)}, \mathbf{M}^{(cd)}] &:= \mathbf{M}^{(ab)}\cdot\mathbf{M}^{(cd)} - \mathbf{M}^{(cd)}\cdot\mathbf{M}^{(ab)} \\
&= \delta_{bc}\mathbf{M}^{(ad)} - \delta_{bd}\mathbf{M}^{(ac)} \\
&\quad - \delta_{ca}\mathbf{M}^{(bd)} + \delta_{ad}\mathbf{M}^{(bc)}.
\end{aligned} \tag{3.43}$$

Therefore $\mathbf{M}^{(ab)} \in \mathrm{so}(8)$.

C. Representation space for OSp(4/8,ℝ)

Combining the above results, we can now construct the direct sum space $(\mathscr{G}_e \otimes \mathscr{E}_8) \oplus \mathscr{T}_g'$ with elements $\mathbf{Z} = \mathbf{z} + \boldsymbol{\xi}^{(++)}$, where $\mathbf{z} \in \mathscr{G}_e \otimes \mathscr{E}_8$ is of the form given in (3.38) and $\boldsymbol{\xi}^{(++)} \in \mathscr{T}_g'$ is a real spinor of the form (3.13).

In component notation we can write

$$Z_A = (z_a, \xi_\alpha^{(++)}), \quad a = 1, 2, \ldots, 8, \quad \alpha = 1, \ldots, 4.$$

We next define a gradation-respecting scalar product by

$$\langle \mathbf{W} | \mathbf{Z} \rangle := \langle \mathbf{w} | \mathbf{z} \rangle + \langle \boldsymbol{\xi}^{(++)} | \boldsymbol{\eta}^{(++)} \rangle, \tag{3.44}$$

where $\langle \mathbf{w} | \mathbf{z} \rangle$ and $\langle \boldsymbol{\xi}^{(++)} | \boldsymbol{\eta}^{(++)} \rangle$ are given by (3.41) and (3.15), respectively.

Clearly, since both $\langle \mathbf{w} | \mathbf{z} \rangle$ and $\langle \boldsymbol{\xi}^{(++)} | \boldsymbol{\eta}^{(++)} \rangle$ are real, so is (3.44).

In terms of the metric tensors (3.16) and (3.37′), the scalar product (3.44) can be expressed as

$$W_A \Lambda^{AB} Z_A = w_a \Delta^{ab} z_b + \xi_\alpha^{(++)} \Gamma^{\alpha\beta}_{(1)} \eta_\beta^{(++)}, \tag{3.45}$$

where Λ^{AB} is the block diagonal metric

$$\Lambda^{AB} = \left[\begin{array}{c|c} \Delta^{ab} & 0 \\ \hline 0 & \Gamma^{\alpha\beta}_{(1)} \end{array} \right]. \tag{3.46}$$

Because the product (3.44) respects gradation, the symplectic transformations, whose generators are given by (3.22), will transform spinors among themselves, and the special orthogonal transformations generated by (3.42) will transform vectors into vectors. Consequently, we only have left to consider those supersymmetric transformations which map spinors into vectors and vice versa. To this end, let $\mathbf{V} \in (\mathscr{G}_e \otimes \mathscr{E}_8) \otimes \mathscr{T}_g' \oplus \mathscr{T}_g' \otimes (\mathscr{G}_e \otimes \mathscr{E}_8)$ denote the transformations $Z_A \mapsto Z'_A = V_A{}^B Z_B$, such that (3.44) remains invariant. Since we are only interested in the subset of $\mathbf{V}$ with graded determinant equal to 1, we can consider infinitesimal transformations of the form

$$V_A{}^B = \mathbb{1}_A{}^B + \epsilon S_A{}^B, \tag{3.47}$$

where ϵ is an infinitesimal real parameter, and $\mathbb{1}_A{}^B$ acts as the identity operator on $(\mathscr{G}_e \otimes \mathscr{E}_8) \oplus \mathscr{T}_g'$.

Invariance of (3.44) implies

$$S_A{}^D W^D \Lambda^{AB} Z_B + W_A \Lambda^{AB} S_B{}^D Z_D = 0. \tag{3.48}$$

Now let $S_A{}^B = \underset{(1)}{S}{}_a{}^\beta + \underset{(2)}{S}{}_\alpha{}^b$, where $\underset{(1)}{S}{}_a{}^\beta \in (\mathscr{G}_e \otimes \mathscr{E}_8) \otimes \mathscr{T}_g'$ and $\underset{(2)}{S}{}_\alpha{}^b \in \mathscr{T}_g' \otimes (\mathscr{G}_e \otimes \mathscr{E}_8)$. Equation (3.48) then becomes

$$\underset{(1)}{S}{}_a{}^\delta \xi_\delta^{(++)} \Delta^{ab} z_b + \xi_\alpha^{(++)} \Gamma^{\alpha\beta}_{(1)} \underset{(2)}{S}{}_\beta{}^d z_d$$
$$+ \underset{(2)}{S}{}_\alpha{}^d w_d \Gamma^{\alpha\beta}_{(1)} \eta_\beta^{(++)} + w_a \Delta^{ab} \underset{(1)}{S}{}_b{}^\delta \eta_\delta^{(++)} = 0. \tag{3.49}$$

Furthermore, since $\underset{(1)}{S}{}_a{}^\delta \xi_\delta^{(++)} = -\xi_\sigma^{(++)} \underset{(1)}{\widetilde{S}}{}^\delta{}_a$, Eq. (3.49) yields the consistent conditions

$$\underset{(1)}{\widetilde{S}}{}^\alpha{}_b \Delta^{bd} = \Gamma^{\alpha\beta} \underset{(2)}{S}{}_\beta{}^d, \tag{3.50a}$$

$$\underset{(2)}{\widetilde{S}}{}^d{}_\alpha \Gamma^{\alpha\beta} = \Delta^{db} \underset{(1)}{S}{}_b{}^\beta, \tag{3.50b}$$

it can be readily verified that the solutions to (3.50) are

$$\underset{(1)}{S}{}_b{}^\alpha = [(i\theta_\gamma \rho_\delta) A^{\gamma\delta}_{(b)}](\eta_\sigma^{(++)} \Gamma^{\sigma\alpha}), \tag{3.51a}$$

$$\underset{(2)}{S}{}_\beta{}^d = -\eta_\beta^{(++)} [(i\theta_\gamma \rho_\delta) A^{\gamma\delta}_{(a)}] \langle \mathbf{A}^{(a)} | \mathbf{A}^{(d)} \rangle. \tag{3.51b}$$

Hence

$$S_A{}^B = [(i\theta_\gamma\rho_\delta)A^{\gamma\delta}_{(a)}][\eta_\sigma^{(++)}\Gamma^{\sigma\beta}_{(1)} - \eta_\alpha^{(++)}\langle \mathbf{A}^{(a)}|\mathbf{A}^{(b)}\rangle]. \quad (3.52)$$

For the purpose of extracting the odd generators of the graded algebra from (3.52) and deriving their anticommutation relations, note that

$$\begin{aligned}\mathbf{S} &= \underset{(1)}{S}{}_a{}^\beta \mathbf{A}^{(a)} \otimes \mathbf{q}_\beta + \underset{(2)}{S}{}_\alpha{}^b \mathbf{p}^\alpha \otimes \mathbf{A}_{(b)} \\ &= [(i\theta_\gamma\rho_\delta)A^{\gamma\delta}_{(a)}]\eta_\beta^{(++)}[\mathbf{A}^{(a)} \otimes (\mathbf{p}^\beta \circ \mathbf{\Gamma}_{(1)}) \\ &\quad - \mathbf{p}^\beta \otimes (\mathbf{A}^{(a)} \cdot \mathbf{\Delta})]. \end{aligned} \quad (3.53)$$

It is evident from this last expression that

$$\mathbf{Q}^{(a\beta)} := [\mathbf{A}^{(a)} \otimes (\mathbf{p}^\beta \circ \mathbf{\Gamma}_{(1)}) - \mathbf{p}^\beta \otimes (\mathbf{A}^{(a)} \cdot \mathbf{\Delta})] \quad (3.54)$$

are the odd generators of the graded algebra that we needed.

In particular, for any two $\mathbf{Q}^{(a\beta)}$, $\mathbf{Q}^{(b\gamma)}$, we have

$$\begin{aligned}&\{\mathbf{Q}^{(a\beta)},\mathbf{Q}^{(b\gamma)}\} \\ &= -\mathbf{A}^{(a)} \otimes (\mathbf{A}^{(b)} \cdot \mathbf{\Delta})\Gamma^{\beta\gamma}_{(1)} - \mathbf{p}^\beta \otimes (\mathbf{p}^\gamma \circ \mathbf{\Gamma}_{(1)})\langle \mathbf{A}^{(a)}|\mathbf{A}^{(b)}\rangle \\ &\quad - \mathbf{A}^{(b)} \otimes (\mathbf{A}^{(a)} \cdot \mathbf{\Delta})\Gamma^{\beta\gamma}_{(1)} - \mathbf{p}^\gamma \otimes (\mathbf{p}^\beta \circ \mathbf{\Gamma}_{(1)})\langle \mathbf{A}^{(b)}|\mathbf{A}^{(a)}\rangle \\ &= -\Gamma^{\beta\gamma}_{(1)}(\mathbf{A}^{(a)} \wedge \mathbf{A}^{(b)}) \cdot \mathbf{\Delta} \\ &\quad - \delta^{ab}(\mathbf{p}^\beta \otimes \mathbf{p}^\gamma + \mathbf{p}^\gamma \otimes \mathbf{p}^\beta) \circ \mathbf{\Gamma}_{(1)}. \end{aligned} \quad (3.55)$$

Thus recalling [cf. Eq. (3.42)] that $(\mathbf{A}^{(a)} \wedge \mathbf{A}^{(b)}) \cdot \mathbf{\Delta}$ are the generators for so(8), while $(\mathbf{p}^\beta \otimes \mathbf{p}^\gamma + \mathbf{p}^\gamma \otimes \mathbf{p}^\beta) \circ \mathbf{\Gamma}_{(1)}$ [cf. (3.22a)] are the generators for sp(4), we see that the anticommutators (3.55) together with the commutators (3.22b) and (3.43) close the graded algebra for osp(4/8), and the vector space $(\mathscr{G}_e \otimes \mathscr{E}_8) \oplus T'_g$, with elements $\mathbf{Z} = \mathbf{z} + \boldsymbol{\xi}^{(++)}$ and inner product defined according to (3.44), is an appropriate representation space for OSp(4/8,$\mathbb{R}$).

Also we note that the representation spaces for the other OSp(4/N,$\mathbb{R}$), $N = 1,2,...,7$, can be easily obtained by taking our fermionic subspace the same as before and requiring that the bosonic basis vectors $\mathbf{A}_{(N+1)},...,\mathbf{A}_{(8)}$, as given in (3.37″), remain invariant under the transformations in SO(N).

Vectors in our bosonic subspaces will, therefore, be given by spinor bilinears of the form

$$\mathbf{w} = \frac{1}{2}\sum_{a=1}^{N}[(i\theta_\alpha\rho_\beta)A^{\alpha\beta}_{(a)}]\mathbf{A}^{(a)}, \quad (3.56)$$

which lie in the hyperplane normal to the vectors $\mathbf{A}^{(b)}$, $b = N+1,...,8$.

We also remark that for the specific case $N = 4$, the projected spinors $\boldsymbol{\xi}^{(++)} \in \mathscr{T}'_g$ can be used for the construction of both the fermionic and bosonic subspaces.

Before ending this section, we feel it is important to include some comments on a technical aspect of our formalism concerned with the additional structure which we had to impose on our original spinors by means of the symplectic product (3.8).

Recall that, although we started with odd Grassmann graded spinors associated with $\mathscr{E}_8$ and a fundamental spinor bilinear given by (2.14), which is invariant under Pin(8), this invariance is broken down to Pin(6) in (3.8), and the association with $\mathscr{E}_8$ as an underlying geometry for the theory is somewhat obscured.

On the other hand we were forced to introduce (3.8) in order to obtain a symplectic metric for our four-dimensional fermionic subspaces. As an end result we now have a spinor structure with Sp(16), Pin(6), and Sp(4) as isometry groups instead of the original ones given by Pin(8) and its subgroups.

Thus it can be argued that while our spinors originated from $\mathscr{E}_8$ together with the new inner product, (3.8) may still be viewed as a natural supporting geometry for OSp(4/N) with N ranging from 1 to 6, such an assumption would not be applicable for extended supersymmetries with $N = 7,8$.

In what follows we outline the basic steps of a procedure by means of which the above objections can easily be answered. Since the details of such an approach entail only minor essential modifications to what we have already developed here, we will be omitting them for the sake of brevity.

Instead of taking $\mathscr{E}_8$ as an underlying geometry for our spinor spaces, start with either of the pseudo-Euclidean spaces $\mathscr{E}_{8,2}$ [signature $(\underbrace{+ + \cdots +}_{8}, - -)$] or $\mathscr{E}_{9,1}$ [signature $(\underbrace{+ + \cdots +}_{9}, -)$].

Note next that for pseudo-Euclidean spaces with h negative signs in the signature, the Cartan conjugate of a spinor $\boldsymbol{\xi}$ is given by

$$\boldsymbol{\xi}^c = i^{(\nu-h)}(H_1 - H_{1'})\cdots(H_{(\nu-h)} - H_{(\nu-h)'})\boldsymbol{\xi}^*, \quad (3.57)$$

instead of (2.17). Also, since

$$(\boldsymbol{\xi}^c)^c = (-1)^{(\nu-h)(\nu-h+1)/2}\boldsymbol{\xi}$$

in these cases, it is clear that self-conjugate spinors may be associated with $\mathscr{E}_{8,2}$ and $\mathscr{E}_{9,1}$, and used for the construction of real subspaces as required in the beginning of this section.

Moreover, (2.15′) implies in addition that our real spinors have an already built-in symplectic metric for the original fundamental bilinear, so we do not need to introduce any further structure into the spinor formalism in order to obtain the required fermionic subspaces.

Specifically for $\mathscr{E}_{8,2}$, these fermionic subspaces result from projecting the odd Grassmann graded real spinors $\boldsymbol{\xi} + \boldsymbol{\xi}^c$, with 2^5 components, into four-dimensional subspaces obtained by successive application of the operators (not unique)

$$\mathscr{P}_1^{(\pm)} = \tfrac{1}{2}[E \pm (H_{1'}H_1 - H_1H_{1'})(H_{2'}H_2 - H_2H_{2'})], \quad (3.58a)$$

$$\mathscr{P}_2^{(\pm)} = \tfrac{1}{2}[E \pm (H_{4'}H_4 - H_4H_{4'})(H_{5'}H_5 - H_5H_{5'})], \quad (3.58b)$$

$$\mathscr{P}_3^{(\pm)} = \tfrac{1}{2}[E \pm (H_{1'}H_1 - H_1H_{1'})S]. \quad (3.58c)$$

Such a procedure allows us to write our original spinor space as a direct sum of eight four-dimensional isomorphic subspaces for which the fundamental bilinear form has *ab initio* a symplectic metric.

As a consequence, the groups of isometries of the fundamental bilinear spinor associated with $\mathscr{E}_{8,2}$ now include Sp(32), Pin(8,2), Spin(8,2), Pin(8), and Sp(4), which contain all the symmetry groups required for the construction of our fermionic subspaces as well as those needed for the bosonic ones with SO(N) internal symmetries, with N having values ranging from 1 to 8.

As far as the bosonic subspaces are concerned, all that we would require for their construction is to resort to Cartan's tensor decomposition that was described in detail previously, but now using pseudovectors of the form $(\boldsymbol{\theta}|H_j\boldsymbol{\rho})$, $j=1,1',...,5,5'$, as a natural ten-dimensional representation space for SO(8,2). Choosing the subspace of these vectors which is invariant under SO(8) results in the bosonic basis which we need for the construction of a representation space for OSp(4/8,$\mathbb{R}$). Further invariance requirements on this subspace similar to those used before, leads us down the ladder of the family of supersymmetries OSp(4/N), $N=1,...,8$.

We believe that the above description should suffice as a guide for those readers interested in rewording the details of our formalism for the cases $\mathscr{C}_{8,2}$ and $\mathscr{C}_{9,1}$. The considerably more complicated computations and the length of formulas that will arise provide the justification for having chosen for the more complete presentation of our ideas the simpler approach based on $\mathscr{C}_8$ (if one keeps in mind the remarks we have just made).

IV. CONCLUSIONS

We have shown that by extending Cartan's theory of spinors to a graded module for which the ring of operators consists of elements of an odd Grassmann subalgebra, one is able to construct a representation space for the orthosymplectic supergroup OSp(4/N,$\mathbb{R}$), where the fermionic subspace was made up of such spinors, while the bosonic vectors were taken from the space of irreducible tensors into which a product of two spinors decomposes under the action of the group of rotations and reversals.

By adopting this procedure it was possible to exhibit the relation between the structures of the fermionic and bosonic subspaces of the graded vector representation space, and their underlying geometry, which originates from a Euclidean space in eight dimensions.

The formalism, we believe, provides some additional insight for investigating the deeper levels of merging which might exist between fermionic and bosonic fields in supersymmetries theories, and the interrelationship between their structures and dimensionalities as determined by a supporting geometry.

For those of us that like to consider spinors as the fundamental builiding blocks from which vectors and tensors are constructed as composite quantities, it is pleasing to see that a procedure can be found by means of which this might be achieved for supersymmetry also, and that the Grassmann gradation of the bosonic subspace is a result of the tensor product of the two odd graded spinors that make up such fields.

APPENDIX A: GENERATORS FOR sp(4) AND G(8,0)⊗ℂ

In Sec. III of the text we obtained explicit tensor representations in terms of a spinor basis for the generators of sp(4). We also remarked there that the spinor operators which lead to the Clifford algebra $G(8,0)$ can be expressed in terms of the Lie algebra of sp(4). Here we present some complementary material on spinorial representations for the infinitesimal generators of the symplectic group in four dimensions, based on Pauli matrices, and also give without proof the formulas that exhibit the form of our spinor operators when written in terms both of Pauli matrices and of the generators of sp(4).

1. Generators for sp(4)

Making use of (3.22a) and (3.14), we get, in terms of right-handed Pauli matrices, the following expressions for the generators of sp(4):

$$\begin{aligned}
&B^{(11)}=\tfrac{1}{2}(I_2+\sigma_3)\otimes(\sigma_1+i\sigma_2),\\
&B^{(22)}=-\tfrac{1}{2}(I_2+\sigma_3)\otimes(\sigma_1-i\sigma_2),\\
&B^{(33)}=\tfrac{1}{2}(I_2-\sigma_3)\otimes(\sigma_1+i\sigma_2),\\
&B^{(44)}=-\tfrac{1}{2}(I_2-\sigma_3)\otimes(\sigma_1-i\sigma_2),\\
&B^{(12)}=\tfrac{1}{2}(I_2+\sigma_3)\otimes\sigma_3,\quad B^{(13)}=\tfrac{1}{2}\sigma_1\otimes(\sigma_1+i\sigma_2),\\
&B^{(14)}=-\tfrac{1}{2}(\sigma_1\otimes\sigma_3+i\sigma_2\otimes I_2),\\
&B^{(23)}=\tfrac{1}{2}(-\sigma_1\otimes\sigma_3+i\sigma_2\otimes I_2),\\
&B^{(24)}=\tfrac{1}{2}\sigma_1\otimes(\sigma_1-i\sigma_2),\\
&B^{(34)}=\tfrac{1}{2}(\sigma_3\otimes\sigma_3-I_2\otimes\sigma_3).
\end{aligned}\tag{A1}$$

In the above we use I_2 to denote the 2×2 identity matrix.

2. Generators for G(8,0)

From (2.10) and the spinor basis isomorphism introduced in Sec. III [just before Eq. (3.28)] it can be seen that matrix representations for H_i ($i=1,1',...,4,4'$) will be antidiagonal in two 8×8 dimensional blocks. By rather straightforward (albeit tedious) computations we find

$$\begin{aligned}
H_1&=\tfrac{1}{4}[i\sigma_2\otimes I_2-i\sigma_2\otimes\sigma_3]\otimes[\sigma_1\otimes I_2-i\sigma_2\otimes\sigma_3]-\tfrac{1}{4}[\sigma_2\otimes(I_2+\sigma_3)]\otimes I_4-\tfrac{1}{4}[\sigma_1\otimes(I_2+\sigma_3)]\otimes[\sigma_3\otimes\sigma_3]\\
&=-\tfrac{1}{2}[(B^{(23)}-B^{(14)})(I_4+B^{(12)}+B^{(34)})]\otimes[(B^{(12)}+B^{(34)})B^{(23)}]-\tfrac{1}{4}[(B^{(23)}-B^{(14)})(I_4-B^{(12)}-B^{(34)})]\\
&\quad\otimes I_4-\tfrac{1}{4}[(B^{(12)}+B^{(34)}-I_4)(B^{(14)}+B^{(23)})]\otimes[B^{(34)}-B^{(12)}],
\end{aligned}\tag{A2a}$$

$$\begin{aligned}
H_{1'}&=\tfrac{1}{4}[i\sigma_2\otimes I_2+i\sigma_2\otimes\sigma_3]\otimes I_4-\tfrac{1}{4}[\sigma_1\otimes I_2+i\sigma_1\otimes\sigma_3]\otimes[\sigma_3\otimes\sigma_3]+\tfrac{i}{4}[\sigma_2\otimes I_2-\sigma_2\otimes\sigma_3]\otimes[\sigma_1\otimes I_2+i\sigma_2\otimes\sigma_3]\\
&=\tfrac{1}{4}[(B^{(23)}-B^{(14)})(I_4-B^{(12)}-B^{(34)})]\otimes I_4-\tfrac{1}{4}[(B^{(23)}+B^{(14)})(I_4-B^{(12)}-B^{(34)})]\otimes[B^{(12)}-B^{(34)}]\\
&\quad+\tfrac{1}{2}[(B^{(23)}-B^{(14)})(I_4+B^{(12)}+B^{(34)})]\otimes[B^{(14)}(B^{(12)}+B^{(34)})],
\end{aligned}\tag{A2b}$$

$$H_2 = -\tfrac{1}{8}[(\sigma_1+i\sigma_2)\otimes(\sigma_1+i\sigma_2)][I_2\otimes\sigma_1 - iI_2\otimes\sigma_2] + \tfrac{1}{4}[i\sigma_1\otimes\sigma_2 - i\sigma_2\otimes\sigma_1][i\sigma_2\otimes\sigma_1 + i\sigma_2\otimes\sigma_2]$$
$$+\tfrac{1}{8}[(\sigma_1-i\sigma_2)\otimes(\sigma_1-i\sigma_2)]\otimes[\sigma_3\otimes\sigma_1 - i\sigma_3\otimes\sigma_2]$$
$$= -\tfrac{1}{2}[(I_4 - B^{(34)} + B^{(12)})(B^{(13)}+B^{(24)})]\otimes[(B^{(12)}-B^{(34)})B^{(24)}] + \tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})(B^{(13)}-B^{(24)})]$$
$$\otimes B^{(44)} + \tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})(B^{(13)}+B^{(24)})]\otimes B^{(22)}, \tag{A2c}$$
$$H_{2'} = \tfrac{1}{8}[(\sigma_1+i\sigma_2)\otimes(\sigma_1+i\sigma_2)]\otimes[\sigma_3\otimes\sigma_1 + i\sigma_3\otimes\sigma_2] + \tfrac{1}{4}[i\sigma_2\otimes\sigma_1 - i\sigma_1\otimes\sigma_2]\otimes[\sigma_2\otimes\sigma_2 - i\sigma_2\otimes\sigma_1]$$
$$-\tfrac{1}{8}[(\sigma_1-i\sigma_2)\otimes(\sigma_1-i\sigma_2)]\otimes[I_2\otimes\sigma_1 + iI_2\otimes\sigma_2]$$
$$= \tfrac{1}{2}[(I_4+B^{(12)}-B^{(34)})(B^{(13)}+B^{(24)})]\otimes[(B^{(34)}-B^{(12)})B^{(13)}]$$
$$+\tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})(B^{(13)}+B^{(24)})]\otimes B^{(11)} + \tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})(B^{(24)}-B^{(13)})]\otimes B^{(33)}, \tag{A2d}$$
$$H_3 = (i/4)[\sigma_1\otimes\sigma_1 - \sigma_2\otimes\sigma_2]\otimes[\sigma_2\otimes I_2] - \tfrac{1}{4}[\sigma_2\otimes\sigma_1 + \sigma_1\otimes\sigma_2]\otimes[\sigma_2\otimes\sigma_3]$$
$$+\tfrac{1}{4}[\sigma_1\otimes\sigma_1+\sigma_2\otimes\sigma_2]\otimes[\sigma_3\otimes\sigma_3] + (i/4)[\sigma_1\otimes\sigma_2 - \sigma_2\otimes\sigma_1]\otimes[\sigma_3\otimes I_2]$$
$$= \tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})B^{(13)}]\otimes[B^{(23)}-B^{(14)}]\otimes[(I_4-B^{(12)}-B^{(34)})]$$
$$+\tfrac{1}{4}[(I_4+B^{(12)}-B^{(34)})B^{(24)}]\otimes[(B^{(12)}-B^{(34)})(I_4+B^{(12)}+B^{(34)})]$$
$$+\tfrac{1}{4}[(I_4+B^{(12)}-B^{(34)})B^{(13)}]\otimes[(B^{(34)}-B^{(12)})(I_4-B^{(12)}-B^{(34)})]$$
$$-\tfrac{1}{4}[(I_4-B^{(12)}+B^{(34)})B^{(24)}]\otimes[(B^{(23)}-B^{(14)})(I_4+B^{(12)}+B^{(34)})], \tag{A2e}$$
$$H_{3'} = -(i/4)[\sigma_1\otimes\sigma_1 - \sigma_2\otimes\sigma_2]\otimes[\sigma_2\otimes I_2] - \tfrac{1}{4}[\sigma_2\otimes\sigma_1+\sigma_1\otimes\sigma_2]\otimes[\sigma_2\otimes\sigma_3]$$
$$+\tfrac{1}{4}[\sigma_1\otimes\sigma_1+\sigma_2\otimes\sigma_2]\otimes[\sigma_3\otimes\sigma_3] + \tfrac{1}{4}[i\sigma_2\otimes\sigma_1 - i\sigma_1\otimes\sigma_2]\oplus[\sigma_3\otimes I_2]$$
$$= -\tfrac{1}{4}[(I_4+B^{(34)}-B^{(12)})B^{(13)}]\otimes[(B^{(23)}-B^{(14)})(I_4-B^{(12)}-B^{(34)})]$$
$$-\tfrac{1}{4}[(I_4+B^{(12)}-B^{(34)})B^{(24)}]\otimes[(B^{(34)}-B^{(12)})(I_4-B^{(12)}-B^{(34)})]$$
$$-\tfrac{1}{4}[(I_4-B^{(34)}+B^{(12)})B^{(13)}]\otimes[(B^{(12)}-B^{(34)})(I_4+B^{(12)}+B^{(34)})]$$
$$+\tfrac{1}{4}[(I_4-B^{(12)}+B^{(34)})B^{(24)}]\otimes[(B^{(23)}-B^{(14)})(I_4-B^{(12)}-B^{(34)})], \tag{A2f}$$
$$H_4 = \tfrac{1}{4}[\sigma_1\otimes I_2 + \sigma_1\otimes\sigma_3]\otimes[\sigma_1\otimes I_2 + i\sigma_2\otimes\sigma_3] + \tfrac{1}{4}[\sigma_1\otimes I_2 - i\sigma_1\otimes\sigma_3]\otimes I_4 + \tfrac{1}{4}[i\sigma_2\otimes I_2 - i\sigma_2\otimes\sigma_3]\otimes[\sigma_3\otimes\sigma_3]$$
$$= \tfrac{1}{2}[(B^{(14)}+B^{(23)})(B^{(12)}+B^{(34)}-I_4)]\otimes[B^{(14)}(B^{(12)}+B^{(34)})] + \tfrac{1}{4}[(B^{(14)}+B^{(23)})(B^{(12)}+B^{(34)}+I_4)]$$
$$\otimes I_4 + \tfrac{1}{4}[(B^{(23)}-B^{(14)})(I_4+B^{(12)}+B^{(34)})]\otimes[(B^{(34)}-B^{(12)})], \tag{A2g}$$
$$H_{4'} = \tfrac{1}{4}[\sigma_1\otimes\sigma_2 + \sigma_1\otimes\sigma_3]\otimes[\sigma_2\otimes I_2 - i\sigma_2\otimes\sigma_3] + \tfrac{1}{4}[\sigma_1\otimes I_2 - \sigma_1\otimes\sigma_3]\otimes I_4 - \tfrac{1}{4}[i\sigma_2\otimes I_2 - i\sigma_2\otimes\sigma_3]\otimes[\sigma_3\otimes\sigma_3]$$
$$= \tfrac{1}{2}[(B^{(14)}+B^{(23)})(B^{(12)}+B^{(34)}-I_4)]\otimes[(B^{(12)}+B^{(34)})B^{(23)}] + \tfrac{1}{4}[(B^{(14)}+B^{(23)})(B^{(12)}+B^{(34)}+I_4)]\otimes I_4$$
$$-\tfrac{1}{4}[(B^{(23)}-B^{(14)})(I_4+B^{(12)}+B^{(34)})]\otimes[(B^{(34)}-B^{(12)})]. \tag{A2h}$$

As pointed out in the discussion following Eqs. (2.13) in Sec. II, the generators of $G(8,0)$ and $G^c = G(8,0)\otimes\mathbb{C}$ are obtained by taking linear combinations of the spinor operators H_i ($i = 1,1',2,2',...,4,4'$) given above. Since these generators are invertible, so are the basis elements of the Clifford algebra, and therefore they generate a multiplicative group $G^*(8,0)$. Also, since the Clifford image of $\mathscr{E}_8$, i_E: $\mathscr{E}_8 \to G(8,0)$, is stable under the twisted adjoint map,[10]

$$\mathrm{Ad}(A)H(\mathbf{x}) = \omega_{\mathscr{G}}(A)H(\mathbf{x})A^{-1}, \quad A\in G^*(8,0), \quad H(\mathbf{x})\in i_E(\mathscr{E}_8), \tag{A3}$$

where $\omega_{\mathscr{G}}$ denotes the degree involution of $G(8,0)$, the group $G^*(8,0)$ is isomorphic to the Clifford group $\Gamma(8)$ of $\mathscr{E}_8$.

Furthermore, $\mathrm{Pin}(8)$ is the subgroup of $\Gamma(8)$ consisting of the elements A which satisfy $\lambda_{\mathscr{G}}(A) = \pm 1$, where the homomorphism $\lambda_{\mathscr{G}}: \Gamma(8)\to\mathbb{R}^*$ (the multiplicative group of $\mathbb{R}$) is defined by

$$AS_{\mathscr{G}}\,\omega_{\mathscr{G}}(A) = \lambda_{\mathscr{G}}(A)e, \tag{A4}$$

and $S_{\mathscr{G}}$ is the involution map $S_{\mathscr{G}}$: $G(8,0)\to G(8,0)^{\mathrm{opp}}$ [$G(8,0)^{\mathrm{opp}}$ is the algebra opposite to $G(8,0)$], while e is the identity element of $G(8,0)$.

If we now note that the restriction of $\mathrm{Ad}(A)$ to i_E ($\mathscr{E}_8$) is an isometry, i.e., $(\tau_A\mathbf{x})\cdot(\tau_A\mathbf{x}) = \mathbf{x}\cdot\mathbf{x}$, where $\tau_A\mathbf{x}$ is the vector in $\mathscr{E}_8$ whose Clifford image is $\mathrm{Ad}(A)H(\mathbf{x})$, then it follows that there exists a surjective homomorphism, Φ: $\Gamma(8)\to\mathrm{O}(8)$, given by $\Phi_{\mathscr{G}}(A) = \tau_A$, and it also follows that the restriction Φ of $\Phi_{\mathscr{G}}$ to $\mathrm{Pin}(8)$, such that Φ: $\mathrm{Pin}(8)\to\mathrm{O}(8)$, is surjective.

Finally we observe that since $\mathrm{Spin}(8)$ is the subgroup of $\mathrm{Pin}(8)$ consisting of elements which satisfy $\omega_{\mathscr{G}}(A) = A$, we have

$$\mathrm{Spin}(8) = \mathrm{Pin}(8)\cap G^0(8,0), \tag{A5}$$

where $G^0(8,0)$ is the subspace of $G(8,0)$ defined by

$$G^0(8,0) = \ker(\omega_{\mathscr{G}} - 1). \tag{A6}$$

The elements of $\mathrm{Spin}(8)$ can be also characterized by the condition $\det \Phi(A) = 1$; i.e., if and only if $\Phi(A)$ is a proper isometry. Thus Φ restricts to the surjective homomorphism

$$\mathrm{Spin}(8) \xrightarrow{\psi} \mathrm{SO}(8). \tag{A7}$$

The above discussion sums up the exact sequence of groups generated by our spinor operators. In particular, we see that the generators of Spin(8) are obtained from the even degree elements of $G(8,0)$, which, when expressed in terms of (A2a)–(A2h), display the relation between the generators of isometries for the fermionic and bosonic subspaces that represent the orthosymplectic symmetries discussed in the text.

APPENDIX B: RELATION BETWEEN THE EXTERIOR AND MATRIX FORMULATIONS OF SPINOR ALGEBRA

In Sec. II of the text we have defined spinors associated with an underlying Euclidean of pseudo-Euclidean space of $n = 2\nu$ dimensions, $\mathscr{E}_n$, as elements of a module of complex p-vectors (with degree p ranging from zero to ν) formed from the basis vectors dual to one of the ν-dimensional isotropic subspaces into which $\mathscr{E}_n$ decomposes.

We thus had that a spinor $\xi \in \mathscr{S}'$ [cf. Eq. (2.7b)] was given by an expression of the form

$$\xi = \sum_{p=0}^{\nu} \sum_{<} \xi_{i_1,\ldots,i_p} \epsilon^{i_1} \wedge \cdots \wedge \epsilon^{i_p}, \tag{B1}$$

where $\{\epsilon^i\}$ is the dual basis isomorphic (via [2.5]) to the isotropic subspace $\mathscr{N}'_\nu$ introduced at the beginning of Sec. II, and the remainder of the notation is defined following Eq. (2.8) in that section. As pointed out there, the set of 2^ν complex coefficients $\xi_{i_1,\ldots,i_p}$ are the components of the spinor ξ and have the property of being totally antisymmetric in their simple indices.

Spinors in $\mathscr{S}'$ are transformed into other spinors in S' via an endomorphism induced by linear operators $H(\mathbf{x})$, where $\mathbf{x}$ is an arbitrary vector in $\mathscr{E}_n$, which act on spinors by means of the Clifford product defined in (2.10) and (2.11).

Note that the space of operators $H(\mathbf{x})$ is spanned by the generators $H_i \equiv H(\mathbf{e}_i)$, $H_{i'} \equiv H(\mathbf{e}'_i)$, $i = 1,\ldots,\nu$, which satisfy the Clifford algebra (2.13). Thus we can write

$$H(\mathbf{x}) = H(\mathbf{r} + \mathbf{r}') = x^i H_i + x'^i H_{i'}. \tag{B2}$$

Furthermore, since the contraction operation $\bar{g}(\xi, r)$ defined by (2.11) is essentially the interior product of ξ with an arbitrary vector $\mathbf{r}$ in the isotropic subspace $\mathscr{N}_\nu$ [cf. Eq. (2.2)] acting on the right, we can express (2.10) as

$$H(\mathbf{x})\xi = \xi \wedge \mathbf{r}' + \xi \llcorner \mathbf{r}, \tag{B3}$$

that is

$$H_i \xi = \xi \llcorner \mathbf{e}_i, \quad H'_i \xi = \xi \wedge \epsilon^i. \tag{B4}$$

Making use of (B4) and Eqs. (2.13) in the text, we can now readily show the relation between our intrinsic formulation of spinor algebra and the one used by Cartan and other authors which is expressed in the perhaps more familiar language of gamma matrices and spinors as one-column vectors.

Consider first a general n-dimensional pseudo-Euclidean space for which the fundamental form has $\not{h}$ positive and $\not{h}$ negative squares. We shall restrict our attention however to spaces of even dimensionality, i.e., $n = 2\nu$, since these are the type of spaces considered in this paper. [To account for spinors associated with odd-dimensional spaces one only has to include an additional direction orthogonal to $\mathscr{E}_n$ and add a further term in the Clifford product (B4) related to this extra dimension and involving the operator defined in (3.4).]

An orthonormal basis for $\mathscr{E}_n$ is $\mathbf{a}_1,\ldots,\mathbf{a}_\nu$, $\mathbf{b}_1,\ldots,\mathbf{b}_\nu$ with

$$\begin{aligned} &\mathbf{a}_i \cdot \mathbf{a}_j = \delta_{ij}, \quad \mathbf{a}_i \cdot \mathbf{b}_j = 0, \quad \text{for } i,j = 1,\ldots,\nu, \\ &\mathbf{b}_i \cdot \mathbf{b}_j = 0, \quad \text{for } i \neq j, \\ &\mathbf{b}_i \cdot \mathbf{b}_i = 1, \quad \text{for } i = 1,\ldots,\nu - h, \\ &\mathbf{b}_i \cdot \mathbf{b}_i = -1, \quad \text{for } i = \nu - h + 1,\ldots,\nu. \end{aligned} \tag{B5}$$

In order to construct the two isotropic subspaces, each of dimension ν, we complexify $\mathscr{E}_n$ and define a basis $\mathbf{e}_1$, $\mathbf{e}_\nu$, $\mathbf{e}'_1,\ldots,\mathbf{e}'_\nu$ for $\mathbb{C}\mathscr{E}_n$ by

$$\begin{aligned} &\mathbf{e}_i = \tfrac{1}{2}(\mathbf{a}_i + i\mathbf{b}_i), \quad \text{for } i = 1,\ldots,\nu - h, \\ &\mathbf{e}_i = \tfrac{1}{2}(\mathbf{a}_i + \mathbf{b}_i), \quad \text{for } i = \nu - h + 1,\ldots,\nu, \\ &\mathbf{e}'_i = \tfrac{1}{2}(\mathbf{a}_i - i\mathbf{b}_i), \quad \text{for } i = 1,\ldots,\nu - h, \\ &\mathbf{e}'_i = \tfrac{1}{2}(\mathbf{a}_i - \mathbf{b}_i), \quad \text{for } i = \nu - h + 1,\ldots,\nu. \end{aligned} \tag{B6}$$

Clearly the subspaces $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$ are isotropic and satisfy the orthogonality conditions (2.1) in the text. Moreover, making use of the isomorphism given by (2.5), we obtain the dual basis ϵ^i to be used in the calculation of the action of the generators $H_{i'}$ on a given spinor. Note that if we choose the coefficients of $\mathbf{e}_i$ and $\mathbf{e}'_i$, for $i = 1,\ldots,\nu - h$, to be complex conjugate to each other, and those of $\mathbf{e}_i$ and $\mathbf{e}'_i$, for $i = \nu - h + 1,\ldots,\nu$, to be real, then our pseudo-Euclidean space will become real even though the bases for the two isotropic subspaces are complex.

Also note that by virtue of the basic anticommutation relations (2.13), the elements $H(\mathbf{a}_i)$, $H(\mathbf{b}_i)$ are representation-free operators corresponding to the gamma matrices.

To be more explicit in establishing the notational relation between spinor theory expressed in the language of exterior algebra and its matrix formulation, we take as an example the spinors associated with the Minkowski space-time $\mathscr{E}_{3,1}$. In this case we have $\nu = 2$, and $\mathbf{a}_i \cdot \mathbf{a}_j = \delta_{ij}$, $\mathbf{a}_i \cdot \mathbf{b}_j = 0$, for $i,j = 1,2$, $\mathbf{b}_i \cdot \mathbf{b}_j = 0$, for $i \neq j$, and $\mathbf{b}_1 \cdot \mathbf{b}_1 = 1$, $\mathbf{b}_2 \cdot \mathbf{b}_2 = -1$. The bases for our isotropic subspaces are $\mathbf{e}_1 = \frac{1}{2}(\mathbf{a}_1 + i\mathbf{b}_1)$, $\mathbf{e}_2 = \frac{1}{2}(\mathbf{a}_2 + \mathbf{b}_2)$ and $\mathbf{e}'_1 = \frac{1}{2}(\mathbf{a}_1 - i\mathbf{b}_1)$, $\mathbf{e}'_2 = \frac{1}{2}(\mathbf{a}_2 - \mathbf{b}_2)$.

A spinor associated with $\mathscr{E}_{3,1}$ is of the form

$$\xi = \xi_0 + \xi_1 \epsilon^1 + \xi_2 \epsilon^2 + \xi_{12} \epsilon^1 \wedge \epsilon^2, \tag{B7}$$

and consequently we get from (B4)

$$H_i \xi = \xi_1 \delta_i^1 + \xi_2 \delta_i^2 + \xi_{12}(\delta_i^2 \epsilon^1 - \delta_i^1 \epsilon^2), \tag{B8a}$$

$$H_{i'} \xi = \xi_0 \epsilon^i + \xi_1 \epsilon^1 \wedge \epsilon^2 \delta_2^i - \xi_2 \epsilon^1 \wedge \epsilon^2 \delta_1^i. \tag{B8b}$$

If we now arrange the spinor coefficients into a one-column matrix such that those with an even number of simple indices come before those with compound indices with an odd number of simple indices, we can write

$$\xi = \begin{pmatrix} \xi_0 \\ \xi_{12} \\ \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \Phi \\ \Psi \end{pmatrix}. \tag{B9}$$

The two component spinor $\Phi = \binom{\xi_0}{\xi_{12}}$ is an example of what Cartan[5] calls a semispinor of the first type, while $\Psi = \binom{\xi_1}{\xi_2}$ is a case of semispinors of the second type. This convention amounts to considering spinors as elements of a vector space with basis $\{\mathbf{l}^\alpha\}$ obtained via the isomorphism $\mathbf{l}^1 \approx \mathbf{1}$, $\mathbf{l}^2 \approx \boldsymbol{\epsilon}^1 \wedge \boldsymbol{\epsilon}^2$, $\mathbf{l}^3 \approx \boldsymbol{\epsilon}^1$, $\mathbf{l}^4 \approx \boldsymbol{\epsilon}^2$, and introducing a dual basis $\{\mathbf{m}_\beta\}$, such that $\mathbf{m}_\beta \circ \mathbf{l}^\alpha = \mathbf{l}^\alpha \circ \mathbf{m}_\beta = \delta^\alpha_\beta$. Thus

$$\boldsymbol{\xi} = \xi_0 \mathbf{l}^1 + \xi_{12} \mathbf{l}^2 + \xi_1 \mathbf{l}^3 + \xi_2 \mathbf{l}^4. \tag{B10}$$

In terms of these bases, and making use of (B8a) and (B8b), the linear operators H_i and $H_{i'}$ may be expressed as tensors of the form

$$\mathbf{H}_1 = \mathbf{l}^1 \otimes \mathbf{m}_3 - \mathbf{l}^4 \otimes \mathbf{m}_2, \quad \mathbf{H}_2 = \mathbf{l}^1 \otimes \mathbf{m}_4 + \mathbf{l}^3 \otimes \mathbf{m}_2, \tag{B11}$$
$$\mathbf{H}_{1'} = -\mathbf{l}^2 \otimes \mathbf{m}_4 + \mathbf{l}^3 \otimes \mathbf{m}_1, \quad \mathbf{H}_{2'} = \mathbf{l}^2 \otimes \mathbf{m}_3 + \mathbf{l}^4 \otimes \mathbf{m}_1.$$

The matrix representation of (B11) follows immediately after noting that the generic form of these tensors is

$$\mathbf{H}_i = (H_i)_\alpha{}^\beta \mathbf{l}^\alpha \otimes \mathbf{m}_\beta, \quad \mathbf{H}_{i'} = (H_{i'})_\alpha{}^\beta \mathbf{l}^\alpha \otimes \mathbf{m}_\beta, \quad i = 1,2,$$

where

$$(H_i)_\alpha{}^\beta = \mathbf{m}_\alpha \circ H_i \circ \mathbf{l}^\beta, \quad (H_{i'})_\alpha{}^\beta = \mathbf{m}_\alpha \circ H_{i'} \circ \mathbf{l}^\beta, \tag{B12}$$

so that $(H_i)_\alpha{}^\beta$ is the matrix element corresponding to the αth row and βth column.

From (B11) and (B12) it is now an easy matter to derive the gamma matrices associated with $\mathscr{E}_{3,1}$, which are equivalent to our linear operators

$$H(\mathbf{a}_1) = H_1 + H_{1'}, \quad H(\mathbf{a}_2) = H_2 + H_{2'},$$
$$H(\mathbf{b}_1) = -i(H_1 - H_{1'}), \quad H(\mathbf{b}_2) = H_2 - H_{2'}.$$

We have that

$$[H(\mathbf{a}_1)_\alpha{}^\beta] = \gamma^1, \quad [H(\mathbf{a}_2)_\alpha{}^\beta] = \gamma^2,$$
$$[H(\mathbf{b}_1)_\alpha{}^\beta] = \gamma^3, \quad [H(\mathbf{b}_2)_\alpha{}^\beta] = \gamma^0.$$

Although the resulting representation is not one commonly used for the Dirac matrices, other representations follow by applying various similarity transformations. In particular, the similarity transformation induced by the matrix

$$S = \left[\begin{array}{cc|cc} 1 & -i & & \\ -i & 1 & \multicolumn{2}{c}{0} \\ \hline & & 1 & -i \\ \multicolumn{2}{c|}{0} & i & -1 \end{array}\right] \tag{B13}$$

is the most convenient one for relating our spinor components to the elementary Weyl spinors with dotted and undotted indices which occur in the spinorial form of the Dirac equation.[11] We thus get

$$S\xi = \left[\begin{array}{cc|cc} 1 & -i & & \\ -i & 1 & \multicolumn{2}{c}{0} \\ \hline & & 1 & -i \\ \multicolumn{2}{c|}{0} & i & 0 \end{array}\right] \begin{pmatrix} \xi_0 \\ \xi_{12} \\ \xi_1 \\ \xi_2 \end{pmatrix} = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \chi^{\dot{1}} \\ \chi^{\dot{2}} \end{bmatrix}, \tag{B14}$$

whence

$$\psi_1 = \xi_0 - i\xi_{12}, \quad \psi_2 = \xi_{12} - i\xi_0,$$
$$\chi^{\dot{1}} = \xi_1 - i\xi_2, \quad \chi^{\dot{2}} = -\xi_2 + i\xi_1. \tag{B15}$$

That is, the undotted spinor components of the Dirac bispinor are given by linear combinations of Cartan's semispinors of the first type, while the dotted components of the bispinor involve linear combinations of semispinors of the second type.

The above discussion suggests a general procedure to be followed in order to establish the relationship between our formalism and the one based on matrices for spinors asociated with other types of Euclidean and pseudo-Euclidean spaces of arbitrary dimensions. It also outlines the method by means of which the components of a Cartan spinor can be related to the spin-tensor components with mixed indices of the relativistic Dirac–Fierz–Pauli field equations for arbitrary spin values.

[1]See, e.g., P. van Niewenhuizen, Phys. Rep. **68**, 189 (1981); Y. Ne'eman, in *Cosmology and Gravitation: Spin, Torsion, Rotation and Supergravity*, edited by P. G. Bergmann and V. De Sabbata (Plenum, New York, 1980).

[2]C. P. Luehr and M. Rosenbaum, J. Math. Phys. **26**, 1828, 1834 (1985).

[3]T. Kugo and P. Townsend, Nucl. Phys. B **221**, 357 (1983); J. W. Moffat, J. Math Phys. **25**, 347 (1984); Z. Hasiewicz and K. Kwasniewski, *ibid.* **26**, 6 (1985).

[4]E. Kähler, Rend. Mat. (3–4) **21**, 425 (1962); W. Graf, Ann. Inst. H. Poincaré **29**, 85 (1978); F. Gürsey, Found. Phys. **13**, 289 (1983); I. M. Benn and R. W. Tucker, Commun. Math. Phys. **89**, 341 (1983); J. Phys. A **16**, 4147 (1983); Commun. Math. Phys. **98**, 53 (1985); J. A. Bullinaria, Phys. Rev. D **33**, 376 (1986).

[5]E. Cartan, *The Theory of Spinors* [(Hermann, Paris, 1966) and (Dover, New York, 1981)].

[6]E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci. **39**, 510 (1953).

[7]R. Coquercaux, University of Texas preprint No. UTTG-21-85.

[8]For a discussion of extended SU(N) conformal supergravity see S. Ferrara, M. Kaku, P. K. Townsend, and P. van Niewenhuizen, Nucl. Phys. B **129**, 125 (1977).

[9]A derivation of $N = 1$ Poincaré, de Sitter, and conformal superalgebras in all space-time dimensions allowing Majorana spinors can be found in J. W. van Holter and A. van Proeyen, J. Phys. A: Math. Gen. **15**, 3763 (1982).

[10]W. Greub, *Multilinear Algebra* (Springer, New York, 1978), 2nd ed.

[11]P. Roman, *Theory of Elementary Particles* (North-Holland, Amsterdam, 1961), 2nd ed.

The Schücking problem[a)]

István Ozsváth and Leland Sapiro
Programs in Mathematical Sciences, The University of Texas at Dallas, P.O. Box 830688, Richardson, Texas 75083-0688

(Received 17 March 1987; accepted for publication 13 May 1987)

The embedding problem for a three-parametric family of homogeneous three-spaces into a higher-dimensional Euclidean space is considered. These three-spaces occur as space sections in cosmological models. After general consideration a certain two-parametric family is embedded into a five-dimensional Euclidean space, deferring the solution of the general case to later papers.

I. INTRODUCTION

A number of very interesting world models of the relativistic cosmology have compact spacelike sections invariant under the transformations of the rotation group. Schücking suggested the embedding of those three-spaces into a higher-dimensional Euclidean space, hoping to obtain a more adequate picture of their geometry.[1] We would like to develop at the outset some language in order to formulate our problem more precisely.

Denote by

$$\xi^1, \xi^2, \xi^3, \xi^4, \tag{1.1}$$

the Cartesian coordinates in the four-dimensional Euclidean space E^4. The equation of the unit sphere S^3 is then

$$(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2 + (\xi^4)^2 = 1. \tag{1.2}$$

By the assignment

$$\xi = \xi^1 + i\xi^2, \quad \eta = \xi^3 + i\xi^4 \tag{1.3}$$

and by forming the unitary matrix

$$A = \begin{pmatrix} \xi & \eta \\ -\bar{\eta} & \bar{\xi} \end{pmatrix} \tag{1.4}$$

we establish a one-to-one correspondence between the points of S^3 and the elements of SU_2,

$$\xi\bar{\xi} + \eta\bar{\eta} = 1 \tag{1.5}$$

being the equation of the unit sphere.

If P_1 and P_2 are points of S^3 and A_1 and A_2 are the corresponding matrices then the matrix multiplication

$$A_1 A_2 = A_3$$

and the correspondence

$$A_3 \to P_3$$

turn S^3 into a Lie group—the universal covering group of O_3, the group of rotations of E^3 around a fixed point.

We introduce the Eulerian angles

$$0 \leqslant x^1 = x \leqslant \pi, \quad 0 \leqslant x^2 = y, \quad x^3 = z \leqslant 2\pi \tag{1.6}$$

as local coordinates by the assignment

$$A = \begin{pmatrix} \xi^1 + i\xi^2 & \xi^3 + i\xi^4 \\ -\xi^3 + i\xi^4 & \xi^1 - i\xi^2 \end{pmatrix} = \begin{pmatrix} \cos(x/2)e^{[i(y+z)/2]} & i\sin(x/2)e^{[i(y-z)/2]} \\ i\sin(x/2)e^{[-i(y-z)/2]} & \cos(x/2)e^{[-i(y+z)/2]} \end{pmatrix}. \tag{1.7}$$

Following Flanders[2] we compute the matrix of the forms

$$\omega = A^{-1}\,dA = \begin{pmatrix} (i/2)(\cos x\,dy + dz) & (i/2)e^{-iz}\,dx - \tfrac{1}{2}\sin x e^{-iz}\,dy \\ (i/2)e^{iz}\,dx + \tfrac{1}{2}\sin x e^{iz}\,dy & -(i/2)(\cos x\,dy + dz) \end{pmatrix} \tag{1.8}$$

each of which is a left invariant one-form of the matrix group defined by (1.7). That is by appropriate numbering we have

$$\begin{aligned} \omega^1 &= \cos z\,dx + \sin x \sin z\,dy, \\ \omega^2 &= -\sin z\,dx + \sin x \cos z\,dy, \\ \omega^3 &= \cos x\,dy + dz, \end{aligned} \tag{1.9}$$

as our invariant one-forms in our coordinate system.

It is easy to see that

$$d\omega^1 = -\omega^2 \wedge \omega^3,$$

$$d\omega^2 = -\omega^3 \wedge \omega^1, \quad d\omega^3 = -\omega^1 \wedge \omega^2. \tag{1.10}$$

The line element on S^3 in this coordinate system is given by

$$\begin{aligned} (ds)^2 &= \tfrac{1}{4}[(\omega^1)^2 + (\omega^2)^2 + (\omega^3)^2] \\ &= \tfrac{1}{4}[(dx)^2 + (dy)^2 + 2\cos x\,dy\,dz + (dz)^2]. \end{aligned} \tag{1.11}$$

If we interpret the assignment (1.7) as

$$\begin{aligned} \xi^1 &= Z^1(x,y,z) = \cos(x/2)\cos[(y+z)/2], \\ \xi^2 &= Z^2(x,y,z) = \cos(x/2)\sin[(y+z)/2], \\ \xi^3 &= Z^3(x,y,z) = -\sin(x/2)\sin[(y-z)/2], \\ \xi^4 &= Z^4(x,y,z) = \sin(x/2)\cos[(y-z)/2], \end{aligned} \tag{1.12}$$

[a)] We dedicate this paper to Professor E. L. Schücking on the occasion of his 60th birthday.

 0022-2488/87/092066-08$02.50

then we have a parametric equation of S^3, or an embedding of it into E^4.

Indeed, as a straightforward calculation shows

$$(ds)^2 = (d\xi^1)^2 + (d\xi^2)^2 + (d\xi^3)^2 + (d\xi^4)^2 = \tfrac{1}{4}[(dx)^2 + (dy)^2 + 2\cos x\, dy\, dz + (dz)^2], \tag{1.13}$$

which is, naturally, not surprising. We are now ready to formulate our problem.

We would like to consider the most general, positive definite metric, invariant under the left translations of the group of rotations. These metrics form a three-parametric family

$$(ds)^2 = p^2(\omega^1)^2 + q^2(\omega^2)^2 + r^2(\omega^3)^2, \tag{1.14}$$

where

$$p, q, r \tag{1.15}$$

are arbitrary parameters. (For example, $p = q = r = 1$ corresponds to a sphere of radius 2.) Denoting the Cartesian coordinates in E^6 by

$$\xi^\alpha, \quad \alpha = 1, 2, 3, 4, 5, 6, \tag{1.16}$$

we are looking for six functions

$$\xi^\alpha = Z^\alpha(x,y,z), \quad \alpha = 1, 2, 3, 4, 5, 6, \tag{1.17}$$

such that

$$\begin{aligned}(ds)^2 &= \sum_{\alpha=1}^{6} (d\xi^\alpha)^2 = \delta_{\alpha\beta} Z^\alpha_{,i} Z^\beta_{,j}\, dx^i\, dx^j \\ &= p^2(\omega^1)^2 + q^2(\omega^2)^2 + r^2(\omega^3)^2 \\ &= \{p^2(\cos z)^2 + q^2(\sin z)^2\}(dx)^2 \\ &\quad + (p^2 - q^2)\sin x \sin 2z\, dx\, dy \\ &\quad + \{[p^2(\sin z)^2 + q^2(\cos z)^2] \\ &\quad \times (\sin x)^2 + r^2(\cos x)^2\}(dy)^2 \\ &\quad + 2r^2 \cos x\, dy\, dz + r^2(dz)^2 \\ &= g_{ij}\, dx^i\, dx^j,\end{aligned} \tag{1.18}$$

where the numbering of the coordinates is as introduced in (1.6) and

$$Z^\alpha_{,i} = \frac{\partial Z^\alpha}{\partial x^i}. \tag{1.19}$$

We turn to this problem after giving the left invariant vector fields in our coordinate system for future reference,

$$\begin{aligned} X_1 &= \cos z \frac{\partial}{\partial x} + \frac{\sin z}{\sin x}\frac{\partial}{\partial y} - \cot x \sin z \frac{\partial}{\partial z}, \\ X_2 &= -\sin z \frac{\partial}{\partial x} + \frac{\cos z}{\sin x}\frac{\partial}{\partial y} - \cot x \cos z \frac{\partial}{\partial z}, \\ X_3 &= \frac{\partial}{\partial z}, \end{aligned} \tag{1.20}$$

with

$$[X_2, X_3] = X_1, \quad [X_3, X_1] = X_2, \quad [X_1, X_2] = X_3. \tag{1.21}$$

II. FUNDAMENTAL EQUATIONS

We expect that, in general, E^6 would be the ambient space. Therefore we have to find six functions

$$\begin{aligned}&\xi^\alpha = Z^\alpha(x^i), \\ &\alpha, \beta = 1,2,...,6; \; i, j = 1, 2, 3\end{aligned} \tag{2.1}$$

satisfying the six nonlinear partial differential equations

$$\delta_{\alpha\beta} Z^\alpha_{,i} Z^\beta_{,j} = g_{ij}, \tag{2.2}$$

where g_{ij} is given by (1.18).

The relevant part of the classical differential geometry—developed by the great masters, is readily available, for instance, in Eisenhart's book[3] or in the beautiful books of Spivak[4]—instructs us to introduce vector fields

$$\eta^\alpha{}_A, \quad A = 4, 5, 6, \tag{2.3}$$

such that

$$\delta_{\alpha\beta} Z^\alpha_{,i} \eta^\alpha{}_A = 0, \quad A = 4, 5, 6 \tag{2.4}$$

and

$$\delta_{\alpha\beta} \eta^\alpha{}_A \eta^\beta{}_B = \delta_{AB}, \quad A, B = 4, 5, 6, \tag{2.5}$$

and integrate the *linear* system

$$\begin{aligned} Z^\alpha_{;ij} &= b^A_{ij} \eta^\alpha{}_A, \\ \eta^\alpha_{A,j} &= -b_{Alj} g^{lm} Z^\alpha_{,m} + \epsilon_A{}^{BC} \mu_{Bj} \eta^\alpha{}_C, \end{aligned} \tag{2.6}$$

where

$$\mu_{Aj} = \tfrac{1}{2} \epsilon_A{}^{BC} V_{BCj} \tag{2.7}$$

and V_{BCj} are defined as in Eisenhart.[3] The semicolon denotes covariant differentiation with respect to the metric (1.18);

$$b_{Aij} \tag{2.8}$$

are symmetric tensors and

$$\mu_{Ai} \tag{2.9}$$

are vectors on our three-space. The integrability conditions of (2.6) are

$$\begin{aligned} R_{ijkl} &= b^A{}_{ik} b_{Ajl} - b^A{}_{il} b_{Ajk}, \\ b_{Aij;k} - b_{Aik;j} &= \epsilon_A{}^{BC}(\mu_{Bk} b_{Cij} - \mu_{Bj} b_{Cik}), \\ \mu_{Aj,k} - \mu_{Ak,j} &+ \epsilon_A{}^{BC}(\mu_{Bj}\mu_{Ck} + b_{Blj} b_{Cmk}\, g^{lm}) = 0, \end{aligned} \tag{2.10}$$

being the Gauss, Codazzi–Mainardi, and Ricci equations, respectively. R_{ijkl} are the components of the Riemann tensor of our three-space. In order to find the as yet unknown coefficients b_{Aij} and μ_{Aj} in (2.6) we have to find a solution of (2.10) by the given R_{ijkl}. A formidable task. The fact that our three-space is homogeneous, that its metric is invariant under the left translations of our group, comes to our rescue. The metric tensor is invariant, therefore the Riemann tensor is invariant. We now *assume* that the tensors b and the vectors μ are also invariant. We make this assumption, and explore its consequences, rather than to attempt to construct a proof of this invariance.

To exploit the invariance of our tensors and vectors we use the invariant one-forms (1.9) and the invariant vector fields (1.20) in order to span the tensor fields over our manifold, since the invariant tensor fields have *constant* coefficients with respect to the tensor products of the invariant one-forms and vector fields.

We use

$$g_{ab},\ g^{ab},\ R_{abcd},\ b_{Aab},\ \mu_{Aa} \tag{2.11}$$

to denote the "frame components" of our tensor fields, adhering to the convention that letters from the beginning of the alphabet

$$a,\ b,\ c=1,\ 2,\ 3 \tag{2.12}$$

are "frame indices," and letters like

$$i,\ j,\ k=1,\ 2,\ 3 \tag{2.13}$$

would denote coordinate indices.

We shall see that our assumption reduces Eqs. (2.10) to *algebraic equations*.

III. GAUSS, CODAZZI–MAINARDI, AND RICCI EQUATIONS

We now develop the Gauss, Codazzi–Mainardi, and Ricci equations in terms of our frame. Using the invariant vector fields (1.20) we introduce the Koszul connection

$$\nabla_{X_a} X_b = \Gamma_{ab}{}^{f} X_f \tag{3.1}$$

in order to carry out covariant differentiation. The frame components of the covariant derivative of a tensor b_{ab} and a vector μ_a are given by the expressions

$$X_c b_{ab} - b_{fb}\Gamma_{ca}{}^{f} - b_{af}\Gamma_{cb}{}^{f} \tag{3.2}$$

and

$$X_c \mu_a - \mu_f \Gamma_{ca}{}^{f}, \tag{3.3}$$

respectively, which reduce to

$$- b_{fb}\Gamma_{ca}{}^{f} - b_{af}\Gamma_{cb}{}^{f} \tag{3.4}$$

and

$$- \mu_f \Gamma_{ca}{}^{f}, \tag{3.5}$$

respectively, in case of invariant tensor and vector fields. The requirement, that the connection is torsion-free and metric, that is, that the torsion tensor and the covariant derivative of the metric vanish, leads to

$$\Gamma_{abc} = \tfrac{1}{2}(C_{bca} + C_{cab} - C_{abc}), \tag{3.6}$$

where the C's are the structure constants of the group

$$[X_a, X_b] = C^{f}{}_{ab} X_f \tag{3.7}$$

and the raising and lowering of the indices are carried out with the help of the frame components of the metric tensor.

The components of the Riemann tensor are given by

$$R^{a}{}_{bcd} = X_c \Gamma_{db}{}^{a} - X_d \Gamma_{cb}{}^{a} + \Gamma_{cf}{}^{a}\Gamma_{db}{}^{f} - \Gamma_{df}{}^{a}\Gamma_{cb}{}^{f} - \Gamma_{fb}{}^{a} C^{f}{}_{cd}. \tag{3.8}$$

These, however, since the Γ's are constant, reduce to

$$R_{abcd} = \Gamma_{cfa}\Gamma_{db}{}^{f} - \Gamma_{dfa}\Gamma_{cb}{}^{f} - \Gamma_{fba} C^{f}{}_{cd}. \tag{3.9}$$

The Ricci tensor components and the Ricci scalar are given by

$$R_{ab} = \Gamma_{fa}{}^{g}\Gamma_{gb}{}^{f} - \Gamma_{fg}{}^{f}\Gamma_{ab}{}^{g} \tag{3.10}$$

and

$$R = R^{f}{}_{f}, \tag{3.11}$$

respectively.

Equations (2.10) take the form

$$R_{abcd} = b^{A}{}_{ac} b_{Abd} - b^{A}{}_{ad} b_{Abc}, \tag{3.12a}$$

$$X_c b_{Aab} - X_b b_{Aac} + b_{Aaf} C^{f}{}_{bc} - b_{Afb}\Gamma_{ca}{}^{f} + b_{Afc}\Gamma_{ba}{}^{f} = \epsilon_A{}^{BC}(\mu_{Bc} b_{Cab} - \mu_{Bb} b_{Cac}), \tag{3.12b}$$

$$X_c \mu_{Ab} - X_b \mu_{Ac} + \mu_{Af} C^{f}{}_{bc} = -\epsilon_A{}^{BC}\{\mu_{Bb}\mu_{Cc} + b_{Bfb} b_{Cgc} g^{fg}\},\quad A = 4, 5, 6, \tag{3.12c}$$

and reduce to algebraic equations in case of invariant tensors, having constant components. We want to write out these equations in full detail for the case of invariant b and μ. In our case the nonvanishing components of the structure constant tensor are

$$C^{1}{}_{23} = C^{2}{}_{31} = C^{3}{}_{12} = 1 \tag{3.13}$$

and the components of the metric are

$$g_{ab} = \mathrm{diag}(p^2 \quad q^2 \quad r^2) \tag{3.14}$$

and

$$g^{ab} = \mathrm{diag}(1/p^2 \quad 1/q^2 \quad 1/r^2) \tag{3.15}$$

as we noticed earlier. As a consequence of all this the components of the Koszul connection are given by

$$\begin{aligned}
\Gamma_{231} &= \tfrac{1}{2}(p^2 - q^2 + r^2),\\
\Gamma_{321} &= -\tfrac{1}{2}(p^2 + q^2 - r^2),\\
\Gamma_{312} &= \tfrac{1}{2}(p^2 + q^2 - r^2),\\
\Gamma_{132} &= -\tfrac{1}{2}(-p^2 + q^2 + r^2),\\
\Gamma_{123} &= \tfrac{1}{2}(-p^2 + q^2 + r^2),\\
\Gamma_{213} &= -\tfrac{1}{2}(p^2 - q^2 + r^2),
\end{aligned} \tag{3.16}$$

and

$$\begin{aligned}
\Gamma_{23}{}^{1} &= (1/2p^2)(p^2 - q^2 + r^2),\\
\Gamma_{32}{}^{1} &= -(1/2p^2)(p^2 + q^2 - r^2),\\
\Gamma_{31}{}^{2} &= (1/2q^2)(p^2 + q^2 - r^2),\\
\Gamma_{13}{}^{2} &= -(1/2q^2)(-p^2 + q^2 + r^2),\\
\Gamma_{12}{}^{3} &= (1/2r^2)(-p^2 + q^2 + r^2),\\
\Gamma_{21}{}^{3} &= -(1/2r^2)(p^2 - q^2 + r^2).
\end{aligned} \tag{3.17}$$

The nonvanishing components of the Riemann tensor, Ricci tensor, and Ricci scalar are in turn,

$$\begin{aligned}
R_{2323} &= (1/4p^2)\{2p^2(-p^2 + q^2 + r^2) - (p^2 - q^2 + r^2)(p^2 + q^2 - r^2)\},\\
R_{3131} &= (1/4q^2)\{2q^2(p^2 - q^2 + r^2) - (p^2 + q^2 - r^2)(-p^2 + q^2 + r^2)\},\\
R_{1212} &= (1/4r^2)\{2r^2(p^2 + q^2 - r^2) - (-p^2 + q^2 + r^2)(p^2 - q^2 + r^2)\},
\end{aligned} \tag{3.18}$$

$$\begin{aligned}
R_{11} &= -(1/2q^2r^2)(p^2 - q^2 + r^2)(p^2 + q^2 - r^2),\\
R_{22} &= -(1/2r^2p^2) \times (p^2 + q^2 - r^2)(-p^2 + q^2 + r^2),\\
R_{33} &= -(1/2p^2q^2)(-p^2 + q^2 + r^2)(p^2 - q^2 + r^2),
\end{aligned} \tag{3.19}$$

$$
\begin{aligned}
R &= (1/2p^2q^2r^2) \\
&\quad \times (p^4+q^4+r^4-2q^2r^2-2r^2p^2-2p^2q^2) \\
&= (1/2p^2q^2r^2)(2(p^4+q^4+r^4)-(p^2+q^2+r^2)^2) \\
&= -(1/2p^2q^2r^2)(p+q+r)(-p+q+r) \\
&\quad \times (p-q+r)(p+q-r),
\end{aligned} \tag{3.20}
$$

respectively.

Introducing the notation

$$
(b_{4ab}) = \begin{bmatrix} A & F & E \\ F & B & D \\ E & D & C \end{bmatrix}, \quad (\mu_{4a}) = (X\ Y\ Z), \tag{3.21}
$$

$$
(b_{5ab}) = \begin{bmatrix} a & f & e \\ f & b & d \\ e & d & c \end{bmatrix}, \quad (\mu_{5a}) = (x\ y\ z), \tag{3.22}
$$

$$
(b_{6ab}) = \begin{bmatrix} \alpha & \varphi & \epsilon \\ \varphi & \beta & \delta \\ \epsilon & \delta & \gamma \end{bmatrix}, \quad (\mu_{6a}) = (\xi\ \eta\ \zeta), \tag{3.23}
$$

we can give Eqs. (3.12) in full detail,

$$
\begin{aligned}
R_{2323} &= BC-D^2+bc-d^2+\beta\gamma-\delta^2 \\
&= (1/4p^2)[2p^2(-p^2+q^2+r^2) \\
&\quad -(p^2-q^2+r^2)(p^2+q^2-r^2)], \\
R_{3131} &= AC-E^2+ac-e^2+\alpha\gamma-\epsilon^2 \\
&= (1/4q^2)[2q^2(p^2-q^2+r^2) \\
&\quad -(p^2+q^2-r^2)(-p^2+q^2+r^2)], \\
R_{1212} &= AB-F^2+ab-f^2+\alpha\beta-\varphi^2 \\
&= (1/4r^2)[2r^2(p^2+q^2-r^2) \\
&\quad -(-p^2+q^2+r^2)(p^2-q^2+r^2)], \\
R_{2331} &= DE-CF+de-cf+\delta\epsilon-\gamma\varphi=0, \\
R_{2312} &= DF-BE+df-be+\delta\varphi-\beta\epsilon=0, \\
R_{3112} &= EF-AD+ef-ad+\epsilon\varphi-\alpha\delta=0;
\end{aligned} \tag{3.24}
$$

$$
\begin{aligned}
&A-(1/2q^2)(p^2+q^2-r^2)B-(1/2r^2)(p^2-q^2+r^2)C \\
&\quad = (z\varphi-y\epsilon)-(\zeta f-\eta e), \\
&(1/2p^2)(3p^2+q^2-r^2)F = (z\beta-y\delta)-(\zeta b-\eta d), \\
&(1/2p^2)(3p^2-q^2+r^2)E = (z\delta-y\gamma)-(\zeta d-\eta c), \\
&\quad \vdots \\
&(1/2q^2)(p^2+3q^2-r^2)F = (x\epsilon-z\alpha)-(\xi e-\zeta a), \\
&-(1/2p^2)(p^2+q^2-r^2)A+B \\
&\quad -(1/2r^2)(-p^2+q^2+r^2)C \\
&\quad = (x\delta-z\varphi)-(\xi d-\zeta f), \\
&(1/2q^2)(-p^2+3q^2+r^2)D = (x\gamma-z\epsilon)-(\xi c-\zeta e), \\
&\quad \vdots \\
&(1/2r^2)(p^2-q^2+3r^2)E = (y\alpha-x\varphi)-(\eta a-\xi f), \\
&(1/2r^2)(-p^2+q^2+3r^2)D = (y\varphi-x\beta)-(\eta f-\xi b), \\
&-(1/2p^2)(p^2-q^2+r^2)A \\
&\quad -(1/2q^2)(-p^2+q^2+r^2)B+C \\
&\quad = (y\epsilon-x\delta)-(\eta e-\xi d);
\end{aligned} \tag{3.25}
$$

$$
\begin{aligned}
&a-(1/2q^2)(p^2+q^2-r^2)b-(1/2r^2)(p^2-q^2+r^2)c \\
&\quad = (\zeta F-\eta E)-(Z\varphi-Y\epsilon), \\
&(1/2p^2)(3p^2+q^2-r^2)f = (\zeta B-\eta D)-(Z\beta-Y\delta), \\
&(1/2p^2)(3p^2-q^2+r^2)e = (\zeta D-\eta C)-(Z\delta-Y\gamma), \\
&\quad \vdots \\
&(1/2q^2)(p^2+3q^2-r^2)f = (\xi E-\zeta A)-(X\epsilon-Z\alpha), \\
&-(1/2p^2)(p^2+q^2-r^2)a+b \\
&\quad -(1/2r^2)(-p^2+q^2+r^2)c \\
&\quad = (\xi D-\zeta F)-(X\delta-Z\varphi), \\
&(1/2q^2)(-p^2+3q^2+r^2)d = (\xi C-\zeta E)-(X\gamma-Z\epsilon), \\
&\quad \vdots \\
&(1/2r^2)(p^2-q^2+3r^2)e = (\eta A-\xi F)-(Y\alpha-X\varphi), \\
&(1/2r^2)(-p^2+q^2+3r^2)d = (\eta F-\xi B)-(Y\varphi-X\beta), \\
&-(1/2p^2)(p^2-q^2+r^2)a \\
&\quad -(1/2q^2)(-p^2+q^2+r^2)b+c \\
&\quad = (\eta E-\xi D)-(Y\epsilon-X\delta);
\end{aligned} \tag{3.26}
$$

$$
\begin{aligned}
&\alpha-(1/2q^2)(p^2+q^2-r^2)\beta-(1/2r^2)(p^2-q^2+r^2)\gamma \\
&\quad = (Zf-Ye)-(zF-yE), \\
&(1/2p^2)(3p^2+q^2-r^2)\varphi = (Zb-Yd)-(zB-yD), \\
&(1/2p^2)(3p^2-q^2+r^2)\epsilon = (Zd-Yc)-(zD-yC), \\
&\quad \vdots \\
&(1/2q^2)(p^2+3q^2-r^2)\varphi = (Xe-Za)-(xE-zA), \\
&-(1/2p^2)(p^2+q^2-r^2)\alpha+\beta \\
&\quad -(1/2r^2)(-p^2+q^2+r^2)\gamma \\
&\quad = (Xd-Zf)-(xD-zF), \\
&(1/2q^2)(-p^2+3q^2+r^2)\delta = (Xc-Ze)-(xC-zE), \\
&\quad \vdots \\
&(1/2r^2)(p^2-q^2+3r^2)\epsilon = (Ya-Xf)-(yA-xF), \\
&(1/2r^2)(-p^2+q^2+3r^2)\delta = (Yf-Xb)-(yF-xB), \\
&-(1/2p^2)(p^2-q^2+r^2)\alpha \\
&\quad -(1/2q^2)(-p^2+q^2+r^2)\beta+\gamma \\
&\quad = (Ye-Xd)-(yE-xD);
\end{aligned} \tag{3.27}
$$

$$
\begin{aligned}
X &= -(y\zeta-z\eta)-(1/p^2)(f\epsilon-e\varphi) \\
&\quad -(1/q^2)(b\delta-d\beta)-(1/r^2)(d\gamma-c\delta), \\
Y &= -(z\xi-x\zeta)-(1/p^2)(e\alpha-a\epsilon) \\
&\quad -(1/q^2)(d\varphi-f\delta)-(1/r^2)(c\epsilon-e\gamma), \\
Z &= -(x\eta-y\xi)-(1/p^2)(a\varphi-f\alpha) \\
&\quad -(1/q^2)(f\beta-b\varphi)-(1/r^2)(e\delta-d\epsilon), \\
&\quad \vdots \\
x &= -(\eta Z-\zeta Y)-(1/p^2)(\varphi E-\epsilon F) \\
&\quad -(1/q^2)(\beta D-\delta B)-(1/r^2)(\delta C-\gamma D),
\end{aligned}
$$

$$y = -(\zeta X - \xi Z) - (1/p^2)(\epsilon A - \alpha E) - (1/q^2)(\delta F - \varphi D) - (1/r^2)(\gamma E - \epsilon C), \quad (3.28)$$

$$z = -(\xi Y - \eta X) - (1/p^2)(\alpha F - \varphi A) - (1/q^2)(\varphi B - \beta F) - (1/r^2)(\epsilon D - \delta E),$$

$$\vdots$$

$$\xi = -(Yz - Zy) - (1/p^2)(Fe - Ef) - (1/q^2)(Bd - Db) - (1/r^2)(Dc - Cd),$$

$$\eta = -(Zx - Xz) - (1/p^2)(Ea - Ae) - (1/q^2)(Df - Fd) - (1/r^2)(Ce - Ec),$$

$$\zeta = -(Xy - Yx) - (1/p^2)(Af - Fa) - (1/q^2)(Fb - Bf) - (1/r^2)(Ed - De).$$

A solution of this formidable system furnishes the frame components b_{Aab} and μ_{Aa}.

IV. EXPLICIT FORM OF THE BASIC EQUATIONS

As we have seen in the previous section, due to the homogeneity of our three-space, Eq. (2.10) reduced to algebraic equations, and we have to integrate only the *linear* partial differential equations (2.6).

We wanted to rewrite these equations in terms of our frame and develop them more fully.

Using the notation

$$Z^\alpha{}_a = X_a Z^\alpha \quad (4.1)$$

(applying the vector fields X_a to the functions Z^α) we can rewrite those equations as

$$X_a Z^\alpha{}_b = \Gamma_{ab}{}^f Z^\alpha{}_f + b^A{}_{ab}\eta^\alpha{}_A, \qquad X_a\eta^\alpha{}_A = -b_{Aa}{}^f Z^\alpha{}_f + \epsilon_A{}^{BC}\mu_{Ba}\eta^\alpha{}_C, \quad (4.2)$$

where

$$b_{Aa}{}^b = b_{Aaf}\, g^{bf}. \quad (4.3)$$

Using (3.17) we have

$$X_1 Z^\alpha{}_1 = b^A{}_{11}\eta^\alpha{}_A,$$
$$X_1 Z^\alpha{}_2 = (1/2r^2)(-p^2+q^2+r^2)Z^\alpha{}_3 + b^A{}_{12}\eta^\alpha{}_A,$$
$$X_1 Z^\alpha{}_3 = -(1/2q^2)(-p^2+q^2+r^2)Z^\alpha{}_2 + b^A{}_{13}\eta^\alpha{}_A,$$
$$X_2 Z^\alpha{}_2 = b^A{}_{22}\eta^\alpha{}_A,$$
$$X_2 Z^\alpha{}_3 = (1/2p^2)(p^2-q^2+r^2)Z^\alpha{}_1 + b^A{}_{23}\eta^\alpha{}_A,$$
$$X_3 Z^\alpha{}_3 = b^A{}_{33}\eta^\alpha{}_A, \quad (4.4)$$

and

$$X_a\eta^\alpha{}_A = -b_{Aa}{}^f Z^\alpha{}_f + \epsilon_A{}^{BC}\mu_{Ba}\eta^\alpha{}_C,$$

or in full detail

$$X_a\eta^\alpha{}_4 = -b_{4a}{}^f Z^\alpha{}_f + \mu_{5a}\eta^\alpha{}_6 - \mu_{6a}\eta^\alpha{}_5,$$
$$X_a\eta^\alpha{}_5 = -b_{5a}{}^f Z^\alpha{}_f + \mu_{6a}\eta^\alpha{}_4 - \mu_{4a}\eta^\alpha{}_6, \quad (4.5)$$
$$X_a\eta^\alpha{}_6 = -b_{6a}{}^f Z^\alpha{}_f + \mu_{4a}\eta^\alpha{}_5 - \mu_{5a}\eta^\alpha{}_4.$$

V. A SPECIAL CASE

In this section we find a certain two-parametric family of solutions of (3.4)–(3.28), characterized by the vanishing of the Ricci scalar. We specialize the fundamental equations for this case. Setting

$$\mu_{Aa} = 0, \quad A = 4,\ 5,\ 6;\ a = 1,\ 2,\ 3 \quad (5.1)$$

and

$$b_{6ab} = 0, \quad a,\ b = 1,\ 2,\ 3, \quad (5.2)$$

one sees immediately that only the diagonal components of b_{4ab} and b_{5ab} can be different from zero. Our system collapses to

$$BC + bc = (1/4p^2)(2p^2(-p^2+q^2+r^2) - (p^2-q^2+r^2)(p^2+q^2-r^2)),$$
$$AC + ac = (1/4q^2)(2q^2(p^2-q^2+r^2) - (p^2+q^2-r^2)(-p^2+q^2+r^2)), \quad (5.3)$$
$$AB + ab = (1/4r^2)(2r^2(p^2+q^2-r^2) - (-p^2+q^2+r^2)(p^2-q^2+r^2)),$$

and

$$A - (1/2q^2)(p^2+q^2-r^2)B - (1/2r^2)(p^2-q^2+r^2)C = 0,$$
$$-(1/2p^2)(p^2+q^2-r^2)A + B - (1/2r^2)(-p^2+q^2+r^2)C = 0, \quad (5.4)$$

and

$$a - (1/2q^2)(p^2+q^2-r^2)b - (1/2r^2)(p^2-q^2+r^2)c = 0,$$
$$-(1/2p^2)(p^2+q^2-r^2)a + b - (1/2r^2)(-p^2+q^2+r^2)c = 0. \quad (5.5)$$

In order to avoid the trivial case of S^3 we have to insist that the determinant

$$1 - (1/4p^2q^2)(p^2+q^2-r^2)^2 = -(1/4p^2q^2)((p^2+q^2-r^2)^2 - 4p^2q^2) = -(1/4p^2q^2)(p^4+q^4+r^4-2q^2r^2 - 2r^2p^2 - 2p^2q^2)$$

should vanish. This happens precisely when the Ricci scalar (3.20) of our space vanishes. We have therefore the following theorem.

Theorem: The subfamily of (1.14), for which the Ricci scalar vanishes, can be embedded into E^5.

This condition is satisfied, for instance, if

$$r = p + q. \quad (5.6)$$

Our equations reduce to

$$BC + bc = 2q(p+q),$$
$$CA + ca = 2p(p+q),$$
$$AB + ab = -2pq, \quad (5.7)$$
$$[1/(p+q)]C = (1/p)A + (1/q)B,$$
$$[1/(p+q)]c = (1/p)a + (1/q)b,$$

having the solution

$$A = p, \quad B = q, \quad C = 2(p+q), \qquad a = \sqrt{3}p, \quad b = -\sqrt{3}q, \quad c = 0. \quad (5.8)$$

That is

$$b_{4ab} = \mathrm{diag}(p \quad q \quad 2(p+q)),$$
$$b_{5ab} = \mathrm{diag}(\sqrt{3}p \quad -\sqrt{3}q \quad 0), \tag{5.9}$$
$$b_{4a}{}^{b} = \mathrm{diag}(1/9 \quad 1/q \quad 2/(p+q)),$$
$$b_{5a}{}^{b} = \mathrm{diag}(\sqrt{3}/p \quad -\sqrt{3}/q \quad 0).$$

Substituting (5.6) and (5.9) into (4.4) and (4.5) we have

$$X_1 Z_1 = p\eta_4 + \sqrt{3}p\eta_5, \tag{5.10a}$$
$$X_1 Z_2 = [q/(p+1)]Z_3, \tag{5.10b}$$
$$X_1 Z_3 = -[(p+q)/q]Z_2, \tag{5.10c}$$
$$X_2 Z_2 = q\eta_4 - \sqrt{3}q\eta_5, \tag{5.10d}$$
$$X_2 Z_3 = [(p+q)/p]Z_1, \tag{5.10e}$$
$$X_3 Z_3 = 2(p+q)\eta_4, \tag{5.10f}$$
$$X_1 \eta_4 = -(1/p)Z_1, \tag{5.11a}$$
$$X_2 \eta_4 = -(1/q)Z_2, \tag{5.11b}$$
$$X_3 \eta_4 = -[2/(p+q)]Z_3, \tag{5.11c}$$

and

$$X_1 \eta_5 = -(\sqrt{3}/p)Z_1, \tag{5.12a}$$
$$Z_2 \eta_5 = (\sqrt{3}/q)Z_2, \tag{5.12b}$$
$$X_3 \eta_5 = 0. \tag{5.12c}$$

We dropped the index α in order to simplify our notation.

We integrate these equations in the next section.

VI. EMBEDDING OF THE SUBFAMILY

Before starting to integrate Eqs. (5.10)–(5.12) we introduce another piece of simplified notation: If F is a function of x,y,z we denote the partial derivatives as

$$\frac{\partial F}{\partial x} = F_x, \quad \frac{\partial F}{\partial y} = F_y, \quad \frac{\partial F}{\partial z} = F_z,$$

respectively.

Since the vector field X_3 has the simple form $X_3 = \partial/\partial z$ in our coordinate system [see (1.20)] we determine the z dependence of our functions first. Equations (5.10f) and (5.11c),

$$Z_{zz} = 2(p+q)\eta_4 \quad \text{and} \quad \eta_{4z} = -[2/(p+q)]Z_z$$

imply

$$Z = S\cos 2z + T\sin 2z + U, \tag{6.1}$$

where S, T, and U are arbitrary functions of x and y only. Substituting (6.1) into (5.10c) and (5.10e) we find partial differential equations for S, T, and U,

$$S_y - 2\cos xT + \sin xT_x = 0, \tag{6.2a}$$
$$T_y + 2\cos xS - \sin xS_x = 0, \tag{6.2b}$$
$$U_x = [(p-q)/(p+q)]S_x, \tag{6.2c}$$
$$U_y = [(p-q)/(p+q)]\sin xT_x, \tag{6.2d}$$

respectively.

Equation (5.10b) leads to

$$(\sin x)^2 S_{xx} + S_{yy} - \tfrac{3}{2}\sin 2x\, S_x + (2(\sin x)^2 + 4(\cos x)^2)S = 0, \tag{6.3a}$$
$$(\sin x)^2 T_{xx} + T_{yy} - \tfrac{3}{2}\sin 2xT_x + (2(\sin x)^2 + 4(\cos x)^2)T = 0, \tag{6.3b}$$
$$T_{xx} + T = 0. \tag{6.3c}$$

These equations are integrability conditions of (6.2).

Observe that the sum of (5.10a), (5.10d), and (5.10f),

$$(X_1X_1 + X_2X_2 + X_3X_3)Z = 3(p+q)\eta_4 + \sqrt{3}(p-q)\eta_5. \tag{6.4}$$

A straightforward calculation shows that

$$3(p+q)\eta_4 + \sqrt{3}(p-q)\eta_5 = Z_{xx} + \cot xZ_x + [1/(\sin x)^2] \times [Z_{yy} - 2\cos xZ_{zy} + Z_{zz}]. \tag{6.5}$$

Substituting (6.1) into (6.5) and using (6.2) and (6.3) we find

$$3(p+q)\eta_4 + \sqrt{3}(p-q)\eta_5 = -6(S\cos 2z + T\sin 2z) + 2[(p-q)/(p+q)](S_{xx} + S) \tag{6.6}$$

and from (5.10f)

$$2(p+q)\eta_4 = -4(S\cos 2z + T\sin 2z). \tag{6.7}$$

Therefore

$$\eta_4 = -[2/(p+q)](S\cos 2z + T\sin 2z), \tag{6.8a}$$
$$\eta_5 = [2/\sqrt{3}(p+q)](S_{xx} + S). \tag{6.8b}$$

We now integrate Eqs. (6.2) and (6.3). Equation (6.3c) implies

$$T = f\cos x + g\sin x, \tag{6.9}$$

where f and g are functions of y only.

Substituting into (6.3b) we obtain

$$\frac{d^2 f}{dy^2} + 4f = 0, \quad \frac{d^2 g}{dy^2} + g = 0, \tag{6.10}$$

that is,

$$f = a\cos 2y + b\sin 2y, \quad g = c\cos y + d\sin y, \tag{6.11}$$

where a, b, c, d, are arbitrary constant five-vectors. Therefore

$$T = (a\cos 2y + b\sin 2y)\cos x + (c\cos y + d\sin y)\sin x. \tag{6.12}$$

We now integrate (6.2) in order to obtain S and U as functions of x and y. The results of this integration are as follows:

$$S = \tfrac{1}{4}a(3+\cos 2x)\sin 2y - \tfrac{1}{4}b(3+\cos 2x)\cos 2y + \tfrac{1}{2}c\sin 2x\sin y - \tfrac{1}{2}d\sin 2x\cos y + \tfrac{1}{4}e(1-\cos 2x), \tag{6.13}$$

where e is the fifth constant five-vector, and

$$U = [(p-q)/(p+q)]\{-\tfrac{1}{4}a(1-\cos 2x)\sin 2y + \tfrac{1}{4}b(1-\cos 2x)\cos 2y + \tfrac{1}{2}c\sin 2x\sin y - \tfrac{1}{2}d\sin 2x\cos y + \tfrac{1}{4}e(1-\cos 2x)\} + F, \tag{6.14}$$

where F is an arbitrary constant five-vector.

At this stage all our equations are satisfied. We now turn our attention to the constant five-vectors entering our expressions as arbitrary constants of integration.

Since η_4 given by (6.8) has to be a unit vector, S and T therefore have to be orthogonal vectors of length $(p+q)/2$, which in turn implies that

$$a=[(p+q)/2]A,\quad b=[(p+q)/2]B,$$
$$c=[(p+q)/2]C,$$
$$d=[(p+q)/2]D,\quad e=[(p+q)/2]\sqrt{3}E$$

have to hold, where A, B, C, D, E are mutually orthogonal unit five-vectors.

Substituting into (6.1) we see that

$$\begin{aligned}Z=&A\{[(p+q)/8](3+\cos 2x)\sin 2y\cos 2z+[(p+q)/2]\cos x\cos 2y\sin 2z-[(p-q)/8](1-\cos 2x)\sin 2y\}\\&+B\{-[(p+q)/8](3+\cos 2x)\cos 2y\cos 2z+[(p+q)/2]\cos x\sin 2y\sin 2z+[(p-q)/8](1-\cos 2x)\cos 2y\}\\&+C\{[(p+q)/4]\sin 2x\sin y\cos 2z+[(p+q)/2]\sin x\cos y\sin 2z+[(p-q)/4]\sin 2x\sin y\}\\&+D\{-[(p+q)/4]\sin 2x\cos y\cos 2z+[(p+q)/2]\sin x\sin y\sin 2z-[(p-q)/4]\sin 2x\cos y\}\\&+E\{[(p+q)/8]\sqrt{3}(1-\cos 2x)\cos 2z+[(p-q)/8]\sqrt{3}(1-\cos 2x)\}+F,\end{aligned}\tag{6.15}$$

showing that our three-space is three-dimensional hypersurface embedded into E^5. There is a certain degree of ambiguity in picking b_{4ab} and b_{5ab} but the rest of it is uniquely defined up to Euclidean motion in E^5.

Setting $F=0$ and choosing the unit vectors A, B, C, D, E in the direction of the coordinate axis we find

$$\begin{aligned}\xi^1=Z^1(x,y,z)=&[(p+q)/8]\\&\times(3+\cos 2x)\sin 2y\cos 2z\\&+[(p+q)/2]\cos x\cos 2y\sin 2z\\&-[(p-q)/8](1-\cos 2x)\sin 2y,\\\xi^2=Z^2(x,y,z)=&-[(p+q)/8]\\&\times(3+\cos 2x)\cos 2y\cos 2z\\&+[(p+q)/2]\cos x\sin 2y\sin 2z\\&+[(p-q)/8](1-\cos 2x)\cos 2y,\\\xi^3=Z^3(x,y,z)=&[(p+q)/4]\sin 2x\sin y\cos 2z\\&+[(p+q)/2]\sin x\cos y\sin 2z\\&+[(p-q)/4]\sin 2x\sin y,\\\xi^4=Z^4(x,y,z)=&-[(p+q)/4]\sin 2x\cos y\cos 2z\\&+[(p+q)/2]\sin x\sin y\sin 2z\\&-[(p-q)/4]\sin 2x\cos y,\\\xi^5=Z^5(x,y,z)=&[(p+q)/8]\sqrt{3}(1-\cos 2x)\cos 2z\\&+[(p-q)/8]\sqrt{3}(1-\cos 2x)\end{aligned}\tag{6.16}$$

or introducing

$$\xi=\xi^1+i\xi^2,\quad \eta=\xi^3+i\xi^4,\quad \zeta=\xi^5\tag{6.17}$$

we have

$$\begin{aligned}\xi=&[(p+q)/2]\{\cos x\sin 2z-(i/4)\\&\times((3+\cos 2x)\cos 2z-[(p-q)/(p+q)]\\&\times(1-\cos 2x))\}e^{2iy},\\\eta=&[(p+q)/2]\{\sin x\sin 2z-(i/2)\sin 2x\\&\times(\cos 2z+[(p-q)/(p+q)])\}e^{iy},\end{aligned}\tag{6.18}$$

$$\begin{aligned}\zeta=&[(p+q)/8]\sqrt{3}(1-\cos 2x)\cos 2z\\&+[(p-q)/8]\sqrt{3}(1-\cos 2x),\end{aligned}$$

as our final result.

Straightforward calculation shows that

$$\begin{aligned}(ds)^2=&d\xi\,d\bar{\xi}+d\eta\,d\bar{\eta}+(d\zeta)^2\\=&\{p^2(\cos z)^2+q^2(\sin z)^2\}(dx)^2\\&+(p^2-q^2)\sin x\sin 2z\,dx\,dy\\&+\{(p^2(\sin z)^2+q^2(\cos z)^2)(\sin x)^2\\&+(p+q)^2(\cos x)^2\}(dy)^2+2(p+q)^2\\&\times\cos x\,dy\,dz+(p+q)^2(dz)^2,\end{aligned}\tag{6.19}$$

that is, (6.15) is indeed an embedding of our subfamily into E^5.

There is a distinguished member of this family at $p=q$. If we disregard the constant conformal factor p^2 then the line element is given by

$$\begin{aligned}(ds)^2=&(dx)^2+((\sin x)^2+4(\cos x)^2)(dy)^2\\&+8\cos x\,dy\,dz+4(dz)^2\end{aligned}\tag{6.20}$$

and Eq. (6.18) simplifies to

$$\begin{aligned}\xi&=\{\cos x\sin 2z-(i/4)(3+\cos 2x)\cos 2z\}e^{2iy},\\\eta&=\{\sin x\sin 2z-(i/2)\sin 2x\cos 2z\}e^{iy},\\\zeta&=(\sqrt{3}/4)(1-\cos 2x)\cos 2z.\end{aligned}\tag{6.21}$$

Observe the remarkable fact that

$$\xi\bar{\xi}+\eta\bar{\eta}+(\zeta)^2=1\tag{6.22}$$

showing that (6.20) is actually embedded into S^4. It would be nice to find—if possible—another equation of the Cartesian coordinates of the form

$$F(\xi,\bar{\xi},\eta,\bar{\eta},\zeta)=0.\tag{6.23}$$

We could then say that the three-space (6.20) is the intersection of S^4 with (6.23). Equation (6.20) is a member of another one-parametric family given by

$$(ds)^2 = (\omega^1)^2 + (\omega^2)^2 + r^2(\omega^3)^2$$
$$= (dx)^2 + ((\sin x)^2 + r^2(\cos x)^2)(dy)^2$$
$$+ 2r^2 \cos x \, dy \, dz + r^2(dz)^2. \quad (6.24)$$

This family has an additional symmetry generated by the vector field $\partial/\partial z$ in our coordinate system. The space sections of the Taub solution[5] have this symmetry.

VII. MISCELLANEOUS REMARKS

Is there any other subfamily that could be embedded into E^5 or is everything else "too curved" for this?

If the ambient space is E^6, however, then the calculations are harder at every stage.

The next task seems to be to embed (6.24).

This family, due to higher symmetry, has other interesting features also.

We think that no member of the family (1.14), except S^3, can be embedded into E^4.

Is there a subfamily of (1.4), that can be embedded into S^4?

ACKNOWLEDGMENTS

We are grateful to Professor A. Trautman for his interest and to Dorothy Carmack for typing the manuscript.

[1]I. Ozsváth and E. L. Schücking, Ann. Phys. (NY) **55**, 166 (1969).
[2]H. Flanders, *Differential Forms* (Academic, New York, 1963).
[3]L. Ph. Eisenhart, *Riemannian Geometry* (Princeton U. P., Princeton, 1926).
[4]M. Spivak, *A Comprehensive Introduction to Differential Geometry* (Publish or Perish, Wilmington, 1979).
[5]A. Taub, Ann. Math. **53**, 472 (1951).

The global Utiyama theorem in Einstein–Cartan theory

Ugo Bruzzo
Dipartimento di Matematica, Università di Genova, Via L. B. Alberti 4, 16132 Genova, Italy

(Received 3 July 1986; accepted for publication 8 April 1987)

A global formulation of Utiyama's theorem for Einstein–Cartan-type gravitational theories regarded as gauge theories of the group of space-time diffeomorphisms is given. The local conditions for the Lagrangian to be gauge invariant coincide with those found by other authors [A. Pérez-Rendón Collantes, "Utiyama type theorems," in *Poincaré Gauge Approach to Gravity. I*, Proceedings Journées Relativistes 1984; A. Pérez-Rendón and J. J. Seisdedos, "Utiyama type theorems in Poincaré gauge approach to gravity. II, " *Preprints de Mathematicas*, Universidad de Salamanca, 1986] in Kibble's and Hehl's approaches.

I. INTRODUCTION

This paper deals with gravitational theories that can be regarded as gauge theories of the group of space-time diffeomorphisms.[1] In particular, we shall consider "metric-affine" theories, in the sense that the field variables are a Lorentzian metric Φ on the space-time four-dimensional manifold M and a linear connection Γ on the coframe bundle $L(M)$.[2] The variational principle is constrained by the requirement that Γ is metric with respect to Φ, i.e., $D_\Gamma \Phi = 0$.

Instead of using Φ as one of the field variables, one can as well use a vierbein (or coframe) field $s = \{e^\alpha\}$, namely, a section of $L(M)$, provided that Φ is recovered *a posteriori* (so to say, after solving the field equations) by regarding s as an orthonormal coframe. Since coframes differing by a Lorentz transformation yield the same Φ, in this case one must also require the theory to be invariant under a Lorentz gauge.

In this paper we consider a generic gravitational theory as a gauge theory of the diffeomorphism group, whose field variables are the coframe field and a linear connection; we describe a suitable geometrical setting for giving a global characterization of the Lagrangians which are gauge invariant, in the sense that they are invariant under space-time diffeomorphisms and local Lorentz transformations (the precise mathematical meaning of these invariances will be clarified later on).

An interesting result of this analysis is that the local conditions of gauge invariance appear to be the same as in Kibble's and Hehl's approaches,[3–4] where Einstein–Cartan-type theories are regarded as gauge theories of the Poincaré group.

II. A GEOMETRICAL SETTING FOR UTIYAMA'S THEOREM

In this section we review some differential geometric constructions, due mainly to Garcia[5] (see also Ref. 6), which allow a very precise and elegant treatment of Utiyama's theorem[5–7] (let us recall that Utiyama's theorem provides a characterization of the Lagrangians invariant under internal gauge transformations). Let M be a paracompact, connected, oriented manifold, P a principal bundle over M with structure group G and projection $\pi: P \to M$; let $\mathbf{g}$ denote the Lie algebra of G. A connection Γ on P can be regarded as a splitting

$$\Gamma: TM \to Q(M) \tag{2.1}$$

of the exact sequence

$$0 \to \mathrm{Ad}\, P \overset{\mathrm{incl}}{\to} Q(M) \overset{\pi_*}{\to} TM \to 0,$$

where TM is the tangent bundle of M, $Q(M)$ is the bundle of G-invariant vector fields on P, and $\mathrm{Ad}\, P$ is the subbundle of the vertical vectors in $Q(M)$, usually called the adjoint bundle. Thus a connection Γ is a section of the affine bundle $C(M)$, whose fiber C_x over $x \in M$ is the set of homomorphisms $\Gamma_x: T_x M \to Q_x(M)$ such that $\pi_* \circ \Gamma_x = \mathrm{id}$.

The space $\mathscr{S}[\mathrm{Ad}\, P]$ of sections of $\mathrm{Ad}\, P$ is a Lie $\mathscr{F}$-algebra (where $\mathscr{F}$ is the ring of smooth functions on M) and is called the *gauge algebra*; its elements, which are vertical vector fields on P, are usually called "infinitesimal gauge transformations." An $A \in \mathscr{S}[\mathrm{Ad}\, P]$ can be represented by a vertical vector field X_A on $C(M)$ in a way that can be intuitively described as follows, a more precise definition being given in Ref. 5. $\mathscr{S}[\mathrm{Ad}\, P]$ can be regarded as the Lie algebra of the group $\mathrm{Aut}^V(P)$ of the vertical automorphisms of P, which acts on $C(M)$ by pullback, thus defining a mapping $\mathrm{Aut}^V(P) \times C(M) \to C(M)$. Differentiating this mapping with respect to the first argument, one gets a mapping $\mathscr{S}[\mathrm{Ad}\, P] \to \mathrm{vert}\; TC(M)$, $A \mapsto X_A$. The local expression of X_A is the following: if $\{D_b, b = 1,...,\dim \mathbf{g}\}$ is a basis of $\mathbf{g}$, $A = A^b D_b$ is a section of $\mathrm{Ad}\, P$, and $\{x^i, \Gamma_i^b\}$ are coordinates on $C(M)$, then

$$X_A = \left(\frac{\partial A^b}{\partial x^i} + c^b_{ad} \Gamma^a_i A^d\right) \frac{\partial}{\partial \Gamma^b_i}, \tag{2.2}$$

where the c^b_{ad} are the structure constants of $\mathbf{g}$. The mapping $A \mapsto X_A$ is a morphism of Lie $\mathbf{R}$-algebras.

The rule that associates the curvature to a connection defines a map $\Omega: J^1C(M) \to \Lambda^2 T^*M \otimes \mathrm{Ad}\, P$, where $J^1C(M)$ is the bundle of jets of sections of $C(M)$.[8] It is easily shown that Ω is a surjective bundle morphism. The gauge algebra $\mathscr{S}[\mathrm{Ad}\, P]$ can be represented on $\Lambda^2 T^*M \otimes \mathrm{Ad}\, P$ in the following way: if f is a function on $\Lambda^2 T^*M \otimes \mathrm{Ad}\, P$ linear on the fibers, and $A \in \mathscr{S}[\mathrm{Ad}\, P]$, then Z_A is the vertical vector field on $\Lambda^2 T^*M \otimes \mathrm{Ad}\, P$ such that

$$(Z_A f)(u) = -f([A,u]) \quad \forall u \in \Lambda^2 T^*M \otimes \mathrm{Ad}\, P. \tag{2.3}$$

Introduced local coordinates $\{x^i, R^{\ b}_{ik}\}$ on $\Lambda^2 T^*M \otimes \mathrm{Ad}\, P$, the local expression of Z_A reads

$$Z_A = -c^d_{ab}\, A^a R^{\ b}_{ik} \frac{\partial}{\partial R^{\ d}_{ik}}. \tag{2.4}$$

Definition: A function L: $J^1C(M) \to \mathbf{R}$ (a Lagrangian function) is said to be *gauge invariant* if

$$\mathscr{L}_{j^1X_A} L = 0 \quad \text{for all } A \in \mathscr{S}[\mathrm{Ad}\, P],$$

where $\mathscr{L}$ denotes the Lie derivative and j^1 the jet extension of vector fields on $C(M)$ to vector fields on $J^1C(M)$.[8]

Theorem (Utiyama): A Lagrangian function L is gauge invariant if and only if $L = \bar{L} \circ \Omega$, where $\bar{L}$ is a function on $\Lambda^2 T^*M \otimes \mathrm{Ad}\, P$ invariant under the representation (2.3) of the gauge algebra $\mathscr{S}[\mathrm{Ad}\, P]$, i.e.,

$$\mathscr{L}_{Z_A} \bar{L} = 0. \qquad \square$$

In less formal terms, we could say that a Lagrangian is invariant with respect to a local gauge if and only if it is a G-invariant function of the curvature.

III. THE GAUGE ALGEBRA OF EINSTEIN–CARTAN THEORY

In this section we wish to describe a geometrical setup for characterizing the set of gauge invariant Lagrangians for general relativity regarded as the gauge theory of the diffeomorphisms group. According to the discussion of Sec. I, we assume the field variables to be sections of a bundle E over M obtained as the fibered product[9] of $L(M)$ with $C(M)$. Here $L(M)$ is the bundle of linear frames on M, while $C(M)$ is the affine bundle of connections on $L(M)$.[5] The Lagrangian L is a function on J^1E; the action functional is a mapping I: $\mathscr{S}[E] \to \mathbf{R}$ (where $\mathscr{S}[E]$ is the space of sections of E) defined as

$$I(\sigma) = \int_M \mathrm{vol}(e)\,(j^1\sigma)^* L \quad \forall \sigma \in \mathscr{S}[E], \tag{3.1}$$

with $j^1\sigma$ being the jet extension[5] of the section $\sigma(x) = (e(x), \Gamma(x))$, and $\mathrm{vol}(e)$ the volume form given by the coframe form $\{e^\alpha, \alpha = 1,...,4\}$.

Now we must define suitable global actions of the diffeomorphism group $\mathrm{Diff}(M)$ and of $O(3,1)$ on E and impose the invariance of (3.1). Let θ be the canonical (soldering) form of $L(M)$, and $\beta \in \mathrm{Diff}(M)$. There exists a canonical lift of β to a diffeomorphism $\hat{\beta}$ of $L(M)$ such that the couple $(\beta, \hat{\beta})$ is an automorphism of $L(M)$; $\hat{\beta}$ is determined by the condition $\hat{\beta}^*\theta = \theta$.

The simplest way to define a lift of β to the bundle $C(M)$ of connections of $L(M)$ is to intend a connection Γ according to (2.1). Then a lift $\tilde{\beta}$ of β to $C(M)$ is defined by requiring the following diagram to be commutative:

$$\begin{array}{ccc} T_xM & \xrightarrow{\ \Gamma_x\ } & Q_x(M) \\ \downarrow{\scriptstyle \beta_*} & & \downarrow{\scriptstyle \hat{\beta}_*} \\ T_{\beta(x)}M & \xrightarrow{\ \tilde{\beta}(\Gamma_x)\ } & Q_{\beta(x)}(M) \end{array}.$$

Combining the two lifts to $C(M)$ and $L(M)$ one gets a representation of $\mathrm{Diff}(M)$ into $\mathrm{Diff}(E)$ which induces by differentiation a representation of TM (the gauge algebra of the diffeomorphism group) into TE.

Now let us turn our attention to the gauge algebra of "local Lorentz transformations." A metric Φ on M determines a reduction κ: $O(M) \to L(M)$, where $O(M)$ is the bundle of coframes on M orthonormal with respect to Φ. The Lorentz gauge algebra is to be identified with $\mathscr{S}[\mathrm{Ad}\, O(M)]$. This algebra can be naturally injected into $\mathscr{S}[\mathrm{Ad}\, L(M)]$ in the following way: let κ': $O(3,1) \to \mathrm{Gl}(4)$ be the Lie group morphism associated with the reduction κ. Any $A \in \mathscr{S}[\mathrm{Ad}\, O(M)]$ can be written as

$$A(u) = A^b(x) D^*_b \quad \text{with } u \in O(M) \text{ and } x = \mu(u), \tag{3.2}$$

where μ: $O(M) \to (M)$ is the bundle projection and $\{D_b\}$ is a basis of the Lorentz algebra $\mathrm{so}(3,1)$; the index b runs over $\mathrm{so}(3,1)$. The above mentioned injection is realized by mapping A to $A^b(x)(\kappa'_* D_b)^*$. Now $\mathscr{S}[\mathrm{Ad}\, O(M)]$ is represented into $\mathscr{S}[\mathrm{vert}\, TL(M)]$ by means of the maps

$$\mathrm{Ad}\, L(M) \xrightarrow{\mathrm{incl}} Q(M) \xrightarrow{\mathrm{incl}} TL(M).$$

This gauge algebra is represented into vert $TC(M)$ using the techniques summarized in Sec. II. Collecting all the results so far discussed, we have that to each vector field

$$Y = Y^i(x) \frac{\partial}{\partial x^i}$$

on M there corresponds a vector field X_Y on E, and to any section A of $\mathscr{S}[\mathrm{Ad}\, O(M)]$ a vertical vector field X_A on E. These correspondences are morphisms of Lie $\mathbf{R}$-algebras. Introduced in E local coordinates systems $\{x^i, \Gamma^b_j, e^\alpha_j\}$, the local expressions of these vector fields read

$$X_Y = Y^i \frac{\partial}{\partial x^i} - \frac{\partial Y^h}{\partial x^j}\left(e^\alpha_h \frac{\partial}{e^\alpha_j} + \Gamma^b_h \frac{\partial}{\Gamma^b_j}\right), \tag{3.3a}$$

$$X_A = \left(\frac{\partial A^b}{\partial x^k} + c^b_{ad} A^a \Gamma^d_k\right) \frac{\partial}{\partial \Gamma^b_k} - e^\alpha A^b c^\beta_{b\alpha} \frac{\partial}{\partial e^\beta_i}. \tag{3.3b}$$

Here the c^a_{bd} are the structure constants of the Lorentz algebra, while the $c^\beta_{b\alpha}$ are the generators of the natural action of $O(3,1)$ over $\mathbf{R}^4$.

IV. UTIYAMA'S THEOREM

The "gravitational" Utiyama's theorem asserts that the conditions of invariance under space-time diffeomorphism and local Lorentz transformations force the Lagrangian to depend upon the field variables (coframe, connection and their derivatives) only through the coframe, curvature, and torsion forms; moreover, the Lagrangian must be a Lorentz invariant function and the associated energy-momentum tensor must vanish identically.

In this section we describe a suitable geometrical setting for expressing these conditions.

(i) Consider the bundle $\mathbf{Curv} = \Lambda^2 T^*M \otimes \mathrm{Ad}\, L(M)$; the law that to any connection associates its curvature, combined with the natural projection $J^1E \to J^1C(M)$, gives a surjective bundle morphism Ω: $J^1E \to \mathbf{Curv}$. Introducing local coordinates $\{x^i, e^\alpha_k, \Gamma^b_k, H^{\ \alpha}_{jk}, B^{\ b}_{jk}\}$ in J^1E and $\{x^i, R^{\ b}_{jk}\}$ in $\mathbf{Curv}$, Ω is described by

$$R^{\ b}_{jk} = B^{\ b}_{jk} - B^{\ b}_{kj} + c^b_{ad} \Gamma^a_j \Gamma^d_k.$$

$\mathscr{S}[\mathrm{Ad}\, O(M)]$ is represented into vert $T(\mathbf{Curv})$ according to Sec. II, while TM is represented into $T(\mathbf{Curv})$ as follows. Any $\beta\in\mathrm{Diff}(M)$ acts on **Curv** by pullback, mapping $u\in\mathbf{Curv}$ to $\beta^{*-1}(u)$. Considering a one-parametric group of diffeomorphisms and differentiating $\beta^{*-1}(u)$ with respect to the parameter, one gets a mapping $TM\to T(\mathbf{Curv})$.

(ii) In complete analogy, we consider the bundle $\mathbf{Tor}=\Lambda^2T^*M\otimes TM$; the law that to any connection and any section of $L(M)$ (coframe) associates the pullback of the torsion form through that section defines again a surjective bundle morphism $T: J^1E\to\mathbf{Tor}$, which, after introducing coordinates $\{x^i, T^\alpha_{jk}\}$ on **Tor**, is described by

$$T^\alpha_{jk}=H^\alpha_{jk}-H^\alpha_{kj}+c^\alpha_{b\beta}(\Gamma^b_j e^\beta_k-\Gamma^b_k e^\beta_j).$$

Again, $T(M)$ is represented into $T(\mathbf{Tor})$ as in $T(\mathbf{Curv})$, while $\mathscr{S}[\mathrm{Ad}\, O(M)]$, according to the methods outlined in Sec. II, is represented into vert $T(\mathbf{Tor})$ as follows. Let $\cdot$ denote the natural product $(\mathrm{Ad}\, O(M))_x\times T_xM\to T_xM$, let $A\in\mathscr{S}[\mathrm{Ad}\, O(M)]$, $u\in\mathbf{Tor}$, f: $\mathbf{Tor}\to\mathbf{R}$ a function linear on the fibers. A is represented by the vertical vector field $\hat{Z}_A$ on **Tor** defined by $(\hat{Z}_A f)(u)=-f(A\cdot u)$. In local coordinates,

$$\hat{Z}_A=A^b c^\alpha_{b\beta}T^\beta_{jh}\frac{\partial}{\partial T^\alpha_{jh}}.$$

(iii) Let us now consider the fibered product $\mathbf{F}=\mathbf{Curv}\times\mathbf{Tor}\times L(M)$ together with the surjective bundle morphism $\Delta: J^1E\to\mathbf{F}$, $\Delta=\Omega\times T\times\mathrm{pr}$, where $\mathrm{pr}: J^1E\to L(M)$ is the natural projection. We shall call Δ the *curvature-torsion* morphism. Collecting the representations of TM and $\mathscr{S}[\mathrm{Ad}\, O(M)]$ on **Curv**, **Tor**, and $L(M)$ that we have so far discussed, we obtain representations $T(M)\to T(\mathbf{F})$, $Y\mapsto Z_Y$, and $\mathscr{S}[\mathrm{Ad}\, O(M)]\to\mathrm{vert}\, T(\mathbf{F})$, $A\mapsto Z_A$. The first representation is a morphism of Lie $\mathscr{F}$-algebras, while the second one is only a Lie $\mathbf{R}$-algebra morphism. After introducing in **F** local coordinates $\{x^i, R^{\ b}_{jk}, T^\alpha_{jk}, e^\alpha_j\}$, the representing vectors read

$$Z_Y=Y^i\frac{\partial}{\partial x^i}-\frac{\partial Y^h}{\partial x^j}\left(T^\alpha_{hk}\frac{\partial}{\partial T^\alpha_{jk}}+R^{\ b}_{hk}\frac{\partial}{\partial R^{\ b}_{jk}}+e^\alpha_h\frac{\partial}{\partial e^\alpha_j}\right),$$

$$Z_A=-A^b\left(c^a_{bd}R^{\ d}_{ik}\frac{\partial}{\partial R^{\ a}_{ik}}+c^\alpha_{b\beta}T^\beta_{jh}\frac{\partial}{\partial T^\alpha_{jh}}+c^\alpha_{b\beta}e^\beta_j\frac{\partial}{\partial e^\alpha_j}\right).$$

Definition: A Lagrangian function $L: J^1E\to\mathbf{R}$ is said to be *locally Lorentz invariant* if

$$\mathscr{L}_{j^1X_A}L=0\quad\forall A\in\mathscr{S}[\mathrm{Ad}\, O(M)];\tag{4.1a}$$

it is said to be *generally invariant* if

$$\mathscr{L}_{j^1X_Y}L=0\quad\text{for all vector fields } Y \text{ on } M.\tag{4.1b}$$

Theorem: L is both local Lorentz and generally invariant if and only if the following conditions hold simultaneously:

(i) $\bar{L}$ filters through the curvature-torsion morphism, i.e., $L=\bar{L}\circ\Delta$, where $\bar{L}$: $\mathbf{F}\to\mathbf{R}$;

(ii) L is invariant under the action of $\mathscr{S}[\mathrm{Ad}\, O(M)]$, i.e.,

$$\mathscr{L}_{Z_A}\bar{L}=0\quad\forall A\in\mathscr{S}[\mathrm{Ad}\, O(M)];$$

(iii) $\bar{L}$ is invariant under the action of TM, i.e.,

$$\mathscr{L}_{Z_Y}\bar{L}=0\quad\text{for all vector fields } Y \text{ on } M.$$ □

Remarks: In local coordinates the conditions (ii) and (iii) read

$$c^a_{bd}R^{\ d}_{ik}\frac{\partial\bar{L}}{\partial R^{\ a}_{ik}}+c^\alpha_{b\beta}T^\beta_{jh}\frac{\partial\bar{L}}{\partial T^\alpha_{jh}}+c^\alpha_{b\beta}e^\beta_j\frac{\partial\bar{L}}{\partial e^\alpha_j}=0;\tag{4.2}$$

$$\frac{\partial\bar{L}}{\partial x^i}=0;\tag{4.3a}$$

$$T^\alpha_{hk}\frac{\partial\bar{L}}{\partial T^\alpha_{jk}}+R^{\ b}_{hk}\frac{\partial\bar{L}}{\partial R^{\ b}_{jk}}+e^i_h\frac{\partial\bar{L}}{\partial e^i_j}=0.\tag{4.3b}$$

Moreover, the condition (4.3b) is equivalent to the fact that the canonical energy-momentum tensor associated to the Lagrangian density $L\cdot\mathrm{vol}(e)$ vanishes identically.

Proof: The proof is the same as in the Yang–Mills case[5]; namely, one writes the conditions (4.1) in local coordinates, thus obtaining a system of differential equations whose local solution is a function on **F** satisfying (4.3). The fact that Δ is a surjective bundle morphism implies that these local conditions are globally equivalent to those stated in the theorem. □

The above mentioned set of local conditions is the same that Pérez-Rendón and Seisdedos find, in a different geometrical setting, for Kibble's and Hehl's approaches.[3,4] This fact is not trivial since those approaches rely on quite different geometrical constructions.

V. CONCLUSIONS

In this paper we have considered the Einstein–Cartan theory as the gauge theory of the group of space-time diffeomorphisms; moreover, in order to allow the use of a coframe (vierbein) field as a field variable instead of the metric form, we have also considered the presence of a Lorentz gauge. Locally, the conditions of gauge invariance force the Lagrangian to depend upon the field variables (coframe, connection) and their derivatives only through the coframe, curvature, and torsion forms; moreover, the Lagrangian cannot depend on the space-time position, and the canonical energy-momentum tensor of the gauge fields must vanish identically. These local conditions have been given a global formulation.

It should be noted that, in the framework of a recently proposed variational formalism on supermanifolds,[10,11] these results can be generalized to (global) superspace formulations of supergravity theories.[12]

ACKNOWLEDGMENTS

The decision to write this paper was made after discussions with Professor A. Pérez-Rendón and Professor J. J. Seisdedos, which are a pleasure to acknowledge. I wish also to thank the referee for valuable suggestions.

This work is based on research supported by the Italian Research Council (CNR), the National Group for Mathematical Physics (GNFM) of CNR, the Consejo Superior de Investigación Cientifica (CSIC), and the Italian Ministry of Public Education through the research project "Geometria and Fisica."

[1]A. Trautman, in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980).

[2]It should be remarked that we consider as a primary object the linear connection Γ on the coframe bundle, and not the associated connection ∇ on the tangent bundle *TM* (holonomic connection in Hehl's terminology). The relationship between the two connections is $\Gamma^\alpha_\beta \wedge e^\beta = \nabla_i e^\alpha \wedge dx^i$.

[3]A. Pérez-Rendon, "Utiyama type theorems in the Poincaré gauge approach to gravity," in *Proceedings Encontro Fisica Matematica* (Coimbra, 1984).

[4]A. Pérez-Rendon and J. J. Seisdedos, "Utiyama type theorems in Poincaré gauge approach to gravity. II," in *Géométrie et Physique, Travaux en Cours*, No. 21, edited by Y. Choquet-Bruhat, B. Coll, R. Kerner, and A. Lichnerowicz (Hermann, Paris, 1987).

[5]P. L. Garcia, J. Differ. Geom. **12**, 209 (1977).

[6]D. Bleecker, *Gauge Theory and Variational Principles* (Addison–Wesley, Reading, MA, 1981).

[7]R. Utiyama, Phys. Rev. **96**, 191 (1956).

[8]A. Trautman, Commun. Math. Phys. **6**, 248 (1967).

[9]D. Husemoller, *Fibre Bundles* (McGraw-Hill, New York, 1966).

[10]U. Bruzzo and R. Cianci, J. Math. Phys. **28**, 786 (1987).

[11]U. Bruzzo, "Lagrangian gauge theories on supermanifolds," *Proceedings of the XIVth Conference on Differential Geometric Methods in Mathematical Physics*, Salamanca, 1985, Lecture Notes in Mathematics (Springer, Berlin, to be published).

[12]U. Bruzzo, "Global Utiyama's theorem in curved superspace," *Proceedings of the XVth Conference on Differential Geometric Methods in Theoretical Physics*, Clausthal, 1986 (World Scientific, Singapore, to be published).

Rigorous estimates for a computer-assisted KAM theory

Alessandra Celletti and Luigi Chierchia[a)]
Forschungsinstitut für Mathematik, ETH-Zentrum, CH-8092 Zürich, Switzerland

(Received 11 August 1986; accepted for publication 6 May 1987)

Nonautonomous Hamiltonian systems of one degree of freedom close to integrable ones are considered. Let ϵ be a positive parameter measuring the strength of the perturbation and denote by ϵ_c the critical value at which a given KAM (Kolmogorov–Arnold–Moser) torus breaks down. A computer-assisted method that allows one to give rigorous lower bounds for ϵ_c is presented. This method has been applied in Celletti–Falcolini–Porzio (to be published in Ann. Inst. H. Poincaré) to the Escande and Doveil pendulum yielding a bound which is within a factor 40.2 of the value indicated by numerical experiments.

I. INTRODUCTION

A problem that has been extensively investigated, both in physics and mathematics, is the stability of invariant surfaces for perturbed integrable systems.[1,2]

Roughly speaking, most of the invariant surfaces for an integrable system are preserved under perturbation if the strength ϵ of the perturbation is sufficiently small. But when ϵ exceeds a certain critical value ϵ_c, these smooth surfaces disappear.

We are interested in analytical tools that allow one to give rigorous and nevertheless realistic lower bounds for ϵ_c in the case of Hamiltonian systems.

For relatively simple dynamical systems, such as holomorphic mappings of the plane ("Siegel's center problem") or some special examples of area preserving diffeomorphisms of an annulus, rather complete results are now available.[3–9] To the best of our knowledge, the methods used in obtaining these results have not been extended to Hamiltonian flows for which the only general tools rely on classical perturbation theory and on KAM theory.[10–16]

For concreteness, and in view of an application we will mention below, we will consider only nonautonomous Hamiltonian systems with one degree of freedom. We remark, however, that our techniques extend easily to the general higher-dimensional situation.

To be more precise, let us consider a Hamiltonian

$$H_0 \equiv h_0(A) + \epsilon f_0(A,\phi,t), \quad \epsilon \geqslant 0,$$

which is a real analytic function defined on a complex domain of the form $D_{R_0}(A_0)\times S^2_{\xi_0}$, where $A_0 \in \mathbb{R}$, $\xi_0>0$, $R_0>0$, $D_{R_0}(A_0)$ is the complex disk

$$\{A\in\mathbb{C}: \ |A-A_0|\leqslant R_0\},$$

and $S^2_{\xi_0}$ is the two-dimensional complex strip

$$\{(\phi,t)\in\mathbb{C}^2: \ |\mathrm{Im}\,\phi|\leqslant\xi_0, \ |\mathrm{Im}\,t|\leqslant\xi_0\}.$$

We assume that the perturbation f_0 has a period 2π both in the "angular" variable ϕ and in the time variable t. In other words, the phase space of our system is the product of a real interval with the standard two-dimensional torus $\mathbb{T}^2 \equiv \mathbb{R}^2/2\pi\mathbb{Z}^2$.

The integrable part h_0 is assumed to be nondegenerate, i.e., for any $A\in D_{R_0}(A_0)$,

$$h_0'' \equiv \frac{d^2h_0}{dA^2}(A) \neq 0.$$

Finally the center A_0 is assumed to be such that the frequency $\omega \equiv h_0'(A_0)$ verifies the Diophantine condition

$$|\omega\nu_1+\nu_2|^{-1}\leqslant C\,|\nu_1|^{\tau} \tag{1}$$

for some $C>0$, $\tau\geqslant1$, and every $(\nu_1,\nu_2)\equiv\nu\in\mathbb{Z}^2$ with $\nu_1\neq0$. For $\epsilon=0$ the torus $\mathfrak{T}^{(0)}(\omega)\equiv\{A_0\}\times\mathbb{T}^2$ is invariant for h_0 and the flow is simply given by

$$(A_0,\phi_0,t_0)\xrightarrow{S_t}(A_0,\phi_0+\omega t,t_0+t). \tag{2}$$

From KAM theory one knows that, for ϵ sufficiently small, there exists, in an ϵ-neighborhood of $\mathfrak{T}^{(0)}(\omega)$, a (unique) analytic torus $\mathfrak{T}^{(\epsilon)}(\omega)$ invariant for H_0 and on which the flow is still given, in suitable coordinates, by (1). Numerical experiments (see e.g., Refs. 17–20 as well as Refs. 1 and 2) have shown that these KAM tori break down when, as mentioned above, ϵ reaches a critical value ϵ_c. We remark that the lower bounds obtained from standard KAM theory have always turned out to be several order of magnitude away from the numerical evidence.

In this paper we are concerned with the problem of obtaining "reasonable" (i.e., "in reasonable agreement with numerical evidence") rigorous lower bounds ϵ_* of ϵ_c so as to insure the existence of KAM tori for $\epsilon<\epsilon_*$.

The method that we are going to present is based on the scheme used by Arnold in his proof of the theorem on conservation, under perturbations, of quasiperiodic motions.[11] We recall briefly this scheme.

One constructs a sequence of Hamiltonians H_j of the form

$$h_j(A';\epsilon)+\epsilon^{2^j}f_j(A',\phi',t';\epsilon), \quad t'=t,$$

defined in shrinking domains $D_{R_j}(A_j)\times S^2_{\xi_j}$. The centers $A_j\in\mathbb{R}$ are chosen so as to keep the frequencies fixed, i.e., $h_j'(A_j)=\omega$. The Hamiltonian H_{j+1}, for $j=0,1,\ldots$, is obtained from H_j with the aid of a real analytic canonical transformation

$$C_j: \ D_{R_{j+1}}(A_{j+1})\times S^2_{\xi_{j+1}}\to D_{R_j}(A_j)\times S^2_{\xi_j}$$

close to the identity transformation. In order to construct C_j,

[a)] Permanent address: Dipartimento di Matematica, IIa Universita' di Roma, 00173 Rome, Italy.

 0022-2488/87/092078-09$02.50

there are a certain number of smallness conditions (in the literature usually referred to as "inductive hypotheses") that ϵ has to satisfy. If ϵ is small enough one can show that all the inductive hypotheses are verified and that, in a suitable sense, H_j becomes, as j goes to infinity, closer to an integrable Hamiltonian. From this one can conclude that the invariant torus $\mathfrak{T}^{(\epsilon)}(\omega)$ is obtained as a limit of the composed transformations

$$C_0 \cdot C_1 \cdot \cdots \cdot C_{j-1}(\{A_j\} \times \mathbb{T}^2).$$

The inductive hypotheses consist of a set of estimates needed to control all the quantities entering in the scheme sketched above. These estimates involve also arbitrary choices of various auxiliary parameters. It is natural to try to obtain better stability estimates by varying these parameters. It will turn out that, in our situation, the dependence of the estimates on the auxiliary parameters is very simple so that, in concrete applications, it will be easy to make good choices. There is, however, a delicate choice that concerns the amount of shrinking of the analyticity domain in the periodic variables ϕ and t. We will discuss this point in detail in Appendix C below.

For the purpose of this introduction, let us denote by $\mathscr{I}_j$ the set of specific conditions that will form our inductive hypotheses at the jth perturbative step leading from H_j to H_{j+1}. In this context, the weakest condition that ϵ has to satisfy in order to insure the existence of the KAM tori $\mathfrak{T}^{(\epsilon)}(\omega)$ is

$$\epsilon < \epsilon_\infty \equiv \sup \{\epsilon > 0\colon\ \mathscr{I}_j \text{ are satisfied for every } j = 0,1,2,...\}.$$

Of course, such a condition has little practical interest since it involves checking an infinite number of estimates. So, we will introduce, for any preassigned integer j_0, a new set of estimates $\mathscr{I}^*_{j_0}$, which will imply all the $\mathscr{I}_j$ for $j > j_0$. Then, for any j_0,

$$\epsilon_{j_0} \equiv \sup \{\epsilon > 0\colon\ \mathscr{I}_0, \mathscr{I}_1, ..., \mathscr{I}_{j_0}, \mathscr{I}^*_{j_0} \text{ are satisfied}\}$$

will provide a concrete lower bound for ϵ_c. In fact, with our choices, ϵ_{j_0} will form a strictly increasing sequence in j_0, so that one obtains better lower bounds by taking larger values of j_0. Now, for $j_0 \sim 20$ the number of elementary operations (i.e., additions, multiplications,..., taking powers, exponentials,...) needed in checking that $\epsilon < \epsilon_{j_0}$ is of the order of 10^5–10^6. Thus, in carrying out these estimates, one is naturally led to the use of computers.

Our method has been applied in Ref. 21, in conjunction with other rigorous numerical computations, in order to give rigorous stability estimates in the following case. Let

$$H_0 \equiv A^2/2 + \epsilon(\cos \phi + \cos (\phi - t)),$$

and consider the stability of the golden-mean torus, i.e., the torus which for $\epsilon = 0$ is given by $\{A_0\} \times \mathbb{T}^2$ with $A_0 \equiv \omega \equiv (1 + \sqrt{5})/2$. In Ref. 20, Escande and Doveil gave numerical, as well as (nonrigorous) theoretical evidence in order to show that this torus disappears for $\epsilon = \epsilon_c \sim 1/29.41$. In Ref. 21 it is proved that the golden-mean torus exists and is analytic for $\epsilon < \epsilon_* \equiv \frac{1}{130}$, and comparing with the experimental results one has

$$\epsilon_c/\epsilon_* \leqslant 40.2.$$

In Ref. 21, it is also pointed out that the best result it was possible to obtain by replacing the method of this paper with the more standard KAM techniques gives a lower bound $\epsilon_{**} = \frac{1}{7930}$ for which $\epsilon_c/\epsilon_{**} \leqslant 2458$.

We conclude this introduction by remarking that the role of computers in obtaining the bounds discussed above is merely to perform lengthy operations with real numbers. By now it is well known how to employ computers in the evaluation of rigorous estimates using, for example, "interval analysis." For more information on this point we refer the reader to Refs. 22–24 and to the literature cited there.

The content of the rest of the paper is as follows. Section II contains the inductive scheme, Sec. III the KAM theorem, Sec. IV the inductive hypotheses $\mathscr{I}^*_{j_0}$, and Sec. V rigorous numerical estimates; and Appendix A contains the self-contained description of the KAM algorithm constructed in this paper, Appendix B the implicit function theorems and a transcedental inequality, and Appendix C the choice of the analyticity-loss sequence $\{\delta_j\}$.

II. INDUCTIVE SCHEME

In this section, maintaining the notations and assumptions of Sec. I, we show how to construct the canonical transformation C_j.

Let us denote by F_j, G_j, and L_j upper bounds on, respectively, $\sup|f_j|$, $\sup|h_j''|$, and $\sup|h_j''|^{-1}$, where the supremum is taken over the domains of definitions $D_{R_j}(A_j) \times S^2_{\xi_j} \equiv D_j \times S^2_{\xi_j}$ and the bars denote the standard norm of complex numbers.

The analyticity assumptions imply the following estimates on the Fourier coeffiecients of f_j:

$$|\hat{f}_{j,\nu}(A)| \leqslant F_j e^{-\xi_j \|\nu\|}, \quad \nu = (\nu_1,\nu_2) \in \mathbb{Z}^2, \tag{3}$$

where for integer vectors ν, $\|\nu\| \equiv |\nu_1| + |\nu_2|$, and

$$\hat{f}_{j,\nu}(A) \equiv \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} f_j(A,\phi,t) e^{-i(\nu_1\phi + \nu_2 t)}\, d\phi\, dt.$$

Another fundamental property of holomorphic functions, which we will often use, is the following. If g is holomorphic on a (smooth) domain D, then for any subdomain $D' \subset D$ one has

$$\sup_{D'} |g'| \leqslant \sup_{D} |g| \ [\operatorname{dist}(\partial D', \partial D)]^{-1}. \tag{4}$$

This estimate follows easily from Cauchy's integral formula for g' taking as contour of integration a circle of radius $\operatorname{dist}(\partial D', \partial D)$ and center $z_0 \in D'$.

Now, following Arnold, we fill in the technical details necessary to carry over the scheme sketched in Sec. I.

Cutoff: Let us split the Fourier expansion of f_j in the following way:

$$f_j = \hat{f}_{j,0} + f_j^{(1)} + f_j^{(2)},$$

where

$$f_j^{(1)} \equiv \sum_{0<\|\nu\|<N_j} \hat{f}_{j,\nu}(A) e^{i(\nu_1\phi + \nu_2 t)},$$

$$f_j^{(2)} \equiv \sum_{\|\nu\|>N_j} \hat{f}_{j,\nu}(A) e^{i(\nu_1\phi + \nu_2 t)},$$

with N_j, to be exactly determined later, such that $f_j^{(2)} \sim O(\epsilon^{2^j})$.

Hamilton–Jacobi perturbative step: Following classical perturbation theory[14] we remove (formally) the perturbation to order $O(\epsilon^{2^{j+1}})$ via the generating function

$$\tilde{\Phi}_j \equiv A'\phi + \epsilon^{2^j}\Phi_j(A',\phi,t;\epsilon),$$

$$\Phi_j \equiv \sum_{0<\|\nu\|<N_j} \frac{\hat{f}_{j,\nu}(A')}{-i[\nu_1 h_j'(A') + \nu_2]} e^{i(\nu_1\phi + \nu_2 t)}.$$

In this case the new integrable part will be

$$h_{j+1}(A') \equiv h_j(A') + \epsilon^{2^j}\hat{f}_{j,0}(A').$$

Analyticity loss in the action variables and the $(j+1)$th approximation to the invariant torus: To make rigorous the formal step described above we have to take care of the small divisors appearing in Φ_j and to do this we have to restrict the analyticity domain in the new action variables. Let $\gamma > 1$ and N_j be such that

$$R_{j+1}' \equiv (1 - 1/\gamma)(CG_jN_j^{1+\tau})^{-1} \leqslant R_j. \tag{5}$$

From (5) it follows easily that for each $A' \in D_{R_{j+1}'}(A_j)$ and $\|\nu\| \leqslant N_j$, $\nu_1 \neq 0$,

$$|\nu_1 h_j'(A') + \nu_2|^{-1} \leqslant \gamma C |\nu_1|^{\tau}. \tag{6}$$

Using an elementary implicit function theorem (Lemma 1 of Appendix B), we can determine the $(j+1)$th approximation to the ω-torus: If, for some $\gamma_1 > 1$,

$$[\gamma_1^2/(\gamma_1 - 1)]\epsilon^{2^j}F_jL_jR_j^{-2} < 1, \tag{7}$$

then there exists a unique $A_{j+1} \in B_{[(\gamma_1-1)/\gamma_1]R_j}(A_j)$ such that

$$h_{j+1}'(A_{j+1}) \equiv h_j'(A_{j+1}) + \epsilon^{2^j}\hat{f}_{j,0}'(A_{j+1}) = h_j'(A_j) = \omega;$$

moreover one has

$$|A_{j+1} - A_j| \leqslant \gamma_1 \epsilon^{2^j} F_j L_j R_j^{-1}. \tag{8}$$

The numbers γ and γ_1 are the first "auxiliary parameters" (Sec. I) which we introduce in order to have complete control of the quantities entering in the estimates.

Now, defining

$$R_{j+1} \equiv R_{j+1}' - \gamma_1\epsilon^{2^j}F_jL_jR_j^{-1},$$

it follows from (8) and (5) that

$$D_{j+1} \equiv D_{R_{j+1}}(A_{j+1}) \subset D_{R_{j+1}'}(A_j) \subset D_j.$$

In order to have complete control on R_{j+1} we require

$$\gamma_1\epsilon^{2^j}F_jL_jR_j^{-1} < (1 - 1/\gamma_1)(1/\gamma_2)(CG_jN_j^{1+\tau})^{-1}, \quad 1/\gamma_2 \equiv 1 - 1/\gamma, \tag{9}$$

and obtain the bounds

$$(\gamma_3 CG_jN_j^{1+\tau})^{-1} \leqslant R_{j+1} < (\gamma_2 CG_jN_j^{1+\tau})^{-1}, \tag{10}$$

with $\gamma_3 \equiv \gamma_1\gamma_2$. Notice that (5) and (9) imply (7).

Control of grad Φ_j and analyticity loss in the periodic variables: In order to control the derivatives of Φ_j we have to restrict the analyticity domain in the (ϕ,t)-variables. Let $\delta_j < \xi_j$ and $A' \in D_{R_{j+1}'}(A_j)$, $(\phi,t) \in S^2_{\xi_j - \delta_j}$. Using (6) and the estimates (3) and (4) we obtain

$$\left|\frac{\partial\Phi_j}{\partial\phi}(A',\phi,t)\right| \leqslant \gamma CF_j \sum_{\nu\neq 0} |\nu_1|^{1+\tau} e^{-\delta_j\|\nu\|} \equiv k_j^{(1)}CF_j, \tag{11}$$

$$\left|\frac{\partial\Phi_j}{\partial A'}(A',\phi,t)\right| < \lambda_j'\gamma CF_jR_j^{-1}\left(\sum_{\nu\neq 0}|\nu_1|^{\tau}e^{-\delta_j\|\nu\|} + \sum_{\nu_2\neq 0}\frac{e^{-\delta_j|\nu_2|}}{|\nu_2|}\right) + \gamma^2C^2F_jG_j\sum_{\nu\neq 0}|\nu_1|^{1+2\tau}e^{-\delta_j\|\nu\|}$$
$$\equiv k_j^{(2)}CF_jR_j^{-1} + k_j^{(3)}C^2F_jG_j, \tag{12}$$

where λ_j' denotes a (strict) upper bound on $[(1 - R_{j+1}'/R_j)]^{-1}$. Analogously one gets

$$\left|\frac{\partial^2\Phi_j}{\partial\phi\,\partial A'}\right| \leqslant k_j^{(4)}CF_jR_j^{-1} + k_j^{(5)}C^2F_jG_j \tag{13}$$

with

$$k_j^{(4)} \equiv \lambda_j'\gamma \sum_{\nu\neq 0}|\nu_1|^{1+\tau}e^{-\delta_j\|\nu\|},$$

$$k_j^{(5)} \equiv \gamma^2 \sum_{\nu\neq 0}|\nu_1|^{2+2\tau}e^{-\delta_j\|\nu\|}.$$

The jth canonical transformation: At a fixed time t, the canonical transformation generated by $\tilde{\Phi}_j$ is obtained by inverting the functions of mixed variables

$$A = A' + \epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}(A',\phi,t), \quad \phi' = \phi + \epsilon^{2^j}\frac{\partial\Phi_j}{\partial A'}(A',\phi,t). \tag{14}$$

First of all we have to make sure that $A \in D_j$ if $A' \in D_{j+1}$. Using (8), (10), and (11) it is readily checked that this will be achieved by requiring

$$(\gamma_2CG_jN_j^{1+\tau})^{-1} + \gamma_1\epsilon^{2^j}F_jL_jR_j^{-1} + k_j^{(1)}\epsilon^{2^j}CF_j \leqslant R_j. \tag{15}$$

Notice that (15) implies (5), thus (15) and (9) imply (5) and (7). Now, using (13) and (10), it is readily checked that

$$\epsilon^{2^j}(k_j^{(4)}CF_jR_j^{-1} + k_j^{(5)}C^2F_jG_j) \leqslant 1 \tag{16}$$

implies the injectivity on D_{j+1} of the first map in (14), which can therefore be inverted in the form

$$A \in \tilde{D}_j \to A + \epsilon^{2^j}\tilde{\Gamma}_j(A,\phi,t) = A' \in D_{j+1},$$

where $\tilde{D}_j$ denotes the image of D_{j+1} under the direct map. Moreover

$$\sup|\tilde{\Gamma}_j| \leqslant \sup\left|\frac{\partial\Phi_j}{\partial\phi}\right|.$$

In order to invert the second map in (14) we have to allow another analyticity loss in the angle variables (however, in practical applications, this second analyticity loss will turn out to be irrelevant with respect to the first one). Let $\dot{\delta}_j > 0$ be such that $\xi_{j+1} \equiv \xi_j - \delta_j - \dot{\delta}_j > 0$. Then, using another elementary implicit function theorem (Lemma 2 of Appendix B), we have that if

$$\epsilon^{2^j}(k_j^{(2)}CF_jR_j^{-1}+k_j^{(3)}C^2F_jG_j)\dot{\delta}_j^{-1}\leqslant 1 \tag{17}$$

the second map in (14) is inverted by

$$\phi'\in S^1_{\xi_{j+1}}\to\phi'+\epsilon^{2^j}\Delta_j(A',\phi',t)=\phi\in S^1_{\xi_j-\delta_j}$$

with

$$\sup|\Delta_j|\leqslant\sup\left|\frac{\partial\Phi_j}{\partial A'}\right|.$$

We can finally define the canonical transformation C_j and its inverse $\widetilde{C}_j$,

$$\widetilde{C}_j:\ (A,\phi,t)\to(A',\phi',t)=(A+\epsilon^{2^j}\widetilde{\Gamma}_j(A,\phi,t),\ \phi+\epsilon^{2^j}\widetilde{\Delta}_j(A,\phi,t),\ t),$$

$$C_j:\ (A',\phi',t)\to(A,\phi,t)=(A'+\epsilon^{2^j}\Gamma_j(A',\phi',t),\ \phi'+\epsilon^{2^j}\Delta_j(A',\phi',t),\ t),$$

with

$$\widetilde{\Delta}_j(A,\phi,t)\equiv\frac{\partial\Phi_j}{\partial A'}(A+\epsilon^{2^j}\widetilde{\Gamma}_j(A,\phi,t),\ \phi,\ t),$$

$$\Gamma_j(A',\phi',t)\equiv\frac{\partial\Phi_j}{\partial\phi}(A',\ \phi'+\epsilon^{2^j}\Delta_j(A',\phi',t),\ t);$$

the domain of holomorphy in the new variables being D_{j+1} and $S^2_{\xi_{j+1}}$.

Estimates on H_{j+1}: The new Hamiltonian $H_{j+1}(A',\phi',t;\epsilon)$ is given by

$$h_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)+\epsilon^{2^j}f_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi},\phi,t\right)+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial t},$$

where Φ_j is evaluated at (A',ϕ,t) with $\phi=\phi'+\epsilon^{2^j}\Delta_j(A',\phi',t)$ so that $f_{j+1}(A',\phi',t)$ will be defined by

$$\epsilon^{-2^{j+1}}\left[h_j\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)-h_j(A')\right]+\epsilon^{-2^j}\left[\hat{f}_{j,0}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)-\hat{f}_{j,0}(A')\right]+\epsilon^{-2^j}\sum_{0<\|\nu\|<N_j}\hat{f}_{j,\nu}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)e^{i(\nu_1\phi+\nu_2t)}+\epsilon^{-2^j}\sum_{\|\nu\|>N_j}\hat{f}_{j,\nu}\left(A'+\epsilon^{2^j}\frac{\partial\Phi_j}{\partial\phi}\right)\cdot e^{i(\nu_1\phi+\nu_2t)}+\epsilon^{-2^j}\frac{\partial\Phi_j}{\partial t}.$$

Denoting by λ_j an upper bound on $[1-R_j^{-1}(R_{j+1}+\epsilon^{2^j}k_j^{(1)}CF_j)]^{-1}$, a straightforward computation yields

$$\sup_{D_{j+1}\times S^2_{\xi_{j+1}}}|f_{j+1}(A',\phi',t)|\leqslant G_j(CF_jk_j^{(1)})^2+k_j^{(6)}CF_j^2R_j^{-1}+F_j\epsilon^{-2^j}\sum_{\|\nu\|>N_j}e^{-\delta_j\|\nu\|}, \tag{18}$$

where

$$k_j^{(6)}\equiv\lambda_j[k_j^{(1)}/(1-e^{-\delta_j})^2].$$

At this point we choose N_j. Let $\alpha>0$ be a new auxiliary parameter and set

$$N_j\equiv\delta_j^{-1}[\log Q_j+2\log(k_j^{(7)}+\log Q_j)],$$

where

$$Q_j\equiv(\alpha k_j^{(8)}\epsilon^{2^j}C^2F_jG_j)^{-1},\quad k_j^{(8)}\equiv(k_j^{(1)})^2/4\beta_j\delta_j^{-1},$$
$$B_j\equiv e^{-\delta_j}/(1-e^{-\delta_j}),\quad k_j^{(7)}\equiv(\beta_j+1)\delta_j.$$

With these definitions it is easy to see that (Lemma 3, Appendix B)

$$F_j\epsilon^{-2^j}\sum_{\|\nu\|>N_j}e^{-\delta_j\|\nu\|}\leqslant\alpha G_j(CF_jk_j^{(1)})^2 \tag{19}$$

provided

$$16e^{-k_j^{(7)}}(\alpha k_j^{(8)}\epsilon^{2^j}C^2F_jG_j)\leqslant 1. \tag{20}$$

Notice that since $\delta_j<1$, $k_j^{(7)}<(1-e^{-1})^{-1}$ and $16e^{-k_j^{(7)}}>e$; hence (20) implies $N_j\geqslant\delta_j^{-1}$.

Denoting by P_j an upper bound on $\epsilon^{2^j}C^2F_jG_j$ and using (18), (19), and (10) we get the basic recurrence relation

$$P_{j+1}\equiv\begin{cases}P_0^2[\sigma_0+\tau_0/(CGR)], & j=0,\\ P_j^2[\sigma_j+\tau_jN_{j-1}^{1+\tau}], & j\geqslant 1,\end{cases}$$

with

$$\sigma_j\equiv\frac{G_{j+1}}{G_j}(1+\alpha)(k_j^{(1)})^2,\quad \tau_j\equiv\begin{cases}G_1k_0^{(6)}/G_0, & j=0,\\ G_{j+1}\gamma_3k_j^{(6)}/G_j, & j\geqslant 1.\end{cases}$$

Next we indicate the necessary bounds on h_j'',

$$\sup|h_j''|\leqslant G_j+\epsilon^{2^j}\lambda_j^2F_jR_j^{-2}\equiv G_{j+1},$$
$$\sup|h_j''|^{-1}\leqslant L_j(1-\epsilon^{2^j}\lambda_j^2L_jF_jR_j^{-2})^{-1}\equiv L_{j+1},$$

provided

$$\epsilon^{2^j}\lambda_j^2L_jF_jR_j^{-2}<1. \tag{21}$$

As for the λ's one can set

$$\lambda_j'=\begin{cases}[1-(\gamma_2CGRN_0^{1+\tau})^{-1}]^{-1}, & j=0,\\ [1-\gamma_1\cdot(N_{j-1}/N_j)^{1+\tau}]^{-1}, & j\geqslant 1,\end{cases}$$

$$\lambda_j=\begin{cases}[1-(\gamma_2CGRN_0^{1+\tau})^{-1}-k_0^{(1)}\epsilon CFR^{-1}]^{-1}, & j=0,\\ [1-\gamma_1\cdot(N_{j-1}/N_j)^{1+\tau}-P_jk_j^{(1)}\gamma_3N_{j-1}^{1+\tau}]^{-1}, & j\geqslant 1.\end{cases}$$

III. KAM THEOREM

We need now to fix the analyticity-loss sequences $\{\delta_j\}$ and $\{\dot{\delta}_j\}$. First notice that we must have

$$\sum_0^\infty(\delta_j+\dot{\delta}_j)<\xi. \tag{22}$$

Let δ be an auxiliary parameter such that $0<\delta<\xi$. Set

$$\delta_j\equiv\delta/2^{j+1},\quad j\geqslant 0,$$

$$\dot{\delta}_j\equiv\begin{cases}\epsilon k_0^{(2)}CFR^{-1}+k_0^{(3)}P_0, & j=0,\\ P_j(\gamma_3k_j^{(2)}N_{j-1}^{1+\tau}+k_j^{(3)}), & j\geqslant 1,\end{cases}$$

and require the condition

$$\sum_0^j\dot{\delta}_n<\xi-\delta. \tag{23}$$

Then it is clear that (22) and (17) are automatically verified.

Remark: In principle any sequence $\{\delta_j\}$ such that

$$\delta_j > 0, \quad \sum \delta_j < \xi, \quad \sum 2^{-j} \log \delta_j^{-1} < \infty$$

is admissible. Our choice is related to the "quadratic" character of the inductive scheme that we are following. For a fuller discussion, see Appendix C. As in Sec. I, let us denote by $\mathscr{I}_j$ the inductive hypotheses (9), (15), (16), (20), (21), and (23) and by ϵ_∞ the number sup $\{\epsilon > 0$: $\mathscr{I}_j$ are verified for every integer $j = 0,1,2,\ldots\}$.

KAM Theorem: If $\epsilon < \epsilon_\infty$ then the map

$$(\phi',t')\in\mathbb{T}^2$$

$$\to J(\phi',t') \equiv \lim_{j\to\infty} C_0 \cdot \cdots \cdot C_{j-1}(A_j,\phi',t')\in\mathbb{R}\times\mathbb{T}^2$$

yields an (analytic) embedding of $\mathbb{T}^2$ into the (generalized) phase space of H_0 so that $J(\mathbb{T}^2)$ is invariant for the flow S_t generated by H_0 and

$$S_t(J(\phi',t')) = J(\phi' + \omega t, t + t'). \tag{24}$$

Proof of (24): Since $C_0 \cdot \cdots \cdot C_{j-1}$ is a canonical transformation, denoting by $S_t^{(j)}$ the flow generated by H_j, we have

$$\begin{aligned} S_t(J(\phi',t')) &\equiv \lim_{j\to\infty} S_t(C_0 \cdot \cdots \cdot C_{j-1}(A_j,\phi',t')) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(S_t^{(j)}(A_j,\phi',t')) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(A_j + O(\epsilon^{2^j}F_j\delta_j^{-1})t, \\ &\quad \phi' + \omega t + O(\epsilon^{2^j}F_jG_j\delta_j^{-1})t^2 \\ &\quad + O(\epsilon^{2^j}F_jR_j^{-1})t,\ t' + t) \\ &= \lim C_0 \cdot \cdots \cdot C_{j-1}(A_j,\phi' + \omega t, t' + t) \\ &= J(\phi' + \omega t, t' + t). \end{aligned}$$

The step before the last identity follows from the inductive hypotheses.

IV. THE INDUCTIVE HYPOTHESES $\mathscr{I}^*_{j_0}$

Assume that the inductive hypotheses are satisfied for $0 \leqslant j \leqslant j_0$, $j_0 > 1$; we present now a method to control them for $j > j_0$. To do this we have to simplify the conditions at the expense of stronger requirements.

Remark: This step, in standard KAM theory, is taken at $j_0 = 0$ and is one of the reasons for the inaccuracy of standard estimates.

Let us start by imposing

$$\begin{aligned} &\left(\frac{k_{j-1}^{(7)}}{k_j^{(7)}}\right)^2 \frac{Q_{j-1}}{Q_j} \\ &\equiv \left(\frac{k_{j-1}^{(7)}}{k_j^{(7)}}\right)^2 \frac{k_j^{(8)}}{k_{j-1}^{(8)}} [P_{j-1}(\sigma_{j-1} + \tau_{j-1}N_{j-2}^{1+\tau})] \leqslant 1, \\ &\quad j > j_0. \end{aligned} \tag{25}$$

Since $k_j^{(7)} \downarrow 1$ one sees that Q_j is increasing in j so that (25) implies easily

$$N_{j-1}/N_j \leqslant \tfrac{1}{2}, \quad j > j_0. \tag{26}$$

Next we split condition (15) in two pieces: Let γ_4 be a new auxiliary parameter such that $\gamma_4 < \gamma_1^{-1}2^{1+\tau}$ and let $\gamma_5 \equiv (1 - \gamma_4^{-1})$. If we require

$$\gamma_4(\gamma_1(1/2^{1+\tau}) + P_jk_j^{(1)}\gamma_3N_{j-1}^{1+\tau}) \leqslant 1, \tag{27}$$

$$\gamma_5P_j\gamma_1\gamma_3^2L_jG_jN_{j-1}^{2+2\tau} \leqslant 1, \tag{28}$$

it is clear that, by (26) and the choice of γ_4 and γ_5, (15) is recovered. Moreover

$$\lambda_j \leqslant \lambda \equiv \gamma_4/(\gamma_4 - 1), \quad \lambda_j' \leqslant \lambda' \equiv \gamma_5/(\gamma_5 - 1).$$

Finally we strengthen (21) and to do this we introduce the last auxiliary parameter. Let $l > 1$ and require

$$[l/(l-1)]P_j(\lambda\gamma_3)^2G_jL_jN_{j-1}^{2+2\tau} \leqslant 1. \tag{29}$$

This is done so that, since $G_jL_j \geqslant 1$, one gets

$$L_{j+1}/L_j \leqslant 1, \quad G_{j+1}/G_j \leqslant g \equiv 2 - 1/l. \tag{30}$$

At this point we need a simple upper bound on N_j $(j > j_0)$. To do this we disregard "logarithmic corrections": Use $\sigma_j + \tau_jN_{j-1}^{1+\tau} > (k_j^{(1)})^2$ to get, for $j \geqslant 1$,

$$P_{j+j_0}^{1/2^j} > P_{j_0}\Psi_j^*, \quad \Psi_j^* \equiv \Psi_j^*(j_0) \equiv \prod_1^j{}_k\ (k_{j_0+k-1}^{(1)})^{2^{1-k}}. \tag{31}$$

Now use (31) to check

$$N_{j+j_0} \leqslant 4^{j+1}\chi_{j+1}, \quad j \geqslant 1, \tag{32}$$

with

$$\begin{aligned} 0 &< \chi_{j+1} \\ &\equiv \chi_{j+1}(P_{j_0}) \equiv 2^{j_0-1}\delta^{-1} \log\left[(\alpha k_{j_0+j}^{(8)})^{1/2^j}P_{j_0}\Psi_j^*\right]^{\tau^{-1}} \\ &\qquad + (1/2^{j-1}) \log\{k_{j_0+j}^{(7)} \\ &\qquad + 2^j \log\left[\alpha k_{j_0+j}^{(8)})1/2^jP_{j_0}\Psi_j^*\right]^{-1}\}. \end{aligned}$$

Finally using (32) we obtain the estimate

$$P_{j_0+j}^{1/2^j} \leqslant P_{j_0}\Psi_j \tag{33}$$

with

$$\Psi_1 \equiv (\sigma_{j_0} + \tau_{j_0}N_{j_0-1}^{1+\tau})^{1/2},$$

$$\Psi_2 \equiv \Psi_1(\sigma'_{j_0+1} + \tau'_{j_0+1}N_{j_0}^{1+\tau})^{1/4},$$

$$\Psi_j \equiv \Psi_2 \prod_3^{j-1} \left[\sigma'_{j_0+k} + \tau'_{j_0+k}(4^k\chi_k)^{1+\tau}\right]^{1/2^{k+1}}, \quad j \geqslant 3,$$

and

$$\sigma_j' \equiv g(1+\alpha)(k_j^{(1)})^2,$$

$$\tau_j' \equiv g\gamma_3\lambda k_j^{(1)}(1 - e^{-\delta_j})^{-2}, \quad j > j_0.$$

Notice that Ψ_j^*, χ_j, and Ψ_j converge monotonically and very rapidly as $j \uparrow \infty$; we will denote the corresponding limits by Ψ^*, χ, and Ψ,

$$\Psi_j^* \uparrow \Psi^*, \quad \chi_j \downarrow \chi, \quad \Psi_j \uparrow \Psi.$$

We are now in a position to control easily all the inductive hypotheses [(9), (27), (28), (16), (20), (29), (23), and (25)] for $j \geqslant j_0 + 1$. Consider, for example, (9), which can be rewritten as

$$[\gamma_1/(\gamma_1 - 1)]\gamma_3^2P_jG_jL_j(N_{j-1}N_j)^{(1+\tau)} < 1. \tag{34}$$

Using (33), (30), and (32) one sees that, for $j = j_0 + n$ and $n \geqslant 2$, (34) is implied by

$$P_{j_0}\theta_n^{(1)}\psi_n \leqslant 1, \tag{35}$$

where

$$\theta_n^{(1)} \equiv [(\gamma_1/(\gamma_1-1))\gamma_3^2 G_{j_0} L_{j_0}(gl)^n (4^{2n+1}\chi_n\chi_{n+1})^{1+\tau}]^{1/2^n}.$$

Now, it is not hard to see that $\theta_n^{(1)} \downarrow 1$ and that the function $n \to \theta_n^{(1)}\psi_n$ $(n \geqslant 2)$ has a unique maximum that will be achieved for some value $n = n_1^*$. Therefore (9) will be implied, for any $j \geqslant j_0 + 2$, by

$$P_{j_0}\theta_{n_1^*}^{(1)}\psi_{n_1^*} \leqslant 1.$$

Completely analogous considerations apply to the rest of the inductive hypotheses; for a complete and explicit list of all the conditions entering in $\mathscr{I}_{j_0}^*$, see Appendix A.

V. RIGOROUS NUMERICAL ESTIMATES

The condition $\epsilon < \epsilon_\infty$ in the KAM theorem of Sec. III can now be replaced by the more practical condition $\epsilon < \epsilon_{j_0}$, where j_0 is any integer greater than 2 and, as in Sec. I,

$$\epsilon_{j_0} \equiv \sup\{\epsilon > 0: \mathscr{I}_0, \mathscr{I}_1, \ldots, \mathscr{I}_{j_0}, \mathscr{I}_{j_0}^* \text{ are verified}\}.$$

From the preceding sections it follows that ϵ_{j_0} is a strictly increasing function of j_0, so that, in concrete applications, one is interested in taking j_0 as large as possible. Already for j_0 greater than, say, 5, it will be readily realized that, in order to check that $\epsilon < \epsilon_{j_0}$, the use of a computer becomes necessary (in applications a reasonable choice might be $j_0 \sim 30$; compare Ref. 21. In this case one can proceed as follows.

Let us denote by $\boldsymbol{\alpha}$ the set of auxiliary parameters $\{\delta, \alpha, \gamma, \gamma_1, \gamma_4, l\}$ and, to stress that the estimates depend on the choice of $\boldsymbol{\alpha}$, let us replace ϵ_{j_0} by and $\epsilon_{j_0}(\boldsymbol{\alpha})$. One can then write a program that, for any choice of $\boldsymbol{\alpha}$, checks if a given number ϵ verifies or not the conditions $\mathscr{I}_0, \mathscr{I}_1, \ldots \mathscr{I}_{j_0}, \mathscr{I}_{j_0}^*$. By "trial and error," it will be easy to find a (close) lower estimate, $\epsilon_*(\boldsymbol{\alpha})$ of $\epsilon_{j_0}(\boldsymbol{\alpha})$. At this point, varying $\boldsymbol{\alpha}$, one can "maximize" $\epsilon_*(\boldsymbol{\alpha})$ so as to obtain the final result. Because of the simple dependence of ϵ_{j_0} on $\boldsymbol{\alpha}$, this latter operation will turn out to be rather straightforward.

Important remark: Our method, as well as all KAM theorems, deals with very general situations and, *a fortiori*, does not exploit the peculiarities of the system at hand; such peculiarities might include the geometry of the phase space, singularities in the action variables, special properties in Fourier space, symmetries, etc. Thus, before applying our method, one might use the more flexible finite-order perturbation theory to conjugate the given Hamiltonian to a new one with a smaller perturbation and which, in general, having lost all its special properties, will be closer to a "generic" Hamiltonian. For a detailed discussion and illustration of these ideas we refer the reader to Ref. 21.

ACKNOWLEDGMENTS

The authors are deeply indebted to Professor G. Gallavotti for having started them in this topic and for having been a continuous source of inspiration and help. They wish also to thank Professor J. Moser for helpful comments.

Finally, they thank respectively, the Department of Planetary Sciences and the Department of Mathematics of the University of Arizona for the kind hospitality during the period in which part of this work has been completed.

APPENDIX A: SELF-CONTAINED DESCRIPTION OF THE KAM ALGORITHM CONSTRUCTED IN THIS PAPER

Let

$$H_0(A,\phi,t;\epsilon) \equiv h_0(A) + \epsilon f_0(A,\phi,t), \quad (A,\phi,t) \in B_{R_0}(A_0) \times \mathbb{T}^2,$$

where

$$B_{R_0}(A_0) \equiv \{A \in \mathbb{R}: |A - A_0| \leqslant R_0\}, \quad \mathbb{T}^2 \equiv \mathbb{R}^2/2\pi\mathbb{Z}^2$$

and $A_0 \in \mathbb{R}$ is such that $\omega \equiv h_0'(A_0)$ satisfies

$$|\omega\nu_1 + \nu_2|^{-1} \leqslant C|\nu_1|^\tau$$

for any $(\nu_1,\nu_2) \in \mathbb{Z}^2$, $\nu_1 \neq 0$ and for some $C, \tau > 0$.

Assume that H_0 can be extended to a holomorphic function on

$$D_0 \times S_0 \equiv D_{R_0}(A_0) \times S_{\xi_0}^2$$
$$\equiv \{A \in \mathbb{C}: |A - A_0| \leqslant R_0\}$$
$$\times \{(z_1,z_2) \in \mathbb{C}^2: |\mathrm{Im}\, z_i| \leqslant \xi_0, \ i = 1,2\}$$

and denote by F_0, G_0, L_0 upper bounds on, respectively,

$$\sup_{D_0 \times S_0} |f_0|, \quad \sup_{D_0} |h_0''|, \quad \sup_{D_0} |h_0''|^{-1}.$$

Finally, let $j_0 \geqslant 2$ be a fixed integer.

1. KAM theorem (compare Secs. III and IV)

If $\mathscr{I}_j$ $(j = 0,1,\ldots,j_0)$ and $\mathscr{I}_{j_0}^*$ are the inductive hypotheses described below and if $\epsilon < \epsilon_{j_0} \equiv \sup\{\epsilon > 0$: $\mathscr{I}_0, \mathscr{I}_1, \ldots, \mathscr{I}_{j_0}$ and $\mathscr{I}_{j_0}^*$ are verified$\}$ then there exists an analytic torus ϵ close to $\{A_0\} \times \mathbb{T}^2$ invariant for the flow generated by H_0. On such a torus the flow is given (in suitable coordinates) by

$$(\phi', t') \in \mathbb{T}^2 \to (\phi' + \omega t, t' + t).$$

The rest of this appendix is devoted to the description of the conditions $\mathscr{I}_j$ and ,$\mathscr{I}_{j_0}^*$. These conditions are expressed in terms of recursive objects. To introduce such objects we start with the following.

2. Definition of the auxiliary parameters

Let δ, α, γ, γ_1, γ_4, l, be such that $\delta < \xi_0$, $\alpha > 0$, $\gamma > 1$, $\gamma_1 > 1$, $\gamma_4 < \gamma_1^{-1} 2^{1+\tau}$, $l > 1$.

Now, define

$$\gamma_2 \equiv (1 - 1/\gamma)^{-1}, \quad \gamma_3 \equiv \gamma_1\gamma_2,$$
$$\gamma_5 \equiv (1 - 1/\gamma_4)^{-1}, \quad g \equiv 2 - 1/l,$$
$$\lambda \equiv \gamma_4(\gamma_4 - 1)^{-1}, \quad \lambda' \equiv \gamma_5(\gamma_5 - 1)^{-1}.$$

Then, for $j \geqslant 0$, we set

$$\delta_j \equiv \delta/2^{j+1}, \quad s_j(\rho) \equiv \sum_{\nu \in \mathbb{Z}^2} |\nu_1|^\rho e^{-\delta_j\|\nu\|} \quad (\rho > 0),$$
$$k_j^{(1)} \equiv \gamma s_j(1+\tau), \quad k_j^{(3)} \equiv \gamma^2 s_j(1+2\tau),$$
$$k_j^{(5)} \equiv \gamma^2 s_j(2+2\tau), \quad \beta_j \equiv e^{-\delta_j}(1 - e^{-\delta_j})^{-1},$$
$$k_j^{(7)} \equiv (\beta_j + 1)\delta_j,$$
$$k_j^{(8)} \equiv (k_j^{(1)})^2/(4\beta_j\delta_j^{-1}).$$

Next we will introduce the recursive quantities P_j, Q_j, N_j, λ_j, λ_j', $k_j^{(2)}$, $k_j^{(4)}$, $k_j^{(6)}$, G_j, L_j, σ_j, τ_j. These quantities will be computable according to the following "computational sequence" ("$\cdots \to X \to Y$" means "from the set of quantities

X and the quantities known before the computation of X one can compute the set of quantities Y"):

$\epsilon, C, F_0, R_0, G_0, L_0 \to$

$P_0 \to Q_0 \to N_0 \to \lambda_0 \to \lambda'_0 \to k_0^{(2)}, k_0^{(4)}, k_0^{(6)}, G_1, L_1, \to \sigma_0, \tau_0 \to$

$P_1 \to Q_1 \to N_1 \to \lambda_1, \lambda'_1 \to k_1^{(2)}, k_1^{(4)}, k_1^{(6)}, G_2, L_2 \to \sigma_1, \tau_1 \to$

$\vdots$

$P_j \to Q_j \to N_j \to \lambda_j, \lambda'_j \to k_j^{(2)}, k_j^{(4)}, k_j^{(6)}, G_{j+1}, L_{j+1} \to \sigma_j, \tau_j \to .$

3. Definition of the recursive quantities

We have

$$P_0 \equiv \epsilon C^2 F_0 G_0, \quad Q_0 \equiv (\alpha k_0^{(8)} P_0)^{-1},$$
$$N_0 \equiv \delta_0^{-1}[\log Q_0 + 2\log(k_0^{(7)} + \log Q_0)],$$
$$\lambda_0 \equiv [1 - (\gamma_2 C G_0 R_0 N_0^{1+\tau})^{-1} - k_0^{(1)} \epsilon C F_0 R_0^{-1}]^{-1},$$
$$\lambda'_0 \equiv [1 - (\gamma_2 C G_0 R_0 N_0^{1+\tau})^{-1}]^{-1},$$
$$k_0^{(2)} \equiv \lambda'_0 \gamma \left(s_0(\tau) + 2 \sum_{n=1}^{\infty} \frac{e^{-\delta_0 n}}{n} \right),$$
$$k_0^{(4)} \equiv \lambda'_0 \gamma s_0(1+\tau), \; k_0^{(6)} \equiv \lambda_0 k_0^{(1)} (1 - e^{-\delta_0})^{-2},$$
$$G_1 \equiv G_0 + \epsilon \lambda_0^2 F_0 R_0^{-2}), \quad L_1 \equiv L_0 (1 - \epsilon \lambda_0^2 L_0 F_0 R_0^{-2})^{-1},$$
$$\sigma_0 \equiv (G_1/G_0)(1+\alpha)(k_0^{(1)})^2, \quad \tau_0 \equiv G_1 k_0^{(6)}/G_0.$$

For $j \geqslant 1$ we set

$$P_j \equiv \begin{cases} P_0^2[\sigma_0 + \tau_0/(C G_0 R_0)^{-1}], & j=1, \\ P_{j-1}^2[\sigma_{j-1} + \tau_{j-1} N_{j-2}^{1+\tau}], & j \geqslant 2, \end{cases}$$
$$Q_j \equiv (\alpha k_j^{(8)} P_j)^{-1},$$
$$N_j \equiv \delta_j^{-1}[\log Q_j + 2\log(k_j^{(7)} + \log Q_j)],$$
$$\lambda_j = [1 - \gamma_1 \cdot (N_{j-1}/N_j)^{1+\tau} - P_j k_j^{(1)} \gamma_3 N_{j-1}^{1+\tau}]^{-1},$$
$$\lambda'_j = [1 - \gamma_1 \cdot (N_{j-1}/N_j)^{1+\tau}]^{-1},$$
$$k_j^{(2)} \equiv \lambda'_j \gamma \left(s_j(\tau) + 2 \sum_{n=1}^{\infty} \frac{e^{-\delta_j n}}{n} \right),$$
$$k_j^{(4)} \equiv \lambda'_j \gamma s_j(1+\tau), \quad k_j^{(6)} \equiv \lambda_j k_j^{(1)} (1 - e^{-\delta_j})^{-2},$$
$$G_{j+1} \equiv G_j [1 + (\lambda_j \gamma_3)^2 P_j N_j^{2+2\tau}],$$
$$L_{j+1} \equiv L_j [1 - (\lambda_j \gamma_3)^2 P_j N_j^{2+2\tau} G_j L_j]^{-1},$$
$$\sigma_j \equiv (G_{j+1}/G_j)(1+\alpha)(k_j^{(1)})^2,$$
$$\tau_j \equiv (G_{j+1}/G_j)\gamma_3 k_j^{(6)}.$$

Remark: At the moment, some of the above quantities may be ill defined but this will not be the case as soon as the correspondent conditions $\mathcal{J}_j$ are verified.

4. The inductive hypotheses $\mathcal{J}_j$ ($0 \leqslant j \leqslant j_0$)

The following set of inequalities, (A1)–(A5), constitute the set of inductive hypotheses $\mathcal{J}_j$ with $j = 0,1,...,j_0$:

$$[\gamma_1/(\gamma_1 - 1)]\gamma_3 \epsilon F_0 L_0 R_0^{-1} (C G_j N_j^{1+\tau}) < 1 \quad (j=0),$$
$$[\gamma_1/(\gamma_1 - 1)]\gamma_3^2 P_j G_j L_j (N_{j-1} N_j)^{1+\tau} < 1 \quad (1 \leqslant j \leqslant j_0), \tag{A1}$$
$$(\gamma_2 C G_0 N_0^{1+\tau})^{-1} + \gamma_1 \epsilon F_0 L_0 R_0^{-1} + k_0^{(1)} \epsilon C F_0 \leqslant R_0 \quad (j=0),$$
$$\gamma_1 [(N_{j-1}/N_j)^{1+\tau} + \gamma_3^2 P_j G_j L_j N_{j-1}^{2+2\tau} + \gamma_2 k_j^{(1)} P_j N_{j-1}^{1+\tau}] \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A2}$$
$$\epsilon k_0^{(4)} C F_0 R_0^{-1} + k_0^{(5)} P_0 \leqslant 1 \quad (j=0),$$
$$[\gamma_3 k_j^{(4)} N_{j-1}^{1+\tau} + k_j^{(5)}] P_j \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A3}$$
$$16 e^{-k_j^{(7)}} \alpha k_j^{(8)} P_j \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A4}$$
$$\epsilon \lambda_0^2 L_0 F_0 R_0^{-2} < 1 \quad (j=0),$$
$$(\gamma_3 \lambda_j)^2 P_j G_j L_j N_{j-1}^{2+2\tau} \leqslant 1 \quad (1 \leqslant j \leqslant j_0), \tag{A5}$$
$$k_0^{(2)} \epsilon C F_0 R_0^{-1} + k_0^{(3)} P_0 + \sum_{n=1}^{j} (k_n^{(2)} N_{n-1}^{1+\tau} + k_n^{(3)}) P_n \leqslant \xi - \delta. \tag{A6}$$

[(A1)–(A6) correspond to, respectively, (9), (15), (16), (20), (21), and (23) of Secs. II and III.]

5. The inductive hypotheses $\mathcal{J}_{j_0}^*$

In order to describe the set of conditions in $\mathcal{J}_{j_0}^*$ we need the following definitions:

$$\sigma'_j \equiv g(1+\alpha)(k_j^{(1)})^2, \quad \tau'_j \equiv g\gamma_3 \lambda k_j^{(1)} (1 - e^{-\delta_j})^{-2}, \quad j > j_0, \quad \Psi_n^* \equiv \prod_{m=1}^{n} (k_{j_0+m-1}^{(1)})^{1/2^{m-1}},$$
$$\chi_n \equiv 2^{j_0 - 1} \delta^{-1} \log[(\alpha k_{j_0+n-1}^{(8)})^{1/2^{n-1}} P_{j_0} \Psi_n^*]^{-1} + (1/2^{n-2}) \log\{k_{j_0+n-1}^{(7)} + 2^{n-1} \log[\alpha k_{j_0+n-1}^{(8)})^{1/2^{n-1}} P_{j_0} \Psi_n^*]^{-1}\},$$
$$\Psi_1 \equiv (\sigma_{j_0} + \tau_{j_0} N_{j_0-1}^{1+\tau})^{1/2}, \quad \Psi_2 \equiv \Psi_1 (\sigma'_{j_0+1} + \tau'_{j_0+1} N_{j_0}^{1+\tau})^{1/4}, \quad \Psi_n \equiv \Psi_2 \prod_{k=2}^{n-1} [\sigma'_{j_0+k} + \tau'_{j_0+k} (4^k \chi_k)^{1+\tau}]^{1/2^{k+1}}, \quad n \geqslant 3,$$
$$\tilde{k}_{j_0+n}^{(2)} \equiv \lambda' \gamma s_{j_0+n}(\tau), \quad \tilde{k}_{j_0+n}^{(4)} \equiv \lambda' \gamma s_{j_0+n}(1+\tau),$$
$$\theta_n^{(1)} \equiv [(\gamma_1/(\gamma_1 - 1))\gamma_3^2 G_{j_0} L_{j_0} (gl)^n (4^{2n+1} \chi_n \chi_{n+1})^{1+\tau}]^{1/2^n}, \quad \theta_n^{(2)} \equiv [\gamma_3 k_{j_0+n}^{(1)} (4^n \chi_n)^{1+\tau} (1 - \gamma_4 \gamma_1 / 2^{1+\tau})^{-1}]^{1/2^n},$$
$$\theta_n^{(3)} \equiv [\gamma_1 \gamma_3^2 \gamma_5 G_{j_0} L_{j_0} (lg)^n (4^n \chi_n)^{2+2\tau}]^{1/2^n}, \quad \theta_n^{(4)} \equiv [\gamma_3 \tilde{k}_{j_0+n}^{(4)} (4^n \chi_n)^{1+\tau} + k_{j_0+n}^{(5)}]^{1/2^n}, \quad \theta_n^{(5)} \equiv \left[16 \alpha k_{j_0+n}^{(8)} e^{-k_{j_0+n}^{(7)}}\right]^{1/2^n},$$

$$\theta_n^{(6)} \equiv \left[\left(k_{j_0+n}^{(7)}/k_{j_0+n+1}^{(7)} \right)^2 \left(k_{j_0+n+1}^{(8)}/k_{j_0+n}^{(8)} \right) \left[\sigma'_{j_0+n} + \tau'_{j_0+n}(4^n\chi_n)^{1+\tau} \right] \right]^{1/2^n},$$

$$\theta_n^{(7)} \equiv \left[(l/(l-1))(\lambda\gamma_3)^2 G_{j_0} L_{j_0} (gl)^n (4^n\chi_n)^{2+2\tau} \right]^{1/2^n}, \quad \theta_n^{(8)} \equiv \left[\tilde{k}_{j_0+n}^{(2)} 4^n \chi_n + k_{j_0+n}^{(3)} \right]^{1/2^n}.$$

Now, denoting by Ψ the limit of the Ψ_n's, one has that

$$\Psi_n \uparrow \Psi \text{ and } \theta_n^{(i)} \downarrow 1 \quad (i = 1,2,...,8);$$

moreover the functions $n \to \Psi_n \theta_n^{(i)}$ $(n \geqslant 2)$ have a unique maximum achieved at some value $n = n_i^*$.

The following set of inequalities, (A7)–(A14), constitutes the set of inductive hypotheses $\mathscr{J}_{j_0}^*$:

(A1) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_1^*} \theta_{n_1^*}^{(1)} \leqslant 1, \tag{A7}$$

$$\gamma_4 \gamma_1 2^{-1(1+\tau)} + \gamma_3 P_{j_0+1} k_{j_0+1}^{(1)} N_{j_0}^{1+\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_2^*} \theta_{n_2^*}^{(2)} \leqslant 1, \tag{A8}$$

$$\gamma_1 \gamma_3^2 \gamma_5 P_{j_0+1} G_{j_0+1} L_{j_0+1} N_{j_0}^{2+2\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_3^*} \theta_{n_3^*}^{(3)} \leqslant 1, \tag{A9}$$

(A3) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_4^*} \theta_{n_4^*}^{(4)} \leqslant 1, \quad n \geqslant 2, \tag{A10}$$

(A4) with $j = j_0 + 1$,

$$P_{j_0} \Psi_{n_5^*} \theta_{n_5^*}^{(5)} \leqslant 1, \quad n \geqslant 2, \tag{A11}$$

$$(k_j^{(7)}/k_{j+1}^{(7)})^2 (k_{j+1}^{(8)}/k_j^{(8)}) [P_j(\sigma_j + \tau_j N_{j-1}^{1+\tau})] \leqslant 1 \quad \text{with } j = j_0, j_0 + 1,$$

$$P_{j_0} \Psi_{n_6^*} \theta_{n_6^*}^{(6)} \leqslant 1, \quad n \geqslant 2, \tag{A12}$$

$$[l/(l-1)] P_{j_0+1} (\gamma\lambda_3)^2 G_{j_0+1} L_{j_0+1} N_{j_0}^{2+2\tau} \leqslant 1,$$

$$P_{j_0} \Psi_{n_7^*} \theta_{n_7^*}^{(7)}, \quad n \geqslant 2, \tag{A13}$$

$$k_0^{(2)} \epsilon C F_0 R_0^{-1} + k_0^{(3)} P_0 + \sum_{m=1}^{j_0+1} (k_m^{(2)} N_{m-1}^{1+\tau} + k_m^{(3)}) P_m + \sum_{m=2}^{n_8^*-1} (P_{j_0} \Psi_m \theta_m^{(8)})^{2^m} + \sum_{m=n_8^*}^{\infty} (P_{j_0} \Psi_{n_8^*} \theta_{n_8^*}^{(8)})^{2^m} \leqslant \xi_0 - \delta. \tag{A14}$$

Remark: Because the convergence of Ψ_n and $\theta_n^{(i)}$ to their limits takes place at a very fast rate, it is clear that to find explicitly the values n_i^* in concrete applications is not a difficult task.

APPENDIX B: IMPLICIT FUNCTION THEOREMS AND A TRANSCENDENTAL INEQUALITY

Lemma 1: Let I be the interval $(x_0 - r, x_0 + r)$, let g be a continuous function on I, and let h be a differentiable function on I.

If $(\sup_I |g|) \cdot (\sup_I |h'|^{-1}) < r$ then there exists a unique point $x_1 \in I$ s.t.

$$h(x_1) + g(x_1) = h(x_0).$$

Moreover $|x_1 - x_0| \leqslant (\sup |g|) \cdot (\sup |h'|^{-1})$.

Proof: The map $x \in I \to h^{-1} \circ (h(x_0) - g(x))$ is a contraction from I into I.

Lemma 2: Let g be a holomorphic map on S_ξ and denote by $|\cdot|_\xi$ the sup norm on S_ξ. If

$$\max\{|g'|_\xi, |g|_\xi \delta^{-1}\} < 1$$

then the map $z \in S_\xi \to z + g(z)$ is one-to-one from S_ξ onto $S_{\xi-\delta}$ and the inverse map $z' \in S_{\xi-\delta} \to z' + h(z') \in S_\xi$ satisfies $|h|_{\xi-\delta} \leqslant |g|_\xi$.

Proof: Injectivity is plain from

$$|z + g(z) - [z' + g(z')]| \geqslant |z - z'| (1 - |g'|_\xi), \quad z, z' \in S_\xi.$$

To prove surjectivity let $w \in S_{\xi-\delta}$. Then the map

$$z \in B \equiv \{z \in C : |z - w| < \delta\} \to w - g(z)$$

is a contraction from B into itself.

Lemma 3: If $e^{x_0} a \geqslant 16$ then $e^x (x + x_0)^{-1} \geqslant a$ for any $x \geqslant \log a + 2 \log(x_0 + \log a)$.

The proof is elementary and is omitted.

APPENDIX C: ON THE CHOICE OF THE ANALYTICITY-LOSS SEQUENCE $\{\delta_j\}$

The size of the perturbation f_{j+1} at the $(j+1)$th stage is given inductively by $P_{j+1} = P_j^2(\sigma_j + \tau_j N_{j-1}^{1+\tau})$, N_{j-1} being a logarithmic correction in P_{j-1}. If we disregard such logarithmic correction we get $P_{j+1} \cong P_j^2 \sigma_j$.

Let us assume, for the moment, that $\xi_0 < 1$. Then $\delta_j < 1$ for each j and

$$\sigma_j \cong s\delta_j^{-n}, \quad \text{some } s > 0 \text{ and } n \in \mathbb{Z},$$

so that

$$P_{j+1} \cong P_j^2 \sigma_j \cong P_j^2 s \delta_j^{-n} \cong P_0^{2^{j+1}} \prod_{0}^{j}{}_k (s\delta_k^{-n})^{2^{j-k}}.$$

From this one deduces that the best (up to the above logarithmic corrections) choice of $\{\delta_j\}$ is the one that minimizes the functional

$$\prod_{0}^{\infty} \delta_k^{-1/2^k}$$

over sequences satisfying $\Sigma\delta_k = \xi_0$. This is an easy minimum problem that can be immediately solved using Lagrange multipliers obtaining

$$\delta_k = \xi_0/2^{k+1}.$$

Now, if $\xi_0 \geqslant 1$, one can replace the auxiliary parameters α of Sec. V by $\alpha' \equiv \{\alpha, j', \delta_0, \delta_1, \ldots, \delta_{j'}\}$, where j' and $\delta_0, \ldots, \delta_{j'}$ are new auxiliary parameters such that

$$\xi' \equiv \xi_0 - \sum_{j=0}^{j'} \delta_j < 1.$$

Then, for $j > j'$ one can repeat the above argument and set $\delta_{j'+1} \equiv \xi'/2^k$ for $k \geqslant 1$.

[1] *Chaotic Behaviour of Deterministic Systems*, edited by G. Iooss, R. H. G. Helleman, and R. Stora (North-Holland, Amsterdam, 1983).

[2] I. Percival, Lect. Notes Phys. **247**, 12 (1985); R. S. MacKay, *ibid.* **247**, 390 (1985); J. Moser, *ibid.* **247**, 492 (1985).

[3] M. R. Herman, Lectures given at Les Houches, July 1981.

[4] M. R. Herman, Asterisque **1**, 103 (1985).

[5] R. de la Llave, J. Math. Phys. **24**, 2118 (1983).

[6] R. de la Llave and D. Rana, "Proof of accurate bounds in small denominators problems," preprint 1986.

[7] C. Liverani and G. Turchetti, Lett. Nuovo Cimento **39**, 17 (1984).

[8] J. N. Mather, Ergodic Theory Dynamical Systems **2**, 397 (1982).

[9] R. S. MacKay and I. C. Percival, Commun. Math. Phys. **98**, 469 (1985).

[10] A. N. Kolmogorov, Dokl. Akad. Nauk. SSR **98**, 527 (1954).

[11] V. I. Arnold, Russ. Math. Surveys **18**, 9 (1963).

[12] J. Moser, Nach. Akad. Wiss. Göttingen, Math. Phys. Kl. **II**, **1**, 1 (1962).

[13] H. Rüssmann, Lect. Notes Phys. **38**, 598 (1976).

[14] G. Gallavotti, *The Elements of Mechanics* (Springer, New York 1983).

[15] G. Gallavotti, *Perturbation Theory for Classical Hamiltonian Systems, Progress in Physics*, edited by J. Fröhlich (Birkhäuser, Boston, 1982).

[16] A. Giorgilli and L. Galgani, "Rigorous estimates for the series expansions of Hamiltonian perturbation theory," preprint 1986.

[17] B. Chirikov, Phys. Rep. **52**, 263 (1979).

[18] J. M. Greene, J. Math. Phys. **20**, 1183 (1979).

[19] G. Benettin, M. Casartelli, L. Galgani, A. Giorgilli, and J. M. Strelcyn, Nuovo Cimento B **50**, 211 (1979).

[20] D. F. Escande and F. Doveil, J. Stat. Phys. **26**, 257 (1981).

[21] A. Celletti, C. Falcolini, and A. Porzio, "Rigorous numerical stability estimates for the existence of KAM tori in a forced pendulum," to appear in Ann. Inst. H. Poincaré.

[22] O. E. Lanford III, "Computer assisted proofs in analysis," Physica A **124**, 465 (1984).

[23] O. E. Lanford III, "A computer assisted proof of Feigenbaum's conjectures," Bull. Am. Math. Soc. **6**, 427 (1982).

[24] J.- P. Eckmann, H. Koch, and P. Wittwer, "A computer assisted proof of universality in area preserving maps," Mem. Am. Math. Soc. **47**, 289 (1984).

Solitary wave solutions of a system of coupled nonlinear equations

C. Guha-Roy
Department of Mathematics, Jadavpur University, Calcutta 700032, India

(Received 19 August 1986; accepted for publication 13 May 1987)

A class of coupled nonlinear wave equations is presented. It is shown that the coupled equation possesses solitary wave solutions. Some comments are made on the previously obtained solutions of a similar class of equations.

I. INTRODUCTION

Following the work of Hirota and Satsuma,[1] the study of a system of coupled KdV equations has assumed crucial significance. Recently, several studies concerning the coupled KdV equation have been done.[2–6] However, no work on the coupled version of the higher KdV equation seems to have been reported.

In this work, we study a new system of coupled nonlinear wave equations dealing with the coupled version of the combined form of the higher (modified) KdV equation and the KdV equation. We also extract the solitary wave solutions of this coupled equation and make some comments on the previously obtained solutions of a similar class of equations.

II. BASIC EQUATION AND MATHEMATICAL FORMULATION OF THE MODEL

We begin by writing down a set of coupled equations in the form

$$u_t + \alpha v^2 v_x + \beta u^2 u_x + \lambda u u_x + \gamma u_{xxx} = 0, \tag{1a}$$

$$v_t + \delta (uv)_x + \epsilon v v_x = 0, \tag{1b}$$

involving the variables $u(x,t)$ and $v(x,t)$. In (1), $\alpha, \beta, \lambda, \gamma, \delta$, and ϵ are arbitrary parameters. For $v = 0$, Eq. (1) reduces to mixed form of the modified KdV equation and the KdV equation.

Let us consider the following ansatz:

$$u = u(x - ct), \quad v = v(x - ct). \tag{2}$$

A similar set of equations also holds if the argument is of the form $(x + ct)$.

Using (2), Eqs. (1) become

$$-cu_s + \frac{\alpha}{3}(v^3)_s + \frac{\beta}{3}(u^3)_s + \frac{\lambda}{2}(u^2)_s + \gamma u_{sss} = 0, \tag{3a}$$

$$-cv_s + \delta(uv)_s + (\epsilon/2)(v^2)_s = 0, \tag{3b}$$

where s denotes the quantity $(x - ct)$.

Integrating (3b), we get, after rearrangement,

$$u = k/v + c/\delta - (\epsilon/2\delta)v, \tag{4}$$

where the integration constant k may be treated as an arbitrary parameter. However, we have to impose $k = 0$ in order to have a regular u everywhere and in particular when $v \to 0$. It may be noted that $u(s)$ satisfies[4] the following boundary conditions:

$$u, u_s, u_{ss} \to 0 \quad \text{as } |s| \to \infty. \tag{5}$$

Thus (4) reduces to

$$v = (2/\epsilon)(c - \delta u). \tag{6}$$

This shows that $v(s)$ is directly related to $u(s)$ and is different from those obtained by Kawamoto[4] using a coupled KdV equation. In Kawamoto's case $v(s)$ was found to be inversely related to $u(s)$.

We now substitute (6) and (3a) and then integrate the result twice w.r.t. s. We get, on using (5),

$$u_s^2 = \sum_{i=2}^{4} a_i u^i, \tag{7}$$

where the parameters a_2, a_3, and a_4 stand for

$$a_2 = (1/\gamma)(c + 8\alpha\delta c^2/\epsilon^3). \tag{8a}$$

$$a_3 = -(1/3\gamma)(\lambda + 16\alpha c\delta^2/\epsilon^3), \tag{8b}$$

and

$$a_4 = (1/6\gamma)(8\alpha\delta^3/\epsilon^3 - \beta). \tag{8c}$$

It is interesting to note that the classical Boussinesq equation[7] can also be transformed to Eq. (7).

In the following, we are going to investigate the possible solitary wave solutions that follow from (7).

III. SOLITARY WAVE SOLUTIONS

Setting $u = 1/\Psi$, we transform[8] Eq. (7) into

$$\Psi_s^2 = a_2\Psi^2 + a_3\Psi + a_4.$$

This equation has a solitary wave solution of the form

$$u(s) = 2a_2/\{A \cosh[\sqrt{a_2}(s + s_0)] - a_3\} \tag{9}$$

for $a_2 > 0$, $a_3 < 0$, and $a_3^2 > 4a_2a_4$. In (9), $A = \sqrt{a_3^2 - 4a_2a_4}$ and s_0 is an integration constant. It may be remarked that a solution corresponding to $a_3 > 0$ is not physically admissible since in this case the denominator of (9) may vanish leading to a singularity. Further when $a_3 = -2\sqrt{a_2a_4}$, Eq. (7) leads to the kink–antikink solutions as

$$u(s) = (a_2/a_3)[1 \pm \tanh(\sqrt{a_2}/2)(s + s_0)]. \tag{10}$$

Let us now distinguish some interesting cases of (7) that also exhibit solitary wave solutions.

Case → 1: When $\beta = 8\alpha\delta^3/\epsilon^3$, a_4 reduces to zero. We have then, from (7),

$$u_s^2 = a_2u^2 + a_3u^3. \tag{11}$$

If $a_2 > 0$ and $a_3 < 0$, (11) yields the solitary wave solution

$$u(s) = -(a_2/a_3)\text{sech}^2[(\sqrt{a_2}/2)(s + s_0)], \tag{12}$$

 0022-2488/87/092087-02$02.50

for

$$-(a_2/a_3) \geqslant u(s) \geqslant 0.$$

It is worthwhile to note that if one applies a similar procedure as given in Sec. II to extract a traveling-wave solution, the coupled KdV equation,

$$u_t + \alpha v v_x + \beta u u_x + \gamma u_{xxx} = 0, \quad (13a)$$

$$v_t + \delta(uv)_x + \epsilon v v_x = 0, \quad (13b)$$

reduces to the same form as in (11) and the solution turns out to be that given by (12). On the other hand, for the coupled modified KdV equation one has to adjust the parameter λ, which we consider below.

Case→2: When $\lambda = -16\alpha c\delta^2/\epsilon^3$, a_3 vanishes and we have, from (7),

$$u_s^2 = a_2 u^2 + a_4 u^4. \quad (14)$$

Now if $a_2 > 0$ and $a_4 < 0$, we again have the solitary solutions

$$u(s) = \pm\sqrt{|(a_2/a_4)|}\,\text{sech}\,[\sqrt{a_2}(s+s_0)]. \quad (15)$$

It is readily seen that the coupled modified KdV equation of the form

$$u_t + \alpha v_x + \beta u^2 u_x + \gamma u_{xxx} = 0, \quad (16a)$$

$$v_t + \delta(uv)_x + \epsilon v v_x = 0, \quad (16b)$$

can be converted into (14) so as to obtain the solutions (15).

It is needless to mention that the corresponding solutions for $v(s)$ may be obtained by inserting the solutions of $u(s)$ into (6).

IV. APPLICATIONS OF THE SOLUTIONS

The solutions obtained in the previous section bear a close similarity to those following from a restricted class of flow fields dealing with large amplitude Rossby waves. In a recent work,[9] Benney has obtained some interesting solutions of such a class of waves by relaxing the assumption that the nonlinearity is weak. It may be noted that when the nonlinearity is weak, at large times each long wave mode is governed by the standard KdV equation. By considering strong nonlinear counterparts, finite amplitude solitary wave solutions may be obtained that are similar to those obtained in Eqs. (10) and (12).

It may be noted that the steady form of the combined KdV–modified KdV equation has also been analyzed by Wadati.[10] For a detailed study of the properties of large amplitude Rossby waves, we refer to the work of Ref. 9.

V. CONCLUDING REMARKS

We have obtained several forms of solitary wave solutions from a set of coupled nonlinear wave equations. It may be remarked that the nature of these solitary wave solutions depend remarkably on the signs of the parameters a_2, a_3, and a_4. Apart from the solitary wave solutions, we have also obtained the kink–antikink solutions. In addition, a number of applications of such solutions have been discussed.

ACKNOWLEDGMENTS

I wish to express my great appreciation to Professor D. K. Sinha and Dr. B. Bagchi for their helpful discussions and thank the referee for his useful comments.

I am also grateful to the Council of Scientific and Industrial Research (C.S.I.R.), India for financial assistance.

[1]R. Hirota and J. Satsuma, Phys. Lett. A **85**, 407 (1981).

[2]B. A. Kupershmidt, J. Phys. A: Math. Gen. **18**, L571 (1985).

[3]M. Ito, Phys. Lett. A **91**, 335 (1982).

[4]S. Kawamoto, J. Phys. Soc. Jpn. **53**, 1203 (1984); see also E. V. Krishnan, *ibid.* **55**, 3753 (1986).

[5]C. Guha-Roy, B. Bagchi, and D. K. Sinha, J. Math. Phys. **27**, 2558 (1986).

[6]C. Guha-Roy, B. Bagchi, and D. K. Sinha, "Solitary wave solution of a new type of coupled K–dV equation," JU/DM/3/85, preprint, 1985.

[7]S. Kawamoto, J. Phys. Soc. Jpn. **54**, 1701 (1985); see also E. V. Krishnan, *ibid.* **51**, 2391 (1982).

[8]L. Hlavatý, J. Phys. A: Math. Gen. **18**, 1933 (1985).

[9]D. J. Benney, Stud. Appl. Math. **60**, 1 (1979).

[10]M. Wadati, J. Phys. Soc. Jpn. **38**, 673 (1975).

Newman–Penrose constants for zero-rest-mass fields on Minkowskian space-time

P. A. Hogan
Mathematical Physics Department, University College, Dublin, Ireland and School of Theoretical Physics, Dublin Institute for Advanced Studies, Dublin, Ireland

(Received 31 December 1986; accepted for publication 29 April 1987)

It is proved that the Newman–Penrose constants associated with zero-rest-mass spin-1 fields on Minkowskian space-time are inherited from the Newman–Penrose constants associated with the complex wave function from which these fields are constructed. This latter construction is also given and a possible extension of these results to spin-s ($=0,1,2,...$) fields is indicated.

I. INTRODUCTION

This paper is concerned with some of the constants or conserved quantities associated with zero-rest-mass free-fields discovered by Newman and Penrose[1,2]—specifically those associated with spin s ($=0,1,2,...$) fields on Minkowskian space-time. We address the question: can one say more than that the constants exist only because of a special choice of boundary conditions? A study by Goldberg[3] of the relationship between the Newman–Penrose constants and invariant transformations for spin-1 fields on Minkowskian space-time seemed to suggest, with some reservation, a negative answer to this question. We begin here by considering spin-1 fields on Minkowskian space-time. Having established in Sec. III that such fields can be constructed using a complex wave function (or pair of real wave functions), we proceed to prove in the following section that *the Newman–Penrose constants associated with zero-rest-mass spin-1 fields on Minkowskian space-time are inherited from the Newman–Penrose constants associated with the complex wave function from which these fields are constructed.* We thus appear to have an affirmative answer, for spin-1 fields, to the question posed above. In Sec. V we indicate how this result might be extended to fields of integral spin $s\geqslant 1$. Some useful formulas are given in the appendices.

II. NEWMAN–PENROSE CONSTANTS

The line element of Minkowskian space-time in rectangular Cartesian coordinates and time $X^i=(x,y,z,t)$ is given by (taking $c=1$)

$$ds^2=\eta_{ij}\,dX^i\,dX^j=dx^2+dy^2+dz^2-dt^2. \qquad (2.1)$$

Alternatively, in coordinates (θ,ϕ,r,u) with

$$x+iy=re^{i\phi}\sin\theta,\quad z=r\cos\theta,\quad u=t-r, \qquad (2.2)$$

the line element (2.1) reads

$$ds^2=r^2(d\theta^2+\sin^2\theta\,d\phi^2)-2\,du\,dr-du^2. \qquad (2.3)$$

Let $Q(X)$ be a wave function. Thus

$$\Box Q\equiv\frac{\partial^2Q}{\partial x^2}+\frac{\partial^2Q}{\partial y^2}+\frac{\partial^2Q}{\partial z^2}-\frac{\partial^2Q}{\partial t^2}=0. \qquad (2.4)$$

In general $Q(X)$ may be complex, i.e., $Q(X)=U(X)+iV(X)$, where U,V are real-valued wave functions. In terms of the coordinates (θ,ϕ,r,u) introduced in (2.2) the wave equation (2.4) reads

$$\frac{\partial^2Q}{\partial r^2}-2\frac{\partial^2Q}{\partial r\,\partial u}+\frac{2}{r}\left(\frac{\partial Q}{\partial r}-\frac{\partial Q}{\partial u}\right)=-\frac{1}{r^2}\Delta Q, \qquad (2.5)$$

where

$$\Delta Q\equiv\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Q}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2Q}{\partial\phi^2}. \qquad (2.6)$$

Consider now solutions of (2.5) of the form

$$Q=\sum_{n=0}^{\infty}\frac{Q^n(u,\theta,\phi)}{r^{n+1}}, \qquad (2.7)$$

where the functions Q^n, $n=0,1,2,...$, are assumed to possess continuous derivatives of all orders for $-\infty<u<+\infty$, $0\leqslant\theta\leqslant\pi$, $0\leqslant\phi\leqslant 2\pi$, with values at $\phi=0$ the same as at $\phi=2\pi$. In the course of this work, when we make series expansions of the form of (2.7), we shall always assume that the coefficients of the various inverse powers of r are functions having these properties. Substituting (2.7) into (2.5) results in

$$-2(n+1)\dot{Q}^{n+1}=\Delta Q^n+n(n+1)Q^n \quad (n\geqslant 0), \qquad (2.8)$$

where the dot indicates partial differentiation with respect to u. Let $Y_{l,m}(\theta,\phi)$ ($l=0,1,2,...;\ |m|\leqslant l,\ m\in\mathbb{Z}$) be the (spin-weight-0) spherical harmonics. These are solutions of

$$\Delta Y_{l,m}=-l(l+1)Y_{l,m}. \qquad (2.9)$$

Let $\overline{Y}_{l,m}$ denote the complex conjugate of $Y_{l,m}$. Using (2.8) and (2.9) we find

$$-2(n+1)\int\dot{Q}^{n+1}\overline{Y}_{n,m}\,d\omega=0, \qquad (2.10)$$

where $d\omega=\sin\theta\,d\theta\,d\phi$ and the integration is over the unit two-sphere. Thus we have

$${}_0\dot{G}^{n+1}_{n,m}=0, \qquad (2.11)$$

where

$${}_0G^{n+1}_{n,m}=\int Q^{n+1}\overline{Y}_{n,m}\,d\omega \quad (n\geqslant 0,\ |m|\leqslant n). \qquad (2.12)$$

The quantities ${}_0G^{n+1}_{n,m}$ are the Newman–Penrose constants or conserved quantities associated with a (complex) wave function or spin-0 field on Minkowskian space-time.

If $F_{ij}=-F_{ji}$ are the components of a real bivector field

 0022-2488/87/092089-05$02.50

then the anti-self-dual bivector $F_{ij} + i\,{}^*F_{ij}$ can be expanded on a basis of complex, anti-self-dual bivectors N_{ij}, M_{ij}, L_{ij} as [see (A7)],

$$F_{ij} + i\,{}^*F_{ij} = \Phi_0\, N_{ij} + \Phi_1\, L_{ij} + \Phi_2\, M_{ij}, \tag{2.13}$$

where Φ_0, Φ_1, Φ_2 are the components of F_{ij} on a null tetrad [given in our case by (A1)] and have spin weights 1,0 and -1, respectively. For the concept of "spin weight," the spin raising and lowering operators $\eth$ and $\bar{\eth}$, and spin-s spherical harmonics, see Newman and Penrose,[1,2] Goldberg *et al.*,[4] and Penrose and Rindler.[5] In Appendix B we note, from these references, some results we will make use of in this paper. If we assume Φ_0 in (2.13) to have the form

$$\Phi_0 = \sum_{n=0}^{\infty} \frac{\Phi_0^n(u,\theta,\phi)}{r^{n+3}}, \tag{2.14}$$

and then require (2.13) to satisfy Maxwell's vacuum field equations we find, among other things, that Φ_0^n satisfies (cf. Newman and Penrose[2] with whom we have some slight differences in notation and convention)

$$-2(n+1)\dot{\Phi}_0^{n+1} = \bar{\eth}\eth\Phi_0^n + n(n+3)\Phi_0^n \quad (n\geqslant 0). \tag{2.15}$$

Proceeding in the same manner as in the scalar case above, but using the spin-1 spherical harmonics ${}_1Y_{l,m}$ (see Appendix B) rather than the ordinary (spin-0) spherical harmonics, we find from (2.15) that

$${}_1\dot{G}^{n+1}_{n+1,m} = 0, \tag{2.16}$$

where

$${}_1G^{n+1}_{n+1,m} = \int \Phi_0^{n+1}\, {}_1\bar{Y}_{n+1,m}\, d\omega \quad (n\geqslant 0,\ |m|\leqslant n+1). \tag{2.17}$$

The quantities ${}_1G^{n+1}_{n+1,m}$ are the Newman–Penrose constants associated with a spin-1 (Maxwell) field on Minkowskian space-time.

In general if ${}_sQ$ $(s=0,1,2,...)$ is the spin-weight-s component of a spin-s field and if we write

$${}_sQ = \sum_{n=0}^{\infty} \frac{{}_sQ^n(u,\theta,\phi)}{r^{n+2s+1}}, \tag{2.18}$$

then the Newman–Penrose constants are given by

$${}_sG^{n+1}_{n+s,m} = \int {}_sQ^{n+1}\, {}_s\bar{Y}_{n+s,m}\, d\omega, \tag{2.19}$$

where $n=0,1,2,...$, $|m|\leqslant n+s$, and ${}_s\bar{Y}_{n+s,m}$ are the complex conjugate spin-weight-s spherical harmonics. Comparing (2.18) with (2.7) and (2.14) we see that

$${}_0Q^n = Q^n, \quad {}_1Q^n = \Phi_0^n \quad (n\geqslant 0). \tag{2.20}$$

III. CONSTRUCTION OF SPIN-1 FIELD

We first indicate a proof that *every solution of the vacuum Maxwell equations on Minkowskian space-time can be constructed out of a pair of real wave functions and a constant real bivector.*[6]

If $F_{ij} = -F_{ji}$ are the components of a real bivector, in coordinates X^i, on Minkowskian space-time, then, as shown in Appendix A, we can write

$$F_{ij} + i\,{}^*F_{ij} = \phi_0\, n_{ij} + \phi_1\, l_{ij} + \phi_2\, m_{ij}, \tag{3.1}$$

where n_{ij}, l_{ij}, m_{ij} are complex, anti-self-dual bivectors with constant components, while ϕ_0, ϕ_1, ϕ_2 are the components of F_{ij} on the null tetrad k^i, l^i, m^i, $\bar{m}^i$ given by (A3). If the bivector (3.1) satisfies Maxwell's vacuum field equations

$$(F^{ij} + i\,{}^*F^{ij})_{,j} = 0, \tag{3.2}$$

then one finds that these are in fact integrability conditions for the existence of a complex-valued function Q of X^i such that Q is a wave function,

$$\Box Q = 0, \tag{3.3}$$

and

$$\begin{aligned}\phi_0 &= -\tfrac{1}{4}k^i m^j Q_{,ij},\\ \phi_1 &= -\tfrac{1}{4}k^i l^j Q_{,ij} = -\tfrac{1}{4}m^i \bar{m}^j Q_{,ij},\\ \phi_2 &= \tfrac{1}{4}l^i \bar{m}^j Q_{,ij}.\end{aligned} \tag{3.4}$$

The two expressions for ϕ_1 arise on account of (3.3). Substituting (3.4) into (3.1) we find

$$F_{ij} + i\,{}^*F_{ij} = \tfrac{1}{4}(l^a\bar{m}^b m_{ij} - k^a l^b l_{ij} - k^a m^b n_{ij})Q_{,ab}. \tag{3.5}$$

If we define the constant, complex, self-dual bivector

$$K_{ij} - i\,{}^*K_{ij} = \tfrac{1}{4}(\bar{m}_i m_j - m_i\bar{m}_j + l_i k_j - l_j k_i), \tag{3.6}$$

we can show, using (3.3) and the equation

$$m_i\bar{m}_j + \bar{m}_i m_j - k_i l_j - k_j l_i = 2\eta_{ij}, \tag{3.7}$$

that (3.5) may be rewritten as

$$F_{ij} + i\,{}^*F_{ij} = (K_i{}^l - i\,{}^*K_i{}^l)Q_{,lj} - (K_j{}^l - i\,{}^*K_j{}^l)Q_{,li}. \tag{3.8}$$

Writing $Q = U + iV$, where U and V are real-valued wave functions, we obtain

$$F_{ij} = (K_i{}^l U_{,l} + {}^*K_i{}^l V_{,l})_{,j} - (K_j{}^l U_{,l} + {}^*K_j{}^l V_{,l})_{,i}, \tag{3.9}$$

and this field can clearly be derived from the four potential (modulo a gauge transformation)

$$A_i = K_i{}^l U_{,l} + {}^*K_i{}^l V_{,l}. \tag{3.10}$$

Thus every solution of the vacuum Maxwell equations on Minkowskian space-time can be constructed from a pair of real wave functions U, V and a constant real bivector K_{ij}. Maxwell fields with four-potentials of the form (3.10), but with the second term on the right missing, have been discussed by Synge[7] along with some detailed examples. Whittaker[8] demonstrated in 1904 that, in effect, the four-potential due to an arbitrarily moving point charge could be put in the form (3.10). Recently Synge[9] has established (3.9) using an argument based on the Cauchy problem for Maxwell's equations and the wave equation. Using the spinor formalism (3.9) has already been proved by Penrose.[10] If the bivector F_{ij} has spinor components $\phi_{AB} = \phi_{BA}$ then (3.9) is equivalent to

$$\phi_{AB} = \bar{\mu}^{C'}\bar{\mu}^{D'}\nabla_{AC'}\nabla_{BD'}Q, \tag{3.11}$$

for some constant one-spinor μ^A. Furthermore Penrose[10] has extended this result to spin-s fields in general. It is also interesting to note in the present context that Stewart[11] has shown that source-free electromagnetic perturbations of vacuum space-times can be described by a complex scalar field and that this generalizes to the case of gravitational perturbations when the background vacuum space-time is algebraically special.

IV. INHERITANCE OF CONSERVED QUANTITIES

The relationship between the components Φ_0, Φ_1, Φ_2 of a Maxwell field which appear in Sec. II and the components ϕ_0, ϕ_1, ϕ_2 used in Sec. III is given by Eq. (A8). If we substitute into (A8) the expressions (3.4) for ϕ_0, ϕ_1, ϕ_2 in terms of the complex wave function Q and for Q the expression (2.7) we obtain an equation of the form (2.14) with

$$\Phi_0^n = \frac{1}{\sqrt{2}}\left\{\frac{1}{2}\sin\theta\frac{\partial^2 Q^n}{\partial\theta^2} - \frac{1}{2\sin\theta}\frac{\partial^2 Q^n}{\partial\phi^2} + \left(n+\frac{3}{2}\right)\right.$$
$$\times\cos\theta\frac{\partial Q^n}{\partial\theta} - \frac{1}{2}(n+1)(n+2)\sin\theta Q^n$$
$$\left. + i\frac{\partial^2 Q^n}{\partial\theta\,\partial\phi} + i(n+1)\cot\theta\frac{\partial Q^n}{\partial\phi}\right\}. \quad (4.1)$$

Since Q^n satisfies (2.8) we can show directly that this Φ_0^n satisfies (2.15). If we substitute for Φ_0^{n+1} from (4.1) into the Newman–Penrose constants for a spin-1 field (2.17) and then integrate by parts we find

$${}_1G^{n+1}_{n+1,m} = \frac{1}{\sqrt{2}}\int Q^{n+1} W_{n,m}\, d\omega, \quad (4.2)$$

with

$$W_{n,m} = \left[\frac{1}{2}\sin\theta\frac{\partial^2}{\partial\theta^2} - \frac{1}{2\sin\theta}\frac{\partial^2}{\partial\phi^2} - \left(n+\frac{1}{2}\right)\cos\theta\frac{\partial}{\partial\theta}\right.$$
$$- i(n+1)\cot\theta\frac{\partial}{\partial\phi} + i\frac{\partial^2}{\partial\theta\,\partial\phi} - \frac{(n+\frac{3}{2})}{\sin\theta}$$
$$\left. - \frac{1}{2}n(n+1)\sin\theta\right]{}_1\overline{Y}_{n+1,m}. \quad (4.3)$$

Using the complex conjugate of Eq. (B4), giving in particular ${}_1\overline{Y}_{n+1,m}$ in terms of $\overline{Y}_{n+1,m}$, and taking into consideration Eq. (2.9) satisfied by $Y_{l,m}$, we can simplify (4.3) to read

$$W_{n,m} = [(n+1)(n+2)]^{1/2}$$
$$\times\left\{\sin\theta\frac{\partial\overline{Y}_{n+1,m}}{\partial\theta} - (n+1)\cos\theta\,\overline{Y}_{n+1,m}\right\}. \quad (4.4)$$

The well-known recurrence formula for spherical harmonics (cf. Abramowitz and Stegun,[12] p. 334)

$$\sin\theta\frac{\partial\overline{Y}_{n+1,m}}{\partial\theta} - (n+1)\cos\theta\,\overline{Y}_{n+1,m}$$
$$= (n+m+1)\overline{Y}_{n,m}, \quad (4.5)$$

allows us to simplify (4.4) and upon substitution of $W_{n,m}$ back into (4.2) we obtain

$${}_1G^{n+1}_{n+1,m} = (n+m+1)[(n+1)(n+2)/2]^{1/2}\,{}_0G^{n+1}_{n,m}, \quad (4.6)$$

with ${}_0G^{n+1}_{n,m}$ given by (2.12). We thus arrive at the remarkably simple conclusion: *For $n = 0,1,2,\ldots$, $|m|\leqslant n$, the conserved quantities associated with a zero-rest-mass spin-1 field on Minkowskian space-time are inherited from the conserved quantities associated with the complex wave function from which the spin-1 is constructed.* We note from (4.4) that $W_{n,m} = 0$ for $|m| = n+1$, so that

$${}_1G^{n+1}_{n+1,\pm(n+1)} = 0 \quad (n\geqslant 0). \quad (4.7)$$

V. GENERALIZATION TO HIGHER SPIN FIELDS

In addition to the operators $\eth$, $\overline{\eth}$ we find it convenient to introduce two further operators

$$D^{n,s} \equiv -\frac{1}{\sqrt{2}}\left\{(n+1)\sin\theta\frac{\partial}{\partial\theta} - is\frac{\partial}{\partial\phi}\right.$$
$$\left. + (n+1)^2\cos\theta\right\}, \quad (5.1a)$$

$$\overline{D}^{n,s} \equiv -\frac{1}{\sqrt{2}}\left\{n\sin\theta\frac{\partial}{\partial\theta} + is\frac{\partial}{\partial\phi} - n^2\cos\theta\right\}, \quad (5.1b)$$

which act on spin-weight-s functions. These have the following effect on the spin-s spherical harmonics:

$$D^{n,s}\,{}_sY_{n,m} = (n-m+1)\{(n-s+1)(n+s+1)/2\}^{1/2}$$
$$\times{}_sY_{n+1,m}, \quad (5.2a)$$

$$\overline{D}^{n,s}\,{}_sY_{n,m} = -(n+m)\{(n-s)(n+s)/2\}^{1/2}\,{}_sY_{n-1,m}, \quad (5.2b)$$

$$\overline{D}^{n+1,s}D^{n,s}\,{}_sY_{n,m} = -\tfrac{1}{2}\{(n+1)^2-m^2\}$$
$$\times\{(n+1)^2-s^2\}\,{}_sY_{n,m}. \quad (5.2c)$$

One can show, using (5.2a) and (5.2b), that if η is a spin-weight-s function then so are the functions $D^{n,s}\eta$ and $\overline{D}^{n,s}\eta$. Also (5.2c) is a disguised version of (B6). Further properties of $D^{n,s}$ and $\overline{D}^{n,s}$, essential to the argument that follows, are given in Appendix C.

Extensive calculations have suggested the following generalization of our results to spin $s\geqslant 1$ fields: given the spin-weight-0 functions Q^n introduced in (2.7) we define a sequence of functions $\{{}_sQ^n\}$, $s = 0,1,2,\ldots$ of spin weights $s = 0,1,2,\ldots$, respectively, to be substituted into (2.18), by the recurrence formula

$${}_sQ^n = -[(n+1)/\sqrt{2}]\,{}_{s-1}\dot{Q}^{n+1}\sin\theta$$
$$+ (\eth D^{n+s-1,s-1} - D^{n+s-1,s}\eth)\,{}_{s-1}Q^n, \quad (5.3)$$

for $s = 1,2,3,\ldots$, with

$${}_0Q^n = Q^n. \quad (5.4)$$

If we write out (5.3) explicitly we can see that ${}_1Q^n = \Phi_0^n$ given by (4.1). Also we can prove by induction on s, using the formulas (C1)–(C6) that ${}_sQ^n$ satisfies

$$-2(n+1)\,{}_s\dot{Q}^{n+1} = \overline{\eth}\,\eth\,{}_sQ^n + n(n+2s+1)\,{}_sQ^n. \quad (5.5)$$

Thus upon substitution of these ${}_sQ^n$ into (2.19) we confirm that ${}_sG^{n+1}_{n+s,m}$ are constants. Obviously one must verify, using spinors, that these are the Newman–Penrose constants for spin-s fields. To see how they are related to the Newman–Penrose constants for $s = 0$ we substitute for ${}_s\dot{Q}^{n+1}$ from (5.5) into (5.3) and then put ${}_sQ^{n+1}$ into (2.19). An exactly similar procedure to the spin-1 case then yields

$${}_sG^{n+1}_{n+s,m} = (n+m+s)\left\{\frac{(n+2s)(n+2s-1)}{2}\right\}^{1/2}$$
$$\times{}_{s-1}G^{n+1}_{n+s-1,m}, \quad (5.6)$$

for $s=0,1,2,\ldots$, $n=0,1,2,\ldots$, $|m|\leqslant n+s$. Thus, in general, *the conserved quantities* ${}_sG^{n+1}_{n+s,m}$, *for any positive integral value of s, are multiples of the conserved quantities* ${}_0G^{n+1}_{n,m}$ *associated with a complex wave function.*

ACKNOWLEDGMENTS

I wish to thank Professor I. Robinson for informative discussions on the matter discussed in Sec. III of this paper and Professor J. L. Synge for a stimulating correspondence on the same topic. I am grateful to Dr. J. D. McCrea for many helpful discussions during the course of the work.

APPENDIX A: BIVECTORS ON A NULL TETRAD

In terms of the coordinates (θ,ϕ,r,u) introduced in (2.2) define the vector fields

$$K=\frac{\partial}{\partial r},\quad L=\frac{1}{2}\frac{\partial}{\partial r}-\frac{\partial}{\partial u},$$
$$M=\frac{1}{r\sqrt{2}}\left(\frac{\partial}{\partial\theta}+\frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right),\tag{A1}$$
$$\overline{M}=\frac{1}{r\sqrt{2}}\left(\frac{\partial}{\partial\theta}-\frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right).$$

These vectors constitute a null tetrad. In the text we require the components of spin-1 zero-rest-mass fields on this null tetrad. These fields arise in our work in the form of components of tensors given in the coordinates $X^i=(x,y,z,t)$. Hence we require the components of the vector fields (A1) in the coordinates X^i. If these components are denoted K^i, L^i, M^i, $\overline{M}^i$, then they are given by

$$K^i=\cos^2\frac{\theta}{2}k^i+\sin^2\frac{\theta}{2}l^i+\frac{1}{2}e^{-i\phi}\sin\theta\, m^i+\frac{1}{2}e^{i\phi}\sin\theta\,\overline{m}^i,$$
$$L^i=\frac{1}{2}\sin^2\frac{\theta}{2}k^i+\frac{1}{2}\cos^2\frac{\theta}{2}l^i-\frac{1}{4}e^{-i\phi}\sin\theta\, m^i-\frac{1}{4}e^{i\phi}\sin\theta\,\overline{m}^i,$$
$$M^i=-\frac{1}{2\sqrt{2}}\sin\theta\, k^i+\frac{1}{2\sqrt{2}}\sin\theta\, l^i+\frac{1}{\sqrt{2}}e^{-i\phi}\cos^2\frac{\theta}{2}m^i-\frac{1}{\sqrt{2}}e^{i\phi}\sin^2\frac{\theta}{2}\overline{m}^i,\tag{A2}$$
$$\overline{M}^i=-\frac{1}{2\sqrt{2}}\sin\theta\, k^i+\frac{1}{2\sqrt{2}}\sin\theta\, l^i-\frac{1}{\sqrt{2}}e^{-i\phi}\sin^2\frac{\theta}{2}m^i+\frac{1}{\sqrt{2}}e^{i\phi}\cos^2\frac{\theta}{2}\overline{m}^i,$$

where

$$k^i\frac{\partial}{\partial X^i}=\frac{\partial}{\partial z}+\frac{\partial}{\partial t},\quad l^i\frac{\partial}{\partial X^i}=-\frac{\partial}{\partial z}+\frac{\partial}{\partial t},$$
$$m^i\frac{\partial}{\partial X^i}=\frac{\partial}{\partial x}+i\frac{\partial}{\partial y},\quad \overline{m}^i\frac{\partial}{\partial X^i}=\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}.\tag{A3}$$

A basis of complex (constant), anti-self-dual bivectors is given by

$$m_{ij}=m_ik_j-m_jk_i,\quad n_{ij}=\overline{m}_il_j-\overline{m}_jl_i,$$
$$l_{ij}=m_i\overline{m}_j-\overline{m}_im_j+l_ik_j-k_il_j.\tag{A4}$$

If

$$M_{ij}=M_iK_j-M_jK_i,\quad N_{ij}=\overline{M}_iL_j-\overline{M}_jL_i,$$
$$L_{ij}=M_i\overline{M}_j-\overline{M}_iM_j+L_iK_j-K_iL_j,\tag{A5}$$

then using (A2) and (A4) we find

$$\sqrt{2}M_{ij}=e^{-i\phi}\cos^2\frac{\theta}{2}m_{ij}+\frac{1}{2}\sin\theta\, l_{ij}-e^{i\phi}\sin^2\frac{\theta}{2}n_{ij},$$
$$2\sqrt{2}N_{ij}=-e^{-i\phi}\sin^2\frac{\theta}{2}m_{ij}+\frac{1}{2}\sin\theta\, l_{ij}+e^{i\phi}\cos^2\frac{\theta}{2}n_{ij},$$
$$2L_{ij}=-e^{-i\phi}\sin\theta m_{ij}+\cos\theta\, l_{ij}-e^{i\phi}\sin\theta n_{ij}.\tag{A6}$$

If $F_{ij}=-F_{ji}$ are the components of a real bivector in coordinates X^i then the complex, anti-self-dual bivector $F_{ij}+i{}^*F_{ij}$, where ${}^*F_{ij}=\frac{1}{2}\epsilon_{ijkl}F^{kl}$ (with ϵ_{ijkl} the Levi-Civita permutation symbol), can be written

$$F_{ij}+i{}^*F_{ij}=\phi_0n_{ij}+\phi_1l_{ij}+\phi_2m_{ij},$$
$$=\Phi_0N_{ij}+\Phi_1L_{ij}+\Phi_2M_{ij},\tag{A7}$$

where

$$\Phi_0=2\sqrt{2}e^{-i\phi}\cos^2\frac{\theta}{2}\phi_0+2\sqrt{2}\sin\theta\,\phi_1-2\sqrt{2}e^{i\phi}\sin^2\frac{\theta}{2}\phi_2,$$
$$\Phi_1=-e^{-i\phi}\sin\theta\,\phi_0+2\cos\theta\,\phi_1-e^{i\phi}\sin\theta\,\phi_2,\tag{A8}$$
$$\Phi_2=-\sqrt{2}e^{-i\phi}\sin^2\frac{\theta}{2}\phi_0+\sqrt{2}\sin\theta\,\phi_1+\sqrt{2}e^{i\phi}\cos^2\frac{\theta}{2}\phi_2.$$

APPENDIX B: SPIN-WEIGHTED FUNCTIONS

For convenience we list here some properties of spin-weighted functions which can be found in Newman and Penrose[2] and the other sources quoted in Sec. II.

If η is a spin-weight-s function on the unit two-sphere (s may be integral or half-integral), then the operator ð, which raises the spin weight by unity, is defined in terms of the polar coordinates θ, ϕ by

$$\eth\eta=-(\sin\theta)^s\left(\frac{\partial}{\partial\theta}+\frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)(\sin\theta)^{-s}\eta,\tag{B1}$$

while the operator $\overline{\eth}$, which lowers the spin weight by unity, is defined by

$$\overline{\eth}\eta=-(\sin\theta)^{-s}\left(\frac{\partial}{\partial\theta}-\frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)(\sin\theta)^s\eta.\tag{B2}$$

When these operators act on the spherical harmonics $Y_{l,m}(\theta,\phi)$ ($l=0,1,2,\ldots;\ |m|\leqslant l,\ m\in\mathbb{Z}$) they produce the (integral) spin-s spherical harmonics

$${}_sY_{l,m}=\begin{cases}\sqrt{\dfrac{(l-s)!}{(l+s)!}}\,\eth^sY_{l,m} & (0\leqslant s\leqslant l),\\ (-1)^s\sqrt{\dfrac{(l+s)!}{(l-s)!}}\,\overline{\eth}^{-s}Y_{l,m} & (-l\leqslant s\leqslant 0).\end{cases}\tag{B3}$$

These functions form a complete, orthonormal set for each integral s. A spin-weight-s function defined on the unit two-sphere can be expanded in spin-s spherical harmonics (see Goldberg *et al.*[4]).

Finally we note that the ${}_sY_{l,m}$ satisfy

$$\eth\, {}_sY_{l,m} = [(l-s)(l+s+1)]^{1/2}\, {}_{s+1}Y_{l,m}, \tag{B4}$$

$$\bar{\eth}\, {}_sY_{l,m} = -\sqrt{(l+s)(l+s+1)}\, {}_{s-1}Y_{l,m}, \tag{B5}$$

and

$$\bar{\eth}\eth\, {}_sY_{l,m} = -(l-s)(l+s+1)\, {}_sY_{l,m}. \tag{B6}$$

APPENDIX C: FORMULAS INVOLVING $D^{n,s}$ AND $\eth$

When acting on a spin-weight-s function, $D^{n,s}$ and $\eth$ satisfy the following equations:

$$\bar{\eth}(\eth D^{n,s} - D^{n,s+1}\eth) = (\eth D^{n,s-1} - D^{n,s}\eth)\bar{\eth} - 2D^{n,s}, \tag{C1}$$

$$\eth(\eth D^{n,s} - D^{n,s+1}\eth) = (\eth D^{n,s+1} - D^{n,s+2}\eth)\eth, \tag{C2}$$

$$(\eth D^{n,s} - D^{n,s+1}\eth)\bar{\eth}\eth = \bar{\eth}\eth(\eth D^{n,s} - D^{n,s+1}\eth) + 2D^{n,s+1}\eth, \tag{C3}$$

$$(\eth D^{n+1,s} - D^{n+1,s+1}\eth) - (\eth D^{n,s} - D^{n,s+1}\eth) = \frac{1}{\sqrt{2}}\left\{\cos\theta\frac{\partial}{\partial\theta} - (2n+3)\sin\theta - \frac{s}{\sin\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right\}, \tag{C4}$$

$$D^{n,s+1}\eth = (n+1)(\sin\theta/\sqrt{2})\{\bar{\eth}\eth + (n-s)(n+s+1)\} + (n-s)(\eth D^{n,s} - D^{n,s+1}\eth). \tag{C5}$$

Finally, if η is a spin-weight-s function, then

$$\bar{\eth}\eth(\eta\sin\theta) - \sin\theta\,\bar{\eth}\eth\eta = 2\cos\theta\frac{\partial\eta}{\partial\theta} + 2i\cot\theta\frac{\partial\eta}{\partial\phi} - \frac{2s}{\sin\theta}\eta + 2s\eta\sin\theta. \tag{C6}$$

[1]E. T. Newman and R. Penrose, Phys. Rev. Lett. **15**, 231 (1965).
[2]E. T. Newman and R. Penrose, Proc. R. Soc. London Ser. A **305**, 175 (1968).
[3]J. N. Goldberg, J. Math. Phys. **8**, 2161 (1967).
[4]J. N. Goldberg, A. J. Macfarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, J. Math. Phys. **8**, 2155 (1967).
[5]R. Penrose and W. Rindler, *Spinors and Space-Time* (Cambridge U. P., Cambridge, 1984), Vol. 1.
[6]I. Robinson (private communication).
[7]J. L. Synge, *Relativity: The Special Theory* (North-Holland, London, 1965).
[8]E. T. Whittaker, *A History of the Theories of Aether and Electricity* (*Classical Theories*) (Nelson, London, 1958).
[9]J. L. Synge (private communication).
[10]R. Penrose, Proc. R. Soc. London Ser. A **284**, 159 (1965).
[11]J. M. Stewart, Proc. R. Soc. London Ser. A **367**, 527 (1979).
[12]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965).

A search for bilinear equations passing Hirota's three-soliton condition. II. mKdV-type bilinear equations

Jarmo Hietarinta
Department of Physical Sciences, University of Turku, 20500 Turku, Finland

(Received 5 January 1987; accepted for publication 29 April 1987)

In this paper (second in a series) [for part I, see J. Math. Phys. **30**, 1732 (1987)] the search for bilinear equations having three-soliton solutions continues. This time pairs of bilinear equations of the type $P_1(D_x,D_t)F\cdot G=0$, $P_2(D_x,D_t)F\cdot G=0$, where P_1 is an odd polynomial and P_2 is quadratic, are considered. The main results are the following new bilinear systems: $P_1=aD_x^7+bD_x^5+D_x^2D_t+D_y$, $P_2=D_x^2$; $P_1=aD_x^3+bD_t^3+D_y$, $P_2=D_xD_t$; and $P_1=D_xD_tD_y+aD_x+bD_t$, $P_2=D_xD_t$. In addition to these, several models with linear dispersion manifolds were obtained, as before.

I. INTRODUCTION

In a recent paper[1] (hereafter referred to as I) we searched for bilinear equations of the type $P(D_x,D_t)f\cdot f=0$ which also had three-soliton solutions (3SS) of the standard Hirota form. In this paper we report the results of a similar kind of search for *pairs* of bilinear equations on *two* functions

$$P_1(D_x,D_t)F\cdot G=0,\quad P_2(D_x,D_t)F\cdot G=0. \tag{1}$$

In this paper we furthermore assume that P_1 is an odd function in its variables and P_2 is a quadratic even function. There are also other possibilities that should be investigated, but even under the present assumptions we do obtain new results.

The best known nonlinear evolution equation whose bilinear form is of type (1) is, no doubt, the modified Korteweg–de Vries (mKdV) equation

$$v_t-6v^2v_x+v_{xxx}=0. \tag{2}$$

If one introduces a new function w by $v=w_x$ it is also possible to cast (2) in the following form [integrate (2) once and set the constant to zero]:

$$w_t-2w_x^3+w_{xxx}=0. \tag{3}$$

Hirota introduces next a dependent variable transformation[2–4] by expressing w in terms of two new functions f and g:

$$w=-2i\arctan(g/f). \tag{4}$$

When this is substituted to (2) the resulting equations vanish, provided that

$$(D_t+D_x^3)f\cdot g=0,\quad D_x^2(f\cdot f+g\cdot g)=0. \tag{5}$$

This is not yet in the form (1), but if we define $F=f-ig$ and $G=f+ig$ so that

$$w=\log(F/G), \tag{4$'$}$$

then (5) becomes

$$(D_t+D_x^3)F\cdot G=0, \tag{6a}$$

$$D_x^2F\cdot G=0. \tag{6b}$$

For some of the higher-order members of the mKdV hierarchy one can also write a bilinear formulation like (6). For the even equation one always takes (6b) but the odd equation can be[5–7]

$$(D_t+D_x^5)F\cdot G=0, \tag{7}$$

for which, corresponding to (2), we get[6,7]

$$v_t+[v_{xxxx}-10vv_x^2-10v^2v_{xx}+6v^5]_x=0, \tag{8}$$

and

$$(D_t+D_x^7)F\cdot G=0, \tag{9}$$

$$v_t+[v_{xxxxxx}-14v^2v_{xxxx}-56vv_xv_{xxx}-42v_{xx}^2v -70v_{xx}v_x^2+70v_{xx}v^4+140v_x^2v^3-20v^7]_x=0. \tag{10}$$

There are also other equations that can be written as a pair of bilinear equations, but which are not of the particular form studied here. For example, the sine–Gordon equation

$$u_{xt}=\sin u \tag{11}$$

goes over the[3,8]

$$\begin{aligned}&(D_xD_t-1)(F\cdot F-G\cdot G)=0,\\ &D_xD_tF\cdot G=0,\end{aligned} \tag{12}$$

when one uses the substitution $u=4\arctan[(F-G)/(F+G)]$. Now the first equation is not as in (1).

Bilinear equations of type (1) appear also as Bäcklund transformations for single equations.[9] Usually the P_i's in these cases contain both odd and even terms and therefore some of their reductions fit into the form studied here.

Ito[10] tested numerically the three-soliton condition (3SC) for bilinear equations of type (1), where P_1 is odd and P_2 even in the variables. His results with a quadratic P_2 were (6b) together with (6a) or (7) or (8), and the following pairs:

$$\begin{aligned}&(D_x^2D_t-D_x-D_t)F\cdot G=0,\\ &D_xD_tF\cdot G=0\end{aligned} \tag{13}$$

and

$$\begin{aligned}&(D_x^3+aD_t+bD_x+cD_y)F\cdot G=0,\\ &D_xD_tF\cdot G=0.\end{aligned} \tag{14}$$

Ramani[7] studied also the existence of 3SS and found that systems having 3SS also passed the Painlevé test. He obtained in this way the above results and some others with higher degree P_2.

Pairs of bilinear equations on two functions are also contained in the Kyoto-school approach to bilinear formalism.[11] For example, (6) can be obtained as a two-reduction[12]

 0022-2488/87/092094-08$02.50

from the first two equations of the modified KP hierarchy and similarly (14) from the modified DKP hierarchy.

II. THE *N*-SOLITON CONDITION

We will now discuss the conditions for the P_i's that guarantee the existence of standard form N-soliton solutions for (1). We assume that the P_i's have definite parity and are without constant terms:

$$P_i(-X,-Y) = (-1)^i P_i(X,Y),\qquad P_2(0,0) = 0. \tag{15}$$

For the one-soliton solution we take the ansatz

$$F = 1 + ae^{n},\quad G = 1 + be^{n}, \tag{16}$$

with

$$n = px + \Omega t + m. \tag{17}$$

(For generalizations, see Ref. 13.) When we recall the definition of the operators D_x and D_t,

$$D_x^n D_t^m F\cdot G = (\partial_x - \partial_{x'})^n(\partial_t - \partial_{t'})^m \times F(x,t)G(x',t')\,|_{x'=x,\,t'=t} \tag{18}$$

it is easy to show that (16) is a solution of (1) if $b = -a$ and if p and Ω are related by

$$P_1(p,\Omega) = 0. \tag{19}$$

(The possibility $b = a$ leads to a trivial result.) Thus for the pair (1) it is the *odd* equation that gives the *dispersion relation*. The overall factor a can be absorbed into m and we take subsequently $a = 1$, $b = -1$.

For the two-soliton solution we take the natural generalization

$$F = 1 + e^{n_1} + e^{n_2} + Ae^{n_1+n_2},\qquad G = 1 - e^{n_1} - e^{n_2} + Be^{n_1+n_2}, \tag{20}$$

where the n_i's are constructed as before in (17) with the parameters (p_i,Ω_i) satisfying (19). When this is substituted into (1) we find that (20) is a solution if A and B are given by

$$A = B = P_2(p_1 - p_2,\Omega_1 - \Omega_2)/P_2(p_1 + p_2,\Omega_1 + \Omega_2). \tag{21}$$

This procedure does not work for those pairs of parameters for which $P_2(p_1 + p_2,\Omega_1 + \Omega_2) = 0$. For such "resonances"[14] one has to use different methods to obtain the two-soliton solution. We will not go into the details of this but just note that multiplication of F and G by $\exp(-n_2)$ [which does not change w of (4')] can be used to convert n_2 to $-n_2$, which helps sometimes. Systems where (19) factors into linear factors are especially prone to resonances. In the following we will not take these complications into account, and therefore the final results guarantee three-soliton solutions only if a nonresonating set of three parameter pairs can be chosen.

With the above caveats we find that a pair of equations of type (1) with (15) has two-soliton solutions given by (20), and that the dispersion relation (19) is given by the odd equation.

The existence of higher-soliton solutions is again a rather restrictive condition. For the general N-soliton solution (20) is generalized as[10]

$$F = \sum_{\mu=0,1} \exp\Big[\sum_{i>j}^{(N)} A_{ij}\mu_i\mu_j + \sum_{i=1}^{N}\mu_i\Big(n_i + \frac{i\pi}{2}\Big)\Big],\qquad G = \sum_{\mu=0,1} \exp\Big[\sum_{i>j}^{(N)} A_{ij}\mu_i\mu_j + \sum_{i=1}^{N}\mu_i\Big(n_i - \frac{i\pi}{2}\Big)\Big]. \tag{22}$$

Here the n_i's are given as before in (17), the extra $i\pi/2$ takes care of the sign reversal, cf. (20). The parameters p_i and Ω_i in n_i must again satisfy the dispersion relation (19) while the constants A_{ij} are determined as in (21):

$$\exp A_{ij} = -P_2(p_i - p_j,\Omega_i - \Omega_j)/P_2(p_i + p_j,\Omega_i + \Omega_j). \tag{23}$$

If $N > 2$ one obtains two kinds of conditions for the polynomils P_i (Ref. 10):

$$S_{\rm odd}[P_1,P_2,n] \equiv \sum_{\sigma=\pm1} P_1\Big(\sum_{i=1}^{n}\sigma_i p_i, \sum_{i=1}^{n}\sigma_i\Omega_i\Big)\sin\Big(\sum_{i=1}^{n}\sigma_i\frac{\pi}{2}\Big) \times\prod_{i>j}^{(n)} P_2(\sigma_i p_i - \sigma_j p_j, \sigma_i\Omega_i - \sigma_j\Omega_j) = 0, \tag{24}$$

$$S_{\rm even}[P_2,n] \equiv \sum_{\sigma=\pm1} P_2\Big(\sum_{i=1}^{n}\sigma_i p_i, \sum_{i=1}^{n}\sigma_i\Omega_i\Big)\cos\Big(\sum_{i=1}^{n}\sigma_i\frac{\pi}{2}\Big) \times\prod_{i>j}^{(n)} P_2(\sigma_i p_i - \sigma_j p_j, \sigma_i\Omega_i - \sigma_j\Omega_j) = 0, \tag{25}$$

for each $n = 1,\ldots,N$, and for all p_k, Ω_k subject to (19). We note first of all that both equations are satisfied identically for $n = 1,2$. Also $S_{\rm odd}$ is identically zero if n is even, and conversely $S_{\rm even}\equiv 0$ when n is odd. For other combinations Eqs. (24) and (25) are not identities and the first nontrivial conditions are on $S_{\rm odd}[P_1,P_2,3]$ and $S_{\rm even}[P_2,4]$. Note also that $S_{\rm even}$ does not depend on P_1, it enters only implicitly through (19).

In this paper we consider only cases for which P_2 is quadratic. In that case $S_{\rm even}[P_2,4]$ vanishes identically. Thus if the system has 3SS then it will automatically have 4SS as well. Our attention focuses therefore on the condition $S_{\rm odd}[P_1,P_2,3] = 0$. To study this equation we use the same algebraic geometrical approach as in I. We recall that the 3SC can be formulated as

$$S_{\rm odd}[P_1,P_2,3]\in I(V_{P_1,3}), \tag{26}$$

where

$$V_{P_1,3} = \{(x,t)\in\mathbb{C}^6 \,|\, P_1(x_i,t_i) = 0,\ \forall i = 1,2,3\}, \tag{27}$$

and $I(V)$ is the ideal of those polynomials that vanish on V.

The polynomial P_1 can be factorized as

$$P_1(X,T) = \prod_j Q_j(X,T)^{n_j}, \tag{28}$$

where each Q_j is a monic irreducible polynomial. (The possible overall constant has been dropped.) For the purpose of classification it is often useful to group the irreducible factors according to their multiplicity as

$$P_1(X,T) = \prod_{i=1}^{s} P_{1,i}(X,T)^{n_i}, \quad n_i > n_j \quad \text{for } i > j. \tag{29}$$

We introduce the definition of $\surd P_1$:

$$\surd P_1(X,T) = \Pi_j Q_j(X,T) = \Pi_i P_{1,i}(X,T). \tag{30}$$

Then, according to the theorem in Sec. III A of I we can express the 3SC as follows: The bilinear system (1) has three-soliton solutions of type (22) if we can write

$$S_{\rm odd}[P_1,P_2,3] = \surd P_1(X_1,T_1)A + \surd P_1(X_2,T_2)B + \surd P_1(X_3,T_3)C, \tag{31}$$

where A, B, and C are some polynomials in the variables X_1, X_2, X_3, T_1, T_2, and T_3.

To implement the above formulation in a practical way suitable for computer algebra systems we introduce a consistent ordering in the set of monomials $X^m T^n$ (for example, first by $m+n$ and then by m among those with the same $m+n$). Then for each $i = 1,2,3$ we take the leading monomial of $\surd P_1(x_i,t_i)$, let us call it $\surd M_1(x_i,t_i)$, and replace it everywhere in $S_{\rm odd}$ by $\surd M_1(x_i,t_i) - \surd P_1(x_i,t_i)$. It is easy to see that $S_{\rm odd}$ vanishes under this rewriting rule iff it can be written as in (31). Note that the rewriting rule decreases the order of $S_{\rm odd}$ and therefore eventually $\surd M_1$'s can no longer be extracted and the procedure terminates. In REDUCE the rewrite rule is accomplished by a LET statement.[15] For the LET statement it is important that the polynomial $\surd M_1$ be monic.

III. RESULTS

The property of passing the 3SC is clearly invariant under any linear change of variables. To avoid needless repetition we should therefore choose the variables in a systematic way. The first step in the classification process goes according to P_2. As a quadratic function in its variables P_2 has two linear factors, so if these factors are identical we transform P_2 to X^2, while if they are different we transform P_2 to XT. In the former case we still have the freedom of defining the T variable, and we will use it by defining the first factor that is not identical to X as T. In the latter case we can at most scale or reflect the variables.

A. $P_2 = X^2$

The classification process is given in Table I. During it we must keep in mind the full factorization (29); we have used square brackets to isolate the factors that belong to different multiplicities as described in (29).

Since the rewrite rule decreases the order we should as the first step find the acceptable leading monomials. We found the following possibilities:

(1.1): $P_1 = X^{2N+1}$, $\surd P_1 = X^K$, where $0 < K \leqslant 2N+1$ when $N \leqslant 3$ and $0 < K \leqslant 2[(N+1)/3] + 3$ otherwise;

(1.2): $P_1 = X^{2M+1}T^{2M}$, $\surd P_1 = XT$;

(1.3): $P_1 = X^2 T^{2M+1}$, $\surd P_1 = X^K T$, $K = 1,2$;

(1.4): $P_1 = T^{2M+1}$, $\surd P_1 = T^L$, $0 < L \leqslant [2M/3] + 1$.

(Here $[a]$ = integer part of a.) We have checked these results explicitly up to total degree 21 and conjecture that they are true for general N and M. Note the interesting dip in the maximum value of K in (1.1) at $N = 4$.

In the first column of Table I we have given the type of the monomial; (n,m) stands for $X^n T^m$. In the second column we have written the combinations that fit into one of the above forms (1.1)–(1.4) and which can arise as the leading monomial of some P_1 as described in (29) and (30). Combinations that would lead to contradiction with the definitions of the leading monomial or P_1 are not included.

Next we should consider the possible homogeneous generalizations of these. For case (1.1) there are no such possibilities because the first factor different from X would be renamed as T and therefore would be contained in the other cases. Case (1.2) would in principle generalize to

TABLE I. Classification of P_1 when $P_2 = X^2$.

Type	Leading monomial	Possible generalizations	Accepted final result	Generalizations with Y
(3.0)				
	$[X]^3$		$[X]^3$	
	$[X^3]$	$[X^3+T]$	$[X^3+T]$	
(2.1)				
	$[X]^2[T]$		$[X]^2[T]$	
	$[X^2T]$	$[X^2T+aX+bT]$	$[X^2T+aX+bT]$	$[X^2T+Y]$
(1.2)				
	$[T]^2[X]$		$[T]^2[X]$	
(0.3)				
	$[T]^3$		$[T]^3$	
(5.0)				
	$[X]^5$		$[X]^5$	
	$[X]^3[X^2]$	$[X]^3[X^2-1]$	$[X]^3[X^2-1]$	
	$[X]^2[X^3]$	$[X]^2[X^3+T]$	$[X^2][X^3+T]$	
	$[X^2]^2[X]$	$[X^2-1]^2[X]$	$[X^2-1]^2[X]$	
	$[X^5]$	$[X^5+R_3+R_1]$	$[X^5+X^2T+aX+bT]$	$[X^5+X^2T+Y]$
(4.1)				
(3.2)				
	$[X]^3[T]^2$		$[X]^3[T]^2$	

TABLE I. (*Continued.*)

Type	Leading monomial	Possible generalizations	Accepted final result	Generalizations with Y
(2.3)				
	$[X]^2[T]^3$		$[X]^2[T]^3$	
	$[X^2][T]^3$	$[X^2-1][T]^3$	$[X^2-1][T]^3$	
(1.4)				
	$[X][T]^4$		$[X][T]^4$	
(0.5)				
	$[T]^5$		$[T]^5$	
(7.0)				
	$[X]^7$		$[X]^7$	
	$[X]^5[X^2]$	$[X]^5[X^2-1]$	$[X]^5[X^2-1]$	
	$[X]^4[X^3]$	$[X]^4[X^3+T]$	$\cdots$	
	$[X]^3[X^2]^2$	$[X]^3[X^2-1]^2$	$[X]^3[X^2-1]^2$	
	$[X]^3[X^4]$	$[X]^3[X^4+R_2+R_0]$	$[X]^3[X^4+aX^2+1]$	
	$[X^2]^3[X]$	$[X^2-1]^3[X]$	$[X^2-1]^3[X]$	
	$[X]^2[X^5]$	$[X]^2[X^5+R_3+R_1]$	$[X]^2[X^5+aX^3+T]$	
	$[X^2]^2[X^3]$	$[X^2-1]^2[X^3+R_1]$	$[X^2-1]^2[X^3+aX]$	
	$[X^3]^2[X]$	$[X^3+T]^2[X]$	$\cdots$	
	$[X^7]$	$[X^7+R_5+R_3+R_1]$	$[X^7+cX^5+X^2T+aX+bT]$	$[X^7+cX^5+X^2T+Y]$
(6.1)				
(5.2)				
	$[X]^5[T]^2$		$[X]^5[T]^2$	
(4.3)				
(3.4)				
	$[X]^3[T]^4$		$[X]^3[T]^4$	
(2.5)				
	$[X]^2[T]^5$		$[X^2][T]^5$	
	$[X^2][T]^5$	$[X^2-1][T]^5$	$[X^2-1][T]^5$	
(1.6)				
	$[X][T]^6$		$[X][T]^6$	
(0.7)				
	$[T]^7$		$[T]^7$	
	$[T]^5[T^2]$	$[T]^5[T^2-1]$	$\cdots$	
	$[T]^3[T^2]^2$	$[T]^3[T^2-1]^2$	$\cdots$	
	$[T][T^2]^3$	$[T][T^2-1]^3$	$\cdots$	
(9.0)				
	$[X]^9$		$[X]^9$	
	$[X]^7[X^2]$	$[X]^7[X^2+1]$	$[X]^7[X^2+1]$	
	$[X]^6[X^3]$	$[X]^6[X^3+T]$	$\cdots$	
	$[X]^5[X^2]^2$	$[X]^5[X^2+1]^2$	$[X]^5[X^2+1]^2$	
	$[X]^5[X^4]$	$[X]^5[X^4+R_2+R_0]$	$[X]^5[X^4+aX^2+1]$	
	$[X][X^2]^4$	$[X][X^2+1]^4$	$[X][X^2+1]^4$	
	$[X^3]^3$	$[X^3+T]^3$	$\cdots$	
		$[X^3+X]^3$	$[X^3+X]^3$	
	$[X]^3[X^3]^2$	$[X]^3[X^3+T]^2$	$\cdots$	
	$[X]^3[X^2]^2[X^2]$	$[X]^3[X^2+1]^2[X^2+a]$	$[X]^3[X^2+1][X^2+a]$	
	$[X^2]^3[X^3]$	$[X^2+1]^3[X^3+T]$	$\cdots$	
		$[X^2+1]^3[X^3+aX]$	$[X^2+1]^3[X^3+aX]$	
	$[X^4]^2[X]$	$[X^4+R_2+R_0]^2[X]$	$[X^4+aX^2+1]^2[X]$	
(8.1)				
(7.2)				
	$[X]^7[T]^2$		$[X]^7[T]^2$	
(6.3)				
(5.4)				
	$[X]^5[T]^4$		$[X]^5[T]^4$	
(4.5)				
(3.6)				
	$[X]^3[T]^6$		$[X]^3[T]^6$	
(2.7)				
	$[X]^2[T]^7$		$[X]^2[T]^7$	
	$[X^2][T]^7$	$[X^2-1][T]^7$	$[X^2-1][T]^7$	
(1.8)				
	$[X][T]^8$		$[X][T]^8$	
(0.9)				
	$[T]^9$		$[T]^9$	
	$[T]^7[T^2]$	$[T]^7[T^2-1]$	$\cdots$	
	$[T]^5[T^2]^2$	$[T]^5[T^2-1]^2$	$\cdots$	
	$[T][T^2]^4$	$[T][T^2-1]^4$	$\cdots$	
	$[T^3]^3$	$[T^3+R_1]^3$	$\cdots$	

TABLE II. Classification of P_1 when $P_2 = XT$.

Type	Leading monomial	Possible homogeneous generalization	Allowed homog. generalizations	Possible nonhomog. generalizations	Allowed nonhomog. generalizations	Generalizations with Y
(3,0)						
	$[X]^3$	$[X-aT]^3$	$[X]^3$		$[X]^3$	
			$[X-T]^3$		$[X-T]^3$	$[Y]^3$
	$[X]^2[X]$	$[X-aT]^2[X-bT]$	$[X-T]^2[X+T]$		$[X-T]^2[X+T]$	
	$[X^3]$	$[X^3+aX^2T+bXT^2+cT^3]$	$[X^3]$	$[X^3+aX+bT]$	$[X^3+aX+bT]$	$[X^3+Y]$
			$[X^3-T^3]$	$[X^3-T^3+aX+bT]$	$[X^3-T^3+aX+bT]$	$[X^3-T^3+Y]$
(2.1)						
	$[X]^2[T]$	$[X-aT]^2[T]$	$[X]^2[T]$		$[X]^2[T]$	
	$[X^2T]$	$[(X^2+aXT+bT^2)T]$	$[X^2T]$	$[X^2T+aX+bT]$	$[X^2T+aX+bT]$	
			$[XT(X-T)]$	$[XT(X-T)+aX+bT]$	$[XT(X-T)+aX+bT]$	$[XTY+aX+bT]$
(5.0)						
	$[X]^5$	$[X-aT]^5$	$[X^5]$		$[X]^5$	
			$[X-T]^5$		$[X-T]^5$	$[Y]^5$
	$[X]^4[X]$	$[X-aT]^4[X-bT]$	$\cdots$			
	$[X]^3[X]^2$	$[X-aT]^3[X-bT]^2$	$[X-T]^3[X+T]^2$		$[X-T]^3[X+T]^2$	
	$[X]^3[X^2]$	$[X-aT]^3[X^2+bXT+cT^2]$	$\cdots$			
	$[X][X^2]^2$	$[X-aT][X^2+bXT+cT^2]^2$	$\cdots$			
	$[X]^2[X^3]$	$[X-aT]^2[X^3+bX^2T+cXT^2+dT^3]$	$\cdots$			
(4.1)						
	$[X]^4[T]$	$[X-aT]^4[T]$	$[X]^4[T]$		$[X]^4[T]$	
	$[X]^3[XT]$	$[X]^3[(X-aT)T]$	$[X]^3[XT]$	$[X]^3[XT-1]$	$\cdots$	
		$[X-T]^3[(X-a)T]$	$[X-T]^3[XT]$	$[X-T]^3[XT+a]$	$[X-T]^3[XT]$	$[Y]^3[XT]$
	$[X^2]^2[T]$	$[X^2+aXT+bT^2]^2[T]$	$[X^2]^2[T]$	$[X^2-1]^2[T]$	$[X^2-1]^2[T]$	
(3.2)						
	$[X]^3[T]^2$	$[X-aT]^3[T]^2$	$[X^3][T]^2$		$[X]^3[T]^2$	
	$[X]^3[T^2]$	$[X-aT]^3[T^2]$	$[X]^3[T^2]$	$[X]^3[T^2-1]$	$[X]^3[T^2-1]$	
(7.0)						
	$[X]^7$	$[X-aT]^7$	$[X]^7$		$[X]^7$	
			$[X-T]^7$		$[X-T]^7$	$[Y]^7$
	$[X]^6[X]$	$[X]^6[X-T]$	$\cdots$			
		$[X-T]^6[X-bT]$	$\cdots$			
	$[X]^5[X]^2$	$[X]^5[X-T]^2$	$\cdots$			
		$[X-T]^5[X-bT]^2$	$[X-T]^5[X+T]^2$		$[X-T]^5[X+T]^2$	
	$[X]^5[X^2]$	$[X]^5[X^2+\cdots]$	$[X]^5[X^2]$	$[X]^5[X^2-1]$	$[X]^5[X^2-1]$	
		$[X-T]^5[X^2+\cdots]$	$[X-T]^5[(X-T)^2]$	$[X-T]^5[(X-T)^2-1]$	$\cdots$	
	$[X]^4[X]^3$	$[X]^4[X-T]^3$	$\cdots$			
		$[X-T]^4[X-bT]^3$	$\cdots$			
	$[X]^4[X]^2[X]$	$[X]^4[X-T]^2[X-bT]$	$\cdots$			
		$[X-T]^4[X-bT]^2[X-cT]$	$\cdots$			
	$[X]^4[X^3]$	$[X]^4[X^3+\cdots]$	$[X]^4[X^3]$	$[X]^4[X^3+aX+bT]$	$\cdots$	
		$[X-T]^4[X^3+\cdots]$	$\cdots$			
	$[X^2]^3[X]$	$[X^2+\cdots]^3[X]$	$[X^2]^3[X]$	$[X^2-1]^3[X]$	$[X^2-1]^3[X]$	
		$[X^2+\cdots]^3[X-T]$	$[(X-T)^2]^3[X-T]$	$[(X-T)^2-1]^3[X-T]$	$\cdots$	
	$[X]^3[X^2]^2$	$[X]^3[X^2+\cdots]^2$	$[X]^3[X^2]^2$	$[X]^3[X^2-1]^2$	$[X]^3[X^2-1]^2$	
		$[X-T]^3[X^2+\cdots]^2$	$[X-T]^3[(X-T)^2]^2$	$[X-T]^3[(X-T)^2-1]^2$	$\cdots$	
	$[X]^3[X]^2[X^2]$	$[X]^3[X-T]^2[X^2+\cdots]$	$\cdots$			
		$[X-T]^3[X-bT]^2[X^2+\cdots]$	$\cdots$			

TABLE II. (*Continued.*)

Type	Leading monomial	Possible homog. generalization	Allowed homog. generalizations	Possible nonhomog. generalizations	Allowed nonhomog. generalizations	Generalizations with Y
	$[X]^3[X^4]$	$[X^3][X^4+\cdots]$	$[X]^3[X^4]$	$[X]^3[X^4+R_2+R_0]$	$[X]^3[X^4-1]$	
		$[X-T]^3[X^4+\cdots]$	$\cdots$			
	$[X^3]^2[X]$	$[X^3+\cdots]^2[X]$	$[X^3]^2[X]$	$[X^3+R_1]^2[X]$	$\cdots$	
		$[X^3+\cdots]^2[X-T]$	$\cdots$			
	$[X^2]^2[X^3]$	$[X^2+\cdots]^2[X^3+\cdots]$	$[X^2]^2[X^3]$	$[X^2-1]^2[X^3+R_1]$	$[X^2-9]^2[X(X^2+4)]$	
(6.1)						
	$[X]^6[T]$	$[X-aT]^6[T]$	$[X]^6[T]$		$[X]^6[T]$	
	$[X]^5[XT]$		$[X]^5[XT]$	$[X]^5[XT-1]$	$\cdots$	
		$[X]^5[(X-T)T]$	$\cdots$			
		$[X-T]^5[(X-aT)T]$	$[X-T]^5[XT]$	$[X-T]^5[XT-a]$	$[X-T]^5[XT]$	$[Y]^5[XT]$
	$[X]^4[X]^2[T]$	$[X]^4[X-T]^2[T]$	$\cdots$			
		$[X-T]^4[X-aT]^2[T]$	$\cdots$			
	$[X^2]^3[T]$	$[X^2+aXT+bT^2]^3[T]$	$[X^2]^3[T]$	$[X^2-1]^3[T]$	$[X^2-1]^3[T]$	
(5.2)						
	$[X]^5[T]^2$	$[X-aT]^5[T]^2$	$[X]^5[T]^2$		$[X]^5[T]^2$	
	$[X]^5[T^2]$	$[X-aT]^5[T^2]$	$[X]^5[T^2]$	$[X]^5[T^2-1]$	$[X]^5[T^2-1]$	
(4.3)						
	$[X]^4[T]^3$	$[X-aT]^4[T]^3$	$[X]^4[T]^3$		$[X]^4[T]^3$	
	$[X][TX]^3$	$[X][T(X-aT)]^3$	$[X][TX]^3$	$[X][TX-1]^3$	$\cdots$	
		$[X-T][T(X-aT)]^3$	$\cdots$			
	$[X^2]^2[T]^3$	$[X^2+aXT+bT^2]^2[T]^3$	$[X^2]^2[T]^3$	$[X^2-1]^2[T]^3$	$[X^2-1]^2[T]^3$	

$P_1 = (X - aT)T^{2n}$, $\sqrt{P_1} = (X - aT)T$, but this did not pass the 3SC except for $a = 0$; a similar result holds for case (1.3). For case (1.4) no X variable is acceptable because then T^{2n+1} would not be the leading term. Since no homogeneous generalizations were possible we did not include this step in the table.

In the next column we have written the possible nonhomogeneous generalizations, and after that the cases that do pass the 3SC. We considered only those generalizations which kept the polynomial P_1 odd. The classification in the table was first made in two dimensions (X and T), but when a completely arbitrary term of type $aX + bT$ was allowed we tried next the same system with $aX + bT$ replaced by Y. The accepted higher-dimensional generalizations are given in the last column.

The results can be combined as follows.

(i) First we have the genuinely nonlinear system

$$P_1 = aX^7 + bX^5 + X^2T + Y, \quad P_2 = X^2. \tag{32}$$

(If here a and/or b is nonzero they/it can be scaled to 1.) As particular cases (32) contains (6) ($a = 0$, $b = 0$, $T \to X$, $Y \to T$), (7) ($a = 0$, $b = 1$, $T \to 0$, $Y \to T$), and (9) ($a = 1$, $b = 0$, $T \to 0$, $Y \to T$). The full generality of (32) is a new result.

(ii) In addition to the above nonlinear result we obtain, as in I, several sequences that have a dispersion manifold that consists of lines.

(1.A) up to degree 9 any polynomial in X subject to (1.1) above.

(1.B) $X^{2M+1}T^{2N}$ [also X^2T^{2N+1}, but it can be regarded as a special case of the one below].

(1.C) $(X^2 - 1)T^{2M+1}$.

(1.D) T^{2M+1}.

Probably these systems pass the 3SC for any value of N and M; we have checked them to the order 15–20.

B. $P_2 = XT$

In this case the X and T variables are already fixed up to scaling and reflections. As a consequence the classification process is longer; it is described in Table II.

As before we start with the leading monomial of P_1. Using reflections we may assume that it is X^nT^m, $n + m$ odd, with $n > m$. Our results are

(2.1): $P_1 = X^{2N+1}$, $\sqrt{P_1} = X^K$, $0 < K \leqslant [2N/3] + 3$;

(2.2): $P_1 = X^{2N+1}T^{2M}$, $\sqrt{P_1} = XT^L$, $L = 1,2$;

(2.3): $P_1 = X^{2N}T^{2M+1}$, $\sqrt{P_1} = X^KT$, $K = 1,2$.

In this case there are nontrivial homogeneous generalizations. The full analysis of all possibilities is rather tedious. In the third column in Table II we have given the various possibilities, and we have furthermore divided the cases so that one nonzero constant is scaled to -1. Some of the cases turned out to be quite demanding even for the computer, the worst one being $[X - T]^4[X - bT]^2[X - cT]$. (In this case we may assume that $b \neq 1 \neq c$ and $b \neq c$, but then we have no cases that pass 3SC.)

After obtaining the homogeneous results we continue by trying the possible nonhomogeneous generalizations indicated in column 3, the accepted results are given in column 4. It contains the full results in two dimensions.

To get the higher-dimensional results we reconsidered those cases that had a term $aX + bT$ (or its scaled version $X - T$) and tested whether the term could be replaced with Y. [Note that it is not necessary to try in this way polynomials $(X - T)^N(X + T)^M$ for they cannot be scaled to a form where the free constants a and b appear only through $aX + bT$.]

The full results can be classified as follows.

(iii) The nonlinear results can be combined into the following two cases:

$$P_1 = aX^3 + bT^3 + Y, \quad P_2 = XT \tag{33}$$

and

$$P_1 = XTY + aX + bT, \quad P_2 = XT. \tag{34}$$

Note the interesting symmetry: X and T are treated on equal footing while Y is the extra variable. As a special case (34) contains (13) ($Y \to X$, $a = b = -1$) and (33) can be reduced to (14) ($a = 1, b = 0$, $Y \to aT + bX + cY$). Again the full generality is a new result.

(iv) The cases with linear dispersion manifolds are the following.

(2.A) We have several one-dimensional results. At degree 5 any polynomial is acceptable but already at degree 7 we get conditions. In general a degree 7 polynomial can be written as $X[X^2 - a][X^2 - b][X^2 - c]$. We find that a, b, and c cannot all be different from each other and zero, and even for the remaining free pair we get conditions. The acceptable results at this degree are $X^3[X^4 - 1]$ and $X[X^2 - 9]^2[X^2 + 4]$, where we have scaled the result to a convenient form.

(2.B) Y^{2N+1}.

(2.C) $[X - T]^{2N+1}[X + T]^2$.

(2.D) X^NT^M.

(2.E) $Y^{2N+1}XT$.

(2.F) $[X^2 - 1]^NT^{2M+1}$.

We checked these results to a rather high degree and conjecture that they hold for arbitrary N, M.

IV. CONCLUSIONS

In this paper we have studied a subclass of pairs of bilinear equations. The main results are given in Eqs. (32)–(34). Equation (33) can be considered as an extension of the original mKdV equation (6). The results (33) and (34) are interesting because of their X–T symmetry, which their previously known special cases did not possess. All of these results are three dimensional.

As far as the models with linear dispersion relations are concerned the pattern follows the one obtained in I. The meaning of these sequences of polynomials is still open.

One interesting question that can be raised is how, or whether, these models fit in the Kac–Moody algebra approach to bilinear equations.[11] An integral part of that approach has been the assignment of different weights to the variables so that the polynomials P_i are weighted homogeneous. For the models found here such assignments are difficult, because we have either the same variable appearing with different powers, or two different variables appearing with the same power.

Since new results were obtainable even with the present restrictions on P_2, it seems likely that many other models can still be found.

ACKNOWLEDGMENTS

I would like to thank A. Ramani for discussions. The computations were done with IBM 3033 at University of Turku. I would like to thank the staff of the Computing Center for their help.

[1]J. Hietarinta, J. Math. Phys. **30**, 1732 (1987).
[2]R. Hirota, J. Phys. Soc. Jpn. **33**, 1456 (1972).
[3]R. Hirota, Prog. Theor. Phys. **52**, 1498 (1974).
[4]R. Hirota, in *Bäcklund Transformations, the Inverse Scattering Method, and Their Applications*, edited by R. M. Miura (Springer, Berlin, 1976), p. 40.
[5]A. Sym, Phys. Lett. A **65**, 383 (1978).
[6]Y. Matsuno, J. Phys. Soc. Jpn. **49**, 787 (1986).
[7]A. Ramani, "Inverse scattering, ordinary differential equations of Painlevé type and Hirota's bilinear formalism," preprint L.P.T.H.E Université Paris—Sud, Orsay, 1980.
[8]R. Hirota, J. Phys. Soc. Jpn. **33**, 1459 (1972).
[9]R. Hirota, in *Solitons*, edited by R. K. Bullough and P. J. Caudrey (Springer, Berlin, 1980), p. 157.
[10]M. Ito, J. Phys. Soc. Jpn. **49**, 771 (1980).
[11]M. Jimbo and T. Miwa, Publ. RIMS, Kyoto Univ. **19**, 943 (1983).
[12]See Ref. 11, p. 962.
[13]A. Nakamura, J. Phys. Soc. Jpn. **48**, 1365 (1980).
[14]R. Hirota and M. Ito, J. Phys. Soc. Jpn. **52**, 744 (1983).
[15]A. C. Hearn, *REDUCE User's Manual Version 3.2* (Rand, Santa Monica, 1985), Publ. CP78, Rev. 4/85.

Non-Abelian Berry's phase, accidental degeneracy, and angular momentum

Jan Segert[a)]
Physics Department, Princeton University, Princeton, New Jersey 08544

(Received 23 December 1986; accepted for publication 13 May 1987)

The non-Abelian Berry's phase effect for a family of operators $H_0 + \mathbf{k}\cdot\mathbf{V}$ is considered, where H_0 is rotationally invariant, $\mathbf{V}$ is a vector operator, and $\mathbf{k}$ varies over the unit vectors in $\mathbb{R}^3$. The parameter space is the two-sphere. The time evolution in the adiabatic limit is given by a connnection on a fiber bundle over the two-sphere. All connections consistent with the rotational symmetry are classified, and the time evolution is explicitly calculated for a nondegenerate Hamiltonian, as well as for a Hamiltonian with a double degeneracy. In the nondegenerate case, the connection is uniquely determined by the symmetry. In the doubly degenerate case, the connection is in some instances not determined by the symmetry. The case of approximate degeneracy is also discussed. A possible experimental test of this effect using optical pumping is described elsewhere.

I. INTRODUCTION

The time evolution of a quantum system governed by a slowly changing Hamiltonian can be analyzed geometrically. This was first noted in the work of Berry,[1] and elaborated by Simon.[2] The case of degenerate Hamiltonians, otherwise known as non-Abelian Berry's phase, was considered by Wilczek and Zee.[3] The main results of these papers is that the time evolution is given by the parallel transport in a certain connection on a fiber bundle over the parameter space. We shall call this the Berry connection. In this paper, we shall analyze a certain class of such systems. The systems described here can be experimentally realized in a number of ways. In another paper,[4] we propose an optical pumping experiment using Pb^{208} atoms. The analysis of such experiments requires the results of the work presented below.

We shall consider a family of Hamiltonians parametrized by the unit vectors $\mathbf{k}$ in $\mathbb{R}^3$. To each unit vector $\mathbf{k}$ we associate a Hamiltonian $H_\mathbf{k} = H_0 + \mathbf{k}\cdot\mathbf{V}$, where H_0 is a rotationally invariant operator, and $\mathbf{V}$ is a vector operator. This generalizes the example of Berry of an atom in a magnetic field.[1] For example, the Hamiltonian of an atom in constant colinear electric and magnetic fields is of this form.[4] The results of this paper are, in fact, valid for a more general class of Hamiltonians than those of the form $H_\mathbf{k} = H_0 + \mathbf{k}\cdot\mathbf{V}$. This class, which will be fully characterized, includes such potentially interesting examples as the quadratic Zeeman Hamiltonian for an atom in a strong field.

The rotation group acts transitively on the parameter space, since any unit vector can be rotated into any other unit vector. Thus all the operators in the family are unitarily equivalent and have the same spectrum. Consider first the nondegenerate case. We choose an eigenvalue E_n. Following Berry and Simon, we form a line bundle F_n over parameter space, where the fiber over a point $\mathbf{k}$ of S^2 is the space of scalar multiples of eigenvectors of $H_\mathbf{k}$ with eigenvalue E_n. If the spectrum has a double degeneracy at E_n, we must consider a two-dimensional complex vector bundle over S^2, which we will also denote by F_n. The rotation group acts on the bundle F_n. The Berry connection must be invariant under this action. We shall topologically classify the bundles F_n in the case of no degeneracy, and in the case of a double degeneracy. This degeneracy is an accidental, where we use the term to mean that the symmetry group of the Hamiltonian does not act irreducibly on the degenerate subspace.[5] We will determine the action of the rotation group on the bundles F_n. We then classify all the invariant connections. The classification depends only on the angular momentum quantum numbers of the state or pair of degenerate states under consideration.

In the nondegenerate case, we find that there is only one rotationally invariant connection, which must then be the Berry connection. In the degenerate case, the results are more interesting. If the two degenerate states have consecutive $\mathbf{J}\cdot\mathbf{k}$ quantum numbers, e.g., $m_1 - m_2 = \pm 1$, then there is more than one invariant connection. In fact, there is an uncountable infinity of invariant connections. If the quantum numbers of the two degenerate states do not satisfy this condition, then the connection is unique, and we can immediately identify the Berry connection as the unique invariant connection. If the two degenerate states satisfy the above condition, we have to do some additional work to identify the Berry connection.

The work of Berry and Simon can be divided into two statements: (1) The evolution of a system under adiabatic change of parameters is determined by some connection on a certain bundle over parameter space. (2) The Schrödinger equation uniquely determines this connection. Statement (1) follows almost immediately from the definition of a connection.[4,6] When the connection is already uniquely determined by the symmetry group, as is often the case, then statement (2) contains no information. But if there is more than one connection compatible with the symmetry group, statement (2) has nontrivial content, and is subject to experimental confirmation. We comment here on some recent related papers. The work of Chiao and Wu[7] interprets the polarization rotation of light in a bent optical fiber as a manifestation of Abelian Berry's phase. The Berry connection for this system is uniquely determined by the symmetry. This is a consequence of the results of the present paper. The polarization rotation can alternately be interpreted as a purely classical

[a)] Address after September 1987: Department of Physics, California Institute of Technology, Pasadena, California 91125.

 0022-2488/87/092102-13$02.50

effect,[6,8] and the result is again uniquely determined by the rotational invariance.[6] The results of the classical and quantum derivations are identical, and have been confirmed by the experiments of Tomita and Chiao.[9] This experiment cannot test statement (2), since the symmetry alone is enough to determine the result. Moody, Shapere, and Wilczek[10] have proposed a spin-resonance experiment to look for manifestations of Abelian Berry's phase. The family of Hamiltonians for this system is of the form $H_{\mathbf{k}} = H_0 + \mathbf{k}\cdot\mathbf{V}$, with each $H_{\mathbf{k}}$ nondegenerate. Thus by the results of the present paper, the rotational symmetry again uniquely determines the result, and this experiment also cannot test statement (2). In the same paper, Moody, Shapere, and Wilczek discuss the effects of non-Abelian Berry's phase on the energy levels of a diatomic molecule. This is formally similar to the systems studied in the present paper. We consider a system governed by a Hamiltonian of the above form, $H_0 + \mathbf{k}\cdot\mathbf{V}$. When this Hamiltonian has accidental degeneracies of a certain type, the Berry connection cannot be determined just from the rotational invariance. Then statement (2) is a nontrivial prediction, which is experimentally testable.[4] We shall find the Berry connection explicitly using Simon's prescription. We also study the case of approximate degeneracy. This is important for the analysis of experimental tests.

In Sec. II we fix conventions and collect standard angular momentum results. In Sec. III we review the construction of vector bundles using projection operators. In Sec. IV we study vector bundles over S^2. In Sec. V we determine the topological structure of eigenvector bundles over S^2. In Sec. VI, we determine the action of the rotation group on the eigenvector bundles, and present the results of the invariant connection classification. In Sec. VII we calculate parallel transport in the Berry connection. Section VIII concerns the case of approximate degeneracy. Section IX consists of concluding remarks. The Appendix contains the technical details of the invariant connection classification.

II. CONVENTIONS

We now specify some of our conventions. The group of rotations in $\mathbb{R}^3$ is SO(3), the group of orthogonal real matrices of dimension 3, with positive determinant. A rotation can be specified by giving the axis of the rotation, a unit vector $\mathbf{k}$, and the angle of rotation in the positive sense α. We define the matrices

$$L_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$
$$L_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.1}$$

which act on $\mathbb{R}^3$. The SO(3) matrix corresponding to the rotation is $\exp(\alpha\mathbf{k}\cdot\mathbf{L})$. Recall that the group SU(2) is a double cover of SO(3). The Lie algebra of SU(2) is isomorphic to the Lie algebra of SO(3). We shall take as our basis of the Lie algebra of SU(2) the matrices $((i/2)\sigma_x, (i/2)\sigma_y, (i/2)\sigma_z)$, with the usual convention for Pauli matrices,

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.2}$$

The obvious Lie algebra isomorphism is $L_x \mapsto (i/2)\sigma_x$, $L_y \mapsto (i/2)\sigma_y$, $L_z \mapsto (i/2)\sigma_z$. We shall call SU(2) the rotation group. This lets us treat half-integral spins as well as integral spins. We denote by (iJ_x, iJ_y, iJ_z) the standard basis of a representation of the Lie algebra of SU(2), where the generators (J_x, J_y, J_z) are self-adjoint. The corresponding unitary representations of SU(2) will be denoted by U. We shall denote by R the vector representation of SU(2), the real three-dimensional representation generated by the Lie algebra of SO(3), using the isomorphism $(i/2)\sigma_x \mapsto L_x$, $(i/2)\sigma_y \mapsto L_y$, $(i/2)\sigma_z \mapsto L_z$. This is, in fact, the adjoint representation of SO(3), and is unitarily equivalent, but not identical, to the usual spin-1 representation of SU(2). We list for convenience some standard results,

$$[J_x, J_y] = iJ_z, \quad [J_y, J_z] = iJ_x, \quad [J_z, J_x] = iJ_y, \tag{2.3}$$

$$\mathbf{J}^2|j,m\rangle = j(j+1)|j,m\rangle, \quad J_z|j,m\rangle = m|j,m\rangle, \tag{2.4}$$

$$\langle j,m|J_z|j,m-1\rangle = \tfrac{1}{2}[(j+m)(j-m+1)]^{1/2}, \tag{2.5}$$

$$\langle j,m|J_y|j,m-1\rangle = (i/2)[(j+m)(j-m+1)]^{1/2}. \tag{2.6}$$

A vector operator $\mathbf{V}$ is a triplet of operators (V_x, V_y, V_z) satisfying[5,11]

$$[J_x, V_y] = iV_z, \quad [J_y, V_z] = iV_x, \quad [J_z, V_x] = iV_y. \tag{2.7}$$

We let $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$ denote the standard orthonormal basis of $\mathbb{R}^3$. We shall use the standard conventions for spherical coordinates in $\mathbb{R}^3$,

$$z = r\cos(\theta), \quad x = r\sin(\theta)\cos(\phi), \quad y = r\sin(\theta)\sin(\phi). \tag{2.8}$$

On the unit sphere S^2, the south pole corresponds to the vector $-\hat{\mathbf{z}}$, the north pole to the vector $\hat{\mathbf{z}}$. By the equator of S^2 we shall mean the map $eq: S^1 \to S^2$ by $\phi \mapsto (\pi/2, \phi)$. We shall alternately represent a point on S^2 by $\mathbf{k}$ or (θ, ϕ).

III. MATHEMATICAL PREREQUISITES ON VECTOR BUNDLES

We discuss a characterization of complex vector bundles over a compact connected metric space M, for example a compact finite-dimensional smooth manifold. This characterization is appropriate for the bundles of eigenstates over parameter space. Consider the product space $M \times \mathscr{H}$, where $\mathscr{H}$ is a complex separable Hilbert space. The Hilbert space $\mathscr{H}$ may be finite or infinite dimensional.

Denote by $\mathscr{P}(\mathscr{H})$ the space of compact positive self-adjoint projection operators on $\mathscr{H}$. An operator P is in $\mathscr{P}(\mathscr{H})$ if and only if all the following conditions are satisfied: $PP = P$, $P^\dagger = P$, the range of P is finite dimensional, and $\langle x|P|x\rangle \geqslant 0$ for all $|x\rangle \in \mathscr{H}$. We give $\mathscr{P}(\mathscr{H})$ the operator norm topology. In this topology, $\mathscr{P}(\mathscr{H})$ consists of disjoint components $\mathscr{P}_n(\mathscr{H})$. Here $\mathscr{P}_n(\mathscr{H})$ consists of all projections whose range has dimension n. This follows easily from the spectral theorem for normal operators,[12] and the fact that the group of unitary operators, with the operator norm topology, is connected.

Consider a continuous map $f: M \to \mathscr{P}(\mathscr{H})$. The range of f must be contained in $\mathscr{P}_n(\mathscr{H})$ for some n. So for any two points $s,v \in M$, the dimensions of the projections $f(s), f(v) \in \mathscr{P}(\mathscr{H})$ coincide. Denote by r the range map r: $\mathscr{P}(\mathscr{H}) \to \mathscr{H}$. The range map assigns to a projection operator P the subspace of $\mathscr{H}$ which is not annihilated by P, the set of vectors $x \in \mathscr{H}$ satisfying $Px = x$. The composition $r \circ f$ maps M into subspaces of $\mathscr{H}$ of constant dimension. The graph $\Gamma \subset M \times \mathscr{H}$ of the map $r \circ f$ is a vector bundle A_f over M. The projection map $\pi: A_f \to M$ is the restriction of the projection $\pi: M \times \mathscr{H} \to M$. We shall now demonstrate the local product structure. Each point $s \in M$ has a neighborhood N_s which is contractible to the point s. The restriction $f: N_s \to \mathscr{P}(\mathscr{H})$ is then homotopic to the constant map c_s: $N_s \to f(s)$. The local product structure follows. Let g be another map from M to $\mathscr{P}(\mathscr{H})$. This map specifies another vector bundle A_g over M. The two vector bundles are topologically isomorphic if and only if the maps $f,g: M \to \mathscr{P}(\mathscr{H})$ are homotopic.

There is an analog for real vector bundles which may be more familiar. We shall construct the tangent bundle of the two-sphere S^2. Consider the usual embedding of S^2 as the unit sphere in $\mathbb{R}^3$. A unit vector $\mathbf{k}$ defines a point on S^2. The tangent space at the point $\mathbf{k}$ consists of all vectors orthogonal to $\mathbf{k}$. This is nothing but the range of the two-dimensioinal projection $P_\mathbf{k}$ which acts on a vector $\mathbf{a}$ by $P_\mathbf{k}\,\mathbf{a} = \mathbf{a} - (\mathbf{k}\cdot\mathbf{a})\mathbf{k}$.

In fact, any finite-dimensional complex vector bundle over M can be constructed this way if the dimension of $\mathscr{H}$ is sufficiently large. In particular, if the dimension of $\mathscr{H}$ is infinite, all finite-dimensional vector bundles can be constructed. This follows from a theorem of Serre and Swan.[13] In other words, $\mathscr{P}_n(\mathscr{H})$ is the universal classifying space for $\mathrm{U}(n)$, also known as the infinite Grassmanian. Any bundle over M is the pullback of the universal bundle over the classifying space.[14] In fact, there exists a universal connection on the universal bundle. This follows from the work of Narasimhan and Ramanan.[15] Any Hermitian connection on a complex vector bundle over M can be obtained as the pullback of the universal connection by the appropriate map of M into the classifying space. If M is the parameter space for a family of Hamiltonians, then the spectral projections[12] of a given eigenvalue give a map of M into the classifying space. The pullback of the universal bundle by this map defines the eigenstate bundle, which will be studied in detail in Sec. V. We can also use this map to pull back the universal connection. The resulting connection is simply the Berry connection. We shall not discuss these results in any detail, but we mention that the classification of rotationally invariant connections, given in the Appendix, can alternately be proved using these results. The classification then reduces to a question in representation theory, which is easily answered. We choose instead to present the direct proof in the Appendix, because it uses only elementary techniques.

IV. VECTOR BUNDLES ON THE TWO-SPHERE

We now study the manifold S^2. Define an atlas on S^2 in the following way. Denote by D_π the open disc in R^2 of radius π. Let (θ,ϕ) denote the polar coordinates on D_π, where θ is the radial coordinate, and ϕ the angular coordinate. Let D_A be the S^2 minus the south pole. Let u_A: $D_A \to D_\pi$ by $(\theta,\phi) \mapsto (\theta,\phi)$, where in the first instance (θ,ϕ) are the spherical coordinates on S^2, and in the second instance the polar coordinates on D_π. Similarly, let D_B be the sphere S^2 minus the north pole. We now define the map u_B: $D_B \to D_\pi$ by $(\theta,\phi) \mapsto (\pi - \theta,\phi)$. The choice of coordinates is a little unfortunate, but the map is clearly nonsingular. The intersection $D_A \cap D_B$ is the sphere minus both poles. The image of this region under u_A is the disc with the origin removed, $D_\pi - 0$. The image under u_B is also $D_\pi - 0$. We see that $u_B \circ (u_A)^{-1}$: $(D_\pi - 0) \to (D_\pi - 0)$, by $(\theta,\phi) \to (\pi - \theta,\phi)$.

We examine vector bundles on S^2. First note that each of the coordinate patches defined above is contractible. Recall that any vector bundle over a contractible space is trivial. All the topological information is contained in the transition function. The transition function is a map g: $D_A \cap D_B \to G$, where G is the structure group of the bundle, which is $\mathrm{U}(n)$ in our case. The toplogical structure of the bundle is determined by the homotopy class of the transition function g. Since $D_A \cap D_B$ is homotopic to the circle S^1 around the equator, we need only consider homotopy classes of maps $S^1 \to \mathrm{U}(n)$. This is simply the fundamental group $\pi_1(\mathrm{U}(n))$, which is known to be $\mathbb{Z}$, the integers. It is easy to verify that $\pi_1(\mathrm{U}(n)) = \mathbb{Z}$, corresponding to the winding number. Consider the determinant map det: $\mathrm{U}(n) \to \mathrm{U}(1)$. One can show that two maps g, g': $S^1 \to \mathrm{U}(n)$ are in the same homotopy class if and only if the composition maps $\det \circ g$, $\det \circ g'$: $S^1 \to \mathrm{U}(1)$ are in the same homotopy class, i.e., have the same winding number. One can then conclude that $\pi_2(\mathscr{P}_n(\mathscr{H})) = \mathbb{Z}$ if the dimension of $\mathscr{H}$ is sufficiently large.

The theory of characteristic classes[14,16,17] provides another approach to the classification of vector bundles. The integer invariant, determined above from the winding number, can also be expressed as $-C_1$, where C_1 is the first Chern number of the bundle. This is valid for complex bundles over S^2 of any dimension. The Chern–Weil theory[17,18] gives an expression for the first Chern number as the integral over S^2 of a differential form c_1 of degree 2. This differential form is explicitly constructed from the curvature of a connection on the bundle. Chern showed that although the form c_1 depends on the choice of connection, the integral C_1 of c_1 over S^2 does not, and is always an integer. In the case of a line bundle, the two-form c_1 is equal to the curvature two-form $\mathscr{R}$. For bundles of higher dimension, the curvature form $\mathscr{R}$ is matrix valued, and the Chern class c_1 is the trace of $\mathscr{R}$. The theory of Chern classes is not necessary for what follows, but it facilitates understanding of the rather abstract classification of invariant connections.

V. STRUCTURE OF THE EIGENSTATE BUNDLES

In this section we shall study the family of Hamiltonians $H_\mathbf{k} = H_0 + \mathbf{k}\cdot\mathbf{V}$, where $\mathbf{k}$ is a unit vector in $\mathbb{R}^3$, $\mathbf{V}$ is a vector operator, and H_0 is a rotationally invariant operator. The parameter space is the space of unit vectors in $\mathbb{R}^3$, or the two-dimensional sphere S^2. By definition, $[\mathbf{J},H_0] = 0$, so $[\mathbf{J},H_\mathbf{k}] = [\mathbf{J},\mathbf{k}\cdot\mathbf{V}] = i\mathbf{k}\times\mathbf{V}$. Obviously $H_\mathbf{k}$ is invariant under rotation about the $\mathbf{k}$ axis, $[\mathbf{k}\cdot\mathbf{J},H_\mathbf{k}] = 0$. In fact, this is the only property that we require. All the results of this paper actually hold for a more general class of Hamiltonians $H_{\mathbf{k}}$. Consider a single Hamiltonian $H_{\hat{\mathbf{z}}}$, satisfying $[J_z, H_{\hat{\mathbf{z}}}] = 0$. We obtain a family of Hamiltonians by letting the rotation group act on $H_{\hat{\mathbf{z}}}$. Let g be an element of the rotation group, and $U(g)$ the corresponding unitary operator on the Hilbert space. The Hamiltonian H_z is mapped to the Hamiltonian $H_{\mathbf{k}} = U(g) H_{\hat{\mathbf{z}}} U^{\dagger}(g)$, where the unit vector $\mathbf{k}$ is defined by $\mathbf{k} = R(g)\hat{\mathbf{z}}$. One easily checks that $[\mathbf{J}\cdot\mathbf{k}, H_{\mathbf{k}}] = 0$. We thus obtain a family $H_{\mathbf{k}}$ of Hamiltonians parametrized by the sphere S^2 of unit vectors in $\mathbb{R}^3$. It is easily verified that $H_{\mathbf{k}} = H_0 + \mathbf{k}\cdot\mathbf{V}$ is a family of this type. Since all the Hamiltonians in the family are unitarily related, they have the same spectrum. The representation U of SU(2) on the Hilbert space is generally reducible. We shall be interested only in the discrete spectrum of the Hamiltonian. If the Hamiltonian also has a continuous spectrum, we shall restrict attention to the subspace of Hilibert space spanned by the eigenvectors of the Hamiltonian. We will thus assume from now on that the Hamiltonian has only a discrete spectrum.

Assume that the Hamiltonian $H_{\mathbf{k}}$ is nondegenerate. We now construct a one-dimensional complex vector bundle over S^2 corresponding to each eigenvalue E_n of the Hamiltonian. Let $|\mathbf{k},n\rangle$ be a normalized eigenvector of $H_{\mathbf{k}}$, $H_{\mathbf{k}}|\mathbf{k},n\rangle = E_n|\mathbf{k},n\rangle$, $\langle\mathbf{k},n|\mathbf{k},n\rangle = 1$. This is unique only up to a phase. Construct the one-dimensional spectral projections $P_{\mathbf{k},n} = |\mathbf{k},n\rangle\langle\mathbf{k},n|$. This is determined uniquely, the phase ambiguity cancels. Consider the map $f_n: S^2 \to \mathscr{P}_1(\mathscr{H})$ by $\mathbf{k} \to P_{\mathbf{k},n}$. Each f_n defines a one-dimensional complex vector bundle F_n over S^2. Since $[\mathbf{J}\cdot\mathbf{k}, H_{\mathbf{k}}] = 0$, $|\mathbf{k},n\rangle$ is an eigenvector of $\mathbf{J}\cdot\mathbf{k}$. We denote by m the eigenvalue, $\mathbf{J}\cdot\mathbf{k}|\mathbf{k},n\rangle = m|\mathbf{k},n\rangle$. The span of the vectors $|\mathbf{k},n\rangle$ for all $\mathbf{k}$ defines a subspace $\mathscr{H}_n$ of $\mathscr{H}$. The rotation group acts irreducibly on $\mathscr{H}_n$. The dimension of an irreducible representation of SU(n) determines the spin quantum number j, by $\dim \mathscr{H}_n = 2j+1$. Each eigenvalue E_n of the Hamiltonian carries the angular momentum quantum numbers (j,m), determined in the above manner. Usually, the operator $\mathbf{V}\cdot\mathbf{k}$ can be considered a small perturbation to the rotationally invariant H_0, and is obvious how to define (j,m). In general, however, since $[\mathbf{J}^2, H_{\mathbf{k}}] \neq 0$, we need the above procedure to determine j.

We will now determine the topological structure of such a line bundle. We choose one of the eigenvalues E_n, and omit the n subscript in the following. Recall that every bundle is trivial when restricted to D_A or D_B. We will give a trivialization and find the transition function. We first fix a product structure for the restriction to D_A. Make a specific choice of phase for $|\hat{\mathbf{z}},n\rangle$, which will remain fixed from now on. We will follow the usual conventions and call this vector $|j,m\rangle$. Here (θ,ϕ) are the spherical coordinates on S^2, and $P_{(\theta,\phi)} = P_{\mathbf{k}}$, where $\mathbf{k} = \cos(\theta)\hat{\mathbf{z}} + \sin(\theta)\cos(\phi)\hat{\mathbf{x}} + \sin(\theta)\sin(\phi)\hat{\mathbf{y}}$. We define a local section of F_n over D_A as follows:

$$|A_{(\theta,\phi)}\rangle = \exp(-i\theta(-\sin(\phi)J_x + \cos(\phi)J_y))|j,m\rangle . \tag{5.1}$$

Similarly, we define a section $|B_{(\theta,\phi)}\rangle$ over D_B,

$$|B_{(\theta,\phi)}\rangle = \exp(i(\pi-\theta)(-\sin(\phi)J_x + \cos(\phi)J_y)) \times \exp(-i\pi J_x)|j,m\rangle . \tag{5.2}$$

One can easily check that these are indeed local sections of the bundle by verifying $P_{(\theta,\phi)}|A_{(\theta,\phi)}\rangle = |A_{(\theta,\phi)}\rangle$, and, likewise, for $|B_{(\theta,\phi)}\rangle$. A global section does not generally exist, as we will see. We now find the transition function relative to this trivialization. The transition function g_{BA} satisfies $g_{BA}(\theta,\phi)|A_{(\theta,\phi)}\rangle = |B_{(\theta,\phi)}\rangle$. This can be expressed as $g_{BA}(\theta,\phi) = |B_{(\theta,\phi)}\rangle\langle A_{(\theta,\phi)}|$. The projection $P_{(\theta,\phi)}$ acts as the identity operator, so

$$\begin{aligned} g_{BA}(\theta,\phi) &= P_{(\theta,\phi)} g_{BA}(\theta,\phi) \\ &= |A_{(\theta,\phi)}\rangle\langle A_{(\theta,\phi)}|B_{(\theta,\phi)}\rangle\langle A_{(\theta,\phi)}| \\ &= \langle A_{(\theta,\phi)}|B_{(\theta,\phi)}\rangle P_{(\theta,\phi)} . \end{aligned} \tag{5.3}$$

We need to compute the matrix element $\langle A_{(\theta,\phi)}|B_{(\theta,\phi)}\rangle$,

$$\langle A_{(\theta,\phi)}| = \langle j,m|\exp(i\theta(-\sin(\phi)J_x + \cos(\phi)J_y)) , \tag{5.4}$$

$$|B_{(\theta,\phi)}\rangle = \exp(i(\pi-\theta)(-\sin(\phi)J_x + \cos(\phi)J_y)) \times \exp(-i\pi J_x)|j,m\rangle , \tag{5.5}$$

$$\langle A_{(\theta,\phi)}|B_{(\theta,\phi)}\rangle = \langle j,m|\exp(i\pi(-\sin(\phi)J_x + \cos(\phi)J_y)) \times \exp(-i\pi J_x)|j,m\rangle , \tag{5.6}$$

$$\langle A_{(\theta,\phi)}|B_{(\theta,\phi)}\rangle = \langle j,m|\exp(-2i\phi J_z)|j,m\rangle .$$

We then find

$$g_{BA}(\theta,\phi) = \exp(-i2m\phi)P_{(\theta,\phi)} . \tag{5.7}$$

We need to compute the winding number of the map of the equator into U(1), $g_{BA}: S^1 \to \mathrm{U}(1)$. The winding number is clearly $-2m$.

We now consider a Hamiltonian for which the eigenvalue E_n is doubly degenerate. Now the bundle F_n is two dimensional, given by a map $f_n: S^2 \to \mathscr{P}_2(\mathscr{H})$. It is clear that the bundle F_n is just the direct sum of two one-dimensional bundles with the appropriate quantum numbers. The subspace $\mathscr{H}_n$ spanned by the eigenvectors need not be an irreducible subspace under the action of SU(2). If it is irreducible, we shall say the two vectors are in the same j multiplet. We again omit the n index.

Since $[\mathbf{J}\cdot\mathbf{k}, H_{\mathbf{k}}] = 0$, we can choose normalized eigenvectors $|\mathbf{k},1\rangle$, $|\mathbf{k},2\rangle$ satisfying $H_{\mathbf{k}}|\mathbf{k},i\rangle = E_n|\mathbf{k},i\rangle$, $\mathbf{J}\cdot\mathbf{k}|\mathbf{k},i\rangle = m_i|\mathbf{k},i\rangle$, and $\langle\mathbf{k},1|\mathbf{k},2\rangle = 0$. Define $P_{\mathbf{k},1} = |\mathbf{k},1\rangle\langle\mathbf{k},1|$, $P_{\mathbf{k},2} = |\mathbf{k},2\rangle\langle\mathbf{k},2|$, $P_{(\theta,\phi)} = P_{\mathbf{k}} = P_{\mathbf{k},1} + P_{\mathbf{k},2}$. The rest is analogous to the one-dimensional case. We can find the transition function for the bundle exactly as above. The transition function g_{BA} is in this case a map into U(2), given by

$$g_{BA}(\theta,\phi) = \exp(-i2m_1\phi)P_{(\theta,\phi),1} + \exp(-i2m_2\phi)P_{(\theta,\phi),2} . \tag{5.8}$$

We could also note that the bundle is the direct sum of two line bundles, and obtain the U(2) matrix g_{BA} by imbedding $\mathrm{U}(1)\times\mathrm{U}(1)$ diagonally in U(2). It is easy to see that the result is the same. The topological classification requires determining the winding number of the map $\det\circ g_{BA}$: $S^1 \to \mathrm{U}(1)$. This is easily seen to be equal to $-2(m_1+m_2)$. This is also clear from the fact that the Chern class of a sum of bundles is the sum of Chern classes.

VI. CLASSIFICATION OF ROTATIONALLY INVARIANT CONNECTIONS

We now discuss the action of the rotation group on the bundles defined above, and the classification of the connections which are invariant under this action. In the one-dimensional case, most of the results can be deduced from the rotational invariance of the curvature and the Chern–Gauss–Bonnet theorem. In the two-dimensional case, we appeal to a generalization of a theorem of Wang.[19] The details are discussed in the Appendix.

We discuss now the action of the rotation group on the eigenvector bundles. We first consider the one-dimensional complex vector bundles corresponding to a nondegenerate eigenstate with quantum numbers (j,m). The fiber over a point $\mathbf{k}\in S^2$ is a one-dimensional subspace of a finite-dimensional complex vector space, the irreducible subspace of the Hilbert space under the action of the rotation group. Recall that such an irreducible subspace has dimension $(2j+1)$. The entire bundle space is the $(2j+1)$-dimensional complex vector space, the fiber is a one-dimensional complex vector space, and the base space is S^2. The structure group is U(1). The action of SU(2) on the bundle space is the action of the corresponding irreducible representation. The action induced on the base space is the usual action of the rotation group on S^2. The discussion for the two-dimensional vector bundles corresponding to a doubly degenerate eigenstate is similar. Let $(j_i < m_i)$, $i = 1, 2$ be the quantum numbers of the two degenerate states. The bundle is just the direct sum of two one-dimensional bundles. The fiber is now a two-dimensional complex vector space, the base space is again S^2, and the entire bundle space is a $(2j_1 + 2j_2 + 2)$-dimensional vector space, the sum of the bundle spaces of the two corresponding one-dimensional bundles. The structure group is $U(1)\times U(1)$. The action of the rotation group on the vector bundle is clear. It is straightforward to deduce the action of the rotation group on the corresponding principal bundles. We remark that the action of the rotation group is transitive on the U(1) principle bundle corresponding to the nondegenerate eigenstate, but is not transitive on the $U(1)\times U(1)$ principal bundle corresponding to the doubly degenerate eigenstate.

We consider the one-dimensional case. Choose an eigenvalue E_n of the Hamiltonian, and denote by (j,m) the corresponding angular momentum quantum numbers. Construct the bundle F_n as above. Consider a connection $\mathscr{A}$ on F_n, and let $\mathscr{R}$ denote its curvature. The curvature is a two-form taking values in the Lie algebra of U(1), which we identify with $\mathbb{R}$. The curvature is then an ordinary two-form on S^2. The Chern–Gauss–Bonnet theorem states that the integral of $2\pi\mathscr{R}$ over S^2 is equal to C_1, the first Chern number of the bundle F_n. We have determined above that $C_1 = 2m$. A rotationally invariant connection has a rotationally invariant curvature form. The only rotationally invariant two-forms on S^2 are constant multiples of the area two-form η, $\eta = \sin(\theta)\, d\theta\wedge d\phi$ in spherical coordinates. The integral of η over S^2 is simply the area of S^2, which is 4π. Thus we find $\mathscr{R} = m\eta$.

The curvature is, in this case, sufficient to determine holonomy. Consider a region $\Omega\subset S^2$, such that the boundary $\gamma = \partial\Omega$ is a smooth simple closed curve on S^2. Then the holonomy for the path γ is simply the integral of $\mathscr{R}$ over Ω. This follows from the definition of curvature as the holonomy of a small path. We recover Berry's result that the holonomy for γ is equal to m times the area of Ω. The curvature is not quite sufficient to determine the connection. There are connections with the same curvature which are not rotationally invariant. There is, however, a unique rotationally invariant connection. One could show this explicitly by noting that in this case two connections with the same curvature are related by a principal bundle automorphism, or gauge transformation. A rotationally invariant connection remains invariant under gauge transformation if and only if the gauge transformation commutes with the action of the rotation group. One may verify that there are no nontrivial gauge transformations commuting with the rotation group action. We will instead rely on the classification results of the Appendix to prove the uniqueness of the invariant connection.

The two-dimensional case is more complicated. The curvature $\mathscr{R}$ is now a matrix-valued two-form. It is much more complicated to determine all such two-forms compatible with rotational invariance. Even if we could determine the curvature form, the holonomy is not easy to find. In the U(1) case we had a formula relating the holonomy for γ to the integral of $\mathscr{R}$ over Ω. There is no analogous formula for the U(2) case. The holonomy may be different for two paths γ, γ' which enclose, respectively, regions Ω, Ω' of the same area, but different shape. We must then appeal to the methods of the Appendix.

We now present the results of the invariant connection classification. In the one-dimensional case, we find there is a unique invariant connection. In the two-dimensional case, the results depend on the quantum numbers of the two states under consideration. Let (j_i,m_i) be the angular momentum quantum numbers of the two degenerate states. Since each one-dimensional bundle has a unique invariant connection, the direct sum of the bundles must have at least one invariant connection, corresponding to the sum of the connections on the line bundles. We shall say that a connection splits if it is a sum of connections on each of the line bundles. The holonomy of a connection which splits is $U(1)\times U(1)\subset U(2)$. It is clear that there is only one invariant connection which splits, since the invariant connections on the line bundles are unique. The results of the classification are the following. If $m_1 - m_2 \neq \pm 1$, then there is a unique invariant connection. This connection must then split. If $m_1 - m_2 = \pm 1$, then there is a two-parameter family of invariant connections, indexed by $(\alpha,\beta)\in\mathbb{R}^2$. The connection corresponding to $\alpha = \beta = 0$ splits. No other connection splits. There is now a nontrivial subgroup of the gauge transformations which commutes with the action of the rotation group. We find that two invariant connections (α,β) and (α',β') are gauge equivalent if and only if $\alpha^2 + \beta^2 = \alpha'^2 + \beta'^2$. The connection (α,β) is thus equivalent to a connection of the form $(\sqrt{\alpha^2+\beta^2},0)$. Up to gauge equivalence, the set of invariant connections can then be parametrized by $\mathbb{R}^+$, the non-negative real numbers.

This classification has important consequences. In the one-dimensional case, we do not need to do any work to find the Berry connection. The Berry connection is just the unique invariant connection. In the two-dimensional case, we can immediately determine the Berry connection if $m_1 - m_2 \neq \pm 1$. The Berry connection is again the unique invariant connection. When $m_1 - m_2 = \pm 1$, we have to actually do some work to find the Berry connection. This is unfortunate, but not tragic. We shall do this in the following section.

VII. THE BERRY CONNECTION

In this section we determine parallel transport in the Berry connection. We use Simon's prescription,[2] generalized by Wilczek and Zee[3] to the degenerate case. We do the two-dimensional case for arbitrary (j_i, m_i). Recall that the split connection is unique for any bundle. We shall find that if $j_1 \neq j_2$, then the Berry connection splits for any m_1, m_2. For $m_1 - m_2 \neq \pm 1$, this follows from the classification, but if $m_1 - m_2 = \pm 1$, this is a nontrivial result. Using this result and the classification, we see that the only possible case where the Berry connection might not split is $j_1 = j_2$, $m_1 - m_2 = \pm 1$. We will, in fact, find that if the two eigenvectors are in the same j multiplet, and $m_1 - m_2 = \pm 1$, then the Berry connection never splits. Note that $j_1 = j_2$ is a weaker statement than saying that the states are in the same j multiplet.

Choose an eigenvalue E_n, and define $P_{\mathbf{k}}$ as before. Suppose we have an eigenstate of $H_{\mathbf{k}}$ which we wish to transport to an eigenstate of $H_{\mathbf{k}'}$, where $\mathbf{k}' = \mathbf{k} + \epsilon\mathbf{a}$, and $\mathbf{a}$ is in the tangent space at $\mathbf{k}$, $\mathbf{a}\cdot\mathbf{k} = 0$, and ϵ is small. We thus want a partial isometry $U_{\mathbf{k}',\mathbf{k}}$ from the range of $P_{\mathbf{k}}$ to the range of $P_{\mathbf{k}'}$, $U_{\mathbf{k}',\mathbf{k}} P_{\mathbf{k}} = P_{\mathbf{k}}$, $P_{\mathbf{k}'} U_{\mathbf{k}',\mathbf{k}} = U_{\mathbf{k}',\mathbf{k}}$. When $\mathbf{k} = \mathbf{k}'$, this should simply be the identity on the range of $P_{\mathbf{k}}$, $U_{\mathbf{k},\mathbf{k}} = P_{\mathbf{k}}$. It is easy to check that if $\mathbf{k}' = \mathbf{k} + \epsilon\mathbf{a}$, $U_{\mathbf{k}',\mathbf{k}} = P_{\mathbf{k}'} P_{\mathbf{k}}$ has the desired properties to first order in ϵ, $U_{\mathbf{k}',\mathbf{k}}{}^\dagger U_{\mathbf{k}',\mathbf{k}} = P_{\mathbf{k}} + o(\epsilon^2)$. According to Simon,[2] and Wilczek and Zee,[3] this is the correct prescription for the Berry connection. To parallel transport along an arbitrary path γ on S^2, starting at $\mathbf{k}_i$ and ending at $\mathbf{k}_f$, choose $N+1$ points ($\mathbf{k}_i = \mathbf{k}_0, \mathbf{k}_1, \ldots, \mathbf{k}_n, \ldots, \mathbf{k}_{N-1}, \mathbf{k}_N = \mathbf{k}_f$) equally spaced along γ, in the end taking the limit $N \to \infty$. The partial isometry $U_\gamma = P_{\mathbf{k}_N} P_{\mathbf{k}_{N-1}} \cdots P_{\mathbf{k}_n} \cdots P_{\mathbf{k}_1} P_{\mathbf{k}_0}$ maps the range of $\mathbf{k}_i$ to the range of $\mathbf{k}_f$. This determines parallel transport along γ in the Berry connection.

Since the action of the rotation group SU(2) on S^2 is transitive, there exists for any two unit vectors $\mathbf{k}$, $\mathbf{k}'$ an $a \in \mathrm{SU}(2)$ such that $R(a)\colon \mathbf{k} \to \mathbf{k}'$. Let $G_{\mathbf{k}}$ denote the isotropy group of the vector $\mathbf{k}$, the set of $h \in \mathrm{SU}(2)$ satisfying $R(h)$: $\mathbf{k} \to \mathbf{k}$. The elment a above is not unique, since $h' a h \in \mathrm{SU}(2)$ also maps $\mathbf{k} \to \mathbf{k}'$ if $h \in G_{\mathbf{k}}$, $h' \in G_{\mathbf{k}'}$. Consider the path γ, and take again $N+1$ equally spaced along the path. We can choose $(a_1, a_2, \ldots, a_n, \ldots, a_n)$, where $a_n \in \mathrm{SU}(2)$, and $R(a_n)$: $\mathbf{k}_{n-1} \to \mathbf{k}_n$. We can choose these so all the a_n are close to the identity e of SU(2). It is convenient to define $h_n = a_n a_{n-1} \cdots a_2 a_1$, so that $R(h_n)\colon \mathbf{k}_0 \to \mathbf{k}_n$. Recall that U denotes the representation of SU(2) on the Hilbert space of states $\mathscr{H}$. Define $U_n = U(h_n)$, then $P_{\mathbf{k}_n} = U_n P_{\mathbf{k}_0} U_n{}^\dagger$.

It is straightforward to write a formula for U_γ,

$$U_\gamma = U_N P_{\mathbf{k}_0} U_N{}^\dagger U_{N-1} P_{\mathbf{k}_0} U_{N-1}{}^\dagger U_{N-2} \times P_{\mathbf{k}_0} \cdots P_{\mathbf{k}_0} U_2{}^\dagger U_1 P_{\mathbf{k}_0} U_1{}^\dagger P_{\mathbf{k}_0} . \tag{7.1}$$

Now

$$U_n{}^\dagger U_{n-1} = U(h_n{}^{-1}) U(h_{n-1}) = U(h_{n-1}{}^{-1} a_n{}^{-1} h_{n-1}) . \tag{7.2}$$

Thus $U_n{}^\dagger U_{n-1}$ is always close to the identity operator. We write U_γ as

$$U_\gamma = U_N A_N A_{N-1} \cdots A_n \cdots A_2 A_1 , \tag{7.3}$$

where $A_n = P_{\mathbf{k}_0} U_n{}^\dagger U_{n-1} P_{\mathbf{k}_0}$. Each A_n is a partial isometry from the range of $P_{\mathbf{k}_0}$ onto itself. We can write $U_n{}^\dagger U_{n-1} = \exp(i\mathbf{G}_n \cdot \mathbf{J})$ for some vector $\mathbf{G}_n$. Since $U_n{}^\dagger U_{n-1}$ is close to the identity operator, we can take $\mathbf{G}_n$ to have small norm, and approximate $U_n{}^\dagger U_{n-1} = \mathbf{1} + i\mathbf{G}_n \cdot \mathbf{J} + o(|\mathbf{G}_n|^2)$. Since we are interested in the limit $N \to \infty$, we need only consider the first-order term.

It is convenient to assume that $\mathbf{k}_i = \mathbf{k}_0 = \hat{\mathbf{z}}$. We wish to evaluate expressions of the form

$$A_n = P_{\hat{\mathbf{z}}} (1 + i\mathbf{G}_n \cdot \mathbf{J}) P_{\hat{\mathbf{z}}} . \tag{7.4}$$

By definition, we have

$$P_{\hat{\mathbf{z}}} = |j_1, m_1\rangle\langle j_1, m_1| + |j_2, m_2\rangle\langle j_2, m_2|, \tag{7.5}$$

$$\mathrm{P}_{\hat{\mathbf{z}}} J_z P_{\hat{\mathbf{z}}} = m_1 |j_1, m_1\rangle\langle j_1, m_1| + m_2 |j_2, m_2\rangle\langle j_2, m_2| , \tag{7.6}$$

$$P_{\hat{\mathbf{z}}} J_x P_{\hat{\mathbf{z}}} = |j_1, m_1\rangle\langle j_1, m_1| J_x |j_2, m_2\rangle\langle j_2, m_2| + |j_2, m_2\rangle\langle j_2, m_2| J_x |j_1, m_1\rangle\langle j_1, m_1| . \tag{7.7}$$

Recalll that $\langle j_1, m_1| J_x |j_2, m_2\rangle$ vanishes if $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$ are not in the same j multiplet. If $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$ are in the same j multiplet, the matrix element vanishes if $m_1 - m_2 \neq \pm 1$. The same statement holds for matrix elements of J_y. If these matrix elements vanish, the Berry connection splits, and is determined by the classification. We can now see that the only instance in which the Berry connection might not split is when $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$ are in the same j multiplet, and $m_1 - m_2 = \pm 1$. We shall from now on use the shorthand $j_1 = j_2$ to mean that the states are in the same j multiplet, and define $j = j_1 = j_2$. We need only consider the case $m_1 - m_2 = 1$, and define $m = \frac{1}{2}(m_1 + m_2)$. We then have $m_1 = m + \frac{1}{2}$, $m_2 = m - \frac{1}{2}$. We will use the shorthand $|m_+\rangle = |j, m + \frac{1}{2}\rangle$, $|m_-\rangle = |j, m - \frac{1}{2}\rangle$. In the following equations, all matrices act on the two-dimensional complex space spanned by $|m_+\rangle$ and $|m_-\rangle$:

$$P_{\hat{\mathbf{z}}} = |m_+\rangle\langle m_+| + |m_-\rangle\langle m_-| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{1} , \tag{7.8}$$

$$\begin{aligned} P_{\hat{\mathbf{z}}} J_z P_{\hat{\mathbf{z}}} &= m(|m_+\rangle\langle m_+| + |m_-\rangle\langle m_-|) \\ &\quad + \tfrac{1}{2}(|m_+\rangle\langle m_+| - |m_-\rangle\langle m_-|) \\ &= m\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = m\mathbf{1} + \frac{1}{2}\sigma_z , \end{aligned} \tag{7.9}$$

$$\begin{aligned} P_{\hat{\mathbf{z}}} J_x P_{\hat{\mathbf{z}}} &= \tfrac{1}{2}[(j + m + \tfrac{1}{2})(j - m + \tfrac{1}{2})]^{1/2} \\ &\quad \times (|m_+\rangle\langle m_-| + |m_-\rangle\langle m_+|) \\ &= \tfrac{1}{2}[(j + m + \tfrac{1}{2})(j - m + \tfrac{1}{2})]^{1/2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= [(j + m + \tfrac{1}{2})(j - m + \tfrac{1}{2})]^{1/2} \tfrac{1}{2}\sigma_x , \end{aligned} \tag{7.10}$$

$$P_{\hat{z}}\, J_y\, P_{\hat{z}} = \tfrac{1}{2}[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}$$
$$\times(-i|m_+\rangle\langle m_-| + i|m_-\rangle\langle m_+|)$$
$$= \tfrac{1}{2}[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$$
$$= [(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\,\tfrac{1}{2}\sigma_y\,. \quad (7.11)$$

We now study a class of paths for which we can compute parallel transport explicitly. These are the paths which are the orbit of a point on S^2 under rotation about an axis. Let **G** be a unit vector. Consider the rotation about the **G** axis by angle t. Let a denote the corresponding element of SU(2), $a = \exp(-it\,\mathbf{G}\cdot\boldsymbol{\sigma})$. For simplicity, we assume that the initial point of the path is $\mathbf{k}_i = \hat{\mathbf{z}}$. The final point of the path is given by $\mathbf{k}_f = R(a)\hat{\mathbf{z}}$. The action on vectors in the Hilbert space is by $\exp(-it\,\mathbf{G}\cdot\mathbf{J})$. We recall the rule (7.3) for parallel transport. Consider $t_N = t/N$. Parallel transport is given by the operator U_γ,

$$U_\gamma = \exp(-it\,\mathbf{G}\cdot\mathbf{J})T, \quad (7.12)$$

$$T = \lim_{N\to\infty} (A(t_N))^N. \quad (7.13)$$

We write A for A_n, since all are the same, and let t_N denote the dependence on N,

$$A(t_N) = P_{\hat{z}} + it_N P_{\hat{z}}(\mathbf{G}\cdot\mathbf{J})P_{\hat{z}} + o(t_N)^2. \quad (7.14)$$

We recall an elementary formula,

$$\lim_{n\to\infty} (1 + x/n)^n = \exp(x)\,. \quad (7.15)$$

In our case this gives

$$T = P_{\hat{z}} \exp(itP_{\hat{z}}(\mathbf{G}\cdot\mathbf{J})P_{\hat{z}})P_{\hat{z}}\,, \quad (7.16)$$

$$U_\gamma = \exp(-it\,\mathbf{G}\cdot\mathbf{J})P_{\hat{z}} \exp(itP_{\hat{z}}\mathbf{G}\cdot\mathbf{J}\,P_{\hat{z}})P_{\hat{z}}\,. \quad (7.17)$$

We are interested in the probability of a transition between eigenstates of $\mathbf{J}\cdot\mathbf{k}$. The projection $P_{\hat{z},+}$ projects onto eigenstates of $\mathbf{k}_i\cdot\mathbf{J}$ with eigenvalue $m+\frac{1}{2}$. The operator $\exp(-it\,\mathbf{G}\cdot\mathbf{J})P_{\hat{z},-}\exp(it\,\mathbf{G}\cdot\mathbf{J})$ projects onto eigenstates of $\mathbf{k}_f\cdot\mathbf{J}$ with eigenvalue $m-\frac{1}{2}$. We define

$$C' = \exp(-it\,\mathbf{G}\cdot\mathbf{J})P_{\hat{z},-} \exp(it\,\mathbf{G}\cdot\mathbf{J})\,U_\gamma P_{\hat{z},+}\,. \quad (7.18)$$

We consider $C'|m_+\rangle$. First note that $P_{\hat{z},+}|m_+\rangle = |m_+\rangle$, then U_γ gives the result of parallel transporting to $\mathbf{k}_f = R(a)\hat{\mathbf{z}}$, and the remaining part projects onto the eigenstate of $\mathbf{k}_f\cdot\mathbf{J}$ with eigenvalue $m-\frac{1}{2}$. The transition probability between eigenstates is thus given by $W = \langle m_+|C'^\dagger C'|m_+\rangle$. Equivalently, W is the operator norm of $C'^\dagger C'$. It is convenient to define

$$C = P_{\hat{z},-} \exp(itP_{\hat{z}}(\mathbf{G}\cdot\mathbf{J})P_{\hat{z}})P_{\hat{z},+}\,, \quad (7.19)$$

which has the property $C^\dagger C = C'^\dagger C'$. Note also $P_{\hat{z}}CP_{\hat{z}} = C$. This property lets us restrict attention to the two-dimensional range of $P_{\hat{z}}$, the subspace $P_{\hat{z}}\mathscr{H}\subset\mathscr{H}$. This has a basis $|m_+\rangle$, $|m_-\rangle$.

We represent operators as matrices in the obvious way,

$$P_{\hat{z}} = \mathbf{1},\quad P_{\hat{z},+} = \tfrac{1}{2}(\mathbf{1}+\sigma_z),\quad P_{\hat{z},-} = \tfrac{1}{2}(\mathbf{1}-\sigma_z)\,, \quad (7.20)$$

$$P_{\hat{z}}\, J_z P_{\hat{z}} = \tfrac{1}{2}(m\mathbf{1}+\sigma_z)\,, \quad (7.21)$$

$$P_{\hat{z}}\, J_x P_{\hat{z}} = \tfrac{1}{2}[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\sigma_x\,, \quad (7.22)$$

$$P_{\hat{z}}\, J_y P_{\hat{z}} = \tfrac{1}{2}[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\sigma_y\,. \quad (7.23)$$

We need to exponentiate these matrices. For any vector **B**, we have

$$\exp(i\mathbf{B}\cdot\boldsymbol{\sigma}) = \cos(|\mathbf{B}|)\mathbf{1} + \sin(|\mathbf{B}|)(i\mathbf{B}\cdot\boldsymbol{\sigma}/|\mathbf{B}|)\,. \quad (7.24)$$

Now $P_{\hat{z}}\mathbf{G}\cdot\mathbf{J}\,P_{\hat{z}} = B_0\mathbf{1} + \mathbf{B}\cdot\boldsymbol{\sigma}$, where

$$B_0 = \tfrac{1}{2}tmG_z\,, \quad (7.25)$$

$$B_x = \tfrac{1}{2}t\,[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}G_x\,, \quad (7.26)$$

$$B_y = \tfrac{1}{2}t\,[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}G_y\,, \quad (7.27)$$

$$B_z = \tfrac{1}{2}tG_z\,. \quad (7.28)$$

We find

$$C = \exp(iB_0)P_{\hat{z},-}(\cos(|\mathbf{B}|)\mathbf{1} + \sin(|\mathbf{B}|)(i\mathbf{B}\cdot\boldsymbol{\sigma}/|\mathbf{B}|))P_{\hat{z},+}\,. \quad (7.29)$$

Using the identities

$$P_{\hat{z},-}\sigma_z P_{\hat{z},+} = 0\,,\quad P_{\hat{z},-}\sigma_x P_{\hat{z},+} = \tfrac{1}{2}(\sigma_x - i\sigma_y)\,,$$
$$P_{\hat{z},-}\sigma_y P_{\hat{z},+} = (i/2)(\sigma_x - i\sigma_y)\,, \quad (7.30)$$

we find

$$C = \exp(iB_0)[\sin(|\mathbf{B}|)/|\mathbf{B}|](B_x\sigma_x - iB_y\sigma_y)\,. \quad (7.31)$$

It is now easy to compute the transition probability W. Recall W is equal to the operator norm of $C^\dagger C$,

$$W = [(B_x^2 + B_y^2)/|\mathbf{B}|^2]\sin^2(|\mathbf{B}|)\,. \quad (7.32)$$

We express **G** in spherical coordinates, $G_x = \sin(\theta)\cos(\phi)$, $G_y = \sin(\theta)\sin(\phi)$, $G_z = \cos(\theta)$. Then

$$B_x = \tfrac{1}{2}t\sin(\theta)\cos(\phi)[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\,, \quad (7.33)$$

$$B_y = \tfrac{1}{2}t\sin(\theta)\sin(\phi)[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\,, \quad (7.34)$$

$$B_z = \tfrac{1}{2}t\cos(\theta)\,, \quad (7.35)$$

$$|\mathbf{B}| = \tfrac{1}{2}t\,[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})\sin^2(\theta) + \cos^2(\theta)]^{1/2}\,, \quad (7.36)$$

$$\frac{B_x^2 + B_y^2}{|\mathbf{B}|^2} = \frac{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)}{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta) + \cos^2(\theta)}\,. \quad (7.37)$$

The transition probability W is given by

$$W = A\sin^2(\tfrac{1}{2}\Omega t) = \tfrac{1}{2}A\,(1-\cos(\Omega t))\,, \quad (7.38)$$

where A and Ω depend on θ,

$$A = \frac{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)}{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta) + \cos^2(\theta)}\,, \quad (7.39)$$

$$\Omega = [(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})\sin^2(\theta) + \cos^2(\theta)]^{1/2}\,. \quad (7.40)$$

We see that the transition probability is an oscillating function of t, which is the angle of rotation about the axis **G**. We check some basic properties, Ω should be real. Note that $-(j-\frac{1}{2})\leqslant m\leqslant(j-\frac{1}{2})$, so

$$\Omega^2 = ((j+\tfrac{1}{2})^2 - m^2 - 1)\sin^2(\theta) + 1 \geqslant 0\,. \quad (7.41)$$

Next, we see that as $\theta \to 0$, $A \to 0$, so $W \to 0$. W is invariant under $\theta \to (\pi - \theta)$, and also under $t \to -t$.

VIII. APPROXIMATE DEGENERACY

We now consider the case of approximate degeneracy. This is important in the analysis of experimental data, and is further discussed elsewhere.[4] Consider a nondegenerate eigenvector of the Hamiltonian $H_{\mathbf{k}}$. We denote the time variable by τ. The Hamiltonian is time dependent, $H(\tau)$ the time dependence is specified by the motion of unit vector $\mathbf{k}(\tau)$ on the parameter space S^2. We shall denote by $\mathbf{k}_i$ the initial point of the path, and by $\mathbf{k}_f$ the final point. We denote by $U(\tau)$ the unitary time evolution operator. This satisfies the Schrödinger equation

$$i\frac{dU}{d\tau} = HU. \tag{8.1}$$

Denote the time derivative of $\mathbf{k}$ by $\mathbf{k}'$. Here $\mathbf{k}'(\tau)$ is a vector tangent to S^2 at the point $\mathbf{k}(\tau)$. Let ω be the length of $\mathbf{k}'$, $\omega(\tau) = |\mathbf{k}'(\tau)|$.

We wish to apply the adiabatic theorem.[20] We first note that since all the j multiplets are mutually orthogonal, the adiabatic parameter change we are considering cannot cause mixing between different j multiplets. We need only consider one j multiplet at a time. We shall from now on assume that all states are in the same j multiplet. This reduces the problem to finite dimensions. The adiabatic theorem in this setting is quite simple. Suppose that $\mathbf{k}(0) = \mathbf{k}_i = \hat{\mathbf{z}}$. Consider a normalized eigenstate $|m\rangle$ of $H_{\hat{\mathbf{z}}}$, with $H_{\hat{\mathbf{z}}}|m\rangle = E_m|m\rangle$. Consider the quantity

$$\delta_m = \min_{-j<m'<j} |E_m - E_{m'}|, \tag{8.2}$$

where δ_m is the splitting between $|m\rangle$ and the state in the same j multiplet closest to it in energy. The adiabatic theorem in this case states that

$$\mathbf{k}_f\cdot\mathbf{J}\, U(\tau)|m\rangle = m\alpha U(\tau)|m\rangle, \tag{8.3}$$

where α is a complex number with unit modulus, if the following condition is satisfied:

$$\sup_{0<x<\tau} \omega(x) \ll \delta_m. \tag{8.4}$$

In fact, we can simplify even further. Let $\mathbf{k}'$ be close to $\mathbf{k}$. Then an eigenstate of $\mathbf{k}\cdot\mathbf{J}$ with eigenvalue m is orthogonal to all eigenstates of $\mathbf{k}'\cdot\mathbf{J}$ except, possibly, those with eigenvalues m, $m \pm 1$. We can then define

$$\delta_m = \min(E_m^-, E_m^+), \tag{8.5}$$

where $E_m^- = |E_m - E_{m-1}|$, $E_m^+ = |E_m - E_{m+1}|$, and the adiabatic theorem as stated in (8.3) still holds.

We are interested in approximate degeneracy between two states $|m_1\rangle$, $|m_2\rangle$ in the same j multiplet. By approximate degeneracy, we mean that the energy difference between the states is the same order of magnitude as ω. From the adiabatic theorem above, we see that no mixing between eigenstates of $\mathbf{k}\cdot\mathbf{J}$ is possible unless $m_1 - m_2 = \pm 1$. This was already apparent for the case of exact degeneracy from the classification of invariant connections. We define as before $m_+ = m_1 = m + \frac{1}{2}$, $m_- = m_2 = m - \frac{1}{2}$. Define δ as the smallest splitting between states, excluding the approximately degenerate states,

$$\delta = \min(\min_{k<m_-} E_k^-, \min_{n>m_+} E_n^+). \tag{8.6}$$

Let $\epsilon = |E_{m_+} - E_{m_-}|$. We suppose that $\omega \ll \delta$ for all τ. The adiabatic theorem then determines up to a phase the time evolution of all states except $|m_+\rangle$ and $|m_-\rangle$. All eigenstates of $\mathbf{k}_i\cdot\mathbf{J}$, except possibly $|m_+\rangle$ and $|m_-\rangle$, evolve into eigenstates of $\mathbf{k}_f\cdot\mathbf{J}$. We can then deduce from the unitarity of the time evolution that the states $|m_+\rangle$ and $|m_-\rangle$ can evolve only into linear combinations of states with $\mathbf{k}_f\cdot\mathbf{J}$ eigenvalues m_+ and m_-. If $\epsilon \ll \omega$, then the states are effectively exactly degenerate, and time evolution is governed by Berry's connection. If $\epsilon \gg \omega$, then the states are effectively nondegenerate, and the adiabatic theorem tells us that there is no mixing between states of different $\mathbf{J}\cdot\mathbf{k}$ eigenvalues. We are, however, interested in the intermediate case.

We shall again consider paths generated by rotation about a unit vector $\mathbf{G}$. Let τ be the time variable, and ω the rate of rotation. The meaning of ω is slightly different than in the previous section, but this should cause no difficulties. The angle of rotation is $t = \omega\tau$ as before. We can formally give an exact solution for the time evolution operator $U(\tau)$. The Schrödinger equation states

$$i\frac{dU}{d\tau} = HU. \tag{8.7}$$

When H is time independent, this has the familiar solution $U(\tau) = \exp(-i\tau H)$. When H is time dependent, the situation is more complicated. Suppose the time dependence of H is given by $H(\tau) = \exp(-i\tau A)H(0)\exp(i\tau A)$. Then it is easy to check that the following is the solution for the time evolution operator U:

$$U(\tau) = \exp(-i\tau A)\exp[-i\tau(H - A)). \tag{8.8}$$

We are specifically interested in the case $A = \omega\mathbf{G}\cdot\mathbf{J}$. Since we need only consider the finite-dimensional restriction to a given j multiplet, we can, in principle, find an exact expression for $U(\tau)$. In practice, this is a little difficult. The simplest interesting case is $j = 1$, where we can have a degenerate pair without the whole multiplet being degenerate. We must then exponentiate three-dimensional matrices, which is most easily done by first diagonalizing. The secular equation is cubic, and explicit formulas exist for the roots. This gets very messy very quickly. We can greatly simplify by using the adiabatic theorem.

We study the time evolution operator in the case of approximate degeneracy. The adiabatic theorem tells us that $U(\tau)$ maps a nondegenerate eigenstate of $\mathbf{k}_i\cdot\mathbf{J}$ with eigenvalue m' into an eigenstate of $\mathbf{k}_f\cdot\mathbf{J}$ with the same eigenvalue. Now the operator $\exp(-i\tau\omega\mathbf{G}\cdot\mathbf{J})$ also maps a nondegenerate eigenstate of $\mathbf{k}_i\cdot\mathbf{J}$ with eigenvalue m' into an eigenstate of $\mathbf{k}_f\cdot\mathbf{J}$ with the same eigenvalue. So we see that $\exp[-i\tau(H - \omega\mathbf{G}\cdot\mathbf{J}))$ must map a nondegenerate eigenstate of $\mathbf{k}_i\cdot\mathbf{J}$ into itself, up to a phase. This tells us that the operator $\exp[-i\tau(H - \omega\mathbf{G}\cdot\mathbf{J}))$ is diagonal in the basis of eigenstates of $\mathbf{k}_i\cdot\mathbf{J}$, except possibly on the pair of degenerate states. In other words, an eigenstate of $\mathbf{k}_i\cdot\mathbf{J}$ with eigenvalue not equal to m_+ or m_- is mapped onto itself times a phase.

Recall that $P_{\hat{z}}$ is the two-dimensional projection onto the approximately degenerate subspace, the subspace spanned by $|m_+\rangle$ and $|m_-\rangle$. It follows that

$$[P_{\hat{z}}, \exp(-i\tau(H-\omega\mathbf{G}\cdot\mathbf{J}))]=0\,. \tag{8.9}$$

Restricted to the degenerate subspace, the time evolution operator then takes the form

$$U(\tau)=\exp(-i\tau\omega\mathbf{G}\cdot\mathbf{J})P_{\hat{z}} \times \exp(-i\tau P_{\hat{z}}(H-\omega\mathbf{G}\cdot\mathbf{J})P_{\hat{z}})P_{\hat{z}}\,. \tag{8.10}$$

We now proceed exactly as in the previous section. It is convenient to use $t=\tau\omega$. We must now consider the operator

$$tP_{\hat{z}}(\mathbf{G}\cdot\mathbf{J}-H/\omega)P_{\hat{z}}=B_0\mathbf{1}+\mathbf{B}\cdot\boldsymbol{\sigma}\,. \tag{8.11}$$

Here we define $\Delta=(E_+-E_-)/\omega$, $\hat{E}=(E_++E_-)\omega$, then the coefficients are

$$B_0=\tfrac{1}{2}t(mG_z-\hat{E})\,, \tag{8.12}$$

$$B_x=\tfrac{1}{2}tG_x[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\,, \tag{8.13}$$

$$B_y=\tfrac{1}{2}tG_y[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})]^{1/2}\,, \tag{8.14}$$

$$B_z=\tfrac{1}{2}t(G_z-\Delta)\,. \tag{8.15}$$

We note that when $\Delta=0$, all the coefficients except B_0 reduce to the values in the previous section. The B_0 just gives the overall phase, which was neglected in the previous section. Alternately, we could redefine our energy scale so $\hat{E}=0$. The value of B_0 does not affect the transition probability W. Expressing $\mathbf{G}$ in polar coordinates as before, we find

$$|\mathbf{B}|=\tfrac{1}{2}t[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})\times\sin^2(\theta)+(\cos(\theta)-\Delta)^2]^{1/2}\,, \tag{8.16}$$

$$\frac{B_x^2+B_y^2}{|\mathbf{B}|^2}=\frac{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)}{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)+(\cos(\theta)-\Delta)^2}\,. \tag{8.17}$$

We then find a similar expression for the transition probability W,

$$W=A\sin^2(\tfrac{1}{2}\Omega t)=\tfrac{1}{2}A(1-\cos(\Omega t))\,, \tag{8.18}$$

where A and Ω depend now on Δ as well as θ,

$$A=\frac{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)}{(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin^2(\theta)+(\cos(\theta)-\Delta)^2}\,, \tag{8.19}$$

$$\Omega=[(j+m+\tfrac{1}{2})(j-m+\tfrac{1}{2})\sin^2(\theta)+(\cos(\theta)-\Delta)^2]^{1/2}. \tag{8.20}$$

The possible transition probability is maximum at $\Delta=\cos(\theta)$. The maximum has a value of unity, so a transition probability of unity is possible. Ω has its minimum at this value of Δ. Now as Δ moves away from this value, A decreases, and Ω increases. At $\Delta=0$, the case of exact degeneracy, we recover our previous result. As a function of Δ, A is Lorentzian centered around $\Delta=\cos(\theta)$, with width $(j+m+\frac{1}{2})(j-m+\frac{1}{2})\sin(\theta)$. Thus the effect is harder to observe at small θ, since Δ must be close to the maximum. We now see why at $\Delta=0$, A decreases as θ decreases. In the limit $\theta\to 0$, the above result can be obtained from time-dependent perturbation theory.[11]

IX. CONCLUSIONS

We have studied the family of Hamiltonians of the form $H_0+\mathbf{k}\cdot\mathbf{V}$. If a Hamiltonian in this family is nondegenerate, we have seen that the adiabatic behavior is determined uniquely by the rotational symmetry. Experiments on such systems can only yield limited information. If the Hamiltonian has a double degeneracy, or approximate degeneracy, between two states m_1 and m_2 in the same j multiplet, then the rotational symmetry suffices to determine the adiabatic behavior if and only if $m_1-m_2\neq\pm 1$. If $m_1-m_2=\pm 1$, we need to invoke the Schrödinger equation, in the guise of Simon's prescription for the Berry connection. Experiments on such systems can then yield more information. An example of such an experiment is discussed in the paper.[4]

APPENDIX: INVARIANT CONNECTIONS

We discuss the classification of rotationally invariant connections on the eigenstate bundles over S^2. The results are summarized in the main text. Here we present the technical details of the classification.

Consider a principal fiber bundle P, with structure group H and base space B. Recall that there is a free right action of H on P, and B is the quotient by this action. Suppose, furthermore, that there is a Lie group G of bundle automorphisms of P. This is a left action of G on P, which commutes with the action of the structure group. The G action on P thus induces a G action on B. We will be interested in classifying connections on P which are invariant under the action of G. We shall always assume that the induced G action on B is transitive. The following can be generalized to nontransitive G action on B, but the discussion becomes more cumbersome. We shall assume that P, G, H, and B are all compact.

We shall formulate a generalization of a theorem of Wang.[19] Wang's theorem gives a classification of invariant connections for a special case. Consider a point $y\in B$, and denote by G_y the isotropy group of y, G_y is the subgroup of G which leaves y fixed. Wang gives a classification in the case when G_y acts transitively on the fiber P_y over y. Recall that G is assumed to act transitively on the base B. We shall obtain a result for the case when G_y does not necessarily act transitively on the fiber over y.

Let $h\in H$, and denote by $R_h: P\to P$ the right action of the structure group. Note that $R_{h^{-1}}$ gives a left action of H on P. This action is not by bundle automorphisms, since it clearly does not commute with the action of the structure group, unless the structure group is Abelian. This left action of H does, however, commute with the left action of G on P. We thus have a left action of $G\times H$ on P. It is easy to check that if the G action on B is transitive, then the $G\times H$ action on P is also transitive. The action of $G\times H$ is not by bundle automorphisms, so we cannot use Wang's theorem.

We recall the construction of a connection on a principal fiber bundle.[17] Consider a point $y\in P$, and the tangent space $T_y=T_yP$. Consider the orbit of y under the right action of the structure group H. Recall that the vertical space $V_y \subset T_y$ is defined as the subspace tangent to the orbit of y under the action of the structure group. Here V_y has the same dimension as H. A connection ω is a one-form on P taking values in the vertical subspace of the tangent space, which is invariant under the structure group action. At a point $y \in P$, ω_y is a linear map $\omega_y: T_y \to T_y$, satisfying two conditions: (1) the range of ω_y is V_y, and (2) the restriction of ω_y to V_y is the identity map of V_y, $\omega_y|_V = \mathrm{Id}_V$. The structure group acts by $R_h: P \to P$, where $h \in H$. Suppose $R_h(y) = x$. Then by invariance under the action of the structure group, we mean $\omega_x R'_h = R'_h \omega_y$. Here R'_h denotes the derivative of the map R_h. One may construct a connection by choosing smoothly for each $y \in P$ an ω_y satisfying (1) and (2) above. This is *a priori* not invariant under H. One obtains an H invariant ω by averaging over the action of H, using the Haar measure on H.

If ω is further invariant under the G action on P, we say that ω is a G-invariant connection. G-invariant connections exist. We use a similar argument. Construct at each $y \in P$ an ω_y satisfying (1) and (2). We now average over the Haar measure on $G \times H$. Thus at least one G-invariant connection exists. There may be more, however. Consider a point $y \in P$, and denote by I_y the isotropy group for the left $G \times H$ action. Here I_y is the subgroup of $G \times H$ which leaves y fixed. If ω is a G-invariant connection, then clearly I_y must leave ω_y invariant. Recall that the $G \times H$ action is transitive. Suppose we have at one point $y \in P$ an ω_y, satisfying (1) and (2), which is invariant under I_y. Then ω is defined at all points of P by the action of $G \times H$. We then have the following proposition.

Proposition: Fix a point $y \in P$. The G-invariant connections on P are in one-to-one correspondence with linear maps $\omega_y: T_y \to V_y$ which satisfy (1) and (2), and which are invariant under the action of I_y.

By invariance under the I_y action we mean the following. Let $i \in I_y$, then i: $y \to y$, the derivative map i': $T_y \to T_y$. Recall that ω_y: $T_y \to T_y$. We say that ω_y is invariant under the action of I_y if $i'\omega_y = \omega_y i'$ for all $i \in I_y$. This proposition contains Wang's theorem as a special case.

We shall be interested in principal H-bundles constructed as follows. Let A be a compact Lie group. Consider a manifold S with a transitive left $G \times A \times H$ action. Let $s \in S$, and denote by $\bar{I}_s$ the isotropy group of s under the $G \times A \times H$ action. $\bar{I}_s$ is a subgroup of $G \times A \times H$. Let P be the quotient of S under the A action, $P = S/A$, and $y \in P$ the point corresponding to $s \in S$. Here P admits a left $G \times H$ action, and can thus be considered as a fiber bundle with structure group H, and a group G of bundle automorphism. The structure group acts on the right, so we must reverse the H action appropriately. A point $y \in P$ corresponds to an A orbit in S. Let $(a,g,h) \in \bar{I}_s \subset A \times G \times H$. Then $(g,h) \in I_y \subset G \times H$. Conversely, suppose $(g,h) \in I_y$. Then there exists an $a \in A$ such that $(a,g,h) \in \bar{I}_s$.

The tangent space T_y corresponds to the quotient T_s/N_s, where $N_s \subset T_s$ is tangent to the A orbit of s. Consider the action of the derivative map of $\bar{\imath}'$: $T_s \to T_s$, where $\bar{\imath} \in \bar{I}_s$. It is easy to see that the subspace N_s is preserved by this map, $\bar{\imath}$: $N_s \to N_s$. Thus $\bar{\imath}'$ passes to a map of the quotient space T_s/N_s, denoted by i': $T_y \to T_y$. This corresponds to the derivative of a map $i \in I_y$.

We now consider an example. Let $S = G \times H$. G acts on S by left translation on the G factor, g': $(g,h) \to (g'g,h)$. Here H acts on S by right translation by the inverse on the H factor, h': $(g,h) \to (g,hh'^{-1})$. Let F_{AG}: $A \to G$ and F_{AH}: $A \to H$ be homomorphisms. A acts on S by a: $(g,h) \to (g(F_{AG}a^{-1}),(F_{AH}a)h)$.

Consider the point $s = (e,e) \in G \times H = S$. We wish to find the isotropy group $\bar{I}_s$. An element of $G \times A \times H$ acts on s by

$$(g,a,h): (e,e) \mapsto (g(F_{AG}a^{-1}), (F_{AH}a)h^{-1}). \tag{A1}$$

This element $(g,a,h) \in \bar{I}_s$ if and only if $g = F_{AG}a$ and $h = F_{AH}a$. The isotropy group I_y of the corresponding point $y \in P$ is the set of (g,h) such that there exists an $a \in A$ satisfying the above two conditions.

Consider the tangent space $T_s = T_sS$. This is equal to the direct sum of the tangent spaces $T_eG \oplus T_eH$. We will from now on denote the tangent space at the identity of a Lie group G by $\mathscr{L}_G = T_eG$. Thus every $Z \in T_sS$ can be written as $Z = (X,Y)$, for some $X \in \mathscr{L}_G$ and $Y \in \mathscr{L}_H$. Consider $(g,a,h) \in \bar{I}_s$. We wish to find the derivative of the map at s. We consider $\exp(tZ) = (\exp(tX), \exp(tY)) \in G \times H$, and $(F_{AG}a,a,F_{AH}a) \in \bar{I}_s$. Recall that all elements of $\bar{I}_s$ can be written in this way. The element of $\bar{I}_s$ gives a map $G \times H \to G \times H$ as follows:

$$\begin{aligned}(F_{AG}a,a,F_{AH}a)&: (\exp(tX), \exp(tY))\\ &\mapsto ((F_{AG}a)\exp(tX)(F_{AG}a^{-1}), (F_{AH}a)\\ &\quad\times \exp(tY)(F_{AH}a^{-1})).\end{aligned} \tag{A2}$$

The derivative of this map s is clearly $(X,Y) \mapsto (\mathrm{Ad}(F_{AG}a)X, \mathrm{Ad}(F_{AH}a)Y)$. Here Ad denotes the adjoint representation of a Lie group on its Lie algebra.

The null space N_s is the tangent along the A orbit of $s \in S$. Let $W \in \mathscr{L}_A$. Then $\exp(tW) \in A$. Under the A action on $S = G \times H$, this gives

$$\exp(tA): (e,e) \mapsto (F_{AG}\exp(-tA), F_{AH}\exp(tA)). \tag{A3}$$

Now F_{AG}': $\mathscr{L}_A \to \mathscr{L}_G$, F_{AH}': $\mathscr{L}_A \to \mathscr{L}_H$. The null space N_s is spanned by vectors of the form $(-F_{AG}'W, F_{AH}'W)$, where $W \in \mathscr{L}_A$. In other words, N_s is the range of the map $(-F_{AG}' \oplus F_{AH}')$: $\mathscr{L}_A \to \mathscr{L}_G \oplus \mathscr{L}_H$.

We now consider the nondegenerate eigenstate bundle over S^2. This is a complex line bundle, so the structure group is U(1). The SU(2) action is transitive on the principal bundle P. In fact, it is easy to check that $P = \mathrm{SU}(2) \times_{\mathrm{U}(1)} \mathrm{U}(1)$. The action of U(1) on the U(1) factor depends on the m quantum number. We now elaborate. Let $G = \mathrm{SU}(2)$, $H = \mathrm{U}(1)$, $A = \mathrm{U}(1)$. A basis for $\mathscr{L}_G = \mathscr{L}_{\mathrm{SU}(2)}$ is given by L_x, L_y, L_z, where $L_x = (i/2)\sigma_x$, $L_y = (i/2)\sigma_y$, $L_z = (i/2)\sigma_z$. These are all anti-Hermitian, and have the commutation relations $[L_x, L_y] = L_z$. Let $R = -i$ be a basis for $\mathscr{L}_A = \mathscr{L}_{\mathrm{U}(1)}$, $\exp(\alpha R) = \exp(-i\alpha)$. Similarly, let Q be a basis for $\mathscr{L}_H = \mathscr{L}_{\mathrm{U}(1)}$. We define the homomorphisms F_{AG} and F_{AH} by their derivatives at the identity,

$$F_{AG}': \mathscr{L}_{\mathrm{U}(1)} \to \mathscr{L}_{\mathrm{SU}(2)}, \quad R \mapsto 2L_z, \tag{A4}$$

$$F_{AH}': \mathscr{L}_{\mathrm{U}(1)} \to \mathscr{L}_{\mathrm{U}(1)}, \quad R \mapsto 2mQ. \tag{A5}$$

Let $s\in S$ be the point $(e,e)\in \mathrm{SU}(2)\times \mathrm{U}(1)$. The tangent space $T_s = \mathscr{L}_{\mathrm{SU}(2)} \oplus \mathscr{L}_{\mathrm{U}(1)}$ is spanned by L_x, L_y, L_z, Q. The null subspace N_s is the range of

$$(-F_{AG}') \oplus F_{AH}' \colon \mathscr{L}_{\mathrm{U}(1)} \to T_s$$

$$R \mapsto -2L_z + 2mQ\,. \tag{A6}$$

Let $\bar{\imath}\in \bar{I}_s$. Then $i = (F_{AG}a,\ a,\ F_{AH}a)$. Suppose that $a = \exp(-\frac{1}{2}tR)$. Recall that the derivative of this map is

$$\bar{\imath}' \colon \mathscr{L}_{\mathrm{SU}(2)} \oplus \mathscr{L}_{\mathrm{U}(1)} \to \mathscr{L}_{\mathrm{SU}(2)} \oplus \mathscr{L}_{\mathrm{U}(1)}\,,$$

$$L_x \mapsto \mathrm{Ad}(\exp(tL_z))L_x\,,$$

$$L_y \mapsto \mathrm{Ad}(\exp(tL_z))L_y\,,$$

$$L_z \mapsto L_z\,, \quad Q \mapsto Q\,. \tag{A7}$$

We consider an infinitesimal action. Let $t = \epsilon \ll 1$. Denote by $\bar{\imath}_\epsilon$ the corresponding element of $\bar{I}_s$. The action of $\bar{\imath}'_\epsilon$ is to first order in ϵ

$$L_x \mapsto L_x + \epsilon L_y\,, \quad L_y \mapsto L_y - \epsilon L_x\,,$$
$$L_z \mapsto L_z\,, \quad Q \mapsto Q\,. \tag{A8}$$

Let L' be the subspace of T_s spanned by L_x and L_y. Let F be the subspace of T_s spanned by L_z and R. Here T_s decomposes as $T_s = L' \oplus F$. The null subspace is a subspace of F, $N_s \subset F$. The vertical subspace, tangent to the structure group action, is also a subspace of F, $V_s \subset F$. Also, $T_s/N_s = L' \oplus F/N_s$, and $V_s/N_s = F/N_s$. Note that an element of $\bar{I}_s$ acts trivially on F, and acts by rotation on L'. In particular, the action on V_s/N_s is trivial.

We need to consider maps $\omega_s \colon T_s/N_s \to V_s/N_s$. Here ω must act as the identity on V_s/N_s. We need maps of $L' \to F/N_s$ which commute with the map $\bar{\imath}'_\epsilon$. First note that the composition $\omega \circ \bar{\imath}'_\epsilon$ maps

$$L_x \mapsto \omega(L_x) + \epsilon\omega(L_y)\,, \quad L_y \mapsto \omega(L_y) - \epsilon\omega(L_x)\,. \tag{A9}$$

The composition $\bar{\imath}'_\epsilon \circ \omega$ maps

$$L_x \mapsto \omega(L_x)\,, \quad L_y \mapsto \omega(L_y)\,. \tag{A10}$$

For these two maps to be equal, we must have

$$\omega(L_x) = 0 = \omega(L_y)\,. \tag{A11}$$

There is then a unique map ω satisfying the necessary conditions. We conclude that there is a unique SU(2) invariant connection for any value of m. This result also follows from an easy application of Wang's theorem.

We now construct a more complicated example. Let $S = G \times B \times H$. The action of $G \times A \times H$ is as follows. G acts on S by left translation on the G factor, g': $(g,b,h) \to (g'g,b,h)$. Here H acts on S by right translation by the inverse on the H factor, h': $(g,b,h) \to (g,b,hh'^{-1})$. Consider the group homomorphisms

$$F_{AG} \colon A \to G\,, \quad F^L_{AB} \colon A \to B\,,$$
$$F^R_{AB} \colon A \to B\,, \quad F_{AH} \colon A \to H\,, \tag{A12}$$

where A acts on S by

$$a \colon (g,b,h) \mapsto (g(F_{AG}a^{-1}),\ (F^L_{AB}a)b(F^R_{AB}a^{-1}),\ (F_{AH}a)h)\,. \tag{A13}$$

Note that the actions of A, G, and H mutually commute, so we indeed have an action of $A \times G \times H$ on S.

Consider the points $s = (e,e,e) \in G \times B \times H = S$. We wish to find the isotropy group $\bar{I}_s$. Let $(a,g,h) \in A \times G \times H$,

$$(a,g,h) \colon (e,e,e) \mapsto (g(F_{AG}a^{-1})\,,$$
$$(F^L_{AB}a)(F^R_{AB}a^{-1})\,,\ (F_{AH}a)h^{-1})\,. \tag{A14}$$

The isotropy group $\bar{I}_s$ is specified by the following conditions: $F^L_{AB}a = F^R_{B}a$, $g = F_{AG}a$, $h = F_{AH}a$. Thus $(g,h) \in I_y$ if and only if there exists an $a \in A$ satisfying the three conditions above.

An element $\bar{\imath} = (a,g,h) \in \bar{I}_s$ maps s to itself. The derivative thus gives a linear map $T_s \to T_s$. The derivative $\bar{\imath}'$ is given by the action of $(a,g,h) \in \bar{I}_s$ on $(\exp(tX)$, $\exp(tY), \exp(tZ)) \in G \times B \times H$, in the limit $t \to 0$. Since $(a,g,h) \in \bar{I}_s$, $F^L_{AB}a = F^R_{AB}a$. We denote the common value as $F_{AB}a$,

$$\bar{\imath} \colon (\exp(tX),\ \exp(tY),\ \exp(tZ)) \mapsto (\bar{g},\bar{b},\bar{h})\,, \tag{A15}$$

where

$$\bar{g} = (F_{AG}a)\exp(tX)(F_{AG}a^{-1})\,,$$
$$\bar{b} = (F_{AB}a)\exp(tY)(F_{AB}a^{-1})\,, \tag{A16}$$
$$\bar{h} = (F_{AH}a)\exp(tZ)(F_{AH}a^{-1})\,.$$

In the limit $t \to 0$, we obtain the derivative map

$$\bar{\imath}' \colon (X,Y,Z) \mapsto (\mathrm{Ad}(F_{AG}a)X,\ \mathrm{Ad}(F_{AB}a)Y,\ \mathrm{Ad}(F_{AH}a)Z)\,. \tag{A17}$$

The null space N is tangent to the orbit of s under A,

$$a \colon (e,e,e) \mapsto ((F_{AG}a^{-1})\,,\ (F^L_{AB}a)(F^R_{AB}a^{-1})\,,\ (F_{AH}a))\,. \tag{A18}$$

Consider $W \in \mathscr{L}_A$, $a = \exp(tW) \in A$. In the limit $t \to 0$, we obtain the derivative of the above map

$$(-F_{AG}') \oplus (F^{L\prime}_{AB} - F^{R\prime}_{AB}) \oplus (F_{AH}') \colon \mathscr{L}_A \to T_s\,. \tag{A19}$$

The range of this map is the null subspace $N_s \subset T_s$. In other words, N_s is generated by vectors of the form

$$(-F_{AG}'W,\ F^{L\prime}_{AB}W - F^{R\prime}_{AB}W,\ F_{AH}'W)\,, \tag{A20}$$

where $W \in \mathscr{L}_A$.

We now wish to consider the bundle corresponding to two degenerate eigenstates. Suppose we have two states with quantum numbers m_1 and m_2. We have seen above that the principal U(1) bundle associated to a line bundle over S^2 is $P = \mathrm{SU}(2) \times_{\mathrm{U}(1)} \mathrm{U}(1)$. Here the U(1) action depends on m. The $\mathrm{U}(1)\times\mathrm{U}(1)$ principal bundle corresponding to the direct sum of the two line bundles can be written as $P = \mathrm{SU}(2) \times_{\mathrm{U}(1)} \mathrm{U}(1)\times\mathrm{U}(1)$, where U(1) acts on $\mathrm{U}(1)\times\mathrm{U}(1)$ in the obvious way. The structure group is $\mathrm{U}(1)\times\mathrm{U}(1)$. We wish to imbed our $\mathrm{U}(1)\times\mathrm{U}(1)$ principal bundle into a U(2) principal bundle. This amounts to constructing the principal bundle

$$P = (\mathrm{SU}(2) \times_{\mathrm{U}(1)} \mathrm{U}(1)\times\mathrm{U}(1)) \times_{\mathrm{U}(1)\times\mathrm{U}(1)} \mathrm{U}(2)\,. \tag{A21}$$

We now elaborate. We shall write P as the quotient of a space S by the action of a group A. We take $G = \mathrm{SU}(2)$, $B = \mathrm{U}(1)\times\mathrm{U}(1)$, $H = \mathrm{U}(2)$. Let $A = (\mathrm{U}(1)\times\mathrm{U}(1)\times\mathrm{U}(1))$. As before, we take L_x, L_y, L_z as a basis for $\mathscr{L}_{SU(2)}$. We denote by $K_0 = (i/2)1$, $K_x = (i/2)\sigma_x$, $K_y = (i/2)\sigma_y$, $K_z = (i/2)\sigma_z$ a basis for $\mathscr{L}_{U(2)}$. Let R_1, R_2, R_3 be the basis for $\mathscr{L}_A = \mathscr{L}_{U(1)\times U(1)\times U(1)}$. Let Q_1, Q_2 be the basis for $\mathscr{L}_B = \mathscr{L}_{U(1)\times U(1)}$. We specify the homomorphisms by giving their derivatives at the identity

$$F_{AG}': \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{SU(2)}, \quad R_1 \mapsto 2L_z, \quad R_2 \mapsto 0, \quad R_3 \mapsto 0, \tag{A22}$$

$$F_{AH}': \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{U(2)}, \quad R_1 \mapsto 0, \quad R_2 \mapsto K_0 + K_z, \quad R_3 \mapsto K_0 - K_z, \tag{A23}$$

$$F_{AB}^{L}{}': \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{U(1)\times U(1)}, \quad R_1 \mapsto 2m_1Q_1 + 2m_2Q_2, \quad R_2 \mapsto 0, \quad R_3 \mapsto 0, \tag{A24}$$

$$F_{AB}^{R}{}': \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{U(1)\times U(1)}, \quad R_1 \mapsto 0, \quad R_2 \mapsto Q_1, \quad R_3 \mapsto Q_2. \tag{A25}$$

We now define a decomposition of the tangent space $T = T_s$. Let $X_1, X_2, ..., X_n \in T$ be a set of vectors. Denote by $\mathscr{C}(X_1, X_2, ..., X_n)$, the linear span of the vectors. This is a subspace of T. Define the following subspaces of T:

$$E = \mathscr{C}(L_z, Q_1, Q_2), \quad L' = \mathscr{C}(L_x, L_y), \quad K' = \mathscr{C}(K_x, K_y), \quad K^c = \mathscr{C}(K_0, K_z), \quad F = E \oplus K^c. \tag{A26}$$

Note that $T = L' \oplus F \oplus K'$. Let $s = (e,e,e)$. We need to determine $\bar{I}_s$. We must first find the subgroup of $A' \subset A$ satisfying $F_{AB}^{L} a = F_{AB}^{R} a$ for all $a \in A'$. The vector space $\mathscr{L}_{A'} \subset \mathscr{L}_A$ satisfies $F_{AB}^{L}{}' W = F_{AB}^{R}{}' W$, for all $W \in \mathscr{L}_{A'}$. Here $\mathscr{L}_{A'}$ is the kernel of the map

$$(F_{AB}^{L}{}' - F_{AB}^{R}{}'): \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{U(1)\times U(1)}, \quad R_1 \mapsto 2m_1Q_1 + 2m_2Q_2, \quad R_2 \mapsto -Q_1, \quad R_3 \mapsto -Q_2. \tag{A27}$$

Consequently $\mathscr{L}_{A'}$ is one dimensional, spanned by the vector $R' = \frac{1}{2}R_1 + m_1R_2 + m_2R_3$. The isotropy group is generated by $\mathscr{L}_{\bar{I}_s}$ which is the image of $\mathscr{L}_{A'}$ under the map

$$\begin{aligned}(-F_{AG}') \oplus (F_{AB}^{L}{}' - F_{AB}^{R}{}') \oplus (F_{AH}'): &\ \mathscr{L}_{A'} \to \mathscr{L}_G \oplus \mathscr{L}_A \oplus \mathscr{L}_H, \\ R' &\mapsto L_z + m_1(K_0 + K_z) + m_2(K_0 - K_z), \\ &\mapsto L_z + (m_1 + m_2)K_0 + (m_1 - m_2)K_z, \\ &\mapsto U. \end{aligned} \tag{A28}$$

Recall $T = \mathscr{L}_G \oplus \mathscr{L}_B \oplus \mathscr{L}_H$. The isotropy group $\bar{I}_s$ acts on T by $\mathrm{Ad}(tU)$. Suppose $t = \epsilon \ll 1$. Denote the corresponding map $\bar{\imath}_\epsilon$: $T \to T$. To first order,

$$\bar{\imath}_\epsilon': T \to T, \quad X \mapsto X + \epsilon[U,X]. \tag{A29}$$

Then $\bar{\imath}_\epsilon'$ acts trivially on E and K^c. The action on L' and K' is given by

$$\begin{aligned} &L_x \mapsto L_x + \epsilon L_y, \quad L_y \mapsto L_y - \epsilon L_x, \\ &K_x \mapsto K_x + \epsilon(m_1 - m_2)K_y, \quad K_y \mapsto K_y - \epsilon(m_1 - m_2)K_x. \end{aligned} \tag{A30}$$

The null space $N = N_s$ is tangent to the orbit of A. Here N is generated by the range of the map

$$\begin{aligned}(-F_{AG}^{L}{}') \oplus (F_{AB}^{L}{}' - F_{AB}^{R}{}') \oplus (F_{AH}'): &\ \mathscr{L}_{U(1)\times U(1)\times U(1)} \to \mathscr{L}_{SU(2)} \oplus \mathscr{L}_{U(1)\times U(1)} \oplus \mathscr{L}_{U(2)}, \\ R_1 &\mapsto -2L_z + 2m_1Q_1 + 2m_2Q_2, \\ R_2 &\mapsto -Q_1 + K_0 + K_z, \\ R_3 &\mapsto -Q_2 + K_0 - K_z. \end{aligned} \tag{A31}$$

Note that $U \in N$. Also $N \subset F$, so N is invariant under $\bar{I}_s$.

We now determine the relation between T/N and V/N. Recall $N \subset E$, consider $O = E/N$. Recall E is five dimensional, N is three dimensional, so the dimension of O is two and

$$T/N = L' \oplus O \oplus K'. \tag{A32}$$

Now consider $V = K^c \oplus K' \subset T$. We need to find $V/N \subset T/N$. One checks that $K^c \oplus N = F$. We thus find

$$V/N = O \oplus K', \quad T/N = L' \oplus V/N. \tag{A33}$$

Recall that the action of $\bar{I}_s$ is trivial on F, and since $O \subset F$, the action is trivial on O. We need to classify maps ω: $T/N \to V/N$, such that ω restricted to $V/N \subset T/N$ is the identity. ω must commute with the action of $\bar{I}_s$. This amounts to classifying maps ω: $L' \to O \oplus K'$ which commute with the action of $\bar{I}_s$,

$$\omega: L' \to O \oplus K', \quad L_x \mapsto O_x + J_x, \quad L_y \mapsto O_y + J_y, \tag{A34}$$

where $O_x, O_y \in O$, and $J_x, J_y \in K'$. Suppose $J_x = \alpha K_x + \beta K_y$, $J_y = \gamma K_x + \delta K_y$.

We now find the action of $\bar{\imath}_\epsilon' \circ \omega$ and $\omega \circ \bar{\imath}_\epsilon'$:

$$\begin{aligned}\bar{\imath}_\epsilon' \circ \omega: &\ L' \to O \oplus K', \\ L_x &\mapsto O_x + \alpha(K_x + \epsilon(m_1 - m_2)K_y) + \beta(K_y - \epsilon(m_1 - m_2))K_x, \\ &\mapsto O_x + (\alpha - \beta\epsilon(m_1 - m_2))K_x + (\beta + \alpha\epsilon(m_1 - m_2))K_y, \end{aligned}$$

$$L_y \mapsto O_y + \gamma(K_x + \epsilon(m_1 - m_2)K_y) + \delta(K_y - \epsilon(m_1 - m_2)K_x),$$
$$\mapsto O_y + (\gamma - \delta\epsilon(m_1 - m_2))K_x + (\delta + \gamma\epsilon(m_1 - m_2))K_y, \tag{A35}$$

$$\omega \circ \bar{\iota}'_\epsilon : L' \to O \oplus K',$$
$$L_x \mapsto O_x + \epsilon O_y + \alpha K_x + \beta K_y + \epsilon\gamma K_x + \epsilon\delta K_y,$$
$$\mapsto O_x + \epsilon O_y + (\alpha + \epsilon\gamma)K_x + (\beta + \epsilon\delta)K_y,$$
$$L_y \mapsto O_y - \epsilon O_x + \gamma K_x + \delta K_y - \epsilon\alpha K_x - \epsilon\beta K_y,$$
$$\mapsto O_y - \epsilon O_x + (\gamma - \epsilon\alpha)K_x + (\delta - \epsilon\beta)K_y. \tag{A36}$$

We demand $\omega \circ \bar{\iota}'_\epsilon = \bar{\iota}'_\epsilon \circ \omega$. This requires

$$O_x = O_y = 0, \quad -\beta(m_1 - m_2) = \gamma, \quad \alpha(m_1 - m_2) = \delta,$$
$$-\delta(m_1 - m_2) = -\alpha, \quad \gamma(m_1 - m_2) = -\beta. \tag{A37}$$

We first notice that $\omega: L' \to K'$, since O_x and O_y vanish. We see that δ and γ are determined by α and β. Combining equations, we find

$$\beta(m_1 - m_2)^2 = \beta, \quad \alpha(m_1 - m_2)^2 = \alpha. \tag{A38}$$

We now see that if $(m_1 - m_2) \neq \pm 1$, then all the parameters are zero, and the only solution is $\omega: L' \to O \oplus K'$ by the zero map. If $(m_1 - m_2) = \pm 1$, then α and β are arbitrary real numbers. We then get solutions

$$\omega: L' \to K',$$
$$L_x \mapsto \alpha K_x + \beta K_y,$$
$$L_y \mapsto \pm(-\beta K_x + \alpha K_y). \tag{A39}$$

The space of invariant connections is parametrized by $(\alpha,\beta) \in \mathbb{R}^2$. The Berry connection for (j,m) corresponds to $\alpha = \frac{1}{2}[(j + m + \frac{1}{2})(j - m + \frac{1}{2})]^{1/2}$, $\beta = 0$. The connection corresponding to $\alpha = \beta = 0$ splits, and no other invariant connection splits.

We should consider two connections as equivalent if they can be brought into one another by gauge transformation. A gauge transformation is a principal bundle automorphism which induces the identity map on the base space. A gauge transformation maps a G-invariant connnection into another G-invariant connection if and only if the gauge transformation commutes with the G action. Since the G action on the base space is transitive, it suffices to determine the action of the gauge transformation on the fiber over one point. Take the point $\hat{\mathbf{z}} \in S^2$. The gauge transformation at $\hat{\mathbf{z}}$ must commute with the action of the isotropy subgroup, the subgroup of G which leaves $\hat{\mathbf{z}}$ fixed. The only such automorphisms are generated by

$$\mathrm{Ad}(\exp(tK_z)): \mathscr{L}_{\mathrm{U}(2)} \to \mathscr{L}_{\mathrm{U}(2)},$$
$$K_x \mapsto \cos(t)K_x + \sin(t)K_y,$$
$$K_y \mapsto -\sin(t)K_x + \cos(t)K_y,$$
$$K_z \mapsto K_z,$$
$$K_0 \to K_0, \tag{A40}$$

where t is a real number. So we see that two invariant connections (α,β) and (α',β') are gauge equivalent if and only if $\alpha^2 + \beta^2 = \alpha'^2 + \beta'^2$. An invariant connection (α,β) is then gauge equivalent to the connection $((\alpha^2 + \beta^2)^{1/2}, 0)$. We see that the gauge equivalence classes of invariant connections are parametrized by $\mathbb{R}^+$, the non-negative real numbers.

[1] M. V. Berry, Proc. R. Soc. London Ser. A **392**, 45 (1984).
[2] B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).
[3] F. Wilczek and A. Zee, Phys. Rev. Lett. **52**, 2111 (1984).
[4] J. Segert, "Non-Abelian Berry's phase effects and optical pumping of atoms," submitted to Ann. Phys.
[5] M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
[6] J. Segert, Phys. Rev. A **36**, 10 (1987).
[7] R. Y. Chiao and Y. S. Wu, Phys. Rev. Lett. **57**, 933 (1986).
[8] F. D. M. Haldane, Opt. Lett. **11**, 730 (1986).
[9] A. Tomita and R. Y. Chiao, Phys. Rev. Lett. **57**, 937 (1986).
[10] J. Moody, A. Shapere, and F. Wilczek, Phys. Rev. Lett. **56**, 893 (1986).
[11] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1977).
[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1980), Vol. I.
[13] M. Karoubi, *K-Theory, An Introduction* (Springer, Berlin, 1974).
[14] J. W. Milnor and J. D. Stasheff, *Characteristic Classes* (Princeton U.P., Princeton, 1974).
[15] M. S. Narasimhan and S. Ramanan, Am. J. Math. **83**, 563 (1961).
[16] R. Bott and L. W. Tu, *Differential Forms in Algebraic Topology* (Springer, Berlin, 1982).
[17] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Wiley, New York, 1969), Vols. I and II.
[18] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds, and Physics, revised edition* (Elsevier, Amsterdam, 1982).
[19] H. C. Wang, Nagoya Math. J. **13**, 1 (1958).
[20] T. Kato, Phys. Soc. Jpn. **5**, 435 (1950).

Physical relations and the Weyl group

Grzegorz Cieciura
Department of Mathematical Methods in Physics, University of Warsaw, Hoza 74, 00-682 Warsaw, Poland
Igor Szczyrba
Department of Mathematics and Applied Statistics, University of Northern Colorado, Greeley, Colorado 80639

(Received 11 November 1985; accepted for publication 6 May 1987)

Let a given particle symmetry be described by a reductive Lie group G. It is proved that the corresponding Weyl group $W(R)$ acts canonically in all zero-weight spaces of G and hence, in particular, on observables. Moreover, it is shown how this $W(R)$ action provides many physical relations, including those believed to be implied by G-transformation properties of observables. The results simplify a testing of symmetries based on various Lie groups (algebras). Their use extends beyond particle physics, e.g., to nuclear physics.

I. INTRODUCTION

Let G be a reductive Lie group describing a symmetry in particle physics, e.g., a grand-unification or n-flavor symmetry. It is usually assumed that physical states correspond to some vectors of complex G representations and observables to some operators acting in these representations' spaces. Considering observables of particular G-transformation properties, i.e., tensor operators of a particular type,[1] one obtains certain relations for expectation values. These relations can then be compared with experimental data.

Our aim is the following. Let $W(R)$ be the Weyl group corresponding to G (R denotes the root system of G, see, e.g., Ref. 2). We prove that in the case of finite-dimensional representations there exists a canonical action of $W(R)$ on observables and that many physical relations are a consequence of this $W(R)$ action rather then of the natural action of G on observables.

All our constructions depend only on the Lie algebra $L(G)$ of G. Therefore we formulate the results in terms of Lie algebras. Let $\mathfrak{G}$ be a reductive (real or complex) Lie algebra, let $\mathfrak{H}$ be its Cartan subalgebra, and let $\underline{\Lambda}$ denote the set of weights of a complex finite-dimensional representation ρ of $\mathfrak{G}$, acting in the vector space V. The analysis of physical and mathematical structures used in descriptions of particle symmetries leads us to a conclusion that the observables acting in V satisfy

$$\rho(\mathfrak{H}) \subset \mathrm{Ob}(V) \subset \rho(\mathfrak{H})',$$

where $\rho(\mathfrak{H})'$ is the commutant of $\rho(\mathfrak{H})$ in End V. It is clear that $\rho(\mathfrak{H})'$ coincides with the zero-weight space of the representation ad ρ of $\mathfrak{G}$, acting in the space End V.

Our main idea consists in proving that in any zero-weight space of the pair $(\mathfrak{G},\mathfrak{H})$ there exists the canonical representation Π of the Weyl group $W(R)$. The representation Π generalizes the natural action of $W(R)$ in the Cartan algebra $\mathfrak{H}$ since $\mathfrak{H}$ is the zero-weight space of the adjoint representation of $\mathfrak{G}$. Next, we show that any $W(R)$-symmetry property of an observable $A \in \rho(\mathfrak{H})'$, e.g., an equation $\Sigma_{w\in W(R)} c_w \cdot \Pi_w(A) = 0$, provides similar equations for the diagonal expectation values of A in any basis consisting of weight vectors, i.e., it provides some physical relations.

For example, if A does not contain the trivial $W(R)$ component then $\Sigma_{w\in W(R)} \Pi_w(A) = 0$ and, hence, we get that the sum of the diagonal expectation values of A, corresponding to any $W(R)$ orbit in the set of weights $\underline{\Lambda}$, is equal to zero. Obviously, the presence or absence of other irreducible $W(R)$ components in A provides more complicated physical relations. This procedure generalizes the results obtained by one of the authors in Ref. 3.

Our results not only facilitate the obtaining of some equations connected with a given observable A, but also enable us to predict many physical relations that can appear in the framework of a considered symmetry. Namely, let X be a $W(R)$ orbit in $\underline{\Lambda}$. We show that decomposing the function space $\mathbb{C}^X := \{f: X \to \mathbb{C}\}$ into the irreducible $W(R)$ components one gets possible equations for diagonal expectation values of the observables. Thus, we reduce the deduction of many physical relations to the analysis of actions of finite groups.

We illustrate both these aspects of our approach by different examples from theories based on the groups SU(n), U(n), or GL(n,C). Let us remark that a given $W(R)$ representation can be, in general, realized in the observables of different $\mathfrak{G}$-transformation properties. Thereby, a given physical relation can be, in general, obtained by means of tensor operators of different types. To make it clear, let us consider the simplest example of mass formulas.

We shall prove that the Coleman–Glashow mass formula[4] appears in SU(3) theory iff the mass operator does not contain the component transformed by the one-dimensional (alternative) representation of $W(R) \simeq S_3$ of signature $(1^3) = (1,1,1)$. But the results of our subsequent paper (Ref. 5) show that the representation (1^3) is contained only in tensor operators of the decimet type. So, this mass formula appears also if the mass operator contains certain terms of the 27-plet type instead of (or in addition to) the adjoint type Gell-Mann mapping D. These two possibilities exist independently of the number of flavors n, e.g., in the meson case with $n \geqslant 4$, the same formulas can be obtained either by means of the singlet type tensor operators and the mapping D or by the u(n) representation of signature (4,2,...,2,0) itself. It concerns, in particular, the mass formula obtained in Ref. 6 for charmed mesons. Similar possibilities exist in the case of the Gell-Mann–Okubo mass formula[4] or its gen-

 0022-2488/87/092115-13$02.50

eralizations.[3] For details see Sec. IV E of this paper and Ref. 5, where the relationship between $\mathfrak{G}$ and $W(R)$ actions is analyzed for three-quark and quark–antiquark $\mathrm{u}(n)$ representations.

Let us also notice that in some experiments we observe first a finite symmetry, and next we extend it to a continuous one. For example, proton–neutron symmetry was replaced by the SU(2)-symmetry group. Our results show that one of the natural ways to extend a finite symmetry group W to a continuous one is the following. We take a semisimple (or reductive) Lie group G such that its Weyl group $W(R)$ is equal to W or at least $W(R)$ contains W.

II. MATHEMATICAL FRAMEWORK OF PARTICLE SYMMETRIES

Let us analyze first some mathematical and physical structures used in descriptions of particle symmetries. We shall deal only with symmetries corresponding to a reductive (real or complex) Lie algebra $\mathfrak{G} = \mathfrak{g} + \mathfrak{c}$, where $\mathfrak{g}$ is semisimple and $\mathfrak{c}$ is Abelian.

A. Cartan algebras

Some elements of $\mathfrak{G}$ are usually identified with physical quantities. Among these quantities a subset of simultaneously measurable ones is chosen. This subset is supposed to reflect the basic conservation laws of the considered theory. For example, in su(3)-flavor (resp. color) theory for basic observables can be taken: electric (resp. color) charge Q and (resp. color) hypercharge Y; in n-flavor theory with $n>3$ the algebra $\mathfrak{G}$ coincides with $\mathrm{u}(n) = \mathrm{su}(n) + \mathrm{u}(1)$ and for basic observables can be taken: electric charge Q, strangeness S, charm C, etc. If the basic observables are interpreted as linear combinations of number operators of particles (e.g., quarks), then *the basic conservation laws are equivalent to the conservation of corresponding combinations of particles*.

In what follows, by a $\mathfrak{G}$-symmetry theory we mean a theory based on the Lie algebra $\mathfrak{G}$, independently of a particle interpretation, e.g., su(3)-flavor theory denotes the classical unitary symmetry theory as well as three-quark flavor theory.

In most theories, one obtains an agreement with experimental data assuming that the basic observables span a Cartan algebra $\mathfrak{H}$ in $\mathfrak{G}$. However, it is sometimes necessary to assume that they span only a Cartan algebra $\widetilde{\mathfrak{H}}$ in a subalgebra $\widetilde{\mathfrak{G}}$ of $\mathfrak{G}$ (the case of a broken symmetry). For example, such a situation arises in the su(6)-spin-flavor theory, where $\widetilde{\mathfrak{G}} = \mathrm{su}(2) + \mathrm{su}(3)$. To simplify the notation, we shall omit the second possibility in general considerations. The case of a broken symmetry and the spin-flavor theory (with arbitrary number of flavors) shall be discussed in Sec. IV C.

A Cartan algebra $\mathfrak{H}$ is, by definition, an Abelian subalgebra in $\mathfrak{G}$ that coincides with its normalizer $N_{\mathfrak{G}}(\mathfrak{H})$, i.e., $\mathrm{ad}\, X(\mathfrak{H}) \subset \mathfrak{H}$, $X \in \mathfrak{G}$, implies that $X \in \mathfrak{H}$. In particular, $\mathfrak{H}$ is a maximal Abelian subalgebra of $\mathfrak{G}$ and, moreover, $\mathfrak{H} = \mathfrak{h} + \mathfrak{c}$, where $\mathfrak{h}$ is a Cartan algebra in $\mathfrak{g}$.

The condition $N_{\mathfrak{G}}(\mathfrak{H}) = \mathfrak{H}$ does not follow if one only assumes that $\mathfrak{H}$ is a maximal Abelian subalgebra and, moreover, it provides important additional consequences. Namely, for an arbitrary finite-dimensional representation (ρ, V) of $\mathfrak{G}$, all operators $\rho(H) \in \mathrm{End}\, V$, $H \in \mathfrak{h}$, are semisimple, i.e., they are diagonalizable or at least their complexifications are diagonalizable. In particular, all operators ad H, $H \in \mathfrak{H}$, are semisimple. The last condition can, in fact, replace the equality $N_{\mathfrak{G}}(\mathfrak{H}) = \mathfrak{H}$.

Since, in analogy to quantum mechanics, physical states are described by vectors (rays) from *complex* V, we see that for theories based on semisimple $\mathfrak{G} = \mathfrak{g}$, the notion of Cartan algebra is adequate for the description of simultaneously measurable quantities. However, if $\mathfrak{c}$ is not trivial then the assumption that its elements represent also some basic physical quantities imposes the following equivalent conditions: (a_1) the operators $\rho(H) \in \mathrm{End}\, V$ are semisimple for $H \in \mathfrak{c}$ (and in consequence for $H \in \mathfrak{H}$), and (a_2) the restriction $(\mathrm{Res}_{\mathfrak{c}}\, \rho, V)$ of ρ to $\mathfrak{c}$ is completely reducible [and in consequence so is (ρ, V)].

Note that if $\mathfrak{G}$ is real then (ρ, V) can be extended to the representation (ρ^{C}, V) of the complexification $\mathfrak{G}^{\mathrm{C}}$ of $\mathfrak{G}$. Therefore, in what follows, *we shall only deal with completely reducible complex representations of complex reductive Lie algebras*.

B. Weight spaces

For a given Cartan algebra $\mathfrak{H}$, the representation space V decomposes into a direct sum of common eigenspaces (weight spaces) of the operators $\rho(H)$, $H \in \mathfrak{H}$, i.e.,

$$V = \dot{\sum_{\lambda \in \underline{\Lambda}}} V(\lambda), \tag{2.1}$$

where $\underline{\Lambda} := \{\lambda \in \mathfrak{H}^* \mid V(\lambda) \neq 0\}$ denotes the set of weights and

$$V(\lambda) = \{v \in V \mid \rho(H) v = \langle \lambda, H \rangle v,\ H \in \mathfrak{H}\}. \tag{2.2}$$

The function $\mathfrak{H}^* \ni \lambda \to \Lambda(\lambda) := \dim V(\lambda) \in \mathbb{Z}$ is called the weight diagram of (ρ, V), and $\underline{\Lambda} = \mathrm{supp}\, \Lambda$. Two representations are equivalent iff their weight diagrams coincide and, moreover, the correspondence $(\rho, V) \to \Lambda$ has the following basic properties. If (ρ, V) is equal to a direct sum $(\rho_0 + \rho_1, V_0 + V_1)$, then

$$(V_0 + V_1)(\lambda) = V_0(\lambda) + V_1(\lambda), \tag{2.3}$$

$$\Lambda = \Lambda_0 + \Lambda_1, \quad \underline{\Lambda} = \underline{\Lambda}_0 \cup \underline{\Lambda}_1. \tag{2.4}$$

For the tensor product of representations ρ_0 and ρ_1, i.e., for $(\rho, V) = (\rho_0 \, \textcircled{T} \, \rho_1,\ V_0 \otimes V_1)$, where $\rho_0 \, \textcircled{T} \, \rho_1(X) := \rho_0(X) \otimes \mathbb{1} + \mathbb{1} \otimes \rho_1(X)$, $X \in \mathfrak{G}$, we have

$$(V_0 \otimes V_1)(\lambda) = \dot{\sum_{\lambda_0 + \lambda_1 = \lambda}} V_0(\lambda_0) \otimes V_1(\lambda_1), \tag{2.5}$$

$$\Lambda = \Lambda_0 * \Lambda_1, \quad \underline{\Lambda} = \underline{\Lambda}_0 + \underline{\Lambda}_1, \tag{2.6}$$

where $*$ denotes the convolution of functions, and $+$ denotes the algebraic sum of subsets. Finally, for the representation $(\rho^\wedge, V^*)$ contragredient to (ρ, V) where $\rho^\wedge(X) = -\rho(X)^*$ (A^* denotes the operator adjoint to $A \in \mathrm{End}\, V$ with respect to the natural duality between V and V^*), we obtain

$$V^*(\lambda) = V(-\lambda)^*, \tag{2.7}$$

$$\Lambda^\wedge(\lambda) = \Lambda(-\lambda), \quad \underline{\Lambda}^\wedge = -\underline{\Lambda}. \tag{2.8}$$

C. Compact forms

In the case of a semisimple complex $\mathfrak{G} = \mathfrak{g}$, the choice of a compact form $\mathfrak{g}_c \subset \mathfrak{g}$ determines, in every representation space V, the (Hermitian) scalar product in which the representation $(\mathrm{Res}_{\mathfrak{g}_c} \rho, V)$ is unitary, i.e., given by anti-Hermitian operators. If the center $\mathfrak{c}$ is not trivial, the situation is slightly more complicated. Namely, for $\mathfrak{G}_c = \mathfrak{g}_c + \mathfrak{c}_r$, where $\mathfrak{c}_r$ is a real form of $\mathfrak{c}$, the following conditions are equivalent: (b_1) there exists a scalar product in V such that $(\mathrm{Res}_{\mathfrak{G}_c} \rho, V)$ is unitary, and (b_2) the elements $\lambda \in \underline{\Lambda}$ assume imaginary values on $\mathfrak{c}_r$, i.e., $\mathfrak{c}^* \cap \underline{\Lambda} \subset \sqrt{-1}\, \mathfrak{c}_r^*$. Thus we see that the choice of a compact form $\mathfrak{G}_c$ leads to quantum mechanics iff the conditions (a_i) and (b_i), $i = 1$ or 2, are fulfilled. Let us notice that if a given reductive $\mathfrak{G}$ is a subalgebra of a semisimple $\mathfrak{g}_0$ then all representations of $\mathfrak{g}_0$ provide, by restriction to $\mathfrak{G}$, representations of the desired properties (a_i) and (b_i). It suggests that a theory based on a reductive Lie algebra $\mathfrak{G}$ should be considered as a preliminary step (before a proper semisimple $\mathfrak{g}_0$ is found).

Moreover, in our opinion, as long as complex representations are used, the whole complex Lie algebra $\mathfrak{G}$ may have a primary physical meaning, whereas a possibility to choose different compact forms $\mathfrak{G}_c$ reflects an additional symmetry of the theory. For example, it may appear that *the choice of* $\mathfrak{G}_c$ *has a definite physical meaning, similarly as in general relativity the choice of a Euclidean subgroup in the Lorentz group is equivalent to the choice of a reference system*. In what follows, this additional symmetry will not be discussed.

Besides a free choice of a real form, there exists a freedom in choosing a Cartan algebra. We are not able to decide if this freedom reflects any additional physical symmetry or if it is only a mathematical redundancy. Leaving this question open, we shall deal with a complex pair $(\mathfrak{G},\mathfrak{H})$, where $\mathfrak{H}$ is a *fixed* Cartan algebra.

D. Observables

Experimental data imply that only eigenstates of basic observables from $\rho(\mathfrak{H})$ can be observed, i.e., only weight vectors describe physical states. In other words, as observables should be considered operators $A \in \mathrm{End}\, V$ which possess an eigenbasis consisting of weight vectors. In consequence, A must belong to the commutant $\rho(\mathfrak{H})' \in \mathrm{End}\, V$. It means we assume that the set $\mathrm{Ob}(V)$ of observables acting in V satisfies

$$\rho(\mathfrak{H}) \subset \mathrm{Ob}(V) \subset \rho(\mathfrak{H})'. \tag{2.9}$$

(Here we admit also observables with a complex spectrum.)

The relation (2.9) can be interpreted in terms of superselection rules corresponding to the decomposition (2.1) of V. Namely, the space $\mathrm{End}\, V = V \otimes V^*$ carries the representation $\mathrm{ad}\, \rho = \rho \otimes \rho^\wedge$ of $\mathfrak{G}$ given by the formula

$$\mathrm{ad}\, \rho(X) \cdot A := [\rho(X), A], \quad X \in \mathfrak{G}, \quad A \in \mathrm{End}\, V. \tag{2.10}$$

Hence, the zero-weight space $(\mathrm{End}\, V)(0)$ of $\mathrm{ad}\, \rho$ coincides with the commutant $\rho(\mathfrak{H})'$. Moreover, according to the formulas (2.5) and (2.7) we have

$$\begin{aligned} \rho(\mathfrak{H})' &= (\mathrm{End}\, V)(0) = (V \otimes V^*)(0) \\ &= \dot{\sum_{\lambda \in \underline{\Lambda}}} V(\lambda) \otimes V^*(-\lambda) \\ &= \dot{\sum_{\lambda \in \underline{\Lambda}}} V(\lambda) \otimes V(\lambda)^* = \dot{\sum_{\lambda \in \underline{\Lambda}}} \mathrm{End}\, V(\lambda) \\ &= \{A \in \mathrm{End}\, V \,|\, A V(\lambda) \subset V(\lambda),\ \lambda \in \underline{\Lambda}\}. \end{aligned} \tag{2.11}$$

Thus $\rho(\mathfrak{H})'$ consists of operators reduced by the decomposition (2.1), i.e., elements of $\rho(\mathfrak{H})'$ commute with all projections $P_\lambda \in \mathrm{End}\, V$, $\lambda \in \underline{\Lambda}$, onto weight spaces $V(\lambda) \subset V$.

We have defined the set of observables $\mathrm{Ob}(V)$ separately for each representation (ρ, V). It might seem to be reasonable to look for a mathematical object that would describe physical quantities only in terms of the pair $(\mathfrak{G},\mathfrak{H})$—independently of its representations. In fact, many physical quantities can be described by elements of the enveloping algebra $\mathfrak{A}(\mathfrak{G})$. However, the mass operator M describing the mixing of scalar mesons fulfills $\rho(\mathfrak{A}(\mathfrak{G})) \not\ni M \in \rho(\mathfrak{H})'$. For details see Appendix A, where an additional characterization of the associative algebra $\rho(\mathfrak{H})''$ is also given.

In what follows, by an isotypic component we shall mean a maximal subrepresentation being a multiple of an irreducible one.

Lemma 1: Let (ρ, V) be a completely reducible representation of $(\mathfrak{G} = \mathfrak{g} + \mathfrak{c},\ \mathfrak{H} = \mathfrak{h} + \mathfrak{c})$, and let $V = \dot{\Sigma}_i V_i$ be the decomposition of V into isotypic components with respect to $(\mathrm{Res}_{\mathfrak{c}}\, \rho, V)$. Then $(\mathrm{Res}_{\mathfrak{g}}\, \rho, V)$ decomposes into the direct sum of representations (ρ_i, V_i) and, moreover,

$$\rho(\mathfrak{H})' = \dot{\sum_i} \rho_i(\mathfrak{h})'.$$

Proof: The spaces V_i are the weight spaces of the representation $\mathrm{Res}_{\mathfrak{c}}\, \rho$. Therefore, $\rho(\mathfrak{c})' = \dot{\Sigma}_i \mathrm{End}\, V_i$ [cf. (2.11)]. It implies that each space V_i is $\mathfrak{g}$-invariant since $\rho(X) \in \rho(\mathfrak{c})'$ for any $X \in \mathfrak{g}$. Moreover, applying (2.3), we get

$$\begin{aligned} \rho(\mathfrak{h})' &= (V \otimes V^*)(0) \\ &= \dot{\sum_i} (V_i \otimes V_i^*)(0) + \dot{\sum_{i \neq j}} (V_i \otimes V_j^*)(0) \\ &= \dot{\sum_i} \rho_i(\mathfrak{h})' + \dot{\sum_{i \neq j}} (V_i \otimes V_j^*)(0). \end{aligned} \tag{2.12}$$

The assertion is true since

$$\rho(\mathfrak{H})' = (\rho(\mathfrak{h}) + \rho(\mathfrak{c}))' = \rho(\mathfrak{h})' \cap \rho(\mathfrak{c})'. \quad \blacksquare$$

Corollary 1: The investigation of a commutant $\rho(\mathfrak{H})'$ reduces to the case of a complex semisimple Lie algebra. ■

From the mathematical point of view, a physical relation, connected with a given observable $A \in \mathrm{Ob}(V)$, is nothing else than an equation for matrix elements of A in a certain physical basis consisting of weight vectors. This basis need not be an eigenbasis of A; e.g., for the magnetic momentum operator μ_3 in su(6) theory, some experimental data are in agreement with diagonal expectation values of μ_3 in a basis that is not an eigenbasis of μ_3, see Sec. IV C.

In the sequel, we shall construct a canonical action of the Weyl group on the commutants $\rho_i(\mathfrak{h})'$ and show how this action provides various physical relations.

III. THE ACTION OF THE WEYL GROUP ON OBSERVABLES

In this section we shall show that the Weyl group $W(R)$ corresponding to a *complex semisimple* Lie algebra $\mathfrak{g}$ acts canonically in the zero-weight space $U(0)$ of an arbitrary $\mathfrak{g}$ representation (π,U). These $W(R)$ representations generalize a natural action of $W(R)$ in the Cartan algebra $\mathfrak{h}=\mathfrak{g}(0)$. Moreover, in the case where U is equal to End V, they provide the $W(R)$ action on $\rho(\mathfrak{h})'=(\mathrm{End}\, V)(0)$ and hence, also on observables.

Let us recall first some basic notions and results concerning $\mathfrak{g}$ and $W(R)$. Our exposition and notation is based mainly on Refs. 2 and 7. See also Refs. 8 and 9.

A. The groups *W(R)* and *A(R)*

Let $R\subset\mathfrak{h}^*$ be the root system of a pair $(\mathfrak{g},\mathfrak{h})$, i.e., a set of nonzero weights of its adjoint representation, let $A(R)\subset\mathrm{GL}(\mathfrak{h}^*)$ denote the group of automorphisms of R. The Weyl group $W(R)\subset A(R)$ is a (normal) subgroup generated by all reflections corresponding to the roots $\alpha\in R$. Both these groups are strictly connected with the Lie group Aut $\mathfrak{g}$ consisting of all automorphisms of $\mathfrak{g}$. We shall need the following facts describing this connection.

(1) The connected component of the unity in Aut $\mathfrak{g}$ coincides with the group of inner automorphisms, i.e., $(\mathrm{Aut}\,\mathfrak{g})_0=\mathrm{Int}\,\mathfrak{g}=\{$the group generated by operators $e^{\mathrm{ad}\,X}$, $X\in\mathfrak{g}\}$.

(2) The Lie algebra $L(\mathrm{Aut}\,\mathfrak{g})$ consists of all derivations of $\mathfrak{g}$. A semisimple Lie algebra has only inner derivations $\mathrm{ad}X$, $X\in\mathfrak{g}$, and its center is trivial. In consequence, $L(\mathrm{Aut}\,\mathfrak{g})=L(\mathrm{Int}\,\mathfrak{g})\simeq\mathfrak{g}$ and the exponential map exp: $\mathfrak{g}\to\mathrm{Int}\,\mathfrak{g}$ is given by $\exp X=e^{\mathrm{ad}\,X}$. Hence, for any $s\in\mathrm{Aut}\,\mathfrak{g}$ we have

$$\mathrm{Ad}_s\, X:=\frac{d}{dt}\Big|_{t=0} s(\exp tX)s^{-1}=\frac{d}{dt}\Big|_{t=0} se^{\mathrm{ad}\,tX}s^{-1}=\frac{d}{dt}\Big|_{t=0}\exp ts(X)=s(X). \tag{3.1}$$

(3) Let $\mathrm{Aut}(\mathfrak{g},\mathfrak{h}):=\{s\in\mathrm{Aut}\,\mathfrak{g}\,|\,s(\mathfrak{h})=\mathfrak{h}\}$ be a subgroup in Aut $\mathfrak{g}$ preserving $\mathfrak{h}$. Its Lie algebra satisfies $L(\mathrm{Aut}(\mathfrak{g},\mathfrak{h}))=\{X\in\mathfrak{g}\,|\,\mathrm{ad}\,X(\mathfrak{h})=\mathfrak{h}\}=N_{\mathfrak{g}}(\mathfrak{h})=\mathfrak{h}$.

(4) For any $\alpha\in R$ and $s\in\mathrm{Aut}(\mathfrak{g},\mathfrak{h})$ we have $\alpha\circ s^{-1}|_{\mathfrak{h}}\in R$. Thus the formula

$$\epsilon(s)\lambda:=\lambda\circ s^{-1}|_{\mathfrak{h}},\quad \lambda\in\mathfrak{h}^*, \tag{3.2}$$

defines a homomorphism ϵ: $\mathrm{Aut}(\mathfrak{g},\mathfrak{h})\to A(R)$.

(5) The homomorphism ϵ is surjective and, moreover, it satisfies

$$\epsilon^{-1}\{W(R)\}=\mathrm{Aut}(\mathfrak{g},\mathfrak{h})\cap\mathrm{Int}\,\mathfrak{g}=:\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}). \tag{3.3}$$

In other words, there exist the following short exact sequences:

$$\begin{array}{ccccccccc} 0 & \to & \ker\epsilon & \hookrightarrow & \mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}) & \to & W(R) & \to & 0,\\ & & \mathrm{id}\downarrow & & \downarrow & & \downarrow & & \\ 0 & \to & \ker\epsilon & \hookrightarrow & \mathrm{Aut}(\mathfrak{g},\mathfrak{h}) & \to & A(R) & \to & 0. \end{array} \tag{3.4}$$

(6) The connected component of the unity $(\mathrm{Aut}(\mathfrak{g},\mathfrak{h}))_0$ of $\mathrm{Aut}(\mathfrak{g},\mathfrak{h})$ fulfills

$$(\mathrm{Aut}(\mathfrak{g},\mathfrak{h}))_0=\ker\epsilon=\exp\mathfrak{h}=e^{\mathrm{ad}\,\mathfrak{h}}\subset\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}). \tag{3.5}$$

(7) There exist the canonical isomorphisms

$\mathrm{Aut}\,\mathfrak{g}/\mathrm{Int}\,\mathfrak{g}\simeq A(R)/W(R)\simeq\mathrm{Aut}\,D$,

where D is the Dynkin diagram of the root system R, and for simple Lie algebras,

$$\mathrm{Aut}\,D=\begin{cases}1, & \text{for } \mathrm{A}_1,\ \mathrm{B}_l\ (l\geqslant 2),\ \mathrm{C}_l\ (l\geqslant 3),\ \mathrm{E}_7,\mathrm{E}_8,\mathrm{F}_4,\mathrm{G}_2 \text{ (types)};\\ Z_2, & \text{for } \mathrm{A}_l\ (l\geqslant 2),\ \mathrm{D}_l\ (l\geqslant 5),\ \mathrm{E}_6;\\ S_3, & \text{for } \mathrm{D}_4.\end{cases}$$

B. The representation of *W(R)* in zero-weight spaces

Let σ: $G\to\mathrm{Int}\,\mathfrak{g}$ be a covering of the group Int $\mathfrak{g}$, i.e., G is a *connected* Lie group such that $L(G)=\mathfrak{g}$ and $\sigma=\mathrm{Ad}$. Clearly,

$$0\to Z(G)\hookrightarrow G\to\mathrm{Int}\,\mathfrak{g}\to 0, \tag{3.6}$$

where $Z(G)$ denotes the center of G, is a short exact sequence. Let us put $G_{\mathfrak{h}}:=\sigma^{-1}\{\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h})\}=\{g\in G\,|\,\mathrm{Ad}_g\mathfrak{h}\subset\mathfrak{h}\}$ and let $\exp_G$ denote the exponential map for G.

Lemma 2: Let the notation be as above. Then the following short exact sequence holds (cf. Ref. 8, Theorem 6.9.6.):

$$0\to\exp_G\mathfrak{h}\hookrightarrow G_{\mathfrak{h}}\overset{\epsilon\circ\sigma}{\to}W(R)\to 0. \tag{3.7}$$

Proof: According to the formula (3.5), $\ker(\epsilon\circ\sigma)=\sigma^{-1}\{\ker\epsilon\}$ coincides with $\sigma^{-1}\{e^{\mathrm{ad}\,\mathfrak{h}}\}$. To show that

$$\sigma^{-1}\{e^{\mathrm{ad}\,\mathfrak{h}}\}=\exp_G\mathfrak{h}, \tag{3.8}$$

let us recall the (nontrivial) known inclusion,

$Z(G)\subset\exp_G\mathfrak{h}$.

[Its proof can be reduced to the case of the *simply connected* universal covering, where even a stronger result holds—see formula (3.13) below.] Now note that for any subgroup $G'\subset G$ such that $Z(G)\subset G'$, the exact sequence (3.6) implies $G'=\sigma^{-1}\{\sigma(G')\}$. Hence, we obtain that

$$\exp_G\mathfrak{h}=\sigma^{-1}\{\sigma(\exp_G\mathfrak{h})\}=\sigma^{-1}\{e^{\mathrm{ad}\,\mathfrak{h}}\}$$

since $\sigma\circ\exp_G X=e^{\mathrm{ad}\,X}$ for any $X\in\mathfrak{g}$. ■

It is worth mentioning (although it is not essential for the sequel) that $\exp_G\mathfrak{h}$ coincides with the connected component of the unity $(G_{\mathfrak{h}})_0$ of $G_{\mathfrak{h}}$. Indeed, from (3.5) and (3.8) it follows that

$$(G_{\mathfrak{h}})_0 \subset \sigma^{-1}\{(\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}))_0\} = \sigma^{-1}\{e^{\mathrm{ad}\,\mathfrak{h}}\} = \exp_G \mathfrak{h}.$$

This inclusion implies the equality of both groups since $\exp_G \mathfrak{h}$ is connected.

The generalization of the natural action of $W(R)$ in the Cartan algebra $\mathfrak{h} = \mathfrak{g}(0)$, for the case of the zero-weight space of an arbitrary $\mathfrak{g}$ representation (π,U) can be done now as follows. We choose a covering group G such that the $\mathfrak{g}$ representation (π,U) is integrable to the representation $(\mathscr{P},U)$ of G, e.g., simply connected G. The formula $\pi\circ\sigma(g)(\cdot) = \mathscr{P}_g \pi(\cdot)\mathscr{P}_g^{-1}$ and the definition (3.2) imply that

$$\mathscr{P}_g U(\lambda) = U(\epsilon\circ\sigma(g)\lambda), \quad g\in G_{\mathfrak{h}}, \quad \lambda\in\underline{\Lambda}, \tag{3.9}$$

where $\underline{\Lambda}$ is the set of weights of (π,U). Thus, in particular, the zero-weight space $U(0)$ is $G_{\mathfrak{h}}$ invariant. Moreover, $\ker(\epsilon\circ\sigma) = \exp_G \mathfrak{h}$ acts trivially on $U(0)$ because $\mathscr{P}_{\exp_G X} = e^{\pi(X)}$. Now, since according to Lemma 2 the group $W(R)$ is isomorphic to $G_{\mathfrak{h}}/\ker(\epsilon\circ\sigma)$, the facts given above imply that $W(R)$ acts on $U(0)$. More precisely, the formula

$$W(R) \ni w \to \Pi_w := \mathscr{P}_g|_{U(0)} \in \mathrm{GL}(U(0)), \quad g\in(\epsilon\circ\sigma)^{-1}\{w\}, \tag{3.10}$$

defines correctly the representation $(\Pi,U(0))$ of the Weyl group $W(R)$.

Let us notice that operators $\mathscr{P}_g$, $g = \exp_G H \in \exp_G \mathfrak{h}$, are scalars on any weight space $U(\lambda)$, $\lambda\in\underline{\Lambda}$, $(\mathscr{P}_g|_{U(\lambda)} = e^{\pi(H)}|_{U(\lambda)} = e^{\langle\lambda,H\rangle}\mathrm{id}_{U(\lambda)})$. Thus, in a way similar to that above, we obtain for any weight $\lambda\in\underline{\Lambda}$, a projective representation $(\Pi^\lambda,U(\lambda))$ of the isotropy subgroup $W_\lambda \subset W(R)$ of the weight λ. In other words, for any weight $\lambda\in\underline{\Lambda}$, the operators

$$\Pi_w^\lambda = \mathscr{P}_w := \mathscr{P}_g|_{U(\lambda)}, \quad g\in(\epsilon\circ\sigma)^{-1}\{w\}, \quad w\in W_\lambda, \tag{3.11}$$

satisfy the relation

$$\mathscr{P}_{w_1}\mathscr{P}_{w_2}\mathscr{P}_{w_1w_2}^{-1} = \mathrm{const}\ \mathrm{id}_{U(\lambda)}, \quad w_1,w_2\in W_\lambda. \tag{3.12}$$

The construction of $W(R)$ action in $U(0)$ given above can be summarized as follows. First, the representation (π,U) of $\mathfrak{g}$ is integrable to a representation of the simply connected Lie group G. The subgroup $G_{\mathfrak{h}} \subset G$ permutes [cf. (3.9)] weight spaces $U(\lambda)$, $\lambda\in\underline{\Lambda}$, according to a natural mapping $\bar{\epsilon}$, which is described by the following diagram (rows form exact short sequences):

$$\begin{array}{ccccccc} 0\to & \exp_G\mathfrak{h} & \hookrightarrow & G_{\mathfrak{h}} & \xrightarrow{\bar\epsilon} & W(R)\to 0 \\ & \downarrow \mathrm{Ad} & & \downarrow \mathrm{Ad} & & \downarrow \mathrm{id} \\ 0\to & e^{\mathrm{ad}\,\mathfrak{h}} & \hookrightarrow & \mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}) & \xrightarrow{\epsilon} & W(R)\to 0. \end{array}$$

In particular, $U(0)$ is $G_{\mathfrak{h}}$ invariant. Moreover, the group $\exp_G\mathfrak{h}$ (connected component of $G_{\mathfrak{h}}$) acts as a scalar in every space $U(\lambda)$, and trivially in $U(0)$. Thus a natural representation of $G_{\mathfrak{h}}/\exp_G \mathfrak{h} \simeq W(R)$ in $U(0)$ is defined.

The construction of the $W(R)$ representation $(\Pi,U(0))$ can be realized without any changes for a complex *reductive* pair $(\mathfrak{G},\mathfrak{H})$. In fact, for $\mathfrak{G} = \mathfrak{g}+\mathfrak{c}$ the group $\mathrm{Int}\,\mathfrak{G} \simeq \mathrm{Int}\,\mathfrak{g}$ acts trivially on the center $\mathfrak{c}$, the group $\mathrm{Aut}_0(\mathfrak{G},\mathfrak{H})$ is isomorphic to $\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h})$ and, moreover, for a representation (π,U) of $(\mathfrak{G},\mathfrak{H})$, the zero-weight space $U(0)$ is contained in the trivial isotypic component $U^{(0)}$ of the representation $\mathrm{Res}_{\mathfrak{c}}\,\pi$. In consequence, the representation (π,U) of $\mathfrak{G}$ provides the same representation $(\Pi,U(0))$ of the Weyl group $W(R)$ as the representation $(\mathrm{Res}_{\mathfrak{g}}\,\pi,U^{(0)})$ of the semisimple part $\mathfrak{g}\subset\mathfrak{G}$. Thus *the generalization for the reductive case is useless*, except for computational reasons (cf. example 1 below and Appendix E).

In the case of classical Lie algebras, the practical computation of the operator Π_w can be achieved by taking an inner automorphism given by a matrix of a simple type, e.g., by a permutation matrix.

Example 1: Let $\mathfrak{g}$ be of the type A_{n-1}, $n>2$, i.e., $\mathfrak{g} = \mathrm{sl}(E)$, where E is the n-dimensional complex vector space. As Cartan algebra $\mathfrak{h}$, we can choose all traceless operators from End E that are diagonal in a fixed basis $\{e_i\}_{i=1}^n$ of E. Let $\{e^i\}_{i=1}^n$ denote the dual basis, and let $\mathscr{S}$ be a Young symmetrizer or, more generally, an element of the algebra generated by permutations of factors in the tensor product $T_q^p(E) := E^{\otimes p}\otimes(E^*)^{\otimes q}$, see, e.g., Ref. 10. It is easy to check that if

$$u = \mathscr{S}\cdot e_{i_1}\otimes\cdots\otimes e_{i_p}\otimes e^{j_1}\otimes\cdots\otimes e^{j_q}$$

belongs to a representation space $U\subset T_q^p(E)$, then $u\in U(0)$ iff there exists an integer k such that $\alpha_1-\beta_1 = \cdots = \alpha_n - \beta_n = k$, where α_i (resp. β_i) denotes the number of indices i_l, $l\in\overline{1,p}$ (resp. j_m, $m\in\overline{1,q}$) equal to i. Moreover, for $w\in W(R)\simeq S_n$ and $u\in U(0)$, we have

$$\Pi_w u = \mathscr{S}(\mathrm{sgn}\,w)^k e_{w(i_1)}\otimes\cdots\otimes e_{w(i_p)} \otimes e^{w(j_1)}\otimes\cdots\otimes e^{w(j_q)}.$$

Indeed, w is equal to $\epsilon(s)$, where $s(X) = gXg^{-1}$, and $g\in G = \mathrm{SL}(E)$ is given by $ge_i = z_i e_{w(i)}$, $z_i\in\mathbb{C}$, $z_1\cdot z_2\cdot\cdots\cdot z_n = \mathrm{sgn}\,w$. Hence, the scalar factor in $\mathscr{P}_g u$ is equal to

$$z_{i_1}\cdot z_{i_2}\cdot\cdots\cdot z_{i_p}\cdot z_{j_1}^{-1}\cdot\cdots\cdot z_{j_q}^{-1} = z_1^{\alpha_1-\beta_1}\cdot\cdots\cdot z_n^{\alpha_n-\beta_n} = (\mathrm{sgn}\,w)^k.$$

■

Analogous computations can be done for other classical Lie algebras and their tensor representations.

Besides the discussed $W(R)$ covering groups $G_{\mathfrak{h}}$, where G is a semisimple or reductive Lie group that covers Int $\mathfrak{g}$, there exist also finite $W(R)$ coverings that satisfy the formula analogous to (3.9). Clearly, if such a finite group is contained in a certain $G_{\mathfrak{h}}$, it provides a $W(R)$ representation equivalent to $(\Pi,U(0))$. See Appendix E and references therein.

For a certain class of representations, the constructed action $(\Pi,U(0))$ of the Weyl group $W(R)$ can be extended (although nonuniquely) onto the whole group of automorphisms $A(R)$. We shall describe this extension and its physical applications in the subsequent paper.[11] Let us also notice that there exists the canonical action Γ of $A(R)$ on the commutant $\mathfrak{A}(\mathfrak{h})'$ of $\mathfrak{A}(\mathfrak{h})$ in the enveloping algebra $\mathfrak{A}(\mathfrak{g})$. Moreover, for any representation (π,U) of $\mathfrak{g}$ and $w\in W(R)$, the operator Γ_w induces the operator Γ_w^π acting in $\pi(\mathfrak{A}(\mathfrak{h})')$. It can be shown that Π_w coincides with Γ_w^π on the space $\pi(\mathfrak{A}(\mathfrak{h})')$, see Appendix B for details.

Now, we shall prove a proposition characterizing those irreducible $\mathfrak{g}$ representations that have a nontrivial zero-weight space. In particular, it suggests that the group Int $\mathfrak{g}$ can be treated as a physical symmetry group corresponding to the Lie algebra $\mathfrak{g}$.

Let $P = P(R) \subset h^*$ [resp. $Q = Q(R) \subset P$] denote the group of weights (resp. radical weights) of the root system R, i.e., elements of $P(R)$ [resp. $Q(R)$] are linear combinations, with integer coefficients, of the fundamental weights $\omega_1,...,\omega_r$ (resp. roots $\alpha \in R$), cf. Ref. 2.

Proposition 1: For a given irreducible representation (ρ, V) of a semisimple Lie algebra $\mathfrak{g}$, the following conditions are equivalent: (i) (ρ, V) is integrable to a representation of Int $\mathfrak{g}$, (ii) 0 is a weight of (ρ, V) and (iii) one (or equivalently every) weight of (ρ, V) belongs to $Q(R)$.

Proof: The equivalence of (ii) and (iii) is the content of an exercise in Ref. 2 (Chap. VIII, Sec. 7, Ex. 3). Our solution of this exercise is given in Appendix C. Let us prove now that (i)$\Leftrightarrow$(iii). It is known (see, e.g., Ref. 8, Sec. 7.4.3.) that if G is the simply connected Lie group corresponding to $\mathfrak{g}$, then its center $Z(G)$ is given by

$$Z(G) = \exp_G(2\pi i Q^{\perp}), \tag{3.13}$$

where $Q^{\perp} = Q(R)^{\perp} := \{H \in \mathfrak{h} | \alpha(H) \in \mathbb{Z} \text{ for any } \alpha \in R\}$ denotes the group associated with Q. The representation (ρ, V) is integrable to a representation $(\mathscr{R}, V)$ of G and, since Int $\mathfrak{g} = G/Z(G)$, condition (i) is equivalent to

$$Z(G) \subset \ker \mathscr{R}. \tag{3.14}$$

On the other hand, the operator $\mathscr{R}(\exp_G(H)) = e^{\rho(H)}$, $H \in \mathfrak{h}$, acts on each weight space $V(\lambda)$, $\lambda \in \underline{\Lambda}$, as the scalar operator $e^{\langle \lambda, H \rangle}$. This fact, together with Eq. (3.13), implies that (3.14) is equivalent to the inclusion $\underline{\Lambda} \subset (Q^{\perp})^{\perp} = Q$. ■

For a given representation (π, U) of $\mathfrak{g}$, we have the unique decomposition $\pi = \pi_0 + \pi_1$, $U = U_0 + U_1$, where U_0 is the direct sum of all irreducible components which have 0 as a weight, and U_1 is the direct sum of all remaining components. Proposition 1 implies two important facts. First, the set of weights $\underline{\Lambda}$ of the representation (π, U) is the *disjoint* union of the corresponding weight sets, i.e., $\underline{\Lambda} = \underline{\Lambda}_0 \cup \underline{\Lambda}_1$, and a weight space

$$U(\lambda) = \begin{cases} U_0(\lambda), & \text{if } \lambda \in Q, \\ U_1(\lambda), & \text{otherwise.} \end{cases}$$

Second, the representation (π_0, U_0) is integrable to the representation $(\mathscr{P}, U_0)$ of the group Int $\mathfrak{g}$. It shows, *a posteriori*, that in our construction of $(\Pi, U_0(0))$, it is sufficient to consider the group $G = \text{Int } \mathfrak{g}$.

IV. PHYSICAL RELATIONS

Let (ρ, V) be a representation of $(\mathfrak{g}, \mathfrak{h})$ and let $\underline{\Lambda}$ denote its set of weights. We shall study now the $W(R)$ representation $(\Pi, U(0))$ in the case where $(\pi, U) = (\text{ad}\,\rho, \text{End } V)$, i.e., the zero-weight space $U(0) = \rho(\mathfrak{h})'$ has the structure of an associative algebra with the decomposition [cf. (2.11)]

$$U(0) = (\text{End } V)(0) = \dot{\sum_{\lambda \in \underline{\Lambda}}} \text{End } V(\lambda). \tag{4.1}$$

Let us notice that for *irreducible* ρ, Proposition 1 implies that every irreducible component of ad ρ has the *nontrivial* zero-weight space. In fact, formulas (2.6) and (2.8) show that the set of weights of ad ρ is equal to $\underline{\Lambda} + (-\underline{\Lambda})$. But for irreducible ρ, the last set is contained in $Q(R)$ since $\lambda_h - \underline{\Lambda} \subset Q$, where λ_h is the highest weight of ρ.

In the previous papers,[3] one of the authors proved that the natural action of $W(R)$ in $\mathfrak{h}^{\otimes n}$ provides many physical relations for a certain class of observables from $\rho(\mathfrak{h})'$. Using the $W(R)$ action Π in $\rho(\mathfrak{h})'$, we are able to obtain in this section an essential refinement and generalization of these results.

A. Properties of $W(R)$ representations in $\rho(\mathfrak{h})'$

Let, as before, $P_\lambda \in \text{End } V$, $\lambda \in \underline{\Lambda}$, denote the projection onto $V(\lambda)$.

Lemma 3: For any $w \in W(R)$, there exists an operator $\mathscr{R}_w \in \text{GL}(V)$ such that

$$\mathscr{R}_w V(\lambda) = V(w\lambda), \quad \lambda \in \underline{\Lambda}, \tag{4.2}$$

$$\Pi_w(A) = \mathscr{R}_w A \mathscr{R}_w^{-1}, \quad A \in U(0). \tag{4.3}$$

In consequence, Π_w is an automorphism of the algebra $U(0) = \rho(\mathfrak{h})'$ and, moreover, $\Pi_w(P_\lambda) = P_{w\lambda}$.

Proof: Let σ: $G \rightarrow \text{Int } \mathfrak{g}$ be a covering of Int $\mathfrak{g}$ such that the representation (ρ, V) is integrable to the representation $(\mathscr{R}, V)$ of G. The mapping $G \ni g \rightarrow \mathscr{P}_g := \mathscr{R}_g(\cdot)\mathscr{R}_g^{-1} \in \text{GL}(U)$ defines the G representation $(\mathscr{P}, U)$ corresponding to the representation (π, U) of $\mathfrak{g}$, i.e.,

$$\mathscr{P}_g(A) = \mathscr{R}_g A \mathscr{R}_g^{-1}, \quad A \in U, \quad g \in G. \tag{4.4}$$

Now, for a given $w \in W(R)$ let us take $g \in (\epsilon \circ \sigma)^{-1}\{w\}$ and let us set $\mathscr{R}_w := \mathscr{R}_g$. The assertion (4.2) is implied by the formula analogous to (3.9), whereas (4.3) follows from (4.4) and (3.10). ■

Lemma 3 implies the following.

(I) For any $\lambda \in \underline{\Lambda}$, the subspace End $V(\lambda) \subset U(0)$ [cf. (4.1)] is W_λ invariant, i.e., we have a representation $(\Pi_\lambda, \text{End } V(\lambda))$ of the isotropy group W_λ.

(II) For any $W(R)$ orbit $[\lambda] := W(R) \cdot \lambda \in W(R) \backslash \underline{\Lambda} = \{\text{set of } W(R) \text{ orbits in } \underline{\Lambda}\}$, the subspace

$$U_{[\lambda]} := \dot{\sum_{\lambda' \in [\lambda]}} \text{End } V(\lambda') \subset U(0)$$

is $W(R)$ invariant and the corresponding representation $(\Pi_{[\lambda]}, U_{[\lambda]})$ is induced by $(\Pi_\lambda, \text{End } V(\lambda))$, i.e.,

$$(\Pi_{[\lambda]}, U_{[\lambda]}) = \text{Ind}_{W_\lambda}^{W(R)}(\Pi_\lambda, \text{End } V(\lambda)). \tag{4.5}$$

(III) Since, according to formula (4.1), we have

$$(\Pi, U(0)) = \dot{\sum_{[\lambda] \in W(R) \backslash \underline{\Lambda}}} (\Pi_{[\lambda]}, U_{[\lambda]}), \tag{4.6}$$

therefore, (4.5) implies that

$$(\Pi, U(0)) = \dot{\sum_{[\lambda] \in W(R) \backslash \underline{\Lambda}}} \text{Ind}_{W_\lambda}^{W(R)}(\Pi_\lambda, \text{End } V(\lambda)). \tag{4.6'}$$

(IV) For any $\lambda \in \underline{\Lambda}$, the representation $(\Pi_\lambda, \text{End } V(\lambda))$ of W_λ is implemented by the projective W_λ representation $(\Pi^\lambda, V(\lambda))$ given by (3.11).

On the other hand, it is a well-known fact that any weight diagram is $W(R)$-invariant, i.e., $\Lambda(w\lambda) = \Lambda(\lambda)$, $\lambda \in \mathfrak{h}^*$, $w \in W(R)$, and hence, in particular, $w(\underline{\Lambda}) = \underline{\Lambda}$. It enables us to define, in the space $\mathbb{C}^{\underline{\Lambda}}$ (functions f: $\underline{\Lambda} \rightarrow \mathbb{C}$), the representation $(T,\mathbb{C}^{\underline{\Lambda}})$ of $W(R)$ by setting

$$(T_w f)(\lambda) := f(w^{-1}\lambda). \tag{4.7}$$

Note that besides the decomposition (4.6) of the representation $(\Pi, \rho(\mathfrak{h})')$, we also have

$$(T,\mathbb{C}^{\underline{\Lambda}}) = \dot{\sum_{X\in W(R)\setminus\underline{\Lambda}}} (T_X,\mathbb{C}^X).$$

Our aim is to show that for a given $W(R)$ orbit $X\subset\underline{\Lambda}$, algebraic invariants provide mappings intertwining the representations (Π_X,U_X) and $(T_X,\mathbb{C}^X)$ of $W(R)$. To this end, we must introduce some additional notations. Let j_λ: $V(\lambda)\to V$, $\lambda\in\underline{\Lambda}$ [resp. p_λ: $V\to V(\lambda)$] denote the natural injection (resp. projection) corresponding to the decomposition (2.1), i.e., $j_\lambda\circ p_\lambda = P_\lambda\in\text{End } V$. Let us set

$$A_\lambda := p_\lambda\circ A\circ j_\lambda\in\text{End } V(\lambda),\quad A\in\rho(\mathfrak{h})',\ \lambda\in\underline{\Lambda}. \tag{4.8}$$

The formula

$$F_A^n(\lambda) := \tau_n(A_\lambda) = \text{tr}(\wedge^n(A_\lambda)), \tag{4.9}$$

where τ_n, $n = 1,2,\ldots$, are algebraic invariants, defines the mappings

$$F^n\colon \rho(\mathfrak{h})'\ni A\to F_A^n\in\mathbb{C}^{\underline{\Lambda}}. \tag{4.10}$$

Proposition 2: Let (ρ,V) be a representation of a semisimple complex pair $(\mathfrak{g},\mathfrak{h})$. Then the mappings F^n given by the formulas (4.8)–(4.10) intertwine (nonlinearly for $n>1$) the representations $(\Pi,\rho(\mathfrak{h})')$ and $(T,\mathbb{C}^{\underline{\Lambda}})$ of the Weyl group $W(R)$.

Proof: For any $w\in W(R)$ and $\lambda\in\underline{\Lambda}$, let us define an isomorphism

$$\mathscr{R}_w(\lambda) := p_{w\lambda}\circ\mathscr{R}_w\circ j_\lambda\in\mathscr{L}(V(\lambda),V(w\lambda)), \tag{4.11}$$

where $\mathscr{R}_w$ is given by Lemma 3. Using the definitions (4.11) and (4.8), together with formulas (4.3), (4.2), and (4.1), we see that for each $w\in W(R)$, $\lambda\in\underline{\Lambda}$, and $A\in\mathfrak{g}(\mathfrak{h})'$, the following equality holds:

$$(\Pi_w(A))_{w\lambda} = \mathscr{R}_w(\lambda)\circ A_\lambda\circ\mathscr{R}_w(\lambda)^{-1}. \tag{4.12}$$

Hence, the definition of the mappings F^n implies that

$$F^n_{\Pi_w(A)}(w\lambda) = F_A^n(\lambda). \tag{4.13}$$

■

Remark 1: According to (4.1), any operator $A\in\rho(\mathfrak{h})' = (\text{End } V)(0)$ is described by independent components $A_\lambda\in\text{End } V(\lambda)$. In consequence, F^n maps surjectively $\rho(\mathfrak{h})'$ onto the set $\{f\in\mathbb{C}^{\underline{\Lambda}} | f(\lambda) = 0 \text{ if } \Lambda(\lambda) < n\}$. In particular, F^1 is a surjection and every restriction of F^n,

$$\text{Res}_X F^n := F^n|_{U_X}: U_X\to\mathbb{C}^X, \tag{4.14}$$

corresponding to a $W(R)$ orbit $X\subset\underline{\Lambda}$, is a surjection for $n\leqslant\Lambda(\lambda) = \dim V(\lambda)$, $\lambda\in X$. ■

For any $A\in\rho(\mathfrak{h})'$ and $\lambda\in\underline{\Lambda}$, let us denote by $\text{Sp}_A(\lambda)$: $=\text{Sp}_{A_\lambda}$ spectrum of the operator A_λ given by (4.8). The intertwining property (4.13) implies the following.

Corollary 2: For each $A\in\rho(\mathfrak{h})'$, $\lambda\in\underline{\Lambda}$, and $w\in W(R)$, it is true that

$$\text{Sp}_{\Pi_w(A)}(w\lambda) = \text{Sp}_A(\lambda).$$ ■

The practical content of Proposition 1 is mainly connected with the mappings $F := F^1$ or $\text{Res}_X F$. Namely, any $W(R)$-symmetry property of a given $A\in\rho(\mathfrak{h})'$, i.e., a relation of the type

$$\sum_{w\in W(R)} c_w\cdot\Pi_w(A) = 0,\quad c_w\in\mathbb{C}, \tag{4.15}$$

gives rise to an analogous property for the function F_A

$$\sum_{w\in W(R)} c_w\cdot F_A(w^{-1}\lambda) = 0,\quad \lambda\in\underline{\Lambda}. \tag{4.16}$$

The last equation is equivalent to the relations

$$\sum_{w\in W(R)} c_w\cdot\text{Res}_X F_{A_X}(w^{-1}\lambda) = 0,\quad \lambda\in X, \tag{4.17}$$

where operators

$$A_X := \dot{\sum_{\lambda\in X}} A_\lambda\in U_X\subset\rho(\mathfrak{h})',\quad X\subset\underline{\Lambda},$$

are constituents of A corresponding to the decomposition of $\underline{\Lambda}$ into $W(R)$ orbits. It is clear that the Eqs. (4.16) and (4.17) provide relations among diagonal matrix elements of $A\in\rho(\mathfrak{h})'$, in any basis consisting of weight vectors. More precisely, for a fixed orbit $X\subset\underline{\Lambda}$, we obtain a physical relation among diagonal expectation values corresponding to this $W(R)$ orbit.

For $n>1$, the mappings F^n: $\rho(\mathfrak{h})'\to\mathbb{C}^{\underline{\Lambda}}$ are not linear and hence the formula (4.13) reflects only very particular properties of a given operator $A\in\rho(\mathfrak{h})'$, e.g., such as $\Pi_w(A) = \pm A^{\pm 1}$. Thereby, the application of the mappings F^n, with $n>1$, seems to be very restricted.

In many cases a symmetry property of type (4.15) can be easily seen. For example, if $A\in\rho(\mathfrak{h})'$ belongs to a nontrivial irreducible subspace, with respect to the action of a *subgroup* $W\subset W(R)$, then $\Sigma_{w\in W}\Pi_w(A) = 0$. More generally, let Θ be a given irreducible representation of W and let $P_\Theta^\Pi\in\text{End}(\rho(\mathfrak{h})')$ denote the projection [acting in $\rho(\mathfrak{h})'$] onto the isotypic component of type Θ. It is known that

$$P_\Theta^\Pi = \frac{1}{|W|}\sum_{w\in W}\bar{\chi}_\Theta(w)\Pi_w,$$

where χ_Θ is the character of Θ. According to Proposition 2, we have $F^n\circ P_\Theta^n = P_\Theta^T\circ F^n$, where P_Θ^T is the projection analogous to P_Θ^Π but acting in $\mathbb{C}^{\underline{\Lambda}}$. Thus if a given operator A does not contain the component transformed by Θ, then

$$P_\Theta^T\circ F_A = 0, \tag{4.18}$$

It is obvious that (4.18) implies nontrivial relations iff $P_\Theta^T\neq 0$, i.e., iff the representation Θ is contained in $(\text{Res}_W\, T,\mathbb{C}^{\underline{\Lambda}})$.

Example 2: Let us consider the Lie algebra sl(3,$\mathbb{C}$) and let $X := S_3\cdot\lambda$ denote the S_3 orbit of a weight $\lambda = (\lambda_1,\lambda_2,\lambda_3)$ such that $\lambda_1>\lambda_2>\lambda_3$. Clearly, the representation $(T,\mathbb{C}^X)$ is equivalent to the left regular representation of $W(R) = S_3$. So, it contains the one-dimensional (alternative) representation of the signature (1^3). In consequence, formula (4.18) provides, for any operator $A\in\rho(\mathfrak{h})'$ which does not have a component transformed by (1^3), the following nontrivial relation:

$$\sum_{w\in S_3^-} F_A(w\lambda) = \sum_{w\in S_3^+} F_A(w\lambda),$$

where $S_3^-\subset S_3$ (resp. S_3^+) consists of all odd (resp. even) permutations. In the case of a mass operator in su(3) theory this equation is known as the Coleman–Glashow mass formula. ■

Now, we shall derive some generalizations of the Coleman–Glashow and Gell-Mann–Okubo mass formulas.

Example 3: Let (ρ,V) be a representation of $\mathrm{sl}(n,\mathbb{C})$ and let us consider an operator $A\in\rho(\mathfrak{h})'$ such that $A = B + C$, where $B\in\rho(\mathfrak{h})'$ is $W(R) = S_n$ invariant, and $C\in\rho(\mathfrak{h})'$ is transformed by the natural representation of S_n in the space

$$\mathbb{C}_0^n := \{x\in\mathbb{C}^n | x_1 + \cdots + x_n = 0\} \simeq \mathfrak{h}\subset \mathrm{sl}(n,\mathbb{C}),$$

i.e., $\mathrm{sgn}\ \mathbb{C}_0^n = (n-1,1)$. The operator B can be decomposed into operators $B_\lambda := b(S_n\cdot\lambda)\cdot \mathrm{id}_{V(\lambda)}$, where $b: S_n\setminus\underline{\Lambda}\to\mathbb{C}$ is a function on the set $S_n\setminus\underline{\Lambda}$ consisting of S_n orbits in $\underline{\Lambda}$. Let $W\subset S_n$ be a subgroup satisfying one of the following equivalent conditions: (1) W acts transitively on the set $\{1,...,n\}$, and (2) a W representation in $\mathbb{C}_0^n$ does not contain the trivial component, i.e., $\Sigma_{w\in W} wx = 0$ for any $x\in\mathbb{C}_0^n$. For example, W can be taken as any cyclic subgroup in S_n generated by a cycle of the length n. Proposition 2 implies that

$$\frac{1}{|W|}\sum_{w\in W} F_A(w\lambda) = b(X)\Lambda(\lambda),$$

where

$$X := S_n\cdot\lambda \text{ and } b(X) = \mathrm{tr}\, B_\lambda/\dim V(\lambda). \tag{4.19}$$

If A describes a (general) charge (e.g., electric charge, isospin, mass, magnetic momentum, etc.), formula (4.19) is equivalent to the following generalized Coleman–Glashow rule: *For all W orbits contained in a given S_n orbit, the corresponding average charge is the same.*

Obviously, if we know the quotients $b(X)/b(Y)$ for different S_n orbits X and Y, we obtain also the relations containing diagonal expectation values corresponding to different S_n orbits. For example, an operator A with $B = b\cdot \mathrm{id}_V$, $b = \mathrm{const}$, satisfies the generalized Gell-Mann–Okubo–Coleman–Glashow rule: *The average charge corresponding to a W orbit does not depend on the orbit.*

In particular, for any (resp. any irreducible) representation (ρ,V) of the Lie algebra $\mathfrak{G} = \mathrm{gl}(n,\mathbb{C})$, the basic observables from $\rho(\mathfrak{H})\subset\rho(\mathfrak{h})'$ fulfill the first (resp. second) rule. Analogous formulas can be easily derived for other classical Lie algebras. ■

B. Reductive case

Let V decompose into a direct sum $V = \Sigma_i V_i$ of subrepresentations ρ_i of semisimple Lie algebra $\mathfrak{g}$. Let $\Lambda = \Sigma_i \Lambda_i$ be the corresponding decomposition of the weight diagram. For a given operator $A\in\rho(\mathfrak{h})'$, let $A^i_j\in(V_i\otimes V_j^*)(0)$ denote a component of A from the decomposition (2.12). Any symmetry property of A gives rise to the same property of each operator A^i_j since the subspaces appearing in (2.12) are $W(R)$ invariant. In consequence, the functions $F_{A_i}\in\mathbb{C}^{\underline{\Lambda}_i}$, where $A_i := A^i_i$, reflecting symmetries of A, carry more information then the function $F_A\in\mathbb{C}^{\underline{\Lambda}}$ itself.

If, moreover, A is contained in $\Sigma_i\rho_i(\mathfrak{h})'$ then $F_{A_i}(\lambda)$, $\lambda\in\underline{\Lambda}_i$, is the sum of certain eigenvalues of the operator A itself (and not only of the operators A_i). The components A^i_j with $i\neq j$ disappear, for instance, if $\mathfrak{g}$ is the semisimple part of a reductive $\mathfrak{G} = \mathfrak{g} + \mathfrak{c}$, the spaces V_i are isotypic components of $\mathrm{Res}_{\mathfrak{c}}\rho$, and the operator A belongs to $\rho(\mathfrak{H})$, $\mathfrak{H} = \mathfrak{h} + \mathfrak{c}$ (compare Lemma 1).

An analogous decomposition of a given $A\in\rho(\mathfrak{h})'$ can be also used in the case of an arbitrary representation (ρ,V) provided A has a symmetry property with respect to the Weyl group $W(\widetilde{R})$ of a reductive pair $(\widetilde{\mathfrak{G}},\widetilde{\mathfrak{H}})$ such that $\widetilde{\mathfrak{G}}\subset\mathfrak{g}$, $\widetilde{\mathfrak{H}}\subset\mathfrak{h}$. In fact, treating A as an element of $\widetilde{\rho}(\widetilde{\mathfrak{H}})'$, where $\widetilde{\rho} := \mathrm{Res}_{\widetilde{\mathfrak{G}}}\,\rho$ is, in general, reducible, we can consider the operators A_i corresponding to the decomposition (2.12) of $\rho(\widetilde{\mathfrak{H}})'$ into $W(\widetilde{R})$-invariant subspaces. Such a situation appears, e.g., in the case of a broken symmetry.

C. The broken symmetry case

A broken symmetry is usually described by two reductive pairs, $(\mathfrak{G},\mathfrak{H})$ and $(\widetilde{\mathfrak{G}},\widetilde{\mathfrak{H}})$ such that $\widetilde{\mathfrak{G}}\subset\mathfrak{G}$, $\widetilde{\mathfrak{G}}\neq\mathfrak{G}$, $\widetilde{\mathfrak{H}}\subset\mathfrak{H}$. Moreover, one assumes that physical states correspond to vectors from a restriction $(\tilde{\rho},V)$ of a $\mathfrak{G}$ representation (ρ,V) to $\widetilde{\mathfrak{G}}$, i.e., $(\tilde{\rho},V) := (\mathrm{Res}_{\widetilde{\mathfrak{G}}}\,\rho,V) = \Sigma_i(\tilde{\rho}_i,V_i)$. Thereby, the $W(\widetilde{R})$-transformation properties of observables from $\tilde{\rho}(\widetilde{\mathfrak{H}})'$ should be analyzed as described above.

Let us recall that there exist two different kinds of embeddings of $(\widetilde{\mathfrak{G}},\widetilde{\mathfrak{H}})$ into $(\mathfrak{G},\mathfrak{H})$. See Ref. 12. Namely, in the case of a regular embedding, the corresponding root systems satisfy $\widetilde{R}\subset R$. In consequence, $W(\widetilde{R})$ is a proper subgroup of $W(R)$ and, moreover, the $W(\widetilde{R})$ representation $\widetilde{\Pi}$ in $\tilde{\rho}(\widetilde{\mathfrak{H}})'$ restricted to $\rho(\mathfrak{H})'$ coincides with the restriction $\mathrm{Res}_{W(\widetilde{R})}\,\Pi$ of the $W(R)$ representation Π in $\rho(\mathfrak{H})'$. For irregular embeddings, the Weyl groups cannot be compared. Indeed, if $W(\widetilde{R})$ were a subgroup of $W(R)$, then it would be generated by a subset of R.

Note also that for both kinds of embeddings, the cases $\widetilde{\mathfrak{H}} = \mathfrak{H}$ and $\widetilde{\mathfrak{H}}\neq\mathfrak{H}$ differ essentially. In the first case, any $\tilde{\rho}_i$-weight space $V_i(\lambda)$, $\lambda\in\underline{\widetilde{\Lambda}}_i\subset\underline{\widetilde{\Lambda}} = \underline{\Lambda}$, is contained in the ρ-weight space $V(\lambda)$, whereas in the second case some weight spaces $V_i(\tilde{\lambda})$, $\tilde{\lambda}\in\underline{\widetilde{\Lambda}}_i$, are not contained in any ρ-weight space $V(\lambda)$, $\lambda\in\underline{\Lambda}$.

Example 4: The classical theory of spin-flavor symmetry uses the algebra $\mathfrak{G} = \mathfrak{g} = \mathrm{sl}(6,\mathbb{C})$ [more precisely its compact form $\mathfrak{g}_c = \mathrm{su}(6)$] and the irreducible $\mathrm{sl}(6,\mathbb{C})$ representation of the signature (3,0,0,0,0,0). The (third component of) magnetic momentum operator μ_3 is built by means of the electric charge operator Q and the spin operator σ_3 in such a way that $\sigma_3\in\mathfrak{h}_c\subset\mathfrak{h}$, where $\mathfrak{h}_c$ is the Cartan algebra of $\mathfrak{g}_c$.

An agreement with experimental data, for the ratio $\mu_3(p)/\mu_3(n)$, can be obtained if one takes, instead of the eigenvalues of $\rho(\mu_3)$, the diagonal matrix elements of $\rho(\mu_3)$ in a basis consisting of weight vectors of the restriction $(\tilde{\rho},V) := (\mathrm{Res}_{\widetilde{\mathfrak{G}}}\,\rho,V)$ to $\widetilde{\mathfrak{G}} = \tilde{\mathfrak{g}}\simeq\mathrm{sl}(2,\mathbb{C}) + \mathrm{sl}(3,\mathbb{C})$. The regular embedding of $\widetilde{\mathfrak{G}}$ into $\mathfrak{G}$ is made by means of the natural representations of these Lie algebras in the space $\mathbb{C}^2\otimes\mathbb{C}^3\simeq\mathbb{C}^6$. Abusing the notation, one usually writes $\mu_3 = \sigma_3\otimes Q\notin\tilde{\mathfrak{h}}\subset\tilde{\mathfrak{g}}$.

The corresponding Weyl group $W(\widetilde{R})$ is isomorphic to $S_2\times S_3$, and we see that for $\rho(\mu_3)\in\rho(\mathfrak{h})'\subset\tilde{\rho}(\tilde{\mathfrak{h}})'$, the following symmetry property is satisfied:

$$\sum_{w\in S_2\times\{1\}} \widetilde{\Pi}_w\,\rho(\mu_3) = 0. \tag{4.20}$$

It means that the states of the opposite spin projections σ_3 provide the opposite projections μ_3, independently of the definition of a physical basis (as ρ or $\tilde{\rho}_i$-weight vectors).

Moreover, the equation $\Sigma_{w\in S_3^+} wQ = 0$ implies the additional relations for the operator $\rho(\mu_3)$. In consequence, using the decomposition of 56-plet of baryons into $\mathrm{sl}(2,\mathbb{C}) + \mathrm{sl}(3,\mathbb{C})$-irreducible subspaces, we obtain the generalized Coleman–Glashow relations for magnetic momenta (compare example 3)

$$\mu_3(p) + \mu_3(\Sigma^-) + \mu_3(\Xi^0) = 0,$$
$$\mu_3(\Delta^+) + \mu_3(\Sigma^{*-}) + \mu_3(\Xi^{*0}) = 0,$$
$$\mu_3(n) + \mu_3(\Sigma^+) + \mu_3(\Xi^-) = 0,$$
$$\mu_3(\Delta^0) + \mu_3(\Sigma^{*+}) + \mu_3(\Xi^{*-}) = 0,$$
$$\mu_3(\Delta^{++}) + \mu_3(\Delta^-) + \mu_3(\Omega^-) = 0,$$

where in every equation all the particles have the same projection σ_3. According to Ref. 13, the recent experimental data are the following (in units $eh/2m_pc$):

$$\mu_3(p) = 2.7928, \quad \mu_3(\Sigma^-) = -1.10 \pm 0.05,$$
$$\mu_3(\Xi^0) = -1.250 \pm 0.014,$$
$$\mu_3(n) = -1.913, \quad \mu_3(\Sigma^+) = 2.379 \pm 0.02,$$
$$\mu_3(\Xi^-) = 1.85 \pm 0.75.$$

Thus, besides the proton–neutron ratio, an agreement of the su(6) theory is rather problematic.

Let us notice, however, that the generalization of this theory for a larger number of flavors does not provide the relations given above. Namely, for $n > 3$, the definition of the operator Q implies that $\mathfrak{G}$ must be taken as the reductive Lie algebra $\mathrm{sl}(2,\mathbb{C}) + \mathrm{sl}(n,\mathbb{C}) + u(1)$ and that $\Sigma_{w\in W} wQ \neq 0$ for any subgroup $W \subset S_n$. Thereby, additional equations, apart from the ones given by the formula analogous to (4.20), can be obtained only for more complicated operators from $\mathfrak{H} \subset \mathfrak{G}$. For example, the operator $\sigma_3 \otimes (Q - mB)$, where B is the baryon number operator and $m = \frac{1}{2}$ (resp. $\frac{1}{2} - 3/2n$) for even (resp. odd) n, satisfies the generalized Coleman–Glashow relations. ■

The examples given above illustrate how one can obtain some physical relations for a given observable. On the other hand, Proposition 2 shows that by studying the representations $(T,\mathbb{C}^X)$, where $X \subset \underline{\Lambda}$ is an orbit of a subgroup $W \subset W(R)$, we get many physical relations that can appear in the framework of a considered symmetry. Before we give some examples of such an application of Proposition 2, let us derive some useful properties of transitive actions of finite groups and corresponding representations in the functions' spaces.

V. DETAILED RESULTS AND EXAMPLES

A. Actions of finite groups

Let W be an arbitrary finite group acting on a finite set X. Let W_{x_0} denote the isotropy group of $x_0 \in X$. If W acts transitively on X, then the mapping $wW_{x_0} \to wx_0$, $w \in W$, provides the isomorphism $W/W_{x_0} \simeq X$ of W actions and, moreover, each isotropy group W_x of $x \in X$ is conjugated with W_{x_0} (since $W_{wx_0} = wW_{x_0}w^{-1}$). On the other hand, W/W_0 is isomorphic to W/wW_0w^{-1} for every subgroup $W_0 \subset W$. Thus we obtain the following lemma.

Lemma 4: There exists one-to-one correspondence among transitive actions of a finite group W and conjugancy classes of subgroups in W. ■

Let $(T,\mathbb{C}^X)$ be the representation of W given by the formula (4.7). It is clear that isomorphic W actions generate equivalent representations. Moreover, the representation $(T,\mathbb{C}^X)$ is equivalent to its contragredient since the formula

$$B(f,g) := \sum_{x\in X} f(x)g(x)$$

defines a nondegenerate W-invariant bilinear form on $\mathbb{C}^X$. Our next lemma is a generalization of the classical Frobenius theorem about the regular representation and provides an efficient irreducibility criterion.

Lemma 5: Let W be a finite group acting transitively on X and let (Θ,E) be a representation of W. Then (i) the intertwining number $c(E,\mathbb{C}^X) := \dim \mathscr{L}_W(E,\mathbb{C}^X)$ is equal to the dimension of the subspace $E_0 \subset E$ consisting of the W_{x_0}-invariant vectors, $x_0 \in X$, and (ii) if $\dim E_0 = 1$ and (Θ,E) can be embedded in $(T,\mathbb{C}^X)$, then (Θ,E) is irreducible.

Proof: (i) Since $c(E,\mathbb{C}^X) = c(\mathbb{C}^X,E)$, we may consider the intertwining operators $\mathscr{T}: \mathbb{C}^X \to E$. Each $\mathscr{T}$ is uniquely determined by the vector $e := \mathscr{T}\cdot 1_{x_0}$, where $1_{x_0} \in \mathbb{C}^X$ is the characteristic function of the subset $\{x_0\} \subset X$. The vector e is W_{x_0} invariant since

$$\Theta_w e = \mathscr{T} \circ T_w \cdot 1_{x_0} = \mathscr{T} \cdot 1_{w^{-1}x_0} = e, \quad w \in W_{x_0}.$$

(ii) It follows immediately from part (i). ■

Remark 2: Part (i) can be also obtained from the Frobenius reciprocity formula, see, e.g., Ref. 14, by noticing that the representation $(T,\mathbb{C}^X)$ is induced by the trivial representation $\mathbf{1}$ of the subgroup W_{x_0},

$$c(\Theta,T) = c(\Theta,\mathrm{Ind}^W_{W_{x_0}} \mathbf{1}) = c(\mathrm{Res}^W_{W_{x_0}} \Theta,\mathbf{1}). \quad ■$$

In the case of a classical Lie algebra, the Weyl group $W(R)$ is an extention of a permutation group S_n and, moreover, the action of S_n on the group of weights $P(R) \subset (1/q)\mathbb{Z}^n$ ($\mathbb{Z}$ integers, $0 < q \in \mathbb{Z}$) can be described by the natural formula,

$$\sigma(x_1,...,x_n) = (x_{\sigma^{-1}(1)},...,x_{\sigma^{-1}(n)}), \quad \sigma \in S_n, \quad x_i \in (1/q)\mathbb{Z}.$$

See Appendix D for details. Therefore, we shall study more closely the natural action of S_n in the Cartesian product Z^n of an arbitrary set Z.

Every S_n orbit $X \subset Z^n$ contains an element of the form $x_0 = (\underbrace{z_1,...,z_1}_{n_1},...,\underbrace{z_k,...,z_k}_{n_k})$, where all the elements $z_1,...,z_k$ are different and $n_1 \geqslant \cdots \geqslant n_k > 0$. We shall call the partition $[n_1,...,n_k]$ of the integer n the type of orbit $X = S_n \cdot x_0 \subset Z^n$.

Lemma 6: For natural action of S_n on the set Z^n, two S_n orbits are isomorphic iff their types coincide. In consequence, each representation $(T,\mathbb{C}^X)$, where X is an S_n orbit in Z^n, can be labeled by the corresponding type of orbit X. ■

A simple application of Lemma 5 is the following useful example.

Example 5: Let us treat the set $\mathbb{C}^n_0 := \{y \in \mathbb{C}^n | y_1 + \cdots + y_n = 0\}$ as the space of the irreducible S_n representation of the signature $(n-1,1)$. Then for any

S_n orbit $X\subset Z^n$ of the type $[n_1,...,n_k]$ we have

$$c(\mathbb{C}_0^n,\mathbb{C}^X)=k-1.$$

In fact, the isotropy group W_{x_0} of the element x_0 given above consists of permutations preserving each of the following subsets of the set $\overline{1,n}$: $\{1,2,...,n_1\}$, $\{n_1+1,...,n_1+n_2\}$,... . Thus an element $y\in\mathbb{C}_0^n$ is W_{x_0} invariant iff $y_1=\cdots=y_{n_1}$, $y_{n_1+1}=\cdots=y_{n_1+n_2}$,... . ■

Finally, let us prove the following lemma.

Lemma 7: Each representation (Θ,E) of the group S_n is equivalent to its contragredient one.

Proof: Every element $\sigma\in S_n$ is conjugate with σ^{-1} since they both have the same decomposition into the disjoint cycles. Thus $\chi_\sigma=\chi_{\sigma^{-1}}=\overline{\chi_\sigma}$, i.e., the characters of Θ and $\Theta^\wedge$ coincide. ■

B. Relations for three-particle and particle–antiparticle states

Let $n\geqslant 3$ and let us consider the S_n orbits in $\mathbb{Z}^n$ [resp. $(1/n)\mathbb{Z}^n$] of the following four types: $[n]$, $[n-1,1]$, $[n-2,1^2]=[n-2,1,1]$, and $[n-3,3]$. These orbits are sufficient to describe one-particle, three-particle, and particle–antiparticle states in theories based on $\mathfrak{G}=\mathrm{gl}(n,\mathbb{C})$ [resp. $\mathrm{sl}(n,\mathbb{C})$]. In fact, $\mathrm{gl}(n,\mathbb{C})$ representations of signatures $(0,...,0)$, $(1,0,...,0)$, $(3,0,...,0)$, $(2,1,0,...,0)$, $(1,1,1,0,...,0)$, or $(1,0,...,0,-1)$ provide S_n orbits only of the types mentioned above.

In what follows, for the simplicity of the notation, we shall denote S_n representations by the corresponding signature or type (see Lemma 6). For the trivial orbit $X=\{\lambda\}$, the representation in $\mathbb{C}^X$ is trivial, i.e., we have $[n]=(n)$.

The representation generated by the orbit of the type $[n-1,1]$ is equivalent to the S_n representation in the space

$$\mathbb{C}^n=\left\{(a_1,...,a_n)\,|\,a_i=b+y_i:\sum_i y_i=0\right\},$$

$$\text{i.e., } [n-1,1]=(n)+(n-1,1). \tag{5.1}$$

The representation $[n-2,1^2]$ coincides with the S_n representation in the space of diagonalless matrices $\{[a_{ij}]\,|\,i\neq j, a_{ij}\in\mathbb{C}\}$ and can be decomposed into the irreducible S_n components as follows:

$$a_{ij}=b+c_{ij}+d_{ij}+e_{ij}+f_{ij},\quad \text{for } n\geqslant 3, \tag{5.2}$$

where $b=\text{const}$; $c_{ij}=x_i-x_j$, $\Sigma_i x_i=0$; $d_{ij}=y_i+y_j$, $\Sigma_i y_i=0$; $e_{ij}=-e_{ji}$, $\Sigma_{i,i\neq j}e_{ij}=0$; $f_{ij}=f_{ji}$, $\Sigma_{i,i\neq j}f_{ij}=0$, hence $f_{ij}=0$ for $n=3$. The irreducibility of the subspaces $\{[e_{ij}]\}$ and $\{[f_{ij}]\}$ follows from Lemma 5. Indeed, $[n-2,1^2]$ corresponds to $W_{x_0}=\{\sigma\in S_n\,|\,\sigma(1)=1,\ \sigma(2)=2\}$ and hence $[a_{ij}]$ is W_{x_0} invariant iff there exist $a_0,...,a_4\in\mathbb{C}$ such that for i, $j\in\overline{3,n}$, $i\neq j$, we have $a_{ij}=a_0$, $a_{i1}=a_1$, $a_{i2}=a_2$, $a_{1i}=a_3$, $a_{2i}=a_4$. So, $[e_{ij}]$ (resp. $[f_{ij}]$) is W_{x_0} invariant if $e_{21}=-e_{21}$ and $e_{i1}=-e_{i2}=-e_{1j}=e_{2j}=[1/(n-2)]e_{12}$, $e_{ij}=0$ (resp. $f_{12}=f_{21}$, $f_{i1}=f_{i2}=f_{j1}=f_{j2}=-[1/(n-2]f_{12}$, $f_{ij}=[2/(n-2)(n-3)]f_{12}$), $i,j\in\overline{3,n}$, $i\neq j$. It can be proved that the space $\{[e_{ij}]\}$ (resp. $\{[f_{ij}]\}$) carries the S_n representation of signature $(n-2,1^2)$ [resp. $(n-2,2)$]. This identification is of no significance for our computations, but we shall use it to simplify the notation. Thus we have

$$[n-2,1^2]=(n)+2\cdot(n-1,1)+(n-2,1^2)+(n-2,2),$$
$$\text{for } n\geqslant 4, \tag{5.3}$$

whereas

$$[1^3]=(3)+2\cdot(2,1)+(1^3).$$

Finally, the representation $[n-3,3]$ is equivalent to the S_n representation in the space of the symmetric matrices $\{[a_{ijk}]\,|\,i,j,k \text{ different}, a_{ijk}=a_{(ijk)}\in\mathbb{C}\}$, and has the following decomposition into irreducible components:

$$a_{ijk}=b+c_{ijk}+d_{ijk}+e_{ijk}, \tag{5.4}$$

where $b=\text{const}$; $c_{ijk}=x_i+x_j+x_k$, $\Sigma_i x_i=0$; $d_{ijk}=f_{ij}+f_{jk}+f_{ki}$, $f_{ij}=f_{ji}$, $\Sigma_{i,i\neq j}f_{ij}=0$; $e_{ijk}=e_{(ijk)}$, $\Sigma_{i,j\neq i\neq k}e_{ijk}=0$; and, moreover, $c_{ijk}=0$ for $n=3$, $d_{ijk}=0$ for $n=3,4$, and $e_{ijk}=0$ for $n=3,4,5$. The irreducibility of the space $\{[e_{ijk}]\}$ can be again seen from Lemma 5. Moreover, it can be shown that this space corresponds to the S_n signature $(n-3,3)$. So, we get

$$[n-3,3]=(n)+(n-1,1)+(n-2,2)+(n-3,3),$$
$$\text{for } n\geqslant 6, \tag{5.5}$$

whereas

$$[3]=(3),\quad [3,1]=(4)+(3,1),$$
$$[3,2]=(5)+(4,1)+(3,2).$$

The obtained formulas for the functions F_A, $A\in\rho(\mathfrak{H})'$, $\mathfrak{H}\subset\mathrm{gl}(n,\mathbb{C})$ enable us to derive immediately some possible physical relations (the more detailed analysis of the considered $\mathrm{gl}(n,\mathbb{C})$ states is given in Ref. 5).

Example 6: Let, for $n=4$, u,d,s,c denote the basic vectors of the representation $(1,0,0,0)$ (one-particle states). In particular, in $u(4)$-flavor theory u,d,s,c are one-quark states. If we consider the three-particle representation (ρ,V) of signature $(3,0,0,0)$ or $(2,1,0,0)$, the weight spaces corresponding to the orbit $S_4\cdot\lambda$, where $\lambda=(2,1,0,0)$, are one dimensional. Moreover, the eigenvalues of an operator $A\in\rho(\mathfrak{H})'$ without the component transformed by the S_4 representation $(2,1^2)$ satisfy three linearly independent equations [see (4.18) and the table of characters of S_4],

$$2\alpha_{uud}-2\alpha_{ddu}+\alpha_{ssu}-\alpha_{uus}+\alpha_{ccu}-\alpha_{uuc}$$
$$+\alpha_{dds}-\alpha_{ssd}+\alpha_{ddc}-\alpha_{ccd}=0,$$
$$\alpha_{uud}-\alpha_{ddu}+2\alpha_{ssu}-2\alpha_{uus}-\alpha_{ccu}+\alpha_{uuc}$$
$$+\alpha_{dds}-\alpha_{ssd}+d_{ccs}-\alpha_{ssc}=0,$$
$$\alpha_{uud}-\alpha_{ddu}+\alpha_{ssu}-\alpha_{uus}+2\alpha_{dds}-2\alpha_{ssd}$$
$$-\alpha_{ddc}+\alpha_{ccd}-\alpha_{ccs}+\alpha_{ssc}=0,$$

where all three-particle eigenstates *uud*, *ddu*, etc., have the same symmetry, i.e., $uud=u\otimes u\otimes d+u\otimes d\otimes u+d\otimes u\otimes u$ is symmetrical if λ is the weight of the representation $(3,0,0,0)$ or, e.g., $uud=u\otimes u\otimes d-u\otimes d\otimes u$ has the mixed symmetry if we consider the representation $(2,1,0,0)$.

The analogous equations for an operator $A\in\rho(\mathfrak{H})'$ without the component transformed by the S_4 representation (2^2) are the following [see the relations for matrix elements in (5.2)]:

$$\alpha_{uud}+\alpha_{ddu}+\alpha_{ssc}+\alpha_{ccs}$$
$$=\alpha_{uus}+\alpha_{ssu}+\alpha_{ddc}+\alpha_{ccd}$$
$$=\alpha_{uuc}+\alpha_{ccu}+\alpha_{dds}+\alpha_{ssd}=0.$$

The case where $A\in(4)+(3,1)\subset\rho(\mathfrak{H})'$ is described by the generalized Coleman–Glashow rule, which implies five linearly independent equations:

$$\alpha_{uud}+\alpha_{dds}+\alpha_{ssc}+\alpha_{ccu}$$
$$=\alpha_{uus}+\alpha_{dds}+\alpha_{ssd}+\alpha_{ccu}$$
$$=\alpha_{uus}+\alpha_{ddu}+\alpha_{ssc}+\alpha_{ccd}$$
$$=\alpha_{uuc}+\alpha_{dds}+\alpha_{ssu}+\alpha_{ccd}$$
$$=\alpha_{uud}+\alpha_{ddc}+\alpha_{ssu}+\alpha_{ccs}$$
$$=\alpha_{uuc}+\alpha_{ddu}+\alpha_{ssd}+\alpha_{ccs}.$$ ■

Let us consider now the representation $(1,0,...,0,-1)$ of $\mathrm{gl}(n,\mathbb{C})$ which is usually used to describe particle–antiparticle states. If we deal with a charge that is independent from the particle–antiparticle reflection (e.g., mass of mesons), we must use operators $A\in\rho(\mathfrak{H})'$ providing symmetrical matrices $[a_{ij}]$. Therefore, according to (5.2) and (5.3), the representation $(n-2,1^2)$ and one of the representations $(n-1,1)$ are excluded. If, moreover, the charge cannot be negative then, for $n>3$, there are possible two fundamental kinds of physical relations for eigenvalues corresponding to the one-dimensional weight spaces $V(\lambda)$, $\lambda\in S_n\cdot(1,0,...,0,-1)$ [compare (5.2)],

$$\sum_{i,i\neq j}a_{ij}=\sum_{i,i\neq k}a_{ik},\quad a_{ij}=a_{ji},$$
$$\text{if } A\in(n)+(n-2,2)\in\rho(\mathfrak{H})', \tag{5.6}$$

and

$$a_{ij}-a_{ik}=a_{lj}-a_{lk},\quad a_{ij}=a_{ji},$$
$$\text{if } A\in(n)+(n-1,1)\subset\rho(\mathfrak{H})' \tag{5.7}$$

(to simplify the notation we do not choose linearly independent equations). Thus we see that if n increases new terms appear in each sum of Eqs. (5.6), whereas in the second case, it increases only the number of the relations in (5.7).

Example 7: Let us consider the flavor theories with 4, 5, or 6 flavors. If we assume that the meson mass operator M (or its square) is of the form $M\in(n)+(n-1,1)\subset\rho(\mathfrak{H})'$, then Eqs. (5.7) imply, e.g., the following linearly independent mass formulas:

$$m_{s\bar d}-m_{s\bar u}=m_{c\bar d}-m_{c\bar u}=m_{b\bar d}-m_{b\bar u}=m_{t\bar d}-m_{t\bar u},$$
$$m_{c\bar s}-m_{c\bar u}=m_{b\bar s}-m_{b\bar u}=m_{t\bar s}-m_{t\bar u},$$
$$m_{d\bar c}-m_{d\bar u}=m_{s\bar c}-m_{s\bar u}=m_{b\bar c}-m_{b\bar u}=m_{t\bar c}-m_{t\bar u},$$
$$m_{d\bar b}-m_{d\bar u}=m_{t\bar b}-m_{t\bar u},$$

where b and t are the fifth and the six quark states, an overbar denotes the antiquark state, and total momenta of all quark–antiquark states from each equation are the same. In other words, for pseudoscalar mesons of the known masses we obtain

$$m_{K^0}-m_{K^+}=m_{D^+}-m_{D^0}=m_{B^0}-m_{B^+},$$
$$m_{D^+}-m_{\pi^+}=m_{F^+}-m_{K^+}.$$

Compare Ref. 6, where the second formula is derived by additional assumptions. The analogous relations hold for vector mesons. According to Refs. 13 and 15, the mass differences (in MeV) are the following (in parentheses for the corresponding vector mesons):

$$m_{K^0}-m_{K^+}=4.01\pm0.13\ \ (6.7\pm1.2),$$
$$m_{D^+}-m_{D^0}=4.7\pm0.3\ \ (2.9\pm2.8),$$
$$m_{B^0}-m_{B^+}=3.4\pm5.8,$$
$$m_{D^+}-m_{\pi^+}=1729.8\pm0.6\ \ (1241\pm3.7),$$
$$m_{F^+}-m_{K^+}=1477.3\pm6\ \ (1217.1\pm16.4).$$

As it was mentioned above, for $M\in(n)+(n-2,2)$, it makes no sense to compare Eqs. (5.6) with experimental data until the final number of flavors is fixed. For example, for $n=4,5,6$, we get, respectively, the following mass formulas:

$$m_{K^0}-m_{K^+}=m_{D^0}-m_{D^+},$$
$$m_{K^0}-m_{K^+}+m_{D^+}-m_{D^0}=m_{B^+}-m_{B^0};$$
$$m_{K^0}-m_{K^+}+m_{D^+}-m_{D^0}+m_{B^0}-m_{B^+}=m_{t\bar u}-m_{t\bar d}.$$

Thus, for $n=6$ and the top mesons $t\bar u$ and $t\bar d$, Eqs. (5.6) and (5.7) predict the mass differences of the opposite signs.

ACKNOWLEDGMENTS

One of the authors, Igor Szczyrba, wishes to express his gratitude to Dr. Hanni Tschinkel and Dr. Gerhard Ecker whose help enabled him to complete this paper during his stay in Austria.

APPENDIX A: OBSERVABLES FROM THE ENVELOPING ALGEBRA

Let $\mathfrak{A}(\mathfrak{G})$ be the enveloping algebra of $\mathfrak{G}$ and let ρ denote also the extension of a representation ρ of $\mathfrak{G}$ onto $\mathfrak{A}(\mathfrak{G})$. As canonical observables, one can consider elements from the set $\{\xi\in\mathfrak{A}(\mathfrak{G})\,|\,\forall(\rho,V):\rho(\xi)\in\rho(\mathfrak{H})'\}$, which coincides with the commutant $\mathfrak{A}(\mathfrak{H})'$ of $\mathfrak{A}(\mathfrak{H})$ in $\mathfrak{A}(\mathfrak{G})$. However, in general, the image $\rho(\mathfrak{A}(\mathfrak{H})')$ is smaller than $\rho(\mathfrak{H})'$. To make it clear, let us analyze relations among the sets $\rho(\mathfrak{H})'$, $\rho(\mathfrak{A}(\mathfrak{G}))$ and $\rho(\mathfrak{A}(\mathfrak{H})')$.

Let $V=\dot\Sigma_i V_i\otimes M_i$ be the decomposition of V into isotypic components, i.e., V_i carries an irreducible representation ρ_i of $\mathfrak{G}$, M_i is a vector space such that $\dim M_i=c(V_i,V)$, and $\rho=\dot\Sigma_i\,\rho_i\otimes\mathrm{id}_{M_i}$. It is clear that

$$\mathrm{End}\,V=\dot\sum_{i,j}(V_i\otimes V_j^*)\otimes(M_i\otimes M_j^*)$$

and hence,

$$\rho(\mathfrak{H})'=(\mathrm{End}\,V)(0)=\dot\sum_i\rho_i(\mathfrak{H})'\otimes\mathrm{End}\,M_i$$
$$+\dot\sum_{i\neq j}(V_i\otimes V_j^*)(0)\otimes(M_i\otimes M_j^*).$$

The irreducibility of ρ_i implies that $\rho_i(\mathfrak{A}(\mathfrak{G}))=\mathrm{End}\,V_i$ and $\rho_i(\mathfrak{A}(\mathfrak{H})')=\rho_i(\mathfrak{H})'$. Thus

$$\rho(\mathfrak{A}(\mathfrak{G}))=\dot\sum_i\mathrm{End}\,V_i\otimes\mathrm{id}_{M_i},$$

whereas

$$\rho(\mathfrak{A}(\mathfrak{H})')=\dot{\sum_i}\rho_i(\mathfrak{H})'\otimes \mathrm{id}_{M_i}.$$

In consequence, we see that the commutant $\rho(\mathfrak{H})'$ coincides with $\rho(\mathfrak{A}(\mathfrak{H})')$ iff $\dim M_1=1$ for any i, and the sets of weights satisfy $\underline{\Lambda}_i\cap\underline{\Lambda}_j=\varnothing$ for $i\neq j$. Now, the octet and singlet representations of su(3) have the common weight $\lambda=0$. Thereby, the mass operator providing a mixing of meson states cannot be described by an element from $\rho(\mathfrak{A}(\mathfrak{H})')$ or even $\rho(\mathfrak{A}(\mathfrak{G}))$.

The additional characterization of the bicommutant $\rho(\mathfrak{H})''$ is given by the following lemma.

Lemma 8: Let (ρ,V) be a completely reducible representation of an Abelian Lie algebra $\mathfrak{H}$. Then, the following associative commutative subalgebras (with unity) in End V are equal: (1) the algebra $\rho(\mathfrak{H})''$ generated by $\rho(\mathfrak{H})$; (2) the image $\rho(\mathfrak{A}(\mathfrak{H}))$ of the enveloping algebra $\mathfrak{A}(\mathfrak{H})$; (3) the algebra generated by the projections $P_\lambda\in\mathrm{End}\ V$, $P_\lambda V=V(\lambda)$, $\lambda\in\underline{\Lambda}$; and (4) $\{A\in\mathrm{End}\ V\,|\,\forall\lambda\in\underline{\Lambda}, AV(\lambda)\subset V(\lambda)$ and $A\,|_{V(\lambda)}=\mathrm{const}\cdot\mathrm{id}_{V(\lambda)}\}$.

Proof: Two first algebras are equal by definition of the extension ρ onto $\mathfrak{A}(\mathfrak{H})$. Two last ones are obviously equal. Eventually, Eq. (2.11) implies that $\rho(\mathfrak{H})''=\{A\in\mathrm{End}\ V\,|\,\forall\lambda\in\underline{\Lambda}, AV(\lambda)\subset V(\lambda)\}'$. But this means that the first algebra coincides with the last one. ■

APPENDIX B: ACTION OF $A(R)$ ON $\mathfrak{A}(\mathfrak{h})'$

For a semisimple $\mathfrak{g}$, the action of Aut $\mathfrak{g}$ on $\mathfrak{g}$ can be extended to the representation of Aut $\mathfrak{g}$ in the enveloping algebra $\mathfrak{A}(\mathfrak{g})$. The derivative of this representation is given by the formula $\mathrm{ad}\,X(\xi)=X\cdot\xi-\xi\cdot X$, $X\in\mathfrak{g}$, $\xi\in\mathfrak{A}(\mathfrak{g})$. Thus, in the case of semisimple $\mathfrak{g}$, Eq. (3.5) implies that $\ker\epsilon$ acts trivially on $\mathfrak{A}(\mathfrak{h})'=(\mathfrak{A}(g))(0)\subset\mathfrak{A}(\mathfrak{g})$ and that $\mathfrak{A}(\mathfrak{h})'$ is $\mathrm{Aut}(\mathfrak{g},\mathfrak{h})$-invariant. It provides the action of the group $A(R)=\mathrm{Aut}(\mathfrak{g},\mathfrak{h})/\ker\epsilon$ [see (3.4)] on $\mathfrak{A}(\mathfrak{h})'$. So we have a representation $A(R)\ni w\rightarrow\Gamma_w\in\mathrm{GL}(\mathfrak{A}(\mathfrak{h})')$. We shall show that for any representation (π,U) of $\mathfrak{g}$ and $w\in W(R)$, the operator Γ_w induces an operator Γ_w^π acting in $\pi(\mathfrak{A}(\mathfrak{h})')$. First, for any $s\in\mathrm{Int}\ \mathfrak{g}$, the representation $\pi\circ s$ is equivalent to π. Indeed, the group Int $\mathfrak{g}$ is generated by operators $e^{\mathrm{ad}\,X}$, $X\in\mathfrak{g}$, and for such elements, the intertwining operator is given by $e^{-\pi(X)}$. Thereby, $\ker(\pi:\mathfrak{A}(\mathfrak{g})\rightarrow\mathrm{End}\ U)$ is Int $\mathfrak{g}$-invariant and hence, $\ker(\pi:\ \mathfrak{A}(\mathfrak{h})'\rightarrow\mathrm{End}\ U)$ is $\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h})/\ker\epsilon=W(R)$-invariant. Thus the definition $\Gamma_w^\pi\pi(\xi):=\pi(\Gamma_w\xi)$, $\xi\in\mathfrak{A}(\mathfrak{h})'$, $w\in W(R)$, is correct.

APPENDIX C: PROOF OF PROPOSITION 1, (ii)⇔(iii)

First, we shall consider the case of simple Lie algebra $\mathfrak{g}$. It follows from Ref. 2 (Chap. VIII, Sec. 7.3) that there exists the subset $\Omega=\{\overline{\omega}_{i_1},\ldots,\overline{\omega}_{i_{N-1}}\}$ in the set of fundamental weights such that (a) elements of the set $\{0\}\cup\Omega$ form the system of representatives for the factor group P/Q; and (b) for every representation (ρ,V) of $\mathfrak{g}$, the set of weights $\underline{\Lambda}$ contains an element from $\{0\}\cup\Omega$.

Condition (a) implies, in particular, that $\overline{\omega}_{i_1},\ldots,\overline{\omega}_{i_{N-1}}\notin Q$. Therefore, if we assume that $\underline{\Lambda}\subset Q$, then it follows from condition (b) that $0\in\underline{\Lambda}$. Thus, (iii) ⇒ (ii). To prove the converse implication, let us recall (see, e.g., Ref. 8. Chap. 7.2.1) that $\lambda_h-\underline{\Lambda}\subset Q$ for any irreducible representation of the highest weight $\lambda_h\in\underline{\Lambda}$. In consequence, if $0\in\underline{\Lambda}$ then we obtain that $\underline{\Lambda}=(\lambda_h-0)-(\lambda_h-\underline{\Lambda})\subset Q-Q=Q$. It completes the proof for simple Lie algebras.

Now, let $\mathfrak{g}=\mathfrak{g}_1+\mathfrak{g}_2+\cdots$ be a semisimple Lie algebra composed of simple Lie algebras $\mathfrak{g}_1$, $\mathfrak{g}_2$, etc. The root system of $\mathfrak{g}$ is of the form $R=j_1(R_1)\cup j_2(R_2)\cup\cdots$, where for any $a=1,2,\ldots$, $R_a\subset\mathfrak{h}_a^*$ denotes the root system of $\mathfrak{g}_a$, and $j_a:\mathfrak{h}_a^*\rightarrow\mathfrak{h}^*$ is the natural injection, i.e., $\mathfrak{h}^*=\dot{\Sigma}_a j_a(\mathfrak{h}_a^*)$. Moreover, the group $Q(R)$ splits into the direct sum of the groups $Q(R_a)$. It is also known that any irreducible representation (ρ,V) of $\mathfrak{g}$ is of form

$$(\rho,V)=(\rho_1 \textcircled{T} \rho_2 \textcircled{T}\cdots,V_1\otimes V_2\otimes\cdots),$$

where (ρ_a,V_a) is an irreducible representation of $\mathfrak{g}_a$. The formula (2.6) implies that $\underline{\Lambda}=j_1(\underline{\Lambda}_1)+j_2(\underline{\Lambda}_2)+\cdots$, where $\underline{\Lambda}_a$ is the set of weights of the representation (ρ_a,V_a). Thus, we obtain that

$$(0\in\underline{\Lambda})\underset{1}{\Leftrightarrow}(0\in\underline{\Lambda}_a,a=1,2,\ldots)\underset{2}{\Leftrightarrow}(\underline{\Lambda}_a\subset Q(R_a),a=1,2,\ldots)$$

$$\underset{3}{\Leftrightarrow}(\underline{\Lambda}\subset Q=Q(R)).$$

In fact, the equivalences 1, 2, and 3 are true since (1) $j_a(\underline{\Lambda}_a)\subset j_a(\mathfrak{h}_a^*)$ and $\mathfrak{h}^*$ is the direct sum of the $j_a(\mathfrak{h}_a^*)$'s; (2) each $\mathfrak{g}_a$ is simple; and (3) $Q(R)$ is the direct sum of the $Q(R_a)$'s.

APPENDIX D: GROUPS $W(R)$, $Q(R)$, AND $P(R)$ FOR CLASSICAL LIE ALGEBRAS

The Weyl groups of the classical Lie algebras are extensions of permutation groups S_n,

$$W(R)\simeq\begin{cases} S_n, & \text{for } A_{n-1},\quad n\geqslant 2,\\ S_n \circledS (Z_2)^n, & \text{for } B_n \text{ and } C_n,\quad n\geqslant 2,\\ S_n\circledS(Z_2)_0^n, & \text{for } D_n,\quad n\geqslant 3,\end{cases}$$

where $Z_2:=\{1,-1\}$, $(Z_2)_0^n$ consists of $\epsilon=(\epsilon_1,\ldots,\epsilon_n)\in(Z_2)^n$ such that $\epsilon_1\cdots\epsilon_n=1$, and $\circledS$ denotes a semisimple product, i.e., $(\sigma,\epsilon)\cdot(\sigma',\epsilon')=(\sigma\circ\sigma,\epsilon\cdot\sigma(\epsilon'))$, $\sigma(\epsilon):=(\epsilon_{\sigma^{-1}(1)}\cdots\epsilon_{\sigma^{-1}(n)})$, $\sigma\in S_n$. The corresponding groups $Q(R)$ and $P(R)$ can be realized as follows:

$$\left.\begin{aligned} Q(R)&=\mathbb{Z}_0^n:=\{x\in\mathbb{Z}^n|x_1\cdots+x_n=0\},\\ P(R)&=(1/n)\{x\in\mathbb{Z}_0^n|x_1\equiv\cdots\equiv x_n\pmod n\},\end{aligned}\right\}\text{ for } A_{n-1},$$

$$\left.\begin{aligned} Q(R)&=\mathbb{Z}^n,\\ P(R)&=\tfrac{1}{2}\{x\in\mathbb{Z}^n|x_1\equiv\cdots\equiv x_n\pmod 2\},\end{aligned}\right\}\text{ for } B_n,$$

$$\left.\begin{aligned} Q(R)&=\{x\in\mathbb{Z}^n|x_1+\cdots+x_n\equiv 0\pmod 2\},\\ P(R)&=\mathbb{Z}^n,\end{aligned}\right\}\text{ for } C_n,$$

$$\left.\begin{aligned} Q(R)&=\{x\in\mathbb{Z}^n|x_1+\cdots+x_n\equiv 0\pmod 2\},\\ P(R)&=\tfrac{1}{2}\{x\in\mathbb{Z}^n|x_1\equiv\cdots\equiv x_n\pmod 2\},\end{aligned}\right\}\text{ for } D_n.$$

The group $W(R)$ acts on $Q(R)$ and $P(R)$ according to the formulas $(\sigma x)_i:=x_{\sigma^{-1}(i)}$ for A_{n-1}, and $((\epsilon,\sigma)x)_i:=\epsilon_i x_{\sigma^{-1}(i)}$ for B_n, C_n, D_n.

Note that in the case of the A_{n-1} type, the $W(R)=S_n$ actions on $Q(R)$ and $P(R)$ coincide with the natural action of S_n in $\mathbb{Z}^n$ only if we describe $\mathfrak{h}^*\simeq\mathbb{C}_0^n$ by means of the gauge condition $x_1+\cdots+x_n=0$. The usually used condition x_n

$=0$ provides more complicated formulas. It is only when we consider the reductive Lie algebra $\mathrm{gl}(n,\mathbb{C})$ that the Weyl group S_n acts automatically on $P(R)=\mathbf{Z}^n$ in the natural way.

APPENDIX E: FINITE COVERINGS OF $W(R)$

Let $(X_\alpha)_{\alpha\in R}$ be a Chevalley system of a complex pair $(\mathfrak{g},\mathfrak{h})$, see, e.g., Ref. 2, Chap. VIII, Sec. 2.4. The elements $e^{\mathrm{ad}\,X_\alpha}e^{\mathrm{ad}\,X_{-\alpha}}e^{\mathrm{ad}}e^{\mathrm{ad}\,X_\alpha}$ generate a finite subgroup $\overline{W}\subset\mathrm{Aut}_0(\mathfrak{g},\mathfrak{h})$ that covers $W(R)$, see Ref. 2, Chap. VIII, exercise 10, in Sec. 5 and the reference therein. Moreover, elements $\exp X_\alpha \exp X_{-\alpha} \exp X_\alpha$ generate[16,17] a finite subgroup $N\subset G_\mathfrak{h}$, where G is a simply connected group such that $L(G)=\mathfrak{g}$. (Note a sign difference in notations used by various authors.) In other words, we have a diagram (rows are not exact)

$$\begin{array}{ccccc} G_\mathfrak{h} & \xrightarrow{\mathrm{Ad}} & \mathrm{Aut}_0(\mathfrak{g},\mathfrak{h}) & \xrightarrow{\epsilon} & W(R) \\ \uparrow & & \uparrow & & \uparrow \mathrm{id.} \\ N & \xrightarrow{\mathrm{Ad}} & \overline{W} & \xrightarrow{\epsilon} & W(R) \end{array}$$

The ratio of orders $|N|/|W(R)|$ is equal to $2^{\dim\mathfrak{h}}$, whereas $|\overline{W}|/|W(R)|=|Q/Q\cap 2P|$ is, in general, smaller. For instance, for $\mathfrak{g}$ of type A_n we have $|N|/|\overline{W}|=2$ if $n\equiv 1(\mathrm{mod}\ 2)$, and $N\simeq\overline{W}$ if $n\equiv 0(\mathrm{mod}\ 2)$. In example 1 of Sect. III B, a group $N\subset\mathrm{SL}(E)$ is generated, e.g., by elements $g_{ij}\in\mathrm{SL}(E)$, $i\neq j$, such that $g_{ij}e_i=-e_j$, $g_{ij}e_j=e_i$, $g_{ij}e_k=e_k$, $i\neq k\neq j$. Using in this example for the z_i's square roots of -1 of higher orders, one can also generate finite $W(R)$ coverings bigger than N.

On the other hand, considering, instead of G, a reductive $\mathrm{Int}\,\mathfrak{g}$ covering group $\widetilde{G}$ $[L(\widetilde{G})=\mathfrak{g}\times\mathfrak{c}]$, we can obtain smaller finite $W(R)$ coverings contained in $\widetilde{G}_\mathfrak{h}$, or even be able to imbed $W(R)$ into $\widetilde{G}_\mathfrak{h}$. In fact, if in example 1 we allow that $\det g=\pm 1$, then elements $g_w\in\mathrm{GL}(E)$, $w\in W(R)\simeq S_n$, such that $g_we_i=e_{w(i)}$, $i\in\overline{1,n}$, generate a subgroup isomorphic with $W(R)$. Clearly, in this case for any weight $\lambda\in\underline{\Lambda}$, the operators $\mathscr{P}_w$, $w\in W(R)$, given by (3.11) provide a usual representation $(\Pi^\lambda,U(\lambda))$ of the isotropy subgroup $W_\lambda\subset W(R)$. Compare Ref. 17, where the representation of a subgroup $N_\lambda\subset N$ which stabilizes $U(\lambda)$ is used.

[1] E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics and Atomic Spectra* (Academic, New York, 1959); L. Michel, *Applications of Group Theory to Quantum Physics, Algebraic Aspects, Lecture Notes in Physics*, Vol. 6 (Springer, Berlin, 1970); J. A. Castilho Alcaras, L. C. Biedenharn, K. T. Hecht, and G. Neely, Ann. Phys. (NY) **60**, 85 (1970); J. D. Louck and L. C. Biedenharn, J. Math. Phys. **11**, 2368 (1970), **14**, 1336 (1973); L. C. Biedenharn, J. D. Louck, E. Chacón, and M. Ciftan, *ibid.* **13**, 1957, 1985 (1972); S. Okubo, *ibid.* **16**, 528 (1975).

[2] N. Bourbaki, *Éléments de Mathématique, Groupes et al algèbres de Lie* (Hermann, Paris, 1975), Chaps. VII and VIII.

[3] I. Szczyrba, J. Math. Phys. **22**, 1289 (1981); Lett. Math. Phys. **5**, 449 (1981); Preprint 127, Sonderforschungbereich 123, Univ. of Heidelberg, (1981).

[4] S. Coleman and S. L. Glashow, Phys. Rev. Lett. **6**, 423 (1961); M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); S. Okubo, Prog. Theor. Phys. **27**, 949 (1962).

[5] G. Cieciura and I. Szczyrba, J. Math. Phys. **28**, 2128 (1987).

[6] M. K. Gaillard, B. W. Lee, and J. L. Rosner, Rev. Mod. Phys. **47**, 277 (1975).

[7] N. Bourbaki, *Éléments de Mathématique, Groupes et algèbres de Lie* (Herman, Paris, 1968), Chaps. I–VI.

[8] M. Goto and F. D. Grosshans, *Semisimple Lie Algebras, Lecture Notes in Pure and Applied Mathematics*, Vol. 38 (Dekker, New York, 1978).

[9] J-P. Serre, *Algèbres de Lie semi-simples complexes* (Benjamin–Cummings, New York, 1965).

[10] H. Weyl, *The Classical Groups, Their Invariants and Representations* (Princeton U. P., Princeton, 1946).

[11] G. Cieciura and I. Szczyrba, "*Physical relations and automorphisms of root systems*," in preparation.

[12] R. Slansky, Phys. Rep. **79**, 1 (1981).

[13] "*Review of particle properties*," Rev. Mod. Phys. **56**, No. 2 (1984).

[14] J-P. Serre, *Linear Representations of Finite Groups* (Springer, New York, 1977).

[15] K. Schubert, (Argus Coll.) Proc. Recontre de Mariond, La Plagne, Jan. 1985.

[16] J. Tits, Théorie des groúpes, "Résumé des cours et travaux," Annuaire du Collège de France, 1980–1981.

[17] R. V. Moody and J. Patera, J. Math. Phys. **25**, 2838 (1984).

Weyl group and tensor operators for hadron-type SU(n) representations

Grzegorz Cieciura
Department of Mathematical Methods in Physics, University of Warsaw, Hoza 74, 00-682 Warsaw, Poland
Igor Szczyrba
Department of Mathematics and Applied Statistics, University of Northern Colorado, Greeley, Colorado 80639

(Received 11 November 1985; accepted for publication 6 May 1987)

Let a particle symmetry be described by a simple Lie algrebra $\mathfrak{g}$ of the type A_{n-1}, i.e., $\mathfrak{g} = \mathrm{sl}(n,\mathbb{C})$ or $\mathfrak{g}$ is a real form of $\mathrm{sl}(n,\mathbb{C})$. For $\mathfrak{g}$ representations describing three-particle or particle–antiparticle states, relationships between two actions, the action of $\mathfrak{g}$ and the action of the corresponding Weyl group S_n, on observables are analyzed. It is shown, in particular, how possible physical relations depend on these two actions. The results enable one to verify quickly if given experimental data can be fitted by means of a $\mathfrak{g}$-symmetry theory.

I. PRELIMINARIES

We continue the analysis of physical relations that can appear in particle symmetries based on reductive Lie algebras. In our previous paper,[1] we showed that from the point of view of physical relations, it is sufficient to consider *complex semisimple* Lie algebras. Let $\mathfrak{g}$ be such an algebra and let $\mathfrak{h}$ be its Cartan subalgebra.

A. Tensor operators

Let (ρ,V) be a finite-dimensional complex representation of $\mathfrak{g}$ describing physical states and let $\mathrm{ad}\,\rho$ be the corresponding $\mathfrak{g}$ action in the space End V. Finally, let $\mathscr{T} \in \mathscr{L}_{\mathfrak{g}}(U,\mathrm{End}\,V)$ denote a linear embedding intertwining a given representation (π,U) of $\mathfrak{g}$ with $(\mathrm{ad}\,\rho,\mathrm{End}\,V)$. Operators from $\mathscr{T}(U) \subset \mathrm{End}\,V$ are called *tensor operators*[2] of the type (π,U). It is usually assumed that each observable has particular $\mathfrak{g}$ transformation properties, i.e., it can be described by a tensor operator of a particular type. See Refs. 2–4.

In Ref. 1 we proved that observables are contained in the zero-weight space $(\mathrm{End}\,V)(0)$ of the representation $(\mathrm{ad}\,\rho,\mathrm{End}\,V)$ and that the Weyl group $W(R)$ (corresponding to $\mathfrak{g}$), acts canonically in any zero-weight space of $(\mathfrak{g},\mathfrak{h})$. This $W(R)$ action Π depends functorially on $\mathfrak{g}$ representations. In particular, every mapping $\mathscr{T}$ restricted to the space $U(0)$ intertwines also the corresponding $W(R)$ actions. For example, the mapping ρ itself defines tensor operators of the adjoint type[5] and $\rho|_{\mathfrak{h}}$ intertwines the natural $W(R)$ actions in $\mathfrak{h}$ and $\rho(\mathfrak{h})$. For an irreducible $\mathfrak{g}$ representation (π,U), the $W(R)$ representation $(\Pi,U(0))$ is, in general, reducible. Thus the decomposition of a given observable $A \in (\mathrm{End}\,V)(0)$ into the components transformed by irreducible $W(R)$ representations refines the decomposition of A into irreducible tensor operators (traditionally used by physicists).

As we showed (Proposition 2 in Ref. 1), many physical relations follow from the $W(R)$ symmetry properties of a given observable A. Since, in general, the same $W(R)$ representation can appear in various subspaces $\mathscr{T}U(0) \subset (\mathrm{End}\,V)(0)$, the same physical relations can be obtained by means of tensor operators of various types.

For example, the results of Ref. 1 (Sec. IV A, Example 2) show that the Coleman–Glashow mass formula in su(3) theory is equivalent to the fact that the mass operator does not contain a component transformed by the alternative S_3 representation of the signature $(1^3) = (1,1,1)$. But for the hadron $\mathrm{sl}(3,\mathbb{C})$ repesentation (ρ,V) of the signature $(2,1,0)$, we have the following decomposition: End $V = \mathbf{1}+\mathbf{8}+\mathbf{8}+\mathbf{10}+\mathbf{10}^*+\mathbf{27}$, where the irreducible $\mathrm{sl}(3,\mathbb{C})$ components are denoted by their dimensions. The zero-weight space of the adjoint representation $\mathbf{8}$ carries the two-dimensional S_3 representation $(2,1)$, i.e., $\mathbf{8}(0) = (2,1)$. It is also clear that $\mathbf{1}(0) = (3)$. We shall see that $\mathbf{10}(0) = \mathbf{10}^*(0) = (1^3)$, whereas $\mathbf{27}(0) = (2,1)+(3)$. Thereby, this mass formula appears iff tensor operators of the decimet type are excluded (and not only if we use tensor operators of the type $\mathbf{1}+\mathbf{8}$).

Thus, in a general situation, *the first problem is to find the decomposition of the zero-weight spaces U(0) into irreducible W(R) components.*

On the other hand, there also exist relations that can be obtained only if one uses tensor operators of a definite type. For example, to have the equidistance rule for masses in the n-flavor theory, one must assume that the mass operator is of the form $c\cdot \mathrm{id}_V + \rho(H)$, $c = \mathrm{const}$, $H \in \mathfrak{h}$, or, just the opposite, to get the mass of a meson equal to the mass of the corresponding antimeson, one has to exclude from the meson mass operator the components $\rho(H)$, $H \in \mathfrak{h}$.

Thus *the second problem is to derive, for irreducible tensor operators from $\mathscr{T}U(0)$, additional relations depending on their tensor type.*

B. Hadron-type representations of $\mathrm{sl}(n,\mathbb{C})$, $n \geqslant 3$

In the present paper we solve the first problem for the irreducible $\mathrm{sl}(n,\mathbb{C})$ representations (π,U) of the following signatures:

$$(p+q,q,\ldots,q,0) \quad p,q \geqslant 0\,,$$

$$(q,\underbrace{1,\ldots,1}_{p},0,\ldots,0) \quad 0 \leqslant p \leqslant n-1, \quad q \geqslant 1\,,$$

$$(\underbrace{2,\ldots,2}_{p},1,\ldots,1,\underbrace{0,\ldots,0}_{q}), \quad 0 \leqslant p+q \leqslant n\,.$$

We denote these representations respectively by $\mathscr{A}^p_q, \mathscr{B}^p_q$, and $\mathscr{C}^p_q$. Moreover, let us put $\mathscr{A}_q := \mathscr{A}^q_q$, $\mathscr{C}_q := \mathscr{C}^q_q$. The capital letters are chosen according to the (general) shape of the corresponding Young schemes: *axe* shape, *boomerang* shape,[6] and *club* shape. In what follows we shall call these respresentations of $\mathrm{sl}(n,\mathbb{C})$ *arms representations*. Obviously, some arms representations coincide:

$$\mathscr{A}_0=\mathscr{B}_1^{n-1}=\mathscr{C}_0=\mathscr{C}_0^n=\mathscr{C}_n^0,$$
$$\mathscr{A}_1=\mathscr{B}_2^{n-2}=\mathscr{C}_1,\quad \mathscr{A}_0^k=\mathscr{B}_k^0,$$
$$\mathscr{A}_1^k=\mathscr{B}_{k+1}^{n-2},\quad \mathscr{C}_k^1=\mathscr{B}_2^{n-k-1}.$$

Arms representations or representations contragredient to them appear, in particular, in the decomposition of the space End V into irreducible $\mathrm{sl}(n,\mathbb{C})$ components if (ρ,V) *is a hadron-type representation* of $\mathrm{sl}(n,\mathbb{C})$ describing three-particle or particle–antiparticle states, i.e., if sgn $V=(3,0,...,0)$ or $(2,1,0,...,0)$ or $(1,1,1,0,...,0)$ or $(2,1,...,1,0)$. In fact, using, e.g, the Young scheme technique[7] we obtain for $n\geqslant 6$

$$\mathrm{End}\,V=\begin{cases}\mathscr{A}_0\dotplus\mathscr{A}_1\dotplus\mathscr{A}_2\dotplus\mathscr{A}_3, & \text{if sgn } V=(3,0,...,0),\quad V=\mathscr{A}_0^3, \quad (1.1)\\ \mathscr{A}_0\dotplus 2\mathscr{A}_1\dotplus\mathscr{A}_2\dotplus\mathscr{B}_3^{n-3}\dotplus(\mathscr{B}_3^{n-3})^*\dotplus\mathscr{C}_2\dotplus\mathscr{D}, & \text{if sgn } V=(2,1,0,...,0),\quad V=\mathscr{B}_2^1, \quad (1.2)\\ \mathscr{C}_0\dotplus\mathscr{C}_1\dotplus\mathscr{C}_2\dotplus\mathscr{C}_3, & \text{if sgn } V=(1,1,1,0,...,0),\quad V=\mathscr{B}_1^2, \quad (1.3)\\ \mathscr{A}_0\dotplus 2\mathscr{A}_1\dotplus\mathscr{A}_2\dotplus\mathscr{B}_3^{n-3}\dotplus(\mathscr{B}_3^{n-3})^*\dotplus\mathscr{C}_2, & \text{if sgn } V=(2,1,...,1,0),\quad V=\mathscr{A}_1, \quad (1.4)\end{cases}$$

where $\mathscr{D}$ of sign $\mathscr{D}=(4,3,2,...,2,1,0)$ denotes the only non-arms-irreducible component. For $n=3$, in the decompositions (1.2)–(1.4) the components $\mathscr{D}$, $\mathscr{C}_3$, $\mathscr{C}_2$, and $\mathscr{C}_1$ do not appear, whereas for $n=4$ (resp. 5), the decomposition (1.3) does not contain $\mathscr{C}_3\dotplus\mathscr{C}_2$ (resp. $\mathscr{C}_3$). Compare Propositions 1 and 5 proved in the sequel. Thus our results concerning arms representations characterize, in particular, the $W(R)\simeq S_n$ action on observables for hadron-type representations of $\mathrm{sl}(n,\mathbb{C})$.

Let us notice that there is no need to analyze separately the S_n-representations $(\mathscr{B}_q^p)^*(0)$. Indeed, the functorial properties of the $W(R)$ action Π imply that for any pair of contragredient $\mathfrak{g}$ representations (π,U) and $(\pi^\wedge,U^*)$, the zero-weight spaces $U(0)$ and $U^*(0)$ carry the contragredient $W(R)$ represenations. But for the group S_n such representations are equivalent (see Ref. 1, Lemma 7). Moreover, since $(\mathscr{A}_q^p)^*=\mathscr{A}_p^q$ and $(\mathscr{C}_q^p)^*=\mathscr{C}_p^q$, it is sufficient to consider the spaces $\mathscr{A}_q^p(0)$ and $\mathscr{C}_q^p(0)$ with $q\geqslant p$.

The main results we obtain about the zero-weight spaces of arms representations are the following. Let $\overline{\mathbb{C}}$ carry the alternative S_n representation of the signature (1^n) and let $S^p(\mathfrak{h})$ [resp. $\wedge^p(\mathfrak{h})$] be the symmetrical (resp. skew-symmetrical) p-fold tensor power of the natural S_n representation in a Cartan algebra $\mathfrak{h}$ of $\mathrm{sl}(n,\mathbb{C})$, i.e., sgn $\mathfrak{h}=(n-1,1)$. Then

$$\mathscr{A}_q^p(0)\simeq\overline{\mathbb{C}}^{\otimes k}\otimes S^{\min(p,q)}(\mathfrak{h}),\quad \text{if } |q-p|=n\cdot k,\quad 0\leqslant k\in\mathbb{Z},$$
$$\mathscr{B}_q^p(0)\simeq\overline{\mathbb{C}}^{\otimes k}\otimes\wedge^p(\mathfrak{h}),\quad \text{if } p+q=n\cdot k,\quad 0\leqslant k\in\mathbb{Z},$$
$$\mathscr{C}_q(0)\simeq\mathbb{C}_0^{n,q},\quad \text{if } 0\leqslant q\leqslant n/2,$$

where for $0\leqslant q\leqslant n/2$, the space $\mathbb{C}_0^{n,q}$ carries irreducible S_n representation of the signature $(n-q,q)$. In all the remaining cases the zero-weight spaces of arms representations vanish. We analyze also properties of (in general reducible) representations $S^p(\mathfrak{h})$ and we show that $\wedge^p(\mathfrak{h})$ is the irreducible S_n representation of the signature $(n-p,1^p)$.

Moreover, for any hadron-type $\mathrm{sl}(n,\mathbb{C})$ representation (ρ,V), we describe explicitly the $\mathrm{sl}(n,\mathbb{C})$-irreducible spaces $\mathscr{T}(U)\subset\mathrm{End}\,V$, compare Ref. 8. Our results show which additional physical relations can be obtained if we fix the S_n-transformation properties of a given observable $A\in(\mathrm{End}\,V)(0)$ and vary its tensor type (π,U).

It turns out that the S_n spectrum in the space $(\mathrm{End}\,V)(0)$ is, in general, bigger than the spectrum of the S_n action in the space $\mathbb{C}^\Delta$, where Δ denote the set of weights of (ρ,V). For example, if $V=\mathscr{B}_2^1$ and $n\geqslant 4$, then in the space $U(0)$ with sgn $U=(4,3,2,...,2,1,0)$, i.e., $\mathscr{T}(U)=\mathscr{D}$, there are contained the S_n representations of the signatures $(n-3,1^3)$ and $(n-3,2,1)$ which do not appear in $\mathbb{C}^\Delta$.

But according to Proposition 2 in Ref. 1, the basic physical relations, which are implied by S_n transformation properties of a given observable A, do not change if we add to A some operators from a S_n-invariant subspace in $(\mathrm{End}\,V)(0)$ which does not appear in C^Δ. Thus, we show which additional (fitting) parameters can be introduced to a considered symmetry theory without any change of the basic physical relations.

Let us notice that as long as *complex representations* are used, our results can be applied for *any simple Lie algebra of the type* A_{n-1}, e.g., for the compact form $\mathrm{su}(n)$ or for $\mathrm{su}(p,q)$, $p+q=n$, $\mathrm{sl}(n,R)$, etc.

C. Notation

In what follows $\mathfrak{g}$ denotes the complex Lie algebra of the type A_{n-1}, $n\geqslant 3$. We shall use the following realization of the pair $(\mathfrak{g},\mathfrak{h})$. Let E be a fixed n-dimensional complex vector space with a given basis $\{e_i\}$, $i\in\overline{1,n}$, and let $\{e^i\}$ denote the dual basis in E^*. The coordinates of a vector $x\in E$ (resp. $\xi\in E^*$) with respect to the chosen basis shall be denoted by $x^1,...,x^n$ (resp. $\xi_1,...,\xi_n$). As $\mathfrak{h}$ we choose all operators in $\mathrm{sl}(E)\simeq\mathrm{sl}(n,\mathbb{C})$ that are diagonal with respect to $\{e_i\}$, i.e.,

$$\mathfrak{h}=\left\{X=\sum_i a_i\, e_i^i\,\middle|\,\sum_i a_i=0\right\},$$

where

$$e_i^j:=e_i\otimes e^j\in E\otimes E^*=\mathrm{End}\,V.$$

Here and in the sequel Σ_i denotes $\Sigma_{1\leqslant i\leqslant n}$. Clearly, $\mathfrak{h}$ is isomorphic to the space $\mathbb{C}_0^n:=\{\lambda=(\lambda_1,...,\lambda_n)\in\mathbb{C}^n|\Sigma_i\lambda_i=0\}$. Thus the spaces $\mathbb{C}$, $\mathbb{C}_0^n$, $\overline{\mathbb{C}}$, and $\overline{\mathbb{C}}_0^n:=\overline{\mathbb{C}}\otimes\mathbb{C}_0^n$ carry, respectively, the irreducible S_n representations of the signatures (n), $(n-1,1)$, (1^n), and $(2,1^{n-2})$.

The standard basis in $\mathbb{C}^n$ shall be denoted by $\epsilon_i:=(0,...,1,...,0)$, $i\in\overline{1,n}$, and its dual basis by $\{\epsilon^i\}$. For any $\lambda\in\mathbb{C}^n$, let us define $\tilde\lambda\in\mathfrak{h}^*$ by the formula

$$\langle\tilde\lambda,X\rangle:=\sum_i\lambda_i a_i,\quad X\in\mathfrak{h}. \quad (1.5)$$

So, $\tilde{\lambda} = \tilde{\mu}$ iff $\lambda - \mu = c \cdot \mathbb{1}$, where $c \in \mathbb{C}$, $\mathbb{1} := (1,\ldots,1) \in \mathbb{C}^n$. The mapping $\lambda \mapsto \tilde{\lambda}$, as well as its restriction to $\mathbb{C}_0^n$, intertwines the representations of S_n. In what follows we shall distinguish λ and $\tilde{\lambda}$ only if it is necessary to avoid misunderstandings.

Moreover, let us introduce the sets

$$\mathscr{L}_p := \{\lambda = (\lambda_1,\ldots,\lambda_n) \in \mathbb{Z}_+^n \mid |\lambda| := \lambda_1 + \cdots + \lambda_n = p\},$$

where $\mathbb{Z}_+ := \{\text{integers} \geqslant 0\}$,

and

$$\pounds_p := \{I \subset \overline{1,n} \mid |I| := (\text{number of elements in } I) = p\}.$$

For any $\lambda \in \mathscr{L}_p$, $\lambda ! := \lambda_1! \cdots \lambda_n!$, and for any $I \in \pounds_p$, let us set $\lambda_I := \Sigma_{i \in I} \epsilon_i \in \mathscr{L}_p \subset \mathbb{C}^n$. Finally, for any finite set X, let $\mathbb{C}^X$ be the set of functions $X \ni x \mapsto a_x \in \mathbb{C}$ [denoted usually by $a = (a_x)$] with the natural structure of $|X|$-dimensional vector space.

II. ARMS REPRESENTATIONS OF sl(n,$\mathbb{C}$)

A. Axe representations $\mathscr{A}_q^p$

Let us set $\Phi_q^p = \Phi_q^p(E) := \mathscr{L}(S^q, S^p) = S^p \otimes (S^q)^*$, where $S^p := S^p(E)$. Let us introduce operators $J_\pm : \Phi_q^p \to \Phi_{q\pm 1}^{p\pm 1}$ (for simplicity of notation we set $\Phi_q^p = 0$ if p or q is negative) by the formulas

$$(J_+A)w = \sum_i e_i \odot A(e^i \lrcorner\, w),$$
$$(J_-A)w = \sum_i e^i \lrcorner\, A(e_i \odot w), \tag{2.1}$$

where $e^i \lrcorner\, w$ denotes the contraction of the tensor $e^i \otimes w$. Clearly, the operators $J_\pm$ do not depend on the choice of a basis in E and commute with the $\mathfrak{g}$ action in $\Sigma_{p,q} \Phi_q^p$. Using the Leibniz rule [$e^i \lrcorner\, w = \partial w / \partial \xi_i$ if elements $w \in S^p(E)$ are treated as the polynomials on E^*] and the Euler formula $\Sigma_i e_i \odot (e^i \lrcorner\, w) = pw$, one can check that

$$[J_-, J_+] = 2H, \quad [H, J_\pm] = \pm J_\pm, \tag{2.2}$$

where $J_\pm$ and H are treated as operators on $\Sigma_{p,q} \Phi_q^p$ and H is given by $HA := \frac{1}{2}(n+p+q)A$, $A \in \Phi_q^p$.

It is also worth mentioning that if Φ_q^p is considered as the space of polynomials on $E^* \dotplus E$ (homogeneous of the degree p in ξ and q in x), then J_+ is the multiplication by $\langle \xi, x \rangle$ whereas

$$J_- = \sum_i \frac{\partial^2}{\partial \xi_i\, \partial x^i}.$$

Let us denote

$$\mathscr{A}_q^p = \mathscr{A}_q^p(E) := \ker(J_- : \Phi_q^p \to \Phi_{q-1}^{p-1}). \tag{2.3}$$

Proposition 1: (a) The space Φ_q^p decomposes into the following direct sum of $\mathfrak{g}$ representations:

$$\Phi_q^p = \sum_r J_+^r\, \mathscr{A}_{q-r}^{p-r}$$
$$\simeq \sum_r \mathscr{A}_{q-r}^{p-r}, \quad \text{where } 0 \leqslant r \leqslant \min(p,q). \tag{2.4}$$

(*b*) *For* $p,q \geqslant 0$, the space $\mathscr{A}_q^p$ carries the irreducible $\mathfrak{g}$ representation of the signature $(p+q, q, \ldots, q, 0)$;

$$\dim \mathscr{A}_q^p = \binom{n+p-1}{p}\binom{n+q-1}{q} - \binom{n+p-2}{p-1}\binom{n+q-2}{q-1}.$$

Proof: (a) The commutation rules (2.2) imply that for $k \geqslant 1$,

$$[J_-, J_+^r] = 2 \sum_{k=1}^r J_+^{k-1} H J_+^{r-k}.$$

Hence, using the definition of H, we obtain that

$$\left[J_-, \frac{1}{r!} J_+^r\right] A = \frac{2}{r!} \sum_{k=1} J_+^{r-1} \frac{1}{2}(n+p+q+2r-2k) A$$
$$= (n+p+q+r-1) \frac{1}{(r-1)!} J_+^{r-1} A, \quad A \in \Phi_q^p. \tag{2.5}$$

The formula (2.4) we want to prove means that every $A \in \Phi_q^p$ decomposes uniquely as follows:

$$A = \sum_{r \geqslant 0} \frac{1}{r!} J_+^r A_r, \quad A_r \in \mathscr{A}_{q-r}^{p-r}. \tag{2.6}$$

For $p=0$ or $q=0$, it is trivially true, thereby we can assume that $p,q \geqslant 1$. Using (2.3) and (2.5), we see that

$$J_-\left(A - \sum_{r \geqslant 1} \frac{1}{r!} J_+^r A_r\right) = J_- A - \sum_{r \geqslant 1} (n+p+q-r-1) \frac{1}{(r-1)!} J_+^{r-1} A_r \tag{2.7}$$

whenever the operators $A_r \in \mathscr{A}_{q-r}^{p-r}$.

Now, if we assume that the decomposition (2.4) takes place for the space Φ_{q-1}^{p-1}, then all the operators $A_r \in \mathscr{A}_{q-r}^{p-r}$ in (2.7) can be chosen in such a way that the right-hand side of (2.7) is equal to zero. It means that the operator $A_0 := A - \Sigma_{r \geqslant 1} (1/r!) J_+^r A_r$ belongs to $\mathscr{A}_q^p$, i.e., we obtain the decomposition (2.6); so the formula (2.4) is proved by induction with respect to $p,q \geqslant 0$.

(b) Obviously, the weight vector $A_h := e_1 \odot \cdots \odot e_1 \otimes e^n \odot \cdots \odot e^n \in S^p \otimes (S^q)^*$, corresponding to the highest weight $\lambda_h = p\epsilon_1 - q\epsilon_n$ belongs to the space $\mathscr{A}_q^p$. It can be checked (Ref. 9, Sec. 45) that A_h is the only primitive vector in $\mathscr{A}_q^p$. It proves the first part of (b), see, e.g., Ref. 10, Chap. VIII, Sec. 6. The formula for dimension holds since $\Phi_q^p \simeq \mathscr{A}_q^p \dotplus \mathscr{A}_{q-1}^{p-1}$. ■

The canonical isomorphism describing the natural S_n action in the zero-weight spaces of axe representations is given by the following.

Proposition 2: The zero-weight space $\mathscr{A}_q^p(0)$ is nontrivial iff $k := |q-p|/n \in \mathbb{Z}_+$ and then

$$\mathscr{A}_q^p(0) \simeq \overline{\mathbb{C}}^{\otimes k} \otimes S^{\min(p,q)}(\mathbb{C}_0^n). \tag{2.8}$$

Proof: For $\lambda, \mu \in \mathbb{Z}_+^n$, let us set

$$e_\lambda := \underbrace{e_1 \odot \cdots \odot e_1}_{\lambda_1} \odot \cdots \odot \underbrace{e_n \odot \cdots \odot e_n}_{\lambda_n},$$

$$e^\mu := \underbrace{e^1 \odot \cdots \odot e^1}_{\mu_1} \odot \cdots \odot \underbrace{e^n \odot \cdots \odot e^n}_{\mu_n}$$

If $\lambda\in\mathscr{L}_p$ (resp. $\mu\in\mathscr{L}_q$), then e_λ (resp. e^μ) spans the one-dimensional weight space $S^p(\tilde\lambda)=(S^p(E))(\tilde\lambda)$ [resp. $(S^q)^*(-\tilde\mu)$]. Thereby, $\tilde{\mathscr{L}}_p:=\{\tilde\lambda\,|\lambda\in\mathscr{L}_p\}$ is the set of weights of the $\mathfrak{g}$ representation S^p. Hence, the zero-weight space of $\Phi^p_q=S^p\otimes(S^q)^*$ is given by

$$\Phi^p_q(0)=\sum_{\tilde\lambda}S^p(\tilde\lambda)\otimes(S^q)^*(-\tilde\lambda),$$

where nontrivial terms appears only for $\tilde\lambda\in\tilde{\mathscr{L}}_p\cap\tilde{\mathscr{L}}_q$. (See Ref. 1, Sec. II.) But for elements from $\mathbb{Z}^n_+$, $\tilde\lambda=\tilde\mu$ holds iff $\mu=\lambda+k\cdot\mathbb{1}$, where $k=(|\mu|-|\lambda|)/n\in\mathbb{Z}$. Thus the intersection $\tilde{\mathscr{L}}_p\cap\tilde{\mathscr{L}}_q$ is either empty [if $p\not\equiv q \pmod n$] or it coincides with $\tilde{\mathscr{L}}_{\min(p,q)}$ (otherwise).

As was mentioned in Sec. I B, it is sufficient to consider the case where $p\leqslant q=p+nk$, $k\in\mathbb{Z}_+$. Then, the space $\Phi^p_q(0)$ has the basis consisting of the elements $\varphi_\lambda:=e_\lambda\otimes e^{\lambda+k\cdot\mathbb{1}}$, $\lambda\in\mathscr{L}_p$. We shall show that S_n acts in $\Phi^p_q(0)$ according to the formula $\sigma\cdot\varphi_\lambda=(\operatorname{sgn}\sigma)^k\varphi_{\sigma(\lambda)}$, $\sigma\in S_n$, i.e.,

$$\Phi^p_q(0)\simeq\bar{\mathbb{C}}^{\otimes k}\otimes S^p(\mathbb{C}^n). \tag{2.9}$$

Indeed, let $g\in G=\mathrm{SL}(E)$ correspond to a given $\sigma\in S_n$, i.e., $ge_i=\gamma_ie_{\sigma(i)}$, $\gamma_i\in\mathbb{C}$, and, in consequence, $g^\wedge e^i=\gamma_i^{-1}e^{\sigma(i)}$. Since $\det g=1$, the coefficients γ_i must satisfy $\gamma_1\cdot\,\cdots\,\cdot\gamma_n=\operatorname{sgn}\sigma$. Thus we obtain

$$\sigma\cdot\varphi_\lambda=g(e_\lambda\otimes e^{\lambda+k\cdot\mathbb{1}})=\gamma^\lambda\cdot e_{\sigma(\lambda)}\otimes\gamma^{-(\lambda+k\cdot\mathbb{1})}e^{\sigma(\lambda+k\cdot\mathbb{1})}$$
$$=\gamma^{-k\cdot\mathbb{1}}\cdot\varphi_{\sigma(\lambda)}=(\operatorname{sgn}\sigma)^k\varphi_{\sigma(\lambda)},$$

where $\gamma^\lambda:=\gamma_1^{\lambda_1}\cdots\gamma_n^{\lambda_n}$. (Compare Example 1 in Ref. 1.)

On the other hand, the decomposition (2.4) implies that $\Phi^p_q\simeq\mathscr{A}^p_q+\Phi^{p-1}_{q-1}$ and hence $\Phi^p_q(0)\simeq\mathscr{A}^p_q(0)+\Phi^{p-1}_{q-1}(0)$. Therefore, the assertion (2.8) follows immediately from (2.9) and the formula $S^p(\mathbb{C}^n)\simeq S^p(\mathbb{C}^n_0)+S^{p-1}(\mathbb{C}^n)$. See the formula (A2) in Appendix A 1, where more detailed information is given about the S_n representations $S^p(\mathbb{C}^n)\simeq\mathbb{C}^{\mathscr{L}_p}=\{\mathscr{L}_p\ni\lambda\mapsto a_\lambda\in\mathbb{C}\}$ and $S^p(\mathbb{C}^n_0)$.

For $p=q$, it is possible to give more precise description of the space $\Phi^p_q=\operatorname{End}S^p$. Namely, in this case let us denote the S_n isomorphism (2.9) by

$$\iota: S^p(\mathbb{C}^n)\xrightarrow{\sim}(\operatorname{End}S^p)(0),$$
$$\iota(a)e_\lambda:=a_\lambda e_\lambda,\quad a=(a_\lambda)\in S^p(\mathbb{C}^n). \tag{2.10}$$

The $\mathfrak{g}$-intertwining operators $J_\pm:\operatorname{End}S^p\to\operatorname{End}S^{p\pm1}$, given by (2.1), induce, by the restriction to the zero-weight spaces, the S_n-intertwining operators $J_\pm: S^p(\mathbb{C}^n)\to S^{p\pm1}(\mathbb{C}^n)$ such that $J_\pm\circ\iota=\iota\circ J_\pm$. It can be verified by the straightforward calculation that these operators are given by the formulas

$$(J_+a)_\lambda:=\sum_i\lambda_ia_{\lambda-\epsilon_i},$$
$$(J_-a)_\lambda:=\sum_i(\lambda_i+1)a_{\lambda+\epsilon_i},\quad a\in S^p(\mathbb{C}^n),\quad \lambda\in\mathscr{L}_{p\pm1}, \tag{2.11}$$

and that $\ker(J_-:S^p(\mathbb{C}^n)\to S^{p-1}(\mathbb{C}^n))$ coincides with $S^p(\mathbb{C}^n_0)$, see Appendix A 1 for details. Thus Proposition 1 implies the following.

Corollary 1: Any $A\in\operatorname{End}S^p$ has the following unique decomposition:

$$A=\sum_{0\leqslant q\leqslant p}\frac{1}{(p-q)!}J_+^{p-q}A_q,\quad A_q\in\mathscr{A}_q, \tag{2.12}$$

and any $a\in S^p(\mathbb{C}^n)$ has the analogous unique decomposition

$$a=\sum_{0\leqslant q\leqslant p}\frac{1}{(p-q)!}J_+^{p-q},\quad a_q\in S^p(\mathbb{C}^n_0). \tag{2.13}$$

Moreover, $a\in(\operatorname{End}S^p)(0)$ iff $A=\iota(a)$, $a\in S^p(\mathbb{C}^n)$, and then $A_q=\iota(a_q)$. ■

Remark 1: The commutation relations (2.2) take place on the space $\Sigma_{p>0}S^p(\mathbb{C}^n)$ if we define H on $S^p(\mathbb{C}^n)$ as the multiplication by $(n/2+p)$. So, the decomposition (2.13) can be also obtained directly [in the same way as the assertion (2.4)]. ■

The definition (2.11) implies the following (useful in applications) formulas for powers of the operators $J_\pm$:

$$\left(\frac{1}{r!}J^r_+a\right)_\lambda=\sum_{\mu\in\mathscr{L}_q}\binom{\lambda}{\mu}a_\mu,\quad\left(\frac{1}{r!}J^r_-a\right)_\lambda=\sum_{\mu\in\mathscr{L}_q}\binom{\mu}{\lambda}a_\mu, \tag{2.14}$$

where $a\in S^q(\mathbb{C}^n)$, $\lambda\in\mathscr{L}_{q\pm r}$,

$$\binom{\lambda}{\mu}:=\binom{\lambda_1}{\mu_1}\cdot\,\cdots\,\cdot\binom{\lambda_n}{\mu_n}\quad\text{and}\quad\binom{\lambda_i}{\mu_i}$$

denotes the binomial coefficient. In fact, to get the inductive implication $(r-1)\Rightarrow(r)$, it is sufficient to check that for $|\lambda|-|\mu|=r$, the equality $\lambda_i\cdot\binom{\lambda_i-1}{\mu_i}=(\lambda_i-\mu_i)\binom{\lambda_i}{\mu_i}$ implies that

$$\frac{1}{r}\sum_i\lambda_i\binom{\lambda-\epsilon_i}{\mu}=\binom{\lambda}{\mu}.$$

Similarly,

$$\frac{1}{r}\sum_i(\lambda_i+1)\binom{\mu}{\lambda+\epsilon_i}=\binom{\mu}{\lambda}\quad\text{if } |\mu|-|\lambda|=r.$$

B. Boomerang representations $\mathscr{B}^p_q$

Let us treat the elements of the space $X^p_q=X^p_q(E):=\Lambda^p\otimes S^q$, where $\Lambda^p:=\Lambda^p(E)$, $S^q:=S^q(E)$, as differential p-forms on E^* with coefficients being homogeneous polynomials of the degree q. Let us introduce the following $\mathfrak{g}$-intertwining operators: $J_\pm: X^p_q\to X^{p\pm1}_{q\mp1}$,

$$J_+\omega=d\omega,\quad J_-\omega=R\lrcorner\,\omega, \tag{2.15}$$

where d denotes exterior derivative and $R=\Sigma_i\xi_i(\partial/\partial\xi_i)$ is the radial vector field on E^*. The intertwiners $J_\pm$, treated as operators on $\Sigma_{p,q}X^p_q$, satisfy

$$J^2_\pm=0,\quad J_+J_-+J_-J_+=(p+q)\cdot\mathrm{id}, \tag{2.16}$$

The last formula follows from the basic properties of the Lie derivative. Moreover, let us set

$$\mathscr{B}^p_q=\mathscr{B}^p_q(E):=\ker(J_-: X^p_q\to X^{p-1}_{q+1}),$$
$$\tilde{\mathscr{B}}^p_q=\tilde{\mathscr{B}}^p_q(E):=\ker(J_+: X^p_q\to X^{p+1}_{q-1}).$$

Proposition 3: (a) The space X^p_q decomposes into the following direct sum of $\mathfrak{g}$ representations:

$$\mathrm{X}^p_q = \mathscr{B}^p_q \dotplus \tilde{\mathscr{B}}^p_q .$$

Moreover, the operators $(1/\sqrt{p+q})J_-$ and $(1/\sqrt{p+q})J_+$ provide mutually inverse isomorphisms between $\tilde{\mathscr{B}}^p_q$ and $\mathscr{B}^{p-1}_{q+1}$.

(b) The space $\mathscr{B}^p_q$ carries the irreducible $\mathfrak{g}$ representation of the signature $(q,\underbrace{1,...,1}_{p},0,...,0)$ if $0 \leqslant p \leqslant n-1$, $q \geqslant 1$, and $\mathscr{B}^p_q = 0$ otherwise;

$$\dim \mathscr{B}^p_q = \frac{n}{n+p}\binom{n-1}{p}\binom{n+q+1}{q-1}.$$

Proof: Part (a) follows immediately from the relation (2.16).

(b) The equality $\ker J_- = \operatorname{im} J_-$ implies that $\mathscr{B}^p_q = J_-(\mathrm{X}^{p+1}_{q-1})$. But on the space X^{p+1}_{q-1}, the operator J_- coincides with the operator Sym_q being the symmetrizer in the q last indices. Since, by the definition $\mathrm{X}^{p+1}_{q-1} = \mathrm{Sym}_{q-1}(\wedge^{p+1} \otimes E^{\otimes(q-1)})$, we obtain that

$$\mathscr{B}^p_q = \mathrm{Sym}_q(\wedge^{p+1} \otimes E^{\otimes(q-1)})$$
$$= \mathrm{Sym}_q \circ \mathrm{Alt}_{p+1}(E^{\otimes(p+q)}),$$

where Alt_{p+1} is the skew symmetrizer in the first $p+1$ indices. Thus the space $\mathscr{B}^p_q$ is the image of an appropriate Young symmetrizer (see Ref. 9 or 11) Finally, $\dim \mathscr{B}^p_q = \Sigma_{0<r\leqslant p}(-)^r \dim \mathrm{X}^{p-r}_{q+r}$. ■

Remark 2: It can be checked that the p-forms

$$\omega_h := \xi_1^{q-1}\left(\sum_i \xi_i \frac{\partial}{\partial \xi_i}\right) \lrcorner\, d\xi_1 \wedge \cdots \wedge d\xi_{p+1}$$

and

$$\tilde{\omega}_h := \xi_1^q \, d\xi_1 \wedge \cdots \wedge d\xi_p$$

are primitive vectors in $\mathscr{B}^p_q$ and $\tilde{\mathscr{B}}^p_q$, respectively. ■

The canonical isomorphism describing the natural S_n action in the zero-weight spaces of boomerang representations is given by the following.

Proposition 4: The zero-weight space $\mathscr{B}^p_q(0)$ is nontrivial iff $k = (p+q)/n \in \mathbb{Z}_+$ and then

$$\mathscr{B}^p_q(0) \simeq \overline{\mathbb{C}}^{\otimes k} \otimes \wedge^p(\mathbb{C}^n_0). \tag{2.17}$$

Proof: For $\mu \in \mathscr{L}_q$ and $I = \{i_1,...,i_p\} \in \pounds_p$, the element

$$\xi^\mu \cdot d\xi_{i_1} \wedge \cdots \wedge d\xi_{i_p} \in \mathrm{X}^p_q$$

belongs to the weight space $\mathrm{X}^p_q(\tilde{\lambda}_I + \tilde{\mu})$. Now, since $\tilde{\lambda}_I + \tilde{\mu} = 0$ iff $\lambda_I + \mu = k \cdot \mathbb{1}$, where $k = (p+q)/n \in \mathbb{Z}_+$, we see that $\mathrm{X}^p_q(0) \neq 0$ iff $k = (p+q)/n \in \mathbb{Z}_+$. In the last case, the space $\mathrm{X}^p_q(0)$ has the basis consisting of the vectors

$$\varphi_{i_1 \cdots i_p} := (\xi_1 \cdot \cdots \cdot \xi_n)^k \frac{d\xi_{i_1}}{\xi_{i_1}} \wedge \cdots \wedge \frac{d\xi_{i_p}}{\xi_{i_p}},$$

where $i_1 < \cdots < i_p$. It can be seen (cf. the proof of Proposition 2) that, for $\sigma \in S_n$,

$$\sigma \varphi_{i_1 \cdots i_p} = (\operatorname{sgn} \sigma)^k \varphi_{\sigma(i_1) \cdots \sigma(i_p)},$$
$$\text{i.e., } \mathrm{X}^p_q(0) \simeq \overline{\mathbb{C}}^{\otimes k} \otimes \wedge^p(\mathbb{C}^n). \tag{2.18}$$

Now, let $\eta \in \overline{\mathbb{C}}^{\otimes k}$ be a fixed nonzero element. The natural mapping

$$\mathrm{X}^p_q(0) \ni \omega = \sum_{i_1 < \cdots < i_p} a_{i_1 \cdots i_p} \varphi_{i_1 \cdots i_p}$$
$$\mapsto \Omega := \eta \otimes \sum_{i_1 < \cdots < i_p} a_{i_1 \cdots i_p} \epsilon_{i_1} \wedge \cdots \wedge \epsilon_{i_p}$$
$$\in \overline{\mathbb{C}}^{\otimes k} \otimes \wedge^p(\mathbb{C}^n))$$

is a S_n intertwiner and transforms the condition $J_- \omega = 0$ into $(\epsilon^1 + \cdots + \epsilon^n) \lrcorner\, \Omega = 0$. But the last condition means that $\Omega \in \wedge^p(\mathbb{C}^n_0)$. (See Appendix A 2, where a detailed description of $\wedge^p(\mathbb{C}^n_0)$ is given.) ■

C. Club representations $\mathscr{C}^p_q$

Let us set $\Psi^p_q = \Psi^p_q(E) := \mathscr{L}(\wedge^q, \wedge^p) = \wedge^p \otimes (\wedge^q)^*$, $\wedge^p := \wedge^p(E)$. Similarly as in Sec. II A, we introduce on the space $\Sigma_{p,q} \Psi^p_q$ the $\mathfrak{g}$-intertwining operators $J_\pm : \Psi^p_q \to \Psi^{p\pm1}_{q\pm1}$,

$$(J_+ A)w := \sum_i e_i \wedge A(e^i \lrcorner\, w),$$
$$(J_- A)w := \sum_i e^i \lrcorner\, A(e_i \wedge w). \tag{2.19}$$

Using the Leibniz rule, for the interior product $\xi \lrcorner\, (\cdot)$ in the Grassmann algebra $\wedge(E)$, and the formula $\Sigma_i e_i \wedge (e^i \lrcorner\, w) = pw$, $w \in \wedge^p(E)$, we obtain the commutation relations of the Lie algebra $\mathfrak{a} := \mathrm{sl}(2,\mathbb{C})$:

$$[J_+, J_-] = 2H, \quad [H, J_\pm] = \pm J_\pm, \tag{2.20}$$

where H is the operator on $\Sigma_{p,q} \Psi^p_q$ given by $HA = \frac{1}{2}(p+q-n)A$, $A \in \Psi^p_q$. Let us set

$$\mathscr{C}^p_q = \mathscr{C}^p_q(E) := \ker(J_- : \Psi^p_q \to \Psi^{p-1}_{q-1}). \tag{2.21}$$

Proposition 5: (a) The space Ψ^p_q decomposes into the following direct sum of $\mathfrak{g}$-representations:

$$\Psi^p_q = \sum_r J^r_+ \mathscr{C}^{p-r}_{q-r} \simeq \sum_r \mathscr{C}^{p-r}_{q-r},$$

where $\max(0, p+q-n) \leqslant r \leqslant \min(p,q)$. Moreover, the space $\mathscr{C}^p_q$ is equal to zero unless

$$p \geqslant 0, \quad q \geqslant 0, \quad p+q \leqslant n, \tag{2.22}$$

(b) If p,q satisfy (2.22) then $\mathscr{C}^p_q$ carries the irreducible $\mathfrak{g}$ representation of the signature $(\underbrace{2,...,2}_{p},1,...,1,\underbrace{0,...,0}_{q})$;

$$\dim \mathscr{C}^p_q = \binom{n}{p}\binom{n}{q} - \binom{n}{p-1}\binom{n}{q-1}.$$

Proof: (a) We shall use some known properties of the algebra $\mathfrak{a} = \mathrm{sl}(2,\mathbb{C})$. Let $D(l)$, $l \in \frac{1}{2}\mathbb{Z}_+$ denote the $(2l+1)$-dimensional irreducible $\mathfrak{a}$ representation given by the matrices

$$[J_-] = \begin{bmatrix} 0 & 2l & & & \\ & 0 & 2l-1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix},$$

$$[H] = \begin{bmatrix} -l & & & & \\ & -l+1 & & & \\ & & \ddots & & \\ & & & l-1 & \\ & & & & l \end{bmatrix}, \tag{2.23}$$

$$[J_+] = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 2 & \ddots & & \\ & & & 0 & \\ & & & 2l & 0 \end{bmatrix}.$$

Moreover, let operators $J_\pm, H \in \text{End}\, V$, satisfying (2.20), provide a finite-dimensional $\mathfrak{g}$ representation in the complex vector space V. As we know, V is a direct sum of irreducible components isomorphic to some representations $D(l)$. Let us denote the weight spaces of V by $V^j := \{v \in V | Hv = jv\}$ and let us put $V_0 := \ker J_-$. The space $V_0^j := V_0 \cap V^j$ consists of the primitive vectors of the lowest weight j (with respect to the generated irreducible subspaces). Obviously, $J_\pm V^j$ is contained in $V^{j\pm 1}$, and V^j (resp. V_0^j) vanishes if $j \notin \frac{1}{2}\mathbb{Z}$ [resp. $j \notin (-\frac{1}{2})\mathbb{Z}_+$]. Moreover, for $l \in \frac{1}{2}\mathbb{Z}_+$ and $r > 2\ell$, we get $J_+^r V_0^{-l} = 0$ and hence the isotypic component in V of the type $D(l)$ is given by

$$\mathfrak{A} V_0^{-l} = \sum_{0 \leqslant r \leqslant 2l} J_+^r V_0^{-l} \simeq D(l) \otimes V_0^{-l},$$

where $\mathfrak{A}$ denotes the universal enveloping algebra of $\mathfrak{a}$. Thus we obtain

$$V = \sum_{l \in (1/2)\mathbb{Z}_+}^{\cdot} \mathfrak{A} V_0^{-l} = \sum_{l,r}^{\cdot} J_+^r V_0^{-l}, \quad l \in \tfrac{1}{2}\mathbb{Z}_+, \quad 0 \leqslant r \leqslant 2l,$$

and, in consequence,

$$V^j = \sum_{r \geqslant \max(0,2j)}^{\cdot} J_+^r V_0^{j-r}, \quad j \in \tfrac{1}{2}\mathbb{Z}.$$

Thereby, for $k \in \mathbb{Z}_+, j \in \frac{1}{2}\mathbb{Z}$, the following decomposition holds:

$$V^j \cap \ker J_+^k = \sum_r^{\cdot} J_+^r V_0^{j-r},$$

where $\max(0,2j) \leqslant r \leqslant k + 2j$. The last formula implies, in particular, that for $0 \leqslant k \leqslant -2j$, the operators J_+^k are injective on the space V^j. See Ref. 10, Chap. VIII, Sec. 1. Now, assertion (a) follows immediately from the facts given above if we set $V := \Sigma_r \Psi_{q-r}^{p-r}$ and, in consequence, for $j = \frac{1}{2}(p + q - n) - r$, we get $V^j = \Psi_{q-r}^{p-r}$, $V_0^j = \mathscr{C}_{q-r}^{p-r}$. Note that, if p,q satisfy (2.22) then J_+^r is injective on $\mathscr{C}_q^p$ for $0 \leqslant r \leqslant n - p - q$, whereas $J_+^r \mathscr{C}_q^p = 0$ for $r > n - p - q$.

(b) This part can be proved in a way similar to that of Proposition 3(b). Let us notice that for p,q satisfying (2.22), $(e_1 \wedge \cdots \wedge e_p) \otimes (e^{n-q+1} \wedge \cdots \wedge e^n)$ is the highest weight vector in $\mathscr{C}_q^p$, and that $\Psi_q^p \simeq \mathscr{C}_q^p + \Psi_{q-1}^{p-1}$. ■

Remark 3: The relations (2.2) in Sec. II A define also a representation of $\mathfrak{g}$. In that case, however, the operator H has a strictly positive spectrum. So, this $\mathfrak{g}$ representation does not have any finite-dimensional subrepresentation. ■

Now, let $\mathbb{C}^{n,p} := \mathbb{C}^{\pounds_p}$ denote the space of complex functions $\pounds_p \ni I \mapsto a_I \in \mathbb{C}$ with the natural S_n action: $(\sigma a)_I := a_{\sigma^{-1}(I)}$, $\sigma \in S_n$. Let us define the operators $J_\pm : \mathbb{C}^{n,p} \to \mathbb{C}^{n,p\pm 1}$,

$$(J_+ a)_I = \sum_{\pounds_p \ni J \subset I} a_J, \quad (J_- a)_I = \sum_{\pounds_p \ni J \supset I} a_J, \qquad (2.24)$$

and let $\mathbb{C}_0^{n,p} := \ker(J_- : \mathbb{C}^{n,p} \to \mathbb{C}^{n,p-1})$, for details see Appendix A 3. The canonical isomorphism describing the natural S_n action in the zero-weight spaces of the club representations is given by the following proposition.

Proposition 6: The zero-weight space $\mathscr{C}_q^p(0)$ is nontrivial iff $0 \leqslant p = q \leqslant n/2$ or $p = 0$, $q = n$ or $p = n$, $q = 0$. Moreover, for $0 \leqslant p \leqslant n/2$,

$$\mathscr{C}_p(0) \simeq \mathbb{C}_0^{n,p}. \qquad (2.25)$$

Proof: For $I = \{i_1, \ldots, i_p\}$ where $i_1 < \cdots < i_p$, let us set $e_I := e_{i_1} \wedge \cdots \wedge e_{i_p}$. Clearly, if $I \in \pounds_p$, $J \in \pounds_q$ then $e_I \otimes e^J \in \Psi_q^p$ is the $\mathfrak{g}$-weight vector of the weight $\tilde{\lambda}$ corresponding to $\lambda := \lambda_I - \lambda_J$. As we know $\tilde{\lambda} = 0$ iff $\lambda_I - \lambda_J = k \cdot \mathbf{1}$, where $k = (p - q)/n \in \mathbb{Z}$. The last equation [together with (2.22)] is satisfied only in the cases mentioned above. To prove the isomorphism (2.25), let us consider the space $\Psi_p^p = \text{End} \wedge^p$. Its zero-weight space $\Psi_p^p(0)$ is spanned by the elements $e_I \otimes e^I$, $I \in \pounds_p$, which are permuted by the Weyl group S_n. It shows that mapping $\iota: \mathbb{C}^{n,p} \to \Psi_p^p(0)$, $\iota(a) e_I := a_I e_I$, $a \in \mathbb{C}^{n,p}$ provides the S_n isomorphism

$$\Psi_p^p(0) = (\text{End} \wedge^p)(0) \simeq \mathbb{C}^{n,p}. \qquad (2.26)$$

Now, note that the $\mathfrak{g}$-intertwining operators $J_\pm : \text{End} \wedge^p \to \text{End} \wedge^{p\pm 1}$ given by (2.19), induce, by restricting to zero-weight spaces, the operators $J_\pm : \mathbb{C}^{n,p} \to \mathbb{C}^{n,p\pm 1}$ satisfying $J_\pm \circ \iota = \iota \circ J_\pm$. It can be easily checked that these operators coincide with the operators defined by (2.24). So, the formula (2.26) implies (2.25). ■

Corollary 2: Any $A \in \text{End} \wedge^p$ has the following unique decomposition:

$$A = \sum_{0 \leqslant q \leqslant \min(p, n-p)} \frac{1}{(p-q)!} J_+^{p-q} A_q, \quad A_q \in \mathscr{C}_q, \qquad (2.27)$$

and any $a \in \wedge^p(\mathbb{C}^n)$ has the analogous unique decomposition

$$a = \sum_{0 \leqslant q \leqslant \min(p, n-p)} \frac{1}{(p-q)!} J_+^{p-q} a_q, \quad a_q \in \mathbb{C}_0^{n,p}. \qquad (2.28)$$

Moreover, $A \in (\text{End} \wedge^p)(0)$ iff $A = \iota(a)$, $a \in \wedge^p(\mathbb{C}^n)$ and then $A_q = \iota(a_q)$. ■

One can prove by induction the following formulas [cf. (2.14)]:

$$\left(\frac{1}{r!} J_+^r a\right)_I = \sum_{\pounds_q \ni J \subset I} a_J, \quad \left(\frac{1}{r!} J_-^r a\right)_I = \sum_{\pounds_q \ni J \supset I} a_J, \qquad (2.29)$$

where $a \in \mathbb{C}^{n,q}$, $I \in \pounds_{q \pm r}$.

III. OBSERVABLES FOR HADRON-TYPE REPRESENTATIONS

Let $\underline{\Lambda}$ denote the set of weights of a $\mathfrak{g}$-representation (ρ, V). Let $V(\lambda)$, $\lambda \in \underline{\Lambda}$, be a weight space. In Ref. 1 we proved that the mapping

$$F: (\text{End}\, V)(0) = \sum_{\lambda \in \underline{\Lambda}}^{\cdot} \text{End}\, V(\lambda) \ni A = \sum_{\lambda \in \underline{\Lambda}}^{\cdot} A_\lambda \mapsto F_A \in \mathbb{C}^{\underline{\Lambda}},$$
$$F_A(\lambda) := \text{tr}\, A_\lambda, \qquad (3.1)$$

intertwines the corresponding actions of the Weyl group. In consequence, we showed how to derive possible physical relations that depend on $W(R)$-symmetry properties of a given observable $A \in (\text{End}\, V)(0)$. However, it is also important to know additional relations that are implied by the tensor operator type of A (pure $\mathfrak{g}$-transformation properties).

TABLE I. Representation End S^3.

q	U_q	$U_q(0)$	$F_A(3\epsilon_i)$	$F_A(2\epsilon_i+\epsilon_j)$, $i\neq j$	$F_A(\epsilon_i+\epsilon_j+\epsilon_k)$, $ijk\neq$
0	$\mathscr{A}_0$	$\{A=\iota(a)=a\cdot \mathrm{id}_V\}$	a	a	a
1	$\mathscr{A}_1$	$\{A=\iota(a)\mid a\in\mathbb{C}_0^n\}$	$3a_{\epsilon_i}$	$2a_{\epsilon_i}+a_{\epsilon_j}$	$a_{\epsilon_i}+a_{\epsilon_j}+a_{\epsilon_k}$
2	$\mathscr{A}_2$	$\{A=\iota(a)\mid a\in S^2(\mathbb{C}_0^n)\}$	$3a_{2\epsilon_i}$	$a_{2\epsilon_i}+2a_{\epsilon_i+\epsilon_j}$	$a_{\epsilon_i+\epsilon_j}+a_{\epsilon_i+\epsilon_k}+a_{\epsilon_j+\epsilon_k}$
3	$\mathscr{A}_3$	$\{A=\iota(a)\mid a\in S^3(\mathbb{C}_0^n)\}$	$a_{3\epsilon_i}$	$a_{2\epsilon_i+\epsilon_j}$	$a_{\epsilon_i+\epsilon_j+\epsilon_k}$

In the sequel, for every hadron-type $\mathfrak{g}$ representation, we describe the restrictions of F to any component $\mathscr{T}_q U_q(0)\subset(\mathrm{End}\,V)(0)$, where $\mathrm{End}\,V=\Sigma_q\mathscr{T}_q(U_q)$ is a decomposition into irreducible tensor operators. This result enables us to derive possible relations depending on the tensor type of a given observable.

A. (Skew) symmetrical sl(n,$\mathbb{C}$) representations

We begin with the case where $V=S^p$ or $V=\wedge^p$. Corollary 1 (resp. 2) provides the decomposition of the space $\mathrm{End}\,S^p$ (resp. $\mathrm{End}\wedge^p$) into irreducible tensor operators. Namely, the space U_q coincides with the realization (2.3) of $\mathscr{A}_q$ [resp. (2.21) of $\mathscr{C}_q$], and the mapping $\mathscr{T}_q$ is equal to $[1/(p-q)!]J_+^{p-q}$, where J_+ is given by (2.1) [resp. (2.19)]. In particular,

$$\mathscr{T}_1=\frac{1}{(p-1)!}J_+^{p-1}\ :\ \mathscr{A}_1=\mathscr{C}_1=\mathfrak{g}\to\mathrm{End}\,V$$

coincides with the mapping ρ. Moreover, formulas (2.14) and (2.29) imply the following corollary.

Corollary 3: An operator $A\in(\mathrm{End}\,S^p)(0)$ [resp. $(\mathrm{End}\wedge^p)(0)$] is a tensor operator of the type $\mathscr{A}_q$, $0\leqslant q\leqslant p$ [resp. $\mathscr{C}_q$, $0\leqslant q\leqslant\min(p,n-p)$] iff

$$F_A(\lambda)=\sum_{\mu\in\mathscr{L}_q}\binom{\lambda}{\mu}a_\mu,\quad a\in S^q(\mathbb{C}_0^n),\quad \lambda\in\mathscr{L}_p$$

$$\left(\text{resp. } F_A(\lambda_I)=\sum_{\pounds_q\ni J\subset I}a_J,\quad a=(a_j)\in\mathbb{C}_0^{n,q},\quad I\in\pounds_p.\right)$$

[Note that the set $\mathscr{L}_p$ (resp. $\pounds_p$) describes the set of weights of the representation S^p (resp. $\wedge^p$).] ■

In Tables I and II, we specify the consequences of Corollary 3 in the case of hadron-type representations $S^3=\mathscr{A}_0^3$ and $\wedge^3=\mathscr{B}_2^1$.

The detailed description of the S_n representations, appearing in Tables I and II, is given in Appendix A. Using it, one can easily get possible relations for eigenvalues of observables.

TABLE II. Representation End $\wedge^3$.

q	U_q	$U_q(0)$	$F_A(\epsilon_i+\epsilon_j+\epsilon_k)$, $ijk\neq$
0	$\mathscr{C}_0$	$\{A=\iota(a)=a\cdot\mathrm{id}_V\}$	a
1	$\mathscr{C}_1$	$\{A=\iota(a)\mid a\in\mathbb{C}_0^n\}$	$a_i+a_j+a_k$
2	$\mathscr{C}_2$	$\{A=\iota(a)\mid a\in\mathbb{C}_0^{n,2}\}$	$a_{ij}+a_{jk}+a_{ki}$
3	$\mathscr{C}_3$	$\{A=\iota(a)\mid a\in\mathbb{C}_0^{n,3}\}$	a_{ijk}

B. Technical remarks

Before studying the cases $V=\mathscr{B}_2^1$ and $V=\mathscr{A}_1=\mathfrak{g}$, we would like to give some useful facts and a convenient notation.

Let W be a $(p-1)$-dimensional vector space spanned by fixed vectors $w_1,\dots,w_p$ such that $w_1+\cdots+w_p=0$. Then, any $(p-1)$-element subset $w_1,\dots,\hat{w}_k,\dots,w_p$ forms a basis in W. However, the distinction of any w_k can be undesirable. We shall call the p-plet $(w_1,\dots,w_p)$ a reper of the space W. Elements of W can be uniquely decomposed with respect to a given reper: $w=\Sigma_i x^i w_i$, provided $\Sigma_i x^i=0$. (In this section Σ_i denotes $\Sigma_{1\leqslant i\leqslant p}$.) The correspondence

$$w\mapsto[w]=\begin{bmatrix}x^1\\ \vdots\\ x^p\end{bmatrix}$$

defines a linear isomorphism that maps the reper $(w_1,\dots,w_p)$ of W onto the reper $(\bar{w}_1,\dots,\bar{w}_p)$ of the space $\mathbb{C}_0^p$, where $\bar{w}_i:=\epsilon_i-(1/p)\Sigma_j\epsilon_j$.

Any endomorphism $Q\in\mathrm{End}\,W$ can be uniquely described by the $p\times p$ matrix $[Q]=[Q_j^i]$ such that $\Sigma_j Q_j^i=0$, $i\in\overline{1,p}$, and $\Sigma_i Q_j^i=0$, $j\in\overline{1,p}$, where $[Qw]=[Q][w]$. It is clear that $[Q\circ P]=[Q]\cdot[P]$ and $[Q+P]=[Q]+[P]$. Moreover, the condition

$$[Q]\cdot\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}=0$$

implies that $\det[Q_j^i+\lambda\delta_j^i]=\lambda\cdot\det(Q+\lambda\cdot\mathrm{id}_W)$. In other words, for $1\leqslant k<p$, the kth algebraic invariants of Q and $[Q]$ coincide, whereas $\det[Q]=0$. It implies, in particular,

$$\mathrm{tr}\,Q=\mathrm{tr}[Q]=Q_1^1+\cdots+Q_p^p,\tag{3.2}$$

$$\det Q=(\text{the sum of the main } (p-1) \text{ minors of } [Q]).\tag{3.3}$$

Let us introduce the action of the group S_p on W by the formula $\sigma w_k=w_{\sigma(k)}$, $\sigma\in S_p$. It provides the irreducible S_p representation equivalent to $\mathbb{C}_0^p$. The corresponding S_p representation in End W fulfills

$$\begin{aligned}\mathrm{End}\,W=W\otimes W^*&=\mathbb{C}_0^p\otimes\mathbb{C}_0^p=S^2(\mathbb{C}_0^p)\dotplus\wedge^2(\mathbb{C}_0^p)\\ &=\mathbb{C}\dotplus\mathbb{C}_0^p\dotplus\mathbb{C}_0^{p,2}\dotplus\wedge^2(\mathbb{C}_0^p),\end{aligned}\tag{3.4}$$

see Lemma 4 in Appendix A 3. In particular, for $p=3$, $\mathbb{C}_0^{p,2}$ vanishes and we have

$$[Q]=\frac{\alpha}{3}\begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{bmatrix} + \begin{bmatrix}\beta_1 & \beta_3 & \beta_2\\ \beta_3 & \beta_2 & \beta_1\\ \beta_2 & \beta_1 & \beta_3\end{bmatrix} + \gamma\begin{bmatrix}0 & 1 & -1\\ -1 & 0 & 1\\ 1 & -1 & 0\end{bmatrix}, \tag{3.5}$$

where $\alpha\in\mathbb{C}$, $(\beta_1,\beta_2,\beta_3)\in\mathbb{C}_0^3$, $\gamma\in\overline{\mathbb{C}}\simeq\Lambda^2(\mathbb{C}_0^3)$. Indeed, a transposition of two elements in the reper (w_1,w_2,w_3) corresponds to a transposition of two rows and corresponding columns in $[Q]$. Such operations (1) do not change α, (2) transpose two (corresponding) coordinates of $(\beta_1,\beta_2,\beta_3)$, and (3) multiply γ by -1.

C. Adjoint sl(n,$\mathbb{C}$) representation

Let us denote the basic vectors in the tensor space $T^p_q(E):=E^{\otimes q}\otimes(E^*)^{\otimes p}$ by $e^{i_1\cdots i_p}_{j_1\cdots j_q}$. For the space

$$V=\mathfrak{g}=\mathrm{sl}(E)=\left\{X=\sum_{ik}X^i_k\,e^k_i\,\middle|\,\sum_i X^i_i=0\right\}\subset T^1_1(E), \tag{3.6}$$

the space End $V=$ End $\mathfrak{g}$ can be realized as follows:

$$\mathrm{End}\,\mathfrak{g}=\left\{A=\sum_{ijkl}A^{ij}_{kl}\,e^{kl}_{ij}\,\middle|\,\sum_i A^{ij}_{il}=0,\ \sum_j A^{ij}_{kj}=0\right\}\subset T^2_2(E), \tag{3.7}$$

where $AX:=\Sigma_{ik}(\Sigma_{jl}A^{ij}_{kl}X^l_j)e^k_i$, i.e., $Ae^j_l=\Sigma_{ik}A^{ij}_{kl}\,e^k_i$. Indeed, $T^1_1(E)$ coincides with $\mathfrak{g}+\langle\mathrm{id}_E\rangle$ and for $A\in T^2_2(E)=\mathrm{End}\,T^1_1(E)$, the condition $\Sigma_i A^{ij}_{il}=0$ (resp. $\Sigma_j A^{ij}_{kj}=0$) means that $AT^1_1(E)\subset\mathfrak{g}$ (resp. $A\langle\mathrm{id}_E\rangle=0$).

The set of weights $\underline{\Lambda}$ of the adjoint representation is the union of the zero-weight and the root system $R=\{\alpha_{ij}\,|\,i\neq j\}$, $\alpha_{ij}:=\epsilon_i-\epsilon_j$. The zero-weight space $\mathfrak{g}(0)=\mathfrak{h}$ is spanned by the reper $(w_1,..,w_n)$, where $w_i:=e^i_i-(1/n)\Sigma_j e^j_j$, whereas $\mathfrak{g}(\alpha_{ij})=\langle e^j_i\rangle$, $i\neq j$. Thereby, we obtain $(\mathrm{End}\,\mathfrak{g})(0)=\Sigma_{\lambda\in\underline{\Lambda}}\mathrm{End}\,\mathfrak{g}(\lambda)=\mathrm{End}\,\mathfrak{g}(0)+\mathrm{End}\,\mathfrak{g}(R)\simeq\mathrm{End}\,\mathbb{C}_0^n+\mathbb{C}^R$, $\mathfrak{g}(R):=\Sigma_{\lambda\in R}\mathfrak{g}(\lambda)$. In other words, any operator $A\in(\mathrm{End}\,\mathfrak{g})(0)$ can be uniquely described by two $n\times n$ matrices: (1) the matrix $[a_{ij}]$ which satisfies $\Sigma_i a_{ij}=0$, $j\in\overline{1,n}$, $\Sigma_j a_{ij}=0$, $i\in\overline{1,n}$ and is the matrix of the operator $A|_{\mathfrak{g}(0)}$ with respect to the reper $(w,...,w_n)$, i.e., $Aw_j=\Sigma_i w_i a_{ij}$; and (2) the diagonalless matrix $[b_{ij}]$, $i\neq j$, given by $Ae^i_j=b_{ij}e^i_j$, $i\neq j$, which characterizes the operator $A|_{\mathfrak{g}(R)}$. Using the vectors e^{ij}_{ij} and e^{ij}_{ji}, which span the zero-weight space in $T^2_2(E)$, we can describe the decomposition given above in the following way:

$$(T^2_2(E))(0)\supset(\mathrm{End}\,\mathfrak{g})(0)=\left\{A=\sum_{ij}a_{ij}\,e^{ij}_{ij}+\sum_{ij,\ i\neq j}b_{ij}\,e^{ij}_{ji}\,\middle|\,\sum_i a_{ij}=0,\ \sum_j a_{ij}=0\right\}.$$

The last formula enables us to describe the decomposition $(\mathrm{End}\,\mathfrak{g})(0)=\Sigma_q\mathscr{T}_qU_q(0)$ and the mapping F on the subspaces $\mathscr{T}_qU_q(0)$. The results are collected in Table III, where we used the following mappings $\mathscr{T}_q:U_q\to\mathrm{End}\,\mathfrak{g}$ [$\mathfrak{g}$ and End $\mathfrak{g}$ are given by (3.6) and (3.7)]:

$$\begin{aligned}
&U_0=\mathbb{C}\ni c\mapsto\mathscr{T}_0(c):=c\cdot\mathrm{id}_V=c\sum_{ij}\Big(e^{ij}_{ji}-\frac{1}{n}e^{ij}_{ij}\Big),\\
&U_1=\mathfrak{g}\ni X\mapsto\mathscr{T}_1(X):=\mathrm{ad}\,X=\sum_{ijk}X^i_j(e^{kj}_{ik}-e^{jk}_{ki}),\\
&U_2=\mathfrak{g}\ni X\mapsto\mathscr{T}_2:=\sum_{ijk}X^i_j\Big(e^{kj}_{ik}+e^{jk}_{ki}-\frac{2}{n}e^{jk}_{ik}-\frac{2}{n}\ \ e^{kj}_{ki}\Big),\\
&U_3=\{A\in\mathrm{End}\,\mathfrak{g}\,|\,A^{ij}_{kl}=A^{(ij)}_{(k,l)}\}\ni A\mapsto\mathscr{T}_3(A):=A,\\
&U_4=\{A\in\mathrm{End}\,\mathfrak{g}\,|\,A^{ij}_{kl}=A^{[ij]}_{(kl)}\}\ni A\mapsto\mathscr{T}_4(A):=A,\\
&U_5=\{A\in\mathrm{End}\,\mathfrak{g}\,|\,A^{ij}_{kl}=A^{(ij)}_{[kl]}\}\ni A\mapsto\mathscr{T}_5(A):=A,\\
&U_6=\{A\in\mathrm{End}\,\mathfrak{g}\,|\,A^{ij}_{kl}=A^{[ij]}_{[kl]}\}\ni A\mapsto\mathscr{T}_6(A):=A.
\end{aligned}\tag{3.8}$$

The mapping $\mathscr{T}_2$ coincides with the Gell-Mann mapping D since $\mathscr{T}_2(X)Y=\{$traceless part of $XY+YX\}$. Note that for $q\geqslant 3$, the function F_A, $A\in U_q(0)$, given in the Table III, is expressed only by means of the matrix $[b_{ij}]$.

In particular, Table III implies that

(i) $\mathscr{T}_0(c)|_{\mathfrak{g}(0)}=-\mathscr{T}_3\circ\iota_3(a)|_{\mathfrak{g}(0)}$,

$\mathscr{T}_0(c)|_{\mathfrak{g}(R)}=n\cdot\mathscr{T}_3\circ\iota_3(a)|_{\mathfrak{g}(R)}$,

if $a=(a_{ij})\in S^2(\mathbb{C}_0^n)$ corresponds to $c\in\mathbb{C}$ by the natural embedding $\mathbb{C}\hookrightarrow S^2(\mathbb{C}_0^n)$, i.e, $a_{ij}=c((1/n)-\delta_{ij})$;

(ii) $\mathscr{T}_2\circ\iota_2(x)|_{\mathfrak{g}(0)}=-(2/n)\mathscr{T}_3\circ\iota_3(a)|_{\mathfrak{g}(0)}$,

$\mathscr{T}_2\circ\iota_2(x)|_{\mathfrak{g}(R)}=\mathscr{T}_3\circ\iota_3(a)|_{g(R)}$,

if $a=(a_{ij})$ corresponds to $x\in\mathbb{C}_0^n$ by the natural embedding $\mathbb{C}_0^n\hookrightarrow S^2(\mathbb{C}_0^n)$, i.e., $a_{ij}=(1-(n/2)\delta_{ij})(x_i+x_j)$;

(iii) $\mathscr{T}_6\circ\iota_6(a)|_{\mathfrak{g}(0)}=\mathscr{T}_3\circ\iota_3(a)|_{\mathfrak{g}(0)}$,

$\mathscr{T}_6\circ\iota_6(a)|_{\mathfrak{g}(R)}=-\mathscr{T}_3\circ\iota_3(a)|_{\mathfrak{g}(R)}$,

if $a\in\mathbb{C}_0^{n,2}\hookrightarrow S^2(\mathbb{C}_0^n)$; see Appendix A 3. Thus, using tensor operators of various types, we can keep the basic physical relations (S_n-transformation properties of a given observable), and simultaneously fit some additional relations. For instance, the mass differences between mesons $\pi_\pm$ and π_0 (or $\rho_\pm$ and ρ_0) can be fitted by using the additional term $\mathscr{T}_3\circ\iota_3(a)$, where $a=(a_{ij})$ is the image of $c\in\mathbb{C}$ or $x\in\mathbb{C}_0^n$, see the embeddings in (i) and (ii) given above.

D. sl(n,$\mathbb{C}$) representation $\mathscr{B}^1_2$

We may assume that $n\geqslant 4$ since for $n=3$, the representation $\mathscr{B}^1_2$ coincides with the adjoint representation. Let $Y:=(1+(2,3))(1-(1,2))$ be the Young symmetrizer. The representation $\mathscr{B}^1_2$ can be realized in the space $V:=YT^3_0(E)\subset T^3_0(E)$ which is spanned by vectors

$$w_{ijk}:=\tfrac{1}{3}Y(e_{ijk}-e_{kij})=\tfrac{2}{3}(e_{ijk}+e_{ikj})-\tfrac{1}{3}(e_{jki}+e_{jik}+e_{kij}+e_{kji}) \tag{3.9}$$

satisfying the equations $w_{ijk}=w_{ikj}$ and $w_{ijk}+w_{jki}+w_{kij}=0$.

In other words

$$V=\left\{v=\sum\xi^{ijk}e_{ijk}=\frac{1}{2}\sum\xi^{ijk}w_{ijk}\,\middle|\,\xi^{ijk}=\xi^{i(jk)},\right.$$

TABLE III. Representation End $\mathfrak{g}$.

q	U_q	$U_q(0)$	$b_{ij}=F_A(\alpha_{ij}),\ i\neq j$	$a_{ij},\ i\neq j$	a_{ii}	$F_A(0)=\sum_i a_{ii}$
0	$\mathscr{A}_0$	$\{c\mid c\in\mathbb{C}\}$	c	$-c/n$	$[(n-1)/n]c$	$(n-1)c$
1	$\mathscr{A}_1$	$\{X=\iota_1(x)=\sum_i x_i e_i^i \mid x\in\mathbb{C}_0^n\}$	x_i-x_j	0	0	0
2	$\mathscr{A}_1$	as above	x_i+x_j	$-\frac{2}{n}(x_i+x_j)$	$\frac{2(n-2)}{n}x_i$	0
3	$\mathscr{A}_2$	$\left\{A=\iota_3(a)=\sum_{ij}a_{ij}e_{ij}^{ij}+\sum_{i\neq j}a_{ij}e_{ji}^{ij}\mid a\in S^2(\mathbb{C}_0^n)\right\}$	$b_{ij}=b_{ji}$	b_{ij}	$-\sum_{j,j\neq i}b_{ij}$	$-\sum_{i\neq j}b_{ij}$
4	$\mathscr{B}_3^{n-3}$	$\left\{A=\iota_4(a)=\sum_{ij}a_{ij}(e_{ij}^{ij}-e_{ji}^{ij}\mid a\in\Lambda^2(\mathbb{C}_0^n)\right\}$	$b_{ij}=-b_{ji};\ \sum_{i,i\neq j}b_{ij}=0$	$-b_{ij}$	0	0
5	$(\mathscr{B}_3^{n-3})^*$	$\left\{A=\iota_5(a)=\sum_{ij}a_{ij}(e_{ij}^{ij}+e_{ji}^{ij}))\mid a\in\Lambda^2(\mathbb{C}_0^n)\right\}$	as above	b_{ij}	0	0
6	$\mathscr{C}_2$	$\left\{A=\iota_6(a)=\sum_{ij}a_{ij}(e_{ij}^{ij}-e_{ji}^{ij})\mid a\in\mathbb{C}_0^{n,2}\right\}$	$b_{ij}=b_{ji};\ \sum_{i,i\neq j}b_{ij}=0$	$-b_{ij}$	0	0

$$\left.\xi^{ijk}+\text{cycl}(ijk)=0\right\}$$

(we shall omit the summation indices unless it causes any misunderstandings). The space V^* can be identified with $YT_3^0(E)$ which is spanned by vectors w^{ijk} satisfying the same equations as the w_{ijk}'s. Moreover, the relations shown in Tables IV and V are fulfilled.

Setting $w_{ijk}^{pqr} := w_{ijk}\otimes w^{pqr}$, we obtain

$$\begin{aligned}\text{End } V=\langle w_{ijk}^{pqr}\rangle &= V\otimes V^*\\ &=\left\{A=\sum A_{pqr}^{ijk}\,e_{ijk}^{pqr}=\frac{1}{4}\sum A_{pqr}^{ijk}\,w_{ijk}^{pqr}\right|\\ &\quad A_{pqr}^{ijk}=A_{p(qr)}^{i(jk)},\\ &\quad \left.A_{pqr}^{ijk}+\text{cycl}(ijk)=0,\quad A_{pqr}^{ijk}+\text{cycl}(pqr)=0\right\}.\end{aligned}\tag{3.10}$$

The set of weights $\underline{\Lambda}$ of $\mathscr{B}_2^1$ is a union of two S_n orbits: $\underline{\Lambda}' := \{\lambda_{ij}\mid i\neq j\}$ and $\underline{\Lambda}'' := \{\lambda_{ijk}\mid ijk\neq\}$, where $\lambda_{ij}=2\epsilon_i+\epsilon_j$ and $\lambda_{ijk}=\epsilon_i+\epsilon_j+\epsilon_k$. The corresponding weight spaces are given by $V(\lambda_{ij})=\langle w_{iji}\rangle$ and $V(\lambda_{ijk})=\langle w_{ijk}, w_{jki}, w_{kij}\rangle$, $\dim V(\lambda_{ijk})=2$. Hence, we see that

$$\begin{aligned}(\text{End } V)\ (0) &= \sum_{\lambda\in\underline{\Lambda}}\text{End } V(\lambda)\\ &=\{A\in\text{End } V\mid A_{pqr}^{ijk}\neq 0\Rightarrow(p,q,r)\in S_3(i,j,k)\},\end{aligned}$$

where S_3 is the permutation group. Using the Tables IV and V, one can check that for any $A\in(\text{End } V)(0)$,

$$Aw_{ijk}=\begin{cases}2(A_{ijk}^{ijk}w_{ijk}+A_{ijk}^{jki}w_{jki}+A_{ijk}^{kij}w_{kij}), & \text{if } ijk\neq, \quad (3.11)\\ 6A_{iji}^{iji}w_{iji}, & \text{if } i=k\neq j, \quad (3.12)\end{cases}$$

holds. The formula (3.11) means that for fixed $ijk\neq$,

$$[Q]=2\begin{bmatrix}A_{ijk}^{ijk} & A_{jki}^{ijk} & A_{kij}^{ijk}\\ A_{ijk}^{jki} & A_{jki}^{jki} & A_{kij}^{jki}\\ A_{ijk}^{kij} & A_{jki}^{kij} & A_{kij}^{kij}\end{bmatrix}\tag{3.13}$$

is the matrix of an operator $A\mid_{V(\lambda_{ijk})}$, $A\in(\text{End } V)(0)$, with respect to the reper $(w_1,w_2,w_3) := (w_{ijk},w_{jki},w_{kij})$ of the space $V(\lambda_{ijk})$ (cf. Ref. 1.—the consequences I–IV of Lemma 3 concerning induced representations).

Let us denote by α_{ijk}, $(\beta_{ijk},\beta_{jki},\beta_{kij})$, and γ_{ijk} the parameters corresponding to the matrix (3.13) at the decomposition (3.5), i.e.,

$$\begin{aligned}A_{ijk}^{ijk}&=\tfrac{1}{3}\alpha_{ijk}+\tfrac{1}{2}\beta_{ijk},\\ A_{jki}^{ijk}&=-\tfrac{1}{6}\alpha_{ijk}+\tfrac{1}{2}\beta_{kij}+\tfrac{1}{2}\gamma_{ijk}.\end{aligned}\tag{3.14}$$

Clearly, a change of the order of the (i,j,k) is equivalent to the permutation of the elements in the reper (w_1,w_2,w_3). Taking into account the S_3-symmetry properties of $[Q]$ described below the decomposition (3.5) and in the formula (3.10), we see that the quantities $\alpha=(\alpha_{ijk})$, $\beta=(\beta_{ijk})$ and $\gamma=(\gamma_{ijk})$, $ijk\neq$, fulfill

$$\begin{aligned}&\alpha_{ijk}=\alpha_{(ijk)},\quad \beta_{ijk}=\beta_{ikj},\\ &\beta_{ijk}+\text{cycl}(ijk)=0,\quad \gamma_{ijk}=\gamma_{[ijk]}.\end{aligned}\tag{3.15}$$

In other words, we label any $A\in(\text{End } V)(0)$ by (α,β,γ) in such a way that $(\sigma\alpha,\sigma\beta,\sigma\gamma)$ corresponds to the operator $\sigma A := \Pi_\sigma(A)$, $\sigma\in S_n$, where $(\sigma\alpha)_{ijk} := \alpha_{\sigma^{-1}(i)\sigma^{-1}(j)\sigma^{-1}(k)}$, etc. Let us notice that to get a complete charaterization of an operator $A\in(\text{End } V)(0)$, we also need the quantity

$$\kappa=(\kappa_{ij}),\quad i\neq j,\quad \kappa_{ij} := 6A_{iji}^{iji},\tag{3.16}$$

describing A on the one-dimensional weight spaces $V(\lambda_{ij})$ [cf. (3.12)].

The correspondence $(\text{End } V)(0)\ni A\mapsto(\alpha,\beta,\gamma,\kappa)$ provides an S_n isomorphism $A=\iota(\alpha,\beta,\gamma,\kappa)$. The formulas (3.15) and (3.16) mean that the quantities α, γ, and κ trans-

TABLE IV. $(ijk\neq)$.

(pqr)	(ijk), (ikj)	(jki), (jik), (kij), (kji)
$\langle w^{pqr},w_{ijk}\rangle$	$\frac{4}{3}$	$-\frac{2}{3}$

TABLE V. $(i\neq j)$.

(pqr)	(iji), (iij)	(jii)
$\langle w^{pqr},w_{iji}\rangle$	$\frac{2}{3}$	$-\frac{4}{3}$

form according to the representations $\mathbb{C}^{n,3}$, $\wedge^3(\mathbb{C}^n)$, and $\wedge^2(\mathbb{C}^n)\dot{+}\mathbb{C}^{n,2}$, respectively. In order to analyze the representation β, it is convenient to consider the space $\hat{D}_n := \{\beta = (\beta_{ijk}) \,|\, \beta \text{ satisfies (3.15) and } \beta_{ijk} = 0 \text{ unless } ijk \neq\}$ and its S_n-invariant subspace

$$D_n := \{\beta \in \hat{D}_n \,|\, \beta^+ = \beta^- = 0\}, \quad \beta^{\pm} := (\beta_{ij}^{\pm}),$$
$$\beta_{ij}^+ := \sum_k \beta_{kij}, \quad \beta_{ij}^- := \sum_k (\beta_{ijk} - \beta_{jik}). \tag{3.17}$$

Any $\beta \in \hat{D}_n$ has the unique decomposition

$$\beta_{ijk} = (2x_i - x_j - x_k) + (a_{ij} + a_{ik}) + (2b_{jk} - b_{ij} - b_{ik}) + d_{ijk}, \quad ijk \neq,$$

where

$$x = (x_i) \in \mathbb{C}_0^n, \quad a = (a_{ij}) \in \wedge^2(\mathbb{C}_0^n),$$
$$b = (b_{ij}) \in \mathbb{C}_0^{n,2}, \quad d = (d_{ijk}) \in D_n.$$

Thus

$$\hat{D}_n = \mathbb{C}_0^n \dot{+} \wedge^2(\mathbb{C}_0^n) \dot{+} \mathbb{C}_0^{n,2} \dot{+} D_n. \tag{3.18}$$

It can be shown that $D_4 = 0$, whereas D_n, $n \geqslant 5$, realizes the irreducible S_n representation of the signature $(n-3,2,1)$. Thus we get [cf. (A5) and (A7)]

$$(\text{End } V)(0) = \begin{cases} 2\mathbb{C} \dot{+} 4\mathbb{C}_0^4 \dot{+} 3\wedge^2(\mathbb{C}_0^4) \dot{+} \wedge^3(\mathbb{C}_0^4) \dot{+} 2\mathbb{C}_0^{4,2}, & n = 4, \\ 2\mathbb{C} \dot{+} 4\mathbb{C}_0^n \dot{+} 3\wedge^2(\mathbb{C}_0^n) \dot{+} \wedge^3(\mathbb{C}_0^n) \dot{+} 3\mathbb{C}_0^{n,2} \dot{+} \mathbb{C}_0^{n,3} \dot{+} D_n, & n \geqslant 5. \end{cases} \tag{3.19}$$

On the other hand, the S_n representation $\mathbb{C}^{\Delta} = \mathbb{C}^{\Delta'} \dot{+} C^{\Delta''}$ can be parametrized by κ and α. Therefore, contrary to $(\text{End } V)(0)$, the space $\mathbb{C}^{\Delta}$ does not contain the representations $\wedge^3(\mathbb{C}_0^n)$ and D_n—they are lost by the mapping F given by (3.1). More precisely, (3.11)–(3.16) and (3.2) imply that $F_A(\lambda_{ij}) = \kappa_{ij}$, $F_A(\lambda_{ijk}) = 2\alpha_{ijk}$, $A \in (\text{End } V)(0)$. Note that (3.3) expresses $\det(A\,|_{V(\lambda_{ijk})})$ by all the parameters α,β,γ.

The detailed results are given in Tables VI and VII. For $q = 0,\ldots,6$, the used spaces U_q are the same as in Table III [see (3.6)–(3.8)], whereas the embeddings $\mathscr{T}_q$ are given in the Table VI (note that $\mathscr{T}_{1=\rho}$). The space $\mathscr{D}$ of signature $(4,3,2,\ldots,2,1,0)$, see decomposition (1.2), coincides with $U_7 := \{A \in \text{End } V \,|\, \Sigma_k A_{pqk}^{ijk} = 0\}$. One can check that an operator $A = \iota(\alpha,\beta,\gamma,\kappa)$ belongs to $U_7(0)$ iff

$$\gamma \in \wedge^3(\mathbb{C}_0^n), \quad \kappa = \tfrac{1}{2}(3\beta^+ + \beta^-), \quad \beta \in \hat{D}_n, \tag{3.20}$$
$$J_-\alpha + 3\beta^+ = 0, \tag{3.21}$$

where

$$(J_-\alpha)_{jk} = \sum_i \alpha_{ijk}.$$

For $n \geqslant 5$, Eq. (3.21) can be transformed, e.g., as follows:

$$\alpha = \tilde{\alpha} - \tilde{\beta}, \quad \tilde{\alpha} \in \mathbb{C}_0^{n,3},$$
$$\tilde{\beta}_{ijk} := \frac{3}{(n-4)}(\beta_{ij}^+ + \beta_{jk}^+ + \beta_{ki}^+) + \frac{3}{(n-3)(n-4)}(\beta_i + \beta_j + \beta_k), \tag{3.22}$$

where $\beta_j := \Sigma_i \beta_{ij}^+ = \Sigma_{ki}\beta_{kij}$. Thus we get that

$$\mathscr{D}(0) = U_7(0) = \wedge^3(\mathbb{C}_0^n) \dot{+} \mathbb{C}_0^{n,3} \dot{+} \hat{D}_n. \tag{3.23}$$

The expressions for α,β,γ, and κ given in Tables VI and VII are derived by means of definition (3.9) and Tables IV and V. Note that for $n = 4$, $\mathbb{C}_0^{4,3} = D_4 = 0$, $\wedge^3(\mathbb{C}_0^4) = \overline{\mathbb{C}}$, and $\wedge^2(\mathbb{C}_0^4) = \overline{\mathbb{C}}_0^4$. We parametrize the irreducible S_4 representations appearing in Table VII as follows: $\mathbb{C} = \{c\}$, $\overline{\mathbb{C}} = \{d\}$, $\mathbb{C}_0^4 = \{x \in \mathbb{C}^4 \,|\, \Sigma_i x_i = 0\}$, $\overline{\mathbb{C}}_0^4 = \{y \in \mathbb{C}^4 \,|\, \Sigma_i y_i = 0\}$, $\mathbb{C}_0^{4,2} = \{a \in \mathbb{C}^{4,2} \,|\, a_{ij} = a_{kl} \text{ if } \{i,j,k,l\} = \overline{1,4}\}$. The skew-symmetric symbol ϵ_{ijk} is defined by $\epsilon_{\sigma(1)\sigma(2)\sigma(3)} = \text{sgn}\,\sigma$, $\sigma \in S_4$. Moreover we use the following convention: $(ijkl)$ runs all the permutations of $(1,2,3,4)$, $\lambda_l := \lambda_{ijk}$, $\alpha_l := \alpha_{ijk}$; summation indices m, p run $\overline{1,4}$, and $\kappa_{ii} = 0$.

The last column in Table VII contains the complete set of relations for the values $F_A(\lambda_{ijk})$ and $F_A(\lambda_{ij})$ (due to our convention the quantificators can be omitted).

APPENDIX A: ON SOME REPRESENTATIONS OF S_n

We analyze here some properties of the S_n representations which appear in the zero-weight spaces of the arms representations of $\text{sl}(n,\mathbb{C})$. For simplicity of the exposition, the proofs of the following lemmas are given in Appendix B.

TABLE VI. Representation $\mathscr{B}_2^1$, $n \geqslant 5$.

q	U_q	$\mathscr{T}_q$	$U_q(0)$	$\alpha_{ijk} = \frac{1}{2}F_A(\lambda_{ijk})$	β_{ijk}	γ_{ijk}	$\kappa_{ij} = F_A(\lambda_{ij})$
0	$\mathscr{A}_0$	$\mathscr{T}_0(c) = c \cdot \text{id}_V = (c/12)w_{ijk}^{ijk}$	$\mathbb{C}$	c	0	0	c
1	$\mathscr{A}_1$	$\mathscr{T}_1(X) = \frac{1}{4}\sum X_j^i(w_{ikl}^{jkl} + w_{kil}^{kjl} + w_{kli}^{klj})$	$\mathbb{C}_0^n$	$x_i + x_j + x_k$	0	0	$2x_i + x_j$
2	$\mathscr{A}_1$	$\mathscr{T}_2(X) = \frac{1}{2}\sum X_j^i(w_{ikl}^{jkl} - w_{kli}^{klj})$	$\mathbb{C}_0^n$	0	$\frac{1}{3}(2x_i - x_j - x_k)$	0	$x_j - x_i$
3	$\mathscr{A}_2$	$\mathscr{T}_3(A) = \frac{3}{8}\sum A_{kl}^{ij}\, w_{mij}^{mkl}$	$S^2(\mathbb{C}_0^n)$	$a_{ij} + a_{jk} + a_{ki}$	$\frac{1}{3}(2a_{jk} - a_{ij} - a_{ik})$	0	$a_{ii} + a_{ij}$
4	$\mathscr{B}_3^{n-3}$	$\mathscr{T}_4(A) = \frac{1}{4}\sum A_{kl}^{ij}\, w_{ijm}^{mkl}$	$\wedge^2(\mathbb{C}_0^n)$	0	$-\frac{1}{3}(a_{ij} + a_{ik})$	$-\frac{1}{3}(a_{ij} + a_{jk} + a_{ki})$	a_{ij}
5	$(\mathscr{B}_3^{n-3})^*$	$\mathscr{T}_5(A) = \frac{1}{4}\sum A_{kl}^{ij}\, w_{mij}^{klm}$	$\wedge^2(\mathbb{C}_0^n)$	0	as above	$\frac{1}{3}(a_{ij} + a_{jk} + a_{ki})$	a_{ij}
6	$\mathscr{C}_2$	$\mathscr{T}_6(A) = \frac{1}{4}\sum A_{kl}^{ij}\, w_{ijm}^{klm}$	$\mathbb{C}_0^{n,2}$	$\frac{1}{3}(a_{ij} + a_{jk} + a_{ki})$	$\frac{1}{9}(a_{ij} + a_{ik} - 2a_{jk})$	0	a_{ij}
7	$\mathscr{D}$	natural embedding	see the formulas	(3.18)–(3.23)			

TABLE VII. Representation $\mathscr{B}_2^1$, $n=4$.

q	$U_q(0)$	$\alpha_l=\frac{1}{2}F_A(\lambda_l)$	β_{ijk}	γ_{ijk}	$\kappa_{ij}=F_A(\lambda_{ij})$	Relations
0	$\mathbb{C}$	c	0	0	c	$\alpha_l=\kappa_{ij}$
1	$\mathbb{C}_0^4$	$-x_l$	0	0	$2x_i+x_j$	$\sum_m \alpha_m=0;\ \ \kappa_{ij}+2\alpha_i+\alpha_j=0$
2	$\mathbb{C}_0^4$	0	$\frac{1}{4}(3x_i+x_l)$	0	x_j-x_i	$\alpha_l=0;\ \ \kappa_{ij}=-\kappa_{ji};\ \ \kappa_{ij}+\kappa_{jk}+\kappa_{ki}=0$
3	$\mathbb{C}+\mathbb{C}_0^4+\mathbb{C}_0^{4,2}$	$2x_l+3c$	$a_{ij}+\frac{1}{4}(3x_i+x_l)$	0	a_{ij}	$2\alpha_l=-\sum_m \kappa_{ml};\ \ \sum_{mp\neq l}\kappa_{mp}=\sum_m(\kappa_{ml}+\kappa_{lm});$ $\kappa_{ij}+\kappa_{jk}+\kappa_{ki}=\kappa_{ji}+\kappa_{kj}+\kappa_{ik}$
4	$\overline{\mathbb{C}}_0^4$	0	$\frac{1}{2}\epsilon_{ijk}\cdot(y_i-y_k)$	$\frac{1}{3}\epsilon_{ijk}\cdot y_l$	$\sum_m \epsilon_{ijm}\cdot y_m$	$\alpha_l=0;\ \ \kappa_{ij}=-\kappa_{ji};\ \ \sum_m \kappa_{im}=0$
5	$\overline{\mathbb{C}}_0^4$	0	as above	$-\frac{1}{3}\epsilon_{ijk}\cdot y_l$	as above	as above
6	$\mathbb{C}_0^{4,2}$	0	$-\frac{1}{3}a_{jk}$	0	a_{ij}	$\alpha_l=0;\ \ \kappa_{ij}=\kappa_{ji};\ \ \sum_m \kappa_{im}=0$
7	$\mathbb{C}_0^4+\overline{\mathbb{C}}_0^4+\overline{\mathbb{C}}$	x_l	$\epsilon_{ijk}\cdot(y_k-y_l)$ $-\frac{1}{12}(3x_i+x_l)$	$\epsilon_{ijk}\cdot d$	$\sum_m \epsilon_{ijm}\cdot y_m$ $+\frac{1}{4}(x_i+3x_j)$	$\kappa_{ij}+\kappa_{ji}+\kappa_{kl}+\kappa_{lk}=0;\ \ \sum_m \alpha_m=0;$ $\sum_m \kappa_{im}=0;\ \ 2\alpha_l=\sum_m \kappa_{ml}$

1. Representations $S^p(\mathbb{C}^n)$

The p-fold symmetric power $S^p(\mathbb{C}^n)$ can be considered as the space of homogeneous polynomials of degree p endowed with the natural S_n action $(\sigma a)(u_1,...,u_n) := a(u_{\sigma(1)},...,u_{\sigma(n)})$, where the variables are denoted by $u=(u_1,...,u_n)$. Polynomials from $S^p(\mathbb{C}^n)$ which are constant on the lines $u+t\cdot\mathbb{1}$, $t\in\mathbb{C}$, form the subspace $S^p(\mathbb{C}_0^n)$. Thus $S^p(\mathbb{C}_0^n)$ is the kernel of the S_n-intertwining operator $J_- := \Sigma_i(\partial/\partial u_i) : S^p(\mathbb{C}^n)\to S^{p-1}(\mathbb{C}^n)$. Obviously, each $a\in S^p(\mathbb{C}^n)$ has the unique decomposition:

$$a(u)=\sum_{0\leqslant q\leqslant p} z^{p-q}a_q(u),\quad a_q\in S^p(\mathbb{C}_0^n), \tag{A1}$$

where $z := u_1+\cdots+u_n$. Therefore, we have the isomorphism

$$S^p(\mathbb{C}^n)\simeq\sum_{0\leqslant q\leqslant p} S^q(\mathbb{C}_0^n)\,. \tag{A2}$$

On the other hand, $S^p(\mathbb{C}^n)$ is isomorphic to $\mathbb{C}^{\mathscr{L}^p}$ since every polynomial can be labeled by its coefficients: $a(u)=\Sigma_\lambda a_\lambda u^\lambda$, $(a_\lambda)\in\mathbb{C}^{\mathscr{L}^p}$, $u^\lambda := u_1^{\lambda_1}\cdot\ \cdots\ \cdot u_n^{\lambda_n}$. So, we may write $a=(a_\lambda)$. In this realization, the S_n action is given by $(\sigma a)_\lambda = a_{\sigma^{-1}(\lambda)}$, whereas the S_n intertwiners $J_- := \Sigma_i\partial/\partial u_i$ and $J_+ := \Sigma_i u_i(u_i(\partial/\partial u_i)+1)$ are expressed by the formulas (2.11).

For small $p\geqslant 0$, it sometimes can be more convenient to use the symmetric coefficients

$$a(u)=(1/p!)\Sigma_{i_1\cdots i_p}a_{i_1\cdots i_p}u_{i_1}\cdot\ \cdots\ \cdot u_{i_p}.$$

They are connected with the a_λ's by the relations

$$a_{i_1\cdots i_p}=\lambda!\,a_\lambda\quad\text{if } \lambda=\epsilon_{i_1}+\cdots+\epsilon_{i_p}\,. \tag{A3}$$

For the symmetric coefficients, we have $(\sigma a)_{i_1\cdots i_p} = a_{\sigma^{-1}(i_1)\cdots\sigma^{-1}(i_p)}$ and

$$(J_-a)_{i_1\cdots i_{p-1}}=\sum_i a_{ii_1\cdots i_{p-1}}\,, \tag{A4}$$

whereas the expression for $b=J_+a$ is slightly more complicated, e.g.,

$$b_{ij}=a_i+a_j,\quad b_{ii}=4a_i\quad\text{for } p=1,\quad i\neq j\,;$$
$$b_{ijk}=a_{jk}+a_{ik}+a_{ij},\quad b_{iij}=4a_{ij}+a_{ii},$$
$$b_{iii}=9a_{ii}\quad\text{for } p=2,\ ijk\neq\,.$$

Remark 4: The decomposition given by (A1) differs from the one given by the formula (2.13) since the operator $(1/k!)J_+^k$ does not coincide with the multiplication by z^k. ■

Lemma 1: Let $c(\cdot,\cdot)$ denote the intertwining number of two representations. Then

(a) $c(S^p(\mathbb{C}_0^n),\mathbb{C})=\mathrm{NS}(2l_2+\cdots+nl_n=p)$,

(b) $c(S^p(\mathbb{C}_0^n),\overline{\mathbb{C}})=\mathrm{NS}(2l_2+\cdots+nl_n=p-\binom{n}{2})$,

(c) $c(S^p(\mathbb{C}_0^n),\mathbb{C}_0^n)=\mathrm{NS}(l_1+2l_2+\cdots+(n-1)l_{n-1}=p\,|\,l_1\neq 0\ (\mathrm{mod}\ n))$,

(d) $c(S^p(\mathbb{C}_0^n),\overline{\mathbb{C}}_0^n)=\mathrm{NS}(l_1+2l_2+\cdots+(n-1)l_{n-1}=p+\binom{n-1}{2}\,|\,l_1+1\neq 0\ (\mathrm{mod}\ n))$,

where $\mathrm{NS}(\mathscr{E})$ [resp. $\mathrm{NS}(\mathscr{E}|\&)$] denote the number of solutions in $l_i\in\mathbb{Z}_+$ of the equation $\mathscr{E}$ (resp. with the additional condition & for the l_i's). ■

For $p=2$ and 3, the decomposition of $S^p(\mathbb{C}_0^n)$ into irreducible components is described in Lemma 4 given below.

2. Representations $\wedge^p(\mathbb{C}^n)$

The element $\sigma\in S_n$ acts on p-fold exterior power $\wedge^p(\mathbb{C}^n)$ permuting the basic vectors: $\sigma(\epsilon_{i_1}\wedge\cdots\wedge\epsilon_{i_p}) = \epsilon_{\sigma(i_1)}\wedge\cdots\wedge\epsilon_{\sigma(i_p)}$. Obviously, the operator $J_-:\wedge^p(\mathbb{C}^n)\to\wedge^{p-1}(\mathbb{C}^n)$, where $J_-\Omega := (\epsilon^1+\cdots+\epsilon^n)\,\lrcorner\,\Omega$, $\Omega\in\wedge^p(\mathbb{C}^n)$ intertwines the S_n action and, moreover, we have $\ker J_- = \wedge^p(\mathbb{C}_0^n)$. Since every $\Omega\in\wedge^p(\mathbb{C}^n)$ has the unique decomposition $\Omega=\Omega_0+(\epsilon_1+\cdots+\epsilon_n)\wedge\Omega_1$, $\Omega_k\in\wedge^{p-k}(\mathbb{C}_0^n)$, we obtain an isomorphism

$$\wedge^p(\mathbb{C}^n)\simeq\wedge^p(\mathbb{C}_0^n)\dotplus\wedge^{p-1}(\mathbb{C}_0^n)\,. \tag{A5}$$

If we describe an element

$$\Omega=\frac{1}{p!}\sum_{i_1\cdots i_p}a_{i_1\cdots i_p}\epsilon_{i_1}\wedge\cdots\wedge\epsilon_{i_p}\in\wedge^p(\mathbb{C}^n)$$

by its skew-symmetric coefficients $a_{i_1\cdots i_p}$, then S_n action is given by $(\sigma a)_{i_1\cdots i_p} = a_{\sigma^{-1}(i_1)\cdots\sigma^{-1}(i_p)}$, $\sigma\in S_n$, whereas

$$(J_-a)_{i_1\cdots i_{p-1}}=\sum_i a_{ii_1\cdots i_p}\,. \tag{A6}$$

Lemma 2: (a) For $0\leqslant p\leqslant n-1$, the representations $\wedge^p(\mathbb{C}_0^n)$ are irreducible and mutually inequivalent.

(b) $\overline{\mathbb{C}}\otimes\wedge^p(\mathbb{C}_0^n)\simeq\wedge^{n-1-p}(\mathbb{C}_0^n)$. ■

3. Representations $\mathbb{C}^{n,p}$

Let the group S_n act in the space $\mathbb{C}^{n,p}:=\mathbb{C}^{\pounds_p}$ by the formula $(\sigma a)_I:=a_{\sigma^{-1}(I)}$, $a=(a_I)\in\mathbb{C}^{\pounds_p}$. Clearly, the one-to-one correspondence, among p-element subsets in $\overline{1,n}$ and their complements, defines the S_n isomorphism between $\mathbb{C}^{n,p}$ and $\mathbb{C}^{n,n-p}$. For the S_n intertwiners $J_\pm$ given by (2.24), the sl(2,C)-commutation relations (2.20) are satisfied if the operator H in $\Sigma_{0<p<n}\mathbb{C}^{n,p}$ is defined by $Ha=(p-n/2)a$, $a\in\mathbb{C}^{n,p}$. Thereby, the kernel $\ker(J_-:\mathbb{C}^{n,p}\to\mathbb{C}^{n,p-1})=:\mathbb{C}_0^{n,p}$ is zero for $p>n/2$, moreover, the intertwiners J^r_+ are injective on $\mathbb{C}_0^{n,p}$ for $0\leqslant r\leqslant n-2p$, and $J^r_+\ \mathbb{C}_0^{n,p}=0$ for $r>n-2p$ [compare the proof of Proposition 5(a)]. Thus we have the decomposition [cf. (2.28)]

$$\mathbb{C}^{n,p}=\sum_{0\leqslant q\leqslant\min(p,n-p)}J_+^{p-q}\,\mathbb{C}_0^{n,q}\simeq\sum_{0\leqslant q\leqslant\min(p,n-p)}\mathbb{C}_0^{n,q}. \tag{A7}$$

Lemma 3: (a) For $0\leqslant p\leqslant n/2$, the representations $\mathbb{C}_0^{n,p}$ are irreducible and mutually inequivalent,

(b) $\dim\mathbb{C}_0^{n,p}=\binom{n}{p}-\binom{n}{p-1}$. ■

Remark 5: It can be proved that the considered irreducible S_n representations have the following signatures:

$$\operatorname{sgn}\wedge^p(\mathbb{C}_0^n)=(n-p,1^p),\quad \operatorname{sgn}\mathbb{C}_0^{n,p}=(n-p,p). \tag{A8}$$

The reader can check that these formulas are in agreement with the Weyl formula describing the dimension of an irreducible S_n representation of the signature $(\alpha_1,...,\alpha_n)$, see, e.g., Ref. 12,

$$N_\alpha=n!\prod_{i<j}\frac{(l_i-l_j)}{l_1!\cdots l_n!},\quad l_i:=\alpha_i+n-i,\quad i\in\overline{1,n}. \ ■$$

The natural embedding $\pounds_p\hookrightarrow\mathscr{L}_p$, $I\mapsto\lambda_I$, induces the S_n intertwiners $\iota_p:\mathbb{C}^{n,p}\hookrightarrow S^p(\mathbb{C}^n)$ and π_p: $S^p(\mathbb{C}^n)\to\mathbb{C}^{n,p}$, where

$$(\iota_pa)_\lambda:=\begin{cases}a_I, & \text{if } \lambda=\lambda_I \text{ for certain } I,\\ 0, & \text{otherwise,}\end{cases}\qquad (\pi_pb)_I:=b_{\lambda_I}.$$

Obviously, $\pi_p\circ\iota_p=\mathrm{id}$ and it is easy to check that $J_-\circ\iota_p=\iota_{p-1}\circ J_-$. Therefore we have also the natural embedding

$$\iota_p:\ \mathbb{C}_0^{n,p}\hookrightarrow S^p(\mathbb{C}_0^n). \tag{A9}$$

Note, however, that ι_p does not intertwine the action of the operator J_+; only a weaker condition $\pi_{p+1}\circ J_+\circ\iota_p=J_+$ is fulfilled.

Lemma 4:

(a) $S^2(\mathbb{C}_0^n)\simeq\mathbb{C}+\mathbb{C}_0^n+\mathbb{C}_0^{n,2}$,

(b) $S^3(\mathbb{C}_0^n)$

$$\simeq\begin{cases}\mathbb{C}+2\mathbb{C}_0^n+\mathbb{C}_0^{n,2}+\mathbb{C}_0^{n,3}+\wedge^2(\mathbb{C}_0^n), & \text{if } n\geqslant5,\\ \mathbb{C}+2\mathbb{C}_0^4+\overline{\mathbb{C}}_0^4, & \text{if } n=4,\\ \mathbb{C}+2\mathbb{C}_0^3+\overline{\mathbb{C}}, & \text{if } n=3.\end{cases}$$ ■

APPENDIX B: PROOFS OF LEMMAS

Lemma 1: (a) Due to the decomposition (A1), we have the isomorphism

$$S^p(\mathbb{C}_0^n)\simeq S^p(\mathbb{C}^n)/zS^{p-1}(\mathbb{C}^n). \tag{B1}$$

Any S_n-invariant vector from the right side space of (B1) can be uniquely represented by the polynomial in variables $z_2,...,z_n$, which are basic S_n-invariant polynomials, i.e., $z_k:=\Sigma_{i_1<\cdots<i_k}u_{i_1}\cdots u_{i_k}$.

(b) Any skew-symmetric polynomial from $S^p(\mathbb{C}^n)$ is a product of the Vandermond polynomial $V(u):=\Pi_{i<j}(u_i-u_j)$ and a symmetric polynomial of the degree $p-\binom{n}{2}$. Thus the assertion is implied by part (a).

(c) and (d) The proofs are entirely similar. Let us prove (d). Since $\overline{\mathbb{C}}^n=\overline{\mathbb{C}}_0^n+\overline{\mathbb{C}}$, it is sufficient to compute $c(S^p(\mathbb{C}_0^n),\overline{\mathbb{C}}^n)$, and next subtract $c(S_p(\mathbb{C}_0^n),\overline{\mathbb{C}})$ given by (b). Any intertwining mapping T: $\overline{\mathbb{C}}^n\to S^p(\mathbb{C}^n)$ has the form $T\eta=\Sigma_i\eta^i\varphi_i$, where $\eta\in\overline{\mathbb{C}}^n$ and $\varphi_1,...,\varphi_n\in S^p(\mathbb{C}^n)$ are such polynomials that $\varphi_{\sigma(i)}(u)=\operatorname{sgn}\sigma\,\varphi_i(u_{\sigma(1)},...,u_{\sigma(n)})$. It means that the system of the polynomials $\varphi_1,...,\varphi_n$ is uniquely determined by the polynomial φ_1 which is skew symmetric in the variables $u_2,...,u_n$. But any such φ_1 is, modulo $z=u_1+\cdots+u_n$, a linear combination of polynomials $\hat V(u)u_1^{l_1}\hat z_2^{l_2}\cdots\hat z_{n-1}^{l_{n-1}}$, where $\hat V(u):=\Pi_{2<i<j}(u_i-u_j)$, $\hat z_k:=\Sigma_{2<i_1<\cdots<i_k}u_{i_1}\cdots u_{i_k}$ and $l_1+2l_2+\cdots+(n-1)l_{n-1}+\binom{n-1}{2}=\deg\varphi_1=p$. In consequence

$$c(S^p(\mathbb{C}_0^n),\overline{\mathbb{C}}^n)=\mathrm{NS}(l_1+2l_2+\cdots+(n-1)l_{n-1}=p-\tbinom{n-1}{2}). \tag{B2}$$

Substituting in the equation appearing in the formula (B2) the number l_1+1 by $n(l_n+1)$, we see that the solutions, for which $l_1+1=0\ (\mathrm{mod}\ n)$, are in one-to-one correspondence with the solutions of the equation $2l_2+\cdots+nl_n=p-\binom{n}{2}$ which provides the intertwining number $c(S^p(\mathbb{C}_0^n),\overline{\mathbb{C}})$. ■

Lemma 2: (b) The exterior product provides a two-linear mapping

$\wedge^p(\mathbb{C}_0^n)\times\wedge^{n-p-1}(\mathbb{C}_0^n)\to\wedge^{n-1}(\mathbb{C}_0^n)$, $\dim\wedge^{n-1}(\mathbb{C}_0^n)=1$,

such that $(\sigma\Omega^p)\times(\sigma\Omega^{n-p-1})=\operatorname{sgn}\sigma(\Omega^p\wedge\Omega^{n-p-1})$, $\sigma\in S_n$, $\Omega^k\in\wedge^k(\mathbb{C}_o^n)$. It implies that $\wedge^{n-p-1}(\mathbb{C}_0^n)$ is contragredient to $\overline{\mathbb{C}}\otimes\wedge^p(\mathbb{C}_0^n)$. But, as we know, the contragredient S_n representations are equivalent.

(a) We shall use the irreducibility criterion proved in Ref. 1 [Lemma 5(ii)]. The representation $\wedge^p(\mathbb{C}^n)$ can be embedded in the space $\mathbb{C}^X$, where $X:=\{(i_1,...,i_p)\in(\overline{1,n})^p|\ i_1,...,i_p\neq\}$ is the S_n orbit of the element $(1,2,...,p)$. Clearly,

$W_0 := \{\sigma\in S_n \,|\, \sigma(i) = i \text{ for } i\in \overline{1,p}\}$ is the isotropy subgroup of $(1,2,...,p)$. One can easily see that a p-form $\Omega\in\wedge^p(\mathbb{C}^n)$ is W_0 invariant iff it can be expressed by $\epsilon_1,...,\epsilon_p,\epsilon_{p+1}+\cdots+\epsilon_n$, i.e., iff

$$\Omega = c_0\epsilon_1\wedge\cdots\wedge\epsilon_p + \left(\sum_{1\leqslant i\leqslant p} c_i\epsilon^i \lrcorner\, \epsilon_1\wedge\cdots\wedge\epsilon_p\right)$$
$$\wedge(\epsilon_{p+1}+\cdots+\epsilon_n),$$

where $c_0,c_1,...,c_p\in\mathbb{C}$. Thus $\Omega\in\wedge^p(\mathbb{C}_0^n)$ is W_0 invariant iff

$$0 = (\epsilon^1+\cdots+\epsilon^n)\lrcorner\,\Omega$$
$$= \sum_{1\leqslant i\leqslant p}[c_0 - (-)^p(n-p)c_i]\epsilon^i\lrcorner\,\epsilon_1\wedge\cdots\wedge\epsilon_p,$$

i.e., iff

$$c_i = (-)^p/(n-p)c_0.$$

So, the only one W_0 invariant ray in $\wedge^p(\mathbb{C}_0^n)$ is spanned by the vector

$$\sum_{1\leqslant i\leqslant n}\epsilon^i\lrcorner\,\epsilon_1\wedge\cdots\wedge\epsilon_p\wedge(\epsilon_{p+1}+\cdots+\epsilon_n).$$

Now, comparing the dimensions of the representations $\wedge^p(\mathbb{C}_0^n)$, $0\leqslant p\leqslant n-1$, we see that only the representations $\wedge^p(\mathbb{C}_0^n)$ and $\wedge^{n-1-p}(\mathbb{C}_0^n)$ could be equivalent. If so, then according to part (b), $\wedge^p(\mathbb{C}_0^n)$ is equivalent to $\overline{\mathbb{C}}\otimes\wedge^p(\mathbb{C}_0^n)$, i.e., the character χ_0^p of $\wedge^p(\mathbb{C}_0^n)$ fulfills: $\chi_0^p(\sigma)=0$ for any odd $\sigma\in S_n$. But from the formula (A5), it follows that

$$\chi_0^p = \sum_{0\leqslant q\leqslant p}(-)^{p-q}\chi^q, \tag{B3}$$

where χ^q is the character of $\wedge^q(\mathbb{C}^n)$. One can easily check that

$$\chi^q(\sigma) = \sum_{i_1<\cdots<i_q}\operatorname{sgn}\begin{pmatrix} i_1 & \cdots & i_q\\ \sigma(i_1) & \cdots & \sigma(i_q)\end{pmatrix} \tag{B4}$$

and hence, $\chi^q(1,2)) = \binom{n-2}{q} - \binom{n-2}{q-2}$. Thus the formula (B3) implies that $\chi_0^p((1,2)) = \binom{n-1}{p}(n-1-2p)/n-1 = 0$ iff $p = n-1-p$. ■

Lemma 3: (a) The group S_n permutes the natural basic vectors in $\mathbb{C}^{\pounds_p}$, which are labeled by subsets $I\in\pounds_p$. Therefore, the character $\overset{p}{\chi}$ of the representation $\mathbb{C}^{n,p}$ is given by

$$\overset{p}{\chi}(\sigma) = (\text{number of } \sigma\text{-invariant subsets } I\in\pounds_p)$$

$$= \sum_{I\in\pounds_p}\delta^I_{\sigma(I)},\quad \sigma\in S_n,$$

where δ is the Kronecker symbol. Hence, we obtain

$$\left(\overset{p}{\chi}\middle|\overset{p}{\chi}\right) = \frac{1}{n!}\sum_{\sigma\in S_n}\sum_{I\in\pounds_p}\sum_{J\in\pounds_p}\delta^I_{\sigma(I)}\,\delta^J_{\sigma(J)}$$
$$= \frac{1}{n!}\sum_{IJ}|I\setminus J|!|J\setminus I|!|I\cap J|!|\overline{I\cup J}|!.$$

For a pair I,J such that $|I\cap J| = k$, we have $|I\setminus J| = |J\setminus I| = (p-k)$ and $|\overline{I\cup J}| = (n-2p+k)$. Moreover, if $0\leqslant p\leqslant n/2$, then k can assume the values $0,1,...,p$. Taking into account, for a given k, the number of different pairs I, J, we get

$$\left(\overset{p}{\chi}\middle|\overset{p}{\chi}\right) = \frac{1}{n!}\sum_{k=0}^{p}(p-k)!^2k!(n-2p+k)!$$
$$\times\binom{n}{k}\binom{n-k}{p-k}\binom{n-p}{p-k}$$
$$= \frac{1}{n!}\sum_{k=0}^{p} n! = p+1.$$

On the other hand, the formula (A7) implies for $0\leqslant p\leqslant n/2$ that $\overset{p}{\chi} = \Sigma_{k=0}^p\overset{k}{\chi}_0$, where $\overset{p}{\chi}_0$ is the character of $\mathbb{C}_0^{n,p}$. In consequence, we obtain

$$(p+1) = \left(\overset{p}{\chi}\middle|\overset{p}{\chi}\right) = \sum_{0\leqslant kl\leqslant p}\left(\overset{k}{\chi}_0\middle|\overset{l}{\chi}_0\right).$$

Since $\left(\overset{k}{\chi}_0\middle|\overset{l}{\chi}_0\right)\geqslant\delta_{kl}$, the last equation means that $\left(\overset{k}{\chi}_0\middle|\overset{k}{\chi}_0\right) = 1$, $\left(\overset{k}{\chi}_0\middle|\overset{l}{\chi}_0\right) = 0$, $k\neq l$.

(b) It follows immediately from (A7) and the formula $\dim\mathbb{C}^{n,p} = \binom{n}{p}$. ■

Remark 6: The irreducibility of $\mathbb{C}_0^{n,p}$, $p\leqslant n/2$, also can be proved by checking that for the isotropy subgroup $W_0\subset S_n$ of the element $\{1,2,...,p\}\in\pounds_p$, the space consisting of W_0-invariant elements in $\mathbb{C}_0^{n,p}$ is one dimensional. On the other hand the irreducibility of $\wedge^p(\mathbb{C}_0^n)$ can be deduced from the formula (A5) by showing that the character χ^q, cf. (B4), satisfies $(\chi^q|\chi^q) = 2$. ■

Lemma 4: For simplicity of notation, let us treat $\mathbb{C}^{n,p}$ as the subspace in $S^p(\mathbb{C}^n)\simeq\mathbb{C}^{\mathscr{L}_p}$ (cf. Appendix A 3 and A1).

(a) For $p=2$, the isomorphism between $\mathbb{C}^{n,2}$ and $S^2(\mathbb{C}_0^n)$ is given by

$$a_{ij} = b_{ij},\quad i\neq j;\quad a_{ii} = -(J\lrcorner b)_i;$$
$$b_{ii} = 0;\quad a\in S^2(\mathbb{C}_0^n),\quad b\in\mathbb{C}^{n,2}.$$

Thus, formula (A7) implies the assertion.

(b) It is sufficient to show that $S^3(\mathbb{C}_0^n)\simeq\mathbb{C}^{n,3}\dotplus\wedge^2(\mathbb{C}^n)$ and next to apply (A5) and (A7). To this end, let us check that the formulas

$$a_{ijk} = b_{ijk};\quad a_{iij} = -\tfrac{1}{2}(J\lrcorner b)_{ij} + c_{ij};$$
$$a_{iii} = \tfrac{1}{2}(J^2\lrcorner b)_i + (J\lrcorner c);\quad ijk\neq, \tag{B5}$$

define a bijective correspondence between elements $a\in S^3(\mathbb{C}_0^n)$ and pairs $(b,c)\in\mathbb{C}^{n,3}\times\wedge^2(\mathbb{C}^n)$. Indeed, each $a\in S^3(\mathbb{C}_0^n)$ has at most one decomposition (B5) since the formulas describing it imply that

$$b_{ijk} = \begin{cases} a_{ijk}, & \text{if } ijk\neq,\\ 0, & \text{otherwise};\end{cases}$$
$$c_{ij} = \tfrac{1}{2}(a_{iij} - a_{iji}). \tag{B6}$$

Moreover, if for a given a we define the b and c by means of (B6) then, according to (A4), we obtain

$$(J\_a)_{ij} = 2a_{iij} - 2c_{ij} + (J\_b)_{ij}, \quad i \neq j;$$

$$(J\_a)_{ii} = a_{iii} + \sum_{j\neq i} a_{jii} .$$

It shows that the formulas (B5) are fulfilled iff $J\_a = 0$, i.e., iff $a \in S^3(\mathbb{C}_0^n)$. ■

[1]G. Cieciura and I. Szczyrba, J. Math. Phys. **28**, 2115 (1987).

[2]E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics and Atomic Spectra* (Academic, New York, 1959); L. Michel, "*Applications of Group Theory to Quantum Physics, Algebraic Aspects,*" *Lecture Notes in Physics*, Vol. 6 (Springer, Berlin, 1970).

[3]J. A. Castilho Alcaras, L. C. Biedenharn, K. T. Hecht, and G. Neely, Ann. Phys. (NY) **60**, 85 (1970).

[4]M. K. Gaillard, B. W. Lee, and J. L. Rosner, Rev. Mod. Phys. **47**, 277 (1975).

[5]J. D. Louck and L. C. Biedenharn, J. Math. Phys. **11**, 2368 (1970); S. Okubo, *ibid.* **16**, 528 (1975).

[6]I. Szczyrba, Lett. Math. Phys. **4**, 249 (1980).

[7]C. Itzykson, M. Nauenberg, Rev. Mod. Phys. **38**, 95 (1966); A. O. Barut, R. Raczka, *Theory of Group Representations and Applications* (PWN, Warsaw, 1980).

[8]L. C. Biedenharn, J. D. Louck, E. Chacón, and M. Ciftan, J. Math. Phys. **13**, 1957, 1985 (1972); J. D. Louck and L. C. Biedenharn, *ibid.* **14**, 1336 (1973); I. Szczyrba, *ibid.* **22**, 1289 (1981).

[9]D. P. Zhelobenko, "*Compact Lie Groups and their Representations, Translations of Mathematical Monographs*, Vol. 40 (Am. Math. Soc., Providence, RI, 1973).

[10]N. Bourbaki, *Élements de Mathématique, Groupes et Algèbres de Lie* (Hermann, Paris, 1975), Chaps. VII and VIII.

[11]H. Weyl, *The Classical Groups, Their Invariants and Representations* (Princeton U. P., Princeton, NJ, 1946).

[12]G. D. James and A. Kerber, *Representation Theory of the Symmetric Groups, Encyclopedia of Mathematics and its Applications*, Vol. 16 (Addison–Wesley, Reading, MA, 1981).

Möller operators in classical relativistic two-particle scattering[a)]

S. De Bièvre
Department of Mathematics, University of Toronto, Toronto, Ontario, M5S 1A1, Canada

(Received 23 January 1987; accepted for publication 29 April 1987)

The scattering problem is formulated in a geometric language adapted to the description of classical relativistic two-body systems. Within the framework of manifestly covariant relativistic particle mechanics, precise, easily verifiable conditions with a clear physical interpretation are given on the interparticle interaction for the Möller operators to exist. The Möller operators are used to define the notion of an asymptotically free presymplectic structure on the evolution space and, using the results obtained, the existence and uniqueness of such a structure is discussed.

I. INTRODUCTION

Since the advent of the no-interaction theorem,[1] a variety of approaches to classical relativistic particle mechanics has been proposed.[2,3] They all use, either implicitly or explicitly, the notions of evolution space and of space of motions, advocated in Ref. 4 for the description of dynamical systems. In Sec. II we shall show how this general geometric framework underlies the different approaches and explicitly identify the ingredients of the general structure in the case of manifestly covariant particle mechanics.[3] It will become clear then that it is generally no longer possible to view the dynamics in those models as generated by the flow of one vector field on a fixed phase space, as in the case of nonrelativistic mechanics, for example. This, in turn, implies that the description of scattering has to be reformulated in the new framework.

We address that problem in Sec. III, where we argue that the asymptotic comparison of the free and the interacting dynamics is most naturally done in the evolution space. Using this point of view, we prove the central result of the paper (Theorem 3.1), which—when applied to manifestly covariant particle mechanics—gives precise and easily verifiable conditions on the dynamics for the Möller operators to exist.

Conditions similar to ours have been used previously[5] to establish, to lowest order in a perturbative expansion, the existence and uniqueness of a symplectic structure "in the past" and "in the future." In Sec. IV, we show how the geometric framework presented in Sec. II, together with the scattering theory of Sec. III lead immediately to a simple definition of these notions. We show moreover that the existence of the Möller operators directly implies the existence of the asymptotic symplectic structures. Their uniqueness becomes, in this framework, a trivial result. We conclude with some comments on asymptotic completeness and its link with the symplectic structure of the theory.

[a)] Part of this work is based on the Ph. D. thesis of the author, submitted to the Department of Physics and Astronomy of the University of Rochester.

II. A GENERAL FRAMEWORK FOR RELATIVISTIC PARTICLE MECHANICS

Let M denote Minkowski space-time. Let $\mathscr{P}$ be the Poincaré group. A particle is a one-dimensional, timelike, and connected submanifold $\gamma_{(k)}$ of M. A *$\mathscr{P}$-invariant two-particle system* is a collection Γ of couples $\gamma = (\gamma_{(1)},\gamma_{(2)})$ of particles which is stable under the action Φ of $\mathscr{P}$ on M; i.e., if $\gamma\in\Gamma$, then $g[\gamma]\in\Gamma$, $\forall g\in\mathscr{P}$, where $g[\gamma]$ denotes the transform of γ under g, implemented by the natural action Φ of $\mathscr{P}$ on M. The simplest example of a $\mathscr{P}$-invariant two-particle system is the free system Γ_0, which consists of all couples of timelike geodesics on M. Particle mechanics is concerned with the explicit construction of physically meaningful two or more particle systems in a mathematically consistent way.

Following Souriau's description of dynamical systems,[4] we claim that a model for the description of two-particle systems should have the following general structure. We define a *$\mathscr{P}$-invariant Hamiltonian two-particle system* as a triple (E,ϵ,ϕ) with the following properties.

(i) (E,ϵ) is a presymplectic manifold[6,7] fibered over two copies of M; i.e., we have surjective submersions $\pi_{(k)}: E\rightarrow M$.

(ii) The characteristic foliation of ϵ (i.e., the foliation generated by the kernel of ϵ) is reducible. We denote the corresponding quotient manifold by Σ and its symplectic form by σ and we write $\rho: E\rightarrow\Sigma$ for the natural projection of E to Σ. Here (E,ϵ) is called the *evolution space* and (Σ,σ) the *space of motions*.

(iii) In order to give a particle interpretation to this structure, one requires the following: for any point $s\in\Sigma$, $\pi_{(k)}(\rho^{-1}(s))$ is a particle for $k=1,2$. In other words, to every point in the space of motions corresponds a couple of particles, or, differently yet, every leaf of the characteristic foliation of ϵ on E projects down to a couple of world lines on M. This identifies a collection Γ of couples of particles with Σ and consequently gives Γ the structure of a differentiable manifold. The points of Σ, or, equivalently, the leaves of the foliation are called the *(generalized) trajectories*.

(iv) ϕ is an action ϕ: $\mathscr{P}\times E\rightarrow E$ of $\mathscr{P}$ on E and satisfies the following.

(a) $\phi_g{}^*\epsilon=\epsilon$, $\forall g\in\mathscr{P}$ (Ref. 6), so that the characteris-

 0022-2488/87/092142-07$02.50

tic foliation of ϵ is equivariant under the action of $\mathscr{P}$. In other words, the leaves of the characteristic foliation of ϵ are mapped into one another by this action. Consequently, $\mathscr{P}$ will also act symplectically on Σ as is readily verified.[4,6]

(b) The following diagram is commutative:

$$\begin{array}{ccc} E & \xrightarrow{\phi_g} & E \\ \pi_{(k)} \downarrow & \Phi_g & \downarrow \pi_{(k)} \\ M & \longrightarrow & M \end{array}.$$

This condition guarantees that the action of $\mathscr{P}$ on Σ is in agreement with the particle interpretation of the points of Σ given above: it is the global version of what is often referred to as the "worldline conditions"[2] and it assures that Γ is a $\mathscr{P}$-invariant two-particle system.

An analysis of the various approaches to relativistic particle mechanics[2] makes it clear that they can all be fit into the above general framework.[8] We work this out explicitly for manifestly covariant particle mechanics. For a more extensive discussion of this formalism, we refer to the literature.[3] We denote by $TM^2 = T(M\times M)$ the tangent bundle to the Cartesian product of Minkowski space-time M with itself. Points in TM^2 are written as $(x,v)\equiv(x^\mu_{(1)}, x^\mu_{(2)}, v^\mu_{(1)}, v^\mu_{(2)})$, where $\mu = 0,1,2,3$. In the following, we are only interested in the open subset of TM^2 where v_1 and v_2 are timelike; we will not introduce a separate notation for it, though. The Poincaré group $\mathscr{P}$ acts naturally on TM^2 by the lift of its action on $M\times M$; we write $(\xi)_{TM^2}$ for the generator of this action corresponding to ξ in the Lie algebra of P. A *manifestly covariant predictive Poincaré-invariant two-particle system* is determined by giving two complete vector fields $X_{(1)}$ and $X_{(2)}$ on TM^2, satisfying the following requirements:

$$\text{(i)}\ X_{(1)} = v^\mu_{(1)} \frac{\partial}{\partial x^\mu_{(1)}} + a^\mu_{(1)} \frac{\partial}{\partial v^\mu_{(1)}}, \tag{2.1}$$

$$X_{(2)} = v^\mu_{(2)} \frac{\partial}{\partial x^\mu_{(2)}} + a^\mu_{(2)} \frac{\partial}{\partial v^\mu_{(2)}}, \tag{2.2}$$

$$\text{(ii)}\ [X_{(k)},(\xi)_{TM^2}] = 0, \quad k = 1,2, \quad \forall\xi\in L(\mathscr{P}), \tag{2.3}$$

$$\text{(iii)}\ [X_{(1)},X_{(2)}] = 0, \tag{2.4}$$

$$\text{(iv)}\ a^\mu_{(1)} v_{(1)\mu} = 0 = a^\mu_{(2)} v_{(2)\mu}. \tag{2.5}$$

Here, the $a^\mu_{(k)}$ are functions on $T(M\times M)$; they specify the dynamics as follows. The world lines are obtained by integrating the differential equations of motion

$$\frac{d^2 x^\mu_{(1)}}{d\tau_{(1)}} = a^\mu_{(1)}(x,v), \tag{2.6}$$

$$\frac{d^2 x^\mu_{(2)}}{d\tau_{(2)}} = a^\mu_{(2)}(x,v), \tag{2.7}$$

for some initial conditions $(x,v)\in T(M\times M)$. The solution of (2.6) gives the world line of particle 1 and the solution of (2.7) gives the world line of particle 2. Now let $(x'_{(1)}, v'_{(1)})$ be a point along the first world line and similarly for $(x'_{(2)}, v'_{(2)})$. We can then use $(x',v') = (x'_{(1)}, v'_{(1)}, x'_{(2)}, v'_{(2)})$ as a new initial condition in (2.6) and (2.7). This will again give two world lines; condition (iii) guarantees that they are identical to the original ones. In view of (iii), the vector fields $X_{(1)}$ and $X_{(2)}$ determine a two-dimensional foliation of TM^2. We want to interpret the leaves of this foliation as the generalized trajectories of the system. However, the quotient of TM^2 by this foliation is 14 dimensional. Since we want a two-particle system to have a 12-dimensional space of motions, we have to somehow eliminate two more degrees of freedom from TM^2 in order to obtain the evolution space of the system. This is made possible by condition (iv); indeed, in view of (iv), the 14-dimensional hypersurface E in TM^2 given by

$$v_{(1)}{}^2 = -m_{(1)}{}^2, \tag{2.8}$$

$$v_{(2)}{}^2 = -m_{(2)}{}^2, \tag{2.9}$$

for some choice of $m_{(1)}, m_{(2)} > 0$ and with $v_{(1)}$ and $v_{(2)}$ future pointing, is invariant under the flow of $X_{(1)}$ and $X_{(2)}$. Consequently E can be used as evolution space of the system. The foliation of E, induced by $X_{(1)}$ and $X_{(2)}$ is reducible; this follows from the completeness of $X_{(1)}$ and $X_{(2)}$ and the observation that the surface $x_{(1)}{}^0 = 0 = x_{(2)}{}^0$ in E intersects every leaf in precisely one point. The invariance of the theory under the Poincaré group is assured by (2.3).

In order to make the above manifestly covariant predictive Poincaré invariant two-particle system into a Poincaré invariant Hamiltonian two-particle system as defined earlier, one proceeds as follows. One looks for a presymplectic form ϵ on E with the following properties:

$$\text{(v)}\ L_{(\xi)_{TM^2}}\epsilon = 0, \quad \forall\xi\in L(\mathscr{P}), \tag{2.10}$$

$$\text{(vi)}\ i(X_{(k)})\epsilon = 0 \quad \text{for } k = 1,2. \tag{2.11}$$

In order to construct a manifestly covariant two-particle system, one has to explicitly construct the functions $a^\mu_{(k)}$ in (2.1) and (2.2). This has been done perturbatively for the electromagnetic and scalar interactions.[9] One is then still faced with the problem of finding a suitable ϵ. The presymplectic form ϵ is in general not uniquely determined by conditions (v) and (vi)[5,10]; so the Hamiltonization of the manifestly covariant two-particle system given by (i)–(iv) is not unique. We will clarify this point in Sec. IV, using our results on the existence of the Möller operators from Sec. III.

In the following section, we will need the description of the free two-particle system Γ_0 in the framework of manifestly covariant particle mechanics, which we now give. The free system is given by (2.1)–(2.9) with $a^\mu_{(k)} = 0$ for $\mu = 0,1,2,3$ and $k = 1,2$. Let ϵ_0 be the restriction to E of the symplectic form ω_0 on $T(M\times M)$,

$$\omega_0 = dx^\mu_{(1)} \wedge dv_{(1)\mu} + dx^\mu_{(2)} \wedge dv_{(2)\mu}. \tag{2.12}$$

Then

$$X^0_{(k)} = v^\mu_{(k)} \frac{\partial}{\partial x^\mu_{(k)}} \quad (k = 1,2) \tag{2.13}$$

and we find, for $k = 1,2$,

$$i(X^0_{(k)})\omega_0 = \tfrac{1}{2}\, dH^0_{(k)} \tag{2.14}$$

with

$$H^0_{(k)} = v^2_{(k)} + m^2_{(k)}. \tag{2.15}$$

As a result, E, the evolution space, defined by (2.8) and

(2.9), becomes a presymplectic manifold when equipped with the presymplectic form ϵ_0, obtained as the restriction of ω_0 to E. Its kernel is spanned by $X^0_{(k)}$, $k = 1,2$. The corresponding space of motions Σ_0 is constructed in the usual way and we have ρ_0: $E \to \Sigma_0$. There is then a unique symplectic form σ_0 on Σ_0 satisfying $\rho_0{}^*\sigma_0 = \epsilon_0$.[4,7]

We close this section by remarking that the traditional Hamiltonian version of nonrelativistic particle mechanics also fits into the geometric scheme proposed at the beginning of this section[4,6,8] (with $\mathscr{P}$ replaced by the Galilei group and M by $\mathbb{R}\times\mathbb{R}^3$). In that case, however, it is customary and convenient to identify (Σ,σ) with $(T^*\mathbb{R}^6,\omega_0)$, where ω_0 is now the canonical symplectic form on $T^*\mathbb{R}^6$, and to view the dynamics as obtained from the flow of the time translation subgroup of the Galilei group, acting on Σ. One could paraphrase the content of the no-interaction theorem by saying that this identification can no longer be made in the relativistic case, except when describing free particles, forcing one to work directly with the more general geometric structure described in this section. As a result, we also need to take a new approach when dealing with the scattering problem, as will be explained in the next section.

III. CONDITIONS FOR THE EXISTENCE OF THE MÖLLER OPERATORS

The scattering problem can be quite generally and concisely defined as the problem of comparing, in a suitably defined asymptotic regime, two dynamical systems. One of those is relatively simple and well understood, and referred to as the free dynamics. The other is the dynamics of the interacting system. The asymptotic comparison is achieved by introducing the so-called Möller operators $\Omega_\pm$.

In the case of Poincaré-invariant two-particle systems, the question can be formulated as follows: Let Γ and Γ_0 be as in Sec. II. Then the Möller operators

$$\Omega_\pm : \Gamma_0 \to \Gamma \tag{3.1}$$

exist, provided for almost all[11,12] $(\gamma_0^{(1)},\gamma_0^{(2)}) \in \Gamma_0$, $\exists!(\gamma_\pm^{(1)},\gamma_\pm^{(2)})\in\Gamma$, such that

$$\lim_{t\to\pm\infty} |\mathbf{x}_\pm^{(k)}(t) - \mathbf{x}^{(k)} - \mathbf{v}^{(k)}t| = 0, \tag{3.2}$$

$$\lim_{t\to\pm\infty} \left|\frac{d}{dt}\mathbf{x}_\pm^{(k)}(t) - \mathbf{v}^{(k)}\right| = 0, \tag{3.3}$$

where we wrote $\gamma_0^{(k)} = \{(t,\mathbf{x}^{(k)} + \mathbf{v}^{(k)}t)\in M\,|\,t\in\mathbb{R}\}$ and $\gamma_\pm^{(k)} = \{(t,\mathbf{x}_\pm^{(k)}(t))\in M\,|\,t\in\mathbb{R}\}$ in some inertial frame of reference on M. It is clear that, if (3.2), (3.3) is true in one such frame, it is true in all of them. One then defines $\Omega_\pm(\gamma_0^{(1)},\gamma_0^{(2)}) = (\gamma_\pm^{(1)},\gamma_\pm^{(2)})\in\Gamma$. Moreover, if $\mathrm{Im}\,\Omega_+ = \mathrm{Im}\,\Omega_-$ (weak asymptotic completeness), one defines the scattering operator S: $\Gamma_0\to\Gamma_0$ by $S = \Omega_+{}^{-1}\Omega_-$. Identifying Γ with Σ and Γ_0 with Σ_0, as in Sec. II, we have

$$\Omega_\pm : \Sigma_0 \to \Sigma. \tag{3.4}$$

Specializing to manifestly covariant particle mechanics, we now wish to discuss under what conditions on "the acceleration fields" $a^\mu_{(k)}$ in (2.1) and (2.2), the existence of $\Omega_\pm$ is guaranteed. To investigate this problem, we need to first reformulate (3.2) and (3.3) directly in terms of the dynamical system as defined in (2.1)–(2.5). The result is given in (3.29). The conditions on the a^μ_k are obtained in Theorem 3.1 and commented on in remark (d) following it.

In order to see more clearly what is involved and to get an intuitive feeling for the interpretation of Theorem 3.1, we first outline a much simplified and familiar problem. We consider the motion of one nonrelativistic particle in one dimension in an outside potential field V and describe the system in terms of an evolution space as follows. Let $E = \mathbb{R}\times T^*\mathbb{R}$ and denote the points of E by (t,x,p). The dynamic vector field for the free particle is

$$X_0 = \partial_t + p\,\partial_x \tag{3.5}$$

and for the particle in the outside potential V is

$$X = \partial_t + p\,\partial_x - V'(x)\partial_p. \tag{3.6}$$

In this case, we can identify $\Sigma \approx \Sigma_0 \approx \{(t,x,p)\in E\,|\,t=0\} \approx T^*\mathbb{R}$, and recover the usual Hamiltonian formulation, with the corresponding approach to the scattering problem.[11] Instead, let us make the following transformation:

$$\varphi:\ (t,x,p)\in\mathbb{R}^3 \to (\tau,x_0,p_0)\in\mathbb{R}^3, \tag{3.7}$$

with

$$\tau = t, \tag{3.8}$$

$$x_0 = x - tp, \tag{3.9}$$

$$p_0 = p, \tag{3.10}$$

so that

$$X_0 = \partial_\tau \tag{3.11}$$

and

$$X = \partial_\tau - V'(x_0 + \tau p_0)(\partial_{p_0} - \tau\,\partial_{x_0}) \equiv \partial_\tau + Y. \tag{3.12}$$

In other words, we "straightened out" the flow of X_0. The flow lines of X_0 on E are now

$$\tau\in\mathbb{R} \to (\tau,x_0,p_0)\in E \tag{3.13}$$

and those of X can be written

$$\tau\in\mathbb{R} \to (\tau,x_0(\tau),p_0(\tau))\in E. \tag{3.14}$$

Upon noticing that we can identify Σ_0 with the surface $\tau = 0$, the question of the existence of $\Omega_\pm$ can be rephrased as follows: do there exist, for almost every $(x_0,p_0)\in\mathbb{R}^2$, flow lines of X (i.e., points in Σ),

$$\tau\in\mathbb{R} \to (\tau,x_0^\pm(\tau),p_0^\pm(\tau))\in E, \tag{3.15}$$

such that

$$\lim_{\tau\to\pm\infty} (x_0^\pm(\tau),p_0^\pm(\tau)) = (x_0,p_0). \tag{3.16}$$

In other words, by trivializing the free evolution through (3.7), we reduced the problem to a study of the large τ behavior of the vector field Y. As is clear from (3.12), this is equivalent to a study of V for large values of x, as expected in this simple case.

We now use the same evolution space approach when dealing with relativistic two-body scattering. Here E is a 14-dimensional manifold on which the two-dimensional foliation generated by the free dynamics [see (2.13)] can be computed explicitly. To do so, we first introduce a new coordinate system on $TM\times M$, defined as follows:

$$\varphi:\ (x_{(k)},v_{(k)})\in TM\times M=\mathbb{R}^{16}\rightarrow(\tau,\sigma,\mathbf{z}_{(k)},v_{(k)})\in\mathbb{R}^2\times\mathbb{R}^{14}, \tag{3.17}$$

where

$$\tau=x^0_{(1)}+x^0_{(2)},\quad \sigma=x^0_{(1)}-x^0_{(2)}, \tag{3.18}$$

$$\mathbf{z}_{(k)}=\mathbf{x}_{(k)}-(\mathbf{v}_{(k)}/v^0_{(k)})x^0_{(k)}\in\mathbb{R}^3, \tag{3.19}$$

where we wrote $x_{(k)}=(\mathbf{x}_{(k)},x^0_{(k)})$ in a fixed inertial frame of reference. One verifies that

$$\frac{\partial}{\partial\tau}=\frac{1}{v^0_{(1)}}X^0_{(1)}+\frac{1}{v^0_{(2)}}X^0_{(2)}, \tag{3.20}$$

$$\frac{\partial}{\partial\sigma}=\frac{1}{v^0_{(1)}}X^0_{(1)}-\frac{1}{v^0_{(2)}}X^0_{(2)}, \tag{3.21}$$

with $X^0_{(k)}$ as in (2.13). Hence ∂_τ and ∂_σ span the leaves L_0 of the free foliation and τ and σ can be used as coordinates on those leaves. Indeed, the latter are the two-dimensional imbedded submanifolds in E determined by [compare (3.13)]

$$L_0\colon (\tau,\sigma)\in\mathbb{R}^2\rightarrow(\tau,\sigma,\mathbf{z}_{(k)},v_{(k)})\in E. \tag{3.22}$$

In other words, the coordinate system (3.17) is particularly well adapted for the description of the free system. Moreover, the parameter τ in (3.18) has a simple interpretation; if σ is kept fixed and we let τ go to $+\infty$, then both $x^0_{(1)}$ and $x^0_{(2)}$ go to $+\infty$, and similarly with $+\infty$ replaced by $-\infty$. So τ gives us the notion of asymptotic past and future on each leaf L_0. Roughly speaking, to investigate the existence of $\Omega_\pm$, we have to compare the two-dimensional leaves L of the foliation of the interacting system with those of the free system, as $\tau\rightarrow\pm\infty$, as we now explain.

Define

$$Z_1\equiv(1/v^0_{(1)})X_{(1)}, \tag{3.23}$$

$$Z_2\equiv(1/v^0_{(2)})X_{(2)}, \tag{3.24}$$

$$Z_a\equiv Z_1+Z_2, \tag{3.25}$$

$$Z_b\equiv Z_1-Z_2, \tag{3.26}$$

with $X_{(k)}$ as in (2.1) and (2.2). Here Z_a and Z_b span the leaves of the foliation of the interacting system. It is then readily verified that Z_a and Z_b are of the form $\partial_\tau+Y_a$ and $\partial_\tau+Y_b$, where Y_a and Y_b are vector fields that do not contain any terms in ∂_τ or ∂_σ. As a result, the leaves of the interacting foliation are of the form [compare with (3.14)]

$$L\colon\ (\tau,\sigma)\in\mathbb{R}^2\rightarrow(\tau,\sigma,\mathbf{z}_{(k)}(\tau,\sigma),v_{(k)}(\tau,\sigma))\in E. \tag{3.27}$$

We can now phrase the existence question for the $\Omega_\pm$ as follows: do there exist, for almost every $(\mathbf{z}_{(k)},v_{(k)})\in\Sigma_0$, leaves L_+ and L_- of the interacting foliation [cf. (3.15)],

$$L_\pm\colon\ (\tau,\sigma)\in\mathbb{R}^2\rightarrow(\tau,\sigma,\mathbf{z}^\pm_{(k)}(\tau,\sigma),v^\pm_{(k)}(\tau,\sigma))\in TM\times M, \tag{3.28}$$

such that [cf. (3.16)]

$$\lim_{\tau\rightarrow\pm\infty}(\mathbf{z}^\pm_{(k)}(\tau,\sigma),v^\pm_{(k)}(\tau,\sigma))=(\mathbf{z}_{(k)},v_{(k)}). \tag{3.29}$$

Taking $\sigma=0$ in (3.29), and using (3.17)–(3.19), together with (3.22), we see that (3.29) is indeed the reformulation of (3.2) and (3.3) on E. Equation (3.29) will hold under appropriate conditions on Y_a and Y_b, i.e., on the four-accelerations $a^\mu_{(k)}$ as $\tau\rightarrow\pm\infty$. Intuitively, one expects that the $a^\mu_{(k)}$ need to decay sufficiently fast for large interparticle separation, i.e., for $(x_{(1)}-x_{(2)})^2\rightarrow\infty$, in analogy with the nonrelativistic case. However, in the latter case, the potential function V does not depend on any other variables, so that this limit is unambiguous; the $a^\mu_{(k)}$, on the other hand, will in general be complicated functions of the four-velocities $v^\mu_{(k)}$ also. One therefore needs to specify carefully how, i.e., along which path in E, to take the limit $(x_{(1)}-x_{(2)})^2\rightarrow\infty$. As further explained in remark (d) after Theorem 3.1, the appropriate procedure is to let $\tau\rightarrow\pm\infty$ for fixed L_0. The following theorem gives us the result we need in a slightly more general setting.

Theorem 3.1: Let

$$Y_a(\tau,\sigma,y)=\sum_{i=1,\dots,n}\alpha_i(\tau,\sigma,y)\partial_i \tag{3.30}$$

and

$$Y_b(\tau,\sigma,y)=\sum_{i=1,\dots,n}\beta_i(\tau,\sigma,y)\partial_i \tag{3.31}$$

be smooth and complete vector fields on $\mathbb{R}^2\times\mathbb{R}^{n-2}$ with $(\tau,\sigma,y)\in\mathbb{R}^2\times\mathbb{R}^{n-2}$ and $\partial_i\equiv\partial_{y^i}$. Suppose Y_a and Y_b satisfy, *for some* $\alpha>1$: for each $(\sigma,y)\in\mathbb{R}^1\times\mathbb{R}^{n-2}$, there exist $R,K,\kappa>0$ such that $\forall(\sigma',y'),(\sigma'',y'')\in B((\sigma,y),R)$, $\forall\tau>\kappa$,

$$\text{(i)}\ \ \|Y_{a,b}(\tau,\sigma',y')\|\leqslant K\tau^{-\alpha}, \tag{3.32}$$

$$\text{(ii)}\ \ \|Y_{a,b}(\tau,\sigma',y'')-Y_{a,b}(\tau,\sigma',y')\|\leqslant K\tau^{-\alpha}\|y''-y'\|. \tag{3.33}$$

Here $\|\cdot\|$ refers to the usual Euclidean norm and $B(\times,\cdot)$ is the closed ball with center "$\times$" and radius "$\cdot$." Define now

$$Z_a=\partial_\tau+Y_a, \tag{3.34}$$

$$Z_b=\partial_\sigma+Y_b, \tag{3.35}$$

and suppose Z_a and Z_b are in involution. Let y_∞ in $\mathbb{R}^{n-2}$ be given. *Then* there exists precisely one leaf L of the foliation generated by Z_a and Z_b such that, upon writing

$$L\colon (\tau,\sigma)\in\mathbb{R}^2\rightarrow(\tau,\sigma,y(\tau,\sigma))\in\mathbb{R}^2\times\mathbb{R}^{n-2} \tag{3.36}$$

one has

$$\lim_{\tau\rightarrow\infty}y(\tau,\sigma)=y_\infty. \tag{3.37}$$

Remarks: (a) A similar result holds, *mutatis mutandis*, for $\tau\rightarrow-\infty$.

(b) Under the assumptions of the theorem, Eqs. (3.32) and (3.33) hold uniformly *on compacta*.

(c) Condition (ii) is a Lipschitz condition; since Y_a and Y_b are smooth, it is a statement about the decay rate as $\tau\rightarrow\infty$ of their derivatives in the y directions.

(d) In order to apply the theorem to the relativistic two-body problem, we set $n=14$, $y=(\mathbf{z}_{(k)},v_{(k)})$. As pointed out earlier, the limit $\tau\rightarrow\pm\infty$ corresponds to the asymptotic past or future. Consequently (3.32) and (3.33) give precise conditions on the behavior of the functions $a^\mu_{(k)}$ as $\tau\rightarrow\pm\infty$ for the Möller operators to exist. In particular, for (3.32) to hold, it is sufficient that, for $|\tau|$ large enough and some K

$$|a^\mu_{(k)}(\tau,\sigma,\mathbf{z}_{(k)},v_{(k)})|<K|\tau|^{-(\alpha+1)}, \tag{3.38}$$

with $\alpha>1$ [and where (3.38) is assumed to hold uniformly *on compacta*]. To further interpret (3.38), recall that the $a^\mu_{(k)}$ represent the four-acceleration of particle (k). Despite their complicated dependence on their arguments, condition (3.38) can be checked quite easily. It suffices to implement the coordinate transformation (3.17)–(3.19) and to study the large τ behavior of the $a^{\mu}_{(k)}$ for fixed σ and $y=(\mathbf{z}_{(k)},v_{(k)})$, i.e., for a fixed free motion of the two particles. One can think of Eq. (3.38) as asserting that the force they would experience if they were interacting decays faster than $|\tau|^{-2}$ in the distant past and future. To make contact with our comment just before the statement of the theorem, observe furthermore the following; if we take $\sigma=0$ in (3.38), it follows from (3.18) that $x^0_{(1)}=\tau/2=x^0_{(2)}$. Hence, from (3.19) and (3.17), it follows that $|\mathbf{x}_{(1)}-\mathbf{x}_{(2)}|\sim|\tau|$ for large τ. So the $a_{(k)}$ decay faster than $|\mathbf{x}_{(1)}-\mathbf{x}_{(2)}|^{-2}$ for large interparticle separation and for fixed $y=(\mathbf{z}_{(k)},v_{(k)})$. Equation (3.33) translates in the present context into the requirement that the partial derivatives of the $a^{\mu}_{(k)}$ also decay faster than $|\tau|^{-2}$. In the case both conditions are satisfied, we say the interaction is *short range*. We conclude that (3.32) and (3.33) do indeed give a precise answer to the question raised just before the statement of the theorem.

As a first step towards proving the theorem, we establish the following lemma.

Lemma 3.2: Under the assumptions of Theorem 3.1, for every $y_\infty\in\mathbb{R}^{n-2}$, there exists precisely one map

$$k':(\tau,\sigma)\in\mathbb{R}^2\to k'(\tau,\sigma)\equiv(\tau,\sigma,k(\tau,\sigma))\in\mathbb{R}^2\times\mathbb{R}^{n-2}, \tag{3.39}$$

such that (i) $\forall\ \sigma\in\mathbb{R}$, the curve

$$\tau\in\mathbb{R}\to k'(\tau,\sigma)\in\mathbb{R}^2\times\mathbb{R}^{n-2} \tag{3.40}$$

is an integral curve of Z_a;

$$\text{(ii)}\ \lim_{\tau\to\infty}k(\tau,\sigma)=y_\infty,\quad \forall\sigma\in\mathbb{R}. \tag{3.41}$$

Remarks: (a) k' is clearly injective. In the proof of the theorem, we will show it is an imbedding.

(b) The method of proof is similar to the one used for proving existence and uniqueness of solutions to ordinary differential equations, except that in the present situation we use initial conditions "at infinity." This kind of argument was first used in Ref. 11.

Proof: Choose $y_\infty\in\mathbb{R}^{n-2}$ fixed. We start by remarking that for fixed $\sigma\in\mathbb{R}$ (3.40) is an integral curve of Z_a satisfying (3.41) if and only if, $\forall\ \tau\in\mathbb{R}$,

$$k(\tau,\sigma)=y_\infty-\int_\tau^\infty Y_a(\tau',\sigma,k(\tau',\sigma))d\tau'. \tag{3.42}$$

To prove the lemma, it is therefore sufficient to show that, given $\sigma\in\mathbb{R}$, there exists a $T>0$ and a unique map

$$h_\sigma:(T,\infty)\to\mathbb{R}^{n-2}, \tag{3.43}$$

such that

$$h_\sigma(\tau)=y_\infty-\int_\tau^\infty Y_a(\tau',\sigma,h_\sigma(\tau'))d\tau'. \tag{3.44}$$

Indeed, since Y_a is complete, h_σ can then uniquely be extended over all $\mathbb{R}$ and k defined by $k(\tau,\sigma)=h_\sigma(\tau)$.

Choose then $\sigma\in\mathbb{R}$ fixed. Consider a continuous function

$$v:\tau\in(\kappa,\infty)\to v(\tau)\in B(y_\infty,r), \tag{3.45}$$

where $r=\min(R,\frac{1}{2})<1$ and both R and κ are as in Theorem 3.1. It then follows from (3.32) that

$$\left|\left|\int_\tau^\infty Y_a(\tau',\sigma,v(\tau'))d\tau'\right|\right|\leqslant K\tau^{-\alpha+1}/(\alpha-1). \tag{3.46}$$

Introduce

$$T=\max\{\kappa,((\alpha-1)^{-1}r^{-1}K)^{1/\alpha-1}\} \tag{3.47}$$

and define, for $v\in C((T,\infty),B(y_\infty,r))$

$$(Fv)(\tau)=y_\infty-\int_\tau^\infty Y_a(\tau',\sigma,v(\tau'))d\tau'. \tag{3.48}$$

It follows from (3.46) and (3.47) that F is a well-defined operator on $C((T,\infty),B(y_\infty,r))$. We now prove it is a strict contraction. Let $v,w\in C((T,\infty),B(y_\infty,r))$. Then

$$\|(Fv)(\tau)-(Fw)(\tau)\|\leqslant\int_\tau^\infty\|Y_a(\tau',\sigma,v(\tau'))-Y_a(\tau',\sigma,w(\tau'))\|d\tau'. \tag{3.49}$$

Using (3.33), (3.49) becomes

$$\|(Fv)(\tau)-(Fw)(\tau)\|\leqslant\int_\tau^\infty K\tau'^{-\alpha}\,d\tau'\|v-w\|_\infty. \tag{3.50}$$

Taking the supremum on both sides and using (3.47), one sees

$$\|Fv-Fw\|_\infty\leqslant r\|v-w\|_\infty. \tag{3.51}$$

Since $r<1$, F is a strict contraction and consequently has a unique fixed point. This proves the lemma.

Proof of Theorem 3.1: With k' as in Lemma 3.2 and comparing (3.41) to (3.37), we see that it is sufficient to prove that Im k' is the leaf L generated by Z_a and Z_b through the point $(0,0,k(0,0))$. The latter can be constructed as follows. First remark that Z_a and Z_b commute; this follows from (3.34), (3.35), and the fact that Z_a and Z_b are in involution. Consequently, the flow F^b of Z_b maps integral curves of Z_a into integral curves of Z_a. We now define

$$\lambda':(\tau,\sigma)\in\mathbb{R}^2\to F^b_\sigma(\tau,0,k(\tau,0))\equiv(\tau,\sigma,\lambda(\tau,\sigma))\in\mathbb{R}^2\times\mathbb{R}^{n-2}. \tag{3.52}$$

Then Im λ' is the leaf L. Notice that λ' is an injective immersion. We now prove that $\lambda'=k'$, thereby proving the theorem. We make use of the result of Lemma 3.2 which asserts that k' is unique. Since λ' is of the form (3.39) and satisfies (3.40), it will be sufficient to prove that

$$\lim_{\tau\to\infty}\lambda(\tau,\sigma)=y_\infty,\quad\forall\sigma\in\mathbb{R}. \tag{3.53}$$

It follows from (3.52) that, $\forall\sigma\in\mathbb{R}$,

$$\lambda(\tau,\sigma)=k(\tau,0)+\int_0^\sigma Y_b(\tau,\sigma',\lambda(\tau,\sigma'))d\sigma'. \tag{3.54}$$

Hence

$$\|\lambda(\tau,\sigma)-y_\infty\|\leqslant\|k(\tau,0)-y_\infty\|+\left|\left|\int_0^\sigma Y_b(\tau,\sigma',\lambda(\tau,\sigma'))d\sigma'\right|\right|. \tag{3.55}$$

Choose $1>\epsilon>0$; then, by (3.41) there exists a $\kappa_1>0$ such that

$$\|k(\tau,0) - y_\infty\| < \epsilon/2, \ \forall \tau > \kappa_1. \tag{3.56}$$

To estimate the second term in (3.55), consider for some fixed value of τ and σ for some $\rho > 1$:

$$u: \sigma' \in [0,\sigma] \to u(\sigma') \in B(k(\tau,0),\rho). \tag{3.57}$$

Since (3.32) and (3.33) hold uniformly on compacta, $\exists\ \kappa_2$, $K > 0$ such that $\forall \tau > \kappa_2$,

$$\left|\left|\int_0^\sigma Y_b(\tau,\sigma',u(\sigma'))d\sigma'\right|\right| \leqslant \sigma K \tau^{-\alpha}. \tag{3.58}$$

Now let

$$\kappa = \max\{\kappa_1,\kappa_2,(\epsilon/2\sigma K)^\alpha\}. \tag{3.59}$$

Then

$$\left|\left|\int_0^\sigma Y_b(\tau,\sigma',u(\sigma'))d\sigma'\right|\right| \leqslant \frac{\epsilon}{2}, \quad \forall \tau > \kappa. \tag{3.60}$$

Comparing (3.60) with the second term in (3.55) we see that the proof will be complete if we can show that, for sufficiently large τ, $\lambda(\tau,\sigma') \in B(k(\tau,0),\rho)$, $\forall \sigma' \in [0,\sigma]$. This is done again with a fixed point argument. Choose $\tau > \kappa$ [see (3.59)] and consider $u \in C([0,\sigma],B(k(\tau,0),\rho))$. Define

$$(Fu)(\sigma') = k(\tau,0) + \int_0^{\sigma'} Y_b(\tau,\sigma'',u(\sigma''))d\sigma''. \tag{3.61}$$

It follows from (3.57) and (3.60) that F is a well-defined map on $C([0,\sigma],B(k(\tau,0),\rho))$ (remember that $\rho > 1 > \epsilon$). We now show F is a contraction. For $u,v \in C([0,\sigma],B(k(\tau,0),\rho))$, one finds readily

$$\|Fu - Fv\|_\infty \leqslant \epsilon/2\|u - v\|_\infty, \tag{3.62}$$

where we used once more the fact that (3.33) holds uniformly *on compacta*. So, since $\epsilon < 1$, F is a strict contraction. Its unique fixed point gives the integral curve $\sigma' \in [0,\sigma] \to \lambda'(\tau,\sigma')$ of Z_b through $(\tau,0,k(\tau,0))$. From (3.55), (3.56), (3.59), and (3.60) we conclude that (3.53) is satisfied. Hence the theorem is proved.

Remark: It is clear from the above proof that the conditions of the theorem could be relaxed as follows. For Y_b, Eq. (3.32) can be replaced by

$$\lim_{\tau \to \infty} Y_b(\tau,\sigma',y') = 0 \tag{3.63}$$

and (3.33) by

$$\|Y_b(\tau,\sigma',y'') - Y_b(\tau,\sigma',y')\| \leqslant f(\tau)\|y'' - y'\|, \tag{3.64}$$

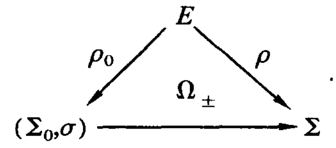
with f some positive function for which $\lim_{\tau\to\infty} f(\tau) = 0$.

IV. ASYMPTOTIC PRESYMPLECTIC STRUCTURES ON *E*

In this section we continue to work within the framework of manifestly covariant particle mechanics, assuming short range forces, so that, as proved in Sec. III, the Möller operators exist. We turn our attention to the problem of determining the presymplectic form ϵ on E, satisfying (2.10) and (2.11). We already pointed out that, even if ϵ exists, it is not necessarily uniquely determined by (2.10) and (2.11). To reduce the remaining freedom in the choice of ϵ, the following additional requirement has been proposed[5,10]: one demands that, as $(x_{(1)} - x_{(2)})^2 \to \infty$, ϵ "approaches" ϵ_0, i.e., ϵ is "asymptotically free." In other words, when the particles are "far apart," ϵ reduces to the free presymplectic form ϵ_0.

As in Sec. III one has to carefully interpret the limit $(x_{(1)} - x_{(2)})^2 \to \infty$. We saw there that it corresponds to letting $|\tau| \to \infty$ on the leaves L_0 of the free foliation. Using the Möller operators, we can now reformulate the additional criterium on ϵ in the form of a simple geometric statement and show to what extent it assures uniqueness of ϵ.

We have the following diagram:



Define, on $\operatorname{Im}\Omega_\pm$, $\sigma_\pm = (\Omega_\pm^{-1})^*\sigma_0$ and, on $\rho^{-1}(\operatorname{Im}\Omega_\pm)$, $\epsilon_\pm = \rho^*\sigma_\pm$. It follows that, on its domain of definition, $\epsilon_\pm$ is a presymplectic form satisfying (2.10) and (2.11). Note that the existence of ϵ_+ and ϵ_- was proved to lowest order in perturbation theory in Ref. 5. Here we reduced the proof of their existence to the existence of the $\Omega_\pm$, which was rigorously established in Sec. III.

Definition 4.1: A presymplectic form ϵ on E, satisfying (2.10) and (2.11), is said to be *asymptotically free* in the future (resp. in the past) if $\epsilon = \epsilon_+$ (resp. $\epsilon = \epsilon_-$) on $\rho^{-1}(\operatorname{Im}\Omega_+)$ [resp. $\rho^{-1}(\operatorname{Im}\Omega_-)$]; ϵ is said to be asymptotically free if it is asymptotically free both in the past and in the future.

It is then clear that, if ϵ and ϵ' are asymptotically free presymplectic forms on E, then ϵ and ϵ' coincide on $\rho^{-1}(\operatorname{Im}\Omega_+)$ and $\rho^{-1}(\operatorname{Im}\Omega_-)$. The criterion of asymptotic freedom does not, however, give information about ϵ on the complement E_B of $E_S \equiv \rho^{-1}(\operatorname{Im}\Omega_+) \cup \rho^{-1}(\operatorname{Im}\Omega_-)$. Here E_B contains the trajectories of the system that correspond to bound states; as a result, the lack of information on ϵ when restricted to E_B from an analysis of the scattering regime of the dynamics should not come as a surprise. Nevertheless, it is proved in Ref. 5 by a formal perturbative calculation that the criterium of asymptotic freedom leads to uniqueness of ϵ on all of E. Our analysis shows that, to prove this result rigorously, one needs additional assumptions of $X_{(1)}$, $X_{(2)}$, and ϵ (note, for example, that analytic dependence on a coupling constant of $X_{(1)}$, $X_{(2)}$, and ϵ is implicit in Ref. 5). We conclude that, by first finding rigorous conditions for the existence of the Möller operators and by defining the asymptotic symplectic structures in terms of them, we have both simplified and clarified the uniqueness question of ϵ.

We close with some remarks on the link between the existence question for ϵ and asymptotic completeness. Let us start by assuming the existence of an asymptotically free ϵ. Then

$$\epsilon_+ = \epsilon_- \quad \text{on } \rho^{-1}(\operatorname{Im}\Omega_+ \cap \operatorname{Im}\Omega_-) \tag{4.1}$$

or

$$\sigma_+ = \sigma_- \quad \text{on } \operatorname{Im}\Omega_+ \cap \operatorname{Im}\Omega_-. \tag{4.2}$$

Systems satisfying (4.1) or (4.2) are called *conservative*.[5] If the system is moreover weakly asymptotically complete ($\operatorname{Im}\Omega_+ = \operatorname{Im}\Omega_-$), then (4.2) is equivalent to symplecticity of the S operator as a map from (Σ_0,σ_0) to (Σ_0,σ_0). We conclude that, in a weakly asymptotically complete theory with an asymptotically free ϵ, S is a symplectic transformation. In that case the results of Ref. 13 apply, where a complete classification of symplectic scattering operators in the relativistic two-body problem is given. We see from the above that a study of weak asymptotic completeness would shed light on the existence question for ϵ as follows: if the system is weakly asymptotically complete, then either S is symplectic $(S^*\sigma_0 = \sigma_0)$ and hence (4.1) and (4.2) hold, or it is not symplectic and then an asymptotically free ϵ does not exist. It would therefore be interesting to find conditions on the $a^{\mu}_{(k)}$ for the system to be weakly asymptotically complete. This question can be expected to be very complicated; already in the nonrelativistic case, a proof of asymptotic completeness in classical two-body systems has only been given for velocity independent forces that are derived from a potential.[11] This means in particular that the symplectic structure of the interacting theory is known *a priori*; it is the symplectic structure inherent in the usual Hamiltonian formulation of nonrelativistic mechanics. In the relativistic case, as pointed out repeatedly, the $a^{\mu}_{(k)}$ are velocity dependent and rather than knowing the symplectic structure of the theory *a priori*, we wish to use knowledge of the scattering operator to study its existence. A further analysis of those problems is therefore needed.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor G. G. Emch for his support and interest during the completion of this work.

The work for this paper was supported in part by a Research Assistantship of the NFWO-Belgium and by the Natural Science and Engineering Research Council (NSERC) Research Grant No. A5206.

[1]P. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, "Relativistic invariance and Hamiltonian theories of interacting particles," Rev. Mod. Phys. **35**, 350 (1963); H. Leutwyler, "A no-interaction theorem in classical relativistic Hamiltonian particle mechanics," Nuovo Cimento **37**, 556 (1965); S. De Bièvre, "Hamiltonian particle mechanics on curved space-times: A no-interaction theorem," J. Math. Phys. **27**, 7 (1986).

[2]E. H. Kerner, *The Theory of Action-At-a-Distance in Relativistic Particle Dynamics. A Reprint Collection* (Gordon and Breach, New York, 1972); "Relativistic action-at-a-distance: Classical and quantum aspects," in *Lecture Notes in Physics*, Vol. 162, edited by J. Llosa (Springer, Berlin, 1982); V. Iranzo. J. Llosa, A. Molina, and F. Marqués, "Comparison of several approaches to the relativistic dynamics of directly interacting particles," Ann. Phys. (NY) **150**, 114 (1983); A. P. Balachandran, G. Marmo, N. Mukunda, J. S. Nilsson, A. Simoni, E. C. G. Sudarshan, and F. Zaccaria, "Unified geometrical approach to relativistic particle dynamics," J. Math. Phys. **25**, 167 (1984).

[3]Ph. Droz-Vincent, "Is interaction possible without heredity?," Lett. Nuovo Cimento **1**, 839 (1969); "Relativistic systems of interacting particles," Phys. Scr. **2**, 129 (1970); J. B. Wray, "Two-particle forces for relativistic Newtonian equations of motion," Phys. Rev. D **1**, 2212 (1970); L. Bel, "Predictive relativistic mechanics," Ann. Inst. H. Poincaré A **14**, 189 (1971); R. Arens, Arch. Rat. Mech. Anal. **47**, 255 (1972).

[4]J. M. Souriau, *Structure des Systèmes Dynamiques* (Dunod, Paris, 1970).

[5]L. Bel and J. Martin, "Formes Hamiltoniennes et systèmes conservatifs," Ann. Inst. H. Poincaré **22**, 173 (1975).

[6]Our notation for differential geometric notions follows closely R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin/Cummings, London, 1978).

[7]See Ref. 6, Secs. 3.2 and 3.3; N. Woodhouse, *Geometric Quantization* (Clarendon, Oxford, 1980), Chaps. I and II; C. Günther, *Präsymplektische Geometrie und Quasikanonische Systeme* (Diplomarbeit, Berlin, 1974).

[8]S. De Bièvre, Ph. D. thesis, University of Rochester, 1986.

[9]L. Bel, A. Salas, and J. M. Sanchez, "Approximate solutions of predictive relativistic mechanics for the electromagnetic interaction I," Phys. Rev. D **7**, 1099 (1973); L. Bel and J. Martin, "Approximate solutions of predictive relativistic mechanics for the electromagnetic interaction II," Phys. Rev. D **8**, 4347 (1973); R. Lapiedra and A. Molina, "Classical predictive electrodynamics of two changes with radiation: General framework. I," J. Math. Phys. **20**, 1308 (1979); R. Lapiedra, F. Marquès, and A. Molina, "II," J. Math. Phys. **20**, 1316 (1979); L. Bel and J. Martín, "Approximate solutions of predictive relativistic mechanics for short-range scalar interactions, "Phys. Rev. D **9**, 2760 (1974).

[10]R. N. Hill and E. H. Kerner, "Unique canonical representation of the inhomogeneous Lorentz group in relativistic particle dynamics," Phys. Rev. Lett. **17**, 1156 (1966).

[11]B. Simon, "Wave operators for classical particle scattering," Commun. Math. Phys. **23**, 37 (1971); H. Narnhofer and W. Thirring, "Classical scattering transformation in classical mechanics," Phys. Rev. A **23**, 1683 (1981).

[12]Typically, one has to exclude a set of measure zero; see Simon in Ref. 11.

[13]S. De Bièvre and G. G. Emch, "Relativistic particle scattering," Commun. Math. Phys. **101**, 539 (1985).

Mass, dual mass, and gravitational entropy

Anne Magnon[a)]
Physics Department, University of Syracuse, Syracuse, New York 13244-1130 and Département de Mathématiques, Université de Clermont-Fd, 63170 Aubière, France

(Received 15 October 1985; accepted for publication 25 March 1987)

A mechanism (mathematical transformation) is considered by which a (Schwarzschild) black-hole singularity can be converted into a (Taub–NUT) (Newman–Unti–Tamburino) wire singularity, or equivalently topological modifications can be induced (e.g., transformation of $S^2 \times R^2$ topology into $S^3 \times R$ and conversely). A topological charge—an invariant of the transformation—emerges as a possible candidate for the description of gravitational entropy in the case of source-free solutions to Einstein's equation with one Killing vector field. For a Schwarzschild black hole this invariant reduces to the area of the event horizon (or equivalently the Bolt charge) and it reduces to the square of the NUT charge (or equivalently the length of the closed timelike orbits) in the case of a Taub–NUT magnetic monopole. These considerations lead to the proposition that, under extreme conditions, gravitational clumping or entropy increase could be described by a modification in the characteristic classes of the space-time manifold due to the onset of nontrivial topological features. Further remarks are presented in view of the role of gravitational magnetic monopoles in quantum gravity, and of a possible relation between the notions of gravitational entropy and arrow of time.

I. INTRODUCTION

The problem of the relation between the notion of entropy and the origin of the arrow of time is one of the most fascinating of contemporary physics. In the early 1970's, with the advent of black-hole physics, baryon and lepton numbers ceased to be conserved, throwing the second law of thermodynamics into jeopardy: if a black hole could conceal the swallowed particles and their physical attributes, it should also do the same for their infalling entropy. Since the only parameters characterizing a black hole are its mass, charge, and spin (m,e,J), an observer standing outside the hole would not have enough information to evaluate the thermodynamic state of the trapped matter. Bekenstein and Hawking[1,2] saved the second law of thermodynamics near a black hole by assigning an entropy S_h to the hole, proportional to the area A of its event horizon: $S_h = \tau A$. Support for such a definition came from the discovery that small perturbations of the hole satisfy an equation which is analogous to the first law of thermodynamics,

$$dm = (\kappa/8\pi)dA + \Omega\, dJ + \phi\, de,$$

where κ, Ω, ϕ, are, respectively, the "surface gravity," the angular velocity, and the electric potential of the hole, all evaluated at the event horizon. Next it was suggested that a temperature

$$T_h = \kappa/8\pi\tau$$

should be assigned to the hole, constancy of (T_h, Ω, ϕ) being appropriate to describe a state of thermal–mechanical–electrical equilibrium. Since classically a black hole does not radiate, support for the introduction of T_h came from Hawking's discovery[3] that holes, at the quantum level, emit all kinds of particles at a temperature $\kappa/2\pi$. More recently, however, it has been objected[4,5] that the Hawking–Bekenstein definition might not be deprived of drawbacks.

(a) The relation between S_h and the physical entropy has only been obtained indirectly via an analogy with thermodynamics (quantum statistical foundations which would bring further support to the definition yet to be forged).

(b) S_h is a single number lacking any space-time distribution outside the horizon, and is not a proper state function because it depends on the horizon, a global entity, which cannot be measured directly on a given space section. Furthermore, one can find null hypersurfaces closely analogous to future event horizons which are deprived of meaningful entropy.

(c) It has also been underlined[6] that the physical origin of this entropy is far from being clear: the formula $S = K \log N$, on which our general understanding of the second law is based, entails the absurdity $S = \infty$ since a bound on the total energy does not suffice to bound the number of possible internal states; for instance, the Oppenheimer–Snyder solutions[7] provide an infinite number of possible internal configurations for a Schwarzschild exterior of fixed mass.

(d) Finally, even if one accepts basing the definition of gravitational entropy on the area of the event horizon, the problem of its relation with the concept of a time arrow remains open, and the task seems to be formidable due to the existence of singularities and nontrivial topology "inside" the hole. Recently, an enlightening synthesis of these difficulties has been proposed by Penrose,[8] who suggests that a key to the problem might lie in a better understanding of the structure of singularities. He further argues that "there should exist a qualitative relation between gravitational clumping and an entropy increase, due to taking up of gravitational potential energy. Hence a high entropy singularity should involve a very large Weyl curvature although so far no mathematical expression of this suggested relation

[a)] Détachée du Ministère des Relations Extérieures, Paris, France.

between Weyl curvature and gravitational entropy has come to light."

In this paper we would like to suggest an approach to the problem. We propose to investigate a relation between modifications of the structure of singularities and modifications in the space-time topology. Our line of arguments will be the following. We shall first (Sec. II) briefly introduce the available definitions of the magnetic mass, a topological charge characteristic of the presence of a gravitational magnetic monopole, and show that these definitions agree. We shall next (Sec. III) restrict ourselves to vacuum solutions to Einstein's equations that admit one Killing vector field and consider a mechanism by which the region of trapped surfaces enclosed within a Schwarzschild horizon can be converted into a (Taub–NUT) wire singularity, or equivalently (Misner's smoothing out procedure) into a modification of the space-time topology (introduction of causality violations). In this process [which intuitively could be viewed as a shrinking of the region enclosed within the event horizon followed by (i) a gluing of the "lines" $r=2m$ in the Schwarzschild diagram, and (ii) an erasing of the resulting "wire singularity at $r=2m$" via a topological modification (from $S^2\times R^2$ to $S^3\times R$)] a new topological charge emerges: a gravitational magnetic monopole has been generated. We then propose a possible candidate for the mathematical description of this process, the Buchdahl,[9] Ehlers,[10] Harrison,[11] Geroch[12] (BEHG) transformation initially presented to generate source-free solutions to Einstein's equation with one Killing vector field, starting from one such solution. Under this transformation, the Schwarzschild solution is converted into a Taub–NUT solution, and the onset of a magnetic mass can be viewed as being induced by that of a wire singularity or equivalently a modification of the space-time topology.[13,14] The entropy trapped inside the Schwarzschild event horizon has been converted into new topological features inducing a causality violation[13] or equivalently new topological charge.

It is also noticed that the BEHG transformation can be reduced to a duality rotation (action of the circle group) on appropriate potentials on the manifold of orbits of the Killing field, inducing twist on these orbits and helicity on the resulting solution to Einstein's equation. It is a charging mapping in the sense of Ref. 14, the total helicity (or flux of the torsion field) being equal to the induced magnetic mass. We shall then consider an invariant of the BEHG transformation and propose a definition of gravitational entropy based on the existence of this invariant for each solution of the BEHG (circle) family. Entropy appears as a topological charge and reduces to its known values in the case of static black holes or Taub–NUT solutions.

II. PRELIMINARIES

Recently boundary conditions suitable for asymptotically Taub–NUT solutions have been considered,[13] leading to an expression of the dual (or magnetic) mass which involves the asymptotic Weyl curvature. On another hand, in the presence of a stationary Killing vector field, another expression of this quantity is available.[14] We shall briefly introduce these definitions and prove that they agree.

A. Stationary space-times

Let (M, g_{ab}, t^a) be a vacuum stationary space-time, i.e., a four-manifold M with a metric g_{ab} of signature $(-,+,+,+)$, solution of $R_{ab}=0$, and t^a a Killing vector field assumed to be (in this section) everywhere timelike and complete. Let $\pi: M\to T$ denote the projection map from M into T, the manifold of orbits of t^a. We assume that T has topology $S^2\times R$ (topology of a wormhole) and we set $t^a t_a = -\lambda$ (the norm of t^a),

$$\omega_a=\lambda^{-1/2}\epsilon_{abcd}t^b\nabla^c t^d$$

(the twist of t^a). There exists on M a curl-free two-form

$$F_{ab}=\nabla_{[a}\lambda^{-1}t_{b]}=\lambda^{-1/2}\epsilon_{abcd}t^d\omega^c, \tag{1}$$

which satisfies $t^aF_{ab}=0$ and $L_tF_{ab}=0$. Hence F_{ab} can be viewed as the pullback to M of a curl-free two-form $\widetilde{F}_{ab}=D_{[a}v_{b]}$ on T, where D denotes the derivative operator compatible with the induced metric $h_{ab}=g_{ab}+\lambda^{-1}t_at_b$. Let S_2 denote a two-sphere surrounding, within T, the nontrivial topological features. Since the second homotopy group of T is nontrivial,

$$\int_{S_2}\widetilde{F}_{ab}\,dS^{ab}$$

need not vanish. Furthermore, since the flux of $\widetilde{F}_{ab}$ through a cap of S_2 bordered by a loop C is given by

$$\int_C v_b\,dS^b,$$

v_b must develop a singularity on S_2. This is characteristic of the presence of a magnetic monopole.[14] The flux

$$N=\int_{S_2}\widetilde{F}_{ab}\,dS^{ab} \tag{2}$$

measures the magnetic charge of the handle.

B. Asymptotically Taub–NUT space-times

Let us now focus on the asymptotically Taub–NUT space-times which have been investigated in Ref. 13. The presence of a Killing vector field is not required here. Such space-times admit a conformal null boundary $\mathscr{I}$ which exhibits the topology of an S^1 fiber bundle over S^2, the Hopf fibers being integral curves of a null vector n^a. Let $\hat{g}_{ab}=\Omega^2g_{ab}$ denote the appropriately rescaled metric which attaches $\mathscr{I}$ to the physical space-time (M,g_{ab}) and $\widehat{\nabla}_a$ the derivative operator which is compatible. When the News function N_{ab} vanishes at $\mathscr{I}$, the magnetic mass $N_{\mathscr{I}}$ can be defined[13] via its action on an asymptotic translation αn^a, $(L_n\alpha=0)$,

$$N_{\mathscr{I}}=\int_C\alpha\epsilon_{abc}\,{}^*K^{cm}l_m\,dS^{ab} \tag{3}$$

where C denotes a cross section of $\mathscr{I}$, ϵ_{abc} and $(m^a,\bar{m}^a,l^a,n^a)$ are, respectively, the induced alternating tensor at $\mathscr{I}$ and the usual Newman–Penrose null tetrad; ${}^*K^{cm}$ is introduced[13] at $\mathscr{I}$, via the pullback of $\Omega^{-1}C^a{}_{bcd}$

(the rescaled Weyl tensor) and is submitted to conditions (4)–(8) listed in Ref. 13. We want to prove that N and $N_{\mathscr{I}}$ agree when the space-time is stationary. Recall that in the presence of an everywhere timelike and complete Killing vector field t_a on (M, g_{ab}) a simple geometrical interpretation of N has been presented[14]: N essentially measures the length of the orbits of t_a (which therefore must be closed), or equivalently the number of times the space-time bundle winds around its fiber. If the expressions of N and $N_{\mathscr{I}}$ are to agree, a similar geometrical interpretation should be available at $\mathscr{I}$. Let us prove that it is so.

Let $\widehat{D}_{[a}w_{b]}$ denote the pullback to $\mathscr{I}$ of $\widehat{\nabla}_{[a}\lambda^{-1}t_{b]}$ where w_b is some covector field on $\mathscr{I}$. Since

$$n^a\widehat{D}_{[a}w_{b]}=0, \quad \mathscr{L}_n\widehat{D}_{[a}w_{b]}=0 \tag{4a}$$

and

$$n^a\widehat{D}_{[a}w_{b]}=\tfrac{1}{2}L_n w_b-\tfrac{1}{2}\widehat{D}_b(w\cdot n), \tag{4b}$$

we can choose w_b such that $w\cdot n=\widehat{C}$ (a constant) (note that w_b is arbitrary up to a gradient). Hence w_b defines a connection on $\mathscr{I}$. Next, we know, from Ref. 15 and formula (6a) in Ref. 13, of the existence of a gauge such that $S_{ab}\equiv R_{ab}-\frac{1}{6}Rg_{ab}=g_{ab}$, degenerate metric on $\mathscr{I}$; this in turn implies that $S_a{}^b=\delta_a{}^b-\nu_a n^b$, where ν_a is some covector field on $\mathscr{I}$. In particular one can identify ν_b with w_b. Under these conditions and since[13]

$$\widehat{D}_{[a}S_{b]}{}^c=\tfrac{1}{4}\epsilon_{abm}{}^*K^{mc}=-\widehat{D}_{[a}\nu_{b]}n^c, \tag{5}$$

$N_{\mathscr{I}}$ appears as a constant multiple of

$$\int_{\mathscr{S}_2}\widetilde{D}_{[a}\tilde{\nu}_{b]}\,dS^{ab}; \tag{6}$$

here $\mathscr{S}$ denotes the manifold of orbits of n^a, $\tilde{\nu}_b$ is a covector field on $\mathscr{S}$ which admits ν_b as its pullback to $\mathscr{I}$, and $\widetilde{D}_a$ is some derivative operator on $\mathscr{S}$. (Recall[13] that for any infinitesimal translation αn^a, at $\mathscr{I}$, α satisfies on $\mathscr{S}$ the rescaling invariant equation

$$\tilde{g}^{mn}\widetilde{D}_m\alpha\widetilde{D}_n\alpha-\tilde{g}^{mn}\widetilde{D}_m\widetilde{D}_n\alpha-\tfrac{1}{2}R\alpha^2=\text{const},$$

hence can be chosen equal to a constant.)

Let Γ denote the lift, to $\mathscr{I}$, of some closed loop γ on $\mathscr{S}$, say originating at some point p on $\mathscr{S}$. Since ν_b defines a unique horizontal lift, on $\mathscr{I}$, we can choose for Γ the closed curve consisting of a horizontal curve originating at P_1 and ending at P_2, two points on the fiber above p, and of that segment of the fiber (vertical piece of Γ) starting at P_2 and ending at P_1. This implies that

$$\int_\Gamma \nu_b\,dS^b=\widehat{C}\int_{P_2}^{P_1}ds \tag{7}$$

where s is, along the fiber, the affine parameter defined by $n^a\widehat{D}_a=\partial/\partial s$. Since $\mathscr{S}$ can be swept by a one-parameter family of loops $\gamma_s(t)$, $0\leqslant s\leqslant 1$, starting and ending (with opposite orientation) at the point p [the trivial loop $\gamma_0(t)\equiv\gamma_1(t)\equiv p$], it is clear from (6) and (7) that, as s varies from 0 to 1, P_2 must describe the whole fiber above p. Hence $N_{\mathscr{I}}$ is essentially a measure of the length of the orbits of n^a (Hopf fibration of $\mathscr{I}$) or equivalently of the number of times $\mathscr{I}$ winds around its fiber. It is to be noticed that ν_b plays, in the above proof, a role analogous to that played by k_b, the null vector field which had to be introduced in Ref. 14 to show that t^a must have closed orbits. This completes the proof.

Theorem: The two definitions of N and $N_{\mathscr{I}}$ are in agreement. In both cases the magnetic mass measures the number of times the (corresponding) bundle winds around its (closed) fiber.

III. BUCHDAHL–EHLERS–HARRISON–GEROCH TRANSFORMATION: A MECHANISM TO GENERATE GRAVITATIONAL (MAGNETIC) MONOPOLES

Recently we introduced[14] a mathematical framework suitable for the description of gravitational magnetic monopoles. In this framework a mechanism was presented by which tensor fields (e.g., metrics) can acquire a weight θ, and derivative operators a torsion $F_{ab}(\theta)$. In this process the key role was played by "charging" mappings f_θ, θ being an element of a cyclic group. The existence of singularities in the potential of $F_{ab}(\theta)$ is a manifestation of the presence of a magnetic monopole, the flux of F_{ab} being a measure of its charge.

On another hand, a transformation is available, the BEHG (Refs. 9–12) transformation, which generates explicit, exact, source-free solutions to Einstein's equation with one Killing vector field, starting from one of them. The resulting family forms a circle. Repeating the transformation merely results in a further rotation within the original circle of solutions. In this section we want to show that this transformation can be considered as a charging mapping f_θ in the sense of Ref. 14. Starting from a particularly chosen solution, the transformation can generate a gravitational magnetic monopole. For instance, if one starts from the Schwarzschild (resp. Kerr) black-hole solution, there is a value of θ for which the resulting charged metric is the Taub–NUT (resp. Kerr–NUT) solution, the corresponding magnetic charge being the NUT parameter. Hence this transformation provides a mechanism by which black-hole singularities and event horizons can be converted into wire singularities (or equivalently causality violations[14]), and conversely. This transformation can thus be viewed as inducing topological modifications [e.g., the Schwarzschild $S^2\times R^2$ topology being converted into $S^3\times R$ topology and conversely (intuitively, the time arrow, which is available on each asymptotic region of the Schwarzschild solution, disappears—closing of the orbits of the timelike Killing vector field—as the region of trapped surfaces shrinks into a wire singularity and gets absorbed into new topological features)].

We shall briefly summarize the BEHG transformation. Let (M, g_{ab}) be a solution of Einstein's equation $R_{ab}=0$, with Killing vector field ξ^a. In this section, the norm and twist of ξ^a are, respectively, $-\lambda=\xi^a\xi_a$ and $\omega_a=\epsilon_{abcd}\xi^b\nabla^c\xi^d=\text{grad}_a\,\omega$. Let $h_{ab}=g_{ab}+\lambda^{-1}\xi_a\xi_b$ be the induced metric on T, the manifold of orbits of ξ^a. Assuming $\lambda\neq 0$, a rescaled metric $\tilde{h}_{ab}$ and complex potential τ can be introduced[12] on T:

$$\tilde{h}_{ab}=\lambda h_{ab}, \tag{8a}$$

$$\tau=\omega+i\lambda. \tag{8b}$$

It then follows[12] that Einstein's vacuum equation $R_{ab}=0$ reduces to

$$\widetilde{R}_{ab}=-2(\tau-\bar{\tau})^{-2}(\widetilde{D}_{(a}\tau)(\widetilde{D}_{b)}\bar{\tau}), \tag{9a}$$

$$\widetilde{D}^2\tau=2(\tau-\bar{\tau})^{-1}(\widetilde{D}_m\tau)(\widetilde{D}_n\tau)\tilde{h}^{mn}, \tag{9b}$$

where $\widetilde{D}_a$ and $\widetilde{R}_{ab}$ are the derivative and Ricci tensor with respect to $\tilde{h}_{ab}$ ($\widetilde{D}^2=\tilde{h}^{ab}\widetilde{D}_a\widetilde{D}_b$). Given a solution $\tilde{h}_{ab}$, τ of (9), a new solution $\tilde{h}'_{ab}=\tilde{h}_{ab}$, $\tau'=\tau'(\tau)$ can be obtained after substitution of $\tilde{h}'_{ab}$ and τ' in (9). The resulting potential τ' is given by

$$\tau'=(a\tau+b)/(c\tau+d), \tag{10}$$

where a, b, c, d are real numbers such that $ad-bc=1$. If one sets

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{11}$$

the resulting metric exhibits a weight θ,

$$\begin{aligned} g'_{ab}(\theta) &= [(\cos\theta-\omega\sin\theta)^2+\lambda^2\sin^2\theta]g_{ab} \\ &+ 2\xi_{(a}[2\alpha_{b)}\cos\theta-\beta_{b)}\sin\theta]\sin\theta \\ &+ \lambda[(\cos\theta-\omega\sin\theta)^2+\lambda^2\sin^2\theta]^{-1}[2\alpha_a\cos\theta-\beta_a\sin\theta][\sin^2\theta][2\alpha_b\cos\theta-\beta_b\sin\theta], \end{aligned} \tag{12}$$

where coefficients α_b and β_b have been introduced in Ref. 12. It has been charged in the sense of Ref. 14. Another way to display the charging mapping f_θ consists in introducing the following potentials:

$$\phi_M=\tfrac{1}{4}(\lambda^2+\omega^2-1),\quad \phi_J=\omega/2\lambda. \tag{13}$$

Substituting into the expression for τ', a straightforward calculation shows that the transformation reduces to

$$f_\theta\begin{pmatrix} \phi_M \\ \phi_J \end{pmatrix}=\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} \phi_M \\ \phi_J \end{pmatrix}. \tag{14}$$

This provides a representation of the cyclic group in the (ϕ_M,ϕ_J) plane. For the Schwarzschild solution, $\phi_M=1/4\lambda(\lambda^2-1)$ and $\phi_J=0$. Hence the mass monopole (in ϕ_M) is nonzero (since $\lambda=-1+2M/r$), while the dual (magnetic) monopole vanishes. A rotation in the (ϕ_M,ϕ_J) plane leads, for $\theta=\pi/4$, to a new solution, with potentials $\phi'_M=0$ and $\phi'_J=-\phi_M$. This solution is precisely the Taub–NUT metric with NUT parameter $N=-M$. This parameter appears clearly as dual to the mass. For the Kerr solution, the action of $f_{\pi/4}$ leads to a solution that, in the absence of dipole angular momentum, must reduce to the Taub–NUT solution: hence this solution is one of the Kerr–NUT solutions. We shall now display the topological charges that can emerge from the action of f_θ. Recall the existence of three real divergence-free vector fields on T,

$$V_1{}^a(\tau-\bar{\tau})^{-2}\tilde{h}^{am}(\widetilde{D}_m\tau+\widetilde{D}_m\bar{\tau}), \tag{15}$$

$$V_2{}^a(\tau-\bar{\tau})^{-2}\tilde{h}^{am}(\bar{\tau}\widetilde{D}_m\tau+\tau\widetilde{D}_m\bar{\tau}), \tag{16}$$

$$V_3{}^a(\tau-\bar{\tau})^{-2}\tilde{h}^{am}(\bar{\tau}^2\widetilde{D}_m\tau+\tau^2\widetilde{D}_m\bar{\tau}) \tag{17}$$

with the following properties: (i) under f_θ, these vector fields are mapped into linear combinations of themselves

$$f_\theta V_1{}^a=V_1{}^a\cos^2\theta-V_2{}^a\sin 2\theta+V_3{}^a\sin^2\theta, \tag{18}$$

$$f_\theta V_2{}^a=\tfrac{1}{2}V_1{}^a\sin 2\theta+V_2{}^a\cos 2\theta-\tfrac{1}{2}V_3{}^a\sin 2\theta, \tag{19}$$

$$f_\theta V_3{}^a=V_1{}^a\sin^2\theta+V_2{}^a\sin 2\theta+V_3{}^a\cos^2\theta; \tag{20}$$

and (ii) the corresponding curl-free two-forms $\epsilon_{abc}V_i{}^c\equiv\widetilde{F}^i{}_{ab}$ ($i=1,2,3$) on T admit the following pull-backs[12] on the space-time (M,g_{ab}):

$$F^1{}_{ab}=\nabla_{[a}\lambda^{-1}\xi_{b]}, \tag{21}$$

$$F^2{}_{ab}=\nabla_{[a}\omega\lambda^{-1}\xi_{b]}-\tfrac{1}{2}\epsilon_{abcd}\nabla^c\xi^d, \tag{22}$$

$$F^3{}_{ab}=\nabla_{[a}(\lambda^{-1}(\omega^2+\lambda^2)\xi_{b]})-2\lambda\nabla_a\xi_b-\omega\epsilon_{abcd}\nabla^c\xi^d. \tag{23}$$

Integrating the forms $\widetilde{F}^i{}_{ab}$ on a two-sphere surrounding the nontrivial topological features leads to various conserved quantities: (i) in the case of a magnetic monopole $\widetilde{F}^1{}_{ab}$ provides the magnetic mass; (ii) in the static case, $\widetilde{F}^3{}_{ab}$ is identically zero, while $\widetilde{F}^2{}_{ab}$ provides the Komar mass; and (iii) consider the Kerr solutions representing the stationary axisymmetric asymptotically flat field outside a rotating massive object; an evaluation of

$$\int_\Sigma F^1{}_{ab}\,d\Sigma^{ab},$$

the flux of $F^1{}_{ab}$ through a cap Σ bordered by a closed loop γ surrounding, in the physical space-time, the rotating source, can be obtained. Choose for γ one of the closed orbits of the rotational Killing vector fields R^a, and consider the timelike two-dimensional cylinder μ generated by the orbits of the stationary Killing vector field t^a, through γ (μ is an integral manifold of the two commuting Killing vector fields). We know, from Ref. 14, that

$$\Delta s=2\lambda^{1/2}\int_\Sigma\nabla_{[a}\lambda^{-1}t_{b]}\,dS^{ab}=2\lambda^{-1/2}\int_\gamma t_a\,dS^a, \tag{24}$$

where s denotes the affine parameter along the orbits of t^a. A direct calculation leads to

$$\Delta s=2\lambda^{-1/2}(R\cdot R)^{-1/2}\cdot(\text{length of }\gamma)\cdot(t\cdot R). \tag{25}$$

In Boyer–Lindquist coordinates, the cylinder will be defined by $r=r_0$, $\phi=\phi_0$. If we further require that the four-velocity of the loop γ be a multiple of t^a, the expression of Δs reduces to

$$\begin{aligned}\Delta s &= 4\pi(4A_0M_0r_0\sin^2\phi_0) \\ &\times(r_0{}^2+A_0{}^2\cos^2\phi_0-2M_0r_0)^{-1/2} \\ &\times(r_0{}^2+A_0{}^2\cos^2\phi_0)^{-1/2}.\end{aligned}$$

This formula shows that, in the absence of wire singularity (or causality violation), the conserved quantity generated by $F^1{}_{ab}$ is a measurement of the state of rotation of the source, or equivalently of the dipole angular momentum of the solution.

Let us return to the BEHG transformation. Let S_2 denote, in the manifold of orbits of ξ^a, a two-sphere surrounding the nontrivial topological features. The associated charges are

$$Q_1 = \int_{S_2} \widetilde{F}^1{}_{ab}\, dS^{ab}, \tag{26}$$

$$Q_2 = \int_{S_2} \widetilde{F}^2{}_{ab}\, dS^{ab}, \tag{27}$$

$$Q_3 = \int_{S_2} \widetilde{F}^3{}_{ab}\, dS^{ab}. \tag{28}$$

Under the action of the (BEHG) circle group, these quantities obey the same transformation rules [(18)–(20)] as the corresponding forms, implying the existence of an invariant of the transformation

$$E = Q_1 Q_3 - (Q_2)^2. \tag{29}$$

This leads to the following theorems.

Theorem 1: The action of $f_{k\pi + \pi/4}$ on any (causally well-behaved) vacuum static solution with nonvanishing Komar mass generates angular momentum. If the static solution is the Schwarzschild black-hole solution, the resulting solution is the Taub–NUT magnetic monopole.

This follows immediately from the fact that $Q_3 = 0$ when ξ^a is hypersurface orthogonal and from formulas (19) and (29).

Corollary: The mapping $f_{k\pi + \pi/4}$ can induce modification in the space-time topology.

This follows immediately from Ref. 14.

Theorem 2: The action of $f_{k\pi + \pi/4}$ on a Kerr solution (with angular momentum dipole) induces an angular momentum monopole. The resulting solution belongs to the Kerr–NUT family.

This follows immediately from formula (19) and Theorem 1.

These results suggest that a relation should exist between the invariant E and the concept of gravitational entropy, bringing support to the following.

Definition: Each solution of Einstein's vacuum equation, with one timelike Killing vector field, will be assigned a gravitational entropy defined by

$$\mathscr{E} = \Big[\int_{S_2} \nabla_{[a} \lambda^{-1} \xi_{b]}\, dS^{ab}\Big]\Big[\int_{S_2} \{\nabla_{[a}(\lambda^{-1}(\omega^2 + \lambda^2)\xi_{b]}) - 2\lambda \nabla_a \xi_b - \omega \epsilon_{abcd} \nabla^c \xi^d\} dS^{ab}\Big] - \Big[\int_{S_2} \{\nabla_{[a} \omega \lambda^{-1} \xi_{b]} - \tfrac{1}{2} \epsilon_{abcd} \nabla^c \xi^d\} dS^{ab}\Big]^2. \tag{30}$$

Theorems 1 and 2 imply in particular that $\mathscr{E}$ reproduces the Bolt charge and NUT charge which have been proposed[19] to describe the gravitational entropy of a Schwarzschild black hole and of a Taub–NUT (Kerr–NUT) magnetic monopole, respectively.

Remarks: Such a formulation has the advantage to cure some of the drawbacks listed in the Introduction: (a) it is obtained within the hyperbolic regime, and does not require the introduction of gravitational instantons in the Euclidean regime; (b) it is not based on null hypersurfaces and could, in principle, be tested on a space section; (c) it emphasizes a possible role of the notion of entropy in the description of modification of the structure of singularities (due to gravitational clumping) and onset of new topological features; (d) the definition could be generalized and adapted to more realistic situations (absence of an isometry, expanding universes); and (e) although the problem of the relation with physical entropy and thermodynamics is left open, further support to such a geometrical approach will be given in our concluding remarks.

IV. CONCLUDING REMARKS

We believe that the following remarks would bring support to the viewpoint developed in this paper according to which the notion of gravitational entropy (entropy increase, time arrow) could be related to (modifications of) the topological structure of the space-time manifold.

(i) In semiclassical electrodynamics, magnetic charge can be introduced as a purely topological construction. Let us outline it briefly. The wave function of a particle, a complex-valued field on the space-time manifold, is defined up to phase. Hence it is more appropriate to introduce a principal S^1 fiber bundle $\mathscr{B}$ over the space-time manifold M, the group structure G on the fiber being induced by the group of complex numbers with modulus unity. The field is thus associated to a cross section of this bundle, the electromagnetic potential serving as a connection,

$$\nabla_\mu = \partial_\mu + A_\mu .$$

This connection in turn provides a notion of horizontal lift for any closed loop γ in M (originating say at point p). Let Γ denote such a lift, a curve originating at point P_1, on the fiber above p, and meeting the fiber again at point P_2. The segment P_2P_1 on the fiber completes Γ into a closed loop $\widetilde{\Gamma}$. The map ψ: $P_1 \to P_2$ defines a subgroup of G, and a subgroup of the holonomy group of the bundle. With each element of the second homotopy group $\pi_2(M)$ of the space-time manifold, there is associated an element of the phase group G describing the change in phase of a field $F_{ab}(\theta)$ as a two-sphere in M [an element of $\pi_2(M)$] is spread by a one-parameter family of closed loops. This phase shift can be evaluated as the flux of $F_{ab}(\theta)$ (the curvature of the connnection on the bundle) and thus be identified with the magnetic charge. Furthermore, this charge is invariant under continuous deformation of the curve Γ in the bundle. This in turn provides a measurement of the following mapping of homotopy groups:

$$\pi_2(M) \to \pi_1(\mathscr{B}).$$

A physical quantity (the Bohm–Aharanov phase shift) turns out to be a purely topological concept.

(ii) On another hand, we have shown (Ref. 14) that in the presence of a magnetic mass, a stationary space-time M (such that T, the manifold of orbits of the timelike Killing vector field, has $S^2 \times R$ topology) must be an S^1 bundle over T. The magnetic mass is a physical quantity associated to the space-time topological structure ($S^3 \times R$ topology). Such topological charges could be used to test the presence of mappings between homotopy groups (associated to the space-time topology), or equivalently the presence of exact sequences[16] of homomorphisms such as

$$\pi_n(M,p) \cdots \to \pi_{n-1}(M,p) \to \cdots \to \pi_0(M,p) \to 0,$$

where each homomorphism in the exact sequence is the kernel of the next. A similar test has already been proposed[17] within the context of Maxwell theory. These ideas bring further support to the viewpoint presented in this paper according to which characteristic classes and associated topological charges might play an important role in the description of the modification of the structure of singularities due to gravitational clumping and, consequently, might shed some light on the concept of entropy.[18]

ACKNOWLEDGMENTS

I would like to thank Professor R. Penrose for suggestions, and P. Mazur and B. Schmidt for helpful comments.

[1] J. D. Bekenstein, Ph.D thesis, Princeton University, 1972; Phys. Rev. D **7**, 2333 (1973).

[2] S. W. Hawking, Phys. Rev. Lett. **26**, 1344 (1971).

[3] S. W. Hawking, "Particles creation by black-holes," DAMTP preprint, Cambridge University, March 1974; "Black-hole explosions?," DAMTP preprint, Cambridge University, January 1974.

[4] D. Wilkins, Gen. Relativ. Gravit. **11**, 45 (1979).

[5] W. Kundt, Nature **259**, 30 (1976).

[6] R. Sorkin, "On the entropy of the vacuum outside a horizon," Tenth International Conference on General Relativity and Gravitation, Padova, July 1983, edited by B. Bertotti, F. de Felice, and A. Pascolini.

[7] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1970).

[8] R. Penrose, "Singularities and time-asymmetry," in General Relativity, an Einstein Centenary Survey, edited by S. W. Hawking, 1979.

[9] M. Buchdahl, Q. J. Math. **5**, (1954).

[10] J. Ehlers, in *Les Théories Relativistes de la Gravitation* (CNRS, Paris, 1959).

[11] B. K. Harrison, J. Math. Phys. **9**, 1744 (1968).

[12] R. Geroch, J. Math. Phys. **12**, 6 (1971).

[13] A. Magnon, J. Math. Phys. **27**, 1059 (1985).

[14] A. Magnon, J. Math. Phys. **27**, 1066 (1985).

[15] R. Geroch, *Asymptotic Structure of Space-Time*, edited by F. P. Esposito and L. Witten (Plenum, New York, 1977).

[16] Sz. T. Hu, *Homotopy Theory* (Academic, New York, 1959).

[17] C. J. S. Clarke, Gen. Relativ. Gravit. **2**, 43 (1971).

[18] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977); Commun. Math. Phys. **66**, 291 (1979).

Colliding gravitational plane waves with noncollinear polarization. I

Frederick J. Ernst
Department of Mathematics, Clarkson University, Potsdam, New York 13676
Alberto García D.
Departamento de Física, CINVESTAV del Instituto Politécnico Nacional, Apdo. 14-740, 07000 México, D. F., Mexico
Isidore Hauser[a)]
1615 Cottonwood Dr., #8, Louisville, Colorado 80027

(Received 1 December 1986; accepted for publication 29 April 1987)

An Ehlers transformation on the Ernst potential for the Nutku–Halil solution [Phys. Rev. Lett. **39**, 1379 (1977)] provides a new solution of the Einstein field equations describing colliding gravitational plane waves with noncollinear polarization, the first of an infinite sequence of solutions that can be generated using techniques described in this paper.

I. INTRODUCTION

Many examples are known of exact global solutions of the vacuum Einstein field equations that describe the collision of two gravitational plane waves[1] with collinear polarization, but few are known that describe the collision of two such waves with noncollinear polarization, where the metric is nondiagonalizable. In the former category are the Khan–Penrose solution[2] and an infinite family of solutions obtained recently by Ferrari and Ibañez[3] using the inverse scattering technique of Belinskii and Zakharov.[4] In the latter category the Nutku–Halil solution[5] and a solution constructed by Chandrasekhar and Xanthopoulos[6] using the region of the Kerr metric[7] interior to the event horizon stand alone. With the exception of the Khan–Penrose and the Nutku–Halil solution, which involve purely impulsive plane waves, all the other known solutions involve a combination of impulsive and shock waves, and their physical significance is not well understood at present.[8]

Some time ago[9] one of the authors of the present paper showed that a double-Harrison[10] (Bäcklund) transformation, when applied to the isotropic Kasner metric, yields the Nutku–Halil colliding wave metric. The derivation was carried out using the Hauser–Ernst[11] homogeneous Hilbert problem (HHP) formulation. It is, of course, well known that the same double-Harrison transformation, when applied to Minkowski space, another Kasner metric, yields the Kerr solution, which Chandrasekhar and Xanthopoulos showed also admits a colliding wave interpretation. It was natural, therefore, to suppose that such a double-Harrison transformation, when applied to the other Kasner metrics, would yield additional colliding wave solutions with noncollinear polarization, each generalizing one of the solutions of Ferrari and Ibañez.

Exploiting a generalization of the Kinnersley–Chitre transformation[12] group structure, we observed in the course of studying this problem some interesting relationships among the colliding wave solutions corresponding to different values of the Kasner parameter. Instead of calculating each metric directly from the corresponding Kasner metric, it was possible to obtain one colliding wave solution from another merely by performing a coordinate transformation together with a simple Ehlers transformation. In fact, given any colliding wave solution, one may use nothing more than a sequence of coordinate and Ehlers transformations to construct an infinite family of additional solutions.

We propose, therefore, to apply this iterative procedure to the two known colliding wave metrics with noncollinear polarization, the Nutku–Halil solution and the Chandrasekhar–Xanthopoulos solution. In this paper we shall demonstrate that this procedure actually does yield new colliding wave solutions by displaying and studying the first new metric so obtained.

In one or more sequels we shall adapt the Hauser–Ernst homogeneous Hilbert problem formalism to the colliding plane wave problem. Among our objectives are the following: (1) deriving a closed form expression for the whole family of solutions with noncollinear polarization corresponding to the Ferrari–Ibañez solutions; (2) identifying which generalized Kinnersley–Chitre transformations produce bona fide colliding wave solutions, and which do not; and (3) identifying some examples of colliding wave solutions that result from something other than a single double-Harrison transformation, together with gauge transformations, on a Kasner metric.

It is our belief that the experience gained through this effort will lead to a better understanding of the interrelationships among colliding wave solutions, and it is our hope that it may also suggest some way to tackle the general "initial value problem" in which data is specified on null surfaces. Xanthopoulos has already solved the corresponding initial value problem for colliding waves with collinear polarization.[13]

II. FERRARI–IBAÑEZ METRICS

In the interaction region the Ferrari–Ibañez metrics can be expressed in the form

[a)] This research was carried out while this author was at the Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional in Mexico City.

$$ds^2 = -\rho^{(n^2-1)/2}(1-x)^{1+n}(1+x)^{1-n}$$
$$\times\{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\}$$
$$+\rho^{1+n}[(1-x)/(1+x)](dx^1)^2$$
$$+\rho^{1-n}[(1+x)/(1-x)](dx^2)^2. \quad (2.1)$$

The determinant

$$\rho^2 := \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix}$$

has the explicit form

$$\rho^2 = (1-x^2)(1-y^2) \quad (2.2)$$

in terms of the x,y coordinates. The well-known Khan–Penrose solution corresponds to the special case $n = 0$.

Metrics with positive and negative values of n are related by the substitution

$$x \to -x, \quad x^1 \to x^2, \quad x^2 \to -x^1. \quad (2.3)$$

More significantly, the metrics with parameter values n and $n' = n + 2$ are related by a coordinate transformation

$$x^1 \to x^{1'} = x^2, \quad x^2 \to x^{2'} = -x^1, \quad (2.4)$$

followed by a simple Ehlers transformation under which

$$g_{11} \to g'_{11} = \rho^4/(g_{11}), \quad g_{22} \to g'_{22} = 1/(g_{22}). \quad (2.5)$$

One may introduce conventional null coordinates u,v through the coordinate transformation

$$x = uV + vU, \quad y = uV - vU, \quad (2.6)$$

where

$$U := +[1-u^2]^{1/2}, \quad V := +[1-v^2]^{1/2}.$$

The region of interaction corresponds to

$$u > 0, \quad v > 0, \quad \rho = 1 - u^2 - v^2 > 0,$$

while the collision of the incident plane waves takes place at

$$u = v = 0.$$

III. A DERIVATION OF THE SIMPLEST NEW COLLIDING WAVE SOLUTION

Colliding plane wave metrics with noncollinear polarization can be described in terms of a complex Ernst potential,[14] say $\mathscr{E} = H_{22}$, the real part of which is $-g_{22}$.

Examples are known[15] of solutions of the vacuum Einstein equations that have singular metrics at $u = 0$ and $v = 0$, where all null tetrad components of the curvature tensor vanish. Such solutions only admit an interpretation in terms of the collision of plane waves, the interaction of which commenced an *infinite* time in the past, for the surfaces $u = 0$ $(v > 0)$ and $v = 0$ $(u > 0)$ are, in reality, at minus null infinity, where the initial data are prescribed.

Solutions of this type have not received much attention, and are generally regarded as things to be avoided as one searches for bona fide colliding plane wave solutions in which the interaction did not commence at an infinite time in the past. It is, therefore, noteworthy that the application of an Ehlers transformation

$$H_{22} \to H'_{22} = 1/(H_{22}) \quad (3.1)$$

to a bona fide colliding wave solution should *in general* give rise to a metric that also admits a bona fide colliding wave interpretation, since the Ehlers transformation cannot produce a metric that is singular at $u = 0$ or at $v = 0$ if that singular behavior was not present in the seed metric. This observation convinces us that there are an enormous number of colliding wave solutions that we are now in a position to derive.

Let us, however, focus our attention upon the solution that results from the application of transformation (3.1) to the Ernst potential of the Nutku–Halil solution. We shall show that when the complex potential H_{22} of the Nutku–Halil metric, the noncollinear generalization of the Khan–Penrose metric, is subjected to the coordinate transformation (2.4), followed by the simple Ehlers transformation (3.1), one obtains a new colliding wave metric, which is a noncollinear generalization of the $n = 2$ Ferrari–Ibañez metric. Repeated application of such coordinate and Ehlers transformations yields solutions corresponding to $n = 4$, 6, 8,... . Similarly, if one begins with the solution of Chandrasekhar and Xanthopoulos, such a procedure yields solutions corresponding to $n = 3$, 5, 7,... .

The interaction region of the Nutku–Halil metric can be expressed[16] in the form

$$ds^2 = -[N_0/(\rho)^{1/2}]\{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\}$$
$$+ [\rho/N_0]\{|1-\xi|^2(dx^1)^2$$
$$- 4\,\mathrm{Im}(\xi)dx^1\,dx^2 + |1+\xi|^2(dx^2)^2\}, \quad (3.2)$$

where

$$N_0 := 1 - \xi\xi^*, \quad \xi := px + iqy,$$

and

$$p^2 + q^2 = 1.$$

When $q \to 0$ and $p \to 1$, the metric (3.2) reduces to the $n = 0$ Ferrari–Ibañez metric. Both the metric (3.2) and the associated curvature tensor are free of singularities except where $\rho \to 0$.

In terms of the x,y coordinates the region of interaction of the colliding plane waves corresponds to

$$0 < x < 1, \quad -x < y < x, \quad \rho = XY > 0,$$

where x plays the role of a "time" coordinate, y plays the role of a "spatial" coordinate, $X := +[1-x^2]^{1/2}$, and $Y := +[1-y^2]^{1/2}$.

The introduction of symbols such as p, q, U, V, X, and Y is motivated to a large extent by the desire to deal with *rational* expressions, which can easily and accurately be manipulated on a microcomputer. It has also been our experience that those microcomputer calculations proceed more efficiently when the various fields are expressed in terms of x and y rather than in terms of the null coordinates u and v.

After the coordinate transformation (2.4) is applied to the metric (3.2), we obtain the following form of the Nutku–Halil metric:

$$ds^2 = -[N_0/(\rho)^{1/2}]\{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\}$$
$$+ [\rho/N_0]\{|1+\xi|^2(dx^1)^2$$
$$+ 4\,\mathrm{Im}(\xi)dx^1\,dx^2 + |1-\xi|^2(dx^2)^2\}. \quad (3.3)$$

The Ernst potential $\mathscr{E} := H_{22}$ of the metric (3.3) is found to be given, *up to an arbitrary constant imaginary term*, by

$$\mathscr{E} = XY - 2(pY - iqX - xY)/(pX - iqY). \qquad (3.4)$$

The real part of $\mathscr{E}$ yields $f := -g_{22}$, while from the imaginary part of $\mathscr{E}$, the twist potential ϕ, one may evaluate the ratio $\omega := g_{21}/g_{22}$ by integrating the well-known[14] equation

$$d\omega = \rho f^{-2} * d\phi, \qquad (3.5)$$

where $*$ is a two-dimensional duality operator with the properties

$$*du = du, \quad *dv = -dv.$$

The g_{11} component of the metric tensor is obtained using the relation

$$g_{11} = (\rho^2 + g_{21}^2)/g_{22}, \qquad (3.6)$$

while the determination of the rest of the metric involves an additional integration.[14] From the Ernst potential $\mathscr{E}$ displayed above one can easily regenerate the Nutku–Halil metric (3.3).

It is to the Ernst potential (3.4) that we applied the Ehlers transformation (3.1). The real part of the new H_{22} yields the new $-g_{22}$, while from the imaginary part the new g_{21}/g_{22} may be obtained. The new g_{11} is obtained using Eq. (3.6), and the coefficient of

$$\{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\}$$

is obtained by noting that the product of this coefficient with g_{22} is invariant under the Ehlers transformation (3.1). The resulting metric has the form

$$ds^2 = -[N/(\rho)^{1/2}]\{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\} + [\rho/N]\{K(dx^1)^2 + 2L\,dx^1\,dx^2 + M(dx^2)^2\}, \qquad (3.7)$$

where K, L, M, and N are polynomials in p, q, x, and y. Introducing the fields

$$A = 2[(1-2px+x^2)(1-y^2) - 2iqy(1-x^2)] - B\rho^2 \qquad (3.8a)$$

and

$$B = 1 - \xi = 1 - px - iqy, \qquad (3.8b)$$

we may express K, L, M, and N in the following form:

$$K = |A|^2, \qquad (3.9a)$$

$$L = \mathrm{Im}(AB^*) = -2qy[\rho^2 - 2(1-px)(x^2-y^2)], \qquad (3.9b)$$

$$M = |B|^2, \qquad (3.9c)$$

and

$$N = \mathrm{Re}(AB^*) = [1 - 4px + 6x^2 - 4px^3 + x^4](1-y^2) + q^2(x^2-y^2)(\rho^2-4). \qquad (3.9d)$$

It is important to note that the last quantity, which plays an important role in the Weyl conform tensor as well as in the metric, vanishes nowhere within the range of coordinates being considered.

IV. COLLINEAR LIMIT AND FURTHER GENERALIZATION OF THE $n=2$ METRIC

When $q \to 0$, the metric (3.7) reduces to

$$ds^2 = -\rho^{3/2}[(1+x)^3/(1-x)] \times \{(dx)^2/(1-x^2) - (dy)^2/(1-y^2)\} + \rho^{-1}[(1-x)/(1+x)](dx^1)^2 + \rho^3[(1+x)/(1-x)](dx^2)^2, \qquad (4.1)$$

which is the $n=2$ solution of Ferrari and Ibañez.

The Ernst potential (3.4) from which we started may be generalized in a number of ways. We have actually considered a generalization involving two additional parameters, one of which corresponds to the inclusion of an arbitrary imaginary constant term in (3.4), and the other of which results from rotating the x^1, x^2 coordinate axes through an arbitrary fixed angle. These extra parameters are, of course, not essential for the Nutku–Halil solution, but we believe they are essential for the solution which results from the application of the Ehlers transformation (3.1). While we have not completed the study of the Weyl tensor for the resulting three-parameter $n=2$ solution, we believe for reasons already stated that solution will, in fact, admit a bona fide colliding wave interpretation. In this paper we shall present the detailed analysis only for the one-parameter $n=2$ solution (3.7).

V. INTERPRETATION OF THE SOLUTION

Let H denote the Heaviside step function. In order to interpret the metric of Eq. (3.7) as a colliding gravitational wave solution we replace u by $uH(u)$ and v by $vH(v)$ in the definitions of x and y as well as in ρ, and we permit the null coordinates u and v to assume negative values.

We identify four regions of the space-time:

region I	$u<0$	$v<0$
region II	$u<0$	$0<v<1$
region III	$0<u<1$	$v<0$
region IV	$u>0$	$v>0$

$$(\rho := 1 - u^2 - v^2 > 0).$$

The analysis begins with region IV, the region of interaction. Because of the symmetry mentioned earlier, it suffices to check the junction conditions across the null surfaces that separate region IV from region II and region III from region I, i.e., across the null surface $u=0$.

The metric is obviously continuous across $u=0$, though not smooth. The first derivative of $uH(u)$ with respect to u is $H(u)$, while the second derivative is a Dirac delta function, which may appear in certain components of the curvature tensor. The metric in region II is independent of u and has the form of a Petrov type N plane wave solution, while the metric in region III is independent of v and also has the form of a Petrov type N plane wave solution. The metric in region I is independent of both u and v and is a portion of Minkowski space.

The evaluation of the curvature tensor[9] of the new space-time is facilitated by the introduction of a null tetrad $\{k,m,t,t^*\}$, where

$$k = -[2N/(\rho)^{1/2}]^{1/2}(dv/V)$$
$$= -[N/2(\rho)^{1/2}]^{1/2}[(dx/X) - (dy/Y)],$$
$$m = [2N/(\rho)^{1/2}]^{1/2}(du/U) \tag{5.1}$$
$$= [N/2(\rho)^{1/2}]^{1/2}[(dx/X) + (dy/Y)],$$
$$t = [\rho/(2N)]^{1/2}(A\,dx^1 + iB\,dx^2).$$

With this choice of null tetrad the only nonvanishing components of the Weyl conform tensor are C_2, C_0, and C_{-2}, where the index denotes the spin weight of the component. (Our C_2 corresponds to the Newman – Penrose Ψ_0, while our C_0 and C_{-2} correspond to Ψ_2 and Ψ_4, respectively.)

Formulas for the null tetrad components of the Weyl conform tensor and the Ricci tensor can be found in many places, including some of our early papers.[14,17] For the present studies we found it somewhat more convenient to use formulas in terms of the fields A and B, which were developed more recently by Ernst.[9] All our calculations were checked using the Grad Student: Rational Calculator symbolic manipulation program.[18]

Our tetrad (5.1) is substantially identical to the tetrad of Ref. 9. The one-forms α and β of Ref. 9 are given by

$$\alpha := (B^*\,dA + A^*\,dB)/(2N) =: (du/U)\alpha_u + (dv/V)\alpha_v,$$
$$\beta := (A\,dB - B\,dA)/(2N) =: (du/U)\beta_u + (dv/V)\beta_v. \tag{5.2}$$

We may infer from Ref. 9 that the null tetrad components of the reduced Ricci tensor

$$S_{ij} := R_{ij} - \tfrac{1}{4}Rg_{ij}$$

and the Ricci scalar R are given by

$$S_{tt^*} + R/4 = -[(U\partial_u)(V\partial_v)(\rho)]/(2\rho^{1/2}N),$$
$$S_{tt} = [(U\partial_u)(\rho\beta_v) + (V\partial_v)(\rho\beta_u)$$
$$- (\alpha_u - \alpha_u^*)(\rho\beta_v) - (\alpha_v - \alpha_v^*)(\rho\beta_u)]/(2\rho^{1/2}N),$$
$$S_{t^*t^*} = (S_{tt})^*, \tag{5.3}$$
$$S_{kk} = [(U\partial_u)^2(\rho) - (\alpha_u + \alpha_u^*)(U\partial_u)(\rho)$$
$$+ 2\rho\beta_u^*\beta_u]/(2\rho^{1/2}N),$$
$$S_{mm} = [(V\partial_v)^2(\rho) - (\alpha_v + \alpha_v^*)(V\partial_v)(\rho)$$
$$+ 2\rho\beta_v^*\beta_v]/(2\rho^{1/2}N),$$
$$S_{tt^*} - R/12$$
$$= (\rho^{1/2}/6N)[(U\partial_u)(\alpha_v + \alpha_v^*) + (v\partial_v)(\alpha_u + \alpha_u^*)$$
$$+ 2\beta_u\beta_v^* + 2\beta_v\beta_u^*],$$

and

$$S_{km} = -S_{tt^*}.$$

The only null tetrad component of the reduced Ricci tensor that involves second derivatives with respect to u is the component S_{kk}, and this only involves the quantity $\rho_{,u,u}$, which has no delta function contribution on the null surface $u = 0$. Similarly, the only null tetrad component of the reduced Ricci tensor that involves second derivatives with respect to v is the component S_{mm}, and this only involves the quantity $\rho_{,v,v}$. Therefore, since the vacuum field equations are satisfied in all four regions I, II, III, and IV, and there is no delta function contribution either at $u = 0$ or at $v = 0$, the vacuum field equations are, in fact, satisfied everywhere.

The curious fact that the field equation $S_{tt} = 0$ takes the form of an Ernst equation,

$$(\mathrm{Re}\,E)\{[(1-x^2)E_{,x}]_{,x} - [(1-y^2)E_{,y}]_{,y}\}$$
$$= (1-x^2)(E_{,x})^2 - (1-y^2)(E_{,y})^2,$$

for the complex potential $E := A/B$ was observed by Chandrasekhar. In our case the complex potential E has the form

$$E = \frac{2[(1-2px+x^2)(1-y^2) - 2iqy(1-x^2)]}{(1-px-iqy)} - \rho^2. \tag{5.4}$$

The nonvanishing null tetrad components of the Weyl tensor are given by

$$C_2 = [-\rho(U\partial_u)\beta_u - \tfrac{3}{2}\beta_u(U\partial_u)(\rho)$$
$$+ 2\rho\alpha_u\beta_u]/(2\rho^{1/2}N),$$
$$C_0 - R/24 - \tfrac{1}{2}S_{tt^*} = -\{\rho^2\beta_u\beta_v^* - \tfrac{1}{4}[(U\partial_u)(\rho)] \tag{5.5}$$
$$\times[(V\partial_v)(\rho)]\}/(2\rho^{3/2}N),$$

and

$$C_{-2} = [-\rho(V\partial_v)\beta_v^* - \tfrac{3}{2}(V\partial_v)(\rho)\beta_v^*$$
$$+ 2\rho\alpha_v^*\beta_v^*]/(2\rho^{1/2}N).$$

Alternatively, one might prefer to follow the example of what was done[19] in order to compute the Weyl conform tensor of the Tomimatsu–Sato solutions. In that case, once C_0 has been evaluated, one proceeds to calculate C_2 and C_{-2} using the vacuum Bianchi identities,

$$\begin{aligned}\beta_v^*(\rho^{3/2}C_2) &= (U\partial_u)(\rho^{3/2}C_0),\\ \beta_u(\rho^{3/2}C_{-2}) &= (V\partial_v)(\rho^{3/2}C_0),\end{aligned} \tag{5.6}$$

where the requirement of exact divisibility of the right-hand side by β_v^* or by β_u, respectively, serves as an excellent check upon the accuracy of one's calculation.

Our results for the Weyl conform tensor components C_2, C_0, and C_{-2} in region IV are contained in the Appendix to this paper. The denominators of C_2, C_0, and C_{-2} consist of a factor N^3 times ρ raised to a power. Since N nowhere vanishes within the range of coordinates being considered, the only curvature singularities are located where $\rho \to 0$, i.e., on the surface where $u^2 + v^2 = 1$.

In region II, where one has a Petrov type N plane wave solution with no u dependence, $C_2 = C_0 = 0$, while in region III, where one has a Petrov type N plane wave solution with no v dependence, $C_0 = C_{-2} = 0$. While C_2 is continuous across the surface $v = 0$ separating region III from region IV, there is a step discontinuity in C_0 and C_{-2}. Similarly there is a step discontinuity in C_0 and C_2 at the null surface $u = 0$ which separates region II from region IV. In addition to these step discontinuities, there is a $\delta(u)$ term in C_2 and a $\delta(v)$ term in C_{-2}. Because our solution is not flat in regions II and III, it involves gravitational shock waves as well as impulsive waves. At this time the only known solutions that are flat in regions II and III are the Khan–Penrose solution and its Nutku–Halil generalization!

Note added in proof

After this work was completed we learned that Ferrari, Ibañez, and Bruni[20] subsequently succeeded in obtaining closed form metrical expressions representing colliding waves with noncollinear polarization and *arbitrary n* value. We in turn have succeeded in generalizing their result, adding one more adjustable parameter. Our solution, which will be described in detail later, has the E potential $E = \rho^n a/b$, where

$$a = X\{(p+p')[(1+x)/(1-x)]^{-(n+1)/2} - (p-p')[(1+x)/(1-x)]^{(n+1)/2}\} + iY\{(q+q')[(1+y)/(1-y)]^{-(n+1)/2} - (q-q')[(1+y)/(1-y)]^{(n+1)/2}\},$$

$$b = X\{(p+p')[(1+x)/(1-x)]^{-(n-1)/2} - (p-p')[(1+x)/(1-x)]^{(n-1)/2}\} - iY\{(q+q')[(1+y)/(1-y)]^{-(n-1)/2} - (q-q')[(1+y)/(1-y)]^{(n-1)/2}\},$$

where $p^2 + q^2 = 1$ and $p'^2 + q'^2 = 1$. When q' is replaced by 0 and p' by 1, one obtains solutions that are equivalent, after correcting misprints in their preprint, to those found by Ferrari, Ibañez, and Bruni.

In the case $n = 0$ the additional parameter q' is inessential and one simply obtains the Nutku–Halil solution. In the case $n = 1$ one obtains both the Chandrasekhar–Xanthopoulos solution[6] and the Ferrari–Ibañez–Bruni $n = 1$ solutions as the respective special cases $(q = 0,\ p = 1)$ and $(q' = 0, p' = 1)$ of the more general Kerr–NUT solution. In the case $n = 2$ one obtains a two-parameter generalization of the solution described in this paper.

ACKNOWLEDGMENTS

F. J. Ernst would like to take this opportunity to thank Professor S. Chandrasekhar for drawing his attention to the fascinating problem of finding new global solutions of the colliding impulsive gravitational plane wave problem. At this time the authors would like to dedicate the first member of the series of new solutions to Professor Chandrasekhar.

One of the authors (FJE) was supported in part by Grant No. PHY-8605958 from the National Science Foundation.

APPENDIX: WEYL CONFORM TENSOR

The evaluation of the Weyl conform tensor in region IV begins with the observation that

$$W_2 := -(2N)^3\rho^{1/2}C_2 = 2\rho[NU(2N\beta_u)_{,u} - UN_{,u}(2N\beta_u) - (2N\alpha_u)(2N\beta_u)] + 3NU\rho_{,u}(2N\beta_u),$$

$$W_0 := -(2\rho^{1/2}N)^3C_0 = \rho^2(2N\beta_u)(2N\beta_v^*) + N^2(x^2-y^2),$$

$$W_{-2} := -(2N)^3\rho^{1/2}C_{-2} = 2\rho[NV(2N\beta_v)_{,v} - VN_{,v}(2N\beta_v) - (2N\alpha_v)(2N\beta_v)] + 3NV\rho_{,v}(2N\beta_v),$$

where

$$2N\beta = A\,dB - B\,dA,$$

$$2N\alpha = A^*\,dB + B^*\,dA,$$

$$A = 2[(1-2px+x^2)(1-y^2) - 2iqy(1-x^2)] - B\rho^2,$$

$$B = 1 - px - iqy,$$

and

$$N = \mathrm{Re}(AB^*) = [1 - 4px + 6x^2 - 4px^3 + x^4](1-y^2) + q^2(x^2-y^2)(\rho^2-4).$$

Because of the similarity of C_{-2} and C_2, one may infer the value of C_{-2} from that of C_2.

Using the Grad Student: Rational Calculator symbolic manipulation program on a 10 MHz 8086 microcomputer, we evaluated the fields W_0 and W_2. Anticipating that we might encounter problems due to the very limited amount of RAM; namely, 360 K bytes, we forced all fields into a format in which they were homogeneous in the auxiliary parameters p_+ and p_-, where

$$1 = p_+^2 + p_-^2, \quad p = p_+^2 - p_-^2, \quad q = 2p_+p_-.$$

In particular, the fields W_2, W_0, and W_{-2} are homogeneous of the eighth degree in p_+,p_-, while the field N is homogeneous of the fourth degree in p_+,p_-. The evaluation of W_0 required 9 min, while the evaluation of W_2 required 34 min. It turned out that no difficulty was experienced due to the limited RAM, although we suspect that the expressions would not have to be much more complicated before such difficulties would be encountered in the evaluation of C_2. The program is capable of exploiting up to 512 K of RAM, but eventually one would be forced to consider the terms of different degree in p_+ and p_- separately, if one wished to continue to perform the calculations on a microcomputer.

Introducing $\rho = XY$, X and Y as much as possible, we found that the expressions for W_0 and W_2 assume the following forms:

$$W_0 = [\rho^6(-11X^2 - 72x + 224) + \rho^4(-X^6 - 8X^4x + 48X^4 + 176X^2x - 448X^2 + 400Y^2x - 576Y^2 - 704x + 1024) + \rho^2(-320Y^4x + 256Y^4 + 640Y^2x - 640Y^2) - 128Y^6x + 128Y^6]p_+^8 - [\rho^7(48i) + \rho^5(32iX^2x - 256iX^2 + 192iY^2x - 384iY^2 - 768ix + 1408i) + \rho^3(-64iX^4x + 320iX^4 + 768iX^2x - 1280iX^2 - 384iY^4x + 384iY^4 + 1280iY^2x - 1280iY^2 - 1024ix + 1024i)]p_+^7p_-$$

$$+[\rho^6(-12X^2+56Y^2-16x-128)+\rho^4(4X^6+16X^4x+32X^2x+64X^2+96Y^4x-64Y^4-224Y^2x$$
$$+96Y^2+128x)+\rho^2(32X^6+128X^4x-256X^4-256X^2x+256X^2+64Y^6x-64Y^6-256Y^4x$$
$$+256Y^4+256Y^2x-256Y^2)]p_+^6p_-^2 \;-[\rho^7(48i)+\rho^5(-32iX^2x-256iX^2+192iY^2x-384iY^2$$
$$-768ix+1408i)+\rho^3(64iX^4x+320iX^4+768iX^2x-1280iX^2-384iY^4x+384iY^4$$
$$+1280iY^2x-1280iY^2-1024ix+1024i)]p_+^5p_-^3 \;+[\rho^6(46X^2+64Y^2-704)$$
$$+\rho^4(-6X^6-96X^4+1408X^2+16Y^6-128Y^4+960Y^2-2048)$$
$$+\rho^2(-64X^6-512X^4+1280X^2)-256X^6]p_+^4p_-^4$$
$$-[\rho^7(-48i)+\rho^5(-32iX^2x+256iX^2+192iY^2x+384iY^2-768ix-1408i)+\rho^3(64iX^4x-320iX^4$$
$$+768iX^2x+1280iX^2-384iY^4x-384iY^4+1280iY^2x+1280iY^2-1024ix-1024i)]p_+^3p_-^5$$
$$+[\rho^6(-12X^2+56Y^2+16x-128)+\rho^4(4X^6-16X^4x-32X^2x+64X^2-96Y^4x-64Y^4+224Y^2x$$
$$+96Y^2-128x)+\rho^2(32X^6-128X^4x-256X^4+256X^2x+256X^2-64Y^6x$$
$$-64Y^6+256Y^4x+256Y^4-256Y^2x-256Y^2)]p_+^2p_-^6 \;-[\rho^7(-48i)+\rho^5(32iX^2x+256iX^2$$
$$+192iY^2x+384iY^2-768ix-1408i)+\rho^3(-64iX^4x-320iX^4+768iX^2x$$
$$+1280iX^2-384iY^4x-384iY^4+1280iY^2x+1280iY^2-1024ix-1024i)]p_+p_-^7$$
$$+[\rho^6(-11X^2+72x+224)+\rho^4(-X^6+8X^4x+48X^4-176X^2x-448X^2-400Y^2x$$
$$-576Y^2+704x+1024)+\rho^2(320Y^4x+256Y^4-640Y^2x-640Y^2)+128Y^6x+128Y^6]p_-^8,$$

$$W_2=[\rho^5(-6X^2x+56X^2+212x-508)+\rho^4(-6X^4y-56X^2xy+184X^2y+228xy-256y)$$
$$+\rho^3(4X^4x-60X^4-288X^2x+800X^2-736Y^2x+1024Y^2+1344x-1984)$$
$$+\rho^2(256Y^2xy-640Y^2y)+\rho(576Y^4x-576Y^4-1280Y^2x+1280Y^2)-768Y^4xy+768Y^4y]p_+^8$$
$$+[\rho^6(36ix-160i)+\rho^5(-60iX^2y-224ixy+336iy)$$
$$+\rho^4(16iX^4+168iX^2x-560iX^2-376iY^2x+680iY^2-976ix+1264i)$$
$$+\rho^3(-48iX^4xy+416iX^4y+1248iX^2xy-2208iX^2y+32iY^2xy+352iY^2y-1792ixy+1536iy)$$
$$+\rho^2(-8iX^6-176iX^4x+784iX^4+1792iX^2x-2944iX^2+704iY^4x-896iY^4+384iY^2x-2304ix+2304i)$$
$$+\rho(768iY^4xy-768iY^4y-512iY^2xy+512iY^2y)-384iY^6x+384iY^6+768iY^4x-768iY^4]p_+^7p_-$$
$$+[\rho^6(168y)+\rho^5(-84X^2x+64X^2-24Y^2x+192Y^2-24x-432)$$
$$+\rho^4(-84X^4y-192X^2xy-48X^2y+256Y^2xy-96Y^2y+96xy-1248y)$$
$$+\rho^3(56X^4x+256X^4+1184X^2x-1856X^2+400Y^4x-528Y^4-1312Y^2x+2112Y^2-896x)$$
$$+\rho^2(-288X^4xy+1440X^4y+2304X^2xy-2304X^2y+320Y^4xy-320Y^4y-2304Y^2xy+2304Y^2y)$$
$$+\rho(-128X^6-704X^4x+1600X^4+1792X^2x-1792X^2-256Y^6x+256Y^6$$
$$+1600Y^4x-1600Y^4-1792Y^2x+1792Y^2)]p_+^6p_-^2$$
$$+[\rho^6(-228ix-160i)+\rho^5(-132iX^2y+144iY^2y-224ixy+688iy)$$
$$+\rho^4(208iX^4+472iX^2x-16iX^2+144iY^4x-192iY^4-488iY^2x+136iY^2+208ix+1264i)$$
$$+\rho^3(-112iX^4xy+96iX^4y+352iX^2xy-1312iX^2y+64iY^4xy-64iY^4y+928iY^2xy$$
$$-928iY^2y-1792ixy+1792iy)+\rho^2(-104iX^6-80iX^4x-816iX^4-1152iX^2x$$
$$-96iY^6x+96iY^6-704iY^4x+704iY^4+2944iY^2x-2944iY^2-2304ix+2304i)$$
$$+\rho(768iX^4xy-1024iX^4y-512iX^2xy+512iX^2y)+384iX^6+768iX^4x-768iX^4]p_+^5p_-^3$$
$$+[\rho^5(-240X^2+640Y^2+152)+\rho^4(496X^2xy-256Y^2xy-384xy)+\rho^3(-136X^4-1984X^2$$
$$-288Y^4-1152Y^2+3968)+\rho^2(64X^4xy-512X^2xy)+\rho(256X^6+1152X^4-2560X^2)$$
$$+1536X^4xy]p_+^4p_-^4 \;+[\rho^6(-228ix+160i)+\rho^5(-132iX^2y+144iY^2y+224ixy+688iy)$$

$$
\begin{aligned}
&+\rho^4(-208iX^4+472iX^2x+16iX^2+144iY^4x+192iY^4-488iY^2x-136iY^2+208ix-1264i)\\
&+\rho^3(112iX^4xy+96iX^4y-352iX^2xy-1312iX^2y-64iY^4xy-64iY^4y-928iY^2xy-928iY^2y+1792ixy\\
&+1792iy)+\rho^2(104iX^6-80iX^4x+816iX^4-1152iX^2x-96iY^6x-96iY^6-704iY^4x-704iY^4\\
&+2944iY^2x+2944iY^2-2304ix-2304i)+\rho(-768iX^4xy-1024iX^4y+512iX^2xy+512iX^2y)\\
&-384iX^6+768iX^4x+768iX^4]p_+^3p_-^5\\
&+[\rho^6(-168y)+\rho^5(84X^2x+64X^2+24Y^2x+192Y^2+24x-432)\\
&+\rho^4(84X^4y-192X^2xy+48X^2y+256Y^2xy+96Y^2y+96xy+1248y)\\
&+\rho^3(-56X^4x+256X^4-1184X^2x-1856X^2-400Y^4x-528Y^4+1312Y^2x+2112Y^2+896x)\\
&+\rho^2(-288X^4xy-1440X^4y+2304X^2xy+2304X^2y+320Y^4xy+320Y^4y-2304Y^2xy-2304Y^2y)\\
&+\rho(-128X^6+704X^4x+1600X^4-1792X^2x-1792X^2+256Y^6x+256Y^6\\
&-1600Y^4x-1600Y^4+1792Y^2x+1792Y^2)]p_+^2p_-^6+[\rho^6(36ix+160i)+\rho^5(-60iX^2y+224ixy+336iy)\\
&+\rho^4(-16iX^4+168iX^2x+560iX^2-376iY^2x-680iY^2-976ix-1264i)\\
&+\rho^3(48iX^4xy+416iX^4y-1248iX^2xy-2208iX^2y-32iY^2xy+352iY^2y+1792ixy+1536iy)\\
&+\rho^2(8iX^6-176iX^4x-784iX^4+1792iX^2x+2944iX^2+704iY^4x+896iY^4+384iY^2x-2304ix-2304i)\\
&+\rho(-768iY^4xy-768iY^4y+512iY^2xy+512iY^2y)-384iY^6x-384iY^6+768iY^4x+768iY^4]p_+p_-^7\\
&+[\rho^5(6X^2x+56X^2-212x-508)+\rho^4(6X^4y-56X^2xy-184X^2y+288xy+256y)\\
&+\rho^3(-4X^4x-60X^4+288X^2x+800X^2+736Y^2x+1024Y^2-1344x-1984)+\rho^2(256Y^2xy+640Y^2y)\\
&+\rho(-576Y^4x-576Y^4+1280Y^2x+1280Y^2)-768Y^4xy-768Y^4y]p_-^8\,.
\end{aligned}
$$

At the present time we do not have a simple elegant representation of the expressions W_2, W_0, and W_{-2}. We believe that it is easier to avoid the introduction of errors if complicated algebraic expressions are transmitted by electronic means. This is likely to become an increasingly important consideration as we obtain ever more involved Weyl tensor results in the future. We propose, therefore, to make available to people who are engaged in research in this field a $5\frac{1}{4}$ in. 360 K floppy disk upon which are recorded results obtained directly from the Grad Student: Rational Calculator program. The results will be in standard ASCII files, which can be read and manipulated using a LISP interpreter and one's own program, or which can be read by the Grad Student: Rational Calculator program and subjected to additional analysis.

[1]P. Szekeres, J. Math. Phys. **13**, 286 (1970).
[2]K. Khan and R. Penrose, Nature **229**, 1575 (1971).
[3]V. Ferrari and J. Ibañez, Gen. Relativ. Gravit. **19**, 405 (1987).
[4]V. A. Belinskii and V. E. Zakharov, Sov. Phys. JETP **50**, 1 (1979).
[5]Y. Nutku and M. Halil, Phys. Rev. Lett. **39**, 1379 (1977).
[6]S. Chandrasekhar and B. C. Xanthopoulos, Proc. R. Soc. London Ser. A **408**, 175 (1986).
[7]R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[8]R. A. Matzner and F. J. Tipler, Phys. Rev. D **29**, 1575 (1984).
[9]F. J. Ernst, "Derivation of Nutku–Halil colliding plane wave solution from isotropic Kasner metric using double-Harrison transformation," in *Gravitational Collapse and Relativity*, edited by H. Sato and T. Nakamura (World Scientific, Singapore, 1987), pp. 141–149.
[10]The double-Harrison transformation which we employed is described by I. Hauser, "On the homogeneous Hilbert problem for effecting Kinnersley–Chitre transformation," in *Solutions of Einstein's Equations: Techniques and Results, Lecture Notes in Physics*, Vol. 205, edited by C. Hoenselaers and W. Dietz (Springer, Berlin, 1984), pp. 128–175.
[11]I. Hauser and F. J. Ernst, J. Math. Phys. **21**, 1126 (1980).
[12]W. Kinnersley and D. M. Chitre, J. Math. Phys. **18**, 1538 (1977); **19**, 1926, 2037 (1978).
[13]B. C. Xanthopoulos, J. Math. Phys. **27**, 2129 (1986).
[14]F. J. Ernst, J. Math. Phys. **15**, 1409 (1974).
[15]We are thinking, in particular, of a solution that can be obtained from the Tomimatsu–Sato solution E of the Ernst equation. See Ref. 6.
[16]S. Chandrasekhar and V. Ferrari, Proc. R. Soc. London Ser. A **396**, 55 (1984).
[17]I. Hauser and F. J. Ernst, J. Math. Phys. **20**, 1041 (1979).
[18]*The Grad Student: Rational Calculator* is a computer program that can be run on any MS-DOS computer such as a Zenith Z-100 or Z-200 or IBM PC or AT. Inquiries should be directed to F. J. Ernst.
[19]J. E. Economou and F. J. Ernst, J. Math. Phys. **17**, 52 (1976).
[20]V. Ferrari, J. Ibañez, and M. Bruni, "Colliding gravitational waves with noncollinear polarization: A class of soliton solutions," preprint, 1987.

Solutions of Einstein's equations for the interior of a stationary axisymmetric perfect fluid

E. Kyriakopoulos
National Technical University of Athens, Physics Laboratory A, Zografou Campus, GR 157 72, Athens, Greece

(Received 19 November 1985; accepted for publication 20 May 1987)

The problem of finding a class of solutions of Einstein's field equations for the interior of a uniformly rotating, axisymmetric perfect fluid is reduced to the integration of a known function. A number of solutions of the class is given explicitly. The equipressure surfaces of all solutions of the class are planes. The class contains solutions with the equation of state $p = \gamma\mu$, $\gamma > 0$, $0 > \gamma \geqslant -1 + 2/\sqrt{5}$, and $\gamma \leqslant -1 - 2/\sqrt{5}$, which are given explicitly.

I. INTRODUCTION

By introducing a new system of coordinates Bonanos and Sklavenitis[1] were able to reduce Einstein's field equations for the interior of a uniformly rotating axisymmetric perfect fluid to a system of six partial differential equations, four of which involve only first derivatives of the metric functions. Four of the six equations are independent. Then they reduced the problem to two second-order partial differential equations. Solutions of Einstein's equations with vanishing magnetic Weyl tensor, whose existence had been proved before,[2] were found explicitly as solutions of these two equations.[3]

In this paper it is shown that we can get a wide class of solutions of the original system of six equations, which satisfy the boundary conditions, by integrating a known function. This is done in Sec. II where it is also shown that for all solutions of the class fluid's equipressure surfaces are planes. Therefore the solutions do not represent the gravitational field of isolated rotating masses.

In Sec. III we derive explicitly a number of solutions of the class. Since the equations of state of these solutions are in one case simple but not realistic and in all other cases complicated we derive in Sec. IV by a different method a new solution whose equation of state is physically interesting. Finally in Sec. IV by calculating the Killing vectors of one solution of the class we find that the axial symmetry is the highest symmetry. In general we expect this to be the case for the solutions of the class. We should mention that the number of known axially symmetric perfect fluid solutions is limited.[4,5]

II. THE CLASS OF SOLUTIONS

To derive our class of solutions we shall follow the notation of Ref. 1. In this work the line element ds^2 for stationary axially symmetric space-time was written in the form

$$ds^2 = \frac{1}{h^2}(dt + A\,d\varphi)^2 - h^2\rho^2\,d\varphi^2 - h^2\Delta^2\left\{P\left(\frac{dh}{h}\right)^2 + Q\,d\rho^2 + 2R\,\frac{dh}{h}\,d\rho\right\}, \quad (1)$$

where

$$\Delta^2 = 1/(PQ - R^2) \quad (2)$$

and t, φ, ρ, and h are the coordinates of the problem. Also in this work the pressure p and the energy density μ, which appear in the energy-momentum tensor

$$T^{ab} = (\mu + p)u^a u^b - pg^{ab} \quad (3)$$

of a perfect fluid, were parametrized as follows:

$$8\pi p = \frac{f}{h^2}, \quad 8\pi(\mu + p) = \frac{f_h}{h} - \frac{2f}{h^2}, \quad (4)$$

where $f = f(h)$ and $f_h = df/dh$. Generally in the paper we shall use the notation $G_x = \partial G/\partial x$ for every coordinate x and every function G. The field equations lead to a system of six equations involving the functions P, Q, R, Δ, A, and f. A solution of the four first-order equations satisfies the other two. We shall find solutions of a system of four equations three of which are first order and the other of second order and then we shall show that these solutions satisfy the remaining first-order equation. We shall choose the following set of four equations[1]:

$$(1/\rho\Delta)\{(P\Delta)_\rho - h(R\Delta)_h\} - 2f = 0, \quad (5)$$

$$\{(\Delta/\rho h^4)(PA_\rho - RhA_h)\}_\rho + h\{(\Delta/\rho h^4)(QhA_h - RA_\rho)\}_h = 0, \quad (6)$$

$$R = \left(\frac{hP_h}{4\rho} - P\frac{A_\rho A_h}{4\rho^2 h^3}\right)\left[1 - \left(\frac{A_h}{2\rho h}\right)^2\right]^{-1}, \quad (7)$$

$$Q = \left[\frac{P_\rho}{2\rho} - f - P\left(\frac{A_\rho}{2\rho h^2}\right)^2\right]\left[1 - \left(\frac{A_h}{2\rho h}\right)^2\right]^{-1}. \quad (8)$$

We shall find solutions of the system of equations (5)–(8) when Δ is independent of ρ, that is, when

$$\Delta = \Delta(h). \quad (9)$$

We define ξ and z by the relations

$$\xi = \rho^2, \quad z = h^2. \quad (10)$$

Then since $f = f(z)$ Eq. (5) becomes

$$(P\Delta - \xi\Delta f)_\xi - z(R\Delta/\sqrt{\xi})_z = 0. \quad (11)$$

Therefore we get

$$P\Delta - \xi\Delta f = z\Gamma_z, \quad (12)$$

$$R\Delta/\sqrt{\xi} = \Gamma_\xi, \quad (13)$$

where Γ is an arbitrary function of ξ and z. Also writing Eq. (6) in the form

 0022-2488/87/092162-05$02.50

$$\left(\frac{\Delta P A_\xi}{z^2} - \frac{\Delta R A_z}{z\sqrt{\xi}}\right)_\xi + z\left(\frac{\Delta Q A_z}{z\xi} - \frac{\Delta R A_\xi}{z^2\sqrt{\xi}}\right)_z = 0, \quad (14)$$

we see immediately that we must have

$$\Delta P(A_\xi/z) - \Delta R(A_z/\sqrt{\xi}) = z^2 H_z, \quad (15)$$

$$\Delta R(A_\xi/z) - \Delta Q(A_z/\sqrt{\xi}) = z\sqrt{\xi}\, H_\xi, \quad (16)$$

where H is an arbitrary function of ξ and z. Solving the above system for A_ξ/z and $A_z/\sqrt{\xi}$ we get

$$A_\xi/z = z\Delta(zQH_z - \sqrt{\xi}RH_\xi), \quad (17)$$

$$A_z/\sqrt{\xi} = z\Delta(zRH_z - \sqrt{\xi}PH_\xi). \quad (18)$$

Expressing R and Q of Eqs. (7) and (8) in terms of ξ and z and subsequently using the above relations to eliminate A_ξ and A_z we get

$$R = (zP_z/2\sqrt{\xi} + z^3\sqrt{\xi}PH_\xi H_z)/(1 + z^4 H_z^2), \quad (19)$$

$$Q = (P_\xi - f + z^2\xi PH_\xi^2)/(1 + z^4 H_z^2). \quad (20)$$

Also from Eqs. (2) and (20) we get

$$\Delta^2 P(P_\xi - f + z^2\xi PH_\xi^2) - (\Delta^2 R^2 + 1)(1 + z^4 H_z^2) = 0. \quad (21)$$

Finally the expressions A_ξ and A_z of Eqs. (17) and (18) must satisfy the relation

$$A_{\xi z} = A_{z\xi}. \quad (22)$$

To find solutions of the system of Eqs. (12), (13), and (17)–(22) we make the ansatz

$$z^2 H_z = h_1(z), \quad (23a)$$

$$H_\xi = 0, \quad (23b)$$

$$P = 1 + \xi p_1(z). \quad (23c)$$

Then Eqs. (12) and (13) give, respectively,

$$\Gamma_z = (1/z)\Delta[1 + \xi(p_1 - f)], \quad (24)$$

$$\Gamma_\xi = z\Delta p_{1z}/2(1 + h_1^2), \quad (25)$$

while using Eq. (13) we find that Eq. (21) is equivalent to the relations

$$\Gamma_\xi^2 = p_1, \quad (26)$$

$$\Delta^2(p_1 - f) = 1 + h_1^2. \quad (27)$$

From Eqs. (25) and (26) we get

$$p_{1z} = \pm[2\Delta(p_1 - f)\sqrt{p_1}]/z. \quad (28)$$

Also the relation $\Gamma_{z\xi} = \Gamma_{\xi z}$, where Γ_z and Γ_ξ are given by Eqs. (24) and (26), leads to Eq. (28). The functions A_ξ and A_z can be easily calculated from Eqs. (17)–(20), (23), (27), and (28). We get

$$A_\xi = zh_1/\Delta, \quad (29)$$

$$A_z = \pm\xi h_1\sqrt{p_1}. \quad (30)$$

Then the relation $A_{\xi z} = A_{z\xi}$ gives

$$\frac{d}{dz}\left(\frac{zh_1}{\Delta}\right) = \pm h_1\sqrt{p_1}. \quad (31)$$

But if Eq. (31) is satisfied Eq. (30) can be integrated. The result is $A = \xi(zh_1/\Delta) + l(\xi)$, where $l(\xi)$ is an arbitrary function of ξ. Also since h_1 and Δ are functions of z only the integration of Eq. (29) gives $A = \xi z h_1/\Delta + \sigma(z)$. Therefore we must have $l = \sigma = c$, where c is an arbitrary constant. Choosing $c = 0$ to satisfy the boundary conditions we get

$$A = \xi z h_1/\Delta. \quad (32)$$

Therefore if we know h_1 and Δ we can obtain A using Eq. (32). We shall find functions Δ, h_1, p_1, and f, which satisfy Eqs. (27), (28), and (31).

From Eqs. (27) and (28) we get

$$(\sqrt{p_1})_z = \pm(1 + h_1^2)/z\Delta. \quad (33)$$

Using the above expression to eliminate p_1 from Eq. (31) and replacing z by w, where

$$w = \ln z, \quad (34)$$

Eq. (31) becomes

$$\frac{d}{dw}\left\{\frac{1}{\Delta}\frac{d}{dw}\left[\ln\left(\frac{h_1}{\Delta}\right)\right] + \frac{1}{\Delta}\right\} = \frac{1 + h_1^2}{\Delta}$$

or

$$\frac{d}{dw}\left(\frac{1 + \eta_w}{\Delta}\right) = \frac{1}{\Delta} + \Delta e^{2\eta}, \quad (35)$$

where

$$\eta = \ln(h_1/\Delta). \quad (36)$$

If we introduce q by the relation

$$\Delta = (1 + \eta_w)e^{q - \eta} \quad (37)$$

and use this expression to eliminate Δ from Eq. (35) we get a second-degree equation of η_w, which involves q and q_w besides η_w. Solving this equation with respect to η_w we get

$$\eta_w = [1/2(e^{2q} - 1)] \times \{1 - q_w - 2e^{2q} \pm [(q_w + 1)^2 + 4(1 - e^{2q})]^{1/2}\}. \quad (38)$$

For any q we can proceed in the calculation of η using Eq. (38). Then we can compute Δ, h_1, p_1, f, and A using Eqs. (37), (36), (31), (27), and (32), respectively, while we can find R and Q using Eqs. (19), (20), and (23). Since q is an arbitrary but known function of w the calculation of Δ, h_1, p_1, f, A, R, and Q is reduced to the integration of a known function. In this way a wide class of solutions can be obtained. We can show that every solution of Eqs. (27), (28), and (31) satisfies the fourth first-order equation[1]

$$(\sqrt{\xi}\Delta R)_\xi - z(\Delta Q)_z + \frac{1}{z^2}\Delta P A_\xi^2 + \frac{1}{\xi}\Delta Q A_z^2 - \frac{2}{z\sqrt{\xi}}\Delta R A_z A_\xi - \frac{1}{2} z f_z \Delta = 0$$

and therefore gives a solution of Einstein's equations.

We shall examine now if the solutions of the class satisfy the boundary conditions. Since in the general case the problem can be reduced to the solution of two partial differential equations involving P, A, and f, the boundary conditions are imposed on these functions. We must have[1]

$$P \to 1 + \xi\tilde{p}_1(z), \quad A \to \xi a_1(z) \quad \text{as } \xi \to 0, \quad (39)$$

where $\tilde{p}_1(z)$ and $a_1(z)$ are arbitrary functions of z. For our class of solutions we have $\tilde{p}_1(z) = p_1(z)$ and $a_1(z) = zh_1/\Delta$, as we see from Eqs. (23c) and (32). Also we must have[1]

$$\tilde{p}_1 - f - (a_1/z)^2 > 0. \tag{40}$$

For our class of solutions we get, if we use Eq. (27),

$$\tilde{p}_1 - f - \left(\frac{a_1}{z}\right)^2 = p_1 - f - \left(\frac{h_1}{\Delta}\right)^2 = \frac{1}{\Delta^2} > 0, \tag{41}$$

that is the relation (40) is satisfied. Therefore all solutions of the class have the proper behavior near the axis. Finally $f(z)$ must lead to a realistic equation of state. The explicit form of $f(z)$, and therefore the equation of state, depends on the specific solution.

Also we shall find the equipressure surfaces of the solutions of the class. To do that we compare[1] the metric of fluids equipressure surfaces

$$d\sigma^2 = h^2(\rho^2\, d\varphi^2 + \Delta^2 Q\, d\rho^2) \tag{42}$$

with the metric $d\sigma_E^2$ of a surface of revolution in Euclidean three-space described in cylindrical coordinates $\tilde{\rho}$, $\tilde{z}$, and φ by the equation $\tilde{z} = \tilde{z}(\rho)$. This metric is

$$d\sigma_E^2 = \left[1 + \left(\frac{d\tilde{z}}{d\tilde{\rho}}\right)^2\right] d\rho^2 + \tilde{\rho}^2\, d\varphi^2. \tag{43}$$

From Eqs. (42) and (43) we obtain

$$h\rho = \tilde{\rho}, \tag{44}$$

$$\Delta^2 Q - 1 = \left(\frac{d\tilde{z}}{d\tilde{\rho}}\right)^2. \tag{45}$$

But from Eqs. (20), (23), and (27) we get for all solutions of the class

$$\Delta^2 Q - 1 = 0. \tag{46}$$

Then Eq. (45) gives

$$\tilde{z} = \text{const}. \tag{47}$$

Therefore the equipressure surfaces are planes.

III. EXPLICIT EXAMPLES

In this section we shall give explicitly a number of solutions of the class. Two such solutions can be obtained if we assume that

$$q_w - e^{2q} + 1 = 0. \tag{48}$$

The general solution of this equation is

$$q = -\tfrac{1}{2}\ln(c_1 e^{2w} + 1) = -\tfrac{1}{2}\ln(c_1 z^2 + 1), \tag{49}$$

where c_1 is an arbitrary constant. Then Eq. (38) gives

$$(\eta_{1,2})_w = \begin{cases} -2, \\ e^{2q}/(1 - e^{2q}) = (1/c_1)e^{-2w}. \end{cases} \tag{50}$$

Therefore we get from Eqs. (34) and (50)

$$\eta_{1,2} = \begin{cases} -2\ln z + c, \\ -(1/2c_1)z^{-2} + c, \end{cases} \tag{51}$$

where c is an arbitrary constant. If we know q and η we can proceed as explained before to calculate Δ, h_1, p_1, f, A, R, and Q. For q and η given by Eqs. (49) and (51) we get the following two solutions.

Solution 1:

$$\Delta_1 = \frac{c_2 z^2}{\sqrt{c_1 z^2 + 1}}, \quad (h_1)_1 = -\frac{1}{\sqrt{c_1 z^2 + 1}},$$
$$P_1 = 1 + \xi \frac{c_1 z^2 + 1}{c_2^2 z^4}, \quad f_1 = -\frac{1}{c_2^2 z^4}, \quad A_1 = -\frac{\xi}{c_2 z}, \tag{52}$$
$$R_1 = -\sqrt{\xi}\,\frac{c_1 z^2 + 1}{c_2^2 z^4}, \quad Q_1 = \frac{c_1 z^2 + 1}{c_2^2 z^4}.$$

Solution 2:

$$\Delta_2 = \frac{\sqrt{c_1 z^2 + 1}}{c_1 c_2 z^2} \exp\left(\frac{1}{2c_1 z^2}\right), \quad (h_1)_2 = \frac{\sqrt{c_1 z^2 + 1}}{c_1 z^2},$$
$$P_2 = 1 + c_2^2 \xi (c_1 z^2 + 1) \exp\left(-\frac{1}{c_1 z^2}\right),$$
$$f_2 = \frac{c_2^2 c_1 z^2}{c_1 z^2 + 1} \exp\left(-\frac{1}{c_1 z^2}\right), \quad A_2 = c_2 \xi z \exp\left(-\frac{1}{2c_1 z^2}\right), \tag{53}$$
$$R_2 = c_1 c_2^2 \sqrt{\xi}\, z^2 \exp\left(-\frac{1}{c_1 z^2}\right),$$
$$Q_2 = \frac{c_1^2 c_2^2 z^4}{c_1 z^2 + 1} \exp\left(-\frac{1}{c_1 z^2}\right),$$

where c_1 and c_2 are arbitrary real constants.

Other solutions can be found as follows: Integrating Eq. (38) and setting

$$q = \tfrac{1}{2}\ln T \tag{54}$$

we get

$$\eta = -w - \tfrac{1}{4}\ln(1 - T) - \int \frac{dw}{2((1/T) - 1)}$$
$$\pm \int \frac{\sqrt{(-(T_w/2T) + 1)^2 + 4(1 - 1/T)}}{2(1/T - 1)}\, dw. \tag{55}$$

We shall find solutions of the above equation for which

$$(-T_w/2T + 1)^2 + 4(1 - 1/T) = T. \tag{56}$$

The solution of (56) is

$$T = [(1 + 2e^{3(w+c)})/(1 - e^{3(w+c)})]^2, \tag{57}$$

where c is a constant. Substituting this expression in Eq. (55), performing the integration, and expressing the final result in terms of z we get the solutions

$$\eta_{3,4} = \begin{cases} \ln[c_1(1 - c_2 z^3)z^{-2}], \\ -\tfrac{1}{4}\ln[c_1(2 + c_2 z^3)z^{11/3}] - (1/18c_2)z^{-3}. \end{cases} \tag{58}$$

Starting from the expressions (54), (57), and (58) and proceeding as before we easily get two solutions.

Solution 3:

$$\Delta_3 = -z^2/c_1(1 - c_2 z^3), \quad (h_1)_3 = -1,$$
$$P_3 = 1 + (\xi/z^4)c_1^2(1 + 2c_2 z^3)^2,$$
$$f_3 = (c_1^2/z^4)[2(2 + c_2 z^3)^2 - 9],$$
$$A_3 = (\xi/z)c_1(1 - c_2 z^3), \tag{59}$$
$$R_3 = -(\sqrt{\xi}/z^4)c_1^2(1 - c_2 z^3)(1 + 2c_2 z^3),$$
$$Q_3 = (c_1^2/z^4)(1 - c_2 z^3)^2.$$

Solution 4:

$$\Delta_4 = \frac{(1-c_2z^3)^2}{3c_1c_2z^3(2+c_2z^3)^{3/4}} z^{11/12} \exp\left(\frac{1}{18c_2z^3}\right),$$

$$(h_1)_4 = \frac{(1-c_2z^3)^2}{3c_2z^3(2+c_2z^3)},$$

$$P_4 = 1 + \xi \frac{c_1^2(1+2c_2z^3)^2}{(1-c_2z^3)^2(2+c_2z^3)^{1/2}} \times z^{-11/6} \exp\left(-\frac{1}{9c_2z^3}\right),$$

$$f_4 = \frac{3c_1^2c_2z^3(2+c_2z^3)^{1/2}[9-2(2+c_2z^3)^2]}{(1-c_2z^3)^4} \times z^{-11/6}\exp\left(-\frac{1}{9c_2z^3}\right), \qquad (60)$$

$$A_4 = \xi \frac{c_1 z}{(2+c_2z^3)^{1/4}} z^{-11/12} \exp\left(-\frac{1}{18c_2z^3}\right),$$

$$R_4 = \sqrt{\xi}\, \frac{3c_1^2c_2z^3(1+2c_2z^3)(2+c_2z^3)^{1/2}}{(1-c_2z^3)^3} \times z^{-11/6}\exp\left(-\frac{1}{9c_2z^3}\right),$$

$$Q_4 = \frac{9c_1^2c_2^2z^6(2+c_2z^3)^{3/2}}{(1-c_2z^3)^4} z^{-11/6}\exp\left(-\frac{1}{9c_2z^3}\right),$$

where c_1 and c_2 are arbitrary real constants.

Additional solutions can be found explicitly as follows: Let us take

$$q = \tfrac{1}{2}(c_1 w + \ln c_2), \qquad (61)$$

where c_1 and c_2 are constants. Substituting the above expression in Eq. (38) and performing the integration we get

$$\eta = \left(c_4 - 1 \pm \sqrt{c_4^2+1}\right) w - (1 \pm 1)\,(c_4/c_1)\ln(c_2 e^{c_1 w} - 1) \pm (2c_4/c_1)\ln\left[\sqrt{c_4^2+1-c_2e^{c_1 w}} - c_4\right]$$
$$\mp (2/c_1)\sqrt{c_4^2+1}\,\ln\left[\sqrt{c_4^2+1-c_2e^{c_1w}} - \sqrt{c_4^2+1}\right] + \ln c_{3\pm}, \qquad (62)$$

where $c_4 = (c_1+2)/4$ and $c_{3\pm}$ are constants. Starting from Eqs. (61) and (62) and proceeding as before we can find two solutions.

Solutions 5 and 6:

$$\Delta_{5,6} = (\sqrt{c_2}/c_{3\pm})z^{c_4 \mp \sqrt{c_4^2+1}}\left(-c_4 \mp \sqrt{c_4^2+1-c_2z^{c_1}}\right)^{(2c_4/c_1)-1}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\pm 2\sqrt{c_4^2+1}/c_1},$$

$$(h_1)_{5,6} = \mp \frac{\sqrt{c_2}z^{c_1/2}}{\sqrt{c_4^2+1-c_2z^{c_1}} \pm c_4},$$

$$P_{5,6} = 1 + \xi(c_{3\pm}^2/c_2)z^{-2c_4 \pm 2\sqrt{c_4^2+1}}\left(\mp\sqrt{c_4^2+1-c_2z^{c_1}} - c_4\right)^{-4c_4/c_1}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\mp 4\sqrt{c_4^2+1}/c_1},$$

$$f_{5,6} = 2c_4(c_{3\pm}^2/c_2)z^{-2c_4 \pm 2\sqrt{c_4^2+1}}\left(\mp\sqrt{c_4^2+1-c_2z^{c_1}} - c_4\right)^{-2/c_1}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\mp 4\sqrt{c_4^2+1}/c_1}, \qquad (63)$$

$$A_{5,6} = \xi c_{3\pm} z^{c_4 \pm \sqrt{c_4^2+1}}\left(\mp\sqrt{c_4^2+1-c_2z^{c_1}} - c_4\right)^{-2c_4/c_1}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\mp 2\sqrt{c_4^2+1}/c_1},$$

$$R_{5,6} = \sqrt{\xi}(c_{3\pm}^2/c_2)z^{-2c_4 \pm 2\sqrt{c_4^2+1}}\left(\mp\sqrt{c_4^2+1-c_2z^{c_1}} - c_4\right)^{-2/c_1}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\mp 4\sqrt{c_4^2+1}/c_1},$$

$$Q_{5,6} = (c_{3\pm}^2/c_2)z^{-2c_4 \pm 2\sqrt{c_4^2+1}}\left(\mp\sqrt{c_4^2+1-c_2z^{c_1}} - c_4\right)^{-(4c_4/c_1)+2}\left(\sqrt{c_4^2+1-c_2z^{c_1}} - \sqrt{c_4^2+1}\right)^{\mp 4\sqrt{c_4^2+1}/c_1},$$

where c_2 is a positive but otherwise arbitrary real constant, the $c_{3\pm}$ are arbitrary real constants, and the upper sign corresponds to one of the solutions and the lower sign to the other.

IV. EQUATION OF STATE AND ANOTHER SOLUTION

Using Eqs. (4) and the expression for f of the various solutions we can find the equation of state of these solutions. For example, for solution 1 we get the equation of state

$$\mu + 11p = 0, \qquad (64)$$

which is not very meaningful physically. However axially, symmetric perfect fluid solutions of Einstein's equations with an equation of state of the form $\mu + 3p = \text{const}$ (see Refs. 5 and 6) or of the form $\mu + 3p = 0$ (Ref. 7) have been reported. The equation of state of the other solutions of Sec. III are complicated and will not be given explicitly.

It is interesting to get solutions of the class with a realistic equation of state. We shall do that now following a way that differs from the one we have followed above. More specifically we shall not use Eq. (38) but from Eqs. (27), (31), and (33) we shall eliminate h_1 and Δ and we shall get a relation between p_1 and f. Then we shall try to find solutions of this equation, whose f gives a reasonable equation of state.

If we set

$$\delta = (h_1/\Delta)^2, \quad \nu^2 = p_1, \qquad (65)$$

Eqs. (27), (31), and (33) take the form

$$\nu^2 - f = \delta + 1/\Delta^2, \qquad (66)$$

$$\partial_w(\ln\delta) + 2 \mp 2\nu\Delta = 0, \qquad (67)$$

$$\nu_w = \pm\Delta(\delta + 1/\Delta^2). \qquad (68)$$

Solving Eqs. (66) and (68) for δ and Δ and substituting in Eq. (67) we get

$$\partial_w[\ln(1 - (\nu^2 - f)/\nu_w^2)] + 2 - f_w/(\nu^2 - f) = 0. \qquad (69)$$

We see that the problem of finding solutions of the class has been reduced to the problem of finding solutions of the above equation, which has two dependent variables.

The equation of state

$$p = \gamma\mu, \tag{70}$$

where γ is a positive constant, is very interesting physically. Using Eqs. (4) we find that the function f which gives the equation of state (70) is

$$f = c_1 z^c = c_1 e^{cw}, \tag{71a}$$

$$c = (3\gamma + 1)/2\gamma, \tag{71b}$$

where c_1 is an arbitrary constant. Substituting the above expression in Eq. (69) we find that this equation takes the form

$$\begin{aligned} &2(v^2 - c_1 e^{cw})^2 v_{ww} + [2(v^2 - c_1 e^{cw}) - c_1 c e^{cw}] v_w^3 \\ &\quad - 2v(v^2 - c_1 e^{cw}) v_w^2 - 2(v^2 - c_1 e^{cw} - c_1 c e^{cw}) \\ &\quad \times (v^2 - c_1 e^{cw}) v_w = 0. \end{aligned} \tag{72}$$

A special solution of this complicated nonlinear second-order differential equation is given by the simple expression

$$v = c_3 e^{(c/2)w} = c_3 z^{c/2}, \tag{73}$$

where c_3 is a constant, provided that c_1, c_3, and c satisfy the relation

$$c_1 c - 2(c_3^2 - c_1) = 0 \tag{74}$$

or the relation

$$c_3^2 c^2 - 4(c_3^2 - c_1) = 0. \tag{75}$$

But if Eqs. (71a), (73), and (75) are satisfied the expression $1 - (v^2 - f)/v_w^2$, whose ln appears in Eq. (69), becomes 0. Therefore the relation (75) cannot be accepted, which means that our solution is given by Eqs. (71), (73), and (74). Calling it solution 7 and using these relations and Eqs. (19), (20), (23), (32), (65), (66), and (68) we find its explicit form, which is the following:

$$\begin{aligned} &\Delta_7 = \frac{7\gamma + 1}{4\gamma c_3} z^{-(3\gamma+1)/4\gamma}, \quad (h_1)_7 = \frac{\sqrt{5\gamma^2 + 10\gamma + 1}}{4\gamma}, \\ &P_7 = 1 + \xi c_3^2 z^{(3\gamma+1)/2\gamma}, \quad f_7 = \frac{4\gamma c_3^2}{7\gamma + 1} z^{(3\gamma+1)/2\gamma}, \\ &A_7 = \xi \frac{c_3\sqrt{5\gamma^2 + 10\gamma + 1}}{7\gamma + 1} z^{(7\gamma+1)/4\gamma}, \\ &R_7 = \sqrt{\xi}\, \frac{4\gamma c_3^2}{7\gamma + 1} z^{(3\gamma+1)/2\gamma}, \quad Q_7 = \frac{16\gamma^2 c_3^2}{(7\gamma + 1)^2} z^{(3\gamma+1)/2\gamma}. \end{aligned} \tag{76}$$

In Eqs. (76) c_3 is an arbitrary real constant. Also since the constant γ appears as a factor in the denominators and the expression $5\gamma^2 + 10\gamma + 1$ must be positive, γ can take any value in the intervals

$$\gamma > 0, \quad 0 > \gamma \geqslant -1 + 2/\sqrt{5}, \quad \gamma \leqslant -1 - 2/\sqrt{5}. \tag{77}$$

The equation of state of solution (76) is given by Eq. (70), where γ takes the above values. Since γ can take any positive value this equation of state is physically meaningful. Cases of particular interest of the above solution are the "(incoherent) radiation" case

$$p = \tfrac{1}{3}\mu, \quad \gamma = \tfrac{1}{3}, \tag{78}$$

and also the "stiff matter" case

$$p = \mu, \quad \gamma = 1, \tag{79}$$

which leads to a sound speed equal to the velocity of light.[4] The "dust" case

$$p = 0, \quad \gamma = 0 \tag{80}$$

is a limiting case, which is excluded. The pressure can be small, but not exactly zero. Exact solutions with equations of state of the form (70) are of interest, for example, in connection with cosmological models, as interior solutions to be matched with vacuum exterior solutions, etc.

Generally we can use Eq. (69) to derive solutions of the class. We can assume a certain form for f (or v) and calculate v (or f) by solving Eq. (69), in which case we have to solve a second- (or first-) order nonlinear differential equation.

V. SYMMETRY OF SOLUTIONS

The fact that the surfaces of constant pressure of all solutions of the class are planes may lead one to suspect that the solutions are probably plane symmetric.[8] To face this problem we recall the well-known fact that the Killing vectors of a solution determine its symmetry in a manner independent of the coordinate system. Therefore to check if a solution of the class is plane symmetric we can calculate its Killing vectors. If we find only two commuting Killing vectors, one timelike and one spacelike, the highest symmetry of the solution is the axial symmetry, that is, the solution is not plane symmetric. We did the above calculation for the solution 7 and we found that it has only two Killing vectors, a timelike and a spacelike ($\xi = \partial_t$, $\eta = \partial_\varphi$). Therefore the solution 7 has no symmetry higher than the axial symmetry. We should point out that the solution 1 for $c_1 = 0$ and solution 3 for $c_2 = 0$ become identical with the solution 7 if $11\gamma + 1 = 0$, $c_2 c_3 + 1 = 0$ and if $11\gamma + 1 = 0$, $c_1 = c_3$, respectively. In general we expect the symmetry of the solutions of the class to be the axial symmetry.

ACKNOWLEDGMENT

I am very grateful to Dr. S. Bonanos for many and very interesting discussions.

[1] S. Bonanos and D. Sklavenitis, J. Math. Phys. **26**, 2275 (1985).
[2] C. B. Collins, J. Math. Phys. **25**, 995 (1984).
[3] D. Sklavenitis, J. Math. Phys. **26**, 2279 (1985).
[4] D. Kramer, H. Stephani, E. Herlt, and M. MacCallum, *Exact Solutions of Einstein's Field Equations* (Cambridge U. P., Cambridge, 1980).
[5] D. Kramer, Class Quantum Gravit. **1**, L3 (1984).
[6] H. D. Wahlquist, Phys. Rev. **172**, 1291 (1968).
[7] P. C. Vaidya, Pramana **8**, 512 (1977). It was shown by E. Herlt and H. Hermann [Exp. Tech. Phys. **28**, 97 (1980)] that Vaidya's solution is a limiting case of Wahlquist's solution.

[8] This was pointed out to me by the referee of the paper, whom I thank.

Charge quantization without superheavy masses in a Kaluza–Klein description of electromagnetism

D. K. Ross
Physics Department, Iowa State University, Ames, Iowa 50011

(Received 8 April 1986; accepted for publication 29 April 1987)

A scalar matter field coupled to general relativity and electromagnetism in a five-dimensional Kaluza–Klein model is considered. The five-dimensional space is assumed to be a fiber bundle as in the usual description of a gauge theory and not a more general manifold. Properly taking this into account allows one to use a Lagrangian density for the scalar field which includes charge quantization but not the unphysical superheavy masses found by other authors. A natural, satisfactory explanation of why charge is quantized results.

I. INTRODUCTION

There is presently no satisfactory explanation of why all observed free particles have a charge that is an integer multiple of the absolute value of the charge on the electron e. Dirac[1] sought such an explanation but was led to a quantization condition for eg, where g is the strength of the magnetic monopole, somewhat to his disappointment. (In the modern language of a fiber bundle model of a U_1 gauge theory, pioneered by Wu and Yang,[2] this quantization condition arises from the requirement that the gauge transformation in the overlap region of the two domains necessary to cover the base space be single valued.)

A promising alternative approach to understanding charge quantization follows from the Kaluza[3]–Klein[4] five-dimensional geometrical unification of general relativity and electromagnetism. This was developed by Einstein, Bargmann, and Bergmann[5,6] who treated the fifth dimension as a compact space. Bergmann[7] showed that the circumference of a closed space in which all self-intersecting geodesics are closed lines without discontinuities of direction is a characteristic constant of that space. This strongly suggests that this circumference should represent some invariant physical quantity. We see below that it is related to electric charge. Note that assuming the fifth dimension to be compact in the original five-dimensional general manifold has little justification and in fact partially dismantles the unification. If we view the five-dimensional space to be a fiber bundle, however, as in the modern description of a gauge theory by Cho,[8] with a U_1 gauge group (isomorphic to a circle) and space-time as the base space, then the above cylindrical space is quite natural.

Klein[9] first got charge quantization in a rather heuristic paper out of a Kaluza–Klein model by defining the momentum conjugate to the fifth dimension and then putting an integer number of de Broglie wavelengths around the compact space. Souriau,[10] in a better approach, considered the Klein–Gordon equation for a scalar field in a general five-dimensional manifold. His work has been used by Chodos and Detweiler[11] and by Gross and Perry.[12] The Lagrangian density can be written as

$$\mathscr{L} = \gamma^{AB} (\partial_A \phi)(\partial_B \phi^\dagger), \tag{1}$$

where γ^{AB} is the metric of the five-dimensional manifold. (A, B take on five values here while μ, ν refer to the usual four-dimensional space-time.) If the $\gamma^{\mu 5}$ components of γ^{AB} are identified with the electromagnetic vector potential A^μ as in the Kaluza–Klein theory and ϕ is assumed to be a periodic function of the coordinate of the fifth dimension x_5, with period equal to the circumference of the fifth dimension, then e is proportional to the reciprocal of the circumference of the fifth dimension and is quantized (see below). Unfortunately the $\gamma^{55}(\partial_5\phi)(\partial_5\phi^\dagger)$ term in (1) leads to superheavy Planck scale masses for all the charged particles. Putting a mass counterterm into (1) by hand to yield small physical masses also does not work. The added mass term would require fine tuning to 20 decimal places[11] and would have to be separately fine tuned for each charged particle. This is clearly very unsatisfactory. If the closed fifth dimension has a dynamical origin as in the work of Chodos and Detweiler,[11] the charged particle would have been a tachyon in the past and would become very massive in the future.

In the present paper, we show that if a Klein–Gordon particle is considered, not in a general five-dimensional Riemannian manifold, but in a five-dimensional fiber bundle with the fiber bundle structure properly taken into account, then (1) is not the only possibility. Using a more appropriate Lagrangian density, which is invariant to any change of basis that does not destroy the fiber bundle, leads to Souriau's charge quantization but without the superheavy masses appearing at all. We thus arrive at a satisfactory explanation of the origin of charge quantization.

II. QUANTIZATION OF ELECTRIC CHARGE WITHOUT SUPERHEAVY MASSES

We will assume that electromagnetism is correctly described by a principal fiber bundle with a U_1 structure group G and space-time for the base space M. The usual four-potentials A_μ are then cross-section-dependent components of the connection form in the fiber bundle. The fiber bundle can be viewed locally, but not globally, as the product of G and M. This elegant fiber bundle description of a gauge theory has been developed by many authors.[8,13–18] Following the nice treatment by Cho[8] and leaving out many details, let us choose a coordinate basis $\xi_\mu = \partial_\mu$ for M, where

$$[\xi_\mu, \xi_\nu] = 0, \tag{2}$$

 0022-2488/87/092167-04$02.50

and a set of n linearly independent left invariant vector fields ξ_i for a basis of G, where

$$[\xi_i,\xi_j] = f^k_{ij}\xi_k . \tag{3}$$

Here n will be 1 for the U_1 electromagnetism case and the structure constants f^k_{ij} vanish for this Abelian group. In much of the following, we will treat the more general non-Abelian case since many of our results hold there. (Uppercase Latin letters will range over $n+4$ values in general, lowercase Greek letters over four space-time indices.) The ξ_i can be naturally mapped into the fiber bundle space as the fundamental vector fields ξ^*_i. Also the ξ_μ can be horizontally lifted into the bundle to become $\hat{\xi}_\mu$. Here $\hat{\xi}_\mu$ and ξ^*_i can be used as a basis for the fiber bundle if we wish. Their commutation relations are

$$[\xi^*_i,\xi^*_j] = f^k_{ij}\xi^*_k , \quad [\xi^*_i,\hat{\xi}_\mu] = 0 , \\ [\hat{\xi}_\mu,\hat{\xi}_\nu] = -F^k_{\mu\nu}\xi^*_k , \tag{4}$$

where $F^k_{\mu\nu}$ are the Yang–Mills fields. Following Cho,[8] we can also introduce a metric for the fiber bundle γ_{AB} invariantly defined to satisfy

$$\gamma_{AB}\hat{\xi}^A_\mu\hat{\xi}^B_\nu = g_{\mu\nu} , \quad \gamma_{AB}\hat{\xi}^A_\mu\xi^{*B}_k = 0 , \\ \gamma_{AB}\xi^{*A}_i\xi^{*B}_k = g_{ik} , \tag{5}$$

where $g_{\mu\nu}$ is the metric of M and g_{ik} the invariant metric of G. For the U_1 case we will take $g_{55}=1$. We could put a scalar field in for g_{55} but such Brans–Dicke[19] type scalar fields seem not to be seen in nature.

The five-dimensional fiber bundle is now a Riemannian manifold on which the Hilbert Lagrangian density R can be defined. Using this in an action principle rather miraculously gives general relativity correctly coupled to the (sourceless) Maxwell equations. The complete Einstein–Maxwell field equations are

$$R_{AB} = 0 \tag{6}$$

except that we have no R_{55} equation in the U_1 case since δg_{55} must always vanish from our above assumption that $g_{55}=1$. We note that we have to be careful not to do something, in this case vary g_{55}, which would be inconsistent with the assumed fiber bundle structure.

Now to have electric charge present, we must have a source field ϕ in addition to the gauge fields above. We will treat a scalar field for simplicity. A Dirac field would be very similar. We want this scalar field to act as a source both for general relativity and for the Maxwell equations. In order to do this, we need a Lagrangian density for the scalar field, $\mathscr{L}_{\text{scalar}}$, which can be added to R *in the five-dimensional fiber bundle*. Variation of $\mathscr{L}_{\text{scalar}}$ with respect to ϕ will then yield the Klein–Gordon equation for the scalar field and variation with respect to γ^{AB} will give a ϕ contribution to the energy momentum tensor. If ϕ is a quantum field, this latter can be treated approximately as a vacuum expectation valued source in the classical general relativity equations. A better approach is to quantize general relativity as well, of course. This really will not concern us here since we are primarily interested in the ϕ field itself. Note that ϕ and $\mathscr{L}_{\text{scalar}}$ live, at least initially, in the five-dimensional fiber bundle in this view. Ordinarily a scalar field is viewed as a cross section of a vector bundle associated to the principal fiber bundle through a representation of G.[16,18] The connection in the principal fiber bundle (gauge covariant derivative) induces a connection in the associated bundle. We will recover this description in the end but only after we integrate over the gauge degrees of freedom in $\mathscr{L}_{\text{scalar}}$ in the fiber bundle to get an effective four-dimensional variational principle.

Treating $\mathscr{L}_{\text{scalar}}$ as an object in the five-dimensional space is exactly along the lines of the approach of Souriau,[10] Chodos and Detweiler,[12] and Gross and Perry.[11] The difference below is that we take account of the fact that this five-dimensional space is a principal fiber bundle and they do not.

We now need an invariant expression for $\mathscr{L}_{\text{matter}}$. Souriau[10] and later authors[11,12] using his approach use (1) for this Lagrangian density.

Souriau[10] uses a metric of the form

$$\gamma^{AB} = \begin{pmatrix} g^{\mu\nu} & -A^\mu \\ -A^\nu & 1+g^{\mu\nu}A_\mu A_\nu \end{pmatrix}, \tag{7}$$

which is equivalent to using ξ_μ from (2) and ξ^*_i as basis vectors in the fiber bundle picture.[8] In this basis, ∂_A in (1) is the usual partial derivative. As alluded to in the Introduction, using (7) in (1) leads to[12]

$$\gamma^{AB}(\partial_A\phi)(\partial_B\phi^\dagger) = |(\partial_\mu + i(n/R^5)A_\mu)\phi|^2 \\ + [n^2/(R^5)^2]|\phi|^2 \tag{8}$$

for the Fourier component of ϕ with x^5 dependence given by e^{inx^5/R^5}, where $2\pi R^5$ is the circumference of the fifth dimension. Equation (8) represents a particle of charge

$$e = (\sqrt{16\pi G}/R^5)n , \tag{9}$$

where n is an integer, and of unphysical superheavy mass

$$m = n/R^5 . \tag{10}$$

If A_μ in (7) is put into conventional units, the properly normalized gauge field is $(16\pi G)^{-1/2}A_\mu$. The fine structure constant is $\alpha = (2L_p/R^5)^2$, where L_p is the Planck length. Thus R^5 must be about 23 times the Planck length.

What is wrong with the Lagrangian density (8)? Can we preserve the charge quantization (9) while eliminating the $\gamma^{55}(\partial_5\phi)(\partial_5\phi^\dagger)$ mass term? Equation (8) is invariant under general coordinate transformations in a general five-dimensional Riemannian manifold. We have a fiber bundle, however, *not* a general five-dimensional manifold. In particular a connection always exists on the fiber bundle. If ξ^*_i is considered the basis of a vertical subspace V_p of the tangent space T_p to the fiber bundle P then a connection is a choice of horizontal subspace H_p such that[8]

(a) T_p is the direct sum of V_p and H_p,

(b) $a\in G$ and $p\in P$, $H_{pa} = R_a H_p$,

where R_a is right multiplication by a,

(c) H_p is smooth on P. (11)

We will exploit the fact that we have a fiber bundle to find a more appropriate Lagrangian density.

We want a Lagrangian density for the scalar field which is invariant under a change of basis of G, invariant under a change of basis of M, and finally invariant under any change of basis of the fiber bundle P which preserves the bundle structure. We also want the usual Klein–Gordon Lagrangian upon integration over the x^5 gauge coordinate. A Lagrangian density that satisfies these conditions is

$$\mathscr{L}_{\text{scalar}} = \gamma^{\alpha\beta}(\hat{\xi}_\alpha\phi)(\hat{\xi}_\beta\phi^\dagger) \tag{12}$$

as we shall now prove. Note that the horizontal lift vector fields $\hat{\xi}_\alpha$ are invariant vector fields $\hat{\xi}_\alpha$ and do *not* depend upon the basis in which they are written, although their components of course do depend upon the choice of basis. If we write these out in terms of the local direct product basis,[8] they are $\hat{\xi}_\alpha = \partial_\alpha - A_\alpha\partial_5$ in the U_1 case and look like gauge covariant derivatives. The $\gamma^{\alpha\beta}$ are the space-time sector components of the fiber bundle metric γ^{AB}. Equation (12) is written in terms of the fiber bundle metric as it must be since it lives in the five-dimensional fiber bundle space.

We note that (12) is at least as natural a Lagrangian density as (1) if one is dealing with a fiber bundle. Both can be viewed as a generalization of the usual Lagrangian density of a scalar field extended to a higher-dimensional fiber bundle. As we shall see below, (12) does not lead to unwanted superheavy masses but (1) does. The fact that we are in a fiber bundle does not rule out (1), but rather opens up the possibility of using (12) instead, thus avoiding the superheavy masses.

Equation (12) as it stands certainly does not look invariant to a change of basis in P since it only involves some of the components of γ^{AB}. However, we must preserve the fiber bundle structure and we will find that any change of basis which does this will leave (12) invariant. We will show, in fact, quite generally even for the non-Abelian case, that $\gamma^{\alpha\beta} = g^{\alpha\beta}$, where $g^{\alpha\beta}$ is the usual space-time metric of the base space. Note that other components of γ^{AB} are not being discussed and $\gamma_{\alpha\beta} \neq g_{\alpha\beta}$ in general, for example. We also verify that at least for the horizontal lift basis and for the local direct product basis in Cho[8] that $\gamma^{\alpha\beta} = g^{\alpha\beta}$. Let us now show this in general.

Consider a general basis $\check{\xi}_\mu, \check{\xi}_i$ in P. We can write this in terms of the horizontal lift vector fields and fundamental lift vector fields as

$$\check{\xi}_\mu = Y^\nu_\mu\hat{\xi}_\nu + Z^i_\mu\xi^*_i, \quad \check{\xi}_i = W^j_i\xi^*_j + X^\mu_i\hat{\xi}_\mu, \tag{13}$$

where the W, X, Y, Z are systems of coefficients. It is easy to show that by rechoosing basis vectors in G and M, we can always write this in the form

$$\check{\xi}_\mu = \hat{\xi}_\mu + Z^i_\mu\xi^*_i, \quad \check{\xi}_i = \xi^*_i + X^\mu_i\hat{\xi}_\mu. \tag{14}$$

Now we use the crucial fact that a connection exists on our fiber bundle. Consider the tangent space T_p to the fiber bundle at point $p \in P$. A subspace of T_p is those vectors which are *only* tangent to the fiber passing through p. This is the vertical subspace V_p mentioned above. A connection is then a choice of H_p satisfying (16). It is clear that $\check{\xi}_i$ can at most be linear combinations of the ξ^*_i. Thus $X^\mu_i = 0$ in (14). Also using H_p we can introduce a connection form[8] satisfying

$$\omega^i(\hat{\xi}_\mu) \equiv \omega^i_A\hat{\xi}^A_\mu = 0, \quad \omega^i(\xi^*_j) \equiv \omega^i_A\xi^{*A}_j\delta^i_j. \tag{15}$$

Thus $Z^i_\mu = \omega^i(\check{\xi}_\mu)$ for self-consistency in (14). Thus quite generally any choice of basis in the fiber bundle can be put into the form

$$\check{\xi}_\mu = \hat{\xi}_\mu + \omega^i(\xi_\mu)\xi^*_i, \quad \check{\xi}_i = \xi^*_i. \tag{16}$$

Now let me introduce dual vector fields to the $\hat{\xi}_\mu$. These are defined as

$$\hat{\omega}^\mu(\hat{\xi}_\nu) \equiv \hat{\omega}^\mu_A\hat{\xi}^A_\nu = \delta^\mu_\nu, \quad \hat{\omega}^\mu(\xi^*_i) \equiv \hat{\omega}^\mu_A\xi^{A*}_i = 0. \tag{17}$$

Now in the general basis (16), the components of $\hat{\xi}_\mu$ and ξ^*_i can be written as

$$\xi^{*A}_i = \delta^A_i \text{ and } \hat{\xi}^\alpha_\mu = \delta^\alpha_\mu. \tag{18}$$

Note that the remaining component of $\hat{\xi}^A_\mu$, namely $\hat{\xi}^i_\mu$, depends upon the basis used and is not needed below. Using (18) we can now find some of the components of ω^i_A and $\hat{\omega}^\mu_A$. Using (18) in (17) gives

$$\hat{\omega}^\mu_A = \delta^\mu_A. \tag{19}$$

Using (18) in (15) gives

$$\omega^i_j = \delta^i_j. \tag{20}$$

Now we have γ_{AB} defined in (5). For our Lagrangian density (12), we need γ^{AB}. This is most conveniently given in terms of the dual vector fields $\hat{\omega}^\mu$ as

$$\gamma^{AB}\hat{\omega}^\mu_A\hat{\omega}^\nu_B = g^{\mu\nu}, \quad \gamma^{AB}\hat{\omega}^\mu_A\omega^i_B = 0, \quad \gamma^{AB}\omega^i_A\omega^k_B = g^{ik}. \tag{21}$$

This is a basis invariant definition and can easily be shown to be consistent with the definition of γ_{AB} in (5) and relations like $\gamma_{AB}\gamma^{BC} = \delta^C_A$. Substituting (19) into (21) then gives

$$\gamma^{\alpha\beta} = g^{\alpha\beta} \tag{22}$$

exactly as we wanted for a general basis in P. Substituting (18) into (5) also gives

$$\gamma_{ik} = g_{ik}, \tag{23}$$

in the general non-Abelian case. All other components of γ^{AB} and γ_{AB} do not behave so nicely under a change of basis.

Using (22), (12) can now be written as

$$\mathscr{L}_{\text{scalar}} = g^{\alpha\beta}(\hat{\xi}_\alpha\phi)(\hat{\xi}_\beta\phi^\dagger). \tag{24}$$

If ϕ is a scalar, it is clear that this is invariant under a change of basis of M since it is a space-time covariant scalar. It is invariant under a change of basis of G since $\hat{\xi}_\alpha$ is an invariant vector field. Also we have shown above that (24) is invariant under any change of basis of P which does not destroy the connection. In fact $g^{\alpha\beta}$ is specified independently of the fiber bundle so that basis invariance in P for the form (24) is obvious. The fact that (12) is invariant to a change in basis in P is not so obvious, but we have shown (12) is equivalent to (24).

If we work (24) out in the local direct product basis where the basis vectors become ordinary partial derivatives, we have

$$\mathscr{L}_{\text{scalar}} = g^{\alpha\beta}(\partial_\alpha\phi - A_\alpha\,\partial_5\phi)(\partial_\beta\phi^\dagger - A_\beta\,\partial_5\phi^\dagger). \tag{25}$$

Assuming that ϕ must be periodic in x^5 with a period equal to the circumference of the fifth dimension (ϕ must return to the same value after going around the closed fifth dimension) gives the usual gauge covariant derivatives with the same charge quantization condition as before, namely (9), for the Fourier component of ϕ given by $\phi_{(n)}(x^\mu)e^{inx^5/R^5}$. (Note that under a "gauge transformation" of the form $x^5 \to x^5 + f(x^\mu)$, this x^5 dependence gives the usual phase change to the wave function.) The difference is that now *we do not get superheavy masses* and in fact no mass at all. An invariant mass term $m^2\phi\phi^\dagger$ can be put in by hand giving the charged particles any mass we like. This is exactly what we want. If we add $\mathscr{L}_{\text{scalar}}$ to the Hilbert action in the five-dimensional fiber bundle and integrate over x^5 to get an effective four-dimensional theory, we end up with exactly the usual Klein–Gordon Lagrangian density for $\phi_{(n)}(x^\mu)$ with the usual gauge covariant derivatives coupled to the usual Hilbert action for $g_{\alpha\beta}$ and the Maxwell Lagrangian density. After the ∂_5 derivatives are taken, the x^5 dependence cancels out between ϕ and $\phi^\dagger$ in (25).

At this point we can view ϕ as given by a cross section of an associated vector bundle if we wish. This is the usual view of matter fields in fiber bundle models of gauge theories. Note that the fact that ϕ originally lives in the fiber bundle itself was important above and led both to the acceptable form of $\mathscr{L}_{\text{scalar}}$ without supermassive modes appearing and to the fact that ϕ has x^5 dependence. This in turn led to a relation between the electric charge and the circumference of the fifth dimension and to the quantization of the charge. We end up with a completely satisfactory explanation of why charge is quantized based on (12) in the fiber bundle.

III. CONCLUSION

Fiber bundles are a natural and elegant language for the description of a gauge theory. We have found a Lagrangian density (12) for a scalar matter field ϕ coupled to general relativity and electromagnetism in a five-dimensional fiber bundle which is invariant under any choice of basis which preserves the fiber bundle structure. This Lagrangian density reduces to the usual Klein–Gordon expression in four dimensions with the circumference of the fifth dimension inversely proportional to the charge and with superheavy masses not appearing. The charge is quantized because ϕ lives in a principal fiber bundle with a U_1 gauge group (isomorphic to a circle). Thus ϕ must be periodic in the x^5 coordinate with period equal to the circumference of the fifth dimension. The fact that the charge is related to this circumference is consistent with the work of Bergmann[7] who showed this circumference to be a constant of the space. The realization that we are dealing with a fiber bundle and not a general five-dimensional manifold allows us to restrict the Lagrangian density to a form without superheavy mass present. The fact that the fiber bundle plays a crucial role here and leads to the first satisfactory explanation for charge quantization, without unphysical elements arising, suggests that the elegant fiber bundle models of gauge theories contain considerable truth value.

[1] P. A. Dirac, Proc. R. Soc. London Ser. A **133**, 60 (1931); Phys. Rev **74**, 817(1948).
[2] C. N. Yang, Phys. Rev. D **12**, 3845 (1975).
[3] T. Kaluza, Sitz. Preuss. Akad. Wiss. Phys. Math. **54**, 966 (1921).
[4] O. Klein, Nature **118**, 516 (1926).
[5] A. Einstein, V. Bargmann, and P. G. Bergmann, *Theodore von Kármán Anniversary Volume* (California Institute of Technology, Pasadena, 1941), p. 212.
[6] A. Einstein and P. Bergmann, Ann. Math. **39**, 683 (1938).
[7] P. G. Bergmann, *Introduction to the Theory of Relativity* (Dover, New York, 1976), p. 273.
[8] Y. M. Cho, J. Math. Phys. **16**, 2029 (1975).
[9] O. Klein, Z. Phys. **37**, 875 (1926).
[10] J.-M. Souriau, Nuovo Cimento **30**, 565 (1963).
[11] A. Chodos and S. Detweiler, Phys. Rev. D **21**, 2167 (1980).
[12] D. J. Gross and M. J. Perry, Nuc. Phys. B **226**, 29 (1983).
[13] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1969).
[14] N. Steenrod, *The Topology of Fibre Bundles* (Princeton U. P., Princeton, 1951).
[15] D. Husemoller, *Fiber Bundles* (Springer, New York, 1966).
[16] M. Daniel and C. Viallet, Rev. Mod. Phys. **52**, 175 (1980).
[17] C. Nash and S. Sen, *Topology and Geometry for Physicists* (Academic, London, 1983).
[18] M. E. Mayer, *Fiber Bundle Techniques in Gauge Theories* (Springer, New York, 1977).
[19] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).

Duality rotations and type D solutions to Einstein equations with nonlinear electromagnetic sources

Humberto Salazar I.,[a] Alberto García D., and Jerzy Plebański[b]
Centro de Investigación y de Estudios Avanzados del IPN, Departamento de Física, Apartado Postal 14-740, 07000 México, D. F., Mexico

(Received 7 May 1985; accepted for publication 1 April 1987)

Within nonlinear electrodynamics of Born–Infeld type allowing for the freedom of duality rotations, explicit type D solutions are constructed. The obtained type D solutions, which generalize the charged Taub–NUT (Newman–Unti–Tamburino) metric with λ, exhaust all solutions within the considered class, under the assumption that the real eigenvectors of the electromagnetic field are aligned along the geodesic and shear-free principal null directions. Various relevant limiting transitions, in particular those of the flat space-time, are studied in some detail.

I. INTRODUCTION

More than fifty years ago Born and Infeld published their relevant papers[1,2] about nonlinear electrodynamics proposing a consistent convergent classical field theory. Further developments[3,4] have led to the conclusion that nonlinear electrodynamics can be obtained from QED—in the limit of high occupation numbers—and therefore can be considered as a classical model of the vacuum polarization processes, concluding finally in a spectacular derivation of the Schwinger Lagrangian.[5]

The basic problems of the theories of Born–Infeld type are the following: the still somewhat unsettled status of the "correct" structural function, and the technical difficulty of devising solutions to nonlinear equations. If the structural function is to be considered as derived from QED it should be presently deduced taking into account the established facts concerned with the unification of the electromagnetic and weak interactions. This likely can be relatively easily done. However, if the Dirac monopoles really exist, the structural function should be derived from a quantum field theory that incorporates both types of charges—magnetic and electric—and that with large magnetic charge, does not allow for a perturbative treatment. With the charges modeled in nonlinear electrodynamics as the singularities of the field, as the only formal idea when magnetic monopole charges are included, there is still the possibility of demanding that the dynamical equations of nonlinear electrodynamics should allow the properly generalized freedom of the duality rotations. We will show in this paper how this can be consistently arranged.

Now, as far as the technical difficulties in handling the nonlinear equations are concerned, indeed, even on the level of special relativity, the exact solutions of the dynamical scheme of Born–Infeld are rather scarce. Most of the literature of the 1930's has dealt only with the spherically symmetric solution corresponding to various *Ansätze* for the structural function. On the level of general relativity we encounter about the same situation; after the early paper,[6] the spherically symmetric case is studied more generally in Ref. 7, although the later paper[8] has shown that the Robinson[9]–Bertotti[10] metric can be considered as a carrier of the Born–Infeld structure.

The basic goal of this paper consists in constructing—by employing more up to date techniques of the theory of exact solutions in general relativity—some nontrivial solutions to the dynamical scheme of nonlinear electrodynamics endowed with the freedom of the duality rotations which permits the inclusion of magnetic monopole charges.

The most likely class of metrics that can relatively easily be shown to be the carrier of Born–Infeld structure can be guessed to be contained within the metrics of type D. We thus explored all known D branches in the assumption of the alignment of real eigenvectors of the electromagnetic field to the shear-free and geodesic principal null directions, concluding that the only "viable" branches from our point of view are the Carter[11] separable $\widetilde{B}^{(\pm)}$ branches and their limiting contractions.

In this situation we have structured this paper as follows. Section II gives the description of nonlinear electrodynamics in general relativity in terms of the null tetrad formalism, and defines the class of nonlinear theories endowed with the freedom of the duality rotations. In Sec. III we show that $\widetilde{B}^{(\pm)}$ branches are indeed carriers of solutions to the dynamical scheme of nonlinear electrodynamics, deriving explicit solutions which generalize the Taub–NUT (Newman–Unti–Tamburino) charged solutions with λ (Refs. 12 and 13) to the case of "nonlinear charges" e and $\check{g}$. Section IV examines the special relativistic limits of the subbranches of our solutions. The final section discusses the completeness of our solutions within the type D solutions and adds some closing remarks.

II. NONLINEAR ELECTRODYNAMICS AND THE DUALITY ROTATIONS

In nonlinear theories of Born–Infeld type the electromagnetic field has the Lagrangian

$$\mathscr{L}_{\mathscr{E}} = -(1/4\pi)\{\tfrac{1}{2}p^{\mu\nu}f_{\mu\nu} - \mathscr{H}(\mathscr{P},\check{\mathscr{Q}})\}, \tag{2.1}$$

where $f_{\mu\nu}$ is a curl ($\check{f}_{\mu\nu}{}^{;\nu} = 0 \leftrightarrow f_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}$) and cor-

[a] Also at Escuela de Ciencias Físico–Matemáticas, Universidad Autónoma de Puebla, Puebla, Mexico.
[b] On leave of absence from the University of Warsaw, Warsaw, Poland.

 0022-2488/87/092171-11$02.50

responds to the intensity of electric field and magnetic induction vectors ($\mathbf{E}$ and $\check{\mathbf{B}}$), while tensor $p^{\mu\nu}$ corresponds to the electric induction and the intensity of the magnetic field vectors ($\mathbf{D}$ and $\mathbf{H}$). The *structual function* $\mathscr{H}$ depends on the two invariants of the tensor $p_{\mu\nu}$,

$$\mathscr{P}:=\tfrac{1}{4}p^{\mu\nu}p_{\mu\nu},\quad \check{\mathscr{Q}}:=\tfrac{1}{4}\check{p}^{\mu\nu}p_{\mu\nu};\qquad \check{p}_{\mu\nu}:=(i/2\sqrt{-g})\epsilon^{\mu\nu\rho\sigma}p_{\rho\sigma}. \tag{2.2}$$

One constrains the admissible structural functions by demanding (i) the correspondence to the linear theory $[\mathscr{H}=\mathscr{P}+O(\mathscr{P}^2,\check{\mathscr{Q}}^2)]$, (ii) the parity conservation $[\mathscr{H}(\mathscr{P},\check{\mathscr{Q}})=\mathscr{H}(\mathscr{P},-\check{\mathscr{Q}})]$, (iii) the positive definiteness of the energy density ($\mathscr{H}_{\mathscr{P}}>0$), and (iv) the requirement of the timelike nature of the energy flux vector ($\mathscr{P}\mathscr{H}_{\mathscr{P}}+\check{\mathscr{Q}}\mathscr{H}_{\check{\mathscr{Q}}}-\mathscr{H}\geqslant 0$), see Ref. 14.

The action $S=\int d_4x\sqrt{-g}(R+2\lambda+\mathscr{L}_{\mathscr{E}})$, extremalized with respect to $g_{\mu\nu}$, A_μ, and $p_{\mu\nu}$ leads to the dynamical equations:

Einstein equations,

$$G_{\mu\nu}=8\pi E_{\mu\nu}+\lambda g_{\mu\nu}, \tag{2.3a}$$

where

$$4\pi E_{\mu\nu}=-f_\mu{}^\lambda p_{\nu\lambda}+g_{\mu\nu}\mathscr{L},\quad \mathscr{L}:=-4\pi\mathscr{L}_{\mathscr{E}}; \tag{2.3b}$$

Faraday equations,

$$\check{f}^{\mu\nu}{}_{;\nu}=0\leftrightarrow f_{\mu\nu}=A_{\mu,\nu}-A_{\nu,\mu}; \tag{2.3c}$$

Maxwell equations,

$$p^{\mu\nu}{}_{;\nu}=0; \tag{2.3d}$$

and additionally—in analogy with the Lorentz theory of electrons—

the "material equations,"

$$f_{\mu\nu}=\mathscr{H}_{\mathscr{P}}p_{\mu\nu}+\mathscr{H}_{\check{\mathscr{Q}}}\check{p}_{\mu\nu}. \tag{2.4}$$

This basic description of the dynamical equations of nonlinear electrodynamics within general relativity can be now easily "translated" into a description in terms of the null tetrad formalism of Debney–Kerr–Schild[15] according to which the metric is given by

$$g=2e^1\otimes e^2+2e^3\otimes e^4,\quad e^2=(\bar{e}^1),\qquad e^3=(\bar{e}^3),\quad e^4=(\bar{e}^4), \tag{2.5}$$

where the $e^a\epsilon\Lambda^1$ have to fulfill the first Cartan structure equations

$$de^a=e^b\wedge\Gamma^a{}_b=\Gamma^a{}_{bc}e^b\wedge e^c, \tag{2.6}$$

with $\Gamma^a{}_b\epsilon\Lambda^1$ satisfying the second structure equations

$$d\Gamma^a{}_b+\Gamma^a{}_s\wedge\Gamma^s{}_b=\tfrac{1}{2}R^a{}_{bcd}e^c\wedge e^d. \tag{2.7}$$

(The tetrad components of a tensorial quantity are determined from the coordinate components according to

$$T^{ab\cdots}{}_{cd\cdots}=e^a_\mu e^b_\gamma\cdots T^{\mu\nu\cdots}{}_{\alpha\beta\cdots}e^\alpha_c e^\beta_d\cdots.)$$

The Riemann curvature components $R^a{}_{bcd}$ may be replaced by the Weyl conformal tensor components, which are characterized by five complex curvature coefficients $C^{(a)}$, $a=1,\ldots,5$, and the components of the traceless Ricci tensor $C_{ab}:=R_{ab}-\frac{1}{4}g_{ab}R$, where $R_{ab}:=R^s{}_{abs}$, and $R=R^a{}_a$.

The Maxwell–Faraday equations of nonlinear electrodynamics can be conveniently stated in terms of differential forms. Indeed, $\check{f}^{\mu\nu}{}_{;\nu}=0$ is equivalent to $f_{\mu\nu}$ being a curl; similarly, $p^{\mu\nu}{}_{;\nu}=0$ is equivalent to $\check{p}_{\mu\nu}$ being a curl. Therefore, the Maxwell–Faraday equations (2.3c) and (2.3d) are equivalent to a complex condition

$$\omega:=\tfrac{1}{2}(f_{\mu\nu}+\check{p}_{\mu\nu})dx^\mu\wedge dx^\nu=\tfrac{1}{2}(f_{ab}+\check{p}_{ab})e^a\wedge e^b\to d\omega=0. \tag{2.8}$$

Expressing $f_{\mu\nu}$ in $E_{\mu\nu}$ in terms of $p_{\mu\nu}$, one easily infers that the Einstein equations (2.3a), in the null tetrad description, amount to

$$G_{ab}=8\pi E_{ab}+\lambda g_{ab}, \tag{2.9}$$

where

$$4\pi E_{ab}=\mathscr{H}_{\mathscr{P}}(-p_{as}p_b{}^s+g_{ab}\mathscr{P})+(\mathscr{P}\mathscr{H}_{\mathscr{P}}+\check{\mathscr{Q}}\mathscr{H}_{\check{\mathscr{Q}}}-\mathscr{H})g_{ab}, \tag{2.10}$$

which, in particular, imply

$$R=-4\lambda-8(\mathscr{P}\mathscr{H}_{\mathscr{P}}+\check{\mathscr{Q}}\mathscr{H}_{\check{\mathscr{Q}}}-\mathscr{H}). \tag{2.11}$$

The material equations in terms of null tetrad components are now

$$\check{f}^{ab}{}_{;b}=0=p^{ab}{}_{;b},\quad f_{ab}=\mathscr{H}_{\mathscr{P}}p_{ab}+\mathscr{H}_{\check{\mathscr{Q}}}\check{p}_{ab}. \tag{2.12}$$

The null tetrad image of the tensorial duality operation (2.2) is defined by

$$\check{p}^{ab}:=-\tfrac{1}{2}\epsilon^{abcd}p_{cd}\to p_{ab}=\check{\check{p}}_{ab}. \tag{2.13}$$

Until now the choice for the null tetrad was left arbitrary in these considerations. Assuming now that the electromagnetic field is *algebraically general*, i.e.,

$$\mathscr{P}+\check{\mathscr{Q}}=\tfrac{1}{4}p^{\mu\nu}p_{\mu\nu}+\tfrac{1}{4}\check{p}^{\mu\nu}p_{\mu\nu}\neq 0\to F+\check{G}=\tfrac{1}{4}f^{\mu\nu}f_{\mu\nu}+\tfrac{1}{4}\check{f}^{\mu\nu}f_{\mu\nu}\neq 0, \tag{2.14}$$

we select the null tetrad in such a manner that out of all independent components of p_{ab} and f_{ab} they are different from zero only

$$p_{34}=D,\quad p_{12}=i\check{H},\quad f_{34}=E,\quad f_{12}=i\check{B}, \tag{2.15}$$

where D, $\check{H}$, E, and $\check{B}$ are real.

Geometrically, the assumption that p_{ab} and f_{ab} have only nontrivial independent components (2.15) means that the null tetrad is chosen to coincide with the common eigenvectors of f_{ab} and p_{ab}.

The invariants (2.14) are now

$$\mathscr{P}+\check{\mathscr{Q}}=-\tfrac{1}{2}(D+i\check{H})^2\neq 0,\qquad F+\check{G}=-\tfrac{1}{2}(E+i\check{B})^2\neq 0, \tag{2.16}$$

so that $(D,\check{H})$ and $(E,\check{B})$ can be interpreted as *independent* parameters of the complex invariants of the electromagnetic field, and therefore invariants as such.

The essential advantage of our choice for the null tetrad is that E_{ab} from (2.10) diagonalizes; out of all independent components of this object only E_{12} and E_{34} are different from zero, and are given by

$$\left.\begin{matrix}4\pi E_{12}\\4\pi E_{34}\end{matrix}\right\}=\mp\tfrac{1}{2}(D^2+\check{H}^2)\mathscr{H}_{\mathscr{P}}+(\mathscr{P}\mathscr{H}_{\mathscr{P}}+\check{\mathscr{Q}}\mathscr{H}_{\check{\mathscr{Q}}}-\mathscr{H}). \tag{2.17}$$

Used in Einstein equations from (2.9)—with $G_{ab} = C_{ab} - \frac{1}{4}g_{ab}R$—this means that the information contained in Einstein equations reduces to

$$C_{ab} = 0, \text{ except for } C_{12} = -C_{34} = -(D^2 + \check{H}^2)\mathscr{H}_{\mathscr{P}}, \tag{2.18}$$

while R is given by (2.11).

Next, within the present assumptions—employing the tetrad duality operation (2.13)—we easily find that the Maxwell–Faraday equations (2.8) reduce to

$$\omega = (D + i\check{B})e^1 \wedge e^2 + (E + i\check{H})e^3 \wedge e^4 \rightarrow d\omega = 0. \tag{2.19}$$

Similarly, one easily finds that under the present assumptions the material equations amount to

$$E + i\check{B} = (\mathscr{H}_{\mathscr{P}} + \mathscr{H}_{\check{\mathscr{Q}}})(D + i\check{H}). \tag{2.20}$$

In this null tetrad description of the dynamical scheme of nonlinear electrodynamics in general relativity, the main weight of the differential structure is shifted onto that *null tetrad* e^a subjected to $C_{ab} = 0$, except for $C_{12} = -C_{34} \neq 0$. This is even more evident if one tries to solve the considered dynamical equations as in Ref. 16 under a strong restrictive assumption that $P + \check{\mathscr{Q}} = \text{const} \neq 0$, and hence all D, $\check{H}$, E, and $\check{B}$ are constants. This subcase studied in Ref. 16 according to the Newman–Penrose (NP) formalism[17] is of some interest, since within its ideology any standard electrovac solution with $F + \check{G} = \text{const} \neq 0$—and necessarily with $\lambda \neq 0$—can be reinterpreted as a solution to the general Born–Infeld scheme.

Now, the studied differential structure, Eqs. (2.11) and (2.17)–(2.20), can be equivalently described in terms of alternative structural functions obtained from $\mathscr{H}$ via Legendre transformations. Indeed, the structural function $\mathscr{H}(P,\check{Q})$, with $\mathscr{P} = -\frac{1}{2}(D^2 - \check{H}^2)$, $\check{\mathscr{Q}} = -i D\check{H}$, can be considered as a function of the independent variables D and $\check{H}$, $\mathscr{H}(D,\check{H})$. The material equations (2.20), in terms of $\mathscr{H}(D,\check{H})$, reduce now to $E = -\mathscr{H}_D$, $\check{B} = \mathscr{H}_{\check{H}}$, therefore

$$d\mathscr{H} = -E\,dD + \check{B}\,d\check{H}. \tag{2.21}$$

Defining the functions

$$\mathscr{M}^{(+)} := \check{B}\check{H} - \mathscr{H}, \quad \mathscr{M}^{(-)} := ED + \mathscr{H},$$
$$\mathscr{M}^{(+)} + \mathscr{M}^{(-)} = DE + \check{B}\check{H}, \tag{2.22}$$
$$\mathscr{L} = -DE + \check{B}\check{H} - \mathscr{H},$$

we have

$$d\mathscr{M}^{(+)} = E\,dD + \check{H}\,d\check{B}, \quad d\mathscr{M}^{(-)} = D\,dE + \check{B}\,d\check{H},$$
$$d\mathscr{L} = -D\,dE + \check{H}\,d\check{B}. \tag{2.23}$$

One easily sees that the so defined $\mathscr{L}$ precisely coincides with $\mathscr{L} = -4\pi\mathscr{L}_{\mathscr{E}}$, see (2.1), i.e., with the original Lagrangian of the theory, which can be considered as a function of $(E,\check{B})$ or, equivalently, of the invariants $(F,\check{G})$.

Most of the early papers on nonlinear electrodynamics work with $\mathscr{L} = \mathscr{L}(F,\check{G})$ as the fundamental structural function.

When one works with the equivalent $\mathscr{H}$ or "energy function" introduced in Ref. 18 considered as the fundamental, one has the technical advantage that $f_{\mu\nu}$'s are given directly by material equations as functions of $p_{\mu\nu}$'s. Of course, the theory is equivalently uniquely defined with any of our "middle of the road" structural functions

$$\mathscr{M}^{(+)} = \mathscr{M}^{(+)}(D,\check{B}) \text{ and } \mathscr{M}^{(-)} = \mathscr{M}^{(-)}(E,\check{H}) \tag{2.24}$$

considered as given and fundamental. In principle, given and known any of the four functions $\mathscr{H}(D,\check{H})$, $L(E,\check{B})$, $\mathscr{M}^{(+)}(D,\check{B})$, and $\mathscr{M}^{(-)}(E,\check{H})$, the remaining three functions can be always calculated.

According to (2.21) and (2.23) the material equations thus have four equivalent representations

$$\begin{aligned} E + i\check{B} &= \left(-\frac{\partial}{\partial D} + i\frac{\partial}{\partial \check{H}}\right)\mathscr{H}, \\ D + i\check{H} &= \left(-\frac{\partial}{\partial E} + i\frac{\partial}{\partial \check{B}}\right)\mathscr{L}, \\ E + i\check{H} &= \left(\frac{\partial}{\partial D} + i\frac{\partial}{\partial \check{B}}\right)\mathscr{M}^{(+)}, \\ D + i\check{B} &= \left(\frac{\partial}{\partial E} + i\frac{\partial}{\partial \check{H}}\right)\mathscr{M}^{(-)}. \end{aligned} \tag{2.25}$$

Working with the structural functions $\mathscr{M}^{(\pm)}$, the field equations amount to the following:

(i) Einstein equations,

$$\begin{aligned} R &= -4\lambda \mp 8\mathscr{M}^{(\pm)} \pm 4(DE + \check{B}\check{H}), \\ C_{12} &= -C_{34} = -(DE + \check{B}\check{H}); \end{aligned} \tag{2.26}$$

(ii) Maxwell–Faraday equations,

$$\omega := (D + i\check{B})e^1 \wedge e^2 + (E + i\check{H})e^3 \wedge e^4 \rightarrow d\omega = 0; \tag{2.27}$$

(iii) material equations,

$$E + i\check{H} = \left(\frac{\partial}{\partial D} + i\frac{\partial}{\partial \check{B}}\right)\mathscr{M}^{(+)} \leftrightarrow D + i\check{B} = \left(\frac{\partial}{\partial E} + i\frac{\partial}{\partial \check{H}}\right)\mathscr{M}^{(-)}. \tag{2.28}$$

The special Bianchi identities, valid with $C_{ab} = 0$ except for $C_{12} = -C_{34}$, are

$$\begin{aligned} &\partial_1{}^{(-)} + (\Gamma_{413} + \Gamma_{314})(\mathscr{M}^{(+)} + \mathscr{M}^{(-)}) = 0, \\ &\partial_2{}^{(-)} + (\Gamma_{423} + \Gamma_{324})(\mathscr{M}^{(+)} + \mathscr{M}^{(-)}) = 0, \\ &\partial_3{}^{(+)} - (\Gamma_{312} + \Gamma_{321})(\mathscr{M}^{(+)} + \mathscr{M}^{(-)}) = 0, \\ &\partial_4{}^{(+)} - (\Gamma_{421} + \Gamma_{412})(\mathscr{M}^{(+)} + \mathscr{M}^{(-)}) = 0, \end{aligned} \tag{2.29}$$

where Γ_{abc} are the connection coefficients defined in (2.6).

At the same time, we find that Maxwell–Faraday equations $d\omega = 0$, worked out by using the first structure equations, amount to four complex conditions

$$\begin{aligned} &\partial_1(E + i\check{H}) + (\Gamma_{314} - \Gamma_{413})(D + i\check{B}) \\ &\quad + (\Gamma_{314} + \Gamma_{413})(E + i\check{H}) = 0, \\ &\partial_2(E + i\check{H}) + (\Gamma_{423} - \Gamma_{324})(D + i\check{B}) \\ &\quad + (\Gamma_{324} + \Gamma_{423})(E + i\check{H}) = 0, \\ &\partial_3(D + i\check{B}) - (\Gamma_{312} + \Gamma_{321})(D + i\check{B}) \\ &\quad - (\Gamma_{312} - \Gamma_{321})(E + i\check{H}) = 0, \end{aligned} \tag{2.30}$$

$$\partial_4(D+i\check{B})-(\Gamma_{421}+\Gamma_{412})(D+i\check{B})$$
$$-(\Gamma_{421}-\Gamma_{412})(E+i\check{H})=0.$$

Nonlinear theory allowing for the freedom of duality rotations: We should like now to determine a subclass of nonlinear theories that admits the freedom of the *duality rotations*, hence being compatible with the existence of magnetic monopole charges.

We begin by observing that with $\check{\phi}_0=\text{const}$, given D, $\check{H}$, E, and $\check{B}$, which fulfill (2.27) and (2.28), the new "duality rotated" objects

$$D'+i\check{B}=e^{i\check{\phi}_0}(D+i\check{B}),\quad E'+i\check{H}':=e^{i\check{\phi}_0}(E+i\check{H}) \tag{2.31}$$

also satisfy Eqs. (2.27) and (2.28). Indeed, as far as the Maxwell–Faraday equations (2.27) are concerned, this statement is trivial; with ω closed, certainly $\omega':=e^{i\check{\phi}_0}\omega$ is closed when $\check{\phi}_0=\text{const}$. On the other hand, the material equations from (2.28) are also invariant under the considered transformation. Now, as far as the Einstein equations are concerned, we easily verify first that under the transformation (2.31) the expression

$$E'D'+\check{B}'\check{H}'=ED+\check{B}\check{H} \tag{2.32}$$

is invariant, and hence so are the components of Einstein equations which involve C_{ab}. On the other hand, the right-hand member of R from (2.26) in the form involving the structural function $\mathscr{M}^{(+)}$, in general *is not* invariant with respect to the duality rotations. It becomes invariant iff the structural function $\mathscr{M}^{(+)}=\mathscr{M}^{(+)}(D,\check{B})$ is such that for *every* $\check{\phi}_0$

$$\mathscr{M}^{(+)}[\cos\check{\phi}_0 D-\sin\check{\phi}_0\check{B},\ \sin\check{\phi}_0 D+\cos\check{\phi}_0\check{B}]$$
$$\equiv\mathscr{M}^{(+)}(D,\check{B}). \tag{2.33}$$

This last condition can be easily seen to constrain the function $\mathscr{M}^{(+)}$ to a function of *one variable* $\frac{1}{2}(D^2+\check{B}^2)$ only

$$\mathscr{M}^{(+)}=\mathscr{M}^{(+)}[\tfrac{1}{2}(D^2+\check{B}^2)]. \tag{2.34}$$

Therefore, with the freedom of the duality rotations defined by the condition that "given e^a, $D+i\check{B}$, and $E+i\check{H}$ which satisfy (2.26)–(2.28), then e^a, $D'+i\check{B}$, and $E'+i\check{H}'$ from (2.31) are also a solution to (2.26)–(2.28) for arbitrary $\check{\phi}_0=\text{const}$," the argument given above clearly constrains nonlinear theories endowed with that freedom to those ones with $\mathscr{M}^{(+)}$ from (2.34).

It will be convenient to understand $\mathscr{M}^{(+)}$ from (2.34) as determined by an arbitrary dimensionless function of a dimensionless variable, $f^{(+)}(x)$, in the form

$$\mathscr{M}^{(+)}:=b^2\mathscr{f}^{(+)}(x),\quad x:=(1/2b^2)(D^2+\check{B}^2)=:x^{(+)}, \tag{2.35}$$

where b is a constant of dimension of electromagnetic field [$\equiv$in gravitational units to (length)$^{-1}$]. Working with this structural function, denoting the derivative of $\mathscr{f}(x)$ with respect to x by superscript ∇, one easily sees that the condition of the correspondence to the linear theory for weak fields now takes the form

$$\mathscr{f}^{(+)}(x)=x+O(x^2)\leftrightarrow\mathscr{f}^{\nabla(+)}(0)=1, \tag{2.36}$$

while the parity conservation is automatically assured. The conditions of positive definiteness of the energy and the time-like character of the energy flux are then easily seen to be equivalent to

$$\mathscr{f}^{(+)}(x)/x\geqslant\mathscr{f}^{\nabla(+)}(x)>0. \tag{2.37}$$

Knowing $\mathscr{M}^{(+)}$ in the form of (2.35), the remaining three structural functions ($\mathscr{M}^{(-)}$, $\mathscr{H}$, and $\mathscr{L}$) can be worked out. In particular, we find that

$$\mathscr{M}^{(-)}=b^2\mathscr{f}^{(-)}(y),\quad y:=(1/2b^2)(E^2+\check{H}^2)=:x^{(-)}, \tag{2.38}$$

where

$$\mathscr{f}^{(-)}(y)=2x\mathscr{f}^{\nabla(+)}(x)-\mathscr{f}^{(+)}(x),\qquad x=x(y),\quad y:=x(\mathscr{f}^{\nabla(+)})^2. \tag{2.39}$$

Notice that

$$x\mathscr{f}^{\nabla(+)}(x)=y\mathscr{f}^{\nabla(-)}(y),\quad \mathscr{f}^{\nabla(+)}(x)\cdot\mathscr{f}^{\nabla(-)}(y)=1,$$
$$\mathscr{f}^{(+)}(x)=2y\mathscr{f}^{\nabla(-)}(y)-\mathscr{f}^{(-)}(y), \tag{2.40}$$

where superscript ∇ denotes the derivatives of $f^{(+)}(x)$ and $f^{(-)}(y)$ with respect to their arguments.

Observe also that the material equations (2.28) now become

$$E+i\check{H}=(D+i\check{B})\mathscr{f}^{\nabla(+)},\qquad D+i\check{B}=(E+i\check{H})\mathscr{f}^{\nabla(-)}. \tag{2.41}$$

Hence $E+i\check{H}$ and $D+i\check{B}$ have a common phase and can be thus parametrized according to

$$E+i\check{H}=(2b^2x^{(-)})^{1/2}e^{i\check{\phi}},\qquad D+i\check{B}=(2b^2x^{(+)})^{1/2}e^{i\check{\phi}}. \tag{2.42}$$

Now, the closure condition of the two-form ω, $d\omega=0$, modulo the special Bianchi identities (2.29), implies

$$i\,d\check{\phi}+\mathscr{f}^{\nabla(-)}[(\Gamma_{314}-\Gamma_{413})e^1+(\Gamma_{423}-\Gamma_{324})e^2]$$
$$-\mathscr{f}^{\nabla(+)}[(\Gamma_{312}-\Gamma_{321})e^3+(\Gamma_{421}-\Gamma_{412})e^4]=0. \tag{2.43}$$

An important example; the Born–Infeld theory: We specialize now $\mathscr{f}^{(+)}(x)$ for the specific case of

$$\mathscr{f}^{(+)}=\sqrt{1+2x}-1, \tag{2.44}$$

compatible with conditions (2.36) and (2.37). It follows from (2.39) that $\mathscr{f}^{(-)}=\mathscr{f}^{(-)}(y)$ is given by

$$\mathscr{f}^{(-)}=1-\sqrt{1-2y}, \tag{2.45}$$

so that

$$\mathscr{M}^{(+)}=b^2\{\sqrt{1+b^{-2}(D^2+\check{B}^2)}-1\},\qquad \mathscr{M}^{(-)}=b^2\{1-\sqrt{1-b^{-2}(E^2+\check{H}^2)}\}; \tag{2.46}$$

while the equivalent structural functions $\mathscr{H}$ and $\mathscr{L}$ can be computed as given by

$$\mathscr{H}=b^2-\sqrt{(b^2+D^2)(b^2-\check{H}^2)}$$
$$\equiv b^2-\sqrt{b^4-2b^2\mathscr{P}+\mathscr{Q}^2},$$

and

$$\mathscr{L}=\sqrt{(b^2-E^2)(b^2+\check{B}^2)}-b^2$$
$$\equiv\sqrt{b^4+2b^2F+\check{G}^2}-b^2$$

$$\leftrightarrow \sqrt{-g}\,\mathscr{L} = b^2\{\sqrt{-\det(g_{\mu\nu} + b^{-1}f_{\mu\nu})} - \sqrt{-\det(g_{\mu\nu})}\}. \quad (2.47)$$

In $\mathscr{L}$ above we recognize the original Born–Infeld Lagrangian proposed exactly 52 years ago.[1,2] According to the second equation of (2.47) $\sqrt{-g}\,\mathscr{L}$ is a *natural* tensor density. The Born–Infeld Lagrangian is *exceptional* as the only one which leads to *single* characteristic cones for the propagation of small perturbations of the electromagnetic field.[14,19] For further developments concerned with characteristics and discontinuities in nonlinear theories see Ref. 20.

The fact that the original Born–Infeld theory is a special case of a theory endowed with the freedom of the duality rotations reinforces our interest in the general class of theories of this type, with an arbitrary—modulo (2.36) and (2.37)—function $\mathscr{f}^{(+)} = \mathscr{f}^{(+)}(x)$, *and b consistently interpreted as the Born constant*, related in the convergent theories to the radius of the electron.

III. EXPLICIT D-TYPE CARTER $\widetilde{B}(\pm)$ SOLUTIONS TO EINSTEIN EQUATIONS WITH NONLINEAR ELECTROMAGNETIC SOURCES

Our objective is now to show that the Carter $\widetilde{B}^{(\pm)}$ branches of D-type metrics are *carriers* of solutions to the dynamical scheme of nonlinear electrodynamics whose structural function allows for the freedom of the duality rotations.

We shall assume that (i) the natural tetrads of the $\widetilde{B}^{(\pm)}$ metrics are aligned along the eigenvectors of the algebraically general electromagnetic field, and (ii) the nonlinearity of the electromagnetic structure is characterized for the $\widetilde{B}^{(+)}$ and $\widetilde{B}^{(-)}$ metrics by the functions $\mathscr{f}^{(+)}(x)$ and $\mathscr{f}^{(-)}(y)$. Of course (ii) is assumed for convenience only; $\mathscr{M}^{(+)} = b^2\mathscr{f}^{(+)}$ and $\mathscr{M}^{(-)} = b^2\mathscr{f}^{(-)}$ describe *the same* nonlinear structure endowed with the freedom of the duality rotations.

Let the Carter $\widetilde{B}^{(\pm)}$ type D metrics, with signature $(+++-)$, be given in the charts $\{x^\mu\} := \{\zeta,\bar{\zeta},r,\tau\}$ and $\{x^\mu\} := (u,v,r,\sigma)$ by

$$g := (r^2 + l^2)dS^{2(\pm)} + \frac{r^2 + l^2}{\mathscr{F}^{(\pm)}}\, dr \underset{s}{\otimes} dr \mp \frac{\mathscr{F}^{(\pm)}}{r^2 + l^2}\, \pi^{(\pm)} \underset{s}{\otimes} \pi^{(\pm)}, \quad (3.1)$$

where

$$dS^{2(\pm)} := \begin{cases} 4\dfrac{d\zeta \underset{s}{\otimes} d\bar{\zeta}}{(1+\epsilon\zeta\bar{\zeta})^2}, \\[2ex] 4\dfrac{du \underset{s}{\otimes} dv}{(1+\epsilon uv)^2}, \end{cases} \quad (3.2)$$

and

$$\pi^{(\pm)} := \begin{cases} d\tau + 2il\dfrac{\bar{\zeta}\, d\zeta - \zeta\, d\bar{\zeta}}{(1+\epsilon\zeta\bar{\zeta})}, \\[2ex] d\sigma + 2l\dfrac{v\, du - u\, dv}{(1+\epsilon uv)}. \end{cases} \quad (3.3)$$

In these expressions, $(+)$ corresponds to the $\widetilde{B}^{(+)}$ metric, while the quantities with sign $(-)$ determine the structure of the $\widetilde{B}^{(-)}$ solution. The real coordinates r, τ, and σ are of dimension of length, while the complex ζ and the real u and v coordinates are dimensionless. The parameter ϵ is a dimensionless constant, and l is the constant of dimension of length. The gravitational constant G and the velocity of light c are equated to the unit. The real analytic functions $\mathscr{F}^{(\pm)} = \mathscr{F}^{(\pm)}(r)$ are of dimension $(\text{length})^2$. In the $\widetilde{B}^{(-)}$ case, in order to assure the signature $(+++-)$, we must require $\mathscr{F}^{(-)} > 0$. In the case of $\widetilde{B}^{(+)}$ the sign of $\mathscr{F}^{(+)}$ remains in principle arbitrary, with the values of r for which $\mathscr{F}^{(+)}(r) = 0$ expected to correspond to the causal horizons.

It can be easily shown that the Carter[11] separable $\widetilde{B}^{(\pm)}$ branches of type D can be brought—without any loss of generality—to the form of (3.1).

In the considered representation of the $\widetilde{B}^{(\pm)}$ metrics, only the sign of the parameter ϵ is relevant. By a scaling transformation ϵ can be brought to the discrete values $1,0,-1$. When $\epsilon^2 = 1$, the curvatures of $dS^{2(\pm)}$ have the unit radii. In subsequent considerations we shall follow in principle this normalization for ϵ, although many forthcoming formulas—in particular those for the natural tetrad and its connections—remain valid apart from this normalization; the last remark is useful when one studies the limiting transitions (contractions) of $\widetilde{B}^{(\pm)}$ metrics. Notices that a formal transformation

$$\zeta \to -u, \quad \bar{\zeta} \to -v, \quad \tau \to i\sigma, \quad \mathscr{F}^{(+)} \to \mathscr{F}^{(-)}, \quad (3.4)$$

which obviously implies $dS^{2(+)} \to ds^{2(-)}$, $\pi^{(+)} \to i\pi^{(-)}$, brings the $\widetilde{B}^{(+)}$ metric into the $\widetilde{B}^{(-)}$ metric. Therefore, if we consider the $\widetilde{B}^{(+)}$ branch as a *complexified* space-time,[21] then both real $\widetilde{B}^{(\pm)}$ metrics can be interpreted as the two different real cuts[22] of the same complex structure. Hence one can restrict oneself to the determination of the $\widetilde{B}^{(+)}$ metric and derive the $\widetilde{B}^{(-)}$ solution according to (3.4); in this way we shall proceed in what follows, omitting for typing reasons the sign $(+)$.

A natural choice of the null tetrads for the $\widetilde{B}^{(+)}$ metric is

$$\left.\begin{matrix} e^1 \\ e^2 \end{matrix}\right\} = -\sqrt{2}\cdot\frac{(r^2+l^2)^{1/2}}{1+\epsilon\zeta\bar{\zeta}}\begin{cases} d\zeta, \\ d\bar{\zeta}, \end{cases}$$
$$\left.\begin{matrix} e^3 \\ e^4 \end{matrix}\right\} = \frac{1}{\sqrt{2}}\left(\frac{r^2+l^2}{\mathscr{F}}\right)^{1/2} dr \pm \left(\frac{\mathscr{F}}{r^2+l^2}\right)^{1/2}\pi. \quad (3.5)$$

The connection one-forms Γ_{ab} computed for these tetrads from the first structure equations amount to

$$\left.\begin{matrix} \Gamma_{42} \\ \Gamma_{31} \end{matrix}\right\} = \frac{\mathscr{F}^{1/2}}{r+il}\cdot\frac{1}{1+\epsilon\zeta\bar{\zeta}}\begin{cases} d\zeta, \\ d\bar{\zeta}, \end{cases}$$
$$\Gamma_{12} + \Gamma_{34} = -\frac{\mathscr{F}}{r^2+l^2}\left[\ln\left(\frac{\mathscr{F}^{1/2}}{r+il}\right)\right]^{\cdot}\pi + \epsilon\frac{\bar{\zeta}\, d\zeta - \zeta\, d\bar{\zeta}}{1+\epsilon\zeta\bar{\zeta}}, \quad (3.6)$$

where dots denote the r derivative.

The second Cartan equations can be written as

$$d\Gamma_{42} + \Gamma_{42} \wedge (\Gamma_{12} + \Gamma_{34}) = \gamma e^3 \wedge e^1,$$
$$d\Gamma_{31} + (\Gamma_{12} + \Gamma_{34}) \wedge \Gamma_{31} = \gamma e^4 \wedge e^2, \quad (3.7)$$
$$d(\Gamma_{12} + \Gamma_{34}) + 2\Gamma_{42} \wedge \Gamma_{31} = \beta e^1 \wedge e^2 + \alpha e^3 \wedge e^4,$$

where

$$\alpha := \left[\frac{\mathscr{F}}{r^2+l^2}\left[\ln\left(\frac{\mathscr{F}^{1/2}}{r+il}\right)\right]^{\cdot}\right]^{\cdot},$$

$$\begin{aligned}\beta &:= \frac{1}{r^2+l^2}\left\{2il\frac{\mathscr{F}}{r^2+l^2}\left[\ln\left(\frac{\mathscr{F}^{1/2}}{r+il}\right)\right]^{\cdot}\right.\\ &\quad\left.+\frac{\mathscr{F}}{(r+il)^2}-\epsilon\right\},\end{aligned} \tag{3.8}$$

$$\gamma := -\frac{1}{2}\frac{1}{r-il}\left[\frac{\mathscr{F}}{(r+il)^2}\right]^{\cdot}.$$

We can now read off from (3.7) the corresponding curvature coefficients, understanding by these the scalar curvature R, the five $C^{(a)}$'s, $a = 1,...,5$, the tetrad components of C_{ab}. Observe that because of $C_{ab} = C_{(ab)}$ and the trace condition $C_{12} + C_{34} = 0$, the last object has, in general, nine independent components only. According to (3.7), the only nontrivial curvature coefficients are

$$C^{(3)} = \tfrac{1}{3}(\alpha+\beta+2\gamma), \quad R = 2(\alpha+\beta-4\gamma), \quad C_{12} = \tfrac{1}{2}(\beta-\alpha) = -C_{34}. \tag{3.9}$$

Now, evaluating these nontrivial curvature coefficients in terms of the structural function $\mathscr{F}$, it is convenient to give it in the form of

$$\mathscr{F} = \epsilon(r^2-l^2) - 2mr - \lambda(\tfrac{1}{3}r^4 + 2l^2r^2 - l^4) + \mathscr{E}, \tag{3.10}$$

where λ is the cosmological constant, m is a mass parameter (Schwarzschild constant for $\widetilde{B}^{(+)}$, and magnetic mass for $\widetilde{B}^{(-)}$), and $\mathscr{E} = \mathscr{E}(r)$ is a real analytic function expected to be related to the electromagnetic sources to Einstein equations. A simple—but rather tedious—computation leads to

$$\begin{aligned} C^{(3)} &= \frac{-2}{(r+il)^3}\left[m + il\left(\epsilon - \tfrac{4}{3}\lambda l^2\right)\right]\\ &\quad+\frac{1}{6}\frac{(r+il)^3}{r^2+l^2}\left[\frac{\mathscr{E}}{(r+il)^3}\right]^{\cdot\cdot},\\ R &= -4\lambda + [1/(r^2+l^2)]\ddot{\mathscr{E}},\\ C_{12} &= -\tfrac{1}{4}[(r^2+l^2)^2]^{-1}\\ &\quad\times[(r^2+l^2)\ddot{\mathscr{E}} - 4r\dot{\mathscr{E}} + 4\mathscr{E}].\end{aligned} \tag{3.11}$$

Of course, with only $C^{(3)} \neq 0$ out of all $C^{(a)}$'s, the $\widetilde{B}^{(\pm)}$ metrics are of the type D, and e^3 and e^4 are the principal (double) null directions, which in both cases are geodesic and shear-free ($\Gamma_{424} = 0 = \Gamma_{313}$, $\Gamma_{422} = 0 = \Gamma_{311}$). In the $\widetilde{B}^{(+)}$ case these directions, taken modulo a real proportionality factor, are

$$\sqrt{2}\left(\frac{r^2+l^2}{\mathscr{F}^{(+)}}\right)^{1/2}\begin{cases}e^3\\ e^4\end{cases} = \frac{r^2+l^2}{\mathscr{F}^{(+)}}\,dr \pm \left(d\tau + 2il\frac{\bar{\zeta}\,d\zeta - \zeta\,d\bar{\zeta}}{1+\epsilon\zeta\bar{\zeta}}\right), \tag{3.12}$$

and they have the common complex expansion $Z = 1/(r+il)$, while in the case of $\widetilde{B}^{(-)}$ the double null directions have vanishing complex expansions $\Gamma_{421} = 0 = \Gamma_{312}$, and they are gradients,

$$\frac{1}{\sqrt{2}}\frac{1+\epsilon uv}{\sqrt{r^2+l^2}}\begin{cases}e^3\\ e^4\end{cases} = \begin{cases}du\,,\\ dv\,.\end{cases} \tag{3.13}$$

Let us now start the integration of the field equation with nonlinear electromagnetic sources for the $\widetilde{B}^{(+)}$ metric. We shall omit the symbol (+) in the structural function,

$$\mathscr{f}^{(+)} =: \mathscr{f}, \quad x^{(+)} =: x, \quad \mathscr{M}^{(+)} =: \mathscr{M}, \quad \check{\phi}^{(+)} =: \check{\phi}. \tag{3.14}$$

Comparing C_{12} and R from (3.11) with their expressions from (2.26c), one arrives at

$$\tfrac{1}{4}[(r^2+l^2)^2]^{-1}[(r^2+l^2)\ddot{\mathscr{E}} - 4r\dot{\mathscr{E}} + 4\mathscr{E}] = (DE + \check{B}\check{H}), \tag{3.15a}$$

$$\tfrac{1}{4}[(r^2+l^2)]^{-1}\ddot{\mathscr{E}} = -2\mathscr{M} + (DE + \check{B}\check{H}). \tag{3.15b}$$

These equations now determine easily the r dependence of the x,

$$x = (1/2b^2)(D^2 + \check{B}^2).$$

Indeed, subtracting Eqs. (3.15a) and (3.15b) and multiplying the result by " $-(r^2+l^2)^2$" we have

$$r\dot{\mathscr{E}} - \mathscr{E} = -2(r^2+l^2)^2 b^2 \mathscr{f}, \tag{3.16}$$

which, differentiated with respect to r, yields

$$r\ddot{\mathscr{E}} = -8r(r^2+l^2)b^2\mathscr{f} - 2(r^2+l^2)^2 b^2 \mathscr{f}^{\triangledown}\dot{x}. \tag{3.17}$$

On the other hand, using $DE + \check{B}\check{H} = 2b^2 x \mathscr{f}^{\triangledown}$ substituting for $\ddot{\mathscr{E}}$ from (3.15b) into (3.17), and canceling by $b^2\mathscr{f}^{\triangledown}$, one arrives at

$$\frac{dx}{x} = \frac{-4r\,dr}{r^2+l^2} \rightarrow x = \frac{e^2+\check{g}^2}{2b^2(r^2+l^2)^2}, \tag{3.18}$$

where we have denoted the positive integration constant by $e^2 + \check{g}^2/2b^2$, anticipating that e and $\check{g}$ will play the roles of electric and magnetic monopole charges, respectively.

Equation (3.16) can now be written in the form

$$r^2((1/r)\mathscr{E})^{\cdot} = -(e^2+\check{g}^2)\mathscr{f}/x, \tag{3.19}$$

and thus integrates easily in the form

$$\mathscr{E} = (e^2+\check{g})r\int_r^\infty \frac{dr}{r^2}\frac{\mathscr{f}(s)}{s}, \quad s := \frac{e^2+\check{g}^2}{2b^2(r^2+l^2)^2}. \tag{3.20}$$

Notice that a contribution from the integration constant of (3.19) to $\mathscr{E}$ of the form const$\cdot r$ can be omitted, being considered as incorporated into the constants m or n of the structural function $\mathscr{F}$ from (3.10).

Now, as far as the electromagnetic field is concerned, with x given by (3.18) and using (2.41) and (2.42), remembering (2.40), we infer from the information deduced from the Einstein equations that

$$\begin{aligned} D + i\check{B} &= [(e^2+\check{g}^2)^{1/2}/(r^2+l^2)]e^{i\check{\phi}},\\ E + i\check{H} &= \mathscr{f}^{\triangledown}[(e^2+g^2)^{1/2}/(r^2+l^2)]e^{i\check{\phi}},\end{aligned} \tag{3.21}$$

where $\check{\phi}$ is real.

The fundamental question arises whether with the so determined electromagnetic quantities from the Einstein equations (i.e., deduced from the energy-momentum tensor that leaves them arbitrary modulo the *variable* duality rotations), a sort of generalized Reinich–Wheeler process,[23,24] can now fix $\check{\phi}$ in such a manner that $d\omega = 0$, assuring thus the validity of Maxwell–Faraday equations $f^{\mu\nu}{}_{;\nu} = 0 = p^{\mu\nu}{}_{;\nu}$.

The answer to this question is positive. Indeed, the two-form

$$\omega = Ae^1\wedge e^2 + Be^3\wedge e^4,\quad A=A(r),\quad B=B(r), \tag{3.22}$$

when written explicitly in terms of tetrads (2.15) and (2.16) is easily seen to the closed iff

$$[(r^2+l^2)A]^\cdot - 2ilB = 0 \leftrightarrow d\omega = 0. \tag{3.23}$$

Moreover, if this condition is met, ω can be represented as an *exact* two-form,

$$\omega = -\frac{1}{2il} d\left[(r^2+l^2)A\left(d\tau + 2il\frac{\bar\zeta\, d\zeta - \zeta\, d\bar\zeta}{1+\epsilon\zeta\bar\zeta}\right)\right]. \tag{3.24}$$

On the other hand, by comparing

$$\omega = (D+i\breve B)e^1\wedge e^2 + (E+i\breve H)e^3\wedge e^4$$

with (3.22) and using (3.21) we have

$$(r^2+l^2)A = \sqrt{e^2+\breve g^2}e^{i\breve\phi},$$
$$B = [(\sqrt{e^2+\breve g^2})/(r^2+l^2)]\mathscr{L}^{\triangledown} e^{i\breve\phi}. \tag{3.25}$$

Assuming thus $\breve\phi = \breve\phi(r)$ we can use the closure condition (3.23). Substituting A and B from (3.25) into (3.23), we obtain

$$\breve\phi = [2l/(r^2+l^2)]\mathscr{L}^{\triangledown}, \tag{3.26}$$

which integrates in the form

$$\breve\phi = \breve\phi_0 - 2l\int_r^\infty \frac{dr}{r^2+l^2}\mathscr{L}^{\triangledown}(s),\quad s:=\frac{e^2+\breve g^2}{2b^2(r^2+l^2)^2}, \tag{3.27}$$

where $\breve\phi_0$ is a real constant.

We decide now to consider e and $\breve g$ as the real independent parameters of our solutions incorporating in them the constant phase $\breve\phi_0$,

$$e + i\breve g =: -\sqrt{e^2+\breve g^2}e^{i\breve\phi_0}. \tag{3.28}$$

According to (3.24), our result for the electromagnetic field is then

$$\omega = \frac{(e+i\breve g)}{2il} d\left\{e^{i\breve\psi(+)}\left(d\tau + 2il\frac{\bar\zeta\, d\zeta - \zeta\, d\bar\zeta}{1+\epsilon\zeta\bar\zeta}\right)\right\},$$
$$\breve\psi := -2l\int_r^\infty \frac{dr}{r^2+l^2}\mathscr{L}^{\triangledown}(s),\quad s:=\frac{e^2+\breve g^2}{2b^2(r^2+l^2)^2}. \tag{3.29}$$

We then establish that the nonlinear electromagnetic generalization of the Carter $\widetilde{B}^{(+)}$ solution is given by

$$g = 4(r^2+l^2)\frac{d\zeta\otimes d\bar\zeta}{(1+\epsilon\zeta\bar\zeta)^2} + \frac{r^2+l^2}{\mathscr{F}^{(+)}}dr^2$$
$$-\frac{\mathscr{F}^{(+)}}{r^2+l^2}\left[d\tau + 2il\frac{\bar\zeta\, d\zeta - \zeta\, d\bar\zeta}{1+\epsilon\zeta\bar\zeta}\right]^2,$$
$$\mathscr{F}^{(+)} = \epsilon(r^2-l^2) - 2mr - \lambda$$
$$\times(\tfrac{1}{3}r^4 + 2l^2r^2 - l^4) + \mathscr{E}^{(+)},$$
$$\mathscr{E}^{(+)} = (e^2+\breve g^2)r\int_r^\infty \frac{dr}{r^2}\frac{\mathscr{L}^{(+)}(s)}{s}, \tag{3.30}$$
$$s := \frac{e^2+\breve g^2}{2b^2(r^2+l^2)^2},$$
$$\omega^{(+)} = \frac{(e+i\breve g)}{2il} d$$
$$\times\left[\exp(i\breve\psi^{(+)})\cdot\left(d\tau + 2il\frac{\bar\zeta\, d\zeta - \zeta\, d\bar\zeta}{1+\epsilon\zeta\bar\zeta}\right)\right],$$
$$\breve\psi^{(+)} = -2l\int_r^\infty \frac{dr}{r^2+l^2}\mathscr{L}^{\triangledown(+)}(s),$$

where $\mathscr{L}^{(+)}(s)$ is an arbitrary function such that $\mathscr{L}^{\triangledown(+)}(0) = 1$, $\mathscr{L}^{(+)}(s)/s \geqslant \mathscr{L}^{\triangledown(+)}(s) > 0$. The curvature quantities that characterize this solution can be easily evaluated from Eqs. (2.24).

Following a similar integration process as in the $\widetilde{B}^{(+)}$ case, or applying the formal transformation (3.4), one obtains the nonlinear electromagnetic generalization of the Carter $\widetilde{B}^{(-)}$ solution, which may be given as

$$g = 4(r^2+l^2)\frac{du\otimes dv}{(1+\epsilon uv)^2} + \frac{r^2+l^2}{\mathscr{F}^{(-)}}dr\otimes dr$$
$$+\frac{\mathscr{F}^{(-)}}{r^2+l^2}\left[d\sigma + 2l\frac{v\,du - u\,dv}{1+\epsilon uv}\right]^2,$$
$$\mathscr{F}^{(-)} = \epsilon(r^2-l^2) - 2rm - \lambda(\tfrac{1}{3}r^4 + 2l^2r - l^4) + \mathscr{E}^{(-)},$$
$$\mathscr{E}^{(-)} = -(e^2+\breve g^2)r\int_r^\infty \frac{dr}{r^2}\frac{\mathscr{L}^{(-)}(s)}{s}, \tag{3.31}$$
$$s = \frac{e^2+\breve g^2}{2b^2(r^2+l^2)^2},$$
$$\omega = \frac{e+i\breve g}{2l} d\left[\exp(i\breve\psi^{(-)})\cdot\left(d\sigma + 2l\frac{v\,du - u\,dv}{1+\epsilon uv}\right)\right],$$
$$\breve\psi^{(-)} = -2l\int_r^\infty \frac{dr}{r^2+l^2}\mathscr{L}^{\triangledown(-)}(s),$$

where $\mathscr{L}^{(-)}(s)$ is an arbitrary function fulfilling the conditions $\mathscr{L}^{\triangledown}(0) = 1$, $\mathscr{L}^{(-)}(s)/s \geqslant \mathscr{L}^{\triangledown(-)}(s) > 0$. The curvature quantities are given by (2.24) with $\mathscr{E}$ replaced by $\mathscr{E}^{(-)}$.

We shall end this section giving some corollaries.

Corollary I: The finite symmetries of $\widetilde{B}^{(\pm)}$ metrics amount to the product of (symmetries of spaces of constant curvature $ds^{2(\pm)}$)$\times R$, or, more specifically, to the Lie groups

$$\widetilde{B}^{(+)}:\begin{cases}\epsilon = 1: & SU(2)\times R \quad (2\leftrightarrow 1)O(3)\times R,\\ \epsilon = 0: & E^{(+)}(3)\times R,\\ \epsilon = 1: & SU(1,1)\times R \quad (2\leftrightarrow 1)O(2,1)\times R,\end{cases} \tag{3.32}$$

and

$$\widetilde{B}^{(-)}:\begin{cases}\epsilon = \pm 1: & SL(2,R)\times R(2\leftrightarrow 1)O(2,1)\times R,\\ \epsilon = 0: & E^{(-)}(3)\times R,\end{cases} \tag{3.33}$$

where $E^{(\pm)}(3)$ are, respectively, the groups of symmetries of the Euclidean and pseudo-Euclidean planes, corresponding to signatures $(+,+)$ and $(+,-)$, which consist of two translation and one—correspondingly trigonometric or hyperbolic—rotation. If $l\neq 0$ and $\epsilon^2 = 1$ the orbits of O(3) [resp. O(2,1)] groups are three dimensional. The three-dimensional nature of the orbits of the O(3) group has been recognized in the case of classical Taub[12]–NUT[13] metrics in Ref. 25 (see also Ref. 26).

Notice that the $\widetilde{B}^{(-)}$ metric admits for $\epsilon = 0$, according to the terminology of Ref. 27, two commuting Killing vectors with *null orbits*.

Corollary II: The $\widetilde{B}^{(+)}$ metrics admit a representation *linear* in the structural function $\mathscr{F}^{(+)}$. Indeed, replacing in (3.30) the coordinate τ by τ' according to

$$\tau = \tau' \pm \int^r \frac{r^2 + l^2}{\mathscr{F}^{(+)}} dr, \tag{3.34}$$

we have

$$g = \frac{4(r^2+l^2)}{(1+\epsilon\zeta\bar{\zeta})^2} d\zeta\, d\bar{\zeta} \pm 2\left[d\tau' + 2il\frac{\bar{\zeta}\, d\zeta - \zeta\, d\bar{\zeta}}{1+\epsilon\zeta\bar{\zeta}}\right] dr - \frac{\mathscr{F}^{(+)}}{r^2+l^2}\left[d\tau' + 2il\frac{\bar{\zeta}\, d\zeta - \zeta\, d\bar{\zeta}}{1+\epsilon\zeta\bar{\zeta}}\right]^2. \tag{3.35}$$

In the case of $\widetilde{B}^{(-)}$ metrics a similar trick does not work because $\mathscr{F}^{(-)} > 0$; it would be applicable, though, if the $\widetilde{B}^{(-)}$ structure were considered as *complexified*.

Corollary III: The $\widetilde{B}^{(+)}$ solution in the case $\epsilon = 0$ admits via the coordinate transformation

$$\tau = t - \int_r \frac{r^2+l^2}{\mathscr{F}^{(+)}} dr + \frac{rz\bar{z}}{r^2+l^2}, \qquad \zeta = \frac{\nu}{\sqrt{2}}\frac{z}{r+il} =: \frac{\nu}{\sqrt{2}} Z, \quad \nu^2 = 1, \tag{3.36}$$

an alternative representation in a chart $\{x^\mu\} = \{z,\bar{z},r,t\}$ according to which the metric assumes a Kerr–Schild[28] form,

$$\begin{aligned} g &= 2dz\, d\bar{z} + 2dr\, dt - \frac{\mathscr{F}^{(+)}}{r^2+l^2}(dt + \bar{Z}\, dz + Z\, d\bar{z} - Z\bar{Z}\, dr)^2, \\ \omega &:= (e + i\check{g})(1/2il) d \times [e^{i\check{\psi}^{(+)}}(dt + \bar{Z}\, dz + Z\, d\bar{z} - Z\bar{Z}\, dr)], \end{aligned} \tag{3.37}$$

with $\mathscr{F}^{(+)}$ and $\check{\psi}^{(+)}$ from (3.30).

This special solution, in the case of the original Born–Infeld Lagrangian (2.47), was found long ago by one of us (J.F.P.).[29]

Corollary IV: The class of $\widetilde{B}^{(\pm)}$ solutions[30] in the original Born–Infeld theory are obtained from (3.30) and (3.31) by setting

$$f^{(+)} = \sqrt{1+2s} - 1, \quad s := [(e^2 + \check{g}^2)/2b^2](r^2+l^2)^{-2}, \tag{3.38}$$

$$f^{(-)} = 1 - \sqrt{1-2s}, \quad s := [(e^2 + \check{g}^2)/2b^2](r^2+l^2)^{-2}, \tag{3.39}$$

respectively.

Corollary V: The nonlinear $\widetilde{B}^{(\pm)}$ metrics contain in the limit $b\to\infty$ the well-known results of the Maxwell theory. Indeed, in the limit $b\to\infty$, it follows in particular that

$$\lim_{b\to\infty} \exp i\check{\psi}^{(\pm)} = (r - il)/r + il), \qquad \lim_{b\to\infty} \mathscr{E}^{(\pm)} = \pm(e^2 + \check{g}^2). \tag{3.40}$$

Corollary VI: The $\widetilde{B}^{(\pm)}$ solutions in the limit $l\to 0$ reduce to the nonlinear electromagnetic generalizations of the Reissner–Nørdstrem (RN) class of metrics with $\epsilon = 1,0,-1$. (The proper RN solution arises in the limiting case of $b\to\infty$, $l\to 0$, and $\epsilon = 1$.)

Corollary VII: The Hamilton–Jacobi equation,

$$(r^2 + l^2)[g^{\mu\nu}S_{,\mu}S_{,\nu} + \kappa^2] = 0, \tag{3.41}$$

for timelike or null geodesics ($\kappa^2 \geqslant 0$), are separable and hence their complete integrals can be obtained by quadratures.

IV. THE SPECIAL RELATIVISTIC LIMITS OF GENERAL RELATIVISTIC SOLUTIONS

The special relativistic limits of our $\widetilde{B}^{(\pm)}$ solutions to the dynamical equations of nonlinear electrodynamics plus the associated Einstein equations in the case of the theory endowed with the freedom of the duality rotations can be obtained in two steps. First, abandoning the gravitational units, we restore the gravitational constant G in the structure $\widetilde{B}^{(\pm)}$ metrics by setting $m\to Gm$, and $\mathscr{E}^{(\pm)}\to G\mathscr{E}^{(\pm)}$. Then, switching off the gravitational field, we execute the limit $G\to 0$, $\lambda\to 0$. As a result of this, $G_{\mu\nu}\to 0$, while the conformal curvature coefficient $C^{(3)}$ from (3.11) reduces to

$$\widetilde{B}^{(\pm)}:\ C^{(3)} = -2il\epsilon/(r+il)^3. \tag{4.1}$$

In the second step we require, additionally,

$$\widetilde{B}^{(\pm)}:\ C^{(3)} = 0 \leftrightarrow \epsilon l = 0. \tag{4.2}$$

This assured, both the conformal curvature tensor and the Einstein tensor of $\widetilde{B}^{(\pm)}$ metrics vanish, and hence these metrics become flat, while the limiting ω's determine some solutions to the dynamical equations of the nonlinear electrodynamics in the flat Minkowski space-time.

Within this program, the sub-branches with $\epsilon = 0$ potentially leave l as a free parameter, while for $\epsilon = 1, -1$ we must assume $l\to 0$. Since in the special relativistic limit the signature has to be $(+++-)$, the only viable limiting solutions are the limits of $\widetilde{B}^{(+)}$ ($\epsilon = 0 \neq l$), $\widetilde{B}^{(+)}$ ($\epsilon = \pm 1$, $l = 0$), and $\widetilde{B}^{(-)}$ ($\epsilon = 1$, $l = 0$).

In the case of the branch $\widetilde{B}^{(+)}$ ($\epsilon = 0 \neq l$), we have an alternative description of our general relativistic solution in the Kerr–Schild form (3.37). With the solution in this form, our special relativistic limit leads to a solution in the flat space-time endowed with the free parameter l, and described by the formulas

$$\begin{aligned} g &= 2\, dz\, d\bar{z} + 2dr\, dt, \\ \omega &= [(e + i\check{g})/2il] d[\exp(i\check{\psi}^{(+)}) \times (dt + Z\, dz + Z\, d\bar{z} - Z\bar{Z}\, dr)], \quad Z := z/(r+il), \\ \check{\psi}^{(+)} &= -2l\int_r^\infty \frac{dr}{r^2+l^2} f^{\nabla(+)}(s), \\ s &:= (e^2 + \check{g}^2)/2b^2(r^2+l^2)^2. \end{aligned} \tag{4.3}$$

The coordinates $\{x^\mu\} = \{z,\bar{z},r,t\}$ are the standard Cartesian null coordinates of the Minkowski space-time; there are no restrictions for the ranges of these coordinates, so that the discussed limiting solution can be understood as defined over the Minkowski space-time with the Euclidean topology of R^4.

The special relativistic limits of $\widetilde{B}^{(+)}(\epsilon = \pm 1, l = 0)$ are given by

$$g = \frac{4r^2}{(1+\epsilon\zeta\bar{\zeta})^2} d\zeta\, d\bar{\zeta} + \epsilon(dr^2 - d\tau^2),$$

$$\omega = -(e + i\check{g}) d[V^{(+)}(r) d\tau$$

$$-(\bar\zeta\, d\zeta - \zeta\, d\bar\zeta)/1 + \epsilon\zeta\bar\zeta)], \tag{4.4}$$

$$V^{(+)}(r) = \int_r^\infty \frac{dr}{r^2} \mathscr{P}^{\nabla(+)}(s), \quad s := \frac{e^2 + \check g^2}{2b^2 r^4}.$$

The coordinate transformations, for $\epsilon = 1$,

$$\zeta = e^{i\phi}\cot(\theta/2),$$
$$\tau = t, \quad x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \tag{4.5}$$
$$z = r\cos\theta, \quad r = (x^2+y^2+z^2)^{1/2},$$

and, for $\epsilon = -1$,

$$\zeta = e^{i\phi}\operatorname{cotanh}(\theta/2), \quad \tau = z,$$
$$x = r\sinh\theta\cos\phi, \quad y = r\sinh\theta\sin\phi, \tag{4.6}$$
$$t = r\cosh\theta, \quad r = (t^2 - x^2 - y^2)^{1/2},$$

bring the metric above to the standard Minkowski line element $g = dx^2 + dy^2 + dz^2 - dt^2$, while the electromagnetic two-forms can be rewritten as

$$\omega(\epsilon = 1) = -(e + i\check g)d[V^{(+)}(r)dt - i\cos\theta\, d\phi]$$
$$= -(e + i\check g)d\left[V^{(+)}(\sqrt{x^2+y^2+z^2})dt - \frac{iz}{\sqrt{x^2+y^2+z^2}}\cdot\frac{x\,dy - y\,dz}{x^2+y^2}\right], \tag{4.7}$$

and

$$\omega(\epsilon = -1)$$
$$= -(e + i\check g)d[V^{(+)}(r)dz + i\cosh\theta\, d\phi]$$
$$= -(e + i\check g)d\left[V^{(+)}(\sqrt{t^2 - x^2 - y^2})dz + \frac{it}{\sqrt{t^2 - x^2 - y^2}}\cdot\frac{x\,dy - y\,dx}{x^2+y^2}\right]. \tag{4.8}$$

Understanding the Minkowski space-time as endowed with the Euclidean topology of R^4, we see that the limiting $\widetilde B^{(+)}$ $(\epsilon = 1, l = 0)$ solution covers the whole space-time and should be interpreted as the nonlinear analog of the Coulomb field generalized by the presence of the magnetic monopole charge.

The limiting solution $\widetilde B^{(+)}$ $(\epsilon = -1, l = 0)$ covers only the regions of the space-time where $t^2 - x^2 - y^2 > 0$, becoming singular along the set of points $t^2 - x^2 - y^2 = 0$.

The limiting $\widetilde B^{(-)}$ $(\epsilon = 1, l = 0)$ solution is given by

$$g = [4r^2/(1+uv)^2]du\,dv + dr^2 + d\sigma^2,$$
$$\omega = -(e + i\check g)d\left[iV^{(-)}(r)d\sigma - \frac{v\,du - u\,dv}{1+uv}\right],$$
$$V^{(-)}(r) = \int_r^\infty \frac{dr}{r^2}\mathscr{P}^{\nabla(-)}(s), \quad s := \frac{e^2+\check g^2}{2b^2r^4}. \tag{4.9}$$

The metric above can be brought to the Minskowski line element by accomplishing the transformation

$$u = e^{\phi}\cot(\theta/2), \quad v = e^{-\phi}\cot(\theta/2), \quad \sigma = z,$$
$$x = r\sin\theta\cosh\phi, \quad y = r\cos\theta, \tag{4.10}$$
$$t = r\sin\theta\sinh\phi, \quad r = (x^2+y^2-t^2)^{1/2}.$$

Then, the ω acquires the form

$$\omega = -(e + i\check g)d[iV^{(-)}(r)dz - \cos\theta\, d\phi]$$
$$= -(e + i\check g)d\left[iV^{(-)}(\sqrt{x^2+y^2-t^2}) - \frac{y}{\sqrt{x^2+y^2-t^2}}\cdot\frac{x\,dt - t\,dx}{x^2-t^2}\right]. \tag{4.11}$$

This limiting solution covers the regions of the space-time where $x^2 + y^2 - t^2 > 0$, becoming singular along $x^2 + y^2 - t^2 = 0$.

The explicit solutions to dynamical equations of nonlinear electrodynamics endowed with the freedom of the duality rotations in the Minkowski space-time, which we have obtained in this section via limiting transitions, are of interest for many reasons. Even in the case of flat space-time—as was mentioned in the Introduction—the explicit analytic solutions to the dynamical scheme of nonlinear electrodynamics are scarce. In the form of (4.3), (4.4), and (4.9) we devised some new solutions that allow for the presence of the magnetic monopole charges. A rather unexpected fact has emerged that in the basic limiting solution (4.7), apart from the nonlinearity of the Born–Infeld scheme, the magnetic monopole contribution to ω enters into our solution precisely in the same form as in the case of the linear theory with

$$\omega = \tfrac{1}{2}(f_{\mu\nu} + \check f_{\mu\nu})dx^\mu \wedge dx^\nu$$
$$= -(e + i\check g)d[(1/r)dt - i\cos\theta\, d\phi]. \tag{4.12}$$

Indeed, with $e = 0$, $\frac{1}{2}f_{\mu\nu}dx^\mu \wedge dx^\nu = -\check g d[\cos\theta\, d\phi]$ is a solution in both linear and nonlinear theories.

Our flat space-time solutions, with the physical interpretation within an easy access, are prototypes of the corresponding general relativistic $\widetilde B^{(\pm)}$ solutions; similarly as in the linear theory the Coulomb field of the electric and magnetic monopole charges, (4.12), is a special relativistic prototype of the Reissner–Nørdstrem solution. Thus, our flat space-time solutions are potentially the key to the physical interpretation of our general relativistic $\widetilde B^{(\pm)}$ solutions obtained via rather formal arguments. In this context we should like to observe that it is rather tempting to follow the point of view of Ref. 31, interpreting (4.7) as the field of nonlinear charges, with e and $\check g$ being endowed with a rest frame, (4.3) as the field of these charges being lightlike, and in the presence of l, $l \neq 0$, the complex field being endowed with some angular momentum, and finally attempting to interpret the (4.8) and (4.11) solutions as the tachyonic fields of nonlinear charges. The authors of this paper, especially after the work of Refs. 32 and 33, are, however, rather skeptical about the physical status of tachyons. We prefer thus to consider our general and special relativistic results just as the technical output of the theory of exact solutions, leaving their possible physical interpretation open to further study.

V. THE COMPLETENESS OF THE D-RESULT FOR THEORIES WITH THE FREEDOM OF THE DUALITY ROTATIONS AND FINAL CONCLUSIONS

Our exact solutions to the dynamical equations of nonlinear electrodynamics endowed with the freedom of the duality rotations plus associated Einstein equations were obtained in Sec. III by postulating that the natural tetrads of the D-type Carter $\widetilde{B}^{(\pm)}$ metrics coincide with the eigenvectors of the algebraically general nonlinear electromagnetic field.

Are there *other* D-type metrics that are also carriers of the discussed dynamical structure? Within the present status of the theory of the D-type metrics, especially after the work of Refs. 27 and 34–37, we are able to answer this question in a negative manner: the D-type $\widetilde{B}^{(\pm)}$ metrics—together with their possible contractions—are the *only* D-type metrics compatible with the dynamical scheme of nonlinear electrodynamics endowed with the structural function $\mathscr{M}^{(+)}\check{\mathscr{M}}^{(+)}[\frac{1}{2}(D^2+\check{B}^2)]$. This statement holds assuming that (i) the two principal null directions are geodesic and shear-free, and (ii) they coincide with the real eigenvectors of the algebraically general nonlinear electromagnetic field.

Observe, however, that if we abandon the proviso "geodesic and shear-free" for the principal null directions of D-metrics aligned along the real eigenvectors of the nonlinear electromagnetic field, then there exists an *exceptional* possibility satisfying the Bianchi identities *outside* the validity of the electromagnetic generalization[38] of the Goldberg–Sachs theorem[39] with

$$\text{(a)}\quad [\mathscr{H}_{\mathscr{P}}(D^2+\check{B}^2)-\tfrac{3}{2}C^{(3)}]\cdot\begin{cases}\Gamma_{424}\\ \Gamma_{313}\end{cases}=0,$$

or (5.1)

$$\text{(b)}\quad [\mathscr{H}_{\mathscr{P}}(D^2+\check{H}^2)+\tfrac{3}{2}C^{(3)}]\cdot\begin{cases}\Gamma_{422}\\ \Gamma_{311}\end{cases}=0.$$

The exceptional D-branches with either the factor of (a) or (b) vanishing have been integrated in Refs. 40 and 41 in the case of the linear electrodynamics. A forthcoming paper[42] of Morales and Plebański reexamines these exceptional branches in the nonlinear case.

As always in the case of a negative result, details of its proof can be of interest for very specialized readers only. For this reason, this we will provide only with an outline of the basic ideas of its proof.

According to the results of Refs. 35–37 *all* D-type metrics which allow the choice of a null tetrad such that out of all independent curvature coefficients only $C^{(3)}$, R, and $C_{12}=-C_{34}$ can be $\neq 0$, either coincide with the general metric

$$g=(1-pq)^{-2}\{(\Delta/P)dp^2+(P/\Delta)(d\tau+q^2\,d\sigma)^2+(\Delta/Q)dq^2-(Q/\Delta)(d\tau-p^2\,d\sigma)^2\},\qquad(5.2)$$
$$P=P(p),\quad Q=Q(q),\quad \Delta:=p^2+q^2$$

or amount to contractions of this metric; this statement also covers the case of D-metrics with the two commuting Killing vectors having in particular null orbits.[27]

The metric (5.2) studied with the (linear) electromagnetic sources has led to the seven parametric families of solutions.[43]

Two basic contractions of the (5.2) metric corresponding physically to switching off either the acceleration parameter, or the rotation (Kerr) parameter, reduce the metric, respectively, to Carter A metric[11] (see also Ref. 44) and to the generalized C-metrics discovered by Kinnersley[45]:

$$g=(\Delta/P)dp^2+(P/\Delta)(d\tau+q^2\,d\sigma)^2+(\Delta/Q)dq^2-(Q/\Delta)(d\tau-p^2\,d\sigma)^2,\qquad(5.3)$$

and

$$ds^2=(p+q)^{-2}\left\{\frac{dp^2}{\mathscr{P}}+\mathscr{P}\,d\sigma^2+\frac{dq^2}{Q}-Q\,d\tau^2\right\}.\qquad(5.4)$$

Now, in order to derive the negative result from the introduction to this section, we have proceeded as follows: for all basic D-metrics (5.2)–(5.4) the tetrads and expressions for $C^{(3)}$, R, and C_{12} are well known in the corresponding charts. By examining the equations $R=-4\lambda-8\mathscr{M}^{(+)}+4(DE+\check{B}\check{H})$ and $C_{12}=-(DE+\check{B}\check{H})$ we have found these conditions *contradictory* for the above quoted metrics, except for the limiting case to the linear theory, $b\to\infty$. In proving the corresponding contradictions exceptional care was given to the alternative of the possible solutions with the basic Killing vectors having null orbits.[27]

On the other hand, the work of this paper clearly demonstrates that the $\widetilde{B}^{(\pm)}$ Carter separable D-branches are carriers of solutions to the dynamical scheme of nonlinear electrodynamics endowed with the freedom of the duality rotations.

The moral of these considerations seems to be that the presence of the acceleration parameter and the rotation parameter prohibits the existence of the D-type solutions for the dynamical scheme of the nonlinear electrodynamics endowed with the freedom of the duality rotations. In this situation it is natural to conjecture that the possible solutions to the discussed dynamical sceheme which would include the rotation and acceleration parameters ought to be already of the G-type. We consider a derivation of such solutions, which would generalize the Kerr–Newman solution for the case of the nonlinear rotating charges as an open challenging problem within the theory of exact solutions in general relativity.

This paper has studied the general relativistic type D solutions to the dynamical scheme of nonlinear electrodynamics endowed with the freedom of the duality rotations, which in particular contains the original Born–Infeld scheme. On the other hand, if one understands nonlinear electrodynamics as derived from QED, the Schwinger Lagrangian[5] (see also earlier papers, Refs. 3 and 4) does not exhibit the invariance with respect to the duality rotations. This is perhaps not surprising; the Lagrangian has been deduced from classical QED which distinguishes the charge e as the (small!) coupling constant, and possible modifications of the properties of the polarized vacuum due to the presence of magnetic monopole charges are thus not taken into account. As t'Hooft[46] pointed out, because of large values of $\check{g}$, the second quantization of the magnetic monopole charges is "extremely hard." It still may turn out that a proper quantum theory of monopolic charges would lead to nonlinear electrodynamics endowed with the freedom of the duality rotations, where the phenomenological "dressed" monopole charges have an assured place. Our general relativistic results seem at least to indicate a formal fact: things are relatively simple if the freedom of the duality rotations is postulated, while when one works with a general structural function which does not allow for that freedom, even in the case of the spherical symmetry, it is not clear whether the inclusion of magnetic monopole charges into the nonlinear analog of Reissner–Nørdstrem solutions is possible.

ACKNOWLEDGMENTS

Useful discussions with Dr. G. Torres del Castillo—to whom we owe the (2.30) transformation—and Dr. B. Mielnik are gratefully appreciated.

[1]M. Born, Nature (London) **132**, 282 (1933); Proc. R. Soc. London, Ser. A **143**, 410 (1934).
[2]M. Born and L. Infeld, Nature (London) **132**, 970, 1004 (1933); Proc. R. Soc. London, Ser. A **144**, 425 (1934); **147**, 522 (1934).
[3]E. Euler and B. Kockel, Naturwissenschaften **23**, 246 (1935).
[4]W. Heisenberg and H. Euler, Z. Phys. **98**, 714 (1936).
[5]J. Schwinger, Phys. Rev. **82**, 664 (1951).
[6]L. Infeld and B. Hoffmann, Phys. Rev. **51**, 765 (1937).
[7]R. Pellicer and R. J. Torrence, J. Math. Phys. **10**, 1718 (1969).
[8]L. E. Morales, Nuovo Cimento B **68**, 55 (1982).
[9]I. Robinson, Bull. Acad. Pol. Sci. Ser. Math. Phys. **7**, 352 (1959).
[10]B. Bertotti, Phys. Rev. **116**, 1331 (1959).
[11]B. Carter, Commun. Math. Phys. **10**, 280 (1968).
[12]A. Taub, Ann. Math. **53**, 472 (1951).
[13]E. Newman, L. Tamburino, and T. Unti, J. Math. Phys. **4**, 915 (1963).
[14]J. F. Plebański, "Lectures on non-linear electrodynamics," Monograph of the Niels Bohr Institute Nordita, Copenhagen, 1968.
[15]G. C. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[16]J. Bičak and J. Slavík, Acta Phys. Pol. B **6**, 489 (1975).
[17]E. T. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962); **4**, 998 (1963).
[18]R. Price, Phys. Rev. D **5**, 2419, 2439 (1972).
[19]G. Boillat, Ann. Inst. H. Poincaré A **5**, 217 (1966); *Propagation des Ondes* (Gauthier-Villars, Paris, 1965).
[20]S. Alarcón G., A. L. Dudley, and J. F. Plebański, J. Math. Phys. **22**, 2835 (1981).
[21]J. F. Plebański and A. Schild, Nuovo Cimento B **25**, 35 (1976).
[22]K. Rozga, Rep. Math. Phys. **11**, 197 (1977).
[23]Y. Rainich, Trans. Am. Math. Soc. **27**, 106 (1925).
[24]C. W. Misner and J. A. Wheeler, Ann. Phys. NY **2**, 525 (1962).
[25]C. W. Misner, J. Math. Phys. **4**, 924 (1963).
[26]J. S. Dowker, Gen. Relativ. Gravit. **5**, 603 (1974).
[27]R. Debever and R. G. McLenaghan, J. Math. Phys. **8**, 1711 (1981).
[28]R. P. Kerr and A. Schild, *Atti Del Convegno Sulla Relativita Generale: Problemi Dell'Energia E Onde Gravitazionali* (Anniversary Volume, Fourth Centenary of Galileo's Birth), edited by G. Barbéra (Firenze, 1965), p. 173; *Applications of Non-linear Partial Differential Equations in Mathematical Physics, Proceedings of Symposia in Applied Mathematics* (Americal Mathematical Society, Providence, RI, 1965), Vol. XVII, p. 199.
[29]J. F. Plebański, report at the Wartburg International Relativity Meeting, Wartburg, 1972; G. F. Torres del Castillo and J. F. Plebański, J. Math. Phys. **26**, 477 (1985).
[30]A. García D., H. Salazar I., and J. F. Plebański, Nuovo Cimento B **84**, 65 (1984).
[31]A. Pérez, Phys. Lett A **31**, 361 (1970).
[32]F. A. E. Pirani, Phys. Rev. D **1**, 3324 (1970).
[33]J. K. Kowalcznski, J. Acta Phys. Pol. B **15**, 101 (1984); Int. J. Theor. Phys. **23**, 27 (1984).
[34]R. Debever, N. Kamran, and R. G. McLenaghan, Phys. Lett. A **93**, 399 (1983).
[35]A. García D. and H. Salazar I., Gen. Relativ. Gravit. **15**, 417 (1983); **16**, 417 (1984); J. Math. Phys. **24**, 2498 (1983).
[36]A. García D., J. Math. Phys. **25**, 1951 (1984).
[37]A. García D. and J. F. Plebański, J. Math. Phys. **23**, 1463 (1982).
[38]L. P. Hughston, R. Penrose, P. Sommers, and M. Walker, Commun. Math. Phys. **27**, 303 (1972).
[39]J. N. Goldberg and R. K. Sachs, Acta Phys. Pol. **22**, Suppl. 13 (1962).
[40]J. F. Plebański and S. Hacyan J. Math. Phys. **20**, 1004 (1979).
[41]A. García D. and J. F. Plebański, J. Math. Phys. **23**, 123 (1982).
[42]L. E. Morales and J. F. Plebański, Nuovo Cimento B **92**, 61 (1986).
[43]J. F. Plebański and M. Demiański, Ann. Phys. (NY) **98**, 98 (1976).
[44]J. F. Plebański, Ann. Phys. (NY) **90**, 196 (1975).
[45]W. Kinnersley and M. Walker, Phys. Rev. D **2**, 1359 (1970).
[46]G. t'Hooft, Phys. Scr. **25**, 133 (1982).

Kinematics and dynamics of conformal collineations in relativity

D. P. Mason and R. Maartens
Centre for Nonlinear Studies and Department of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg 2050, South Africa

(Received 30 December 1986; accepted for publication 6 May 1987)

Anisotropic fluids in general relativity that admit a conformal collineation, a generalization of a conformal motion, are considered. By investigating the kinematic properties of such fluids, and then using the field equations, some recent results on the restrictions imposed by a conformal collineation symmetry are generalized.

I. INTRODUCTION

Recently Duggal and Sharma[1] investigated the dynamic restrictions imposed by a special conformal collineation in a class of anisotropic relativistic fluids. In this paper, using the methods of a previous paper,[2] we generalize the results of Duggal and Sharma. We show that there is an implicit kinematic assumption in their paper, which may be dropped, allowing for a wider range of possibilities. Also, we extend their results to the general class of anisotropic fluids without energy flux.

A conformal collineation[1] is generated by an affine conformal vector field (ACV) ξ^a, characterized by its effect on the metric tensor,[3]

$$\mathscr{L}_\xi g_{ab} = 2\psi g_{ab} + H_{ab}, \quad H_{[ab]} = 0 = H_{ab;c}, \tag{1}$$

where ψ is the scalar and H_{ab} the symmetric, parallel (and therefore Killing) tensor associated with ξ^a. An ACV is a generalization of a conformal Killing vector field (CKV), to which it reduces iff $H_{ab} = Ag_{ab}$, $A = \text{const}$. A special conformal collineation, generated by a special ACV, is characterized by (1) together with

$$\psi_{;ab} = 0, \tag{2}$$

and is therefore a generalization of a special conformal motion. It is easily shown that an ACV is special iff it leaves invariant the curvature tensor $R^a{}_{bcd}$. A special ACV is therefore a particular case of a Ricci collineation vector field,

$$\mathscr{L}_\xi R_{ab} = 0. \tag{3}$$

Following Herrera *et al.*,[4] Duggal and Sharma[1] consider fluids without energy flux, in which there is a preferred direction of pressure anisotropy. Herrera *et al.* used the field equations to investigate the dynamic consequences when such fluids admit a special CKV ξ^a. Duggal and Sharma extend this work to the more general case of a special ACV ξ^a, and show in particular that the stiff equation of state ($p = \mu$) is no longer singled out when ξ^a is orthogonal to the fluid four-velocity u^a. In both of these papers, it is assumed that ξ^a maps fluid flow lines into fluid flow lines, i.e., that $\mathscr{L}_\xi u^a$ is parallel to u^a. That this is not in general the case for CKV is shown by a counterexample given by Maartens *et al.*[2] This counterexample applies also to ACV, of which CKV is a particular case.

In Sec. II we show that for an ACV,

$$\mathscr{L}_\xi u^a = -(\psi - \tfrac{1}{2}H_{bc}u^bu^c)u^a + v^a,$$

where v^a is orthogonal to u^a and involves the vorticity and acceleration of the fluid. We display a proper ACV with $v^a \neq 0$ in an Einstein static fluid space-time. We also give a detailed characterization of the kinematic restrictions imposed by an ACV. In Sec. III we use the field equations to investigate the dynamic restrictions imposed by a special ACV. In particular, we show that the implicit assumption of Duggal and Sharma[1] that $v^a = 0$ amounts to assuming that u^a is an eigenvector of H_{ab}. Although Herrera *et al.*[4] restricted attention to the case where a special CKV maps fluid flow lines into fluid flow lines ($v^a = 0$), that assumption remarkably turns out not to be a further restriction since u^a is trivially an eigenvector of H_{ab} for the case of a special CKV. Furthermore, u^a is also an eigenvector of H_{ab} for the special case $H_{ab} = \gamma R_{ab}$ considered by Duggal and Sharma,[1] so that their results for this special case are not affected by their assumption that $v^a = 0$. In Sec. IV we extend the dynamic results for a special ACV to fluids with arbitrary pressure anisotropy.

II. KINEMATICS

The effect of an ACV on any non-null unit vector X^a is given by

$$\begin{aligned}\mathscr{L}_\xi X^a &= -(\psi + (\epsilon/2)H_{bc}X^bX^c)X^a + Y^a,\\ \mathscr{L}_\xi X_a &= (\psi - (\epsilon/2)H_{bc}X^bX^c)X_a + H_{ab}X^b + Y_a,\end{aligned} \tag{4}$$

where Y^a is some vector orthogonal to X^a, $\epsilon = +1$ if X^a is spacelike and $\epsilon = -1$ if X^a is timelike. The proof of Eqs. (4) is a generalization of that for the case of a CKV (Ref. 2). We decompose $\mathscr{L}_\xi X^a$ as $\mathscr{L}_\xi X^a = \alpha X^a + Y^a$, for some α and Y^a, where $Y^aX_a = 0$. Contracting with X_a, and using $\mathscr{L}_\xi(X_aX^a) = 0$ and $\mathscr{L}_\xi X_a = \mathscr{L}_\xi(g_{ab}X^b)$, we obtain

$$\begin{aligned}\alpha &= -\epsilon X^a\mathscr{L}_\xi X_a\\ &= -\epsilon X^a(2\psi X_a + H_{ab}X^b + g_{ab}\mathscr{L}_\xi X^b),\end{aligned}$$

and Eqs. (4) follow. In general $Y^a \neq 0$: an explicit example of a CKV ($H_{ab} = 0$) with $Y^a \neq 0$ in Robertson–Walker space-time is given in Ref. 2.

Now consider a fluid four-velocity u^a ($u^au_a = -1$) and a unit vector n^a orthogonal to u^a ($n^an_a = 1$, $n^au_a = 0$). From (4) we get

$$\begin{aligned}\mathscr{L}_\xi u^a &= -(\psi - \tfrac{1}{2}H_{bc}u^bu^c)u^a + v^a,\\ \mathscr{L}_\xi u_a &= (\psi + \tfrac{1}{2}H_{bc}u^bu^c)u_a + H_{ab}u^b + v_a,\end{aligned} \tag{5}$$

 0022-2488/87/092182-05$02.50

where $v^a u_a = 0$, and

$$\mathcal{L}_\xi n^a = -(\psi + \tfrac{1}{2}H_{bc}n^b n^c)n^a + m^a,$$
$$\mathcal{L}_\xi n_a = (\psi - \tfrac{1}{2}H_{bc}n^b n^c)n_a + H_{ab}n^b + m_a, \tag{6}$$

where $m^a n_a = 0$. Then (5), (6), and $\mathcal{L}_\xi(n_a u^a) = 0$ give the condition

$$m_a u^a + v_a n^a = -H_{ab}u^a n^b. \tag{7}$$

In Sec. III we will identify n^a with the preferred direction of anisotropy in a particular class of anisotropic fluids. Duggal and Sharma[1] give (5) and (6) with $v^a = 0 = m^a$. By (7), this implies that $H_{ab}u^a n^b = 0$. This kinematic condition will hold if u^a or n^a are eigenvectors of H_{ab}, but in general it will not hold.

We can display an example of a fluid space-time with a proper ACV which does not map fluid flow lines into fluid flow lines. Consider the Einstein static fluid space-time,

$$ds^2 = -dt^2 + (1-r^2)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$$

with $u^a = \delta^a_0$. This space-time admits[5] a proper CKV ($H_{ab} = 0$, $\psi_{;ab} \neq 0$),

$$\xi^a_1 = (1-r^2)^{1/2}\cos t\delta^a_0 - r(1-r^2)^{1/2}\sin t\delta^a_1,$$

and it admits[6] a proper affine collineation vector ($\psi = 0$, $H_{ab} \neq 0$),

$$\xi^a_2 = t\delta^a_0.$$

If we take the combination $\xi^a = \xi^a_1 + \xi^a_2$, we obtain a proper ACV ($\psi_{;ab} \neq 0$ and $H_{ab} \neq 0$),

$$\xi^a = (t + (1-r^2)^{1/2}\cos t)\delta^a_0 - r(1-r^2)^{1/2}\sin t\delta^a_1,$$
$$\psi = -(1-r^2)^{1/2}\sin t,\quad H_{ab} = -2t_{,a}t_{,b}. \tag{8}$$

Now it is clear that ξ^a does not map the fluid flow conformally, since $\mathcal{L}_\xi u^a$ is not parallel to u^a. Hence (8) is an ACV which is not a CKV nor an affine collineation vector, and for which $v^a \neq 0$ in (5).

Now ψ and v^a may be related to kinematic quantities of the fluid by following the approach of Ref. 2. We decompose ξ^a as $\xi^a = \alpha u^a + \beta^a$, where $\alpha = -u_a\xi^a$ and $\beta^a u_a = 0$. Then, using

$$u_{a;b} = \sigma_{ab} + \tfrac{1}{3}\theta h_{ab} + \omega_{ab} - \dot{u}_a u_b,$$

we find

$$\mathcal{L}_\xi u_a = \dot{\alpha}u_a + \alpha[\dot{u}_a + (\log\alpha^{-1})_{;b}h^b{}_a] + \beta^b\dot{u}_b u_a + 2\omega_{ab}\beta^b,$$

which, together with (5), gives upon contraction with u^a and h^{ac} ($= g^{ac} + u^a u^c$),

$$\psi = \dot{\alpha} + \dot{u}_a\xi^a + \tfrac{1}{2}H_{ab}u^a u^b, \tag{9}$$

$$v_a = 2\omega_{ab}\xi^b + \alpha[\dot{u}_a + (\log\alpha^{-1})_{;b}h^b{}_a] - H_{bc}u^b h^c{}_a. \tag{10}$$

There are two special cases of interest.

(a) $\xi^a u_a = 0$: Then $\alpha = 0$, and (9) and (10) reduce to

$$\psi = \dot{u}_a\xi^a + \tfrac{1}{2}H_{ab}u^a u^b, \tag{11}$$

$$v_a = 2\omega_{ab}\xi^b - H_{bc}u^b h^c{}_a. \tag{12}$$

By (12), it follows that if ξ^a is parallel to the vorticity vector ω^a ($\omega_{ab}\omega^b = 0$), or if the vorticity is zero, then ξ^a maps flow lines into flow lines (i.e., $v^a = 0$) iff u^a is an eigenvector of H_{ab}. On the other hand, if u^a is an eigenvector of H_{ab}, then ξ^a maps flow lines into flow lines iff ξ^a is parallel to ω^a or $\omega_{ab} = 0$.

(b) $\xi^a = \xi u^a$: Then $\alpha = \xi$ and $v^a = 0$ (since $\mathcal{L}_\xi u^a$ is clearly parallel to u^a), so that (9) and (10) give

$$\psi = \dot{\xi} + \tfrac{1}{2}H_{ab}u^a u^b, \tag{13}$$

$$\dot{u}_a = -[(\log\xi^{-1})_{;b} - \xi^{-1}H_{bc}u^c]h^b{}_a. \tag{14}$$

By (14), it follows that if u^a is an eigenvector of H_{ab}, then ξ^{-1} is an acceleration potential. The shear σ_{ab} and the expansion θ of the fluid may be obtained by contracting (1) with $h^a{}_c h^b{}_d - \tfrac{1}{3}h^{ab}h_{cd}$ and h^{ab},

$$\sigma_{cd} = (2\xi)^{-1}(h^a{}_c h^b{}_d - \tfrac{1}{3}h^{ab}h_{cd})H_{ab}, \tag{15}$$

$$\theta = 3\xi^{-1}\psi + (2\xi)^{-1}h^{ab}H_{ab}. \tag{16}$$

Equation (15) shows the close relation between H_{ab} and the fluid shear.

III. DYNAMICS

We now consider how the field equations alter the purely kinematic results of Sec. II for the particular case of a *special* ACV. Using (1) and (3), the Lie derivative of the field equations

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = T_{ab} \tag{17}$$

gives an expression for the Lie derivative of an arbitrary energy-momentum tensor,

$$\mathcal{L}_\xi T_{ab} = [2\Lambda\psi + \tfrac{1}{2}((\Lambda - \tfrac{1}{2}T)g^{cd} + T^{cd})H_{cd}]g_{ab} + (\tfrac{1}{2}T - \Lambda)H_{ab}. \tag{18}$$

In this section we consider a fluid with a preferred direction of pressure anisotropy and no energy flux, so that T_{ab} takes the form[1,2]

$$T_{ab} = \mu u_a u_b + p_\parallel n_a n_b + p_\perp p_{ab}, \tag{19}$$

where μ is the total energy density, n^a is the unit vector along the dynamically preferred direction, $p_{ab} = h_{ab} - n_a n_b$ is the projection tensor into the local two-planes of pressure isotropy ($p_{ab}u^b = 0 = p_{ab}n^b$), and $p_\parallel$ and $p_\perp$ are the pressure along and orthogonal to n^a, respectively.[3] When $p_\parallel = p_\perp$, (19) reduces to the energy-momentum tensor for a perfect fluid.

By (1), (5)–(7), and (19), (18) takes the form

$$\begin{aligned}
&[\mathcal{L}_\xi\mu + 2\psi\mu + (\mu + p_\perp)H_{rs}u^r u^s]u_a u_b + [\mathcal{L}_\xi p_\perp + 2\psi p_\perp]p_{ab} + [\mathcal{L}_\xi p_\parallel + 2\psi p_\parallel - (p_\parallel - p_\perp)H_{rs}n^r n^s]n_a n_b\\
&\quad + 2(\mu + p_\perp)u_{(a}v_{b)} + 2(p_\parallel - p_\perp)n_{(a}m_{b)} + p_\perp H_{ab} + 2(\mu + p_\perp)u_{(a}H_{b)t}u^t + 2(p_\parallel - p_\perp)n_{(a}H_{b)t}n^t\\
&\quad = [2\Lambda\psi + \tfrac{1}{4}(\mu - p_\parallel + 2\Lambda)H_{rs}p^{rs} + \tfrac{1}{4}(\mu + 2p_\perp + p_\parallel - 2\Lambda)H_{rs}u^r u^s\\
&\qquad + \tfrac{1}{4}(\mu - 2p_\perp + p_\parallel + 2\Lambda)H_{rs}n^r n^s](p_{ab} - u_a u_b + n_a n_b) - \tfrac{1}{2}(\mu - 2p_\perp - p_\parallel + 2\Lambda)H_{ab}.
\end{aligned} \tag{20}$$

By contracting (20) in turn with the tensors $u^a u^b$, $u^a n^b$, $u^a p^{bc}$, $n^a n^b$ [using (7)], $n^a p^{bc}$, p^{ab}, and $p^{ac}p^{bd} - \frac{1}{2}p^{ab}p^{cd}$, we obtain

$$\mathscr{L}_\xi \mu + 2\psi(\mu + \Lambda) = -\tfrac{1}{4}(\mu - p_\parallel + 2\Lambda)H_{ab}p^{ab} + \tfrac{1}{4}(\mu + 2p_\perp + p_\parallel - 2\Lambda)H_{ab}u^a u^b - \tfrac{1}{4}(\mu - 2p_\perp + p_\parallel + 2\Lambda)H_{ab}n^a n^b, \quad (21)$$

$$\mathscr{L}_\xi p_\parallel + 2\psi(p_\parallel - \Lambda) = \tfrac{1}{4}(\mu - p_\parallel + 2\Lambda)H_{ab}p^{ab} + \tfrac{1}{4}(\mu + 2p_\perp + p_\parallel - 2\Lambda)H_{ab}u^a u^b - \tfrac{1}{4}(\mu - 2p_\perp + p_\parallel + 2\Lambda)H_{ab}n^a n^b, \quad (22)$$

$$\mathscr{L}_\xi p_\perp + 2\psi(p_\perp - \Lambda) = \tfrac{1}{4}(\mu + 2p_\perp + p_\parallel - 2\Lambda)H_{ab}u^a u^b + \tfrac{1}{4}(\mu - 2p_\perp + p_\parallel + 2\Lambda)H_{ab}n^a n^b, \quad (23)$$

$$(\mu + p_\parallel)n^a v_a = -\tfrac{1}{2}(\mu + 2p_\perp + p_\parallel - 2\Lambda)H_{ab}u^a n^b, \quad (24)$$

$$(\mu + p_\perp)p^{ab}v_b = -\tfrac{1}{2}(\mu + 2p_\perp + p_\parallel - 2\Lambda)p^{ab}H_{bc}u^c, \quad (25)$$

$$(p_\parallel - p_\perp)p^{ab}m_b = -\tfrac{1}{2}(\mu - 2p_\perp + p_\parallel + 2\Lambda)p^{ab}H_{bc}n^c, \quad (26)$$

$$(\mu - p_\parallel + 2\Lambda)(p_a{}^c p_b{}^d - \tfrac{1}{2}p_{ab}p^{cd})H_{cd} = 0. \quad (27)$$

Equations (21)–(26) generalize Eqs. (12)–(17) of Ref. 1, which were derived under the implicit assumption that $v^a = 0 = m^a$ (and which have zero cosmological constant Λ). [Note that Eq. (15) of Ref. 1 is identically true by virtue of the kinematic condition (7), which implies $H_{ab}u^a n^b = 0$; this invalidates observation (ii) and part of observation (i) at the end of Sec. II in Ref. 1.]

The following results are readily derived from (24)–(26) and (7):

$$\mu + p_\parallel \neq 0 \neq \mu + p_\perp \text{ and } u^a \text{ an eigenvector of } H_{ab} \Rightarrow v^a = 0; \quad (28)$$

$$v^a = 0 \Rightarrow \mu + 2p_\perp + p_\parallel - 2\Lambda = 0 \text{ or } u^a \text{ an eigenvector of } H_{ab}; \quad (29)$$

$$\mu + p_\parallel \neq 0 \neq p_\parallel - p_\perp \text{ and } n^a \text{ an eigenvector of } H_{ab} \Rightarrow m^a = 0; \quad (30)$$

$$m^a = 0 \Rightarrow \mu - 2p_\perp + p_\parallel + 2\Lambda = 0 \text{ or } n^a \text{ an eigenvector of } H_{ab}. \quad (31)$$

For example, if $m^a = 0$ then (7) implies $v_a n^a = -H_{ab}u^a n^b$, which is substituted in (24) to obtain

$$(\mu - 2p_\perp + p_\parallel + 2\Lambda)H_{ab}u^a n^b = 0,$$

while (26) implies

$$(\mu - 2p_\perp + p_\parallel + 2\Lambda)p^{ab}H_{bc}n^c = 0,$$

so that either $\mu - 2p_\perp + p_\parallel + 2\Lambda = 0$, or $(H_{ab}n^b)u^a = 0 = (H_{ab}n^b)p^{ac}$. This proves (31), and (28)–(30) are proved similarly.

Thus, apart from special cases, the vanishing of v^a (m^a) is equivalent to u^a (n^a) being an eigenvector of H_{ab}. This is the dynamic characterization of the kinematic vectors v^a and m^a, which the field equations impose in the case of a special ACV. There are two important cases when u^a and n^a are eigenvectors of H_{ab}, so that v^a and m^a are forced to vanish by (28) and (30) (provided the physically reasonable energy conditions $\mu + p_\parallel \neq 0 \neq \mu + p_\perp$ are satisfied). First, if the special ACV reduces to a special CKV, then $H_{ab} = Ag_{ab}$ (we can set $A = 0$ by absorbing it into ψ). In this case, the results of Ref. 2 are regained. Second, if $H_{ab} = \gamma R_{ab}$ (so that space-time is Ricci recurrent, by virtue of $H_{ab;c} = 0$), then the field equations (17), with (19), show that u^a and n^a are eigenvectors of H_{ab}. This is the case chosen by Duggal and Sharma.[1] Thus the main results of their paper (on the equation of state) are unaffected by their implicit assumption that $v^a = 0 = m^a$.

Note that the results (28)–(31) hold only for ξ^a a *special* ACV. The ACV (8) in Einstein static space-time is not special ($\psi_{;ab} \neq 0$), and it does not satisfy (28): for this ACV, u^a is an eigenvector of H_{ab} by (8) (since $u_a = -\delta_a^0 = -t_{,a}$), and $p_\parallel = p_\perp = 0$, $\mu \neq 0$ (so that $\mu + p_\parallel \neq 0 \neq \mu + p_\perp$), but $v^a \neq 0$.

Consider now how the field equations restrict the kinematic quantities of the fluid in the two cases considered in Sec. II.

(a) $\xi^a u_a = 0$: If $\mu + p_\parallel \neq 0 \neq \mu + p_\perp$ and u^a is an eigenvector of H_{ab}, then (28) and (12) show that either the vorticity vector ω^a is parallel to the special ACV ξ^a, or the vorticity vanishes,

$$\omega^a = \alpha \xi^a \quad \text{or} \quad \omega_{ab} = 0. \quad (32)$$

(b) $\xi^a = \xi u^a$: Then $v^a = 0$, and (28) and (14) imply that

$$\mu + 2p_\perp + p_\parallel - 2\Lambda = 0 \quad \text{or} \quad \dot{u}_a = -(\log \xi^{-1})_{;b}h^b{}_a. \quad (33)$$

In the case of a perfect fluid ($p_\parallel = p_\perp = p$ and n^a an arbitrary unit vector orthogonal to u^a), contraction of (20) with $h^a{}_c h^b{}_d - \frac{1}{3}h^{ab}h_{cd}$ gives (since $v_a = 0$)

$$\tfrac{1}{2}(\mu - p + 2\Lambda)(h^a{}_c h^b{}_d - \tfrac{1}{3}h^{ab}h_{cd})H_{ab} = 0,$$

so that by (15) we get

$$\mu - p + 2\Lambda = 0 \quad \text{or} \quad \sigma_{ab} = 0. \quad (34)$$

The results (33) and (34) were obtained for a Ricci collineation (of which a special conformal collineation is a particular case) by Oliver and Davis.[7] If $p_\parallel \neq p_\perp$, then (27) and (15) give the generalization of (34),

$$\mu - p_\parallel + 2\Lambda = 0 \quad \text{or} \quad p_a{}^c p_b{}^d \sigma_{cd} = -(\tfrac{1}{2}\sigma_{cd}n^c n^d)p_{ab}.$$

Thus if $\mu - p_\parallel + 2\Lambda \neq 0$, the projection of the shear tensor into the local two-planes of pressure isotropy is isotropic, i.e., the preferred direction of pressure anisotropy is also a preferred shear direction.

Finally, we consider the restrictions on the equations of state, imposed by the field equations for an anisotropic fluid (19) which admits a special conformal collineation. Up to now, we have not used Eqs. (21)–(23), which are the time-like components of the Lie derived field equations (20). These equations, together with the contracted Bianchi identities, show that[1] when $\xi^a = \xi u^a$ or $\xi^a = \xi n^a$ or $\xi^a u_a = 0 = \xi^a n_a$, either the equation of state is fixed or there is a condition on ψ and H_{ab}. Since (21)–(23) do not depend on v^a and m^a, the results of Duggal and Sharma[1] are unaffected by their assumption that $v^a = 0 = m^a$. These results, with cosmological constant included, are

$$\xi^a = \xi u^a \Rightarrow (\mu + 2p_\perp + p_\parallel - 2\Lambda) \times (4\psi + H^a{}_a + 2H_{ab}u^a u^b) = 0, \tag{35}$$

$$\xi^a = \xi n^a \Rightarrow (\mu - 2p_\perp + p_\parallel + 2\Lambda) \times (4\psi + H^a{}_a - 2H_{ab}n^a n^b) = 0, \tag{36}$$

$$\xi^a u_a = 0 = \xi^a n_a \Rightarrow (\mu - p_\parallel + 2\Lambda) \times (4\psi - H_{ab}u^a u^b + H_{ab}n^a n^b) = 0. \tag{37}$$

[Equation (23) of Ref. 1 appears to contain a minor error by the inclusion of the term $H = H^a{}_a$.] By (13) and (16), we can rewrite the result (35) for $\xi^a = \xi u^a$,

$$\mu + 2p_\perp + p_\parallel - 2\Lambda = 0 \quad \text{or} \quad \theta = \tfrac{1}{2}(\log \xi)^{\cdot}. \tag{38}$$

IV. GENERAL ANISOTROPY

For a fluid without energy flux and without a preferred direction of anisotropy, the energy-momentum tensor is

$$T_{ab} = \mu u_a u_b + p h_{ab} + \pi_{ab}, \tag{39}$$

where p is the isotropic pressure and π_{ab} is the anisotropic pressure tensor ($\pi^a{}_a = 0 = \pi_{ab}u^b$). The energy-momentum tensor (39) reduces to (19) when

$$p = \tfrac{1}{3}(p_\parallel + 2p_\perp), \quad \pi_{ab} = (p_\perp - p_\parallel)(\tfrac{1}{3}h_{ab} - n_a n_b).$$

Using (1), (5), and (39) in (18), we obtain the generalized form of (20), and when we contract with $u^a u^b$, h^{ab}, $u^a h^b{}_c$, and $h^a{}_c h^b{}_d - \frac{1}{3}h^{ab}h_{cd}$, we obtain

$$\mathscr{L}_\xi \mu + 2\psi(\mu + \Lambda) = \tfrac{1}{4}(\mu + 3p - 2\Lambda)H_{rs}u^r u^s - \tfrac{1}{4}(\mu - p + 2\Lambda)H_{rs}h^{rs} - \tfrac{1}{2}H_{rs}\pi^{rs}, \tag{40}$$

$$\mathscr{L}_\xi p + 2\psi(p - \Lambda) = \tfrac{1}{4}(\mu + 3p - 2\Lambda)H_{rs}u^r u^s + \tfrac{1}{12}(\mu - p + 2\Lambda)H_{rs}h^{rs} + \tfrac{1}{6}H_{rs}\pi^{rs}, \tag{41}$$

$$(\mu + p)v_a = -\pi_{ab}v^b - \tfrac{1}{2}(\mu + 3p - 2\Lambda)H_{bc}h^b{}_a u^c, \tag{42}$$

$$\mathscr{L}_\xi \pi_{ab} = 2u_{(a}\pi_{b)c}v^c + \tfrac{1}{3}(H_{cd}\pi^{cd})h_{ab} - \tfrac{1}{2}(\mu - p + 2\Lambda)(h_a{}^c h_b{}^d - \tfrac{1}{3}h_{ab}h^{cd})H_{cd}. \tag{43}$$

Now it follows from (42) and (39) that the generalizations of (28) and (29) are

$$\mu + p \neq 0 \quad \text{and} \quad u^a \text{ an eigenvector of } H_{ab} \Rightarrow T_{ab}v^b = -\mu v_a, \tag{44}$$

$$v^a = 0 \Rightarrow \mu + 3p - 2\Lambda = 0 \quad \text{or} \quad u^a \text{ an eigenvector of } H_{ab}. \tag{45}$$

If we make the physically reasonable assumptions that $\mu > 0$ and that spacelike eigenvalues of the energy-momentum tensor are positive (see Ref. 2 for a discussion), then (44) becomes

$$\mu + p \neq 0 \quad \text{and } u^a \text{ an eigenvector of } H_{ab} \Rightarrow v^a = 0.$$

Thus the dynamic results on v^a of Sec. III are carried over for a general anisotropic pressure tensor, i.e., that apart from special cases, the special ACV ξ^a maps fluid flow lines into flow lines ($v^a = 0$) iff u^a is an eigenvector of the affine conformal tensor H_{ab}. The condition that u^a be an eigenvector of H_{ab} is equivalent to the condition that u^a be an eigenvector of $\mathscr{L}_\xi R^a{}_b$ ($R_{ab} \neq 0$), since by (1) and (3),

$$\mathscr{L}_\xi R^a{}_b = -2\psi R^a{}_b - H^a{}_c R^c{}_b,$$

and since u^a is an eigenvector of $R^a{}_b$ for arbitrary π_{ab} (provided the energy flux vanishes), by (17) and (39). Thus u^a is an eigenvector of H_{ab} iff u^a remains an eigenvector of $R^a{}_b$ as $R^a{}_b$ is deformed under ξ^a. We see that the underlying reasons for the dynamic characterization of $v^a = 0$ are the invariance of R_{ab} under ξ^a [Eq. (3)], and the vanishing of the energy flux.

The restrictions on the kinematic quantities given by (32) and (33) also hold for arbitrary π_{ab} [with $2p_\perp + p_\parallel = 3p$ in (33)], as is readily seen by the same arguments used to derive (32) and (33).

Finally, we consider the extension of the results in Sec. III on equations of state to the case of arbitrary π_{ab}, using (40) and (41). A straightforward generalization of the arguments of Ref. 1 (or of Ref. 2, Sec. VI) shows that for $\xi^a = \xi u^a$, (38) is carried over,

$$\mu + 3p - 2\Lambda = 0 \quad \text{or} \quad \theta = \tfrac{1}{2}(\log \xi)^{\cdot}.$$

For $\xi^a u_a = 0$, we obtain

$$\mathscr{L}_\xi(\mu - p) + (\mu - p + 2\Lambda)\xi^a{}_{;a} + 2\pi^{ab}{}_{;a}\xi_b + \pi^{ab}H_{ab} = 0,$$

which, using (40), (41), (1), and (11), reduces to

$$(\mu - p + 2\Lambda)(8\psi + H^a{}_a + 4\dot{u}_a\xi^a) + 2\pi^{ab}H_{ab} + 12\pi^{ab}{}_{;a}\xi_b = 0.$$

Thus no simple generalization of (36) is obtained, and (36) is dependent on the special form of π_{ab} for a preferred direction of anisotropy.

V. CONCLUDING REMARK

The theoretical investigation of conformal motions, initiated by Herrera *et al*,[4] laid the basis for a number of physically applicable special solutions admitting conformal motions.[8,9] Hopefully, the theoretical investigation of the more general conformal collineations, initiated by Duggal and Sharma,[1] will also give rise to physical applications in relativistic fluids.

ACKNOWLEDGMENTS

We thank the referee for valuable comments.

D.P.M. thanks the Foundation for Research Development of the Council for Scientific and Industrial Research, Pretoria, for an Operating Grant.

[1]K. L. Duggal and R. Sharma, J. Math. Phys. **27**, 2511 (1986).

[2]R. Maartens, D. P. Mason, and M. Tsamparlis, J. Math. Phys. **27**, 2987 (1986).

[3]We follow standard geometric conventions and notation, as used in Refs. 1 and 2. Our choice of symbols differs somewhat from that of Ref. 1: our ψ, H_{ab}, n^a, $p_\parallel$, $p_\perp$ correspond to their σ, h_{ab}, S^a, p, $\bar{p}$, respectively.

[4]L. Herrera, J. Jimenez, L. Leal, J. Ponce de Leon, M. Esculpi, and V. Galino, J. Math. Phys. **25**, 3274 (1984).

[5]R. Maartens and S. D. Maharaj, Class. Quant. Gravit. **3**, 1005 (1986).

[6]R. Maartens, J. Math. Phys. **28**, 2051 (1987).

[7]D. R. Oliver and W. R. Davis, Gen. Relativ. Gravit. **8**, 905 (1977); Ann. Inst. H. Poincaré **30**, 339 (1979).

[8]L. Herrera and J. Ponce de Leon, J. Math. Phys. **26**, 778, 2018, 2847 (1985).

[9]M. Esculpi and L. Herrera, J. Math. Phys. **27**, 2087 (1986).

Algebra of operators for a system of composite and elementary particles

A. Zinoun and J. Cortois
Laboratoire de Physique Théorique, Université des Sciences et Techniques de Lille I, 59655 Villeneuve d'Ascq Cédex, France

(Received 7 October 1986; accepted for publication 25 March 1987)

In this paper, a system made up of two different fermion species with two algebraical structures, a bigraded-Lie structure and a Lie superalgebra structure, are shown. The Lie superalgebra structure gives the normal commutation relations, while the bigraded Lie structure gives the abnormal commutation relations. The Lie superalgebra structure seems to be the most convenient for studying systems of composite and elementary particles.

I. INTRODUCTION

It is well known that the simplest case of a Lie algebra generated from bilinear products of creation and annihilation operators which do not change the number of particles, is the case of isospin[1] as it was originally conceived for systems of neutrons and protons. If we include bilinear product which change the number of particles with the isospin operators, this gives rise to the quasispin algebra,[2] the quasispin operators are generalized and defined for any case where the number of quantum states n is even and can be grouped into pairs.

In this paper we show that the set of creation and annihilation operators and their product for a system of composite and "free" particles can have two algebraical structures, namely a Lie superalgebra[3,4] (or equivalently a graded-Lie algebra[5]) structure, or a bigraded-Lie algebra one. Such systems are encountered in astrophysics, chemical kinetics, plasma physics, and other fields. The Lie superalgebra structure gives the "normal commutation"[6] relations (independent fermion fields anticommute), while the bigraded-Lie structure gives the "abnormal commutation" relations[6] (independent fermion fields commute). The well-known choice[6] between commutation relations or anticommutation relations come from the choice between these two algebraical structures. In the bigraded-Lie case only the properties of the constituents appear explicitly, while the Lie superalgebra structure allows us to have explicitly the properties and the presence of bound composites particles in the algebra of observables and thus to have their creation and annihilation operators as dynamical variables as well as the creation and annihilation operators for the constituents. That is why we believe that the Lie superalgebra structure is the best one to use in this field.

II. BIGRADED LIE ALGEBRAICAL STRUCTURE

A. State space

We consider a system of composite and free particles (nuclei, electrons, and atoms); nuclei are considered as elementary particles and atoms are obtained by the association of one nucleus and l electrons.

Let $\mathscr{H}_n$ be the Hilbert state space for one nucleus and J its nuclear spin. If J is a half-integer, then $\mathscr{A}_r\mathscr{H}_n^r$ is the r-fold antisymmetric tensor product of $\mathscr{H}_n$, and

$$\mathscr{F}_a(\mathscr{H}_n)=\bigoplus_{r=0}^{\infty}\mathscr{A}_r\mathscr{H}_n^r \tag{1}$$

is the antisymmetric Fock space over $\mathscr{H}_n$.

If J is an integer, then $\mathscr{S}_r\mathscr{H}_n^r$ is the r-fold symmetric tensor product of $\mathscr{H}_n$, and

$$\mathscr{F}_s(\mathscr{H}_n)=\bigoplus_{r=0}^{\infty}\mathscr{S}_r\mathscr{H}_n^r \tag{2}$$

is the symmetric Fock space over $\mathscr{H}_n$.

Let $\mathscr{H}_e$ be the Hilbert state space for the electron. Then $\mathscr{A}_l\mathscr{H}_e^l$ is the l-fold antisymmetric tensor product of $\mathscr{H}_e$, and

$$\mathscr{F}_a(\mathscr{H}_e)=\bigoplus_{l=0}^{\infty}\mathscr{A}_l\mathscr{H}_e^l \tag{3}$$

is the antisymmetric Fock space over $\mathscr{H}_e$.

The state space for a system of atoms, free nuclei, and free electrons is, when J is a half-integer,

$$\mathscr{F}=\mathscr{F}_a(\mathscr{H}_n)\otimes\mathscr{F}_a(\mathscr{H}_e)=\bigoplus_{r,l}\mathscr{A}_r\mathscr{H}_n^r\otimes\mathscr{A}_l\mathscr{H}_n^l$$

$$=\bigoplus_{(2Jr,2sl)}\mathscr{F}_{(2Jr,2sl)}, \tag{3'}$$

where $2Jr$ is the degree of r nuclei, which means, that when we exchange r nuclei with r other nuclei, the wave function is symmetric or antisymmetric depending upon whether $2Jr$ is even or odd.

Here $2sl$ is the degree of l electrons, s is the electron spin $s=\frac{1}{2}$,

$$\mathscr{F}=\bigoplus_{(2Jr,l)}\mathscr{F}_{(2Jr,l)}. \tag{4}$$

When J is an integer, the state space is

$$\mathscr{E}=\mathscr{F}_s(\mathscr{H}_n)\otimes\mathscr{F}_a(\mathscr{H}_e)=\bigoplus_{r,l}\mathscr{S}_r\mathscr{H}_n^r\otimes\mathscr{A}_l\mathscr{H}_n^l$$

$$=\bigoplus_{2Jr,l}\mathscr{E}_{(2Jr,l)}, \tag{5}$$

and, we limit our study to the space $\mathscr{F}$; the state space $\mathscr{E}$ is treated similarly.

The state space $\mathscr{F}$ is a bigraded space, it is graded according to the $2Jr$ and graded according to electron number. (If $J=\frac{1}{2}$, $\mathscr{F}$ is graded according to nucleus number and graded according to electron number.)

 0022-2488/87/092187-05$02.50

B. Algebra of operators on $\mathscr{F}$

Let $\mathrm{End}_{(2Jr',l')}(\mathscr{F})$ consist of those bilinear maps f of $\mathscr{F}$ into itself such that $f(\mathscr{F}_{2Jr,l}) \subset \mathscr{F}_{(2J(r+r'),l+l')}$.

It is clear that if $f \in \mathrm{End}_{(2Jk,l)}(\mathscr{F})$ and $g \in \mathrm{End}_{(2Jk',l')}(\mathscr{F})$, then the composition $f \cdot g$ lies in $\mathrm{End}_{(2J(k+k'),l+l')}$. The algebra of operators on $\mathscr{F}$ is

$$L = \mathrm{End}(\mathscr{F}) = \bigoplus_{(r_1,l_1)} \mathrm{End}_{(2Jr_1,l_1)}(\mathscr{F}) = \bigoplus_{(r_1,l_1)} L_{(2Jr_1,l_1)}, \tag{6}$$

where L is called a bigraded algebra.

We define a bigraded Lie algebra structure on L if we are given a bilinear map denoted by [,] of $L \times L \to L$ such that if $f \in L_{(2Jr_1,l_1)}$ and $g \in L_{(2Jr_2,l_2)}$,

$$[f,g] = fg - (-1)^{4J^2 r_1 r_2 + l_1 l_2} g f \tag{7}$$

and

$$[f,[g,k]] = [[f,g],h] + (-1)^{4J^2 r_1 r_2 + l_1 l_2}[g,[f,h]], \tag{8}$$

where $f \in L_{(2Jr_1,l_1)}$, $g \in L_{(2Jr_2,l_2)}$, and $h \in L_{(2Jr_3,l_3)}$.

C. Example

For simplicity, we take a quantum mechanical model of a composite particle composed by two types of fermions such that each type of fermion can be represented by only a single state (the two types of fermions are proton and electron).

The state space of this system is

$$\mathscr{F} = \mathscr{F}_{(0,0)} + \mathscr{F}_{(1,0)} + \mathscr{F}_{(0,1)} + \mathscr{F}_{(1,1)}. \tag{9}$$

Denote the creation and annihilation operator for the electron by a^+ and a, and those for the proton by b^+ and b. The space $\mathscr{F}_{(0,0)}$ is spanned by the vacuum state $|0>$, $\mathscr{F}_{(1,0)}$ is spanned by the electron state $|e\rangle = a^+|0\rangle$, $\mathscr{F}_{(0,1)}$ is spanned by the proton state $|p\rangle = b^+|0\rangle$, $\mathscr{F}_{(1,1)}$ is spanned by the proton–electron state $|ep\rangle = a^+b^+|0\rangle$. The state space $\mathscr{F}$ is a four-dimensional space.

The Hamiltonian of this sytem is taken to be

$$H = \alpha_1 a^+ a + \alpha_2 b^+ b + v a^+ b^+ ba.$$

It has four eigenstates, consisting of the vacuum state $|0\rangle$ with eigenvalue zero, the proton state $b^+|0\rangle$ with eigenvalue α_2, the electron state with eigenvalue α_1 and proton–electron state $a^+b^+|0\rangle$ with eigenvalue $\alpha_1 + \alpha_2 + v$; $a^+b^+|0\rangle$ can be thought as a simplified model of a composite particle (here a hydrogen atom). The operators $A^+ = a^+b^+$ and $A = ba$ are the creation and annihilation operators of a hydrogen atom.

The algebra of operators on $\mathscr{F}$ is

$$\mathrm{End}(\mathscr{F}) = L = L_{(-1,-1)} + L_{(-1,1)} + L_{(0,-1)} + L_{(-1,0)} + L_{(0,0)} + L_{(1,0)} + L_{(0,1)} + L_{(1,-1)} + L_{(1,1)}, \tag{10}$$

where $L_{(-1,-1)}$, $L_{(-1,1)}$, $L_{(1,-1)}$, and $L_{(1,1)}$ are one dimensional and are spanned, respectively, by ab, ab^+, a^+b, and b^+a^+; $L_{(0,-1)}$, $L_{(-1,0)}$, $L_{(1,0)}$, and $L_{(0,1)}$ are two dimensional and are spanned, respectively, by $\{b,a^+ab\}$, $\{a,b^+ba\}$, $\{a^+,b^+a^+b\}$, and $\{b^+,b^+a^+a\}$; and $L_{(0,0)}$ is four dimensional and is spanned by I, a^+a, b^+b, b^+a^+ab; where I denotes the identity operator. The algebra L is 16 dimensional.

According to Eq. (7) the proton and the electron creation and annihilation operators verify the following relations:

$$\begin{aligned}
&[a,a] = aa + aa = 0,\\
&[b,b] = bb + bb = 0,\\
&[a,a^+] = aa^+ + a^+a = 1,\\
&[b,b^+] = bb^+ + b^+b = 1,\\
&[a^+,b^+] = a^+b^+ - b^+a^+ = 0,\\
&[a^+,b] = a^+b - ba^+ = 0,\\
&[a,b] = ab - ba = 0,\\
&[a,b^+] = ab^+ - b^+a = 0.
\end{aligned} \tag{11}$$

We have omitted writing the other commutation and anticommutation relations since they can be obtained easily from Eqs. (7) and (11).

Yet, note that the operators A^+ and A are not Bose operators since they satisfy nontrivial commutation relations:

$$\begin{aligned}
&[A,A^+] = AA^+ - A^+A = 1 - a^+a - b^+b,\\
&[A,a^+] = Aa^+ + a^+A = b,\\
&[A,b^+] = Ab^+ + b^+A = a,\\
&[A,a] = [A,b] = 0,\\
&[A,A] = 0.
\end{aligned} \tag{12}$$

The last four equations in (11) are the abnormal commutation relations.

Consequently, the bigraded structure gives the abnormal commutation relations between independent fermion creation and annihilation operators.

A representation of L on $\mathscr{F}$ which agrees with the bigraded structure is

$$a = \begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}, \quad b = \begin{pmatrix}0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix},$$

$$a^+ab = \begin{pmatrix}0&0&0&0\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}, \quad b^+ = \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix},$$

$$ab = \begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, \quad a^+b = \begin{pmatrix}0&0&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, \tag{13}$$

$$I = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}, \quad a^+a = \begin{pmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix},$$

$$b^+b = \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}, \quad b^+a^+ab = \begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}.$$

We have omitted writing the Hermitian conjugate of these operators.

III. LIE SUPERALGEBRA STRUCTURE

A. State space

We have seen [Eq. (4)] that $\mathscr{F} = \oplus_{(2Jr,l)} \mathscr{F}_{(2Jr,l)}$.

Let $\mathscr{F}_{(2Jr_1,l_1)}$ be the subspace of $\mathscr{F}$ representing states of r_1 nuclei and l_1 electrons. We consider bound atoms formed by the association of one nucleus and n electrons. We describe these bound atoms via an orthonormal but incomplete set of wave functions $\Phi_\alpha(X,x_1,...,x_n)$. (Here X refers to the nucleus coordinates including spin, and x to the electron coordinates.) To describe the "free" electrons and nuclei, we choose a complete orthonormal set $\Phi_j(X)$ of one-nucleus wave functions and a similar set $\Phi_i(x)$ of one-electron wave functions. We introduce functions $\Psi_m(X_1\cdots X_{r_1},\ x_1\cdots x_{l_1})$ corresponding to m bound atoms, $r_1 - m$ free nuclei, and $l_1 - m$, free electrons describing a system of r_1 nuclei and l_1 electrons. These functions can be expressible as linear combinations:

$$\begin{aligned}\Psi_m(X_1,...,X_{r_1},\ x_1,...,x_{l_1}) &= \sum_{\substack{\alpha_1\cdots\alpha_m \\ j_{m+1}\cdots j_{r_1} \\ i_{mn+1}\cdots i_{l_1}}} C(\alpha_1\cdots\alpha_m, j_{m+1}\cdots j_{r_1}, i_{mn+1}\cdots i_{l_1}) \\ &\times \Phi_{\alpha_1}(X_1,x_1\cdots x_n)\cdots\Phi_{\alpha_m}(X_m,x_{(m-1)n+1}\cdots x_{mn}) \\ &\times \Phi_{J_{m+1}}(X_{m+1})\cdots\Phi_{j_{r_1}}(X_{r_1}) \\ &\times \Phi_{i_{mn+1}}(x_{mn+1})\cdots\Phi_{i_{l_1}}(x_{l_1}). \end{aligned} \quad (14)$$

The coefficients $C(\alpha_1\cdots\alpha_m,\ j_{m+1}\cdots j_{r_1},\ i_{mn+1}\cdots i_{l_1})$ are totally symmetric or antisymmetric under interchanges of the atomic indices α depending upon whether $2J+n$ is even or odd, and totally antisymmetric under interchanges of the nucleus (electron) indices $j(i)$ (J is half-integer). It is clear that the commutation or the anticommutation relations between the atoms will depend on the sign of $(-1)^{2J+n}$. We want that $(-1)^{2J+n}$ appears explicitly in the commutation or the anticommutation relations which will characterize the algebra of operators on $\mathscr{F}$.

We have $\mathscr{F} = \oplus_{(2Jr_1,l_1)} \mathscr{F}_{(2Jr_1,l_1)}$.

We make $\mathscr{F}$ into a graded vector space by setting

$$\mathscr{F}_i = \oplus_{2Jr_1+l_1=i} \mathscr{F}_{(2Jr_1,l_1)}. \quad (15)$$

Then

$$\mathscr{F} = \oplus_i \mathscr{F}_i. \quad (16)$$

We say that $\mathscr{F}$ is a graded vector space. [Note that the wave functions that correspond to physical states in $\mathscr{F}$ are $\chi_m(X_1\cdots X_{r_1},x_1\cdots x_{l_1}) = \hat{A}\Psi_m(X_1\cdots X_r,x_1\cdots x_{l_1})$, where $\hat{A}$ is the antisymmetrizer operator.]

Here $\mathscr{F}_i$ may be written in the following form:

$$\begin{aligned}\mathscr{F}_i &= \oplus_{2Jr_1+l_1=i} \mathscr{F}_{(2Jr_1,l_1)} \\ &= \oplus_m \hat{A}(\mathscr{F}^a_{(2J+n)m} \otimes \mathscr{F}^b_{2J(r_1-m)} \otimes \mathscr{F}^c_{l_1-mn}), \end{aligned} \quad (17)$$

where $\mathscr{F}^a_{(2J+n)m}$ is the state space for the m bound atoms, the parity of a state in $\mathscr{F}^a_{(2J+n)m}$ is $(-1)^{(2J+n)m}$, and the parity of an atom is $(-1)^{2J+n}$; $\mathscr{F}^b_{2J(r_1-m)}$ is the state space for the (r_1-m) free nucleus, the parity of a state in $\mathscr{F}^b_{2J(r_1-m)}$ is $(-1)^{2J(r_1-m)}$, and the parity of a nucleus is $(-1)^{2J}$; and $\mathscr{F}^c_{l_1-mn}$ is the state space for the (l_1-mn) free, the parity of a state in $\mathscr{F}^c_{l_1-mn}$ is $(-1)^{l_1-mn}$, and the parity of an electron is (-1).

A state in $\mathscr{F}_i$ can be represented by

$$\begin{aligned}|\chi_m\rangle &= \frac{1}{(r_1!l_1!)}\frac{1}{2}\int dX_1\cdots dX_{r_1}\, dx_1\cdots dx_{l_1} \\ &\times \chi_m(X_1\cdots X_{r_1},x_1\cdots x_{l_1})\Psi^+(X_1)\cdots \\ &\times \Psi^+(X_{r_1})\Psi^+(x_1)\cdots\Psi^+(x_{l_1})|0\rangle, \end{aligned} \quad (18)$$

where $\Psi^+(X)$ is the field operator for the nucleus, $\Psi^+(x)$ is the field operator for the electron, and $|0\rangle$ is the vacuum state. From (14) we can write (18) in the following form:

$$\begin{aligned}|\chi_m\rangle &= \frac{(n!)^{m/2}}{(r_1!l_1)^{1/2}}\sum_{\alpha,i,j} d(\alpha,i,j) \\ &\times A^+_{\alpha_1}\cdots A^+_{\alpha_m} b^+_{jm+1}\cdots b^+_{jr_1} a^+_{i_{mn+1}}\cdots a^+_{l_1}|0\rangle, \end{aligned} \quad (19)$$

where

$$\begin{aligned}A^+_{\alpha_i} &= \frac{1}{(n!)^{1/2}}\int dX_i\, dx_{(i-1)n+1}\cdots dx_{i_n} \\ &\times \Phi_{\alpha_i}(X_i,x_{(i-1)n+1}\cdots x_{i_n}) \\ &\times \Psi^+(X_i)\Psi^+(x_{(i-1)n+1})\cdots\Psi^+(x_{i_n}) \end{aligned} \quad (20)$$

is the creation operator of an atom in the state α_i,

$$b^+_j = \int dX_j\, \varphi_j(X_j)\Psi^+(X_j) \quad (21)$$

is the creation operator of a nucleus in the state j, and

$$a^+_i = \int dx_i\, \varphi_i(x_i)\Psi^+(x_i) \quad (22)$$

is the creation operator of an electron in the state i.

The annihilation operators are defined by

$$A_{\alpha_i} = (A^+_{\alpha_i})^+, \quad b_j = (b^+_j)^+, \quad a_i = (a^+_i)^+. \quad (23)$$

Equation (17) can be obtained by another method. Let P_m be the space of all wave functions χ_m; it is a subspace of $\mathscr{F}$.

It is obvious that the space P_m contain not only the states having m atoms but also the states having $m+1$, $m+2$,..., etc. bound atoms, and we have the following relations:

$$P_0 \supset P_1 \supset\cdots\supset P_m \supset\cdots\supset P_q, \quad q = \min(r_1,l_1). \quad (24)$$

We have a sequence of embedded vector spaces, we can define in the usual way the quotient spaces or equivalently the factor spaces.

Let

$$P_j/P_{j+1} = Q_j. \quad (25)$$

Then Q_j is exactly the state space of physical j-atom wave functions, $r_1 - j$ free proton wave functions, and $l_1 - jn$ free electron wave functions. We have then

$$P_0 = Q_0 \oplus Q_1 \oplus\cdots\oplus Q_m \oplus\cdots\oplus Q_q = \mathscr{F}_i. \quad (26)$$

B. Algebra of operators on $\mathscr{F}$

We have seen [Eq. (15)] that

$$\mathscr{F} = \oplus_i \mathscr{F}_i .$$

Let $\mathrm{End}_k(\mathscr{F})$ consist of those linear maps, f, of $\mathscr{F}$ into itself such that $f\mathscr{F}_n \subset \mathscr{F}_{n+k}$. (27)

The algebra of operators on $\mathscr{F}$ is

$$L = \mathrm{End}(\mathscr{F}) = \oplus_k \mathrm{End}_k(\mathscr{F}) = \oplus L_k, \tag{28}$$

where $k = 2Jr + l$. Here L is a graded algebra or equivalently a superalgebra.[3–4]

We say that L is a graded Lie algebra[5] or equivalently a Lie superalgebra, if we are given a bilinear map denoted by $[\ ,\ ]$ of $L \times L \to L$ such that the following three conditions hold:

$$[L_k, L_l] \subset L_{k+l}, \tag{29}$$

$$[f,g] = -(-1)^{kl}[g,f] \quad \text{for } f \in L_k, \quad g \in L_l, \tag{30}$$

and

$$[f,[g,h]] = [[f,g],h] + (-1)^{kl}[g,[f,h]], \quad h \in L_m. \tag{31}$$

Condition (29) says that the bracket multiplication is consistent with the grading. Condition (30) is the graded version of anticommutativity (supersymmetry). Condition (31) is the graded version of Jacobi's identity.

We define a Lie superalgebra structure on $L = \mathrm{End}(\mathscr{F}) = \oplus_k L_k$, when $k = 2Jr + l$ by setting

$$[f,g] = f \cdot g - (-1)^{kp} g \cdot f, \tag{32}$$

for $f \in L_k$ and $g \in L_p$.

It is obvious that conditions (29) and (30) are satisfied and a straightforward verification shows that condition (31) is also satisfied.

C. Example

We take the same quantum mechanical model as we do in Sec. I. The state space for this system is

$$\mathscr{F} = \mathscr{F}_0 \oplus \mathscr{F}_1 \oplus \mathscr{F}_2, \tag{33}$$

where $\mathscr{F}_0 = \mathscr{F}_{(0,0)}$ is spanned by the vacuum state, $\mathscr{F}_1 = \mathscr{F}_{(1,0)} \oplus \mathscr{F}_{0,1}$ is spanned by the electron and the proton states $|e\rangle$ and $|p\rangle$, and $\mathscr{F}_2 = \mathscr{F}_{(1,1)}$ is spanned by the proton–electron state $|ep\rangle$.

The algebra of operator on $\mathscr{F}$ is

$$\mathrm{End}(\mathscr{F}) = L = L_{-2} \oplus L_{-1} \oplus L_0 \oplus L_1 \oplus L_2, \tag{34}$$

where L_{-2} and L_2 are one dimensional and are spanned, respectively, by ab and b^+a^+.

Here L_{-1} and L_1 are four dimensional and are spanned by $\{a, b, a^+ab, b^+ba\}$ and $\{a^+, b^+, b^+a^+a, b^+a^+b\}$, and L_0 is six dimensional and is spanned by $\{I, a^+a, b^+b, b^+a^+ab, b^+a, a^+b\}$.

According to Eq. (32), proton and electron creation and annihilation operators satisfy the following relations:

$$\begin{aligned}
&[a,a] = aa + aa = 0,\\
&[b,b] = bb + bb = 0,\\
&[a,a^+] = aa^+ + a^+a = 1,\\
&[b,b^+] = bb^+ + b^+b = 1,\\
&[a^+,b^+] = a^+b^+ + b^+a^+ = 0,\\
&[a,b^+] = ab^+ + b^+a = 0,\\
&[a^+,b] = a^+b + ba^+ = 0,\\
&[a,b] = ab + ba = 0.
\end{aligned} \tag{35}$$

The last four equations in (35) show that proton and electron creation and annihilation operators anticommute.

The operators A^+, A, a, b, a^+, and b^+ satisfy the following commutation relations:

$$\begin{aligned}
&[A,A^+] = AA^+ - A^+A = 1 - a^+a - b^+b,\\
&[A,a^+] = Aa^+ - a^+A = b,\\
&[A,b^+] = Ab^+ - b^+A = -a,\\
&[A,a] = Aa - aA = 0,\\
&[A,b] = Ab - bA = 0,\\
&[A,A] = AA - AA = 0.
\end{aligned} \tag{36}$$

We have omitted writing the Hermitian conjugate of (36). [Note the difference between (36) and (12).]

Consequently, the Lie superalgebra structure gives the normal commutation relations. A representation of L in $\mathscr{F}$ which agrees with (32) is

$$a = \begin{pmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{pmatrix}, \quad b = \begin{pmatrix} 0&0&1&0\\ 0&0&0&-1\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix},$$

$$a^+ab = \begin{pmatrix} 0&0&0&0\\ 0&0&0&-1\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}, \quad b^+ba = \begin{pmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{pmatrix}, \tag{37}$$

$$ab = \begin{pmatrix} 0&0&0&-1\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}, \quad a^+b = \begin{pmatrix} 0&0&0&0\\ 0&0&1&0\\ 0&0&0&0\\ 0&0&0&0 \end{pmatrix}.$$

We have omitted writing the Hermitian conjugate of (37).

IV. CONNECTION BETWEEN THE TWO STRUCTURES

We have associated to each structure a representation in the state space $\mathscr{F}$ [see (13) and (37)].

These representations can be connected by a linear supertransformation $\hat{L}$, the denomination supertransformation comes from the fact that $\hat{L}$ acts on the space of operators (which is a vector space), not on state space. This supertransformation is exactly the Klein transformation.

Denote by a'', b'', $(b^+a^+a)''$,... the representation of algebra L in the bigraded Lie case, and a', b', $(b^+a^+a)'$,... the representation of L in the Lie superalgebraical case.

We can find a linear supertransformation $\hat{L}$ such that

$$a' = \hat{L}a'', \ b' = \hat{L}b'', \ldots,$$

$$a'_{ij} = L_{ij,kl} a''_{kl}, \quad b'_{ij} = L_{ij,kl} b''_{kl}, \ldots,$$

where

a'_{ij} is a matrix element of a',

a''_{kl} is a matrix element of a''.

The supertransformation $\hat{L}$ can be expressed as a symmetric matrix, with the following non-null elements:

$$L_{ii,ii} = 1 \quad \text{for } i = 1,\ldots,4,$$

$$L_{12,12} = L_{13,13} = L_{23,23} = L_{34,34} = 1,$$

$$L_{14,14} = L_{24,24} = -1.$$

V. CONCLUSION

A system made up of two different fermion species can have two algebraical structures a bigraded-Lie structure and a graded-Lie one.[3–5] The choice between these two structures follows from the standard argument[6] according to which the choice between commutation or anticommutation relations of kinematically independent fermion species is a matter of convention. If one wants to have creation and annihilation operators of composite particles as dynamical variables of the system, the Lie superalgebraical structure is the most convenient since these operators appear explicitly with their parity in the algebra of operators, while in the bigraded Lie structure, only the operators for the constituents appear explicitly. These two structures are connected by a supertransformation which is just the Klein transformation.

[1]J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), pp. 158, 220, and 255.

[2]A. K. Kerman, Ann. Phys. (NY) **12**, 300 (1961); H. G. Lipkin, Nucl. Phys. **26**, 147 (1961).

[3]V. G. Kac, Adv. Math. **26**, 8 (1977).

[4]D. A. Leites, Russian Math. Surveys **35**, 1 (1980).

[5]L. Corwin, Y. Ni'Eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).

[6]R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964), pp. 146ff.

The physical basis of an n g-boson system

Xiao-yan Sen and Z. R. Yu
Department of Physics, Nanjing University, Nanjing, People's Republic of China

(Received 11 September 1986; accepted for publication 7 January 1987)

The mathematical group chain of an n g-boson system is

$$\begin{array}{l} SU(9)\supset SO(9)\supset SO(5)\times SU_1(2)\times SU_2(2) \\ \qquad\qquad\qquad\qquad \cup \\ \qquad\qquad\qquad\qquad SU_a(2)\times SU_b(2). \end{array}$$

In this paper, the mathematical basis of this group chain is obtained by using the briefer method and then constructing the states of n g-bosons with good angular momentum.

I. INTRODUCTION

In the preceding paper,[1] we have shown that the mathematical group chain for n g-bosons is

$$\begin{array}{l} SU(9)\supset SO(9)\supset SO(5)\times SU_1(2)\times SU_2(2) \\ \qquad\qquad\qquad \cup \\ \qquad\qquad\qquad SU_a(2)\times SU_b(2) \end{array}$$

and the corresponding mathematical basis is explicitly given by

$$\begin{aligned} |nv;p,\Lambda\alpha\beta;\Sigma\gamma\delta\rangle = {} & \left(\frac{(2v+7)!!}{2^p p!(2p+2v+7)!!}\right)^{1/2}\frac{2^{\Lambda+\Sigma}}{(2\Lambda)!(2\Sigma)!}\left(\frac{(\Lambda+\alpha)!(\Lambda+\beta)!(\Sigma+\gamma)!(\Sigma\Delta+\delta)!}{2^{\alpha+\beta+\gamma+\delta}(\Lambda-\alpha)!(\Lambda-\beta)!(\Sigma-\gamma)!(\Sigma-\delta)!}\right)^{1/2} \\ & \times(S^+)^p(\lambda_-)^{\Lambda-\alpha}(\nu_-)^{\Lambda-\beta}(\sigma_-)^{\Sigma-\gamma}(\tau_-)^{\Sigma-\delta}\sum_l \Gamma_l'(g_4^+)^{2\Sigma}(Z_1^+)^{\omega-l}(Z_2^+)^l \\ & \times\sum_\Delta \Gamma_\Delta(g_0^+)^{p-2\Lambda-2\Delta}(\beta_0^+)^\Delta(g_2^+)^{2\Lambda}(\sqrt{(2\Lambda)!})^{-1}|0\rangle, \end{aligned} \tag{1}$$

where

$$2\omega=v-p-2\Sigma$$

and

$$\Gamma_l'=\left(\frac{(2\Sigma+1)\omega!(2\Sigma+\omega+1)!(4\Sigma+2p+2\omega+5)!!}{(4\Sigma+2p+4\omega+5)!!}\right)^{1/2}\frac{(2p+3)!!}{l!(\omega-l)(\omega+2\Sigma+1-l)!(2p+2l+3)!!}, \tag{2}$$

$$\Gamma_\Delta=\left(\frac{(2\Lambda+1)!(p-2\Lambda)!(p+2\Lambda+1)!!(p+2\Lambda+2)!!}{2^{2\Lambda+1}(2p+1)!!}\right)^{1/2}\frac{1}{2^\Delta\Delta!(\Delta+2\Lambda+1)!(p-2\Delta-2\Lambda)!}. \tag{3}$$

Note that expression (1) has some differences with (18a) in Ref. 1. Here we have written it in a more compact form.

II. THEORY

In Ref. 1, we have already constructed the physical basis for the $n=2$ g-boson system. Here we give the physical basis for an arbitrary n-boson system.

As described in Ref. 1, in order to form states with good angular momentum, we take a maximum weight state which has a weight value of $M=4v$ and thus a unique value of L, then apply the L_- operator given in Eq. (8) of Ref. 1. It can be easily shown that the states (1) are eigenstates of L_0 with eigenvalue $M=\beta+3\alpha+\delta+7\gamma$. The physical states thus have to be a linear combination of states with the same M, n, and v.

For arbitrary n, the action of the operators appearing in L_- is given as follows:

$$\nu_-|nv;p,\Lambda\alpha\beta;\Sigma\gamma\delta\rangle=\left[\tfrac{1}{2}(\Lambda+\beta)(\Lambda-\beta+1)\right]^{1/2}|nv;p,\Lambda\alpha\beta-1;\Sigma\gamma\delta\rangle, \tag{4a}$$

$$\tau_-|nv;p,\Lambda\alpha\beta;\Sigma\gamma\delta\rangle=\left[\tfrac{1}{2}(\Sigma+\delta)(\Sigma-\delta+1)\right]^{1/2}|nv;p,\Lambda\alpha\beta;\Sigma\gamma\delta-1\rangle, \tag{4b}$$

$$\begin{aligned} & U^{(1/2)(1/2)}_{(-1/2)(1/2)}|nv;p,\Lambda\alpha\beta;\Sigma\gamma\delta\rangle \\ & =\frac{1}{2}\left[\frac{(p-2\Lambda)(p+2\Lambda+3)(\Lambda-\alpha+1)(\Lambda+\beta+1)}{(2\Lambda+1)(2\Lambda+2)}\right]^{1/2}|nv;p,\Lambda+\tfrac{1}{2},\alpha-\tfrac{1}{2},\beta+\tfrac{1}{2};\Sigma\gamma\delta\rangle \\ & \quad-\frac{1}{2}\left[\frac{(p-2\Lambda+1)(p+2\Lambda+2)(\Lambda+\alpha)(\Lambda-\beta)}{2\Lambda(2\Lambda+1)}\right]^{1/2}|nv;p,\Lambda-\tfrac{1}{2},\alpha-\tfrac{1}{2},\beta+\tfrac{1}{2};\Sigma\gamma\delta\rangle, \end{aligned} \tag{4c}$$

 0022-2488/87/092192-02$02.50

$$T^{(10)(1/2)}_{(1/2)(1/2)(1/2);(-1/2)(1/2)}|nv;p\,\Lambda\alpha\beta;\,\Sigma\gamma\delta\rangle$$
$$= \sum_{\substack{i=-1,1\\ j=-\frac12,\frac12\\ k=-\frac12,\frac12}} (-1)^{i'+j+k}\langle p+i\|T\|p\rangle \left[\frac{\Omega^p_{ij}(\Omega^p_{ij}+1)\Theta_{\alpha j}\Theta_{\beta j}\Theta_{\gamma k}\Theta_{\delta k}}{[2\Lambda+2f(j)][2\Lambda+2f(j)+1][p+1+f(-i)](2p+3)(2\Sigma+1)^2}\right]^{1/2}$$
$$\times|nv;p+i,\Lambda+j,\alpha+\tfrac12,\beta+\tfrac12;\Sigma+k,\gamma-\tfrac12,\delta+\tfrac12\rangle, \quad (4d)$$

where

$$f(x)=\begin{cases}0, & x<0,\\ 1, & x>0,\end{cases}\quad i'=\begin{cases}i+1, & i<0,\\ i, & i>0,\end{cases}\quad \Omega^p_{ij}=p+(-1)^{f(i)+1}(i+2f(j))+2\Lambda(-1)^{f(i)+f(j)},$$
$$\Theta_{\rho j}=\Lambda+\rho(-1)^{f(j)+1}+f(j),\quad \rho=\alpha,\beta,\quad \Phi_{\sigma k}=\Sigma+\sigma(-1)^{f(k)}+f(k),\quad \sigma=\gamma,-\delta,$$
$$\langle p\pm1\|T\|p\rangle=\left[\frac{(2\Sigma+\frac32\mp\frac12)(a_\mp d_-\pm b_\mp d_+)}{(2\Sigma+\frac32\pm\frac12)(a_+b_-+a_-b_+)}\right]^{1/2}\left[\frac{(p+\frac32\mp\frac12)(2p+3)}{(p+\frac32\pm\frac32)(2p+3\pm2)}\right]^{1/4},$$

and

$$a_\pm=2\left[\frac{2p+3\pm2}{(p+\frac12\pm\frac12)(p+\frac52\pm\frac12)(2p+3)}\right]^{1/2},$$
$$b_\pm=\left[\frac{(p\pm2\Lambda+1\pm2)(p\pm2\Lambda+2\pm2)}{(2\Lambda+1)(\Lambda+1)}+\frac{(p\pm2\Lambda+1)(p\pm2\Lambda+2)}{\Lambda(2\Lambda+1)}\right]$$
$$\times\frac{\beta}{[(p+\frac12\pm\frac12)(p+\frac52\pm\frac12)(2p+2\pm1)(2p+5\pm1)]^{1/2}},\quad d_\pm=\frac{2\Sigma(\Sigma+1)}{\Sigma(\Sigma+1)\pm\gamma\delta}.$$

As an example, we discuss the case $n=3$. The possible values of L are 12, 10, 9, 8, 7, 6^2, 5, 4^2, 3, 2, 0, and the maximum weight state is

$$|L=12,M=12\rangle=(1/\sqrt6)(g_4^+)^3|0\rangle$$

corresponding to the state

$$|n=3,v=3;p=0,\Lambda=0,\alpha=0,\beta=0;$$
$$\Sigma=\tfrac32,\gamma=\tfrac32,\delta=\tfrac32\rangle. \quad (5)$$

With the use of the previous formulas, the action of L_- on the highest weight state can be calculated to yield

$$L_-|33;0,000;\tfrac32\tfrac32\tfrac32\rangle=2\sqrt3|33;0,000;\tfrac32\tfrac32\tfrac12\rangle$$
$$=\sqrt6(g_3^+)(g_4^+)^2|0\rangle.$$

Since the state on the left-hand side corresponds to the state $|L=12,M=12\rangle$, the state on the right-hand side necessarily corresponds to $|L=12,M=11\rangle$,

$$|L=12,M=11\rangle=(1/\sqrt2)(g_3^+)(g_4^+)^2|0\rangle. \quad (6)$$

By acting the operator L_- once again, we obtain

$$|L=12,M=10\rangle=(8/23)^{1/2}|33;0,000;\tfrac32\tfrac32-\tfrac12\rangle$$
$$-(7/46)^{1/2}|33;1,\tfrac12\tfrac12\tfrac12;111\rangle$$
$$=(8/23)^{1/2}(g_3^+)^2(g_4^+)|0\rangle$$
$$-(7/46)^{1/2}(g_2^+)(g_4^+)^2|0\rangle. \quad (7)$$

The state $|L=10,M=10\rangle$ can be obtained by its normalization and its orthogonality with $|L=12,M=10\rangle$:

$$|L=10,M=10\rangle=(7/46)^{1/2}(g_3^+)^2g_4^+|0\rangle$$
$$+(8/23)^{1/2}(g_2^+)(g_4^+)^2|0\rangle. \quad (8)$$

By using this procedure, we can construct all states of n g-bosons with good angular momentum.

[1]Z. R. Yu, O. Scholten, and H. Z. Sun, J. Math. Phys. 27, 442 (1986).

Exact solutions of the multidimensional classical ϕ^6-field equations obtained by symmetry reduction

P. Winternitz
Centre de Recherches Mathématiques, Université de Montréal, C.P. 6128, succursale A, Montréal, Quebec, Canada H3C 3J7

A. M. Grundland
Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1C 5S7

J. A. Tuszyński
Department of Physics, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X7

(Received 18 November 1986; accepted for publication 15 April 1987)

The often used ϕ^6 model of classical critical phenomena is studied in $(3+1)$-dimensional Minkowski and Euclidean spaces. The Euler–Lagrange equations describing the kinetics of the scalar order parameter are in this case nonlinear Klein–Gordon equations. The method of symmetry reduction is systematically applied to derive all the solutions invariant under subgroups with generic orbits of codimension 1. Whenever the obtained ordinary differential equations have the Painlevé property, they can be transformed to one of two standard forms. These are then solved in terms of elliptic functions or elementary ones. This results in a large number of new exact solutions. Particularly interesting solutions are found in the immediate vicinity of the tricritical point. Our treatment of the ϕ^6 theory is complete only for the four-dimensional spaces M(3,1) and E(4), but many of the results are given for the more general cases of $M(n,1)$ and $E(n+1)$.

I. INTRODUCTION

The purpose of this paper is to obtain new analytical solutions of the equation of motion of the classical ϕ^6-field theory. The Lagrangian density of this theory is given by the expression

$$\mathscr{L}(\mathbf{x}) = \tfrac{1}{2}\partial_\mu \phi\, \partial^\mu \phi - V(\phi), \tag{1.1}$$

$$V(\phi) = a_2\phi^2 + a_4\phi^4 + a_6\phi^6, \tag{1.2}$$

where $\phi(x) = \phi(x_0,x_1,\dots,x_n)$ is a classical scalar field and summation over repeated indices is to be understood.

The equation of motion is the Lagrange–Euler equation, obtained by minimizing the corresponding action. For the Lagrangian density, Eq. (1.1), the equation of motion is

$$\Box_\epsilon \phi = -2(a_2\phi + 2a_4\phi^3 + 3a_6\phi^5), \tag{1.3}$$

where

$$\Box_\epsilon \equiv \frac{\partial^2}{\partial x_0^2} + \epsilon \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \quad (\epsilon = \pm 1),$$

is the Laplace–Beltrami operator in Minkowski space $M(n,1)$ $(\epsilon = -1)$ or in Euclidean space $E(n+1)$ $(\epsilon = +1)$. Also, in Eq. (1.3) a_2, a_4, and a_6 are constants and $a_6 \neq 0$. We shall call Eq. (1.3) the nonlinear Klein–Gordon equation (NLKGE).

The method used consists essentially of two steps, applicable to a large class of partial differential equations (PDE's). The first step is that of *symmetry reduction* from the considered PDE to an ordinary differential equation (ODE). In particular, for the NLKGE (1.3) this is achieved by letting

$$\phi(x) = \rho(x)\mathrm{F}(\xi(x)), \tag{1.4}$$

where ρ and ξ are explicitly given by symmetry considerations. The function $F(\xi)$ satisfies an ODE obtained by substituting Eq. (1.4) into the NLKGE, Eq. (1.3).

The second step is to integrate, if possible, the obtained ODE's. In degenerate cases, as we shall see below, the ODE simplifies either to an algebraic equation or to a first-order ODE that can be integrated directly. For most symmetry variables ξ and multipliers ρ, however, we obtain a nonlinear second-order ODE. In this case we perform a *singularity analysis* in order to establish whether the equation is of the Painlevé type (no moving critical points).[1–4] In the affirmative case we then reduce the ODE to one of the 50 standard forms[1] established by Painlevé and Gambier.[4] It turns out then that we can always integrate the resultant equation in terms of either elementary functions or Jacobi elliptic functions.

In general, the symmetry group of the NLKGE, Eq. (1.3), is simply the corresponding Poincaré group $P(n,1)$ for $M(n,1)$ or Euclidean group $P(n+1,0)$ for $E(n+1)$. However, if the right-hand side of Eq. (1.3) is a homogeneous polynomial, i.e., $a_2 = a_4 = 0$, the group is larger. Indeed, in addition to translations, rotations, and Lorentz boosts the symmetry group of the equation $\Box_\epsilon \phi = a\phi^m$ includes dilations and hence is the corresponding similitude group: $\mathrm{Sim}(n,1)$ for $\epsilon = -1$ or $\mathrm{Sim}(n+1)$ for $\epsilon = +1$. This additional symmetry will allow us to obtain many new analytical solutions of the NLKGE, Eq. (1.3), for $a_2 = a_4 = 0$, i.e.,

$$\Box_\epsilon \phi = a\phi^5, \quad a = -6a_6, \quad \epsilon = \pm 1. \tag{1.5}$$

For $m = (n+3)/(n-1)$ the invariance group is well known to be even larger,[5] namely the entire conformal group of space-time. Since we concentrate below on the case $m = 5$, $n = 3$, not satisfying the above relation, we do not

 0022-2488/87/092194-19$02.50

dwell upon this point here. The NLKGE, Eq. (1.3), is of considerable interest in classical and quantum field theory and also in condensed matter theory. In all these areas it is considered in connection with various critical phenomena.

In the context of classical and quantum field theory, NLKGE's typically arise as equations of motion for scalar meson fields with the Lagrangian density (1.1). The nonlinear field potential $V(\phi)$ depends on parameters whose values may cause it to become multistable, e.g., "mass" a_2, coupling constants a_4 and a_6 in the case of Eq. (1.3). In particular, massive or massless theories with polynomial potentials $V(\phi)$ demonstrate a variety of interesting nonperturbative features. In classical theories $\phi(\mathbf{x})$ is a c-number field which can be considered a first-order approximation, or can be subsequently quantized in order to be useful in quantum theories, where $\phi(\mathbf{x})$ is an operator field. For reviews of quantum field theory applications of classical solutions, see, e.g., Jackiw[6] and Rajaraman.[7] Among the interesting features that we wish to mention are the following.

(i) The existence of a degenerate vacuum leading to the possibility of spontaneous symmetry breaking.

(ii) The presence of exact, "topologically stable" solutions that cannot be obtained using perturbation techniques. These may have the form of kinks [$\phi(\xi)$ goes from one constant value ϕ_1 at $\xi \to -\infty$, to a different one ϕ_2 at $\xi \to +\infty$], solitary waves [localized solutions such that $\phi(\xi)$ approaches the same constant ϕ_0 at $\xi \to \pm\infty$, where ϕ_0 may or may not be 0], various types of algebraic solutions and, finally, periodic solutions which are usually expressed in terms of Jacobi or Weierstrass elliptic functions.[8] These solutions can, in turn, serve as the basis for further perturbative calculations, perturbing around them, rather than around solutions of linearized equations.

(iii) Theories with polynomial self-interactions, at least in $(1+1)$ or $(2+0)$ dimensions share some of the interesting properties of completely integrable field theories. A pertinent example of such a theory is obtained for $V(\phi) = 1 - \cos\phi$ in Eq. (1.2). The equation of motion is then the sine–Gordon equation

$$\phi_{x_0x_0} - \phi_{x_1x_1} = \sin\phi, \qquad (1.6)$$

which can be solved by inverse scattering techniques.[9] The set of solutions includes kinks, multikinks, and periodic solutions (cnoidal waves). However, there is an important difference between the solitary waves and kinks of polynomial theories and those of completely integrable theories. While both types may be stable with respect to various perturbations, solitons (the latter type) are also stable with respect to mutual interactions: they survive collisions amongst each other. The integrability properties of soliton theories do not easily generalize to more than two dimensions. The methods employed in this paper, however, do not rely on integrability and are indeed applicable in $(n+1)$ dimensions.

In the context of relativistic field theory $\phi(\mathbf{x})$ is in general a complex field (a scalar meson wave function), a_2 is a mass, and a_4 and a_6 are coupling constants corresponding to different types of nonlinear self-interactions. In this case we have $\epsilon = -1$, on the other hand, the $\epsilon = +1$ case is also of interest, since it leads to imaginary time solutions, i.e., instantons which provide evidence for quantum mechanical tunneling.[6,7]

In condensed matter theory a Hamiltonian density of the Landau–Ginzburg–Wilson[10] form

$$\mathscr{H} = (m/2)\phi_t^2 + (D/2)(\nabla\phi)^2 + a_2\phi^2 + a_4\phi^4 + a_6\phi^6 \qquad (1.7)$$

is often derived as a continuum limit of an appropriate lattice Hamiltonian. Here, ϕ is an order parameter whose average value is zero in the disordered phase and nonzero in the ordered phase. On a microscopic scale, ϕ can either be discrete valued (spin) or continuous (elongation). The potential energy of the lattice Hamiltonian has an anharmonic on-site part and a harmonic (for elongations) or an Ising (for spins) off-site part. A coarse graining procedure yields Eq. (1.7) where D defines the surface energy (in spin systems $D>0$ is typical of antiferromagnetism, $D<0$ of ferromagnetism). The Euler–Lagrange equations based on the Lagrangian corresponding to the Hamiltonian of Eq. (1.7) lead to the NLKGE, Eq. (1.3), where we set $\epsilon = -\operatorname{sgn}(D)$; $x_0 = m^{-1/2}t$ and $(x_1,x_2,x_3) = |D|^{-1/2}(x,y,z)$. Thus both signs $\epsilon = \pm 1$ in Eq. (1.3) are relevant in condensed matter applications (without the need of continuing the equations to imaginary time). Following Landau,[11] the temperature behavior is given by $a_2 = \alpha(T - T_c)$. If $a_4 > 0$, the transition is of second order while for $a_4 < 0$ it is of first order. The condition $a_4 = 0$ corresponds to a very important case of a tricritical point on the phase diagram where a line of first-order phase transitions intersects with a line of second-order phase transitions. The values of a_4 may only be altered by external fields or by structural changes in the system. The transition temperature is obtained as T_c for $a_4 > 0$ and $T_c^* = T_c + a_4^2/4\alpha a_6$ for $a_4 < 0$, at which the order parameter experiences a discontinuity $\Delta\phi = \pm(-a_4/6a_6)^{1/2}$. Metastable states exist in the range of temperatures: $T_c < T < T_0^* = T_c + a_4^2/3\alpha a_6$ thus giving rise to a thermal hysteresis. The plot $V(\phi)$ at various temperatures is shown in Fig. 1 for $a_4 > 0$ and in Fig. 2 for $a_4 < 0$. In principle ϕ is a

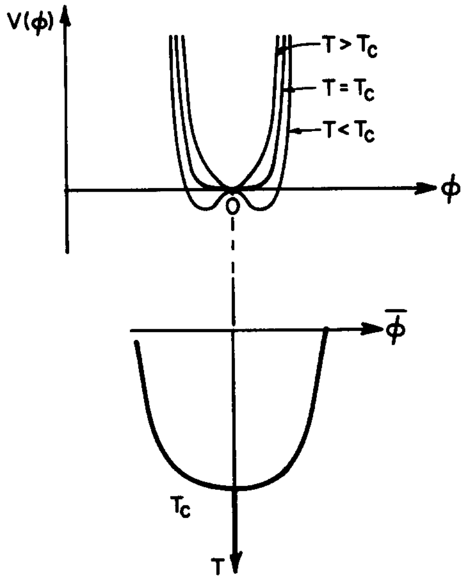


FIG. 1. The form of the potential $V(\phi)$ and the corresponding plot of $\bar{\phi}(T)$ for $a_4 > 0$.

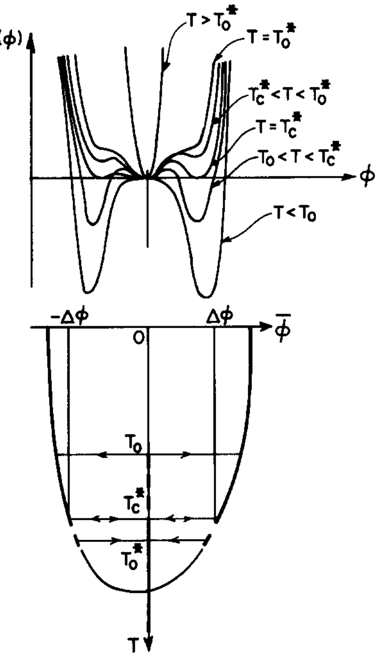


FIG. 2. The form of the potential $V(\phi)$ and the corresponding plot of $\phi(T)$ for $a_4 < 0$.

real or complex function of **x**, but in specific applications it is either one or the other, depending on its physical meaning (e.g., ϕ is real in structural phase transitions, where it represents elongation, while ϕ is complex in superconductivity where it represents the Cooper pair's wave function).

The polynomial form of $V(\phi)$ is due to the basic assumption of Landau's thermodynamic theory of phase transitions, [11] i.e., that close to criticality the thermodynamic potential is an analytic function of the order parameter and hence can be expanded in a convergent power series. For symmetry reasons (e.g., time-reversal invariance in the case of magnets) the expansion may be restricted to even powers only.

The number of terms to be kept in the expansion depends on physical considerations. Since this is a theory of symmetry changes at criticality, group theoretical considerations will be decisive. For a more general situation of an n-component order parameter a truncation criterion has been proposed,[12] requiring that sufficiently many terms be kept to allow spontaneous symmetry breaking to the smallest (generic) subgroup of the symmetry group G of the disordered phase. Another reason for truncating the series at a ϕ^p term, relevant in both field theories and condensed matter physics, is given by renormalization theory.[10] A particular ϕ^p-model is renormalizable when the dimension of physical space-time equals $n_c(p) = 2p/(p-2)$. Thus, the ϕ^6-model is renormalizable when $n_c = 3$, while the ϕ^4-model is renormalizable when $n_c = 4$. The situation in other dimensions than n_c is determined by the type of cutoffs used. In field theories ultraviolet cutoffs restrict the usefulness of the ϕ^6-model to $n \leqslant 3$ while in condensed matter infrared cutoffs imply that $n \geqslant 3$. For a given n, the terms of the series expansion of $V(\phi)$ with powers greater than p can, at most, affect critical amplitudes but not critical exponents. They are called irrelevant operators. Thus the ϕ^6-term is a small correction when $n \geqslant 4$ but not when $n = 3$.

Finally, we wish to mention that the special case of $a_2 = a_4 = 0$ in Eq. (1.7) is of particular interest, since it describes the simplest example of a multicritical point, i.e., a tricritical point, occurrences of which have been spotted in a vast array of physical systems.[13,14] For reviews of the formalism of critical phenomena and the great variety of different fields where those phenomena occur we refer to the literature.[13,15,16] In particular, Gordon[17] studies a one-dimensional model which simulates a tricritical structural phase transition. This model leads to Eq. (1.3) with $n = 1, \epsilon = -1$, and $a_4 = 0$. The actual transition takes place when $a_2 \to 0$. As mentioned above, we find numerous new solutions of the NLKGE, (1.3), for $a_2 = a_4 = 0$, i.e., Eq. (1.5). These can be used to study, e.g., perturbatively, the approach to a tricritical point which can be viewed as the "meeting point" of three different types of potential wells: single, double, and triple.

There already exists a sizable literature on exact solutions of classical ϕ^6-field theories, i.e, solutions of the NLKGE, Eq. (1.3).[17–23] They are mainly restricted to the $(1+1)$-dimensional case, or equivalently, to translation wave solutions of the form $\phi(\xi)$, where $\xi = k_\mu x^\mu$ in $(n+1)$ dimensions. An exception is Ref. 24 where a spherically symmetric static solution of Eq. (1.5) is found for $n = 3$. The literature on classical and quantum ϕ^4-field theories is quite extensive[5,25–30] and is also largely restricted to $(1+1)$ dimensions. Studies in higher dimensions, to our knowledge, have largely been either approximate or numerical ones[31,32] and have not engaged in a direct construction of exact classical solutions.

As stated above, our aim is to perform an exhaustive study of group theoretical reductions of the NLKGE, Eq. (1.3), in $(n+1)$ dimensions, to an ODE which is to be solved analytically whenever possible. Since we are interested in applications in both relativistic field theory and condensed matter physics, we shall allow solutions to be complex, but will discuss reality properties as well. Singular solutions will also be considered. A trivial comment is that a characteristic velocity c in Eq. (1.3) has been set equal to $c = 1$. In relativistic field theory c is the velocity of light in vacuum; in other applications it may represent some other characteristic velocity of the system (e.g., the speed of sound in a crystal). Hence "tachyonic" solutions corresponding to a velocity $v > 1$ (space independent in a specific frame of reference) will not necessarily be unphysical.

Section II of this paper is devoted to the actual symmetry reduction of Eqs. (1.3) and (1.5) to one of many possible ODE's. In Sec. III we analyze the reduced ODE's and determine which of them have the Painlevé property. Those that do pass the Painlevé test are brought to one of two standard forms: (P XXIX, or P XXX)[1] which can be solved exactly. In Sec. IV we present the exact solutions to Eqs. (1.3) and (1.5) and discuss some of their properties.

II. SYMMETRY REDUCTION FOR THE EQUATION OF MOTION

A. The symmetry group of the NLKGE

As stated in the Introduction, our aim is to apply the method of symmetry reduction in a systematic manner so as to obtain new analytical solutions of the NLKGE (1.3). We shall concentrate on the cases when the PDE (1.3) is reduced to an ODE, or in some cases, actually to an algebraic equation.

The process of symmetry reduction from a PDE to an ODE (or to a lower-dimensional PDE) is a standard one.[33–38] It consists of the following steps.

(1) Find the symmetry group G of the equation and its Lie algebra L. Standard algorithms exist for doing this.[33–35] Moreover computer programs, written in REDUCE,[39] MACSYMA,[40] or other symbolic languages, exist that greatly facilitate the task of determining the symmetry algebra L of a system of differential equations. If G is assumed to be a group of local point transformations, acting on the space of independent and dependent variables and transforming solutions of a scalar equation among each other, then the general element of the Lie algebra L has the form

$$X = \eta_i(\mathbf{x},\phi)\frac{\partial}{\partial x_i} + \Psi(\mathbf{x},\phi)\frac{\partial}{\partial \phi}, \qquad (2.1)$$

where $\eta_i(\mathbf{x},\phi)$ and $\Psi(\mathbf{x},\phi)$ are known (they will involve arbitrary constants and in some cases arbitrary functions[41]).

(2) Find all subgroups $G_i \subset G$ having generic orbits of codimension 1 in the space of independent variables $\{\mathbf{x}\}$ (or of codimension $1<k<n$ if we are interested in reducing to a PDE with k independent variables). This amounts to classifying the appropriate subalgebras of L into conjugacy classes under the action of the invariance group G, and choosing a representative of each class.

Methods for classifying subalgebras of Lie algebras have recently been developed and are, at least for finite-dimensional Lie algebras, not difficult to apply.[42–44]

(3) Consider each subgroup G_i (representing a conjugacy class) separately and find the invariants of its action on the space $\{\mathbf{x},\phi\}$ of independent and dependent variables. This can be done by solving a system of first-order linear partial differential equations

$$X_a H(\mathbf{x},\phi) = 0, \quad a = 1,\ldots,k, \qquad (2.2)$$

where $\{X_1,\ldots,X_k\}$ is a basis of the subalgebra L_i and each X_a is a first-order differential operator of the form (2.1). In the case when L_i is the Lie algebra of a subgroup $G_i \subset G$, having generic orbits of codimension 1 in $\{\mathbf{x}\}$, the general solution of (2.2) will be an arbitrary function of two elementary invariants $I_1(\mathbf{x},\phi)$ and $I_2(\mathbf{x},\phi)$. Setting

$$I_1(\mathbf{x},\phi) = c_1, \quad I_2(\mathbf{x},\phi) = c_2 \qquad (2.3)$$

(where c_i are constants) we obtain a symmetry variable

$$\xi = \xi(\mathbf{x}) \qquad (2.4)$$

by eliminating ϕ from (2.3) (if possible). Solving (2.3) for ϕ we obtain an expression for $\phi(\mathbf{x})$, involving a function $F(\xi)$. This function will satisfy an ODE in ξ, obtained by substituting the expression for $\phi(\mathbf{x})$ into the original PDE.

In many cases of interest, in particular in all cases arising in the present paper, a simplification occurs. The invariance group of the equation involves only transformations of the form

$$\mathbf{x}' = \Lambda_g(\mathbf{x}), \quad \phi' = \Omega_g(\mathbf{x}) + \Sigma_g(\mathbf{x})\phi. \qquad (2.5)$$

Thus, the new independent variables $\mathbf{x}'$ depend only on the old independent variables $\mathbf{x}$, the new dependent variable ϕ' depends linearly on the old one. Equivalently, we have

$$\eta_i = \eta_i(\mathbf{x}), \quad \Psi = \alpha(\mathbf{x}) + \beta(\mathbf{x})\phi \qquad (2.6)$$

in all the infinitesimal operators (2.1). In this case Eqs. (2.3) yield $\xi = \xi(\mathbf{x})$, $\phi(\mathbf{x}) = \rho(\mathbf{x})F(\xi) + \sigma(\mathbf{x})$, where $F(\xi)$ satisfies an ODE. Moreover, if $\alpha(\mathbf{x}) = 0, \beta = \text{const}$ we find $\sigma(\mathbf{x}) = 0$; if $\alpha(\mathbf{x}) = \beta(\mathbf{x}) = 0$ we also have $\rho(\mathbf{x}) = 1$. In all cases $\rho(\mathbf{x})$ and $\sigma(\mathbf{x})$, as well as $\xi(\mathbf{x})$ are known functions determined by the symmetry. In the case considered in this paper we always obtain

$$\xi = \xi(\mathbf{x}), \quad \phi(\mathbf{x}) = \rho(\mathbf{x})F(\xi). \qquad (2.7)$$

(4) Solve the ODE for $F(\xi)$, substitute into (2.7) and thus obtain particular solutions of the original PDE.

Let us now return to the NLKGE (1.3). Applying the standard algorithm (and the MACSYMA program[40]) we find that for general values of a_2, a_4, and a_6 the symmetry group G is the Poincaré group $P(n,1)$ for $\epsilon = -1$ and the Euclidean group $P(n+1,0)$ for $\epsilon = +1$. These are the isometry groups of the corresponding flat spaces and they leave invariant much more general equations,[37,38] namely,

$$H(\Box_\epsilon \phi, (\nabla\phi)^2, \phi) = 0, \qquad (2.8)$$

where H is an arbitrary sufficiently smooth function of three variables. A larger group is obtained only in the special case

$$a_2 = a_4 = 0, \quad -6a_6 \equiv a \neq 0, \qquad (2.9)$$

namely the isometry group extended by dilations, i.e., the corresponding similitude groups $\mathrm{Sim}(n,1)$ for $\epsilon = -1$, or $\mathrm{Sim}(n+1,0)$ for $\epsilon = +1$. As mentioned in the Introduction, for $n = 2$ the symmetry group is the conformal group $\mathrm{Conf}(2,1) \sim O(3,2)$, or $\mathrm{Conf}(3,0) \sim O(4,1)$, but we restrict ourselves below to $n = 3$.

A standard basis for the Lie algebra $\mathrm{sim}(n,1)$ is given by the rotations M_{ab}, Lorentz boosts M_{0a}, translations P_μ, and dilation D, realized by the differential operators

$$\begin{aligned} M_{ab} &= x_a\partial_b - x_b\partial_a, \\ M_{0a} &= -x_0\partial_a - x_a\partial_0, \quad a,b = 1,\ldots,n, \\ P_\mu &= \partial_\mu, \quad D = x_\mu\partial_\mu - \frac{\phi}{2}\frac{\partial}{\partial\phi}, \quad \mu = 0,1,\ldots,n. \end{aligned} \qquad (2.10)$$

For the Euclidean similitude algebra $\mathrm{sim}(n+1,0)$ the corresponding basis is

$$\begin{aligned} M_{\mu\nu} &= x_\mu\partial_\nu - x_\nu\partial_\mu, \quad P_\mu = \partial_\mu, \\ D &= x_\mu\partial_\mu - \frac{\phi}{2}\frac{\partial}{\partial\phi}, \qquad \mu,\nu = 0,1,\ldots,n. \end{aligned} \qquad (2.11)$$

In both cases we have put $\partial_\mu \equiv \partial/\partial_{x_\mu}$ and summation over repeated indices is understood.

We shall treat two cases separately. The first is symmetry reduction using subgroups of the corresponding isometry groups. In this case we will always have $\rho(\mathbf{x}) = 1$ in (2.7). The second involves the use of subgroups of the corresponding similitude group, not contained in the isometry group, i.e., directly involving a dilation.

A systematic study of symmetry reduction to lower-dimensional PDE's would require the knowledge of all subgroups of the symmetry group, having generic orbits of codimension k in the space of independent variables, with $1 \leqslant k \leqslant n$. We are reducing to an ODE, hence we only need $k=1$. All subgroups of $P(n,1)$ and $P(n+1,0)$ having generic orbits of codimension 1 in $M(n,1)$ and $E(n+1)$, respectively, are known, as are their invariants, for arbitrary dimension n.[37] The subgroups of similitude groups are more difficult to classify and are known only for small dimensions. For Minkowski space $M(n,1)$ the subgroups of Sim(2,1) are found in Ref. 44, and those of Sim(3,1) in Ref. 42. For Euclidean spaces only the subgroups of Sim(3,0) exist in the literature,[45] those of Sim(4,0) having orbits of codimension 1 are easy to obtain and are presented below. For the problem at hand, namely Eq. (1.5), we restrict ourselves to four-dimensional spaces (when using the corresponding similitude groups), although many of the results are valid for arbitrary n.

B. Symmetry reduction by subgroups of the isometry group

1. Euclidean space E(n+1)

Recent papers were devoted to symmetry reduction for Eq. (2.8) [containing (1.3) as a special case] in both $E(n+1)$ and $M(n,1)$.[37,38] In the case of $E(n+1)$ the only codimension 1 symmetry variables are the obvious ones, namely the radii of various spheres or cylinders,

$$\xi = r_k = (x_0^2 + x_1^2 + \cdots + x_k^2)^{1/2}, \quad 0 \leqslant k \leqslant n. \quad (2.12)$$

They are invariants of the subgroups

$$O(k+1) \otimes P(n-k,0), \quad (2.13)$$

corresponding to the Lie algebras

$$\{M_{\alpha\beta}\} \oplus \{M_{ab}, P_a\}, \quad \alpha,\beta = 0,1,\ldots,k, \quad a,b = k+1,\ldots,n. \quad (2.14)$$

Applying general results[37,38] to the special case of Eq. (1.3) we find that the substitution of (2.7) with $\rho = 1$ into (1.3) provides the following ODE for $F(\xi)$:

$$F_{\xi\xi} + (k/\xi) F_\xi = -2\lambda (a_2 F + 2a_4 F^3 + 3a_6 F^5), \quad 0 \leqslant k \leqslant n, \quad (2.15)$$

with $\xi = r_k$ as in (2.12) and $\lambda = 1$.

The properties of this equation and solutions for various values of k are discussed in Sec. III.

2. Minkowski space M(n,1)

The situation for Minkowski space $M(n,1)$ is somewhat richer. Again, all codimension 1 symmetry variables are known.[37,38] Reductions analogous to those obtained in Euclidean space are provided by the subgroups

$$O(k,1) \otimes P(n-k,0) \quad \text{and} \quad O(k+1) \otimes P(n-k-1,1) \quad (2.16)$$

corresponding to the Lie algebras

$$\{M_{\alpha\beta}\} \oplus \{M_{ab}, P_a\}, \quad \alpha,\beta = 0,1,\ldots,k, \quad a,b = k+1,\ldots,n$$

and

$$\{M_{ab}\} \oplus \{M_{\rho\sigma}, P_\rho\}, \quad a,b = 1,\ldots,k+1, \quad \rho,\sigma = 0,k+2,\ldots,n,$$

respectively. The invariants are, respectively,

$$\xi = \tau_k = (x_0^2 - x_1^2 - \cdots - x_k^2)^{1/2}, \quad 0 \leqslant k \leqslant n \quad (2.17)$$

and

$$\xi = r_k = (x_1^2 + \cdots + x_{k+1}^2)^{1/2}, \quad 0 \leqslant k \leqslant n-1. \quad (2.18)$$

Further invariants, having no analogy in Euclidean space, are given by subgroups involving lightlike translations. These invariants and corresponding Lie algebras are

$$\xi = x_0 + x_1 \quad \{M_{ab}, M_{0a} - M_{1a}, P_a, P_0 - P_1, \; a,b = 2,\ldots,n\}, \quad (2.19)$$

$$\xi = x_2 + p \ln(x_0 + x_1) \quad \{M_{ab}, M_{0a} - M_{1a}, P_a, M_{01} + pP_2, P_0 - P_1, \; p \in \mathbb{R}, \; p \neq 0, \; a,b = 3,\ldots,n\}, \quad (2.20)$$

$$\xi = x_2 + (x_0 + x_1)^2/4 \quad \{M_{ab}, M_{0a} - M_{1a}, P_a, M_{02} - M_{12} + P_0 + P_1, P_0 - P_1, \; a,b = 3,\ldots,n\}. \quad (2.21)$$

An important phenomenon occurring in spaces with indefinite metric, in particular $M(n,1)$, is that of "degenerate symmetry variables."[37,38] As opposed to ordinary codimension 1 symmetry variables, the degenerate ones involve arbitrary functions of a null variable, say $x_0 + x_n$. They are obtained from invariants of Euclidean subgroups of $P(n,1)$, by performing a Euclidean transformation, with coefficients depending on $x_0 + x_n$. Thus, if $\Phi(x) = F(\xi_k)$ with $\xi_k = (x_1^2 + \cdots + x_{k+1}^2)^{1/2}$ with $0 \leqslant k \leqslant n-2$ is a solution of Eq. (2.8), or more specifically of (1.3), then so is

$$\phi^D(\mathbf{x}) = F(\xi_k^D), \quad \xi_k^D = \left\{ \sum_{i=1}^{k+1} (\mathbf{x} - \mathbf{B}, \mathbf{A}_i)^2 \right\}^{1/2}, \quad (2.22)$$

where $\mathbf{A}_i$ and $\mathbf{B}$ are any vector functions of $x_0 + x_n$, satisfying

$$(\mathbf{A}_i, \mathbf{A}_j) = -\delta_{ij}.$$

Moreover, $F(\xi_k)$ and $F(\xi_k^D)$ satisfy the same ODE. In $M(n,1)$ there will be $n-1$ different types of degenerate symmetry variables, reducing the considered PDE to an ODE. Thus in $M(2,1)$ any degenerate symmetry variable is conjugate under $P(2,1)$ to

$$\xi_1 = x_1 + f(x_0 + x_2). \quad (2.23)$$

In $M(3,1)$ two types exist, represented by

$$\xi_1 = x_1 \cos\theta + x_2 \sin\theta + f, \quad (2.24)$$

$$\xi_2 = [(x_1 + f_1)^2 + (x_2 + f_2)^2]^{1/2}. \quad (2.25)$$

In (2.24) and (2.25), f, θ, f_1, and f_2 are arbitrary twice differentiable functions of $(x_0 + x_3)$.

Ordinary (nondegenerate) codimension 1 symmetry variables may be special cases of degenerate ones. This occurs when the corresponding symmetry algebra contains $P_0 - P_n$ as an ideal. Thus, the variables (2.20) and (2.21) in $M(3,1)$ are equivalent to (2.24) with $\theta = 0$ and $f = \mu \ln(x_0 + x_1)$, or $\theta = 0$, $f = (x_0 + x_1)^2/4$, respectively.

For $\xi = x_0 + x_1$ the substitution (2.7) with $\rho = 1$ reduces the NLKGE (1.3) to the algebraic equation

$$a_2 F + 2a_4 F^3 + 3a_6 F^5 = 0, \quad (2.26)$$

having constant solutions. In all other cases we obtain the

ODE (2.15) with $\lambda = \pm 1$, depending on the choice of ξ. The values of λ and k for each variable ξ are indicated in Table I for both $E(n+1)$ and $M(n,1)$. Symmetry variables that are special cases of degenerate variables are listed under the same number as the degenerate ones. Thus No. 3a and No. 3b are special cases of No. 3. The subalgebra of the isometry algebra leading to a specific nondegenerate symmetry variable is indicated in column 2 of Table I.

The symmetry variables ξ introduced in $E(n+1)$ and $M(n,1)$, and summarized in Table I, as well as the corresponding Lie algebras, are representatives of conjugacy classes. Applying the corresponding groups $P(n+1,0)$ or $P(n,1)$ to the representative variables, we obtain the general form of the variable, without changing the form of the reduced equation. Thus, e.g., $x_1 + f(x_0 + x_n)$ in $M(n,1)$ represents the class of variables $(\mathbf{A},\mathbf{x}) + f[(\mathbf{B},\mathbf{x})]$, where $\mathbf{A}$ and $\mathbf{B}$ are constant vectors satisfying $\mathbf{A}^2 = -1$, $\mathbf{B}^2 = 0$, $(\mathbf{A},\mathbf{B}) = 0$.

C. Symmetry reduction by subgroups involving dilations

We now restrict ourselves to the case when $a_2 = a_4 = 0$ in (1.3) and put $-6a_6 \equiv a$. The equation we study is hence the one given in (1.5) and its symmetry group is $\mathrm{Sim}(n+1,0)$ for $\epsilon = 1$ and $\mathrm{Sim}(n,1)$ for $\epsilon = -1$. For practical reasons we restrict ourselves to $n = 3$ (the subgroup classification is not available for $n > 3$).

1. Euclidean space E(4)

We simplify notation with respect to the general case (2.11) and use the following basis for sim(4):

$$\begin{aligned} &L_1 = -x_2\partial_3 + x_3\partial_2, \quad K_1 = x_1\partial_0 - x_0\partial_1, \\ &L_2 = -x_3\partial_1 + x_1\partial_3, \quad K_2 = x_2\partial_0 - x_0\partial_2, \\ &L_3 = -x_1\partial_2 + x_2\partial_1, \quad K_3 = x_3\partial_0 - x_0\partial_3, \\ &P_\mu = \partial_\mu, \quad D = x_1\partial_1 + x_2\partial_2 + x_3\partial_3 + x_0\partial_0 - \tfrac{1}{2}\phi\partial_\phi, \\ &\qquad \mu = 0,1,2,3. \end{aligned} \tag{2.27}$$

We are interested in subgroups of $\mathrm{Sim}(4,0)$ having generic orbits of codimension 1 in $E(4)$. The corresponding subalgebras must be of dimension $d \geqslant 3$. Subalgebras of $p(4,0)$ have already been considered, hence we restrict ourselves to subalgebras containing the dilation operator D in some form. Using methods developed elsewhere[42–45] it is easy to show that only the five types of subalgebras of $\mathrm{sim}(4,0)$, listed in column 2 of Table II (No. 2–6), lead to codimension 1 symmetry variables.

Consider first the algebras Nos. 2 and 3 of Table II: $\{\tilde{D} = D + bL_3, K_3, P_3, P_0\}$, $b \geqslant 0$. To find the invariants (2.3) in this case we must solve the corresponding equations (2.2), i.e.,

$$\begin{aligned} &P_0H(\mathbf{x},\phi) = 0, \quad P_3H(\mathbf{x},\phi) = 0, \\ &K_3H(\mathbf{x},\phi) = 0, \quad \tilde{D}H(\mathbf{x},\phi) = 0. \end{aligned} \tag{2.28}$$

The first two equations imply $H = H(x_1,x_2,\phi)$. The third is then satisfied identically. The last equation in (2.28) is

$$\tilde{D}H = \left[(x_1 + bx_2)\frac{\partial}{\partial x_1} + (x_2 - bx_1)\frac{\partial}{\partial x_2} + x_3\frac{\partial}{\partial x_3} + x_0\frac{\partial}{\partial x_0} - \frac{1}{2}\phi\frac{\partial}{\partial \phi} \right] H(x_1,x_2,\phi) = 0. \tag{2.29}$$

The characteristic system for (2.29) is

$$\frac{dx_1}{x_1 + bx_2} = \frac{dx_2}{x_2 - bx_1} = -2\frac{d\phi}{\phi}. \tag{2.30}$$

TABLE I. Reduction of the equation $\Box_\epsilon \phi = f(\phi)$ to the ODE $\ddot{F} + (k/\xi)\dot{F} = \lambda f(F)$ in $E(n+1)$ ($\epsilon = +1$) and $M(n,1)$ ($\epsilon = -1$) by means of subgroups of the isometry groups; $\phi(x) = F(\xi)$; $\mathbf{A}_i$ and $\mathbf{B}$ are arbitrary vector functions of $x_0 + x_n$.

		$E(n+1)$		
No.	Algebra	ξ	k	λ
1	$M_{\mu\nu} \oplus \{M_{ab},P_a\}$ $\mu,\nu = 0,1,\dots,k;\ a,b = k+1,\dots,n$	$(x_0^2 + x_1^2 + \cdots + x_k^2)^{1/2}$	$0 \leqslant k \leqslant n$	1
		$M(n,1)$		
1	$M_{\mu\nu} \oplus \{M_{ab},P_a\}$ $\mu,\nu = 0,1,\dots,k;\ a,b = k+1,\dots,n$	$(x_0^2 - x_1^2 - \cdots - x_k^2)^{1/2}$	$0 \leqslant k \leqslant n$	1
2	$M_{ij} \oplus \{M_{\alpha\beta},P_\alpha\}$ $i,j = 1,\dots,k+1;\ \alpha,\beta = 0,k+2,\dots,n$	$(x_1^2 + x_2^2 + \cdots + x_{k+1}^2)^{1/2}$	$0 \leqslant k \leqslant n-1$	-1
3	Degenerate variables	$\left\{\sum_{i=1}^{k+1} (\mathbf{A}_i,\mathbf{x} - \mathbf{B})^2\right\}^{1/2}$ $(\mathbf{A}_i,\mathbf{A}_j) = -\delta_{ij},\quad 1 \leqslant i,j \leqslant k+1$	$0 \leqslant k \leqslant n-2$	-1
3a	$M_{ab},\ M_{0a} - M_{na},\ P_a,$ $M_{0n} + pP_1,\ P_0 - P_n;\ a,b = 2,\dots,n-1$	$x_1 + p\ln(x_0 + x_n)$	0	-1
3b	$M_{ab},\ M_{0a} - M_{na},\ P_a,$ $M_{01} + M_{n1} + P_0 + P_n,\ P_0 - P_n;$ $a,b = 2,\dots,n-1$	$x_1 + \frac{1}{4}(x_0 + x_n)^2$	0	-1

TABLE II. Reduction of the equation $\Box_\epsilon \phi = a\phi^5$ ($\epsilon = 1$) in E(4) to an ODE; $\phi(\mathbf{x}) = \rho F(\xi)$.

No.	Algebra	ρ	ξ	ODE
1	$M_{\mu\nu} \oplus \{M_{ab}, P_a\}$ $\mu,\nu = 0,1,\ldots,k$ $a,b = k+1,\ldots,n$	$(-a)^{-1/4}$	$(x_0^2 + x_1^2 + \cdots + x_k^2)^{1/2}$, $0 \leqslant k \leqslant n$	$\ddot F + (k/\xi)\dot F + F^5 = 0$
2	D, K_3, P_0, P_3	$[-4a(x_1^2 + x_2^2)]^{-1/4}$	$\frac{1}{2}\arctan(x_2/x_1)$	$\ddot F + F + F^5 = 0$
3	$D + bL_3, K_3, P_0, P_3$ $b > 0$	$[-a[(b^2+1)/b^2]$ $\times (x_1^2 + x_2^2)]^{-1/4}$	$-[b/(b^2+1)][\arctan(x_2/x_1)$ $+ (b/2)\ln(x_1^2 + x_2^2)]$	$\ddot F + \dot F + [(b^2+1)/4b^2]F + F^5 = 0$
4	D, L_3, P_0	$[-ax_3^2/4]^{-1/4}$	$(x_1^2 + x_2^2)/x_3^2$	$\xi(1+\xi)\ddot F + (2\xi + 1)\dot F + \frac{3}{16}F + F^5 = 0$
5	D, L_3, K_3	$[-a(x_0^2 + x_3^2)/4]^{-1/4}$	$[(x_1^2 + x_2^2)/(x_0^2 + x_3^2)]$	$\xi(1+\xi)\ddot F + (\frac{3}{2}\xi + 1)\dot F + \frac{1}{16}F + F^5 = 0$
6	D, L_1, L_2, L_3	$[-ax_0^2/4]^{-1/4}$	$(x_1^2 + x_2^2 + x_3^2)/x_0^2$	$\xi(1+\xi)\ddot F + (2\xi + \frac{3}{2})\dot F + \frac{3}{16}F + F^5 = 0$

Solving (2.30) we obtain

$$\Phi = A(x_1^2 + x_2^2)^{-1/4}F(\xi), \qquad (2.31)$$
$$\xi = B[\arctan(x_2/x_1) + (b/2)\ln(x_1^2 + x_2^2)],$$

where A and B are some normalization constants. Substituting (2.31) into (1.5) with $\epsilon = 1$, we obtain the corresponding ODE. For $b = 0$ we choose $A = (-4a)^{-1/4}$, $B = \frac{1}{2}$ and obtain the ODE in row 2 of Table II. For $b > 0$ we choose $A = [-a(b^2+1)/b^2]^{-1/4}$, $B = -b(b^2+1)$ and obtain the equation of row 3 (the case $b < 0$ is equivalent to $b > 0$).

The remaining three cases are completely analogous; the calculations are somewhat simpler. The results are all presented in Table II which also includes, in row 1, the reductions obtained by subgroups of the Euclidean group (for n arbitrary).

2. Minkowski space M(3.1)

We proceed in the same manner as for E(4), but the results are much richer. We use a basis analogous to (2.27), namely

$$L_1 = -x_2\partial_3 + x_3\partial_2, \quad K_1 = -(x_1\partial_0 + x_0\partial_1),$$
$$L_2 = -x_3\partial_1 + x_1\partial_3, \quad K_2 = -(x_2\partial_0 + x_0\partial_2),$$
$$L_3 = -x_1\partial_2 + x_2\partial_1, \quad K_3 = -(x_3\partial_0 + x_0\partial_3), \qquad (2.32)$$
$$P_\mu = \partial_\mu,$$
$$D = x_1\partial_1 + x_2\partial_2 + x_3\partial_3 + x_0\partial_0 - \frac{1}{2}\phi(\partial/\partial_\phi).$$

The subalgebras of sim(3,1) are known,[42] but we are interested in those that are not contained in the Poincaré algebra p(3.1). We run through all such subalgebras corresponding to groups with generic orbits of codimension 1 in M(3,1) and find the two invariants (2.3) in each case. The result can always be written in the form (2.7) with $\rho(\mathbf{x})$ and $\xi(\mathbf{x})$ known. Substituting into (1.5) with $\epsilon = -1$ we obtain one of the following equations for $F(\xi)$: (i) an algebraic equation yielding F directly, (ii) a first-order ODE that we can solve, or (iii) a second-order ODE in one of several possible forms. Whenever this equation can be transformed into an equation with the Painlevé property[1–5] we can solve it in terms of elementary or elliptic functions.

The calculations involved are all quite straightforward,

TABLE III. Reduction of the equation $\Box_\epsilon \phi = a\phi^5$ ($\epsilon = -1$) in M(3,1) to an algebraic equation or a first-order ODE; $\phi(x) = \rho F(\xi)$.

No.	Algebra	ρ	ξ	ODE	Solution $F(\xi)$
1	$D + K_3, L_1 - K_2,$ $P_2, P_0 - P_3$	$x_1^{-1/2}$	$x_0 + x_3$	$\cdots$	$(-3/4a)^{1/4}$
2	$D + K_3, L_2 + K_3 + \epsilon P_2,$ $P_0 - P_3$	$[x_1 + \epsilon x_2(x_0 + x_3)]^{-1/2}$	$x_0 + x_3$	$\cdots$	$(-3/4a)^{1/4}(\xi^2 + 1)^{1/4}$
3	$D + K_3, L_3, P_0 - P_3$	$(x_1^2 + x_2^2)^{-1/4}$	$x_0 + x_3$	$\cdots$	$(-1/4a)^{1/4}$
4	$L_3, L_1 - K_2, L_2 + K_1,$ $P_1, P_2, P_0 - P_3$	1	$x_0 + x_3$	$\cdots$	0
5	$D + K_3, L_3, P_1, P_2$	$(x_0 - x_3)^{-1/4}$	$x_0 + x_3$	$\dot F + aF^5 = 0$	$[4a(\xi - \xi_0)]^{-1/4}$
6	$D + K_3, L_2 + K_1, P_2$	$(x_1^2 + x_3^2 - x_0^2)^{-1/4}$	$x_0 + x_3$	$4\xi\dot F + F - 4aF^5 = 0$	$[-(4a/\xi_0)(\xi - \xi_0)]^{-1/4}$
7	$D + K_3, L_3, L_1 - K_2,$ $L_2 + K_1$	$(x_1^2 + x_2^2 + x_3^2 - x_0^2)^{-1/4}$	$x_0 + x_3$	$4\xi\dot F + 3F - 4aF^5 = 0$	$[(-4a/3\xi_0^3)(\xi^3 - \xi_0^3)]^{-1/4}$
8	$D + K_3, L_1 - K_2 + P_2,$ $L_2 + K_1 + P_1$	$[x_1^2 + x_2^2 + x_3^2 - x_0^2 - x_3$ $+ x_0 - 2x_2^2/(1 + x_0 + x_3)]^{-1/4}$	$x_0 + x_3$	$4(\xi^2 - 1)\dot F + (3\xi - 1)F$ $- 4a(\xi + 1)F^5 = 0$	$[2a\xi(\xi + 1) - a(\xi + 1)(\xi^2 - 1)$ $\times \ln[(1 + \xi)/(1 - \xi)]$ $+ c(\xi^2 - 1)(\xi + 1)]^{-1/4}$
9	$D - K_3, L_2 + K_1, P_2$	$(x_0 + x_3)^{-1/4}$	$((x_1^2 + x_3^2 - x_0^2)/(x_0 + x_3))^{1/2}$	$\dot F + 2a\xi F^5 = 0$	$[4a(\xi^2 - \xi_0^2)]^{-1/4}$
10	$D - K_3, L_3, L_1 - K_2,$ $L_2 + K_1$	$(x_0 + x_3)^{-1/4}$	$((x_1^2 + x_2^2 + x_3^2 - x_0^2)/(x_0 + x_3))^{1/2}$	$3\dot F + 2a\xi F^5 = 0$	$[4a(\xi^2 - \xi_0^2)/3]^{-1/4}$

TABLE IV. Reduction of the equation $\square_\epsilon \phi = a\phi^5$ ($\epsilon = -1$) in M(3,1) to a second-order ODE; $\phi(x) = \rho F(\xi)$, h, θ, f, f_1, and f_2 are arbitrary functions of $x_0 + x_3$, b is a constant, $\epsilon_a = \pm 1$ ($a = 0,1,2$).

No.	Algebra	ρ	ξ	ODE
1	$L_1, L_2, L_3, P_1, P_2, P_3$	$(-a)^{-1/4}$	x_0	$\ddot F + F^5 = 0$
2	$L_1, K_2, K_3, P_2, P_3, P_0$	$(a)^{-1/4}$	x_1	
3	Degenerate variable	$(a)^{-1/4}h^{1/2}$	$h(x_1 \cos\theta + x_2 \sin\theta + f)$	
3a	$L_1 - K_2, K_3 + bP_1, P_2, P_0 - P_3$	$h = 1$	$\theta = 0, f = b \ln(x_0 + x_3)$	
3b	$L_1 - K_2, L_2 + K_1 + P_0 + P_3,$ $P_2, P_0 - P_3$	$h = 1$	$\theta = 0, f = \frac{1}{4}(x_0 + x_3)^2$	
3c	$D + K_3 + \epsilon_0(P_0 + P_3),$ $P_2, P_0 - P_3$	$h = e^{-\epsilon_0(x_0+x_3)/2}$	$\theta = 0, f = 0$	
3d	$D + b(L_2 + K_1), L_1 - K_2,$ $P_2, P_0 - P_3$	$h = (x_0 + x_3)^{-1/2}$	$\theta = 0, f = b(x_0 + x_3)\ln(x_0 + x_3)$	
3e	$D + bK_3, L_1 - K_2, P_2,$ $P_0 - P_3, b \neq 1$	$h = (x_0 + x_3)^{1/(b-1)}$	$\theta = 0, f = 0$	
4	L_3, K_3, P_1, P_2	$(-a)^{-1/4}$	$[x_0^2 - x_1^2]^{1/2}$	$\ddot F + (1/\xi)\dot F + F^5 = 0$
5	L_3, K_3, P_0, P_3	$(a)^{-1/4}$	$(x_1^2 + x_2^2)^{1/2}$	
6	Degenerate variable	$(a)^{-1/4}h^{1/2}$	$h[(x_1 + f_1)^2 + (x_2 + f_2)^2]^{1/2}$	
6a	$D + K_3 + \epsilon_0(P_0 + P_3), L_3, P_0 - P_3$	$h = e^{-\epsilon_0(x_0+x_3)/2}$	$f_1 = f_2 = 0$	
7	K_1, K_2, L_3, P_3	$(-a)^{-1/4}$	$(x_0^2 - x_1^2 - x_2^2)^{1/2}$	$\ddot F + (2/\xi)\dot F + F^5 = 0$
8	L_1, L_2, L_3, P_0	$(a)^{-1/4}$	$(x_1^2 + x_2^2 + x_3^2)^{1/2}$	
9	$L_1, L_2, L_3, K_1, K_2, K_3$	$(-a)^{-1/4}$	$(x_0^2 - x_1^2 - x_2^2 - x_3^2)^{1/2}$	$\ddot F + (3/\xi)\dot F + F^5 = 0$
10	D, K_3, P_0, P_3	$[4a(x_1^2 + x_2^2)]^{-1/4}$	$\frac{1}{2}\arctan(x_2/x_1)$	$\ddot F + F + F^5 = 0$
11	Degenerate variable	$\{4a[(x_1 + f_1)^2 + (x_2 + f_2)^2]\}^{-1/4}$	$\frac{1}{2}\{\arctan[(x_2 + f_2)/(x_1 + f_1)] - \theta\}$	
11a	$D + K_3, L_3 + \epsilon_0(P_0 + P_3), P_0 - P_3$	$f_1 = f_2 = 0$	$\theta = \epsilon_0(x_0 + x_3)/2$	
11b	$D + \frac{1}{2}K_3, L_1 - K_2 + P_0 + P_3, P_0 - P_3$	$f_1 = 0, f_2 = \frac{1}{2}(x_0 + x_3)^2$	$\theta = 0$	
12	$D + bL_3, K_3, P_0, P_3, b > 0$	$[[a(b^2+1)/b^2](x_1^2 + x_2^2)]^{-1/4}$	$-[b/(b^2+1)][\arctan(x_2/x_1)$ $+ (b/2)\ln(x_1^2 + x_2^2)]$	$\ddot F + \dot F + [(b^2+1)/4b^2]F$ $+ F^5 = 0$
13	Degenerate variable	$\{[a(b^2+1)/b^2][(x_1 + f_1)^2$ $+ (x_2 + f_2)^2]\}^{-1/4}$	$-[b/(b^2+1)]\{\arctan[(x_2 + f_2)/(x_1 + f_1)]$ $+ (b/2)\ln[(x_1 + f_1)^2 + (x_2 + f_2)^2] - \theta\}$	
13a	$D + K_3 + \epsilon_1(P_0 + P_3),$ $L_3 + \epsilon_2(P_0 + P_3), P_0 - P_3$	$f_1 = f_2 = 0$	$b = -\epsilon_1\epsilon_2, \theta = -\epsilon_2(x_0 + x_3)/2$	
13b	$D + pK_3, L_3 + qK_3, P_0 - P_3,$ $p \neq 1, q \neq 0$	$f_1 = f_2 = 0$	$q = b/(b^2+1), p = 1 - [b^3/2(b^2+1)^2]$ $\theta = (1/q)\ln(x_0 + x_3)$	
14	D, L_3, P_0	$(ax_3^2/4)^{-1/4}$	$(x_1^2 + x_2^2)/x_3^2$	$\xi(1+\xi)\ddot F + (2\xi + 1)\dot F$ $+ \frac{3}{16}F + F^5 = 0$
15	D, L_3, P_3	$(-ax_0^2/4)^{-1/4}$	$-(x_1^2 + x_2^2)/x_0^2$	
16	D, K_1, P_3	$(ax_2^2/4)^{-1/4}$	$(x_1^2 - x_0^2)/x_2^2$	
17	D, L_3, K_3	$(a(x_3^2 - x_0^2)/4)^{-1/4}$	$(x_1^2 + x_2^2)/(x_3^2 - x_0^2)$	$\xi(1+\xi)\ddot F + (\frac{3}{2}\xi + 1)\dot F$ $+ \frac{1}{16}F + F^5 = 0$
18	D, L_1, L_2, L_3	$(-ax_0^2/4)^{-1/4}$	$-(x_1^2 + x_2^2 + x_3^2)/x_0^2$	$\xi(1+\xi)\ddot F + (2\xi + \frac{3}{2})\dot F$ $+ \frac{3}{16}F + F^5 = 0$
19	D, K_1, K_2, L_3	$(ax_3^2/4)^{-1/4}$	$(x_1^2 + x_2^2 - x_0^2)/x_3^2$	
20		$(-(2q+1)/a)^{1/4}(x_0 + x_1)^{q/2},$ $q \neq -\frac{1}{2}$	$(x_0^2 - x_1^2 - \cdots - x_k^2)(x_0 + x_1)^q$	$\ddot F + (3q + k)/(2q+1)(1/\xi)\dot F$ $+ F^5 = 0$
20a	$D + [(1+q)/q]K_1, L_1, P_2, P_3$		$k = 1$	
20b	$D + [(1+q)/q]K_1, L_3 + K_2, P_3$		$k = 2$	
20c	$D + [(1+q)/q]K_1, L_1, L_2 - K_3,$ $L_3 + K_2$		$k = 3$	
21		$[(-\epsilon/4a)[(2k-3)^2/(x_0 + x_1)]]^{1/4}$	$[(2k-3)/4][\ln(x_0 + x_1)$ $+ \epsilon[(x_0^2 - x_1^2 - \cdots - x_k^2)/(x_0 + x_1)]]$	$\ddot F + \dot F + F^5 = 0$
21a	$D - K_1 + bL_1 + \epsilon(P_0 - P_1), P_2, P_3$	$k = 1$		
21b	$D - K_1 + \epsilon(P_0 - P_1), L_3 + K_2, P_3$	$k = 2$		
21c	$D - K_1 + bL_1 + \epsilon(P_0 - P_1),$	$k = 3$		
22	$D + \frac{1}{2}K_3, L_1 - K_2 + P_0 + P_3, P_1$	$(9/4a)^{1/4}[x_2 - [(x_0 + x_3)^2/4]]^{-1/2}$	$[6(x_3 - x_0) + 6x_2(x_0 + x_3) - (x_0 + x_3)^3]$ $\times\{8[x_2 - \frac{1}{4}(x_0 + x_3)^2]^{3/2}\}^{-1}$	$(1 + \xi^2)\ddot F + \frac{7}{3}\xi\dot F + \frac{1}{3}F + F^5 = 0$

though long and cumbersome. Many of them, in particular the derivation of the ODE's from the PDE (1.5), once $\rho(\mathbf{x})$ and $\xi(\mathbf{x})$ are found, were performed on a computer using a MACSYMA program, written for this purpose.

The first two cases are summarized in Table III. Subalgebras Nos. 1,...,4 all lead to algebraic equations; the first three of them give nontrivial solutions. Subalgebras Nos. 5,...,10 of Table III lead to first-order ODE's given in column 5. The solutions are in the last column.

The second-order ODE's that we obtain are all presented in Table IV. This table is organized in the same manner as Table II for the Euclidean case E(4) and indeed the equations corresponding to cases Nos. 1,...,19 coincide with equations in Table II. The last three equations (reductions Nos. 20–22) are specific for M(3,1).

The phenomenon of degenerate symmetry variables, already discussed in connection with Table I, reoccurs here in a richer form (because of the additional dilation invariance). Dropping all details and restricting ourselves to M(3,1) [rather than M(n,1)] the effect of degenerate symmetry variables amounts to the following. Let $\phi_0(x_1,x_2)$ be a solution of Eq. (1.5). Then

$$\begin{aligned} \phi(\mathbf{x}) &= h^{1/2}\phi_0(\xi_1,\xi_2), \\ \xi_1 &= h\,[x_1\cos\theta + x_2\sin\theta + f_1], \\ \xi_2 &= h\,[-x_1\sin\theta + x_2\cos\theta + f_2], \end{aligned} \tag{2.33}$$

where h,f_1,f_2, and θ are arbitrary functions of $x_0 + x_3$, is also a solution. Moreover, if $\phi_0(x_1,x_2)$ has the form

$$\phi_0(x_1,x_2) = \rho(x_1,x_2)F_0(\xi(x_1,x_2)),$$

where F_0 satisfies an ODE, then ϕ has the form

$$\phi(\mathbf{x}) = h^{1/2}\rho(\xi_1,\xi_2)F(\xi(\xi_1,\xi_2)),$$

where F satisfies the same ODE as F_0.

As in the case of reductions by subgroups of the isometry group P(3,1), it often happens that codimension 1 symmetry variables are special cases of degenerate ones. The subgroup providing the reduction then contains $P_0 - P_3$ as an ideal (without containing P_0 and P_3 separately).

In Table IV we present all the codimension 1 symmetry variables and also the corresponding degenerate ones. Special cases of the degenerate variables are given the same number as the degenerate one; thus 3a–3e are all special cases of No. 3. Reductions Nos. 1–9 all come from subgroups of P(3,1), Nos. 10–22 involve dilations.

Comparing the Euclidean and Minkowski cases, we see the following differences.

(1) A single reduction in E(4) often corresponds to several reductions in M(3,1), because of the difference between spacelike and timelike variables.

(2) Degenerate variables occur in M(3,1) (but not in spaces with a definite metric).

(3) New types of reductions occur in M(3,1), leading to algebraic equations, first-order ODE's and new types of second-order ODE's.

The ODE's listed in Table IV (and Table II) will be analyzed in Sec. III below and solved analytically whenever possible. They can all be cast into one generic form, namely

$$(\alpha_1 + \alpha_2\xi + \alpha_3\xi^2)F_{\xi\xi} + (\beta_1 + \beta_2\xi)F_\xi + \gamma F + (\mu_1 + \mu_2\xi)F^5 = 0, \tag{2.34}$$

where the constants α_i,β_i,γ, and μ_i can be read off from Table II and Table IV for E(4) and M(3,1), respectively.

III. ANALYSIS OF THE REDUCED EQUATIONS

A. General comments and Painlevé analysis

The purpose of this section is to obtain explicit solutions of the reduced equations of Sec. II. All algebraic equations and first-order differential equations that occurred have already been solved (the results are in Table III). We are interested in solving the second-order differential equation (2.15) with λ and k as in Table I and the equations presented in Tables II and IV. All equations occurring in E(4) (Table II) also occur in M(3,1) (Table IV), so we concentrate on the latter case.

We shall treat all the obtained second-order ODE's for the function $F(\xi)$ in a uniform manner.

(1) We subject the considered ODE for $F(\xi)$ to the "Painlevé test" [2] in order to determine whether it, or an equation for $H(\xi) = [F(\xi)]^q$, where q is a positive integer, satisfies certain necessary conditions for having the Painlevé property.

(2) If the result of this test is positive and if the equation (for H) does actually have the Painlevé property (rather than only satisfying the necessary conditions) then we can, by a transformation of the dependent and independent variables, transform it into its "standard form." The standard form for a second-order ODE, linear in the second derivative H'' rational in H' and H and analytical in ξ, belongs to one of the 50 classes, established by Painlevé and Gambier[4] and listed, e.g., in Ince.[1] The solutions of all the standard equations are known.[1] They are given in terms of solutions of linear equations, in terms of elementary transcendental functions, or elliptic functions, or in terms of one of the six Painlevé transcendents. It turns out that all the ODE's obtained in Sec. II that pass the Painlevé test can be solved in terms of elliptic functions or elementary ones. Equations that do not have the Painlevé property will in general have moving logarithmic singularities (or worse) and we have no systematic method for integrating them.

We recall here that an ODE has the Painlevé property if its general solution has no moving critical points, i.e., no singularities, other than poles, the position or character of which depends on the initial conditions.

The Painlevé test for ODE's is quite elementary.[2] The general solution of the considered ODE is represented as a series

$$F(\xi) = \sum_{j=0}^{\infty} f_j\tau^{j+\alpha}, \quad \tau \equiv \xi - \xi_0, \tag{3.1}$$

where ξ_0 is an arbitrary constant and f_j are constants to be determined. The series is substituted into the equation and coefficients of independent powers in τ are compared. For the equation to have the Painlevé property, the following conditions are necessary: (i) the number α is a negative integer. (ii) A recursion relation is obtained for f_j that has the form

$$P(j)f_j=\Phi(f_0,f_1,\dots,f_{j-1},\xi_0), \tag{3.2}$$

where $P(j)$ is a polynomial in j with $m-1$ positive integer roots (m is the order of the equation; in our case $m=2$). These values of j are called "resonance values" and f_j at a resonance cannot be determined. (iii) At each resonance the "resonance condition"

$$\Phi(f_0,f_1,\dots,f_{j-1},\xi_0)=0 \tag{3.3}$$

must be satisfied for all values of ξ_0. The solution (3.1) then involves n arbitrary constants, namely ξ_0 and the $n-1$ resonance values f_j.

If α turns out to be a negative rational number, $\alpha=-p/q$ (where p and q are mutually prime positive integers) then we set $H(\xi)=[F(\xi)]^q$ and perform the Painlevé test for $H(\xi)$.

The procedure is entirely algorithmic and we perform it using a MACSYMA program.[3] In all cases the expansion for the function $F(\xi)$ satisfying one of the second-order equations of Sec. II leads to the value $\alpha=-\frac{1}{2}$. Hence we always first make the substitution

$$F(\xi)=[H(\xi)]^{1/2}. \tag{3.4}$$

Equation (2.15) is reduced to

$$H_{\xi\xi}=(1/2H)H_\xi^2-(k/\xi)H_\xi - 4\lambda(a_2H+2a_4H^2+3a_6H^3). \tag{3.5}$$

Applying the Painlevé test to this equation we find $\alpha=-1$ and a resonance occurs at $j=3$. The resonance condition (3.3) is satisfied precisely in the following cases: (i) $k=0$, (ii) $k=2$, $a_2=a_4=0$, (iii) $k=3$, $a_2=a_6=0$ (however, we are only interested in the case $a_6\neq 0$).

Let us now concentrate on the equations obtained from subalgebras of sim(4,0) or sim(3,1), involving dilations.

All of these equations can be written in the form (2.34), where the values of the constants can be read off from entries Nos. 10–22 of Table IV (and also Nos. 2–6 of Table II). The Painlevé test in all cases provides the value $\alpha=-\frac{1}{2}$ in (3.1); the relevant tranformation is hence (3.4). Among the considered equations the only ones that pass the Painlevé test are

$$\ddot F+F+F^5=0, \tag{3.6}$$

$$\xi(1+\xi)\ddot F+(2\xi+\tfrac{3}{2})\dot F+\tfrac{3}{16}F+F^5=0, \tag{3.7}$$

$$\ddot F+[(3q+k)/(2q+1)](1/\xi)\dot F+F^5=0 \quad \text{for } q=-k/3,\ k-2, \text{ and } 4-3k, \tag{3.8}$$

$$(1+\xi^2)\ddot F+\tfrac{7}{3}\xi\dot F+\tfrac{1}{3}F+F^5=0. \tag{3.9}$$

The transformation (3.4) takes these equations into

$$\ddot H=\dot H^2/2H-2(H+H^3), \tag{3.6'}$$

$$\ddot H=\frac{\dot H^2}{2H}-\frac{1}{\xi(1+\xi)}\left[\left(2\xi+\frac{3}{2}\right)\dot H+\frac{3}{8}H+2H^3\right], \tag{3.7'}$$

$$\ddot H=\frac{\dot H^2}{2H}-\left[\frac{3q+k}{2q+1}\frac{1}{\xi}\dot H+2H^3\right], \tag{3.8'}$$

$$\ddot H=\frac{\dot H^2}{2H}-\frac{1}{1+\xi^2}\left(\frac{7}{3}\xi\dot H+\frac{2}{3}H+2H^3\right), \tag{3.9'}$$

respectively.

B. Reduction of the Painlevé type equations to their standard forms

Let us first consider the nonlinear Klein–Gordon equation (1.3) for a_2, a_4, and a_6 arbitrary. Reductions are possible by subgroups of the isometry group only. They all lead to Eq. (2.15) which is of Painlevé type only for $k=0$. In this case we can integrate immediately (first multiplying by F_ξ) and obtain

$$F_\xi^2=-2\lambda(a_0+a_2F^2+a_4F^4+a_6F^6), \tag{3.10}$$

where a_0 is an integration constant. The equation is thus reduced to quadratures; we shall discuss solutions in terms of elliptic and elementary functions below. Notice that Eq. (3.10) is obtained for arbitrary values of n and both for Euclidean and Minkowski spaces.

Other Painlevé type equations are obtained only for $a_2=a_4=0$ ($a_6\neq 0$ by assumption). For arbitrary $n\geqslant 2$ and both for E$(n+1)$ and M$(n,1)$ we have Eq. (2.15) with $k=2$, i.e.,

$$F_{\xi\xi}+(2/\xi)F_\xi=\lambda aF^5, \quad -6a_6\equiv a. \tag{3.11}$$

The transformation (3.4) takes (3.11) into

$$H_{\xi\xi}=(1/2H)H_\xi^2-(2/\xi)H_\xi+2\lambda aH^3, \tag{3.12}$$

a special case of (3.8'). The equations that remain to be integrated are (3.6)–(3.9). We write them all in a unified manner as

$$H_{\xi\xi}=(1/2H)H_\xi^2+RH_\xi+SH+TH^3. \tag{3.13}$$

Comparing with the Painlevé list of 50 canonical equations, reproduced by Ince,[1] we see that (3.13) is of "canonical type III." Hence, whenever it has the Painlevé property, it can, by a transformation of the type[1]

$$H(\xi)=\lambda(\xi)W(\eta), \quad \eta=\eta(\xi) \tag{3.14}$$

be reduced to one of the equations P XVII–P XXXV. Moreover, the specific "candidates" are the Painlevé equations P XXIX, or P XXX. Putting (3.14) into (3.13) we obtain

$$\ddot W=\frac{1}{2W}(\dot W)^2+T\frac{\lambda^2}{\dot\eta^2}W^3 +\frac{1}{2\lambda^2\dot\eta^2}(\dot\lambda^2-2\lambda\ddot\lambda+2R\lambda\dot\lambda+2S\lambda^2)W -\frac{1}{\lambda\dot\eta^2}(\dot\lambda\dot\eta+\lambda\ddot\eta-R\lambda\dot\eta)\dot W, \tag{3.15}$$

where the dots denote differentiation with respect to the argument.

Comparing with the standard forms, we see that we must let

$$T(\lambda^2/\dot\eta^2)=\tfrac{3}{2}, \tag{3.16}$$

$$\dot\lambda\dot\eta+\lambda\ddot\eta-R\lambda\dot\eta=0, \tag{3.17}$$

$$(1/2\lambda^2\dot\eta^2)(\dot\lambda^2-2\lambda\ddot\lambda+2R\lambda\dot\lambda+2S\lambda^2)\equiv\mu/2, \tag{3.18}$$

where $\mu(\xi)$ satisfies

$$\mu=\begin{cases}0, & \text{for P XXIX,}\\ \mu_0\neq 0, & \text{for P XXX}\end{cases}$$

($\mu_0=$ constant). Equations (3.16) and (3.17) determine λ and η, namely

$$\dot{\eta} = (2T/3)^{1/2}\lambda, \tag{3.19}$$

$$\lambda = \lambda_0 T^{-1/4} \exp\left(\frac{1}{2}\int R\, d\xi\right) \tag{3.20}$$

and we obtain

$$\mu = \frac{3}{2T\lambda^2}\left\{\frac{\ddot{T}}{2T} - \frac{9\dot{T}^2}{16T^2} - \dot{R} + \frac{3}{4}R^2 - \frac{R\dot{T}}{4T} + 2S\right\}. \tag{3.21}$$

The only freedom is the choice of the constant λ_0. We use λ_0 to normalize μ to $\mu = \pm 1$ if $\mu \neq 0$ and to normalize η if $\mu = 0$.

The results of the Painlevé analysis can be summarized as follows. While most of the ODE's obtained by the symmetry reduction of the NLKGE (1.3) do not have the Painlevé property, quite a few special cases do.

If $(a_2,a_4) \neq (0,0)$ in (1.3) then the only reduction leading to an ODE with the Painlevé property is due to translational invariance and leads to the ODE (2.15) with $k = 0$, $\lambda = \pm 1$. On the other hand, for $a_2 = a_4 = 0$, $a_6 \neq 0$, the results are much richer. In addition to (2.15) we obtain many other reductions to Painlevé-type equations, summarized in Tables V and VI. We see that just three versions of essentially one ODE occur, namely

$$\ddot{W} = (1/2W)\dot{W}^2 + \tfrac{3}{2}W^3 + (\mu/2)W, \quad \mu = 0,1,-1. \tag{3.22}$$

This equation can be once integrated and its first integral for $\dot{W} \neq 0$ is

$$\dot{W}^2 = W^4 + CW + \mu W^2, \quad \mu = 0,1, \text{ or } -1, \tag{3.23}$$

where C is an arbitrary, possibly complex, constant.

Constant nonzero solutions of Eq. (3.22) are also of interest and are equal to

$$W = \pm(-\mu/3)^{1/2}, \quad \mu = \pm 1. \tag{3.24}$$

All reductions of the NLKG equation (1.5) leading to the equations (3.23) in Euclidean space E(4) are summarized in Table V. The number in brackets in column 1 indicates the position of the corresponding reduction in Table II. Similarly, the reductions of (1.5) to the same equations (3.23) in Minkowski space M(3,1) are summarized in Table VI. In column 1 the number in brackets refers to the position of the reduction in Table IV.

To unify the presentation let us reduce Eq. (2.15) with $k = 0$ to a first-order ODE. Multiplying by F_ξ, integrating once, letting

$$F(\xi) = (-8\lambda a_6)^{-1/4}[W(\xi)]^{1/2}, \tag{3.25}$$

and assuming $\dot{F} \neq 0$, we obtain the first-order ODE

$$\dot{W}^2 = W(W^3 + \alpha W^2 + \beta W + \gamma) \tag{3.26}$$

with

$$\alpha = -8\lambda a_4/(-8\lambda a_6)^{1/2}, \quad \beta = -8\lambda a_2, \quad \gamma \in \mathbb{C} \tag{3.27}$$

(γ is an integration constant).

In order to obtain explicit solutions of the NLKGE (1.3) or (1.5) we must now integrate Eq. (3.26) and the special cases (3.23) of this equation. The solutions are, of course, well known. In general they can be expressed in terms of elliptic functions, in special cases they reduce to elementary functions. We proceed to discuss the solutions.

C. Solutions of the Painlevé type equations

Let us start with the most general of the equations involved, namely (3.26). Various versions of this equation are discussed in any book on elliptic functions (see, e.g., Refs. 8 and 46). We shall just summarize the results that we need for the purpose of this paper.

We denote the polynomial on the right-hand side of (3.26)

$$P(W) = W(W - W_1)(W - W_2)(W - W_3) \tag{3.28}$$

with

$$\begin{aligned} W_1 + W_2 + W_3 &= -\alpha, \\ W_1W_2 + W_2W_3 + W_3W_1 &= \beta, \\ W_1W_2W_3 &= -\gamma \end{aligned} \tag{3.29}$$

(the roots W_i are constants).

Equation (3.26) can be simplified by a fractional linear transformation of the dependent variable

$$W(\xi) = (\rho Z(\xi) + \sigma)/(\mu Z(\xi) + \nu), \tag{3.30}$$

where ρ, σ, μ, and ν are constants.

If all four roots of $P(W)$ are distinct we choose ρ, σ, μ, and ν so as to transform the zeros at $W = 0$, W_1, W_2, and W_3 into zeros at $Z = \pm 1$ and $Z = \pm M$, where M is some constant. If three zeros are distinct, we transform the double

TABLE V. Reduction of the equation $\Box_\epsilon \phi = a\phi^5$ ($\epsilon = +1$) in E(4) to the ODE $\dot{W}^2 = W^4 + CW + \mu W^2$; $\phi(x) = \sigma[W(\eta)]^{1/2}$, C is a constant.

No.	Algebra	$\sigma(x)$	$\eta(x)$	μ
P1 (1, $k=0$)	$L_1, K_2, K_3, P_2, P_3, P_0$	$(-4a/3)^{-1/4}$	x_1	0
P2 (1, $k=2$)	L_1, L_2, L_3, P_0	$((4a/3)(x_1^2+x_2^2+x_3^2))^{-1/4}$	$\ln((x_1^2+x_2^2+x_3^2))^{1/2}$	1
P3 (6)	D, L_1, L_2, L_3	$((4a/3)(x_1^2+x_2^2+x_3^2))^{-1/4}$	$\ln\left(\frac{(x_1^2+x_2^2+x_3^2+x_0^2)^{1/2}-x_0}{(x_1^2+x_2^2+x_3^2+x_0^2)^{1/2}+x_0}\right)^{1/2}$	1
P4 (2)	D, K_3, P_0, P_3	$((4a/3)(x_1^2+x_2^2))^{-1/4}$	$\arctan(x_2/x_1)$	-1

TABLE VI. Reduction of the equation $\square_\epsilon \phi = a\phi^5$ ($\epsilon = -1$) in M(3,1) to the ODE $\dot{W}^2 = W^4 + CW + \mu W^2$; $\phi(x) = \sigma[W(\eta)]^{1/2}$, h, θ, f, f_i are functions of $x_0 + x_3$, $\epsilon_0 = \pm 1$, b, and C are constants.

No.	Algebra	$\sigma(x)$	$\eta(x)$	μ
P1 (1)	$L_1, K_2, K_3. P_0, P_2, P_3$	$(-3/4a)^{1/4}$	x_1	0
P2	Degenerate	$(-3/4a)^{1/4}h^{1/2}$	$h(x_1\cos\theta + x_2\sin\theta + f)$	0
P3 (20a, $q=1$)	$D+2K_1, L_1, P_2, P_3$	$(-1/4a)^{1/4}(x_0-x_1)^{-1/3}$	$(x_0^2-x_1^2)^{1/6}(x_0+x_1)^{1/3}$	0
P4 (20b, $q=-\frac{2}{3}$)	$D-\frac{1}{2}K_1, L_3+K_2, P_3$	$(1/4a)^{1/4}(x_0+x_1)^{-1/3}$	$(x_1^2+x_2^2-x_0^2)^{1/2}(x_0+x_1)^{-2/3}$	0
P5 (20b, $q=-2$)	$D+\frac{1}{2}K_1, L_3+K_2, P_3$	$(1/4a)^{1/4}(x_1^2+x_2^2-x_0^2)^{-1/6}(x_0+x_1)^{-1/3}$	$(x_1^2+x_2^2-x_0^2)^{1/6}(x_0+x_1)^{-2/3}$	0
P6 (20b, $q=-2$)	D, L_1, L_2-K_3, L_3+K_2	$(3/4a)^{1/4}(x_0+x_1)^{-1/2}$	$(x_1^2+x_2^2+x_3^2+x_0^2)^{1/2}(x_0+x_1)^{-1}$	0
P7 (20c, $q=-5$)	$D+\frac{4}{3}K_1, L_1, L_2-K_3,$ L_3+K_1	$(3/4a)^{1/4}(x_1^2+x_2^2+x_3^2-x_0^2)^{-1/6}(x_0+x_1)^{-5/6}$	$(x_1^2+x_2^2+x_3^2-x_0^2)^{1/6}(x_0+x_1)^{-5/3}$	0
P8 (22)	$D+\frac{1}{2}K_3, P_1, L_1-K_2+P_0+P_3$	$(-12/a)^{1/4}[4x_2-(x_0+x_3)^2]^{-1/2}(1+\xi^2)^{-1/6}$, $\xi = \dfrac{6(x_3-x_0)+6x_2(x_0+x_3)-(x_0+x_3)^3}{[4x_2-(x_0+x_3)^2]^{3/2}}$	$\int_s^1(1-u^3)^{-1/2}\,du$, $s=(1+\xi^2)^{-1/3}$	0
P9 (8)	L_1, L_2, L_3, P_0	$(-3/4a)^{1/4}(x_1^2+x_2^2+x_3^2)^{-1/4}$	$\ln(x_1^2+x_2^2+x_3^2)^{1/2}$	1
P10 (7)	K_1, K_2, L_3, P_3	$(3/4a)^{1/4}(x_0^2-x_1^2-x_2^2)^{-1/4}$	$\ln(x_0^2-x_1^2-x_2^2)^{1/2}$	1
P11 (20a, $q=-1$)	$D+cL_1, P_2, P_3$	$(-3/4a)^{1/4}(x_0^2-x_1^2)^{-1/4}$	$\ln[(x_0^2-x_1^2)^{1/2}(x_0-x_1)^{-1}]$	1
P12 (20c, $q=1$)	$D+2K_1+cL_1, L_2-K_3,$ L_3+K_1	$(-9/4a)^{1/4}(x_1^2+x_2^2+x_3^2-x_0^2)^{-1/4}$	$\ln(x_1^2+x_2^2+x_3^2-x_0^2)^{1/2}(x_0+x_1)$	1
P13 (18)	D, L_1, L_2, L_3	$(-3/4a)^{1/4}(x_1^2+x_2^2+x_3^2)^{-1/4}$	$\ln\left[\dfrac{(x_0^2-x_1^2-x_2^2-x_3^2)^{1/2}-x_0}{(x_0^2-x_1^2-x_2^2-x_3^2)^{1/2}+x_0}\right]^{1/2}$	1
P14 (19)	D, K_1, K_2, L_3	$(-3/4a)^{1/4}(x_1^2+x_2^2-x_0^2)^{-1/4}$	$\ln\left[\dfrac{(x_1^2+x_2^2+x_3^2-x_0^2)^{1/2}-x_3}{(x_1^2+x_2^2+x_3^2-x_0^2)^{1/2}+x_3}\right]^{1/2}$	1
P15	D, K_3, P_0, P_3	$[-4a(x_1^2+x_2^2)/3]^{-1/4}$	$\arctan(x_2/x_1)$	-1
P16 (11)	Degenerate	$\{-4a[(x_1+f_1)^2+(x_2+f_2)^2]/3\}^{-1/4}$	$\arctan((x_2+f_2)/(x_1+f_1))-\theta$	-1

zero to $Z\to\infty$, the other two to $Z=0$ and $Z=1$. If two are distinct we transform one to $Z\to\infty$, one to $Z\to 0$. If all four coincide, we integrate directly. The existence of at least one multiple zero of P leads to solutions of (3.26) in terms of elementary functions.

Directly from (3.26) we see that $W(\xi-\xi_0)$ is a solution, then so is $W(-\xi+\xi_0)$ and we shall not specify this sign ambiguity each time. The following possibilities occur.

(1) *Constant solutions:* For these we return to the original NLKGE (1.3) and obtain

$$\phi = \epsilon_1\left[\frac{-a_4+\epsilon_2(a_4^2-3a_2a_6)^{1/2}}{3a_6}\right]^{1/2},$$

$$\epsilon_1,\epsilon_2 = \pm 1 \quad (\text{or } \phi = 0). \tag{3.31}$$

(2) *Quadruple root of P(W):* $W_1 = W_2 = W_3 = 0$,

$$W = 1/(\xi-\xi_0). \tag{3.32}$$

(3) *One triple root, one simple one:* (a) $W_2 = W_3 = 0$, $W_1 \neq 0$:

$$W = W_1/[1-\tfrac{1}{4}W_1^2(\xi-\xi_0)^2]. \tag{3.33}$$

These solutions vanish asymptotically in both directions

$$\lim_{\xi\to\pm\infty} W = 0 \tag{3.34}$$

and have simple poles at

$$\xi = \xi_0 \pm 2/W_1. \tag{3.35}$$

(b) $W_1 = W_2 = W_3 \neq 0$:

$$W = W_1(\xi-\xi_0)^2/[(\xi-\xi_0)^2-4W_1^{-2}]. \tag{3.36}$$

These solutions also have two simple poles at the values (3.35); however, they approach a nonzero limit at infinity

$$\lim_{\xi\to\pm\infty} W = W_1 \neq 0. \tag{3.37}$$

(4) *Two double roots:* $W_3 = 0$, $W_1 = W_2 \neq 0$,

$$W = W_1[1-e^{W_1(\xi-\xi_0)}]^{-1}. \tag{3.38}$$

The asymptotic behavior of W is (for Re $W_1 > 0$)

$$\lim_{\xi\to-\infty} W = W_1, \quad \lim_{\xi\to+\infty} W = 0, \tag{3.39}$$

and W is singular for $\xi = \xi_0$.

(5) *One double root, two simple ones:* (a) $W_3 = 0 \neq W_1 \neq W_2 \neq 0$:

$$W = \frac{2W_1W_2}{(W_1 - W_2)\cosh\sqrt{W_1W_2}(\xi - \xi_0) + W_1 + W_2}. \tag{3.40}$$

The asymptotic behavior of W is (for $W_1W_2 > 0$)

$$\lim_{\xi \to \pm\infty} W = 0. \tag{3.41}$$

(b) $0 \neq W_2 = W_3 \neq W_1 \neq 0$:

$$W = W_1W_2 \tanh^2[\sqrt{W_2(W_2 - W_1)}/2](\xi - \xi_0) \times \{W_2 \tanh^2[\sqrt{W_2(W_2 - W_1)}/2(\xi - \xi_0) + W_1 - W_2]\}^{-1}, \tag{3.42}$$

$$\lim_{\xi \to \pm\infty} W = W_2 \neq 0. \tag{3.43}$$

(6) *Four distinct simple roots:* $W_1 \neq W_2 \neq W_3 \neq W_1$, $W_1W_2W_3 \neq 0$. In (3.30) we put

$$-\sigma = \rho = 1,$$
$$\nu = \frac{1}{W_1}\left\{-1 + \epsilon_0\left[\frac{(W_1 - W_2)(W_1 - W_3)}{W_2W_3}\right]^{1/2}\right\},$$
$$\mu = \frac{1}{W_1}\left\{1 + \epsilon_0\left[\frac{(W_1 - W_2)(W_1 - W_3)}{W_2W_3}\right]^{1/2}\right\}, \tag{3.44}$$
$$\epsilon_0 = \pm 1,$$

and transform Eq. (3.26) into the standard form

$$\dot{Z}^2 = A(1 - Z^2)(M^2 - Z^2), \tag{3.45}$$

where

$$M = \{W_1(W_2 + W_3) - 2W_2W_3 + 2\epsilon_0[W_2W_3(W_1 - W_3)(W_1 - W_2)]^{1/2}\}/W_1(W_3 - W_2), \tag{3.46}$$
$$A = \tfrac{1}{4}[W_1(W_2 + W_3) - 2\epsilon_0[W_2W_3(W_1 - W_3)(W_1 - W_2)]]^{1/2}. \tag{3.47}$$

The general solution of Eq. (3.45) can be written in one of three, *a priori* equivalent, forms

$$Z = \mathrm{sn}\,(\sqrt{A}\,M(\xi - \xi_0), 1/M), \tag{3.48a}$$
$$Z = \mathrm{cn}\,(\sqrt{-A(1 - M^2)}(\xi - \xi_0),\ (1 - M)^{-1/2}), \tag{3.48b}$$
$$Z = \mathrm{dn}\,(\sqrt{-A}\,(\xi - \xi_0),\ (1 - M^2)^{1/2}). \tag{3.48c}$$

The Jacobi elliptic functions in (3.48) are particularly convenient if their modulus m [$m = 1/M$, $(1 - M^2)^{-1/2}$, or $(1 - M^2)^{1/2}$ resp.] satisfies

$$m \in \mathbb{R}, \quad 0 < m^2 < 1. \tag{3.49}$$

In this case they have one real and one purely imaginary period. Thus, assuming we have $M^2 \in \mathbb{R}$, we would use (3.48a) for $M^2 > 1$, (3.48b) for $M^2 < 0$, and (3.48c) for $0 < M^2 < 1$.

All of the solutions obtained above depend on one integration constant ξ_0. In general ξ_0 is complex, though often specific reality properties are imposed by the underlying physics. Note that if we replace $\xi_0 \to \xi_0 - i\pi/W_1$ in (3.38), we change the sign in front of the exponential. If we replace $\xi_0 \to \xi_0 - i(\pi/2)(W_1W_2)^{-1/2}$ in (3.40) we change $\cosh(W_1W_2)^{1/2}(\xi - \xi_0)$ into $i\sinh(W_1W_2)^{1/2}(\xi - \xi_0)$. The replacement $\xi_0 \to \xi_0 - i\pi[W_2(W_2 - W_1)]^{-1/2}$ takes the tanh function in (3.42) into the coth one. Note that (3.33) behaves like an "algebraic singular solitary wave," (3.36) like an "algebraic kink." For W_1 real, (3.38) is a kink. For $W_1W_2 > 0$ (3.40) is a solitary wave (bump) vanishing at infinity, for $W_1W_2 < 0$ it is a periodic solution. Similarly for $W_2(W_2 - W_1) > 0$ (3.42) is a solitary wave that does not vanish at infinity. Again, this solution is periodic for $W_2(W_2 - W_1) < 0$.

The function $F(\xi)$ of (3.25), directly related to the solution of the NLKGE (1.3) has, in some cases, a different asymptotic behavior. Thus, for instance, (3.42) yields the solution

$$F(\xi) = (-8\lambda a_6)^{-1/4} \times \sqrt{W_1W_2}\tanh\frac{\sqrt{W_2(W_2 - W_1)}}{2}(\xi - \xi_0) \times\left[W_2\tanh^2\frac{\sqrt{W_2(W_2 - W_1)}}{2}(\xi - \xi_0) + W_1 - W_2\right]^{-1/2}, \tag{3.50}$$

which behaves like a kink

$$\lim_{\xi \to -\infty} F(\xi) = -(-8\lambda a_6)^{-1/4}\sqrt{W_2},$$
$$\lim_{\xi \to +\infty} F(\xi) = +(-8\lambda a_6)^{-1/4}\sqrt{W_2}, \tag{3.51}$$

rather than a bump.

Let us now consider the special cases occurring in the reductions of the NLKGE (1.5), when the above results simplify.

(A) The Painlevé XXIX equation and its first integral (3.23) with $\mu = 0$. This corresponds to the reduction P1 of Table V and P1–P8 of Table VI. The roots W_i $(i = 1,2,3)$ of the polynomial P(W) in (3.28) in this case are

$$W_k = W_0 \exp i(\omega + [2(k-1)/3]\pi), \quad k = 1,2,3,$$
$$W_0, \omega \in \mathbb{R}, \quad 0 \leqslant W_0 < \infty, \quad 0 \leqslant \omega < 2\pi. \tag{3.52}$$

Among the solutions (3.31)–(3.48) the ones that occur are the following.

(i) *Algebraic solutions* with one simple pole. These occur when $C = 0$ in (3.23) (with $\mu = 0$) and hence when $W_0 = 0$ in (3.52). They correspond to

$$W = \pm 1/(\eta - \eta_0), \quad \eta_0 = \text{const.} \tag{3.53}$$

(ii) *Doubly periodic solutions:* These occur for $W_0 \neq 0$ in

(3.52). In this case (3.46)–(3.48) simplify and we obtain

$$M = i(2-\sqrt{3}), \quad A = -(w_0 e^{i\omega}/2)^2(1+2\sqrt{3}). \tag{3.54}$$

In order for the modulus m to satisfy (3.49) we choose the solution (3.48b), i.e. "cnoidal waves."

We have

$$Z = \mathrm{cn}\,\{W_0 e^{i\omega}[3\sqrt{3}-4]^{1/2}(\eta-\eta_0),\ \tfrac{1}{2}[2+\sqrt{3}]^{1/2}\} \tag{3.55}$$

and using (3.30) and (3.44), we have

$$W = W_0 e^{i\omega}\{(Z-1)/[(\sqrt{3}+1)Z+(\sqrt{3}-1)]\}. \tag{3.56}$$

Notice that $W(\eta)$ has poles at $\eta = \eta_c + 4nK$, where

$$Z(\eta_c) = -(\sqrt{3}-1)/(\sqrt{3}+1) \sim -0.2679. \tag{3.57}$$

Moreover, if the argument

$$\Psi \equiv W_0 e^{i\omega}[3\sqrt{3}-4]^{1/2}(\eta-\eta_0)$$

in (3.55) is real, then Z is real and oscillates between 1 (for $\Psi = 4mK$), 0 [for $\Psi = (2m+1)K$], and -1 [for $\Psi = (4m+2)K$], where $4K$ is the real period of cn $(\Psi,[2+\sqrt{3}]^{1/2}/2)$. For Ψ pure imaginary, Z is again real, it has poles for $\Psi = (2n+1)iK'$ and we have $Z(4niK') = 1$, $Z((4n+2)iK') = -1$, where $4K'$ is the corresponding imaginary period. The solution $W(\phi)$ is regular at the poles, so the only singularities of $W(\eta)$ are due to the vanishing of the denominator in (3.56), i.e., they occur at the points η_0 satisfying (3.57). The situation is summarized in Table VII (No. 1).

(B) The Painlevé XXX equation and its first integral (3.23) with $\mu = 1$. This corresponds to reductions P2 and P3 of Table V and P9–P14 of Table VI. In all these cases the variable $\eta(\mathbf{x})$ has the form

$$\eta = \ln \zeta \tag{3.58}$$

and we shall express solutions in terms of ζ directly. Equations (3.29) for the roots of $P(W)$ simplify to

$$W_1 + W_2 + W_3 = 0, \quad W_1W_2 + W_2W_3 + W_3W_1 = 1,$$
$$W_1W_2W_3 = -C. \tag{3.59}$$

The three roots can be parametrized in terms of one complex constant v,

$$W_1 = -2v, \quad W_2 = v + i(1+3v^2)^{1/2}, \quad W_3 = v - i(1+3v^2)^{1/2} \tag{3.60}$$

with

$$C = 2v(4v^2+1).$$

Only some of the solutions (3.31)–(3.48) occur in this case.

(i) *Constant solutions:* Directly from (3.22) with $\mu = 1$ we have

$$W = \pm i\sqrt{\tfrac{1}{3}} \quad (\text{or } W = 0). \tag{3.61}$$

(ii) *Solutions of type (3.40):* These occur for $v = C = 0$. Permuting the roots in (3.60) we have $W_1 = i$, $W_2 = -i$, $W_3 = 0$ and (3.40) reduces to

$$W = -2i[\zeta_0\zeta/(\zeta^2+\zeta_0^2)], \quad \zeta_0 = \text{const.} \tag{3.62}$$

Thus the "solitary wave" solution (3.40) gives rise to the algebraic solution (3.62). Notice that W still satisfies (3.41) and could hence be interpreted as an algebraic solitary wave.

(iii) *Solutions of type (3.42):* From (3.60) we obtain, setting $W_2 = W_3$, that

$$W_1 = -2i\epsilon_1\sqrt{\tfrac{1}{3}}, \quad W_2 = W_3 = i\epsilon_1\sqrt{\tfrac{1}{3}},$$
$$C = -i\epsilon_1 2\sqrt{\tfrac{1}{3}}/3, \quad \epsilon_1 = \pm 1. \tag{3.63}$$

The solitary wave (3.42) turns into a trigonometrically periodic solution, namely

$$W = \frac{-2i\epsilon_1}{\sqrt{3}}\,\frac{\tan^2(\frac{1}{2}\ln(\zeta/\zeta_0))}{\tan^2(\frac{1}{2}\ln(\zeta/\zeta_0))+3}. \tag{3.64}$$

Notice that this solution is regular for $\ln(\zeta/\zeta_0)$ real.

(iv) *Doubly periodic solutions:* Substituting (3.60) into

TABLE VII. Solutions of the equation $\dot{W}^2 = W^4 + CW + \mu W^2$ in terms of Jacobi elliptic functions.

μ	No.	v^2	M^2	A	W	Z
0	1	$\cdots$	$M^2 = -(7-4\sqrt{3}) < 0$	$-(W_0e^{i\omega}/2)^2(1+2\sqrt{3})$	(3.56)	$\mathrm{cn}[W_0e^{i\omega}\sqrt{3\sqrt{3}-4}(\eta-\eta_0),(\sqrt{2+\sqrt{3}})/2]$
1	2a	$0<v^2<\infty$	$M^2<0$	$A<0$	(3.67)	$\mathrm{cn}[\sqrt{-A(1-M^2)}(\eta-\eta_0),\sqrt{1/(1-M^2)}]$
1	2b	$-\infty<v^2<-\frac{1}{3}$	$M^2<0$	$A<0$		
1	3a	$(-4+\sqrt{5})/22\leqslant v^2<0$	$0<M^2\leqslant M_0^2<1$	$A\leqslant 0$	(3.67)	$\mathrm{dn}[\sqrt{-A}(\eta-\eta_0),\sqrt{1-M^2}]$
1	3b	$-\frac{1}{12}<v^2<(-4+\sqrt{5})/22$	$0<M_0\leqslant M^2\leqslant 1$	$A>0$		
1	4a	$(-4-\sqrt{5})/22\leqslant v^2<-\frac{1}{4}$	$1<M^2\leqslant M_1^2$	$A\geqslant 0$	(3.67)	$\mathrm{sn}[\sqrt{A}M(\eta-\eta_0),1/M]$
1	4b	$-\frac{1}{3}<v^2<-(4-\sqrt{5})/22$	$1<M_1^2<M^2<\infty$	$A<0$		
-1	5a	$\frac{1}{3}<v^2<\infty$	$M^2<0$	$A<0$	(3.79)	$\mathrm{cn}[\sqrt{-A(1-M^2)}(\eta-\eta_0),\sqrt{1/(1-M^2)}]$
-1	5b	$-\infty<v^2<0$	$M^2<0$	$A<0$		
-1	6	$\frac{1}{4}<v^2<\frac{1}{3}$	$0<M^2<1$	$A<0$	(3.79)	$\mathrm{dn}[\sqrt{-A}(\eta-\eta_0),\sqrt{1-M^2}]$
-1	7	$0<v^2<\frac{1}{12}$	$1<M^2$	$A<0$	(3.79)	$\mathrm{sn}[\sqrt{A}M(\eta-\eta_0),1/M]$

(3.44)–(3.47), we obtain

$$M = i\,\frac{6v^2+1-[(4v^2+1)(12v^2+1)]^{1/2}}{2v(3v^2+1)^{1/2}}, \qquad (3.65)$$

$$A = -\{2v^2+[(4v^2+1)(12v^2+1)]^{1/2}\}/2, \qquad (3.66)$$

and

$$W(\eta) = 2v\sqrt{4v^2+1}(1-Z(\eta))/[(\sqrt{12v^2+1}+\sqrt{4v^2+1})Z(\eta)+(\sqrt{12v^2+1}-\sqrt{4v^2+1})]. \qquad (3.67)$$

In (3.67) $Z(\eta)$ is a Jacobi elliptic function and $\eta = \ln \zeta$ as in (3.58).

A priori, v is an arbitrary complex number, so that M^2 and the modulus of $Z(\eta)$ is complex. We restrict our considerations to the case when $v^2 \in \mathbb{R}$. We then see that M^2 is real for $v^2 > -\frac{1}{12}$ and for $v^2 < -\frac{1}{4}$. Following the discussion below formulas (3.48) and (3.43), and performing some elementary calculations, we arrive at the situation summarized in Table VII as cases Nos. 2–4.

When the pertinent solution is given by the Jacobi cn function (cnoidal waves) the argument is real as long as $\eta = \ln \xi/\xi_0$ is real. The function $\mathrm{cn}(\chi,m)$ is real and oscillates between $Z = \pm 1$ for $\chi \in \mathbb{R}$, $0 < m^2 < 1$. The function $W(\eta)$ of (3.67) has poles for $\eta = \eta_c$ satisfying

$$Z(\eta_c) = (-\sqrt{12v^2+1}+\sqrt{4v^2+1}) \times [(\sqrt{12v^2+1}+\sqrt{4v^2+1})]^{-1}. \qquad (3.68)$$

In the region where the pertinent solution is No. 3, i.e., given by the function $\mathrm{dn}(\chi,m)$, we can have $\sqrt{-A}$ real or (in a different subregion) pure imaginary. The function $\mathrm{dn}(\chi,m)$ is real and positive [oscillating between $M \leqslant Z(\eta) \leqslant 1$] along the real axis, real and oscillating between $-\infty < Z < \infty$ for χ imaginary.

Finally $\sqrt{A}\,M$ in the sn solution can be real or imaginary. If the argument χ is real, we have $-1 \leqslant \mathrm{sn}(\chi) \leqslant 1$, for χ imaginary $\mathrm{sn}(\chi,1/M)$ is itself imaginary and has poles at $\chi_k = (2k+1)iK'$, where $2iK'$ is the imaginary period.

(C) The Painlevé XXX equation and its first integral (3.23) with $\mu = -1$. This corresponds to the reduction P4 of Table V and P15, P16 of Table VI. The variable $\eta - \eta_0$ occurring in the solutions will have the form

$$\eta - \eta_0 = \arctan \zeta - \theta, \qquad (3.69)$$

where θ is either an arbitrary function of $x_0 + x_3$ [in M(3,1)] or a constant [in E(4)]; ζ is given in the mentioned tables. Equations (3.29) reduce to

$$W_1 + W_2 + W_3 = 0, \quad W_1W_2 + W_2W_3 + W_3W_1 = -1, \quad W_1W_2W_3 = -C. \qquad (3.70)$$

The roots and C are expressed in terms of one complex constant as

$$W_1 = -2v, \quad W_2 = v + i(3v^2-1)^{1/2}, \quad W_3 = v - i(3v^2-1)^{1/2}, \quad C = 2v(4v^2-1). \qquad (3.71)$$

Among the solutions (3.31)–(3.48) the following ones occur.

(i) *Constant solutions:*

$$W = \pm\sqrt{\tfrac{1}{3}}. \qquad (3.72)$$

(ii) *Solutions of the type (3.40):* Putting $v = 0$ in (3.71) and permuting the roots, we have

$$W_1 = 1, \quad W_2 = -1, \quad W_3 = 0, \quad C = 0. \qquad (3.73)$$

Equation (3.40) reduces to $W = \pm[\cos(\eta - \eta_0)]^{-1}$. Using (3.69) we obtain a solution in terms of ζ,

$$W = \pm(1+\zeta^2)^{1/2}/(\cos\theta + \zeta\sin\theta). \qquad (3.74)$$

Thus a solitary wave solution gives rise to an algebraic solution in ζ.

(iii) *Solutions of type (3.42):* We set $v^2 = \frac{1}{3}$ and have

$$W_1 = -2\sqrt{\tfrac{1}{3}}, \quad W_2 = W_3 = \sqrt{\tfrac{1}{3}}, \quad C = 2\sqrt{\tfrac{1}{3}}/3, \qquad (3.75)$$

$$W = \frac{2}{\sqrt{3}}\,\frac{\tanh^2[(\eta-\eta_0)/2]}{3-\tanh^2[(\eta-\eta_0)/2]}, \qquad (3.76)$$

with $\eta - \eta_0$ given in (3.69).

(iv) *Doubly periodic solutions:* Substituting (3.71) into (3.44)–(3.48) we obtain

$$M = i\{6v^2-1-[(4v^2-1)(12v^2-1)]^{1/2}\}/2v\sqrt{3v^2-1}, \qquad (3.77)$$

$$A = -\tfrac{1}{2}(2v^2+[(4v^2-1)(12v^2-1)]^{1/2}), \qquad (3.78)$$

$$W(\eta) = -2v\sqrt{4v^2-1}[Z(\eta)-1]/[(\sqrt{12v^2-1}+\sqrt{4v^2-1})Z(\eta)+(\sqrt{12v^2-1}-\sqrt{4v^2-1})], \qquad (3.79)$$

with η as in (3.69). In general, for v complex we have M^2 complex and there is no reason to give preference to any one of the three types of solutions $Z(\eta)$ of Table VII (Nos. 5–7). Restricting to $v^2 \in \mathbb{R}$ we list all cases when M^2 is real, together with the appropriate choice of Jacobi elliptic function, in Table VII. We find $A < 0$ in all cases. It is a standard matter to discuss the reality conditions, singularities, etc. of the solutions.

IV. DISCUSSION OF THE SOLUTIONS

A. Solutions in Euclidean space

In the general case of the NLKGE (1.3) in Euclidean space, when $(a_2,a_4) \neq (0,0)$, the only exact solutions that we are able to obtain by symmetry reduction are those due to translational invariance. They have the form

$$\phi(\mathbf{x}) = (-8a_6)^{-1/4}[W(\xi)]^{1/2}, \qquad \xi = (\mathbf{A},x+\mathbf{B}), \quad \mathbf{A}^2 = 1, \tag{4.1}$$

where **A** and **B** are constant vectors in $E(n+1)$. The variable ξ is conjugate, under translations and rotations, to $\xi = x_1$.

Thus, $\phi(\mathbf{x})$ can be constant, as in (3.31), or is given by (4.1), where $W(\xi)$ is any of the solutions (3.32)–(3.48) of Eq. (3.26). All these solutions essentially "live" in one space dimension.

All other solutions that we obtain in $E(n+1)$ pertain to the more special NLKGE (1.5). Thus we have $a_2 = a_4 = 0$ and we have put $a \equiv -6a_6 \neq 0$. Subgroups of the symmetry group having generic orbits of codimension 1 lead to the reductions of Table II. The obtained ODE's have the Painlevé property only in the cases P1–P4 of Table V. The equations of Table V were solved in Sec. III C. In some cases the form of these solutions suggests generalizations that provide further solutions. The results can be summed up as the following exact solutions of the NLKGE (1.5) (with $\epsilon = 1$):

$$\phi(\mathbf{x}) = ((3-2k)/4a)^{1/4}(x_0^2 + x_1^2 + \cdots + x_k^2)^{-1/4}, \qquad 0 \leqslant k \leqslant n, \tag{4.2}$$

$$\phi(\mathbf{x}) = (-3c^2/a)^{1/4}(x_1^2 + x_2^2 + x_3^2 + c^2)^{-1/2}, \tag{4.3}$$

$$\phi(\mathbf{x}) = (-3(1-c^2)/4a)^{1/4} \times [(x_0^2 + x_1^2 + x_2^2 + x_3^2)^{1/2} + cx_0]^{-1/2}, \tag{4.4}$$

$$\phi(\mathbf{x}) = [-a(x_1^2 + x_2^2 + x_3^2)]^{-1/4} \times \frac{\tan[\frac{1}{2}\ln c(x_1^2 + x_2^2 + x_3^2)^{1/2}]}{\{3 + \tan^2[\frac{1}{2}\ln c(x_1^2 + x_2^2 + x_3^2)^{1/2}]\}^{1/2}}, \tag{4.5}$$

$$\phi = [-a(x_1^2 + x_2^2 + x_3^2)]^{-1/4} \times \frac{\tan[\frac{1}{2}\ln(\zeta/\zeta_0)]}{\{3 + \tan^2[\frac{1}{2}\ln(\zeta/\zeta_0)]\}^{1/2}}, \qquad \zeta = \frac{[x_0^2 + x_1^2 + x_2^2 + x_3^2]^{1/2} - x_0}{[x_0^2 + x_1^2 + x_2^2 + x_3^2]^{1/2} + x_0}, \tag{4.6}$$

$$\phi = [a(x_1^2 + x_2^2)]^{-1/4}\tanh[(\eta - \eta_0)/2] \times [3 - \tanh^2[(\eta - \eta_0)/2]]^{-1/2}, \qquad \eta = \arctan(x_2/x_1). \tag{4.7}$$

In (4.2) we obtain solutions for arbitrary integer values of k satisfying $0 \leqslant k \leqslant n$, even though Eq. (2.15) has the Painlevé property only for $k = 0$ and $k = 2$. Solutions (4.3) and (4.5) come from P2 of Table V. Notice that for $c^2 > 0$ (4.3) is a localized spherically symmetric static solution. Solutions (4.4) and (4.6) come from P3 of Table V. The corresponding variable ζ is best interpreted in spherical coordinates. Letting

$$x_0 = r\cos\alpha, \quad x_1 = r\sin\alpha\sin\beta\sin\gamma,$$

$$x_2 = r\sin\alpha\sin\beta\cos\gamma, \quad x_3 = r\sin\alpha\cos\beta,$$

we find

$$\zeta = ((r - x_0)/(r + x_0))^{1/2} = \tan(\alpha/2), \qquad r = (x_0^2 + x_1^2 + x_2^2 + x_3^2)^{1/2}. \tag{4.8}$$

Notice also that η of (4.7) is a polar angle in the (x_1,x_2) plane. Solution (4.7) is multivalued since it depends in a nonperiodic manner on this polar angle.

All other exact solutions that we have obtained in E(4) involve Jacobi elliptic functions. They all have the form

$$\phi(\mathbf{x}) = \sigma[W(\eta)]^{1/2}, \tag{4.9}$$

where $\sigma(\mathbf{x})$ and $\eta(\mathbf{x})$ are given in Table V for the four different cases $P_1,...,P_4$ that occur. For P1 the solutions are summarized in row No. 1 of Table VII, for P2 and P3 in rows Nos. 2–4 of the same table, for P4 in rows Nos. 5–7. The reduction P4 of Table V provides multivalued solutions, unless the constant v [see (3.70) and (3.77)–(3.79)] is so chosen that the relevant period of the Jacobi elliptic function involved happens to coincide with that required to obtain $W(\eta) = W(\eta + 2\pi)$.

B. Solutions in Minkowski space

For $(a_2,a_4) \neq (0,0)$ the only reductions, leading to Painlevé type equations, involve translational invariance, as in the Euclidean case. We have

$$\phi(\mathbf{x}) = (-8\lambda a_6)^{-1/4}[W(\xi)]^{1/2} \tag{4.10}$$

and for M(3,1), ξ takes one of two forms.

$$\text{(i) } \xi = (\mathbf{A}_1,\mathbf{x})\cos\theta + (\mathbf{A}_2,\mathbf{x})\sin\theta + f, \quad \lambda = -1,$$

$$(\mathbf{A}_i,\mathbf{A}_k) = -\delta_{ik}, \quad (\mathbf{A}_i,\mathbf{L}) = 0, \quad \mathbf{L}^2 = 0, \tag{4.11}$$

where $\mathbf{A}_1$, $\mathbf{A}_2$, and **L** are constant vectors and θ and f are scalar functions to the null variable $(\mathbf{L},\mathbf{x})$.

$$\text{(ii) } \xi = (\mathbf{A},\mathbf{x} + \mathbf{B}), \quad \mathbf{A}^2 = 1, \quad \lambda = +1 \tag{4.12}$$

(**A** and **B** are constant vectors). As in the Euclidean case $\phi(\mathbf{x})$ can be a constant, given by (3.31), or it is given by (4.10), where $W(\xi)$ is any of the solutions (3.32)–(3.48) of (3.26).

For the homogeneous case $a_2 = a_4 = 0$, we obtain many new exact solutions in terms of elementary functions, or Jacobi elliptic ones. They are obtained by making explicit the results of Sec. III on Painlevé-type equations. In some cases we were able to fit the results into families of solutions, this in turn made it possible to "guess" some particular solutions of ODE's of Table IV that do not have the Painlevé property. Here we shall just list the results; they can easily be checked directly.

1. Algebraic solutions

$$(1)\ \phi(\mathbf{x}) = (-(3-2k)/4a)^{1/4} \times (x_1^2 + x_2^2 + \cdots + x_{k+1}^2)^{-1/4}, \qquad 0 \leqslant k \leqslant n-1. \tag{4.13}$$

For $k = 0$ this is solution No. 1 of Table III, for $k = 1$, No. 3 of Table III, and also P15 of Table VI for $W =$ const. For $k = 2$ this solution is provided by P9 and P13 (both for $W =$ const).

In M(3,1) we have $k = 0,1,2$. Related degenerate variables give further solutions in M(3,1) for $k = 0$ and 1, respectively.

$$(2)\ \phi(\mathbf{x}) = (-3/4a)^{1/4} \times (x_1\cos\theta + x_2\sin\theta + f)^{-1/2}, \tag{4.14}$$

(3) $\phi(\mathbf{x}) = (-1/4a)^{1/4}$

$$\times[(x_1+f_1)^2+(x_2+f_2)^2]^{-1/4}, \quad (4.15)$$

where θ, f, f_1, and f_2 are arbitrary functions of x_0+x_3.

Solution (4.14) is provided by P2, special cases thereof by No. 2 of Table III, P1, and the nonconstant algebraic solutions of P15 and P16. Solution (4.15) is provided by P16 for $W=$ const.

(4) $\phi(\mathbf{x}) = ((3-2k)/4a)^{1/4}$

$$\times(x_0^2-x_1^2-\cdots-x_k^2)^{-1/4}, \quad 0\leqslant k\leqslant n. \quad (4.16)$$

The case $k=1$ comes from P11 with $W=$ const, $k=2$ from No. 9 of Table III and also P10 for $W=$ const, $k=3$ from No. 10 of Table III and also P12 and P14 for $W=$ const.

(5) $\phi(\mathbf{x}) = (3c/a)^{1/4}(x_1^2+x_2^2+x_3^2+c)^{-1/2}$ (4.17)

[see P9 where W is the solution (3.62)].

(6) $\phi(\mathbf{x}) = (3c/a)^{1/4}(x_0^2-x_1^2-x_2^2-c)^{-1/2}$ (4.18)

[see P10 where W is the solution (3.62)].

(7) $\phi(\mathbf{x}) = ([(3-2k)/4a]c)^{1/4}$

$$\times\{(x_0^2-x_1^2-\cdots-x_k^2)\times[(x_0+x_1)^{2k-3}+c]\}^{-1/4}, \quad 1\leqslant k\leqslant n \quad (4.19)$$

(see Table III, Nos. 5–7 for $k=1$, 2, and 3, respectively).

(8) $\phi(\mathbf{x}) = ((3-2k)/4a)^{1/4}((x_0^2-x_1^2-\cdots-x_k^2)^{1/2}$

$$+c(x_0+x_1)^{k/3})^{-1/2}, \quad 1\leqslant k\leqslant n \quad (4.20)$$

(see P3, P4, and P6 for $k=1$, 2, and 3, respectively).

(9) $\phi(\mathbf{x}) = ((3-2k)/4a)^{1/4}((x_0^2-x_1^2-\cdots-x_k^2)^{1/2}$

$$+c(x_0^2-x_1^2-\cdots-x_k^2)^{1/3}\times(x_0+x_1)^{(3k-4)/3})^{-1/2}, \quad 1\leqslant k\leqslant n \quad (4.21)$$

(see P5 and P7 for $k=2$ and 3, respectively).

(10) $\phi(\mathbf{x}) = ([(9-6k)/a]c)^{1/4}$

$$\times((x_0^2-x_1^2-\cdots-x_k^2)(x_0+x_1)^{k-2}+c(x_0+x_1)^{2-k})^{-1/2}, \quad 1\leqslant k\leqslant n \quad (4.22)$$

[for $k=1$ and 3, see P11 and P12, respectively, with W given by (3.62)].

(11) $\phi(\mathbf{x}) = (3(1-c^2)/4a)^{1/4}$

$$\times((x_1^2+x_2^2+x_3^2-x_0^2)^{1/2}+cx_1)^{-1/2} \quad (4.23)$$

[see P14 with W as in (3.62)].

(12) $\phi(\mathbf{x}) = (-3(1-c^2)/4a)^{1/4}$

$$\times((x_0^2-x_1^2-x_2^2-x_3^2)^{1/2}+cx_0)^{-1/2} \quad (4.24)$$

[see P13 with W as in (3.62)].

In (4.17)–(4.24) c is an arbitrary integration constant. Notice that for $c>0$ (4.17) represents a real, static, nonsingular localized solution in M(3,1). It has recently been discussed by Umezawa.[47] Solution (4.18) is independent of the space variable x_3 (in the chosen frame of reference), is singular on the hyperbolic cylinder $x_0^2-x_1^2-x_2^2=c$ and is real for $x_0^2-x_1^2-x_2^2-c>0$, $(3c/a)>0$, pure imaginary otherwise.

2. Elementary nonalgebraic solutions

One elementary nonalgebraic (involving a logarithm) solution is given in No. 8 of Table III, namely

(13) $\phi(x) = (x_1^2+x_2^2+x_3^2-x_0^2+x_0-x_3-2x_2^2/(1+x_0+x_3))^{-1/4}$

$$\times(2a\zeta(\zeta+1)-a(\zeta+1)(\zeta^2-1)\ln[(1+\zeta)/(1-\zeta)]+c(\zeta^2-1)(\zeta+1))^{-1/4}, \quad \zeta=x_0+x_3. \quad (4.25)$$

All other solutions of this type in M(3,1) are obtained from the results of Table VI. More specifically, the reductions P9–P14 yield such solutions when (3.64) is substituted for W, as do P15 and P16 when W is given by (3.76).

A family of solutions in M(n,1) is given by

(14) $\phi(\mathbf{x}) = (-(3-2k)/a)^{1/4}(\tan(\frac{1}{2}\ln c\zeta_k(x_0+x_1)^{k-2})/\{\zeta_k[3+\tan^2(\frac{1}{2}\ln c\zeta_k(x_0+x_1)^{k-2})]\}^{1/2})$,

$$\zeta_k = (x_0^2-x_1^2-\cdots-x_k^2)^{1/2}, \quad 1\leqslant k\leqslant n. \quad (4.26)$$

The cases $k=1$, 2, and 3, correspond to P11, P10, and P12, respectively.

(15) $\phi(\mathbf{x}) = (1/a)^{1/4}(x_1^2+x_2^2+x_3^2)^{-1/4}(\tan[\frac{1}{2}\ln c(x_1^2+x_2^2+x_3^2)^{1/2}]/\{3+\tan^2[\frac{1}{2}\ln c(x_1^2+x_2^2+x_3^2)^{1/2}]\}^{1/2})$ (4.27)

(see P9 of Table VI).

(16) Reduction P13 provides the solution

$$\phi(\mathbf{x}) = (1/a)^{1/4}(x_1^2+x_2^2+x_3^2)\tan(\tfrac{1}{2}\ln c\zeta)\times\{3+\tan^2(\tfrac{1}{2}\ln c\zeta)\}^{-1/2}, \quad (4.28)$$

$$\zeta = [(x_0^2-x_1^2-x_2^2-x_3^2)^{1/2}-x_0]^{1/2}\times[(x_0^2-x_1^2-x_2^2-x_3^2)^{1/2}+x_0]^{-1/2}. \quad (4.29)$$

The variable ζ of (4.29) has a simple meaning in spherical coordinates. Thus, inside the forward light cone of the origin we let

$$x_0 = r\cosh\alpha, \quad x_1 = r\sinh\alpha\sin\theta\cos\phi,$$
$$x_2 = r\sinh\alpha\sin\theta\sin\phi, \quad x_3 = r\sinh\alpha\cos\theta.$$

We then have

$$\zeta = i\tanh(\alpha/2). \tag{4.29'}$$

(17) $$\theta(\mathbf{x}) = (-1/a)^{1/4}(x_0^2 - x_1^2 - x_2^2)^{-1/4} \times(\tan(\tfrac{1}{2}\ln c\zeta)/[3+\tan^2(\tfrac{1}{2}\ln c\zeta)]^{1/2}). \tag{4.30}$$

$$\zeta = \{((-x_0^2 + x_1^2 + x_2^2 + x_3^2)^{1/2} - x_3) \times[((-x_0^2 + x_1^2 + x_2^2 + x_3^2)^{1/2} + x_3)]^{-1}\}^{1/2} \tag{4.31}$$

(see P14 of Table VI).

The variable (4.31) is best understood in terms of hyperbolic coordinates. This time, outside the light cone of the origin, we put

$$x_0 = \rho\sinh\alpha\cosh\beta, \quad x_1 = \rho\sinh\alpha\sinh\beta\cos\phi,$$
$$x_2 = \rho\sinh\alpha\sinh\beta\sin\phi, \quad x_3 = \rho\cosh\alpha,$$

and find

$$\zeta = i\tanh(\alpha/2). \tag{4.31'}$$

(18) The degenerate variable of P16 in Table VI, in conjunction with (3.79), provides the solutions

$$\phi(\mathbf{x}) = \{-a[(x_1+f_1)^2 + (x_2+f_2)^2]\}^{-1/4} \times\tanh\tfrac{1}{2}[\arctan\{(x_2+f_2)/(x_1+f_1)\} - \theta] \times\{3 - \tanh^2\tfrac{1}{2}[\arctan\{(x_2+f_2)/(x_1+f_1)\} - \theta]\}^{-1/2}, \tag{4.32}$$

where f_1, f_2, and θ are arbitrary functions of $x_0 + x_3$.

Reduction P15 provides a special case of (4.32) obtained by putting $f_1 = f_2 = 0$, $\theta = \text{const}$.

As in the Euclidean space case of solution (4.7), the Minkowski space solution (4.32) is multivalued. Indeed, it involves a nonperiodic function ($\tanh x$) of the polar angle in the (x_1, x_2) plane.

3. Solutions in terms of Jacobi elliptic functions

Each reduction of Table VI provides solutions of the NLKGE (1.5) in terms of Jacobi elliptic functions. Comparing with Table VII, we see that P1–P8 provide solutions $\phi(\mathbf{x}) = \sigma[W(\eta)]^{1/2}$ with W expressed, preferably, in terms of the function $\text{cn}(x,m)$. The reductions P9–P14, and also P15 and P16, give solutions in terms of $\text{cn}(x,m)$, $\text{sn}(x,m)$, or $\text{dn}(x,m)$, depending on the values of the parameter v^2 involved.

V. CONCLUSIONS

A few words should be said about the "integrability" or "nonintegrability" of the NLKGE (1.3) and its particular case (1.5). These equations are not of the type that can be solved by inverse scattering, or other linear techniques,[9,48] in $(1+1)$ dimensions, still less in $(n+1)$.

We have been able to obtain a large number of exact particular solutions, because the method of symmetry reduction in many cases leads to ODE'S with the Painlevé property.

A Painlevé test, very similar to the one described in Sec. III, can be performed directly for PDE's,[49,50] without first performing a symmetry reduction. Equation (1.3), and also (1.5), fail the test; the solutions of these PDE's are not necessarily single valued in the neighborhood of their singularity surfaces. This is reflected in the fact that generically speaking, the reduced ODE's of Tables II and IV do not have the Painlevé property, only the exceptional ones, listed in Tables V and VI, do.

This paper is entirely devoted to the problem of finding the exact solutions, summarized in Sec. IV. A sequel, in which we study physical applications of the solutions, as outlined in the Introduction, is in preparation.

ACKNOWLEDGMENTS

We thank D. W. Rand for performing a large number of MACSYMA calculations that were used in this paper and also P. Winternitz, Jr. whose computer graphs are reflected in Table VII.

The research reported in this paper was partly supported by the Natural Sciences and Engineering Research Council of Canada, that of P. Winternitz also by the Fonds FCAR du Gouvernement du Québec.

[1]E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956).
[2]M. J. Ablowitz, A. Ramani, and H. Segur, J. Math. Phys. **21**, 715 (1980).
[3]D. Rand and P. Winternitz, Comp. Phys. Commun. **42**, 359 (1986).
[4]P. Painlevé, Acta Math. **25**, 1 (1902); B. Gambier, *ibid.* **33**, 1 (1910).
[5]S. Fubini, Nuovo Cimento A **34**, 521 (1976); V. de Alfaro, S. Fubini, and G. Furlan, *ibid.* **34**, 569 (1976).
[6]R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[7]R. Rajaraman, *Solitons and Instantons* (North-Holland, Amsterdam, 1982).
[8]A. Erdelyi, W. Magnus, F. Oberhettinger, and F. C. Tricomi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 2.
[9]M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981).
[10]D. J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena* (McGraw-Hill, New York, 1978).
[11]L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1980).
[12]S. Galam and J. L. Birman, Phys. Lett. A **98**, 125 (1983).
[13]S. K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, 1976).
[14]A. Aharony, "Multicritical points," in *Critical Phenomena*, edited by F. J. W. Hahne, *Lecture Notes in Physics*, Vol. 186 (Springer, Berlin, 1983).
[15]H. Haken, "Synergetics: Basic concepts and mathematical tools, in cooperative effects," *Progress in Synergetics*, edited by H. Haken (North-Holland, Amsterdam, 1974).
[16]R. Blinc and B. Zeks, *Soft Modes in Ferroelectrics and Antiferroelectrics* (North-Holland, Amsterdam, 1974).
[17]A. Gordon, Physica B **122**, 321 (1983).
[18]S. N. Behera and A. Khare, Pramana **15**, 245 (1980).
[19]S. N. Behera and A. Khare, Lett. Math. Phys. **4**, 153 (1980).
[20]Y. Ishibashi and I. Suzuki, J. Phys. Soc. Jpn. **53**, 1093 (1984).
[21]K. Parlinski and P. Zielinski, Z. Phys. B **44**, 317 (1981).
[22]P. B. Burt, Phys. Rev. Lett. **32**, 1080 (1974); Lett. Nuovo Cimento **13**, 26 (1975).
[23]D. Boyanovsky and L. Masperi, Phys. Rev. D **21**, 1550 (1980).
[24]G. Ciecura and A. M. Grundland, J. Math. Phys. **25**, 3460 (1984).
[25]J. A. Krumhansl and J. R. Schrieffer, Phys. Rev. B **11**, 3535 (1975).
[26]D. J. Scalapino, M. Sears, and R. A. Farrell, Phys. Rev. B **6**, 3409 (1972).
[27]J. F. Currie, J. A. Krumhansl, A. R. Bishop, and S. E. Trullinger, Phys.

Rev. B **22**, 477 (1980).
[28]R. F. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D **10**, 4130 (1979).
[29]J. Goldstone and R. Jackiw, Phys. Rev. D **11**, 1486 (1975).
[30]A. R. Bishop, "Solitons and physical perturbations," in *Solitons in Action*, edited by K. Lonngren and A. Scott (Academic, New York, 1978).
[31]A. R. Bishop and J. A. Krumhansl, Phys. Rev. B **12**, 2824 (1975).
[32]T. Schneider and E. Stoll, Phys. Rev. B **13**, 1216 (1976).
[33]L. V. Ovsiannikov, *Group Analysis of Differential Equations* (Academic, New York, 1982).
[34]G. W. Bluman and J. D. Cole, *Similarity Methods for Differential Equations* (Springer, New York, 1974).
[35]P. Olver, *Applications of Lie Groups to Differential Equations* (Springer, New York, 1986).
[36]P. Winternitz, in "Nonlinear phenomena," *Lecture Notes in Physics*, Vol. 189 (Springer, Berlin, 1983).
[37]A. M. Grundland, J. Harnad, and P. Winternitz, J. Math. Phys. **25**, 791 (1984).
[38]A. M. Grundland, J. Harnad, and P. Winternitz, in *Wave Phenomena, Modern Theory and Applications* (North-Holland, Amsterdam, 1984).
[39]F. Schwarz, Comput. Phys. Comm. **27**, 179 (1982); Computing **34**, 91 (1985).
[40]B. Champagne and P. Winternitz, Preprint CRM-1278, Montréal, 1985.
[41]D. David, N. Kamran, D. Levi, and P. Winternitz, Phys. Rev. Lett. **55**, 2111 (1985); J. Math. Phys. **27**, 1225 (1986).
[42]J. Patera, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **16**, 1615 (1975).
[43]J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **18**, 2259 (1977).
[44]J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus, Can. J. Phys. **54**, 950 (1976).
[45]J. Patera, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **17**, 717 (1976).
[46]P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists* (Springer, Berlin, 1971).
[47]M. Umezawa, Prog. Theor. Phys. **75**, 474, 953, 1231 (1986).
[48]M. J. Ablowitz and A. S. Fokas, in "Nonlinear phenomena," *Lecture Notes in Physics*, Vol. 189 (Springer, Berlin, 1983).
[49]J. E. Weiss, M. Tabor, and G. Carnavale, J. Math. Phys. **24**, 522 (1983).
[50]J. E. Weiss, J. Math. Phys. **24**, 1405 (1983).

Renormalization schemes and renormalization group functions in Yang–Mills theories with massive fermions

H. Caprasse and M. Hans[a)]
Physique Théorique et Mathématique, Université de Liège[b)]

(Received 12 December 1985; accepted for publication 6 May 1987)

The connection between renormalization schemes (RS's) and the renormalization group (RG) functions for a massive Yang–Mills theory is investigated. The RS's are defined in a manner independent of the regularization procedure. The RS transformations are defined in such a way that it is clear that they form a group. It is shown that to a given set of RG functions corresponds an infinite number of RS's. The subgroup of RS transformations which leave invariant the (mass-shell) MS-RG functions is carefully described. Gauge invariance, regularity of the theory when $m \to 0$ and mass decoupling are imposed and the corresponding indeterminations of RS's are given. It is seen that a RS which fulfills simultaneously the above conditions does not exist.

I. INTRODUCTION

In massless field theories the connection between renormalization schemes (RS's) and the renormalization group[1] (RG) functions is well known. Up to an arbitrary constant the RS is fixed once the RG β function coefficients are given. This property had a great impact on our way of handling calculations in QCD and, in particular, Stevenson relied upon it in his proposed optimization method.[2]

Here we study the connection between RS's and the RG functions for a massive Yang–Mills theory. This connection turns out to be much more intricate than for the massless case. It is due not only to the fact that there are now several RG functions but also to the existence of a mass scale associated to the massive fermions.

We formulate the problem in such a way that it becomes clear that the set of all RS transformations form a group. We study the structure of this group with respect to its action on the set of RG functions. This allows us to show that to each set of RG functions correspond an *infinite* number of RS's. This is true, in particular, if all RG functions are chosen equal to zero. We also add several additional realistic boundary conditions and show how the indetermination is reduced. In particular, gauge invariance, regularity when $m \to 0$, and decoupling when $m \to \infty$ cannot be fulfilled simultaneously.

In Sec. II we give a precise definition of a RS and introduce the RG functions.

In Sec. III we give an appropriate formulation of a change of RS, we give the group of RS transformations, and finally we write the equations which express the change of RG functions.

In Sec. IV we discuss in detail the properties of the system of equations given in Sec. III. We discover equivalence classes among the RS's. They contain all RS's which correspond to *one given* set of RG functions.

In Sec. V, we show that this degeneracy is described by the subgroup of RS transformations leaving this set of RG functions invariant. It does not depend on the particular set we use. For the (mass-shell) MS-RG functions, we introduce four sets of particular solutions. To each set corresponds a function which contains the degeneracy.

In Secs. VI–VIII we add to our previous system of equations boundary conditions which express the regularity in mass of the theory (when $m \to 0$ and $m \to \infty$) and the gauge dependence of a RS. We deduce the extent to which the degeneracy is reduced.

II. THE BASIC INGREDIENTS

A. Definition of an RS

It is by now well known (see, for instance, Ref. 3) that renormalization consists in the subtraction of divergences existing in the power series expansions in the coupling constant of a few independent irreducible Green's functions. In a typical Yang–Mills theory we need to consider the gauge boson, fermion, and Faddeev–Popov propagators together with the various vertex functions. They are not all independent because of the Slavnov–Taylor identities.

Let us denote by S_Γ a set of independent irreducible Green's functions. We define a RS independently of the regularization procedure used[4] in the following way. If μ is the renormalization scale we demand that all Γ in S_Γ

$$\Gamma|_{p^2 = -\mu^2} = \tau_\Gamma(\alpha,\lambda,\eta). \tag{1}$$

Here, α is the coupling, λ the gauge parameter, and $\eta = m/\mu$ with m a generic symbol for the various fermion masses.

We can use a particular set of irreducible Green's functions, a unique renormalization scale μ, and a particular renormalization point for three- or four-point functions and still have the most general RS because the functions τ_Γ are arbitrarily chosen. Indeed, to any RS corresponds a unique set of τ_Γ which are determined by taking $p^2 = -\mu^2$ into the Green's functions Γ. The functions τ_Γ take care of all arbitrariness.

From now on we shall take QCD in the Lorentz gauge as our prototype. The following considerations can neverthe-

a) Research assistant, National Fund for Scientific Research (Belgium).
b) Postal address: Institut de Physique au Sart Tilman, Bâtiment B.5, B-4000 Liège 1, Belgium.

 0022-2488/87/092213-07$02.50

less be applied to all theories which satisfy multiplicative renormalization conditions like (2) and (3) below. These are quite general since they involve a mass parameter and a gauge parameter. Of course, restrictions to theories without mass or gauge parameters is straightforward.

B. The RG functions

The renormalization constants are defined by the relations[4]

$$\alpha_B = Z_\alpha \alpha \mu^\epsilon, \quad m_B = Z_m m, \quad \lambda_B = Z_\lambda \lambda, \tag{2}$$

and for the matter field by

$$\psi_B = Z_2^{1/2} \psi \tag{3}$$

and an analogous equation for the Faddeev–Popov field. In these equations, μ denotes the regularization scale and the index B defines the bare quantities.

The RG functions are

$$\beta = \mu \frac{d\alpha}{d\mu}, \; \gamma = \frac{\mu}{m} \frac{dm}{d\mu}, \; \gamma_\lambda = \mu \frac{d\lambda}{d\mu}, \; \gamma_\psi = \frac{\mu}{Z_2} \frac{dZ_2}{d\mu}. \tag{4}$$

All these functions are finite, depend on α, λ, η, and also of the RS.

III. THE LINK BETWEEN RS'S AND THE RG FUNCTIONS

A. The change of RS

The change of RS is introduced by a change of the τ_Γ functions defined in (1). The Z functions are consequently modified but the bare quantities stay the same, so that

$$Z_\alpha \alpha = Z'_\alpha \alpha', \quad Z_m m = Z'_{m'} m', \quad Z_\lambda \lambda = Z_{\lambda'} \lambda', \quad Z_2 S = Z'_2 S'. \tag{5}$$

Primes denote the corresponding quantities in the new RS while S is the residue at the pole of the Fourier transform of the matter propagator,

$$\int d^4x\, e^{-ipx} \langle 0 | T\bar{\psi}(x)\psi(0)|0\rangle = \frac{S(p^2)}{\not p - m(p^2)}. \tag{6}$$

Starting from the MS scheme[5] and deleting the primes, we deduce from (5) the relations between the parameters in the MS scheme and those of a general RS,

$$\alpha_{\rm MS} = \alpha A_0(\alpha,\lambda,\eta), \quad \eta_{\rm MS} = \eta M_0(\alpha,\lambda,\eta), \quad \lambda_{\rm MS} = \lambda L_0(\alpha,\lambda,\eta), \quad S_{\rm MS} = S B_0(\alpha,\lambda,\eta), \tag{7}$$

when A_0, M_0, L_0, and B_0 are the convergent parts of the Z constants.

The functions A_0, M_0, L_0, and B_0 in relation (7) represent a change of RS starting from the MS scheme and define the final RS. However, it is straightforward to show that a change of RS starting *from any scheme* $(\text{RS})_0$ gives rise to the same equations [with MS replaced by $(\text{RS})_0$ and $A_0,\dots, B_0$ replaced by some functions we call f, g, h, l]. The functions f, g, h, l do not depend on the starting scheme. With this in mind we rewrite (7) in the compact form,

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}), \quad s'(\mathbf{x}') = s(\mathbf{x})\cdot l(\mathbf{x}), \tag{8}$$

where

$$\mathbf{x} = (\alpha,\lambda,\eta), \quad \mathbf{x}' = (\alpha',\lambda',\eta'), \quad \mathbf{f} = (\alpha f(\alpha,\lambda,\eta),\ \lambda g(\alpha,\lambda,\eta),\ \eta h(\alpha,\lambda,\eta)). \tag{9}$$

The p^2-dependence of s is implicit. The functions $\mathbf{f}(\mathbf{x})$ and $s(\mathbf{x})$ so introduced generate a *true* group (and not merely a groupoid[6]). Its multiplication law under composition is

$$[\mathbf{g}(\mathbf{x}),k(\mathbf{x})]\circ[\mathbf{f}(\mathbf{x}),l(\mathbf{x})] = [\mathbf{g}(\mathbf{f}(\mathbf{x})),k(\mathbf{f}(\mathbf{x}))\cdot l(\mathbf{x})] \tag{10}$$

and every transformation has an inverse since in perturbation theory the inverse functions of f, g, h, and l always exist. Here $\mathbf{f}(\mathbf{x}) = \mathbf{x}$, $l(\mathbf{x}) = 1$ is of course the unit transformation.

B. The change of RG functions

From (4) and (7) we deduce the differential equations relating the RG functions in the new scheme to those in the MS scheme. We obtain

$$\beta_{\rm MS} = \left(f + \alpha \frac{\partial f}{\partial \alpha}\right)\beta + \alpha \frac{\partial f}{\partial \lambda}\gamma_\lambda + \alpha \frac{\partial f}{\partial \eta}\Gamma, \tag{11}$$

$$\gamma_{\lambda\,\rm MS} = \lambda \frac{\partial g}{\partial \alpha}\beta + \left(g + \lambda \frac{\partial g}{\partial \lambda}\right)\gamma_\lambda + \lambda \frac{\partial g}{\partial \eta}\Gamma, \tag{12}$$

$$\Gamma_{\rm MS} = \eta \frac{\partial h}{\partial \alpha}\beta + \eta \frac{\partial h}{\partial \lambda}\gamma_\lambda + \left(h + \eta \frac{\partial h}{\partial \eta}\right)\Gamma, \tag{13}$$

$$\gamma_{\psi\rm MS}\cdot l = l\cdot\gamma_\psi - \left(\frac{\partial l}{\partial \alpha}\beta + \frac{\partial l}{\partial \lambda}\gamma_\lambda + \frac{\partial l}{\partial \eta}\Gamma\right), \tag{14}$$

where

$$\Gamma = \mu \frac{d\eta}{d\mu} = \eta(\gamma - 1) \tag{15}$$

has been introduced to write the equations in the most elegant way.

These equations are restricted by the fact that β, γ, γ_λ, and γ_ψ must have the expansions

$$\begin{aligned}
\beta(\alpha,\lambda,\eta) &= \alpha^2\beta^{(2)}(\lambda,\eta) + \alpha^3\beta^{(3)}(\lambda,\eta) + \cdots,\\
\gamma_\lambda(\alpha,\lambda,\eta) &= \alpha\gamma_\lambda^{(1)}(\lambda,\eta) + \alpha^2\gamma_\lambda^{(2)}(\lambda,\eta) + \cdots,\\
\Gamma(\alpha,\lambda,\eta) &= \eta[-1 + \alpha\gamma^{(1)}(\lambda,\eta) + \alpha^2\gamma^{(2)}(\lambda,\eta) + \cdots],\\
\gamma_\psi(\alpha,\lambda,\eta) &= \alpha\gamma_\psi^{(1)}(\lambda,\eta) + \alpha^2\gamma_\psi^{(2)}(\lambda,\eta) + \cdots,
\end{aligned} \tag{16}$$

and if t represents one of the functions f, g, h, l, it has the expansion

$$t(\alpha,\lambda,\eta) = 1 + t_1(\lambda,\eta)\alpha + t_2(\lambda,\eta)\alpha^2 + \cdots. \tag{16'}$$

Equations (11)–(14), (16), and (16′) allow us to study precisely the relationship between the f, g, h, l functions and the RG functions β, γ_λ, Γ, γ_ψ.

IV. DISCUSSION OF THE RELATIONSHIP BETWEEN THE RS AND RENORMALIZATION GROUP FUNCTIONS

In this section we solve the system of differential equations introduced, i.e., we determine the set of solutions $t(\alpha,\lambda,\eta)$ from the knowledge of the set of renormalization functions. For the coefficients of the α expansions of the MS scheme RG functions we use the notations $b^{(k)}$, $C_\lambda^{(k)}$, $C^{(k)}$, and $C_\psi^{(k)}$.

The system (11)–(14) is solved recursively using (16) and (16′). To first order we get

$$\eta\frac{\partial f_1}{\partial\eta}=\beta^{(2)}-b^{(2)},\quad \eta\frac{\partial g_1}{\partial\eta}=\frac{\gamma_\lambda^{(1)}-C_\lambda^{(1)}}{\lambda},$$
$$\eta\frac{\partial h_1}{\partial\eta}=\gamma^{(1)}-C^{(1)},\quad \eta\frac{\partial l_1}{\partial\eta}=C_\psi^{(1)}-\gamma_\psi^{(1)}, \tag{17}$$

and, in general, to order k,

$$\eta\frac{\partial f_k}{\partial\eta}=(k\beta^{(2)}-2b^{(2)})f_{k-1}+\gamma_\lambda^{(1)}\frac{\partial f_{k-1}}{\partial\lambda}+\gamma^{(1)}\eta\frac{\partial f_{k-1}}{\partial\lambda}+\cdots, \tag{18}$$

$$\eta\frac{\partial g_k}{\partial\eta}=\left[(k-1)\beta^{(2)}+\frac{\gamma_\lambda^{(1)}}{\lambda}-\frac{\partial C_\lambda^{(1)}}{\partial\lambda}\right]g_{k-1}+\gamma_\lambda^{(1)}\frac{\partial g_{k-1}}{\partial\lambda}+\gamma^{(1)}\eta\frac{\partial g_{k-1}}{\partial\eta}-\frac{C_\lambda^{(1)}}{\lambda}f_{k-1}+\cdots, \tag{19}$$

$$\eta\frac{\partial h_k}{\partial\eta}=[(k-1)\beta^{(2)}+\gamma^{(1)}-C^{(1)}]h_{k-1}+\gamma_\lambda^{(1)}\frac{\partial h_{k-1}}{\partial\lambda}+\gamma^{(1)}\eta\frac{\partial h_{k-1}}{\partial\eta}-C^{(1)}f_{k-1}+\cdots, \tag{20}$$

$$\eta\frac{\partial l_k}{\partial\eta}=[(k-1)\beta^{(2)}-\gamma_\psi^{(1)}+C_\psi^{(1)}]l_{k-1}+\gamma_\lambda^{(1)}\frac{\partial l_{k-1}}{\partial\lambda}+\gamma^{(1)}\eta\frac{\partial l_{k-1}}{\partial\eta}+C_\psi^{(1)}f_{k-1}+\lambda\frac{\partial C_\psi^{(1)}}{\partial\lambda}g_{k-1}+\cdots, \tag{21}$$

where the terms not written contain coefficients of order less than or equal to $k-2$.

These equations are recursively integrated and at each order appears an integration constant which is an arbitrary function of the gauge parameter λ. At order k, if the arbitrary function associated with the function t (i.e., f, g, h, or l) is called $T_k(\lambda)$, then

$$t_k(\lambda,\eta,T_k(\lambda),\ldots,T_1(\lambda))=t_k^0(\lambda,\eta,T_{k-1}(\lambda),\ldots,T_1(\lambda))+T_k(\lambda). \tag{22}$$

Therefore there is an infinite number of RS's corresponding to one set of RG functions. This degeneracy can be described by the arbitrary functions $T_k(\lambda)$ with $k=1,2,\ldots$.

The discussion above is valid for any RG functions. In particular, we can take

$$\beta=\gamma=\gamma_\psi=\gamma_\lambda=0. \tag{23}$$

The corresponding schemes are μ-independent. There are an infinite number of those RS's. In each of them one can define a coupling $\hat\alpha$, a gauge parameter $\hat\lambda$, and a mass $\hat m$ independent of μ. Any physical quantity can be expressed in terms of these parameters and to any finite order of perturbations will depend on the other parameters which describe the RS's. If any optimization is attempted it must be done on these parameters and not on the RG function coefficients which are all equal to zero.

V. THE ORIGIN OF THE DEGENERACY

In this section we explain the origin of the degeneracy we have described: what happens is that among the RS transformations many do not change the RG functions. The set of all RS corresponding to the same set of RG functions form an equivalence class. All classes are in one-to-one correspondence with each other.

The general transformation law of RG functions under a change of RS can be written in a compact form. Taking the derivative of (8) with respect to μ we have

$$\boldsymbol{\beta}'(\mathbf{f}(\mathbf{x}))=\frac{d\mathbf{f}}{d\mathbf{x}}\cdot\boldsymbol{\beta}(\mathbf{x}), \tag{24}$$

$$\gamma'_\psi(\mathbf{f}(\mathbf{x}))l(\mathbf{x})=\gamma_\psi(\mathbf{x})l(\mathbf{x})+\frac{dl}{d\mathbf{x}}\cdot\boldsymbol{\beta}(\mathbf{x}), \tag{24'}$$

where

$$\boldsymbol{\beta}=(\beta,\gamma_\lambda,\Gamma)$$

and $d\mathbf{f}/d\mathbf{x}$ is the Jacobian matrix of the transformation (8).

Transformations leaving the RG functions invariant ($\boldsymbol{\beta}'=\boldsymbol{\beta}$, $\gamma'_\psi=\gamma_\psi$) have the following properties. (1) If $[\mathbf{f}(\mathbf{x}),l(\mathbf{x})]$ is a solution, $[\mathbf{f}(\mathbf{f}(\mathbf{x})),l(\mathbf{f}(\mathbf{x}))l(\mathbf{x})]$ is also a solution (from the properties of the Jacobian matrix). (2) If $[\mathbf{f}(\mathbf{x}),l(\mathbf{x})]$ and $[\mathbf{g}(\mathbf{x}),\mathbf{k}(\mathbf{x})]$ are two solutions, so is their composition (10); transformations of RS leaving a set of RG functions invariant therefore define a *subgroup* of the renormalization group. (3) If $[\mathbf{f}(\mathbf{x},\nu),l(\mathbf{x},\nu)]$ is a family of solutions parametrized by ν, then it is still a solution if one makes the substitution

$$\nu\to\hat\nu(\mathbf{x}) \tag{25}$$

for any function $\hat\nu(\mathbf{x})$ satisfying

$$\boldsymbol{\beta}\cdot\frac{d\hat\nu}{d\mathbf{x}}=0 \tag{26}$$

or, equivalently,

$$\mu\frac{d\hat\nu}{d\mu}=0. \tag{27}$$

All invariance subgroups are in one-to-one correspondence. It is interesting to study more carefully the invariance subgroup associated with the MS-RG functions $r_{\rm MS}(\alpha,\lambda,\eta)$. The most general transformation is constructed from four *particular* solutions. From Eqs. (11)–(15) with $\beta=\beta_{\rm MS},\ldots,\gamma_\psi=\gamma_{\psi\rm MS}$ and the particular properties of these functions,

$$\frac{\partial\beta_{\rm MS}}{\partial\lambda}=\frac{\partial\beta_{\rm MS}}{\partial\eta}=\frac{\partial\gamma_{\rm MS}}{\partial\lambda}=\frac{\partial\gamma_{\rm MS}}{\partial\eta}=\frac{\partial\gamma_\psi^{\rm MS}}{\partial\eta}=\frac{\partial\gamma_\lambda^{\rm MS}}{\partial\eta}=0, \tag{28}$$

we extract easily the following solutions: (1) $\mathbf{f}(\mathbf{x})=\mathbf{x}$ and $l(\mathbf{x})$ is an η independent function; it has to satisfy

$$\mu\frac{dl}{d\mu}=0; \tag{29}$$

(2) $f=1$, $g=1$, $l=1$, but $h(x)$ is a function independent either of η, or of λ; it satisfies also

$$\mu\frac{dh}{d\mu}=0; \tag{30}$$

(3) $f=1$, $h=1$, but $g=g(\alpha,\lambda)$ and $l=l(\alpha,\lambda)$ (because

$\gamma_\psi^{\rm MS}$ depends on λ) are η-independent solutions; (4) $f=f(\alpha)$, and thus $g=g(\alpha,\lambda)$, $h=h(\alpha)$, $l=l(\alpha,\lambda)$ (all RG functions depend on α); they define a solution independent of η.

These solutions can be chosen to be independent. Thanks to property (1) and the fact that the particular solutions are not the identity, it is easy to generate for each of them an infinite number of other solutions of the same form[7] by composing the solution with itself (iteration). These iterations can be interpolated to form a one-parameter family (ν_i for solution i), such that the composition of a solution with ν_i^1 and a solution with ν_i^2 is a solution with $\nu_i^1+\nu_i^2$. Each family is thus an Abelian subgroup of the RG.

Using property (3), we replace ν_i by $\hat{\nu}_i(\mathbf{x})$. The forms of the solutions are still the same but the latter define now for each i a non-Abelian subgroup of the RG. Indeed, composition of a solution with $\hat{\nu}_i^1$ and a solution with $\hat{\nu}_i^2$ is

$$\hat{\nu}_i^1(\mathbf{f}_i(\mathbf{x},\hat{\nu}_i^2))+\hat{\nu}_i^2. \tag{31}$$

It is a solution because it verifies (27). Manifestly, these subgroups are non-Abelian except for $i=1$ because

$$\mathbf{f}_1(\mathbf{x},\hat{\nu})=\mathbf{x} \quad \text{and} \quad l_1(\mathbf{x},\hat{\nu})=[l_1(\mathbf{x})]^{\hat{\nu}}. \tag{32}$$

Using property (2) we generate the general solution by composing the four families of solutions. This general solution depends on four arbitrary functions: $\hat{\nu}_i$ satisfying (27). In the Appendix, we show that the general solution of (27) depends on a sequence of arbitrary functions of λ (one for each order). The general solution obtained by composition is therefore the most general: we recover the same degree of arbitrariness as in (22). More precisely, the arbitrary functions of λ contained in $\hat{\nu}_4$ are in a one-to-one relationship with the sequence of the $F_k(\lambda)$. We use the notation

$$\hat{\nu}_4 \leftrightarrow \{F_k(\lambda)\}. \tag{33}$$

Taking into account this relation, we can show that the arbitrary functions of λ contained in $\hat{\nu}_3$ are in a one-to-one relationship with the $G_k(\lambda)$,

$$\hat{\nu}_3 \leftrightarrow \{G_k(\lambda)\}. \tag{34}$$

Taking into account Eqs. (33) and (34), there is a one-to-one relationship between

$$\hat{\nu}_2 \leftrightarrow \{H_k(\lambda)\} \tag{35}$$

and finally with (33)–(35) we have

$$\hat{\nu}_1 \leftrightarrow \{L_k(\lambda)\}. \tag{36}$$

In conclusion, it is possible to generate the most general solution starting from four very particular solutions using the properties of the system. It is useful to calculate the degree of arbitrariness in any circumstances. For instance, in a theory without mass and gauge parameters, solutions (2) and (3) disappear, solution (1) is the identity [see (29)], and $\hat{\nu}$ can only be a constant. In that theory, arbitrariness is reduced to a single constant $\hat{\nu}_4=\zeta$ and the invariance subgroup is Abelian.

The analysis we have done of Eqs. (11)–(14) is the most general one in the sense that the constraints imposed on them are the weakest ones [see (20)]. In the following sections we shall examine what becomes of the analysis of (11)–(14) under various additional boundary conditions.

VI. MASS BEHAVIOR

We require that the theory, i.e., both the RG functions and the RS transformations, be regular when $m\to 0$.[8,9]

First, Eqs. (11)–(14) admit solutions only if Eqs. (17) are satisfied, i.e., if

$$\beta^{(2)}\to b^{(2)}, \quad \gamma^{(1)}\to C^{(1)}, \quad \gamma_\lambda^{(1)}\to C_\lambda^{(1)}, \quad \gamma_\psi^{(1)}\to C_\psi^{(1)}, \tag{37}$$

when $m\to 0$.[10]

Second, Eqs. (18)–(21) still admit several solutions but the degeneracy is restricted by the fact that the functions $\hat{\nu}_i$ not only satisfy (27) but must satisfy the equations

$$\eta\frac{\partial\hat{\nu}_i}{\partial\eta}\to 0. \tag{38}$$

The solution of (38) is written in the Appendix in the general case. It depends only on a sequence of arbitrary *constants*. Symbolically, the arbitrariness of the general solution is reduced to

$$\hat{\nu}_4\leftrightarrow\{F_k\}, \quad \hat{\nu}_3\leftrightarrow\{G_k\}, \quad \hat{\nu}_2\leftrightarrow\{H_k\}, \quad \hat{\nu}_1\leftrightarrow\{L_k\}. \tag{39}$$

In the case of QED it takes a very simple form because there we have the relations

$$\alpha_B\mu^{-\epsilon}/\lambda_B=\alpha/\lambda, \quad \gamma_\lambda=(\lambda/\alpha)\beta, \tag{40}$$

i.e., α/λ does not depend on μ and verify (38). Therefore, here $\hat{\nu}$ can be an arbitrary function of α/λ and indeed

$$\begin{aligned} &F_k(\lambda)=F_k/\lambda^{k-1}, \quad H_k(\lambda)=H_k/\lambda^k,\\ &L_k(\lambda)=L_k/\lambda^k. \end{aligned} \tag{41}$$

In QCD one gets analogously

$$\begin{aligned} &F_k(\lambda)=F_k(\lambda+d/c)^{-(k-1)(b/c)},\\ &G_k(\lambda)=G_k(\lambda+d/c)^{-k(b/c)}\cdot(c+d/\lambda),\\ &H_k(\lambda)=H_k(\lambda+d/c)^{-k(b/c)},\\ &L_k(\lambda)=L_k(\lambda+d/c)^{-k(b/c)}, \end{aligned} \tag{42}$$

where

$$c=\tfrac{13}{24}C_2(G)-\tfrac{1}{3}T(R)\cdot NF, \quad d=-\tfrac{1}{8}C_2(G). \tag{43}$$

VII. THE GAUGE DEPENDENCE

Here we study the gauge dependence of the different functions. Physical and bare quantities are gauge independent. This means that

$$\frac{d\alpha_B}{d\lambda}=0=\frac{dm_B}{d\lambda}, \tag{44}$$

the derivatives being taken at fixed μ.

From (2) we find we can define two new RG functions[11]

$$\rho=\frac{d\alpha}{d\lambda}, \quad \sigma=\frac{1}{m}\frac{dm}{d\lambda}, \tag{45}$$

which admit the power expansions

$$\begin{aligned} &\rho=\rho^{(2)}(\lambda,\eta)\alpha^2+\rho^{(3)}(\lambda,\eta)\alpha^3+\cdots,\\ &\sigma=\sigma^{(1)}(\lambda,\eta)\alpha+\sigma^{(2)}(\lambda,\eta)\alpha^2+\cdots. \end{aligned} \tag{46}$$

Notice that, as is well known,

$$\rho=0=\sigma \tag{47}$$

in the MS scheme. These equations express the gauge invariance of the scheme.

In general, we can choose ρ and σ and look for the corresponding RS. To Eqs. (11)–(14) expressing the change of RS we must join the equations

$$\rho_{\rm MS} = 0 = \left(f + \alpha \frac{\partial f}{\partial \alpha}\right)\rho + \left(\alpha\eta \frac{\partial f}{\partial \eta}\right)\sigma + \alpha \frac{\partial f}{\partial \lambda}, \quad (48)$$

$$\sigma_{\rm MS} = 0 = \frac{\partial h}{\partial \alpha}\rho + \left(h + \eta \frac{\partial h}{\partial \eta}\right)\sigma + \frac{\partial h}{\partial \lambda}. \quad (49)$$

The discussion of Sec. IV can be redone. Beside Eqs. (18)–(21) we get

$$\frac{\partial f_k}{\partial \lambda} = -k\rho^{(2)} f_{k-1} - \sigma^{(1)}\eta \frac{\partial f_{k-1}}{\partial \eta} + \cdots, \quad (50)$$

$$\frac{\partial h_k}{\partial \lambda} = -(k-1)\rho^{(2)} h_{k-1} - \sigma^{(1)}\left[h_{k-1} + \eta \frac{\partial h_{k-1}}{\partial \eta}\right] + \cdots. \quad (51)$$

These equations will allow us to tell something about the λ-dependence of f and h.

However, before we discuss this point we notice that there exist constraints between ρ, σ and β, γ. The reason is that $(d/d\lambda)|_\mu$ and $d/d\mu$ are not independent derivatives. If we introduce the derivative

$$\mu D_\mu \equiv \mu \frac{d}{d\mu}\bigg|_\lambda = \mu \frac{\partial}{\partial \mu} + \bar\beta \frac{\partial}{\partial \alpha} + \bar\gamma m \frac{\partial}{\partial m}, \quad (52)$$

where

$$\bar\beta = \mu D_\mu \alpha = \mu \frac{d\alpha}{d\mu} - \mu \frac{d\lambda}{d\mu}\frac{d\alpha}{d\lambda} = \beta - \gamma_\lambda \rho, \quad (53)$$

$$\bar\gamma = \frac{\mu}{m} D_\mu m = \frac{\mu}{m}\frac{dm}{d\mu} - \mu \frac{d\lambda}{d\mu} \cdot \frac{1}{m}\frac{dm}{d\lambda} = \gamma - \gamma_\lambda \sigma, \quad (54)$$

we have

$$\left[\mu D_\mu, \frac{d}{d\lambda}\right] = 0 \quad (55)$$

and thus

$$\mu D_\mu \rho = \frac{d\bar\beta}{d\lambda}, \quad \mu D_\mu \sigma = \frac{d\bar\gamma}{d\lambda}. \quad (56)$$

Expanding each RG function in powers of α we find

$$\frac{\partial \beta^{(k+1)}}{\partial \lambda} + \eta \frac{\partial \rho^{(k+1)}}{\partial \lambda} = \cdots, \qquad \frac{\partial \gamma^{(k)}}{\partial \lambda} + \eta \frac{\partial \sigma^{(k)}}{\partial \lambda} = \cdots, \quad (57)$$

where the terms not written contain coefficients of lower orders.

Integrating Eqs. (18), (50) and (20), (51) requires the equality of the crossed derivatives. But the two constraints obtained are nothing but (57). There is no other constraint. So, there exist RS's corresponding to

$$\beta = \gamma = \gamma_\lambda = \gamma_\psi = \rho = \sigma = 0. \quad (58)$$

In those RS's, we have

$$\mu \frac{d\hat\alpha}{d\mu} = 0 = \frac{d\hat\alpha}{d\mu} \quad (59)$$

and

$$\mu \frac{d\hat m}{d\mu} = 0 = \frac{d\hat m}{d\mu}. \quad (60)$$

Here $\hat\alpha$ and $\hat m$ can be *physical* quantities.

The discussion of Sec. V can also be redone. We just have to add ρ and σ to the other RG functions. For instance, to Eq. (24) we must add

$$\boldsymbol{\rho}'(\mathbf{f}(\mathbf{x})) \# \frac{d\mathbf{f}}{d\mathbf{x}} \cdot \boldsymbol{\rho}(\mathbf{x}), \quad (61)$$

where $\boldsymbol{\rho} = (\rho, 1, \eta\sigma)$ and # means proportional. With this compact form it is obvious that the transformations of RS which do not change any RG function of a given set define a subgroup of the RG. The discussion in the MS scheme is modified by the two equations

$$\frac{\partial f}{\partial \lambda} = \frac{\partial h}{\partial \lambda} = 0 \quad (62)$$

expressing the invariance of ρ and σ ($=0$). The particular solution (2) has to be independent of λ and $\hat v_2$, $\hat v_4$ must verify

$$\frac{d\hat v_2}{d\lambda} = \frac{d\hat v_4}{d\lambda} = 0. \quad (63)$$

Consequently, these v depend only on a sequence of arbitrary *constants* and the total degeneracy is reduced to

$$\hat v_4 \leftrightarrow \{F_k\}, \quad \hat v_3 \leftrightarrow \{G_k(\lambda)\}, \quad \hat v_2 \leftrightarrow \{H_k\}, \quad \hat v_1 \leftrightarrow \{L_k(\lambda)\}. \quad (64)$$

Thus, for instance, there are still an infinite number of gauge independent RS leading to the same RG functions.

VIII. COMBINING CONSTRAINTS

Here we examine the consequences on our previous equations of several constraints.

First we require both gauge independence and regularity when $m \to 0$ or ∞ of all RG functions and all RS's. The constraints (37) and (56) are obtained but *there exists a new one*. One substitutes $k = 2$ into (16) to get

$$\eta \frac{\partial f_2}{\partial \eta} = \gamma^{(1)} \eta \frac{\partial f_1}{\partial \eta} + \gamma_\lambda^{(1)} \frac{\partial f_1}{\partial \lambda} + 2f_2(\beta^{(2)} - b^{(2)}) + \beta^{(3)} - b^{(3)}. \quad (65)$$

Since $\beta^{(2)} \to b^{(2)}$ and $\eta(\partial f_{1,2}/\partial\eta) \to 0$ when $m \to 0$ we find

$$\gamma_\lambda^{(1)} \frac{\partial f_1}{\partial \lambda} = b^{(3)} - \beta^{(3)}, \quad (66)$$

but from (50) we also get

$$\frac{\partial f_1}{\partial \lambda} = -\rho^{(2)}. \quad (67)$$

Therefore, when $m \to 0$, we get

$$\beta^{(3)} - \gamma_\lambda^{(1)} \rho^{(2)} = b^{(3)}. \quad (68)$$

Several solutions to Eqs. (11)–(14), (48), and (49) still exist. Indeed, one can calculate the degeneracy using the general decomposition. Solution (2) has to be the identity since it cannot be simultaneously independent of λ and regular if $m \to 0$. Here $\hat v_4$ has to verify

$$\eta \frac{\partial \hat v_4}{\partial \eta} \to 0, \quad \mu \frac{d\hat v_4}{d\mu} = 0, \quad \frac{d\hat v_4}{d\lambda} = 0. \quad (69)$$

Only a constant can do that; $\hat{v}_4 = \xi$; $\hat{v}_3$ and $\hat{v}_1$ verify the same equations as in Sec. VI. They do not depend on η and λ. The final degeneracy is thus

$$\hat{v}_4 \leftrightarrow \xi, \quad \hat{v}_3 \leftrightarrow \{G_k\}, \quad \hat{v}_1 \leftrightarrow \{F_k\}. \tag{70}$$

The arbitrariness above cannot be eliminated with boundary conditions compatible with the massless case. The reason is that there do not exist RG functions associated to Z_2 and Z_λ to restrict further $\hat{v}_3$ and $\hat{v}_1$.

But we can look for mass-dependent RS's satisfying both regularity for $m \to 0$ and decoupling for $m \to \infty$. Here MOM is such a scheme.[12] With these constraints the arbitrariness *disappears*. Indeed, there exist no transformations (regular if $m \to 0$ and ∞) other than **1** leaving $\boldsymbol{\beta}_{\text{MOM}}$ and $\gamma_{\psi\text{MOM}}$ invariant in *non*-Abelian theory. This can be seen from Eqs. (17)–(21) with all RG functions replaced by those of MOM.[13]

$$\textit{First order: } \eta \frac{\partial f_1}{\partial \eta} = 0 \text{ (idem for } g,h,l); \tag{71}$$

$$\textit{Second order: } \frac{\partial f_1}{\partial \lambda} = 0; \tag{72}$$

$$\left\{1 + \lambda \frac{b^{(2)}}{C_\lambda^{(1)}} - \frac{\lambda}{C_\lambda^{(1)}} \frac{\partial C_\lambda^{(1)}}{\partial \lambda}\right\} g_1 + \lambda \frac{\partial g_1}{\partial \lambda} = f_1, \tag{73}$$

$$\left\{1 + \lambda \frac{\underline{b}^{(2)}}{\underline{C}_\lambda^{(1)}} - \frac{\lambda}{\underline{C}_\lambda^{(1)}} \frac{\partial \underline{C}_\lambda^{(1)}}{\partial \lambda}\right\} g_1 + \lambda \frac{\partial g_1}{\partial \lambda} = f_1, \tag{74}$$

where

$$b^{(2)} = \lim_{\eta \to 0} \beta^{(2)}_{\text{MOM}}, \quad \underline{b}^{(2)} = \lim_{\eta \to \infty} \beta^{(2)}_{\text{MOM}}. \tag{75}$$

From (73) and (74) in a non-Abelian theory, one deduces

$$f_1 = g_1 = 0. \tag{76}$$

Similarly one can show that $l_1 = 0$ and recursively that all f_k, g_k, h_k, l_k are equal to zero.

Thus there exists only *one* regular and decoupling RS corresponding to given $\boldsymbol{\beta}$, γ_ψ functions. Its functions ρ and σ are calculable, so that they cannot be chosen. Consequently, *there exists no regular, decoupling, and gauge independent scheme.*

Of course, the $\boldsymbol{\beta}$, γ_ψ functions of a regular decoupling scheme are constrained. They have to satisfy

$$\beta^{(2)} \to b^{(2)}, \quad \gamma^{(1)} \to C^{(1)}, \quad \gamma_\lambda^{(1)} \to C_\lambda^{(1)}, \quad \gamma_\psi^{(1)} \to C_\psi^{(1)}, \tag{77}$$

if $\eta \to 0$, and

$$\beta^{(2)} \to \underline{b}^{(2)}, \quad \gamma^{(1)} \to \underline{C}^{(1)}, \quad \gamma_\lambda^{(1)} \to \underline{C}_\lambda^{(1)}, \quad \gamma_\psi^{(1)} \to \underline{C}_\psi^{(1)}, \tag{78}$$

if $\eta \to \infty$. But they are not the only constraints: there exist to each order a new constraint on the integral of each function ($\boldsymbol{\beta}, \gamma_\psi$) with respect to η as a function of λ. The RS functions of MOM satisfy of course all these constraints but there are many other solutions, so that MOM is not the only regular decoupling scheme. The gauge dependence of these functions cannot be chosen and this is why ρ and σ cannot be put equal to zero.

IX. CONCLUSIONS

We have given in our work a general description of the RS transformations. We have been able to introduce a group structure which makes transparent the correspondence between RS's and RG transformations.

Our main results are that (1) there are an infinite number of RS corresponding to a given set of RG functions; each set of RG functions defines a class of RS's; the degeneracy of all classes is the *same* and *is equal to the one of the subgroup of the RG transformations* which leave the set of MS-RG functions invariant; (2) the degree of arbitrariness can be lowered by adding new RG functions or imposing boundary conditions; in this last case, constraints appear on the set of RG functions; (3) to eliminate any degeneracy one can impose both regularity and decoupling properties, but the RG functions have to satisfy an interpolation constraint at *each* order (so that we cannot parametrize them); and (4) there exists no regular, decoupling, gauge invariant scheme in a non-Abelian theory.

APPENDIX: THE BEHAVIOR OF $\hat{v}$ UNDER VARIOUS BOUNDARY CONDITIONS

We show how one can obtain the expression of $\hat{v}$ under the various boundary conditions discussed in the text.

We start from Eq. (29) for $\hat{v}$, i.e.,

$$\mu \frac{d\hat{v}}{d\mu} = 0 \tag{A1}$$

or

$$\beta_{\text{MS}} \frac{\partial \hat{v}}{\partial \alpha} + \Gamma_{\text{MS}} \frac{\partial \hat{v}}{\partial \eta} + \gamma_{\lambda\,\text{MS}} \frac{\partial \hat{v}}{\partial \lambda} = 0. \tag{A2}$$

Writing

$$\hat{v} = \hat{v}^{(0)} + \hat{v}^{(1)} \alpha + \cdots, \tag{A3}$$

we get the equations

$$\eta \frac{\partial \hat{v}^{(0)}}{\partial \eta} = 0, \tag{A4}$$

$$\eta \frac{\partial \hat{v}^{(k)}}{\partial \eta} = b^{(2)}(k-1)\hat{v}^{(k-1)} + C_\lambda^{(1)} \frac{\partial \hat{v}^{(k-1)}}{\partial \lambda} + C^{(1)} \eta \frac{\partial \hat{v}^{(k-1)}}{\partial \eta} + \cdots. \tag{A5}$$

The solution is constructed recursively and depends on an arbitrary function $R_k(\lambda)$ to each order k.

We introduce the various boundary conditions, (1) $\hat{v}$ is regular when $\eta \to 0$,

$$R_k(\lambda) = R_k \left[\exp - \left(\int \frac{b^{(2)}}{C_\lambda^{(1)}(\lambda)} d\lambda\right)\right]^k; \tag{A6}$$

$$(2)\ \partial \hat{v}/\partial \lambda = 0, \quad R_k(\lambda) \equiv R_k; \tag{A7}$$

(3) the above two boundary conditions are simultaneously fulfilled. Equation (A2) in the limit $m \to 0$ implies that $\hat{v}$ for $\eta \to 0$ is just an arbitrary constant. The η-dependence is fixed by (A5), so $\hat{v}$ is equal for all η to that arbitrary constant.

[1] A. Peterman, Phys. Rep. C **53**, 157 (1979).
[2] P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981); Nucl. Phys. B **203**, 472 (1982).

[3]P. Pascual and R. Tarrach, "QCD: Renormalization for the practitioner," *Lecture Notes in Physics*, Vol. 194 (Springer, Berlin, 1984).
[4]M. Hans, Lett. Nuovo Cimento **41**, 592 (1985).
[5]G.'t Hooft and M. Veltman, Nucl. Phys. B **44**, 184 (1972).
[6]M. Astaud and B. Jouvet, Nuovo Cimento A **63**, 5 (1969).
[7]Because, if some function f, g, h, or l is equal to 1, it stays so.
[8]N. K. Nielsen, Nucl. Phys. B **124**, 109 (1977).
[9]Alternatively, we can require the Green's functions to be regular when $m \to \infty$ (decoupling property); B. A. Ovrut and H. J. Schnitzer, Nucl. Phys. B **179**, 381 (1981); **189**, 509 (1981); W. Bemreuther, Z. Phys. C **20**, 331 (1983).
[10]If $m \to \infty$, $b^{(2)},...,C_\psi^{(1)}$ must be replaced by the RG functions of the MS scheme without the massive particles ($\underline{b}^{(2)},...,\underline{C}_\psi^{(1)}$).
[11]D. J. Gross, in *Methods of Field Theory, Les Houches 1975* (North-Holland, Amsterdam, 1976).
[12]The momentum subtraction scheme (MOM) has been discussed by T. Applequist and J. Carrazone, Phys. Rev. D **11**, 2856 (1975).
[13]The mass dependence of β modifies these equations with terms like $\eta(\partial\beta^{(2)}/\partial\eta)\,h_{k-1}$ but it is not important here.

Quantum kinematics of strings: Quantization on $C^\infty(M^3, \mathbb{R}^2)/C^\infty(M^3, \mathrm{GL}(2,\mathbb{R}))$

László Szabó
Institute for Theoretical Physics, Eötvös University, Budapest, Hungary

(Received 31 December 1986; accepted for publication 15 April 1987)

Global quantum kinematics is discussed on the configuration manifold $C^\infty(M^3,\mathbb{R}^2)/C^\infty(M^3,\mathrm{GL}(2,\mathbb{R}))$ with respect to the possible relevance to the string theory.

I. INTRODUCTION

The configuration space for a string (particle) theory is

$$Q = \mathrm{Emb}(\Sigma,M)/\mathrm{Diff}(\Sigma),$$

where M is a three-dimensional Riemann space and Σ is a one-dimensional manifold topologically equivalent to S^1 or $\mathbb{R}^1$ dependently on whether the string is closed or open. As it is known, this space is an infinite-dimensional locally convex manifold.[1] Because of the complicated global structure of this space the *correct canonical commutation relations are not the Heisenberg relations* associated with a linear configuration space.

The straightforward generalization of canonical quantization for an arbitrary (nonlinear) configuration manifold needs a so-called canonical group G, which is a connected Lie group acting symplectically, transitively, and effectively. This group is destined to play the same role as the Weyl group in the standard quantization procedure on $\mathbb{R}^n$ and the Lie algebra $L(G)$ defines the basic commutation relations.[2] One can derive particular Hilbert space representations of the quantized system through the unitary representations of the canonical group. Unfortunately a correct canonical group for $\mathrm{Emb}(\Sigma,M)/\mathrm{Diff}(\Sigma)$ has not been found yet.

II. ENLARGED CONFIGURATION MANIFOLD

We enlarge the configuration manifold such that it has a more simple structure and the quantization procedure can be fulfilled.

If a function $f \in C^\infty(M,\mathbb{R}^2)$ is transversal to $0 \in \mathbb{R}^2$, i.e., $0 \in \mathbb{R}^2$ is not a critical value of f, then $f^{-1}(0) \subset M$ is a submanifold. This fact would provide a parametrization of all one-dimensional submanifolds in M. Unfortunately the space of smooth transversal functions is not even a manifold.

Let us cancel the condition of transversality and suppose that the configuration space is the whole $C^\infty(M,\mathbb{R}^2)$. This enlarged space contains configurations corresponding to various subsets in M. One can interpret the one-dimensional objects as open or closed strings and the zero-dimensional subsets as random heaps of pointlike particles.

Two functions f and g mean the same physical configuration iff $f^{-1}(0) = g^{-1}(0)$, that is we can introduce a gauge symmetry group $C^\infty(M,\mathrm{GL}(2,\mathbb{R}))$ as follows: $f \sim g$ iff there exists $\eta \in C^\infty(M,\mathrm{GL}(2,\mathbb{R}))$ such that

$$f(x) = \eta(x)\cdot g(x).$$

III. QUANTIZATION ON $C^\infty(M,\mathbb{R}^2)$

For quantization one needs a canonical group $\mathscr{G}$ and a momentum map P: $L(\mathscr{G}) \mapsto C^\infty(T^*Q,\mathbb{R})$ which causes the following diagram to commute[2]:

$$\begin{array}{ccc} C^\infty(T^*Q,\mathbb{R}) & \longrightarrow & \mathrm{Hamilton}(T^*Q) \\ & & \mathrm{Fund}(A) \\ P \uparrow & & \uparrow \updownarrow A \\ & L(\mathscr{G}) & \end{array} \tag{1}$$

One can verify that a good canonical group is

$$\mathscr{G} = C^\infty[M, L^*(A(2,\mathbb{R})) \otimes_s A(2,\mathbb{R})],$$

where $A(2,\mathbb{R})$ denotes the affine group over $\mathbb{R}^2$. The following group action is transitive and effective:

$$[\ell_{\{(\eta,Y),(a,A)\}}(\omega,r)](p) = \{\ell^{*}_{(a(p),A(p))}{}^{-1}[\omega(p) + df^{(\eta(p),Y(p))}_{r(p)}(p)],\ \ell_{(a(p),A(p))}r(p)\},$$

where $(\omega,r) \in T^*(C^\infty(M,\mathbb{R}^2))$, $p \in m$ and

$$f^{(\eta,Y)}:\ r \in \mathbb{R}^2 \mapsto \sum_\sigma Y^h_\sigma r_b + \eta_\sigma.$$

It is also symplectic since it is a composition of a diffeomorphism on the configuration manifold and an additive translation along the fibers by a closed one-form.[3] The momentum function[2,3] is defined as follows:

$$\begin{aligned} P_{(\epsilon,X)}&:\ (\omega,r) \mapsto \int_M (\epsilon_a(p) + X^b_a(p) r_b(p)\omega^a(p))dp, \\ P_{(\eta,Y)}&:\ (\omega,r) \mapsto \int_M f^{(\eta(p),Y(p))}(r(p))dp, \end{aligned} \tag{2}$$

where $(\epsilon,X) \in T_{(0,1)}C^\infty(M,A(2,\mathbb{R}))$ and $(\eta,Y) \in T_{(0,0)}C^\infty(M, L^*(A(2,\mathbb{R})))$.

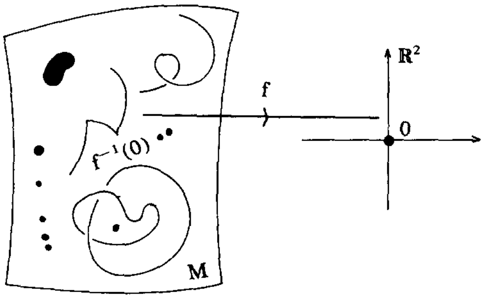


FIG. 1. The inverse image $f^{-1}(0)$ corresponds to a collection of pointlike particles, closed and open strings, and higher-dimensional heaps.

 0022-2488/87/092220-03$02.50

Let us compute the Poisson brackets to show that diagram (1) is commutative,

$$\{P_{(\epsilon,X)},P_{(\tilde{\epsilon},\tilde{X})}\}$$
$$=\int_M [X^s_a(p)\omega^a(p)(\tilde{\epsilon}(p)+\tilde{X}^b_s(p)r_b(p)) - (\epsilon_s(p)+X^b_s(p)r_b(p))\tilde{X}^s_a(p)\omega^a(p)]dp$$
$$=\int_M [(X^s_a(p)\tilde{\epsilon}_s(p)-(\tilde{X}^s_a(p)\epsilon_s(p)) + [X,\tilde{X}]^b_a(p)r_b(p)\omega^a(p)]dp = P_{[(\epsilon,X),(\tilde{\epsilon},\tilde{X})]},$$
$$\{P_{(\eta,Y)},P_{(\tilde{\eta},\tilde{Y})}\}=0,$$
$$\{P_{(\epsilon,X)},P_{(\eta,Y)}\}$$
$$=\int_M\Big[\sum_\sigma(\epsilon_s(p)+X^b_s(p)r_b(p))Y^s_\sigma(p)\Big]dp$$
$$=\int_M\Big[\sum_\sigma\epsilon_s(p)Y^s_\sigma(p)+Y^s_\sigma(p)X^b_s(p)r_b(p)\Big]dp$$
$$=P_{[(\epsilon,X),(\eta,Y)]},$$

where we used the following Lie product:

$$[\{(\eta,Y),\ (\epsilon,X)\},\ \{(\tilde{\eta},\tilde{Y}),\ (\tilde{\epsilon},\tilde{X})\}]$$
$$=\{(\eta,Y)(\tilde{\epsilon},\tilde{X})-(\tilde{\eta},\tilde{Y})(\epsilon,X),\ [(\epsilon,X),\ (\tilde{\epsilon},\tilde{X})]\}.$$

IV. PARTICULAR HILBERT SPACE REPRESENTATION

Consider the ultralocal unitary representations[2,4] of canonical group $\mathscr{G}$. First we study the unitary representations τ of $L^*(A(2,\mathbb{R}))\otimes_s A(2,\mathbb{R})$ in a Hilbert space H. The ultralocal representation is given in the direct integral space

$$\mathscr{H}=\int_M^\otimes H_p\,dp$$

as follows:

$$[T(\tilde{g})\Psi](p)=\tau(\tilde{g}(p))\Psi(p),\quad \Psi\in\mathscr{H},\quad \tilde{g}\in\mathscr{G}.$$

One can use Mackey's technique[5] for semidirect product group $L^*(A(2,\mathbb{R}))\otimes_s A(2,\mathbb{R})$. The character group of $L^*(A(2,\mathbb{R}))$ is

$$\chi_{(\epsilon,X)}(\eta,Y)=\exp i\,\mathrm{Tr}\left\{\begin{bmatrix}0&0\\ \eta&Y\end{bmatrix}\begin{bmatrix}0&0\\ \epsilon&X\end{bmatrix}\right\}.$$

The natural $A(2,\mathbb{R})$ action on char$\{L^*(A(2,\mathbb{R}))\}$ is

$$\chi_{(\epsilon,X)}\overset{\ell_{(a,A)}}{\longmapsto}\chi_{(\epsilon,X)(-A^{-1}a,A^{-1})}.$$

The orbits in char$\{L^*(A(2,\mathbb{R}))\}$ are

$$\{\chi_{(\epsilon+Xa,XA)}\}_{(a,A)\in A(2,\mathbb{R})}.$$

One can classify the orbits according to the rank of matrix X. By way of example suppose that it has maximal rank. Thus the orbit is $\mathbb{R}^2\times\mathrm{GL}(2,\mathbb{R})$. The affine group $A(2,\mathbb{R})$ acts freely and the little group is trivial.

According to the Mackey theorem the space of representation is $\mathscr{L}^2(\mathbb{R}^2\times\mathrm{GL}(2,\mathbb{R}),\ dg)$, where dg denotes the induced Haar measure. The representation of $L^*(A(2,\mathbb{R}))\otimes_s A(2,\mathbb{R})$ is

$$[\tau_{\{(\epsilon,X),(a,A)\}}f](\eta,Y)$$
$$=\exp i\,\mathrm{Tr}\left\{\begin{bmatrix}0&0\\ \eta&Y\end{bmatrix}\begin{bmatrix}0&0\\ \epsilon&X\end{bmatrix}\right\}f((\eta,Y)(a,A)).$$

Finally the space of ultralocal representation is

$$\mathscr{H}=\int_M^\otimes[\mathscr{L}^2(\mathbb{R}^2\times\mathrm{GL}(2,\mathbb{R}),\ dg)]_p\,dp$$
$$=\mathscr{L}^2(M\times\mathbb{R}^2\times\mathrm{GL}(2,\mathbb{R}),\ dg\otimes dp).$$

The representation is defined as

$$[T((\epsilon,X),(a,A))\Psi](p,(\eta,Y))$$
$$=\exp i\,\mathrm{Tr}\left\{\begin{bmatrix}0&0\\ \eta&Y\end{bmatrix}\begin{bmatrix}0&0\\ \epsilon(p)&Y(p)\end{bmatrix}\right\}\times\Psi[p,(\eta,Y)(a(p),A(p))].\qquad(3)$$

V. DIRAC CONSTRAINT EQUATIONS

Fortunately the gauge group $C^\infty(M,\mathrm{GL}(2,\mathbb{R}))$ is a subgroup of canonical group $C^\infty[M,L^*(A(2,\mathbb{R}))\otimes_s A(2,\mathbb{R})]$, therefore we have a natural representation of the gauge group in the Hilbert space of states.

Consider the Lie algebra element

$$\begin{bmatrix}A_{11}(x)&A_{12}(x)\\ A_{21}(x)&A_{22}(x)\end{bmatrix}\in L\,[C^\infty(M,\mathrm{GL}(2,\mathbb{R}))].$$

From (2) we have

$$[\hat{A}\Psi]\left[x,\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix},\begin{bmatrix}Y_{12}&Y_{12}\\ Y_{21}&Y_{22}\end{bmatrix}\right]$$
$$=\Big[[Y_{11}A_{11}+Y_{12}A_{21}]\frac{\partial}{\partial Y_{11}}+[Y_{11}A_{12}+Y_{12}A_{22}]\frac{\partial}{\partial Y_{12}}+[Y_{21}A_{11}+Y_{22}A_{21}]\frac{\partial}{\partial Y_{21}}+[Y_{21}A_{12}+Y_{22}A_{22}]\frac{\partial}{\partial Y_{22}}\Big]\Psi,$$

where

$$\left[x,\begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix},\begin{bmatrix}Y_{11}&Y_{12}\\ Y_{21}&Y_{22}\end{bmatrix}\right]$$

are coordinates on manifold $M\times\mathbb{R}^2\times\mathrm{GL}(2,\mathbb{R})$.

The Dirac constraint condition[6] requires that

$$\hat{A}\Psi=0$$

for any infinitesimal gauge transformation. The four independent Dirac constraint equations are

$$\left[Y_{11}\frac{\partial}{\partial Y_{11}}+Y_{21}\frac{\partial}{\partial Y_{21}}\right]\Psi=0,$$
$$\left[Y_{11}\frac{\partial}{\partial Y_{12}}+Y_{21}\frac{\partial}{\partial Y_{22}}\right]\Psi=0,$$
$$\left[Y_{12}\frac{\partial}{\partial Y_{11}}+Y_{22}\frac{\partial}{\partial Y_{21}}\right]\Psi=0,$$
$$\left[Y_{12}\frac{\partial}{\partial Y_{12}}+Y_{22}\frac{\partial}{\partial Y_{22}}\right]\Psi=0.$$

VI. OPERATORS OF GENERALIZED COORDINATES AND MOMENTUMS

From the momentum map (2) we know that the following Lie algebra elements correspond to the coordinates of configuration manifold:

$$Z_1:\ x\in M\mapsto\left[\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1&0\\0&0\end{bmatrix}\right]\in T_{(0,0)}C^\infty\,[M,L^*(A(2,\mathbb{R}))],$$

$$Z_2:\ x\in M\mapsto\left[\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix}\right]\in T_{(0,0)}C^\infty\,[M,L^*(A(2,\mathbb{R}))].$$

The self-adjont generators corresponding to these Lie algebra elements are determined by representation (3),

$$[\hat{f}_1\Psi]\left[x,\begin{bmatrix}\eta_1\\\eta_2\end{bmatrix},\begin{bmatrix}Y_{11}&Y_{12}\\Y_{21}&Y_{22}\end{bmatrix}\right]=Y_{11}\Psi\left[x,\begin{bmatrix}\eta_1\\\eta_2\end{bmatrix},\begin{bmatrix}Y_{11}&Y_{12}\\Y_{21}&Y_{22}\end{bmatrix}\right],$$

$$[\hat{f}_2\Psi]\left[x,\begin{bmatrix}\eta_1\\\eta_2\end{bmatrix},\begin{bmatrix}Y_{11}&Y_{12}\\Y_{21}&Y_{22}\end{bmatrix}\right]=Y_{21}\Psi\left[x,\begin{bmatrix}\eta_1\\\eta_2\end{bmatrix},\begin{bmatrix}Y_{11}&Y_{12}\\Y_{21}&Y_{22}\end{bmatrix}\right].$$

In the same way from (2) we know that Lie algebra elements

$$\left.\begin{array}{l}W_1:\ x\in M\mapsto\left[\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0&0\\0&0\end{bmatrix}\right]\\ W_2:\ x\in M\mapsto\left[\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}0&0\\0&0\end{bmatrix}\right]\end{array}\right\}\in T_{(0,1)}C^\infty\,(M,A(2,\mathbb{R}))$$

correspond to the two generalized momentums. One can compute from (3) the corresponding self-adjont generators

$$\hat{\omega}_1\Psi=-i\left[Y_{11}\frac{\partial\Psi}{\partial\eta_1}+Y_{21}\frac{\partial\Psi}{\partial\eta_2}\right],$$

$$\hat{\omega}_2\Psi=-i\left[Y_{12}\frac{\partial\Psi}{\partial\eta_1}+Y_{22}\frac{\partial\Psi}{\partial\eta_2}\right].$$

VII. CONCLUSIONS

We suggested describing processes of open and closed strings and pointlike particles together in an enlarged configuration amphitheater, which contains other higher-dimensional heaps too, without any interpretation so far. We have taken the initial steps in quantum kinematics on this configuration manifold. Of course there is some arbitrariness in the quantization procedure as for example the choice of canonical group and that of the definition of momentum map, etc. The main motivation of our choice was the analogy to quantization on manifold $C^\infty(M,\mathrm{GL}^+(3,\mathbb{R}))$, which is the configuration manifold of gravity.[2,4]

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Z. Horváth and L. Palla for many helpful discussions.

[1]P. W. Michor, *Manifold of Differentiable Mappings* (Shiva, Orpington, Kent, 1980).

[2]C. J. Isham, "Topological and global aspects of quantum theory," in *Relativity, Groups and Topology II* (North-Holland, Amsterdam, 1984).

[3]R. Abraham and J. E. Marsden, *Foundation of Mechanics* (Benjamin, London, 1978), 2nd ed.

[4]C. J. Isham, "Quantum geometry," in *Quantum Theory of Gravity—A Collection of Essays in Honour of Bryce De Witt's Sixtieth Birthday*, edited by S. Cristensen (Hilger, Bristol, 1984).

[5]G. W. Mackey, *Unitary Group Representations in Physics, Probability and Number Theory* (Benjamin, London, 1978).

[6]K. Sundermayer, "Constrained dynamics," *Lecture Notes in Physics*, Vol. 169 (Springer, New York, 1982).

Collectivity and geometry. V. Spectra and shapes in the two-dimensional case

E. Chacón
Instituto de Física, Universidad Nacional Autonoma de México, Apdo. Postal 20-364, México, D.F. 01000 Mexico

P. O. Hess[a)]
Centro de Estudios Nucleares, Universidad Nacional Autonoma de México, Apdo. Postal 70-543, México, D.F. 04510 Mexico

M. Moshinsky[b)]
Instituto de Fisico, Universidad Nacional Autonoma de México, Apdo. Postal 20-364, México, D.F. 01000 Mexico

(Received 13 June 1986; accepted for publication 30 January 1987)

In the present paper a program for a collective symplectic model in a two-dimensional space based on the sp(4,R) Lie algebra is developed in full. The problem is only of conceptual interest but, as it admits a complete analytic discussion, it provides insights into more realistic collective nuclear models such as the interacting boson approximation (IBA) associated with u(6) and the symplectic model based on sp(6,R).

I. INTRODUCTION AND SUMMARY

In the series of papers with the same general title,[1,2] as well as in other publications[3–5] the authors and their collaborators have analyzed the matrix representation of Hamiltonians in the enveloping algebra of the symplectic groups Sp(2d,R), where d is an integer, with respect to states characterized by irreducible representations (irreps) in the positive discrete series for these groups. The discussion required the matrix representation of the generators of the symplectic Lie algebra sp(2d,R) with respect to the states mentioned, where in turn these generators can be expressed as functions of the coordinates and momenta, or, equivalently, the creation and annihilation operators of an A-body system.

The problem of physical interest in this program corresponds to $d = 3$, i.e., sp(6,R), giving rise to what is known as the symplectic model of the nucleus.[6–12] While a lot of work has been done in this model[1,3,6–12] we feel that a systematic discussion, covering the physical implications of all the subgroup chains of Sp(6,R), has not been achieved, in part because of the complexities associated with the sp(6,R) Lie algebra. Thus the authors and their collaborators have paid considerable attention[2,4,5] to the corresponding problem in a two- rather than a three-dimensional space, i.e., to sp(4,R), in the hope of clarifying there the main conceptual structures before passing to sp(6,R). Thus in the present paper we give a systematic discussion of spectra and shapes for Hamiltonians in the enveloping algebra of sp(4,R).

Our analysis will proceed along the following lines. In Sec. II we indicate that there are seven maximal subalgebras[13] of sp(4,R) of which only five that contain the angular momentum J_0 in a two-dimensional space are of interest to us. We give the generators of these five maximal subalgebras as well as their Casimir operators. This discussion parallels the one in the interacting boson approximation (IBA),[14,15] where the maximal subalgebras of u(6) containing o(3) are u(5), su(3), and o(6).

In Sec. III we discuss the basis states classified by irreps of the maximal subalgebras, and show that the ones associated with the direct sum of two-dimensional symplectic algebras, i.e., sp′(2,R) ⊕ sp″(2,R), do not contain multiplicity indices, and thus can provide the orthonormal basis which we derive explicitly. We determine the matrix elements of the generators of sp(4,R) in this basis and use them in all of the following discussion. A different derivation of these matrix elements was given previously by Hecht and Peterson.[16]

In Sec. IV we discuss collective Hamiltonians in the enveloping algebra of sp(4,R) for the A-body system in a two-dimensional space. We introduce the Jacobi relative coordinates x_{is} and momenta p_{is}, $i = 1,2$, $s = 1,2,...,n = A - 1$, suppressing, as usual,[1] those associated with the center of mass motion. We then consider the corresponding creation and annihilation operators and express the generators of sp(4,R) in terms of them, discussing the behavior of these generators under Hermitian conjugation, rotation, and time reflection. The Hamiltonians we shall consider will be Hermitian polynomials of up to second order in the generators of sp(4,R) that are invariant under rotation (i.e., commute with J_0) and time reflection. We show that these Hamiltonians can be expressed in terms of Casimir operators of maximal subalgebras of sp(4,R) as well as powers of the generators of their Abelian subalgebras. Thus we have Hamiltonians associated with specific chains of subalgebras whose spectra can be given in closed form, as well as others, which we could call transitional, that involve more than one chain and whose spectra has to be calculated numerically. We shall give some examples of the latter spectra in Sec. VI, using the basis characterized by the irreps of sp′(2,R) ⊕ sp″(2,R) as indicated in Sec. III. This program parallels the one of the IBA,[14,15] where the specific chains of subalgebras were u(5), su(3), and o(6).

In Sec. V we turn our attention to the problem of shape of a many-body system in two-dimensional space. We define, in terms of the coordinates x_{is}, the quadrupole tensor, which gives us a measure for the deformation of an A-body system.

[a)] Fellow of the Deutscher Akademischer Austauschdienst.
[b)] Member of El Colegio Nacional.

0022-2488/87/092223-18$02.50

By passing from x_{is} to the Dzublik[17]–Zickendraht[18] system that gives us the coordiantes associated with the intrinsic quadrupole moment, we can define shape operators[19] and obtain their matrix elements with respect to the basis states discussed in Sec. III. In turn from these matrix elements we can obtain the shape parameter β or eccentricity ϵ for eigenstates of definite Hamiltonians.

In Sec. VI we carry out some calculations of conceptual interest. First we discuss the shape of states associated with specific chains of subalgebras both when, from the standpoint of the oscillator shell model, we have closed or half-open shells, as well as for different number of quanta of excitation. We then analyze "transitional nuclei," i.e., those with Hamiltonians in which we have linear combinations of Casimir operators and powers of weight generators of different chains of subalgebras.

Again this is discussed for closed and half-open shells and, when relevant, for different number of quanta of excitation. We also consider the spectra and shape of a Hamiltonian not discussed in Sec. III in which the quadrupole operator components go up to fourth order [which is thus also of the same order in the generators of sp(4,R)]. This Hamiltonian is associated with the strongly deformed potentials discussed by a number of authors.[9,20]

Finally, in the concluding section we review critically all of the previous results and consider their implications for the different collective models such as Bohr–Mottelson, IBA, symplectic, etc.

II. MAXIMAL SUBALGEBRAS OF sp(4,R), THEIR GENERATORS AND CASIMIR OPERATORS

The ten generators of sp(4,R), when written in vector form with spherical components,[5] can be denoted by

$$\mathscr{N}, B_q^\dagger, J_q, B_q, \quad q = 1,0,-1, \tag{2.1}$$

where the raising of the index in B_q is given by the standard rule

$$B^q = (-)^q B_{-q} \tag{2.2}$$

and similarly for $B_q^\dagger, J_q$.

The commutation relations for these generators are

$$[\mathscr{N}, B_q^\dagger] = B_q^\dagger, \tag{2.3a}$$
$$[\mathscr{N}, B_q] = -B_q, \tag{2.3b}$$
$$[\mathscr{N}, J_q] = 0, \tag{2.3c}$$
$$[J_{q''}, B_{q'}] = \varepsilon_{q''q'q} B^q,$$
$$[J_{q''}, J_{q'}] = \varepsilon_{q''q'q} J^q, \quad [J_{q''}, B_{q'}^\dagger] = \varepsilon_{q''q'q} (B^\dagger)^q,$$
$$[B_q, B_{q'}] = 0, \quad [B_q^\dagger, B_{q'}^\dagger] = 0,$$
$$[B_{q''}, B_{q'}^\dagger] = -2\varepsilon_{q''q'q} J^q + (-)^{q''} 2\mathscr{N}\delta_{q'',-q'},$$

where repeated indices q are summed over the values $1,0,-1$ while $\varepsilon_{q''q'q}$ is the antisymmetric tensor in the indices indicated.

Following the type of analysis usual in IBA[14,15] we want now to determine maximal subalgebras (i.e., those that do not admit any other subalgebra between them and the full algebra) of sp(4,R) that contain in turn the o(2) subalgebra, i.e., the generator J_0, associated with the angular momentum in this two-dimensional problem. The maximal subalgebras have been discussed in other publications[13] where it was shown that there were seven of them. Five of these contain J_0 and we proceed to give their generators, the commutation relation for them that follow from (2.3), and the Casimir operators.

A. The sp′(2,R)⊕sp″(2,R) subalgebra

The six generators are I'_q, I''_q, $q = 1,0,-1$, given by

$$I'_1 = \tfrac{1}{2}B_1^\dagger, \tag{2.4a}$$
$$I'_0 = \tfrac{1}{2}(\mathscr{N} + J_0), \tag{2.4b}$$
$$I'_{-1} = \tfrac{1}{2}B_{-1}, \tag{2.4c}$$
$$I''_1 = \tfrac{1}{2}B_{-1}^\dagger, \tag{2.4d}$$
$$I''_0 = \tfrac{1}{2}(\mathscr{N} - J_0), \tag{2.4e}$$
$$I''_{-1} = \tfrac{1}{2}B_1, \tag{2.4f}$$

where from (2.3) their commutation relations are given by

$$[I''_{q'}, I''_{q''}] = 0, \tag{2.5a}$$
$$[I'_{-1}, I'_1] = -I'_0, \tag{2.5b}$$
$$[I'_0, I'_{\pm 1}] = \pm I'_{\pm 1}, \tag{2.5c}$$
$$[I''_{-1}, I''_1] = -I''_0, \tag{2.5d}$$
$$[I''_0, I''_{\pm 1}] = \pm I''_{\pm 1}. \tag{2.5e}$$

The Casimir operators are clearly

$$I'^2 = I'_0(I'_0 - 1) + 2I'_1 I'_{-1}$$
$$= \tfrac{1}{4}(\mathscr{N} + J_0)(\mathscr{N} + J_0 - 2) + \tfrac{1}{2}B_1^\dagger B_{-1}, \tag{2.6a}$$
$$I''^2 = I''_0(I''_0 - 1) + 2I''_1 I''_{-1}$$
$$= \tfrac{1}{4}(\mathscr{N} - J_0)(\mathscr{N} - J_0 - 2) + \tfrac{1}{2}B_{-1}^\dagger B_1, \tag{2.6b}$$

as from (2.3) they commute with all generators of the subalgebra.

The states in this chain can be characterized by the commuting Hermitian operators

$$I'^2, \; I''^2, \; I'_0, \; I''_0, \tag{2.7}$$

which implies from (2.4b) and (2.4f) that they are also characterized by eigenvalues of the angular momentum operator J_0 and the number operator $\mathscr{N}$ which, as will be shown in Sec. IV, is associated with the number of quanta in a harmonic oscillator Hamiltonian for the A-body system.

B. The su(2)⊕u(1) subalgebra

The three generators of su(2) are the J_q, $q = 1,0,-1$, while the single generator of u(1) is the $\mathscr{N}$, all of them appearing in (2.1). From (2.3) the commutation rules are then

$$[J_{q''}, J_{q'}] = \varepsilon_{q''q'q} J^q, \tag{2.8a}$$
$$[\mathscr{N}, J_q] = 0, \tag{2.8b}$$

while the Casimir operator of su(2) is

$$J^2 = J_0(J_0 + 1) - 2J_{-1}J_1. \tag{2.9}$$

The states in this chain can be characterized by the commuting operators

$$\mathscr{N}, \quad J^2, \quad J_0, \tag{2.10}$$

and thus comparing it with (2.7) for the sp′(2,R)

$\oplus$ sp″(2,R) chain we see that we are missing one label. As a result these states, discussed in the two papers of Ref. 4, are nonorthonormal and a complex procedure is required to determine their overlap.

C. The sp(2,R) ⊕ o(2) subalgebra

The three generators of sp(2,R) are I_q, $q = 1,0,-1$, given by

$$I_1 = -(1/\sqrt{2})B_0^\dagger, \tag{2.11a}$$

$$I_0 = \mathscr{N}, \tag{2.11b}$$

$$I_{-1} = (1/\sqrt{2})B_0, \tag{2.11c}$$

with the commutation rules

$$[I_{-1},I_1] = -I_0, \tag{2.12a}$$

$$[I_0,I_{\pm 1}] = \pm I_{\pm 1}, \tag{2.12b}$$

which follow from (2.3). The generator of o(2) is the angular momentum operator J_0 that commutes with the I_q.

The Casimir operator of sp(2,R) is given by

$$I^2 = I_0(I_0-1) + 2I_1I_{-1} = \mathscr{N}(\mathscr{N}-1) - B_0^\dagger B_0. \tag{2.13}$$

The states in this chain can be characterized by the commuting operators

$$J_0,\ I^2,\ I_0, \tag{2.14}$$

and again we see that we are missing one label, giving the nonorthonormal states discussed in Ref. 2.

D. The cm(2) subalgebra

This subalgebra is the equivalent in two dimensions of the collective motion Lie algebra cm(3) originally introduced by Biedenharn *et al.*[8] and Rowe and Rosensteel.[9] We will discuss its physical significance in more detail in Secs. V and VI, but here we want only to give its six generators (to be denoted by Q_q, R_q, $q = 1,0,-1$), their commutation rules and Casimir operators. We have then

$$Q_1 = \tfrac{1}{2}(B_1^\dagger - B_1) - J_1, \tag{2.15a}$$

$$Q_0 = \tfrac{1}{2}(B_0^\dagger + B_0) + \mathscr{N}, \tag{2.15b}$$

$$Q_{-1} = \tfrac{1}{2}(B_{-1}^\dagger - B_{-1}) + J_{-1}, \tag{2.15c}$$

$$R_1 = \tfrac{1}{2}(B_1^\dagger + B_1), \tag{2.15d}$$

$$R_0 = J_0, \tag{2.15e}$$

$$R_{-1} = \tfrac{1}{2}(B_{-1}^\dagger + B_{-1}), \tag{2.15f}$$

where the commutation rules are

$$[Q_{q'},Q_{q''}] = 0, \quad q',q'' = 1,0,-1, \tag{2.16a}$$

$$[J_0,Q_0] = 0, \tag{2.16b}$$

$$[R_1,R_{-1}] = R_0, \tag{2.16c}$$

$$[R_0,R_{\pm 1}] = \pm R_{\pm 1}, \tag{2.16d}$$

$$[R_0,Q_{\pm 1}] = \pm Q_{\pm 1}, \tag{2.16e}$$

$$[R_{\pm 1},Q_0] = -Q_{\pm 1}, \tag{2.16f}$$

$$[R_{\pm 1},Q_{\pm 1}] = 0, \tag{2.16g}$$

$$[R_{\pm 1},Q_{\mp 1}] = -Q_0, \tag{2.16h}$$

and in the last two we take either the upper or the lower signs.

From (2.16) we immediately see that the operators

$$Q^2 = Q_0^2 - 2Q_1Q_{-1}, \tag{2.17}$$

$$W = Q_1R_{-1} - Q_{-1}R_1 + Q_0R_0 \tag{2.18}$$

commute with all the generators (2.15) of cm(2) and thus are the Casimir operators of this Lie algebra.

We note that the operator $\mathscr{N}$ is not contained in the cm(2) subalgebra and thus does not commute with the Casimir operators (2.17) and (2.18). As $\mathscr{N}$ will be relevant for introducing either the oscillator Hamiltonian or the kinetic energy in the A-body problem, we will not be interested in characterizing our states by irreps of cm(2) despite the fact that

$$Q^2,\ W,\ Q_0,\ R_0, \tag{2.19}$$

give a set of four commuting operators that provide a sufficient number of labels for these states.

Note that the cm(2) subalgebra defined by (2.15) is not maximal as we can add to it $\bar{R} \equiv B_0^\dagger - B_0$ and the generators will still close under commutation. Despite this fact we will keep the definitions (2.15), to be consistent with the previous literature on this subject.[8,9]

E. The o(3,1) subalgebra

As sp(4,R) is isomorphic to o(3,2), where the latter has an o(3,1) maximal subalgebra, we expect also this type of subalgebra in the former. If we define

$$K_q \equiv (i/2)(B_q^\dagger - B_q), \quad q = 1,0,-1, \tag{2.20}$$

we immediately check from (2.3) that J_q, K_q close under commutation and that in fact we have

$$[J_{q''},J_{q'}] = \varepsilon_{q''q'q}J^q, \tag{2.21a}$$

$$[J_{q''},K_{q'}] = \varepsilon_{q''q'q}K^q, \tag{2.21b}$$

$$[K_1,K_{-1}] = J_0, \tag{2.21c}$$

$$[K_0,K_{\pm 1}] = \mp J_{\pm 1}, \tag{2.21d}$$

so that they are the generators of an o(3,1) subalgebra.[21]

The Casimir operators are then clearly[21]

$$J^2 - K^2 \equiv \sum_q (-)^q J_q J_{-q} - \sum_q (-)^q K_q K_{-q}, \tag{2.22}$$

$$\mathbf{J}\cdot\mathbf{K} = \sum_q (-)^q J_q K_{-q}, \tag{2.23}$$

which is corroborated by the fact that from (2.21) they commute with all the generators J_q, K_q, $q = 1,0,-1$ of o(3,1).

Again $\mathscr{N}$ is not contained in o(3,1) and, as follows from the discussion at the end of the previous subsection, we will not be interested in characterizing our states by irreps of o(3,1) despite the fact that

$$J^2 - K^2,\ \mathbf{J}\cdot\mathbf{K},\ J_0,\ K_0 \tag{2.24}$$

give a set of four commuting operators that provide a sufficient number of labels for these states.

III. STATES ASSOCIATED WITH THE MAXIMAL SUBALGEBRAS, AND MATRIX REPRESENTATION OF THE GENERATORS OF sp(4,R)

For the discussion of spectra of Hamiltonians in the enveloping algebra of sp(4,R) and the shapes of their eigenfunctions, we require a complete set of states characterized by a definite irrep in the positive discrete series of this Lie algebra. Furthermore we also need the matrix elements of the generators of sp(4,R) with respect to this set of states.

As usual in problems of this type,[2,4,15] it is convenient in turn to characterize the complete set of states by irreps of the maximal subalgebras of sp(4,R) that contain o(2), i.e., the angular momentum operator J_0. We showed in the previous section five of these, but only three,

$$\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R), \tag{3.1a}$$

$$\mathrm{su}(2) \oplus \mathrm{u}(1), \tag{3.1b}$$

$$\mathrm{sp}(2,R) \oplus \mathrm{o}(2), \tag{3.1c}$$

also contain u(1), i.e., the operator $\mathscr{N}$. As this operator will appear in the Hamiltonians we shall consider, either in the form of an A-body oscillator or in relation with the kinetic energy of the system, we will only be interested in the characterization of the complete set of states by the subalgebras in (3.1).

A. States associated with irreps of maximal subalgebras

To get these states we start by dividing the set of ten generators (2.1) of sp(4,R) into three subsets of raising, weight, and lowering type, which are separated by semicolons[4,5]

$$B_q^\dagger,\ J_1;\quad \mathscr{N},\ J_0;\quad B_q,\ J_{-1}. \tag{3.2}$$

The lowest weight state, which we designate by $|ws\rangle$, will now satisfy the equations

$$B_q|ws\rangle = 0,\ q = 1,0,-1, \tag{3.3a}$$

$$J_{-1}|ws\rangle = 0, \tag{3.3b}$$

$$\mathscr{N}|ws\rangle = w|ws\rangle, \tag{3.3c}$$

$$J_0|ws\rangle = -s|ws\rangle, \tag{3.3d}$$

where w,s are integer of half-integer numbers.[4,5] The irrep of sp(4,R) in the positive discrete series is then characterized by

$$(w,s). \tag{3.4}$$

A complete set of states for the irrep (w,s) of sp(4,R) can then be built[4,5] by applying powers of the raising generators $B_1^\dagger, B_0^\dagger, B_{-1}^\dagger, J_1$ to $|ws\rangle$ and, in their simplest form, they can be written as

$$|(ws)\sigma\tau NM) \equiv (B_1^\dagger)^{(1/2)(N+M-\tau-\sigma)}(B_0^\dagger)^\tau \times (B_{-1}^\dagger)^{(1/2)(N-M-\tau+\sigma)} J_1^{s+\sigma}|ws\rangle, \tag{3.5}$$

where the choice of exponents guarantees, from the commutation rules (2.3), that the ket (3.5) is an eigenstate of the weight generators $\mathscr{N}$, J_0, i.e.,

$$\mathscr{N}|(ws)\sigma\tau NM) = (N+w)|(ws)\sigma\tau NM), \tag{3.6a}$$

$$J_0|(ws)\sigma\tau NM) = M|(ws)\sigma\tau NM). \tag{3.6b}$$

The kets (3.5) are not states characterized by definite irreps of the maximal subalgebras in (3.1), but if we wanted to obtain the latter we would only have to consider linear combinations of (3.5) over the indices σ, τ. The N, M would remain fixed as they are related to the commuting operators $\mathscr{N}$, J_0 present in all three maximal subalgebras of (3.1).

We start by considering the maximal subalgebra $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$ of (3.1a) whose Casimir operators are the I'^2, I''^2 of (2.6) with eigenvalues[2]

$$I'^2 \to \lambda'(\lambda'-1), \tag{3.7a}$$

$$I''^2 \to \lambda''(\lambda''-1). \tag{3.7b}$$

It is convenient to express λ', λ'' in terms of the quantum numbers ν, μ by the relations

$$\lambda' = \tfrac{1}{2}(w+\nu+\mu), \tag{3.8a}$$

$$\lambda'' = \tfrac{1}{2}(w+\nu-\mu), \tag{3.8b}$$

where w is one of the numbers related to the irrep of sp(4,R) as indicated in (3.3c) and (3.4). We then show in Appendix A that the complete set of orthonormal states characterized by the irreps (3.7) of I'^2, I''^2, as well as by the eigenvalues of the operators I_0', I_0'' of (2.4b) and (2.4f) or, equivalently, of $\mathscr{N}$, J_0, can be written as

$$|(ws)\nu\mu NM\rangle = \sum_{\sigma,\tau} A_{\sigma,\tau}|(ws)\sigma\tau NM), \tag{3.9}$$

where

$$A_{\sigma\tau} = \left[\frac{(s+\mu)!(s-\mu)!\nu!(2w-3)!\Gamma(w-s+\nu-1)\Gamma(w-s-1)\Gamma(w+s)(w+\mu+\nu-1)(w+\nu-\mu-1)}{(2s)!(2w+\nu-3)!\Gamma(w+\mu-1)\Gamma(w-\mu-1)\Gamma(w+s+\nu)\Gamma(w+(N+M+\nu+\mu)/2)\Gamma(w+(N-M+\nu-\mu)/2)}\right]^{1/2}$$
$$\times\frac{(s-\sigma)!\Gamma(w-1+[(\nu+\mu+\sigma+\tau)/2])\Gamma(w-1+[(\nu-\mu-\sigma+\tau)/2])}{\tau![([(N+M-\nu-\mu)/2])!([(N-M-\nu+\mu)/2])!]^{1/2}}(-1)^{(1/2)(\nu+\mu-\sigma-\tau)}2^{(1/2)(s+\nu+\sigma+\tau)-N}$$
$$\times\sum_r\frac{(-1)^r}{r!(s+\mu-r)!(s+\sigma-r)!([(\nu-\mu-\sigma-\tau)/2]-s+r)!(r-\mu-\sigma)!\Gamma(w-1+[(\nu+\mu+\sigma+\tau)/2]-r)}. \tag{3.10}$$

As indicated in the discussion of the previous section, for $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$ the states can be characterized by the eigenvalues of the four commuting Hermitian operators in (2.7). Thus we have a sufficient number of labels and can obtain the orthonormal states (3.9). We shall use these states in the present paper in preference to those of the two other subalgebras in (3.1) which only have three commuting operators as indicated in (2.10) and (2.13). Nevertheless, for the sake of completeness, we shall also discuss briefly the states associated with the other two maximal subalgebras in (3.1).

For $\mathrm{su}(2) \oplus \mathrm{u}(1)$ of (3.1b) the Casimir operators J^2, $\mathscr{N}$ of (2.9) have eigenvalues

$$J^2 \to j(j+1),\quad \mathscr{N} \to N+w. \tag{3.11}$$

The states characterized by the irreps (3.11) of J^2, $\mathscr{N}$ as well as by the eigenvalue M of J_0 can be denoted by

$$|wN(ls)jM\} \tag{3.12}$$

and are given explicitly in Eq. (3.14) of the first paper of Ref. 4. The l, is a multiplicity index and, as indicated in Eq. (3.15) of the same reference, is restricted by

$$|j-s| \leqslant l \leqslant j+s, \tag{3.13a}$$

$$l = N, N-2, \ldots, 1 \text{ or } 0. \tag{3.13b}$$

It is easy to write the states (3.12) in terms of the kets (3.5) as we did in (3.9) for the subalgebra $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$. We shall not do this explicitly because, as we indicated in the previous paragraph, the presence of a multiplicity index[4] gives a nonorthonormal basis whose overlaps are cumbersome to determine, and thus the states (3.12) are not as useful as (3.9).

For $\mathrm{sp}(2,R) \oplus \mathrm{o}(2)$ of (3.1c) the Casimir operators I^2, J_0 of (2.12) have eigenvalues

$$I^2 \to \lambda(\lambda-1), \quad J_0 \to M, \tag{3.14}$$

where following Ref. 2 we write

$$\lambda = w + \Lambda. \tag{3.15}$$

The states characterized by the eigenvalues of I^2, J_0 as well as by the N related to the eigenvalues of $I_0 = \mathscr{N}$ of (2.11b), can be written as

$$|(ws)q\Lambda NM], \tag{3.16}$$

where q is a multiplicity index. These states were given explicitly in Eqs. (4.42) and (4.52) of Ref. 2 and they can be expressed as linear combinations of the kets (3.5), if we use the relation between the generators of $\mathrm{sp}(4,R)$ and the creation and annihilation operators for the A-body system.[4] Again we shall not do this explicitly for the same reasons given in the previous paragraph for the subalgebra (3.1b).

B. Matrix representation of the generators of sp(4,R)

As we indicated in the previous subsection we shall only be interested in the matrix elements of the ten generators (2.1) of $\mathrm{sp}(4,R)$ with respect to the states (3.9), characterized in turn by the irrep $(\lambda'\lambda'')$ or, equivalently, $(\nu\mu)$ of the $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$ maximal subalgebra. From (3.6) and (3.9) we immediately get that the matrix elements of the weight generators $\mathscr{N}$, J_0 are

$$\langle (ws)\nu'\mu'N'M'|\mathscr{N}|(ws)\nu\mu NM\rangle = (N+w)\delta_{\nu'\nu}\delta_{\mu'\mu}\delta_{N'N}\delta_{M'M}, \tag{3.17a}$$

$$\langle (ws)\nu'\mu'N'M'|J_0|(ws)\nu\mu NM\rangle = M\delta_{\nu'\nu}\delta_{\mu'\mu}\delta_{N'N}\delta_{M'M}. \tag{3.17b}$$

Furthermore, we note from (2.4) that $B_1^\dagger$, B_{-1} and $B_{-1}^\dagger$, B_1 are, respectively, elements of the Lie algebras $\mathrm{sp}'(2,R)$ and $\mathrm{sp}''(2,R)$. Thus we can use the Wigner–Eckart theorem for the $\mathrm{sp}(2,R)$ algebra to determine their matrix elements. As $\mathrm{sp}(2,R)$ is isomorphic to $\mathrm{su}(1,1)$, we can make use of the corresponding analysis of Ui[22] to obtain

$$\langle (ws)\nu'\mu'N'M'|B_{\pm 1}^\dagger|(ws)\nu\mu NM\rangle = -\left[\tfrac{1}{2}(N \pm M + \nu \pm \mu + 2w)(N \pm M - \nu \mp \mu + 2)\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu}\delta_{N'N+1}\delta_{M'M\pm 1}, \tag{3.17c}$$

$$\langle (ws)\nu'\mu'N'M'|B_{\pm 1}|(ws)\nu\mu NM\rangle = \left[\tfrac{1}{2}(N \mp M - \nu \pm \mu)(N \mp M \mp \mu + \nu + 2w - 2)\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu}\delta_{N'N-1}\delta_{M'M\pm 1}. \tag{3.17d}$$

There remain four generators which we designate by

$$T^{(1/2)(1/2)}_{(1/2)(1/2)} \equiv B_0^\dagger, \quad T^{(1/2)(1/2)}_{(1/2)-(1/2)} \equiv -\sqrt{2}J_1, \quad T^{(1/2)(1/2)}_{-(1/2)(1/2)} \equiv \sqrt{2}J_{-1}, \quad T^{(1/2)(1/2)}_{-(1/2)-(1/2)} \equiv B_0. \tag{3.18}$$

As shown in Appendix B this notation is introduced to indicate that the four remaining generators belong to an irreducible tensor $\mathbf{T}^{(1/2)(1/2)}$ of rank $(\frac{1}{2}\frac{1}{2})$ with respect to the Lie algebra $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$. Again using the results of Ui[22] we see that we only need to obtain the reduced matrix element

$$\langle (ws)\nu'\mu'\|\mathbf{T}^{(1/2)(1/2)}\|(ws)\nu\mu\rangle, \tag{3.19}$$

which was achieved in Appendix B with the help of the states (3.9) for special values of the quantum numbers. Thus we obtain the matrix elements

$$\begin{aligned}
&\langle (ws)\nu'\mu'N'M'|B_0^\dagger|(ws)\nu\mu NM\rangle \\
&= \frac{1}{2}\left[\frac{\nu(w-s+\nu-2)(w+s+\nu-1)(2w+\nu-3)(N+M-\mu-\nu+2)(N-M+\mu-\nu+2)}{(w+\mu+\nu-2)(w-\mu+\nu-2)(w+\mu+\nu-1)(w-\mu+\nu-1)}\right]^{1/2}\delta_{\nu'\nu-1}\delta_{\mu'\mu}\delta_{N'N+1}\delta_{M'M} \\
&+ \frac{1}{2}\left[\frac{(\nu+1)(w-s+\nu-1)(w+s+\nu)(2w+\nu-2)(2w+N+M+\nu+\mu)(2W+N-M+\nu-\mu)}{(w+\mu+\nu)(w-\mu+\nu)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu+1}\delta_{\mu'\mu}\delta_{N'N+1}\delta_{M'M} \\
&+ \frac{1}{2}\left[\frac{(\mu+s+1)(s-\mu)(w+\mu-1)(w-\mu-2)(N+M+\nu+\mu+2w)(N-M+\mu-\nu+2)}{(w+\mu+\nu)(w-\mu+\nu-2)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu+1}\delta_{N'N+1}\delta_{M'M} \\
&+ \frac{1}{2}\left[\frac{(\mu+s)(s-\mu+1)(w+\mu-2)(w-\mu-1)(N+M-\mu-\nu+2)(N-M+\nu-\mu+2w)}{(w+\mu+\nu-2)(w-\mu+\nu)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu-1}\delta_{N'N+1}\delta_{M'M},
\end{aligned} \tag{3.20a}$$

$$
\begin{aligned}
&\langle (ws)\nu'\mu'N'M'|J_1|(ws)\nu\mu NM\rangle \\
&= -\frac{1}{2}\left[\frac{\nu(w-s+\nu-2)(w+s+\nu-1)(2w+\nu-3)(N+M-\mu-\nu+2)(N-M+\nu-\mu+2w)}{2(w+\mu+\nu-2)(w-\mu+\nu-2)(w+\mu+\nu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu-1}\delta_{\mu'\mu}\delta_{N'N}\delta_{M'M+1} \\
&\quad+\frac{1}{2}\left[\frac{(\nu+1)(w-s+\nu-1)(w+s+\nu)(2w+\nu-2)(2w+M+N+\nu+\mu)(N-M+\mu-\nu)}{2(w+\mu+\nu)(w-\mu+\nu)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu+1}\delta_{\mu'\mu}\delta_{N'N}\delta_{M'M+1} \\
&\quad-\frac{1}{2}\left[\frac{(\mu+s+1)(s-\mu)(w+\mu-1)(w-\mu-2)(N+M+\nu+\mu+2w)(N-M+\nu-\mu+2w-2)}{2(w+\mu+\nu)(w-\mu+\nu-2)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu+1}\delta_{N'N}\delta_{M'M+1} \\
&\quad-\frac{1}{2}\left[\frac{(\mu+s)(s-\mu+1)(w+\mu-2)(w-\mu-1)(N+M-\mu-\nu+2)(N-M+\mu-\nu)}{2(w+\mu+\nu-2)(w-\mu+\nu)(w+\nu+\mu-1)(w+\nu-\mu-1)}\right]^{1/2}\delta_{\nu'\nu}\delta_{\mu'\mu-1}\delta_{N'N}\delta_{M'M+1},
\end{aligned}
\tag{3.20b}
$$

while those of B_0, J_{-1} can be determined from (3.20a) and (3.20b) by Hermitian conjugation.

We have thus obtained the matrix representation of the generators of sp(4,R) in a basis characterized by irreps of sp'(2,R) $\oplus$ sp"(2,R) and their Abelian subalgebras u(1), o(2) whose generators are $\mathscr{N}$, J_0.

These matrix elements were derived previously by Peterson and Hecht[23] for applications to a different problem, and without the explicit use of the state (3.9). They differ though from those given in (3.17) and (3.20) by phases, as Peterson and Hecht do not use those of Ui for sp(2,R).

We can now turn our attention to the collective Hamiltonians in the enveloping algebra of sp(4,R), to their spectra and to the shape of their eigenfunctions.

IV. COLLECTIVE HAMILTONIANS IN THE ENVELOPING ALGEBRA OF sp(4,R)

For an A-body system in two-dimensional space collective Hamiltonians are associated with functions of the coordinates and momenta that involve sums over the particle index. If we eliminate the center of mass motion, designating by x_{is}, p_{is}, $i = 1,2$, $s = 1,2,...,n = A - 1$, the Jacobi relative coordinates and momenta, this implies[1] that the collective degrees of freedom are functions of x_{is}, p_{is} involving sums over s in the range $s = 1,2,...,n = A - 1$. In previous discussions[2,4] we showed though that the generators of sp(4,R) can be written as functions of x_{is}, p_{is} involving this type of summation. Thus we can express collective Hamiltonians as functions of the generators of sp(4,R) so long as they are Hermitian and remain invariant under rotation (i.e., commute with J_0) and time reflection.

To write the generators (2.1) of sp(4,R) explicitly in terms of x_{is}, p_{is} we first give these variables in terms of circular components, i.e.,

$$x_s \equiv (1/\sqrt{2})(x_{1s} + ix_{2s}), \tag{4.1a}$$

$$\bar{x}_s \equiv (1/\sqrt{2})(x_{1s} - ix_{2s}), \tag{4.1b}$$

$$p_s \equiv (1/\sqrt{2})(p_{1s} + ip_{2s}), \tag{4.1c}$$

$$\bar{p}_s \equiv (1/\sqrt{2})(p_{1s} - ip_{2s}), \tag{4.1d}$$

with s taking the values

$$s = 1,2,...,n, \quad n = A - 1. \tag{4.1e}$$

We introduce the corresponding creation and annihilation operators by the definition

$$\eta_s = (1/\sqrt{2})(x_s - ip_s), \tag{4.2a}$$

$$\bar{\eta}_s = (1/\sqrt{2})(\bar{x}_s - i\bar{p}_s), \tag{4.2b}$$

$$\xi_s = (1/\sqrt{2})(x_s + ip_s), \tag{4.2c}$$

$$\bar{\xi}_s = (1/\sqrt{2})(\bar{x}_s - i\bar{p}_s), \tag{4.2d}$$

where we have taken units in which $\hbar$, the mass of the particles, and an appropriate frequency are taken as 1. These operators satisfy the commutation relation

$$[\xi_s, \eta_t] = 0, \tag{4.3a}$$

$$[\bar{\xi}_s, \bar{\eta}_t] = 0, \tag{4.3b}$$

$$[\bar{\xi}_s, \eta_t] = \delta_{st}, \tag{4.3c}$$

$$[\xi_s, \bar{\eta}_t] = \delta_{st}, \tag{4.3d}$$

and have the Hermitian property

$$\eta_s^\dagger = \bar{\xi}_s, \quad \bar{\eta}_s^\dagger = \xi_s, \quad \xi_s^\dagger = \bar{\eta}_s, \quad \bar{\xi}_s^\dagger = \eta_s. \tag{4.4}$$

We furthermore need to know the behavior of these creation and annihilation observables under the operation of time inversion which we shall denote by T. Clearly the real observables x_{is}, p_{is}, $i = 1,2$, $s = 1,2,...,n$, behave under time inversion as

$$Tx_{is} = x_{is}, \tag{4.5a}$$

$$Tp_{is} = -p_{is}, \tag{4.5b}$$

and thus from (4.1), (4.2), and taking into account that the operation of time inversion implies also a conjugation,[23] we get

$$T\eta_s = \bar{\eta}_s, \quad T\bar{\eta}_s = \eta_s, \quad T\xi_s = \bar{\xi}_s, \quad T\bar{\xi}_s = \xi_s. \tag{4.6}$$

We now define

$$B_1^\dagger = (1/\sqrt{2})\eta_s\eta_s, \tag{4.7a}$$

$$B_0^\dagger = \eta_s\bar{\eta}_s, \tag{4.7b}$$

$$B_{-1}^\dagger = (1/\sqrt{2})\bar{\eta}_s\bar{\eta}_s, \tag{4.7c}$$

$$J_1 = -(1/\sqrt{2})\eta_s\xi_s, \tag{4.7d}$$

$$J_0 = \tfrac{1}{2}(\eta_s\bar{\xi}_s - \bar{\eta}_s\xi_s), \tag{4.7e}$$

$$J_{-1} = (1/\sqrt{2})\bar{\eta}_s\bar{\xi}_s, \tag{4.7f}$$

$$B_1 = -(1/\sqrt{2})\xi_s\xi_s, \tag{4.7g}$$

$$B_0 = \xi_s\bar{\xi}_s, \tag{4.7h}$$

$$B_{-1} = -(1/\sqrt{2})\bar{\xi}_s\bar{\xi}_s, \tag{4.7i}$$

$$\mathscr{N} = \tfrac{1}{2}(\eta_s\bar{\xi}_s + \bar{\eta}_s\xi_s + n), \tag{4.7j}$$

where repeated indices s are summed from 1 to n. We easily check, using (4.3), that $\mathscr{N}$, $B_q^\dagger$, J_q, B_q, $q = 1,0,-1$ of (4.7) satisfy the commutation rules (2.3), and thus we have obtained a realization of the generators of sp(4,R) in terms of the translationally invariant creation and annihilation operators (4.2) of the A-body system. Incidentally, we note from (4.7j), (4.2), and (4.1) that $2\mathscr{N}$ is the Hamiltonian of a $2n$-dimensional oscillator, as was mentioned several times in the previous sections.

We wish now to derive the properties of the generators of sp(4,R) under Hermitian conjugation, rotation (i.e., commutation with J_0), and time reflection. From (4.4) and (4.7) we have

$$(B_q^\dagger)^\dagger = (-1)^q B_{-q}, \tag{4.8a}$$

$$(B_q)^\dagger = (-1)^q B_{-q}^\dagger, \tag{4.8b}$$

$$(J_q)^\dagger = (-1)^q J_{-q}, \tag{4.8c}$$

$$\mathscr{N}^\dagger = \mathscr{N}, \tag{4.8d}$$

while from the commutation rules (2.3) we get

$$[J_0, B_q^\dagger] = qB_q^\dagger, \tag{4.9a}$$

$$[J_0, B_q] = qB_q, \tag{4.9b}$$

$$[J_0, J_q] = qJ_q, \tag{4.9c}$$

$$[J_0, \mathscr{N}] = 0. \tag{4.9d}$$

Finally, from (4.6) we see that under time reflection we get

$$TB_q^\dagger = B_{-q}^\dagger, \tag{4.10a}$$

$$TB_q = B_{-q}, \tag{4.10b}$$

$$TJ_q = -J_{-q}, \tag{4.10c}$$

$$T\mathscr{N} = \mathscr{N}. \tag{4.10d}$$

With the help of these properties of the generators we can now discuss collective Hamiltonians, functions of them, that are Hermitian and invariant under rotation and time reflection.

We start with Hamiltonians that are of first degree in the generators and obviously only

$$\mathscr{N}, \tag{4.11a}$$

$$Q_0 = \tfrac{1}{2}(B_0^\dagger + B_0) + \mathscr{N}, \tag{4.11b}$$

satisfy the requirements mentioned at the end of the previous paragraph. We note that $\mathscr{N}$ is present in the three maximal subalgebras of (3.1) while Q_0 is an element of cm(2).

Turning now our attention to Hamiltonians of second degree in the generators we obviously have that

$$\mathscr{N}^2, \tag{4.12a}$$

$$(\mathscr{N}Q_0 + Q_0\mathscr{N}), \tag{4.12b}$$

$$Q_0^2, \tag{4.12c}$$

also satisfy all the requirements as does

$$J_0^2. \tag{4.12d}$$

We can then go systematically through all possible bilinear combinations and, with the help of (4.8)–(4.10), find [up to the terms (4.11) of first order that can come from commutations] that there are only six more Hamiltonians that satisfy the requirements of Hermiticity and invariance under rotation and time reflection, given by

$$B_1^\dagger B_{-1} + B_{-1}^\dagger B_1, \tag{4.13a}$$

$$J_1 J_{-1} + J_{-1} J_1, \tag{4.13b}$$

$$B_0^\dagger B_0, \tag{4.13c}$$

$$(B_1^\dagger J_{-1} + J_1 B_{-1}) - (B_{-1}^\dagger J_1 + J_{-1} B_1), \tag{4.13d}$$

$$B_1^\dagger B_{-1}^\dagger + B_1 B_{-1}, \tag{4.13e}$$

$$B_0^\dagger B_0^\dagger + B_0 B_0. \tag{4.13f}$$

We wish now to show that all the bilinear Hamiltonians (4.13) can be given in terms of appropriate Casimir operators of the maximal subalgebras combined with the terms (4.12). For this purpose we first need to see which of the Casimir operators discussed in Sec. II are invariant under time reflection, as obviously all are Hermitian and, by definition, commute with J_0.

From (4.10) we note that Casimir operators of sp′(2,R) ⊕ sp″(2,R), i.e., I'^2, I''^2 of (2.6), are not invariant under time reflection, and that in fact

$$TI'^2 = I''^2, \quad TI''^2 = I'^2. \tag{4.14}$$

Thus only the combination $I'^2 + I''^2$ would fulfill all requirements. On the other hand, and again using (4.10), we see that Casimir operators J^2 of su(2) and I^2 of sp(2,R) are invariant under time reflection. For the Casimir operators of cm(2) we have

$$TQ^2 = Q^2, \quad TW = -W, \tag{4.15}$$

and so we can only retain Q^2, while both Casimir operators of o(3,1) are invariant under time reflection.

We see then that following six Casimir operators, or combination of them, of the maximal subalgebras satisfy all the requirements of collective Hamiltonians

$$I'^2 + I''^2, \tag{4.16a}$$

$$J^2, \tag{4.16b}$$

$$I^2, \tag{4.16c}$$

$$Q^2, \tag{4.16d}$$

$$J^2 - K^2, \tag{4.16e}$$

$$\mathbf{J}\cdot\mathbf{K}. \tag{4.16f}$$

It can be easily seen that

$$B_1^\dagger B_{-1} + B_{-1}^\dagger B_1 = 2(I'^2 + I''^2) - (\mathscr{N}^2 + J_0^2) + 2\mathscr{N}, \tag{4.17a}$$

$$J_1 J_{-1} + J_{-1} J_1 = J_0^2 - J^2, \tag{4.17b}$$

$$B_0^\dagger B_0 = \mathscr{N}(\mathscr{N} - 1) - I^2, \tag{4.17c}$$

while the Hamiltonians (4.13d)–(4.13f) can be expressed as linear combinations of the Casimir operators Q^2, $J^2 - K^2$, $\mathbf{J}\cdot\mathbf{K}$ and the operators (4.12).

We have thus Hamiltonians associated with specific subalgebras whose eigenvalues can be given in closed form. We consider first sp′(2,R) ⊕ sp″(2,R) for which the most general Hamiltonian is

$$H = \alpha(I'^2 + I''^2) + \beta\mathscr{N}^2 + \gamma J_0^2 + \delta\mathscr{N}, \tag{4.18}$$

where $\alpha, \beta, \gamma, \delta$ are arbitrary constants and for which, from (3.6)–(3.8), the eigenvalues are

$$E = (\alpha/2)\{[(w+\nu)^2 + \mu^2] - 2(w+\nu)\} + \beta(N+w)^2 + \gamma M^2 + \delta(N+w). \tag{4.19}$$

For su(2) ⊕ u(1) the Hamiltonian is

$$H = \alpha J^2 + \beta\mathscr{N}^2 + \gamma J_0^2 + \delta\mathscr{N} \tag{4.20}$$

with eigenvalues

$$E = \alpha j(j+1) + \beta(N+w)^2 + \gamma M^2 + \delta(N+w), \quad (4.21)$$

while for $sp(2,R) \oplus o(2)$ we just have to replace J^2 by I^2 in the Hamiltonian (4.20) and $J(J+1)$ by $\lambda(\lambda-1)$ in the eigenvalues (4.21).

A similar discussion could be carried out for cm(2) and o(3,1), but as we indicated in previous sections their Casimir operators Q^2, $J^2 - K^2$, $\mathbf{J}\cdot\mathbf{K}$ would have to be supplemented with the operator $\mathscr{N}$ that belongs to the other maximal subalgebras. Thus the Hamiltonian would be of the transitional type, i.e., more than one chain of subalgebras is involved and thus the calculations have to be carried out numerically, as we shall proceed to do in Sec. VI using the matrix representations (3.17) and (3.20) of the generators of $sp(4,R)$.

V. SHAPES ASSOCIATED WITH THE EIGENSTATES OF OUR HAMILTONIANS

While our Hamiltonians are defined in the enveloping algebra of $sp(4,R)$ they correspond to a microscopic description as, through (4.7), (4.2), these generators in turn depend on the coordinates and momenta of the A-body system. We can take advantage of this microscopic picture to determine the shape of the eigenstates of the Hamiltonian, through the discussion of the mass quadrupole tensor for the A-body system.

Using the circular and translationally invariant coordinates x_s, $\bar{x}_s$, $s = 1,2,...,n = A-1$ defined in (4.1a) and (4.1b) the mass quadrupole tensor in the frame of reference fixed in space is given by[2]

$$Q_1 = (1/\sqrt{2})x_s x_s, \quad Q_0 = x_s \bar{x}_s, \quad Q_{-1} = (1/\sqrt{2})\bar{x}_s \bar{x}_s, \quad (5.1)$$

where the repeated indices s are summed over the values $s = 1,...,n$. We see from the relations

$$x_s = (1/\sqrt{2})(\eta_s + \xi_s), \quad \bar{x}_s = (1/\sqrt{2})(\bar{\eta}_s + \bar{\xi}_s) \quad (5.2)$$

that follow from (4.2), and the explicit expression of the generators of $sp(4,R)$ given in (4.7), that the operators Q_q, $q = 1,0,-1$, are exactly the generators of cm(2) given in (2.14a)–(2.14c). Thus we have a physical understanding of the Q_q which in (2.14) were introduced in an abstract fashion.

We would like though to have the quadrupole tensor in the frame of reference fixed in the body and this requires that the $2n$ coordinates x_s, $\bar{x}_s$, $s = 1,2,...,n$, are transformed by the procedure originally introduced by Dzublick *et al.*[17] and by Zickendraht.[18] The new coordinates are ρ_1, ρ_2, which give the length of the principal axis for the A-body system in a two-dimensional space, the angle ϑ, which relates the frame of reference fixed in space with that fixed in the body, and the remaining $2n-3$ variables, which we denote generically by ϕ (Refs. 2 and 3). As was discussed originally in the groups led by Filippov[6] and Vanagas,[7] the variables ϕ are angles that characterize the elements of the defining representation

$$\mathbf{D}(\phi) = \|D_{st}(\phi)\| \quad (5.3)$$

of the n-dimensional orthogonal group $O(n)$ associated with the particle indices $t,s = 1,2,...,n = A-1$. For the particular case of the two-dimensional space it was shown[2] that x_s, $\bar{x}_s$ are then given by

$$x_s = (1/\sqrt{2})e^{i\vartheta}[\rho_1 D^{\,1}_{n-1,s}(\phi) + i\rho_2 D^{\,1}_{ns}(\phi)], \quad (5.4a)$$

$$\bar{x}_s = (1/\sqrt{2})e^{-i\vartheta}[\rho_1 D^{\,1}_{n-1,s}(\phi) - i\rho_2 D^{\,1}_{ns}(\phi)]. \quad (5.4b)$$

In the frame of reference fixed in the body the components Q_q, $q = 1,0,-1$ for the mass quadrupole take then the form

$$Q_{\pm 1} = (1/\sqrt{2})e^{\pm i2\vartheta}\tfrac{1}{2}(\rho_1^2 - \rho_2^2), \quad (5.5a)$$

$$Q_0 = \tfrac{1}{2}(\rho_1^2 + \rho_2^2), \quad (5.5b)$$

where we made use of the orthogonality property of the defining representation of $O(n)$, i.e.,

$$D_{ts}(\phi)D_{t's}(\phi) = \delta_{tt'}, \quad (5.6)$$

where repeated indices are summed from 1 to n.

The expressions (5.5) corroborate the interpretation[2] of ρ_1, ρ_2 as the principal semiaxis of the ellipse of inertia. One measure of the deformation is then the eccentricity

$$\epsilon = (\rho_1^2 - \rho_2^2)^{1/2}/\rho_1, \quad (5.7)$$

while another is given by a parameter β, which corresponds, in this two-dimensional problem, to the β variable of the Bohr–Mottelson model,[24] and can be defined by

$$\beta = (\rho_1^2 - \rho_2^2)/(\rho_1^2 + \rho_2^2). \quad (5.8)$$

Introducing the change of variables

$$\rho_1 = \rho \cos\gamma, \quad (5.9a)$$

$$\rho_2 = \rho \sin\gamma \quad (5.9b)$$

and

$$r = \tfrac{1}{2}\rho^2, \quad (5.9c)$$

$$\theta = 2\gamma + (\pi/2), \quad (5.9d)$$

$$\varphi = 2\vartheta, \quad (5.9e)$$

we see that the eccentricity takes the form

$$\epsilon = [2\sin\theta/(1+\sin\theta)]^{1/2}, \quad (5.10a)$$

while the β becomes

$$\beta = \sin\theta. \quad (5.10b)$$

As $\rho_1 \geqslant \rho_2 \geqslant 0$ (Refs. 2 and 3) we have that γ,θ of (5.9) are, respectively, in the interval

$$0 \leqslant \gamma \leqslant \pi/4, \quad \pi/2 \leqslant \theta \leqslant \pi, \quad (5.11)$$

implying that both ϵ and β take values in the range from 0 to 1. The first value corresponds to a circular shape, i.e., no deformation as $\rho_1 = \rho_2$, while the second is associated with extreme deformation, i.e., $\rho_1 = \rho$, $\rho_2 = 0$.

In the coordinates (5.9) the components of the mass quadrupole take the values

$$Q_{\pm 1} = (1/\sqrt{2})r \sin\theta \, e^{\pm i\varphi}, \quad (5.12a)$$

$$Q_0 = r, \quad (5.12b)$$

from which we have that

$$\sin^2\theta = 2Q_1 Q_{-1}/Q_0^2. \quad (5.13)$$

It would be difficult to calculate the expectation value of $\sin^2\theta$ with respect to the eigenkets $|\Lambda\rangle$ of Hamiltonians of the type discussed in the previous section, because of the Q_0^2

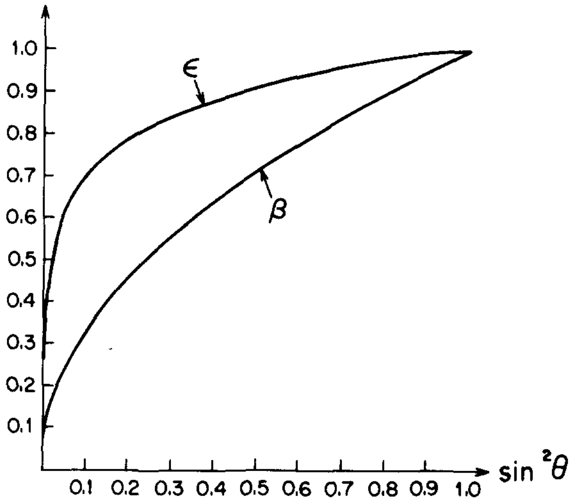


FIG. 1. The deformation β and eccentricity ϵ of Eq. (5.10) as a function of $\sin^2\theta$.

appearing in the denominator. Thus we prefer to estimate the expectation value of $\sin^2\theta$ through that of the ratio

$$\langle\Lambda|2Q_1Q_{-1}|\Lambda\rangle/\langle\Lambda|Q_0^2|\Lambda\rangle. \tag{5.14}$$

The values of ϵ and β can then be evaluated through the functional relations (5.10), and in Fig. 1 we draw both ϵ and β as functions of $\sin^2\theta$.

The operators Q_1, Q_0, Q_{-1} are given as linear combinations of generators of sp(4,R) in (2.14a)–(2.14c) and their matrix elements with respect to the states (3.9) can then be obtained from (3.17) and (3.20). On the other hand, using again the kets $|(ws)v\mu NM\rangle$ of (3.9) the eigenstates of definite Hamiltonians can be written as

$$|\Lambda\rangle = \sum_{v\mu NM} C[(ws)v\mu NM,\Lambda]|(ws)v\mu NM\rangle, \tag{5.15}$$

where the C's are numerical coefficients. Therefore again the evaluation of the ratio of expectation (5.14) requries only the matrix elements (3.17) and (3.20) of the generators of sp(4,R) with respect to the states (3.9).

We are thus in position to determine both the eigenvalues and eigenfunctions of the Hamiltonians discussed in the previous section and also analyze the shape associated with these eigenfunctions. We shall proceed to give some illustrative examples of all the relevant aspects of this problem in the next section.

VI. SPECTRA OF COLLECTIVE HAMILTONIANS IN THE ENVELOPING ALGEBRA OF sp(4,R) AND SHAPE OF THEIR EIGENSTATES

In Eq. (3.17) of Sec. III we gave the matrix elements of the generators of sp(4,R) with respect to a complete and orthonormal set of states (3.9) that were characterized by the irreps of the chain of subalgebras

$$\begin{array}{cccc} \overset{(ws)}{\mathrm{sp}(4,R)} \supset & \overset{\lambda'=(w+v+\mu)/2}{\mathrm{sp}'(2,R)} & \oplus & \overset{\lambda''=(w+v-\mu)/2}{\mathrm{sp}''(2,R)} \\ & \cup & & \cup \\ & \underset{(N+w+M)/2}{\mathrm{o}'(2)} & & \underset{(N+w-M)/2}{\mathrm{o}''(2)} \end{array}, \tag{6.1}$$

where above or below each Lie algebra we indicate the quantum numbers determining the irrep. We then showed in Sec. IV how to use these matrix elements to obtain the eigenvalues of collective Hamiltonians in the enveloping algebra of sp(4,R), and indicated in Sec. V how to determine the shape of the corresponding eigenfunctions.

In this section we proceed to apply the previous analysis to specific cases so as to sharpen our intuition on the type of spectra and shape we can expect for Hamiltonians associated with generators in different subalgebras or with linear combinations of these Hamiltonians. As our space is two dimensional, the present problem is purely conceptual, but we expect that many of the results carry on to the physical situation in three dimensions, and thus would be of interest in the symplectic model of the nucleus.

We start our discussion by introducing an "oscillator shell model" in our two-dimensional space to obtain the values (w,s) characterizing the irrep of sp(4,R) that we can associate with our A-body system.

A. The oscillator shell model and the irreps of sp(4,R)

We give in Fig. 2 the levels of a two-dimensional harmonic oscillator as function of the number quanta that we put in the ordinate and designate by a, and the angular momentum in the plane which we put in the abscissa and designate by b. Clearly

$$|b| = a, a-2, \ldots, 1 \text{ or } 0, \tag{6.2}$$

and for $b \neq 0$ each level is doubly degenerate, i.e., we have the values $\pm b$ is indicated in Fig. 2. If we assume that we fill each level with four particles, associated with the two possibilities of spin and isospin, then with each value of a we have

$$4(a+1) \tag{6.3}$$

particles. If we fill these levels compactly up to a number of quanta c, i.e., we close the shells up to c, and at the $c+1$ level denote by V the number of valence particles, our total num-

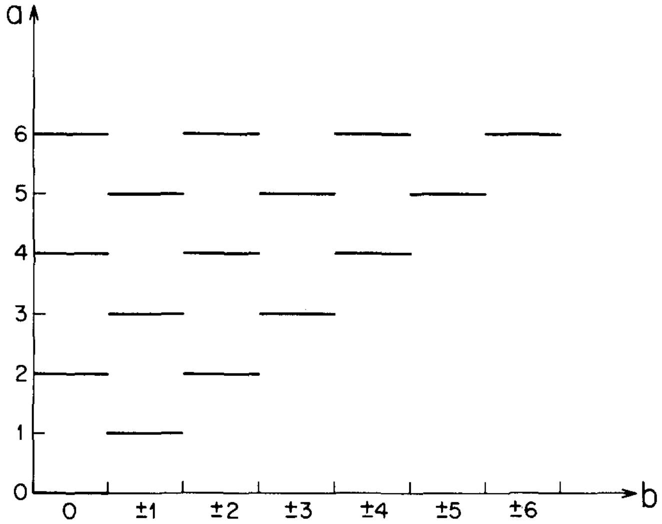


FIG. 2. Levels of the two-dimensional harmonic oscillator as a function of the number of quanta a, and where b is the angular momentum in two dimensions, which can have positive and negative values.

ber of particles is

$$A=4\sum_{a=0}^{c}(a+1)+V=2(c+1)(c+2)+V. \qquad (6.4)$$

The A-body state of the previous paragraph[3] can then be interpreted as the lowest weight state $|ws\rangle$ of (3.3). When we apply the operator $\mathscr{N}$ to $|ws\rangle$ we get the eigenvalue w but, as indicated at the end of the paragraph following Eq. (4.7j), $2\mathscr{N}$ is the translational invariant[1,2] Hamiltonian for A particles in the two-dimensional oscillator and thus, applying $\mathscr{N}$ to the state discussed in the previous paragraph, we obtain

$$w=\frac{1}{2}\left\{4\sum_{a=0}^{c}a(a+1)+(c+1)V+n\right\}$$
$$=\frac{2}{3}c(c+1)(c+2)+\frac{1}{2}(c+1)V+\frac{1}{2}(A-1). \qquad (6.5)$$

To determine s we need to apply the operator J_0 of (4.7e), which corresponds to half the angular momentum of a system of A particles in the two-dimensional oscillator. Clearly only the valence particles contribute and we could get many values of s depending on how we arrange them in the level of $c+1$ quanta. We note from (3.3) that $|ws\rangle$ is also characterized by a definite irrep of the su(2) subalgebra as

$$J^2|ws\rangle=s(s+1)|ws\rangle. \qquad (6.6)$$

We will be interested in the highest possible irrep s of su(2) consistent with the given number of valence particles V, reflecting the corresponding situation for su(3) in the symplectic model of the nucleus.[1,3,9] To get this highest value of s we shall assume for simplicity that V is a multiple of 4, and we fill then with four particles (spin up or down, isospin up or down) all the levels with positive angular momentum starting with $b=c+1$ and going down by groups of 2 until we exhaust the valence particles. We then have

$$2s=4\{(c+1)+(c-1)+\cdots$$
$$+[(c+1)-2(V/4)+2]\}$$
$$=V(c+1)-V[(V/4)-1]. \qquad (6.7)$$

An example we shall discuss in all of the following applications will correspond to filled shells up to an including $c=5$, and taking valence particles in the shell of six quanta, with V going from 0 (closed shell) to 12 (half-filled shell) by jumps of 4. From (6.5) and (6.7) we get then the following values:

V	0	4	8	12
w	181.5	195.5	209.5	223.5
s	0	12	20	24

. (6.8)

Having obtained the irreps (ws) of sp(4,R) from the two-dimensional oscillator shell model, we proceed now to discuss the spectra of different Hamiltonians and the shape of their corresponding eigenfunctions. We start first with the Hamiltonians associated with the maximal subalgebras of sp(4,R) that were mentioned in (3.1).

B. Spectra and shapes for the sp′(2,R)⊕sp″(2,R) subalgebra

The Hamiltonian in this case is given by (4.18) and we will concentrate in particular on its term

$$I'^2+I''^2, \qquad (6.9)$$

whose eigenvalue is, from (4.19), given by

$$E_{\nu\mu}=\tfrac{1}{2}\{w^2+w(2\nu-2)+\nu(\nu-2)+\mu^2\}. \qquad (6.10)$$

To discuss the spectra given by (6.10) we need to have also the range of values for ν,μ when we fix the irrep (ws) of sp(4,R) and the number of collective excitation quanta N and angular momentum M. To obtain this range we note that the eigenstate $|(ws)\nu\mu NM\rangle$ of our Hamiltonian is given by (3.9) in which appear the coefficients $A_{\sigma\tau}$ of (3.10), where the latter have a number of factorials involving w,s,ν,μ,N,M.

For the $A_{\sigma\tau}$ to make sense the arguments of the factorials must be non-negative, from which we derive immediately that

$$\tfrac{1}{2}(N+M-\nu-\mu)\geqslant 0, \qquad (6.11a)$$
$$\tfrac{1}{2}(N-M-\nu+\mu)\geqslant 0, \qquad (6.11b)$$
$$0\leqslant\nu\leqslant N, \qquad (6.11c)$$
$$-s\leqslant\mu\leqslant s, \qquad (6.11d)$$

where the right-hand side of (6.11c) comes from summing (6.11a) and (6.11b). The inequalities (6.11) indicate that if N, M are small numbers, for example in the range between 0 and 10, the same holds for ν,μ and therefore they are small compared with the w's in (6.8). Thus the spectrum is approximately given by

$$E_{\nu\mu}\cong\tfrac{1}{2}\{w^2+w(2\nu-2)\}, \qquad (6.12)$$

and, as w is fixed, this means that we have a vibrational spectrum with respect to the quantum number ν.

The shape of the states $|(ws)\nu\mu NM\rangle$ can be obtained by determining the expectation values of the operators $2Q_1Q_{-1}$ and Q_0^2 with respect to them and taking their ratios. These expectation values follow immediately from (3.17) and are given explicitly in Appendix C. We shall discuss first the state of 0 collective excitation, i.e., $N=0$ for which $\nu=0$ and $\mu=M$. In this case, as can be seen immediately from the results in Appendix C, we have

$$\beta^2=\frac{\langle 2Q_1Q_{-1}\rangle}{\langle Q_0^2\rangle}=\frac{w+s(s+1)-M^2}{w(w+\frac{1}{2})}. \qquad (6.13)$$

In Fig. 3 we make use of (6.13) to graph the deformation parameter β as function of the number V of valence particles for the state $M=0$, considering the shells below six quanta as filled, so we can use (6.8) for the values of w,s. We also give the value of the eccentricity ϵ which is related to β through (5.10).

We note that both β and ϵ increase as we go from closed shells to the middle of the next shell, but this does not necessarily imply a deformation in the classical sense as we see that if we go to higher and higher shells, i.e., $c\to\infty$ we have from (6.5) and (6.7) that the ratio $(s/w)\to 0$ and thus also the β of (6.13) vanishes.

In Table I we give the deformation parameter β for the

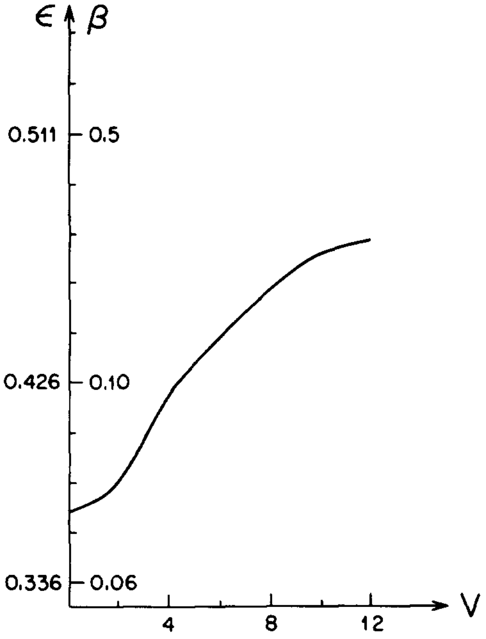


FIG. 3. Deformation β and eccentricity ϵ as a function of the number of valence nucleons V for the $M=0$ state of zero excitation quanta $N=0$. Note the increase with increasing number of V.

states $|(ws)NONO\rangle$, as obtained from the expectation values of $2Q_1Q_{-1}$ and Q_0^2 in Appendix C, for values of N in the range $N=0,1,...,5$ and of the number V of valence particles for $V=0,4,8,12$. The deformation β decreases slightly with N but increases with V, in a similar way as for the case $N=M=0$ discussed in the previous paragraph.

C. Spectra and shapes for the su(2) ⊕ u(1) subalgebra

The most interesting of the Hamiltonians (4.20) associated with su(2) ⊕ u(1) subalgebra is the one corresponding to the Elliott type[25] quadrupole–quadrupole interaction in the three-dimensional case, which is related to the Casimir operators of the su(3) and o(3) Lie algebras. In the two-dimensional case it will be associated with the Casimir operators of su(2) and o(2) and given by

$$J_1J_{-1}+J_{-1}J_1=J_0^2-J^2, \tag{6.14}$$

whose spectrum is

$$M^2-j(j+1). \tag{6.15}$$

Clearly we have rotational bands characterized by j with the energy of the levels increasing with the squares of the angular momentum M, where the latter is in the interval $-j\leqslant M\leqslant j$. Note that the number N of excitation quanta is also an integral of motion, and N and j are related through (3.13).

To get the shape we first must calculate with the help of (3.17) the elements of the finite matrix

$$\|\langle (ws)v'\mu'NM\,|J^2|\,(ws)v\mu NM\rangle\|, \tag{6.16}$$

where $(ws),N,M$ are fixed and v,μ are limited by the inequalities (6.11). By diagonalizing this matrix we obtain its eigenvalues, which of course must be of the type $j(j+1)$ with j integer or half-integer depending on s, but also get the orthogonal matrix whose elements give the coefficients with which to express the eigenstates of (6.14) as linear combinations of the $|(ws)v\mu NM\rangle$ of the type (5.15). The calculation of the shape follows then the steps indicated in Sec. V.

In Table I we give the deformation β of the state $M=0$ of the lowest rotational band, i.e., the maximum eigenvalue j of the matrix (6.16), as function of N in the range $N=0,1,...,5$ of $V=0,4,8,12$. We see that β increases both as function of N and V, and that its values in the case of this rotational spectrum are consistently larger than those for the vibrational one associated with the Hamiltonian $I'^2+I''^2$ of the previous subsection.

In Table II we give also the deformation for closed shells and in the middle of the shell, for $N=3$, $M=0$, but for higher bands, i.e., for j's lower than the maximum value.

D. Spectra and shapes for the sp(2,R) ⊕ o(2) subalgebra

The Hamiltonian associated with the Casimir operator I^2 of sp(2,R) given by (2.12) has the eigenvalue

$$\lambda(\lambda-1)=w^2+w(2\Lambda-1)+\Lambda(\Lambda-1), \tag{6.17}$$

TABLE I. The deformation parameter β as a function of N in the range $N=0,1,...,5$ and of $V=0,4,8,12$. The parameter β is given for the limits sp'(2,R) ⊕ sp"(2,R), su(2) ⊕ u(1), and sp(2,R) ⊕ o(2). For sp'(2,R) ⊕ sp"(2,R) we took the state $|(ws)NONO\rangle$, for su(2) ⊕ u(1) the $M=0$ state of the lowest rotational band, and for sp(2) ⊕ o(2) the $M=0$ state with $\Lambda=0$. The parameter β decreases slightly for sp'(2,R) ⊕ sp"(2,R), increases for su(2) ⊕ u(1), and keeps constant for sp(2,R) ⊕ o(2).

	V	$N=0$	$N=1$	$N=2$	$N=3$	$N=4$	$N=5$
sp'(2,R) ⊕ sp"(2,R)	0	0.074	0.074	0.074	0.074	0.074	0.074
su(2) ⊕ u(1)	0	0.074	0.074	0.096	0.110	0.122	0.133
sp(2,R) ⊕ o(2)	0	0.074	0.074	0.074	0.074	0.074	0.074
sp'(2,R) ⊕ sp"(2,R)	4	0.096	0.095	0.095	0.094	0.094	0.093
su(2) ⊕ u(1)	4	0.096	0.111	0.124	0.136	0.148	0.158
sp(2,R) ⊕ o(2)	4	0.096	0.096	0.096	0.096	0.096	0.096
sp'(2,R) ⊕ sp"(2,R)	8	0.120	0.119	0.118	0.117	0.116	0.115
su(2) ⊕ u(1)	8	0.120	0.133	0.145	0.156	0.166	0.175
sp(2,R) ⊕ o(2)	8	0.120	0.120	0.120	0.120	0.120	0.120
sp'(2,R) ⊕ sp"(2,R)	12	0.128	0.127	0.126	0.125	0.124	0.123
su(2) ⊕ u(1)	12	0.128	0.140	0.150	0.161	0.171	0.180
sp(2,R) ⊕ o(2)	12	0.128	0.128	0.128	0.128	0.128	0.128

TABLE II. Deformation parameter β for the first band heads which are states with $M=0$. The number of excited quanta is $N=3$. Values of β are given for closed ($V=0$) and half-filled shell ($V=12$) as well for the two limiting cases sp(2,R) ⊕ o(2) and su(2) ⊕ u(1). The states of sp(2,R) ⊕ o(2) are abbreviated by $|q\Lambda\rangle$ and those of su(2) ⊕ u(1) by $|lj\rangle$.

	$\|lj\rangle$	su(2) ⊕ u(1)	$\|q\Lambda\rangle$	sp(2,R) ⊕ o(2)
		$V=0$		
1. Band	$\|3,3\rangle$	0.110	$\|0,0\rangle$	0.074
2. Band	$\|1,1\rangle$	0.098	$\|1,2\rangle$	0.127
		$V=12$		
	$\|lj\rangle$	su(2) ⊕ u(1)	$\|q\Lambda\rangle$	sp(2,R) ⊕ o(2)
1. Band	$\|3,27\rangle$	0.161	$\|0,0\rangle$	0.128
2. Band	$\|3,26\rangle$	0.164	$\|0,1\rangle$	0.144
3. Band	$\|3,25\rangle$	0.162	$\|1,1\rangle$	0.144
4. Band	$\|1,25\rangle$	0.159	$\|0,2\rangle$	0.157

where we made use of the relation (3.15). From the discussion in Ref. 2 the Λ [which corresponds to the $(L/2)$ appearing there] takes the values $\Lambda = N, N-1,...,1,0$, and thus for a small number of excitation quanta, e.g., for N between 0 and 10, the w of (6.8) is much larger than Λ and the eigenvalue of I^2 will be given approximately by

$$w^2 + w(2\Lambda - 1). \quad (6.18)$$

As w is fixed we have a vibrational spectra similar to the one appearing for the case $sp'(2,R) \oplus sp''(2,R)$.

For the shape we have to calculate the matrix elements of I^2 with respect to the states $|(ws)\nu\mu NM\rangle$ which we can do with the help of (3.17) and (3.20). For fixed $(ws),N,M$ we have again a finite matrix and the calculation of the deformation parameter β proceeds in analogous fashion to what was discussed in the previous subsection for J^2.

In Table I we give the deformation β of the eigenstates of I^2 with eigenvalue $\Lambda = 0$ and angular momentum $M = 0$ for $N = 0,1,...,5$ and $V = 0,4,8,12$. The β's increase with V but remain unchanged with N, and for this vibrational spectra these are consistently lower than for the rotational one of the $su(2) \oplus u(1)$ subalgebra. In Table II we give the deformations for states with $N = 3$ and some Λ's that are larger than 0.

E. Transitional Hamiltonians involving the subalgebras $su(2) \oplus u(1)$ and $sp(2,R) \oplus o(2)$

We shall now consider the Hamiltonian

$$H = (1-x)(5/w)I^2 + x(J_0^2 - J^2), \quad (6.19)$$

where x is a real parameter in the interval $0 \leqslant x \leqslant 1$, while J_0 is the angular momentum and J^2, I^2 are given by (2.9), (2.12). In (6.19) we multiplied the operator I^2 by $(5/w)$ to eliminate the factor w in its eigenvalue (6.18) and have the Λ appearing there multiplied by a factor of 10, which makes it easier to draw the energy levels as functions of x.

For x in the open interval $0 < x < 1$, the spectrum of the Hamiltonian (6.19) cannot be given in closed form as it involves the Casimir operators of the two different subalgebras $sp(2,R) \oplus o(2)$ and $su(2) \oplus u(1)$. To obtain its eigenvalues we need to calculate and diagonalize a matrix of the type (6.16) in which we replace J^2 by the H of (6.19). The shape of the corresponding eigenfunctions can the be determined by the same procedure outlined in Sec. VI C.

In Figs. 4(a) and 4(b) we give, respectively, the energy levels of H as function of x in the case of closed shells $V = 0$ and half-open ones $V = 12$, in both taking $N = 3$. In the former case $s = 0$ and from (3.13) we see that $j = 1,3$ while in the latter $s = 24$ and thus $j = 27,26,25^2,24^2,23^2,22,21$, where the exponent indicates the multiplicity with which the angular momentum appears when it is larger than 1. At $x = 0$ and $x = 1$, we give, respectively, the irreps Λ and j associated with $sp(2,R)$ and $su(2)$. The lowest levels correspond to the smallest values of Λ and largest of j, which in the latter case imply $j = 3$ for $V = 0$ and $j = 27$ for $V = 12$. We also indicate at the right-hand side of Figs. 4 the angular momentum M associated with the energy levels. The lowest level has been normalized to zero energy and its value has been subtracted from all the others.

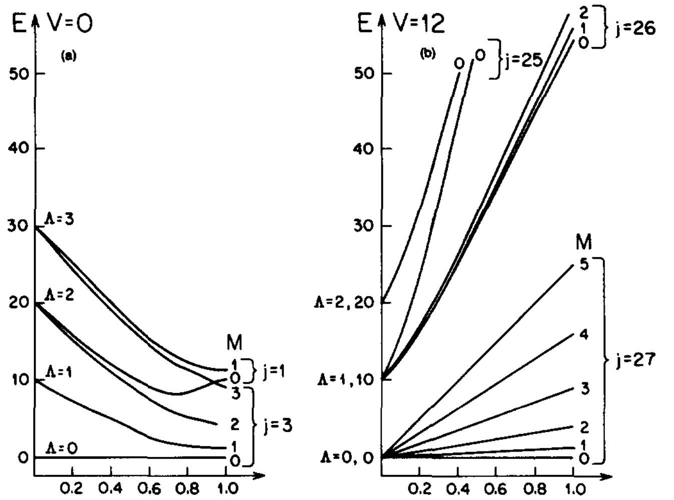


FIG. 4. The spectrum of the transitional Hamiltonian (6.19) as a function of the parameter x for (a) the closed and (b) the half-filled shell. The open shell is the sixth oscillator shell. On the left and right of each curve appear the quantum numbers Λ and j which are irreps of $sp(2,R)$ and $su(2)$ in the limiting cases $x = 0$ and 1. The number of excitation quanta is 3.

In Fig. 5 we give the deformation parameter β and eccentricity ϵ for the lowest energy level of H, when $N = 3$ and $M = 0$, as functions of the number of valence particles V in the two limiting cases $x = 0$ and $x = 1$. For x in the open interval $0 < x < 1$, the corresponding curve will be between the two drawn.

Finally we consider a Hamiltonian of the type (6.19) in which we eliminate the angular momentum term J_0^2. For

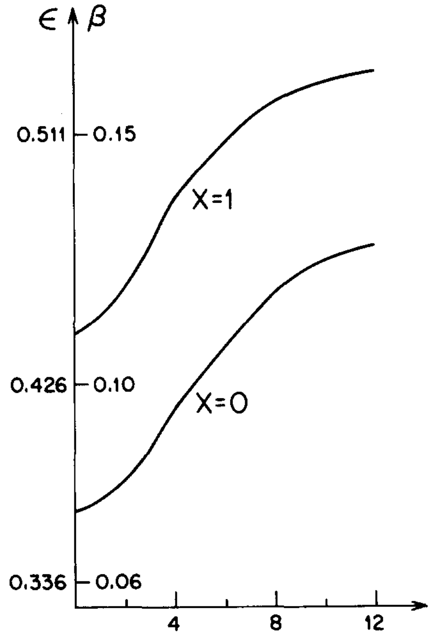


FIG. 5. The deformation β and eccentricity ϵ as a function of the number of valence nucleons V for the lowest $M = 0$ state. The number of excited quanta is $N = 3$. The calculation is done with the transitional Hamiltonian (6.19). The $x = 0$ and $x = 1$ values correspond, respectively, to the "vibrational" and "rotational" limit. There is an increase in $\beta(\epsilon)$ going from closed shell ($V = 0$) to the half-filled shell ($V = 12$). The open shell is the sixth oscillator shell.

FIG. 6. Spectrum of the transitional Hamiltonian (6.19) without the J_0^2 term. The number of excited quanta is $N = 3$ and the number of valence nucleons is $V = 12$. The open shell is the sixth oscillator shell. The energy of the ground state $M = 0$ is normalized to zero. There is a breaking of degeneracy in the transitional region $0 < x < 1$ due to the interaction of I^2 and J^2. The absolute value of the energy E, however, is notably smaller by two orders of magnitude as compared with the Hamiltonian that includes the complete quadrupole–quadrupole interaction, i.e., the J_0^2 term.

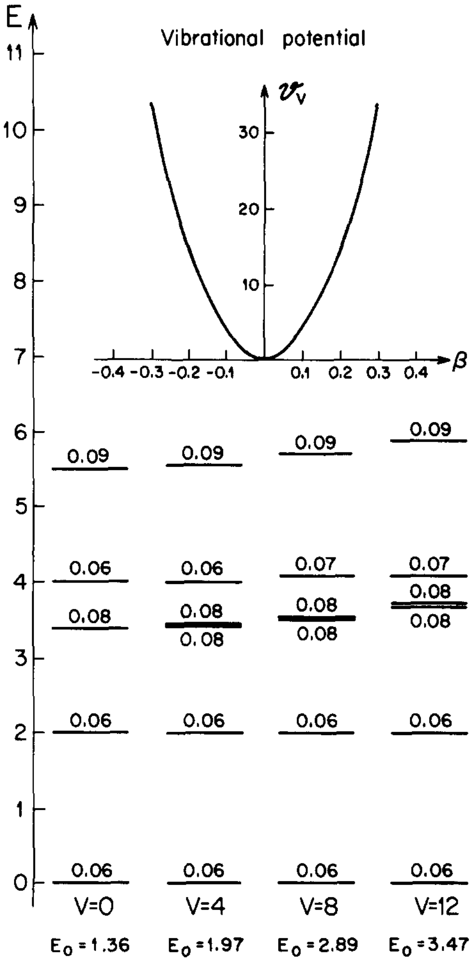


FIG. 7. Spectra as a function of the number of valence nucleons for the transitional Hamiltonian (6.22) with a quadratic (vibrational) cm(2) potential $\mathscr{V}_V$. The classical potential is indicated in the upper half of the figure. Here E_0 gives the ground state energy which increases for increasing number of valence nucleons. A special characteristic is the "computational" degeneracy of the levels, i.e., in numerical calculations no significant breaking of degeneracy was found. Each level, except for $V = 0$, has more than one angular momentum state. For example the ground level contains the states $M = 0$ up to the maximum value of $M = s$ given by (6.8). For $V = 0$ we expect a spectrum similar to the two-dimensional harmonic oscillator, i.e., for zero oscillator quanta $M = 0$, for one $M = 0,1$ for two $M = 0^2,1,2$, etc. Because the cm(2) interaction in (6.22) does not commute with $\mathscr{N}$ we expect a breaking of degeneracy of the oscillator shells. The spectrum for $V = 0$ indeed can be correlated to the two-dimensional harmonic oscillator of Fig. 2 with steps of 2 in the excited quanta a (note that $N = a/2$). Here M is related to b of Fig. 2 by $M = b/2$. The ground state is not degenerate for $V = 0$ and has angular momentum $M = 0$. The next two are $M = 0$ and 1 corresponding to the one excitation quanta $N = 1$. The next states are $M = 0,1$ and $M = 0,2$, where the latter are not plotted in the figure. It gets, however, more difficult to associate a state to a given number of excited quanta due to the cm(2) interaction which mixes them. Above each state the deformation β is indicated. Note that for all states β is small.

$x = 1$ the rotational band at the right-hand side of Fig. 4(b) that is associated with $j = 27$, collapses to a single level, but in the open interval $0 < x < 1$ the interaction between I^2 and J^2 breaks the degeneracy in the angular momenta. In Fig. 6 we illustrate this breaking for the first six angular momenta values $M = 0,1,...,5$ when the number of valence particles is $V = 12$.

Again we normalize our ground state energy to zero, and as the scale of energies in Fig. 6 is one hundreth of that of Fig. 4(b), we see that this breaking is much smaller than the one obtained by the action of the angular momentum term.

F. Transitional Hamiltonians involving the subalgebras cm(2) and u(1)

The subalgebra cm(2) of Sec. II D contains the interesting term

$$2Q_1Q_{-1} = Q_0^2 - Q^2, \tag{6.20}$$

which from (5.1) corresponds to the full physical quadrupole interaction, that contrasts with the one of the Elliott[25] type for the two-dimensional case, i.e., $J_1J_{-1} + J_{-1}J_1$, which is restricted to a single shell of the harmonic oscillator. The term (6.20) cannot be considered as a Hamiltonian by itself, as we must add to it the kinetic energy which, from (4.7j) and (5.1), is given by

$$2\mathscr{N} - Q_0. \tag{6.21}$$

The appearance of $\mathscr{N}$, which is a generator of a u(1) subalgebra, together with powers and products of the generators of cm(2), indicates that our Hamiltonians are of the transitional type and we shall discuss only two of them

$$H_V = \mathscr{N} + y(2Q_1Q_{-1}), \tag{6.22}$$

$$H_R = \mathscr{N} + y'(2Q_1Q_{-1}) + y''(2Q_1Q_{-1})^2, \tag{6.23}$$

where y, y', y'' are some real numbers and the indices V, R for the Hamiltonians indicate that they are of the vibrational or rotational type, as will be shown later when discussing Figs. 7 and 8.

To obtain the eigenvalues and eigenfunctions of the

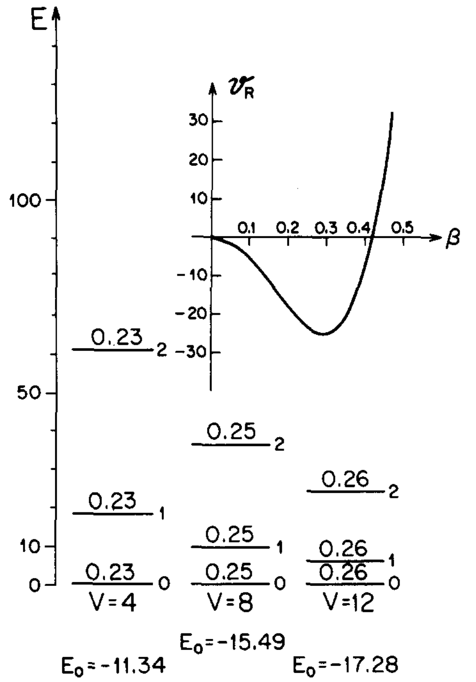


FIG. 8. Spectra as a function of the number of valence nucleons for the transitional Hamiltonian (6.23) with a rotational cm(2) potential $\mathscr{V}_R$ which is indicated in the upper half of the figure. The minimum of the classical potential is chosen to be at $\beta = 0.3$ and at -25 in energy. The ground state is normalized to zero. The ground state energy is denoted by E_0. Note that E_0 increases with decreasing number of valence nucleons V. We give only the lowest band which is rotational for all values of V. The spacing between states is decreasing with increasing V. Above each state, whose angular momentum is indicated at its right-hand side, we give the deformation β. Note that the deformation is big and near to the position of the minimum of the classical potential.

Hamiltonian (6.22), (6.23) we can again make use of the complete set of states $|(ws)\nu\mu NM\rangle$, with respect to which we need to evaluate the elements of the matrix

$$\|\langle(ws)\nu'\mu'N'M|(2Q_1Q_{-1})^m|(ws)\nu\mu NM\rangle\|, \quad (6.24)$$

where $m = 1$ or 2. Note that as $2Q_1Q_{-1}$ does not commute with $\mathscr{N}$ we have different indices N,N' in bra and ket and our matrix is now infinite. For calculations we have to consider N's up to some maximum value N_0 that assures convergence for the lowest lying eigenvalues and eigenfunctions, where the former are obtained by diagonalizing the now finite matrix and the latter can be expressed in the form (5.15) with the coefficients coming from the orthogonal matrix required for the diagonalization of (6.24). From the explicit expression of the eigenfunctions the shape can be calculated by the procedures outlined in Sec. V.

In the lower part of Fig. 7 we have the energy levels of H_V when $y = 10^{-2}$ as function of the number of valence particles $V = 0,4,8,12$. The angular momentum M of these levels is degenerate as discussed in the figure caption and on top of each level we give the value of the deformation parameter. Below the number of valence particles we give the energy of the lowest level which in the figure were normalized to 0. The almost equidistant position of the levels and their low deformation parameter sugget that H_V is a vibrational Hamiltonian.

In the lower part of Fig. 8 we have the energy levels of H_R when $y' = -10^{-2}$, $y'' = 10^{-6}$ as function of the number of valence particles $V = 4,8,12$. We exclude the closed shell case $V = 0$ because the high separation of the levels would not fit in the figure. The angular momentum of the levels is indicated on their right and the deformation parameter β appears above them. Below the number of valence particles we give the energy of the lowest level which in the figure were normalized to 0. The levels follow closely the M^2 rule and that, together with their large deformation parameter β, suggest that H_R is a bona fide rotational Hamiltonian.

The interpretation of H_V and H_R as vibrational and rotational Hamiltonians, that follows from their spectra and shape in the quantum picture, is corroborated by a classical analysis. We note from (5.8)–(5.12) that

$$2Q_1Q_{-1} = r^2 \sin^2\theta = \tfrac{1}{4}\rho^4\beta^2, \quad (6.25)$$

where ρ^2 can be correlated with the mean square radius for the two-dimensional A-body system and β is the deformation parameter. The value of ρ^2 can be estimated by an analysis, discussed in Appendix D, that extends to two-dimensional space the well-known arguments for determining the mean square radius in three dimensions with the help of the nuclear shell model.[20,26] From this analysis one obtains

$$\rho^2 \simeq \tfrac{1}{2}A^{3/2}, \quad (6.26)$$

and thus the classical estimate for the potential energy in H_V, H_R with the above values of y, y', y'', is then given by

$$\mathscr{V}_V = 553\beta^2 \quad (6.27a)$$

$$\mathscr{V}_R = -553\beta^2 + 3058\beta^4, \quad (6.27b)$$

where in both cases we used the A associated with 12 valence particles, i.e., $A = 96$.

The potentials (6.27) are of the type used in the Frankfurt[27] collective model, but for two[28] instead of three dimensions, and they are drawn in the upper part of Figs. 7 and 8 for the $\mathscr{V}_V$ and $\mathscr{V}_R$, respectively. We see that the vibrational potential $\mathscr{V}_V$ has a minimum for $\beta = 0$, i.e., there is no deformation, while the minimum of the rotational potential $\mathscr{V}_R$ occurs at $\beta = 0.3$ which indicates a strong deformation, and agrees in order of magnitude with the values of β predicted in the quantum picture and given above the energy levels in Fig. 8.

We note that the Hamiltonian H_R of (6.23) is no longer of the type of those discussed in Sec. IV, as it is not limited to terms up to second degree in the generators of sp(4,R) but contains a fourth-order term $(2Q_1Q_{-1})^2$. This type of term is the two-dimensional equivalent of those used in the nuclear symplectic model to produce strong deformations.[29]

Having analyzed the spectra of Hamiltonians associated with Casimir operators of different maximal subalgebras, as well as of relevant transitional Hamiltonians involving more than one subalgebra, and determined also the shape of their eigenfunctions, we proceed to discuss the conclusions that follow from all our analyses.

VII. CONCLUSIONS

We first want to note that the discussion in Secs. II and III of this paper parallels closely the analysis followed in collective algebraic models of the nucleus and, in particular, in the IBA. One first introduces a basic algebra, in our case $sp(4,R)$ and in IBA I it is $u(6)$. We then consider the maximal subalgebras that include among their generators the operators of angular momentum. In our case these were the five discussed in Sec. II and for the IBA they are[14,15] $u(5)$, $su(3)$, $o(6)$. One then chooses one of the subalgebras to characterize by its irreps a complete set of states to be used in the calculations. In our case the most convenient one proved to be $sp'(2,R) \oplus sp''(2,R)$, while in the IBA some authors use $su(3)$ (Ref. 14) and others $u(5)$ (Ref. 15). Once one has the complete set of states one can calculate the matrix elements with respect to them of the generators of our basic algebras and of the Casimir operators of their subalgebras, as was done here in (3.17) and for the IBA in Refs. 14 and 15 among many others.

The next step is to study the spectra of Hamiltonians in the enveloping algebras of our problems. There are, both in our present analysis and in the IBA, what we could call pure Hamiltonians, i.e., those associated with a single chain of subalgebras whose spectra is given by a closed formula, and transitional ones which involve several subalgebras, whose spectra has to be calculated numerically using the matrix elements of the Casimir operators of the subalgebras mentioned above.

In our case the subalgebras $sp'(2,R) \oplus sp''(2,R)$, $su(2) \oplus u(1)$, and $sp(2,R) \oplus u(1)$ are the only ones that contain the generator $\mathscr{N}$ of $u(1)$ which, from (6.21), can be associated with the kinetic energy. We shall restrict ourselves to them when comparing our analysis to that of the IBA, and, as shown in Sec. VI B–VI D they give rise to one rotational and two vibrational types of spectra. We also discussed numerically the transitional Hamiltonian that includes the subalgebras $sp(2,R) \oplus o(2)$ and $su(2) \oplus u(1)$, in which we pass continuously from vibrational to rotational spectra.

In the IBA case the subalgebras $u(5)$, $su(3)$, $o(6)$ give rise, respectively, to vibrational, rotational, and γ-unstable types of spectra,[14,15] which have actually been identified in some nuclei, as well as transitional spectra associated with mixtures of the Casimir operators of the above subalgebras.

So far we have been speaking only of spectra, but the appearance of vibrational or rotational types suggests, both in this analysis and in the IBA, that we are dealing, respectively, with systems of small or large deformations.

In the present problem we have though an actual microscopic theory of collective motions, as illustrated in (4.7) by the realization of the generators of $sp(4,R)$ in terms of the creation and annihilation operators of the A-body system. We can then define the components (5.1) of the quadrupole tensor and with its help actually discuss the shape, i.e., the deformation parameter β of the eigenfunctions of our Hamiltonians. We find that the states associated with the rotational levels are more deformed than those of the vibrational ones, *but* if we go to higher and higher shells, where we approach the classical interpretation, in *both* the vibrational and rotational pictures the deformation goes to zero.

Clearly the presence of rotational spectra in an algebraic model is not sufficient to guarantee a strong deformation of the system. Thus in the IBA, and in other collective models based on group theory, analyses have[30] or should be made to show, independently of the rotational characteristics of the spectra, that a strong deformation is present in the eigenstates of the Hamiltonian.

In the analysis of Sec. VI F of this paper we saw that, to get strong deformations, it was essential to have higher powers of the $2Q_1Q_{-1}$ term associated with the Casimir operator of $cm(2)$. These powers are the equivalent, in our dimensional A-body system, of the generalized type of Bohr–Mottelson potential that was introduced by Greiner and collaborators.[27]

We want now to turn our attention to the real symplectic model, i.e., the one in which A nucleons move in the three-dimensional physical space. Much work has been done in this model and its successes and drawbacks have been discussed in other publications.[1,3,6–12,29]

We, in collaboration with other researchers, intend to review this model following step by step the procedure outlined in the present paper for the two-dimensional case. If we succeed in this endeavor, we would consider as finished our program on collectivity and geometry.

ACKNOWLEDGMENTS

The authors would like to thank R. T. Sharp and P. Winternitz for helpful discussions.

APPENDIX A: THE COMPLETE ORTHONORMAL BASIS

The purpose of this appendix is to give an outline of the derivation of the states (3.9) and (3.10), which are an orthonormal basis for the irrep (w,s) of $sp(4,R)$ classified by the chain of subalgebras

$$sp(4,R) \supset sp'(2,R) \times sp''(2,R) \supset o'(2) \times o''(2). \quad \text{(A1)}$$

As discussed in the text, these basis states have a general form

$$|(ws)\nu\mu NM\rangle = \sum_{\sigma\tau} A_{\sigma\tau} |(ws)\sigma\tau NM\rangle, \quad \text{(A2)}$$

with the ket on the right-hand side given in (3.5).

As a first step in our construction we find the member of the basis which has the lowest weight in the irrep (ν,μ) of $sp'(2,R) \oplus sp''(2,R)$. Such state, denoted for shortness as $|LW\rangle$ obeys the equations

$$I'_0|LW\rangle = \tfrac{1}{2}(w+\nu+\mu)|LW\rangle, \quad \text{(A3a)}$$

$$I''_0|LW\rangle = \tfrac{1}{2}(w+\nu-\mu)|LW\rangle, \quad \text{(A3b)}$$

$$I'_{-1}|LW\rangle = 0, \quad \text{(A3c)}$$

$$I''_{-1}|LW\rangle = 0. \quad \text{(A3d)}$$

From (3.6a), (3.6b), and the fact that $2I'_0 = \mathscr{N} + J_0$, $2I''_0 = \mathscr{N} - J_0$, it follows that a state of the form (A2) will satisfy Eqs. (A3a) and (A3b) provided we set $N = \nu$ and $M = \mu$. Therefore $|LW\rangle$ has a realization of this type:

$$|\mathrm{LW}\rangle = \sum_{\sigma,\tau} (-)^{\nu-\tau} 2^{(1/2)(s+\sigma+\nu-\tau)}(s-\sigma)!$$
$$\times A'_{\sigma\tau}(B_1^\dagger)^{(1/2)(\nu+\mu-\sigma-\tau)}(B_0^\dagger)^\tau$$
$$\times (B_{-1}^\dagger)^{(1/2)(\nu-\mu+\sigma-\tau)}(J_1)^{s+\sigma}|ws\rangle, \qquad (A4)$$

where the constant factors standing on the left of $A'_{\sigma\tau}$ were introduced for the sake of later convenience.

Then Eqs. (A3c) and (A3d) give us two recursion relations for the determination of the unknown coefficients $A'_{\sigma\tau}$ in (A4). Explicitly, the recursion formulas are

$$(\nu+\mu-\sigma-\tau+2)(2w+\nu+\mu+\sigma+\tau-4)A'_{\sigma-1,\tau-1} + 2\tau(s+\sigma)A'_{\sigma\tau} + \tau(\tau+1)A'_{\sigma-1,\tau+1} = 0, \qquad (A5)$$

$$(\nu-\mu+\sigma-\tau+2)(2w+\nu-\mu-\sigma+\tau-4)A'_{\sigma+1,\tau-1} + 2\tau(s-\sigma)A'_{\sigma\tau} + \tau(\tau+1)A'_{\sigma+1,\tau+1} = 0, \qquad (A6)$$

and by direct substitution on them we can check that their solution is

$$A'_{\sigma\tau} = C(-1)^{(1/2)(\nu-\mu+\sigma-\tau)}\Gamma\left(w-1+\frac{\nu+\mu+\sigma+\tau}{2}\right)\Gamma\left(w-1+\frac{\nu-\mu-\sigma+\tau}{2}\right)2^{\tau-\nu}(\tau!)^{-1}$$
$$\times\sum_r(-1)^r\left[r!(s+\mu-r)!(s+\sigma-r)!(r-\mu-\sigma)!\left(\frac{\nu-\mu-\sigma-\tau}{2}-s+r\right)!\right.$$
$$\left.\times\Gamma\left(w-1+\frac{\nu+\mu+\sigma+\tau}{2}-r\right)\right]^{-1}. \qquad (A7)$$

In (A7) C is a normalization coefficient which must be given a value such that $\langle\mathrm{LW}|\mathrm{LW}\rangle = 1$. We have found that, with an arbitrary phase factor set equal to 1,

$$C = \left[\frac{(s+\mu)!(s-\mu)!\nu!(2w-3)!\Gamma(w-s+\nu-1)\Gamma(w-s-1)\Gamma(w+s)}{(2s)!(2w+\nu-3)!\Gamma(w+\mu-1)\Gamma(w-\mu-1)\Gamma(w+\mu+\nu-1)\Gamma(w-\mu+\nu-1)\Gamma(w+s+\nu)}\right]^{1/2}. \qquad (A8)$$

We have thus determined the state $|\mathrm{LW}\rangle \equiv |(ws)\nu\mu, N=\nu, M=\mu\rangle$ with lowest weight in the algebra $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$. The general state of the basis can be obtained by application of powers of the raising operators I'_1, I''_1, of each $\mathrm{sp}(2,R)$ algebra on $|\mathrm{LW}\rangle$. Recalling that $I'_1 = \frac{1}{2}B_1^\dagger, I''_1 = \frac{1}{2}B_{-1}^\dagger$, and supplying the appropriate normalization factors, we have

$$|(ws)\nu\mu NM\rangle$$
$$= \left[\frac{\Gamma(w+\nu+\mu)\Gamma(w+\nu-\mu)}{\Gamma(w+(N+M+\nu+\mu)/2)\Gamma(w+(N-M+\nu-\mu)/2)((N+M-\nu-\mu)/2)!((N-M-\nu+\mu)/2)!}\right]^{1/2}$$
$$\times 2^{-(N-\nu)}(B_1^\dagger)^{(1/2)(N+M-\nu-\mu)}(B_{-1}^\dagger)^{(1/2)(N-M-\nu+\mu)}|(ws)\nu\mu, N=\nu, M=\mu\rangle. \qquad (A9)$$

Introducing here (A4), (A7), and (A8) and merging factors we deduce the detailed expression for the $\mathrm{sp}(4,R)$ basis states given in (3.9) and (3.10).

APPENDIX B: MATRIX REPRESENTATION OF GENERATORS OF sp (4,R)

In this appendix we indicate the procedure that we have followed to obtain the matrix elements of the four generators of $\mathrm{sp}(4,R)$ given in (3.18) with respect to the basis states (3.9).

To begin with, in the notation of (3.18) we have that

$$[I'_{\pm 1}, T^{(1/2)(1/2)}_{r'r''}] = (r' \mp \tfrac{1}{2})T^{(1/2)(1/2)}_{r'\pm 1,r''},$$
$$[I'_0, T^{(1/2)(1/2)}_{r'r''}] = r' T^{(1/2)(1/2)}_{r'r''}, \qquad (B1)$$

$$[I''_{\pm 1}, T^{(1/2)(1/2)}_{r'r''}] = (r'' \mp \tfrac{1}{2})T^{(1/2)(1/2)}_{r',r''\pm 1},$$
$$[I''_0, T^{(1/2)(1/2)}_{r'r''}] = r'' T^{(1/2)(1/2)}_{r'r''}, \qquad (B2)$$

where r', r'' can take any of the two values $\frac{1}{2}, -\frac{1}{2}$. Formulas (B1) and (B2) make the four operators $\{T^{(1/2)(1/2)}_{r'r''}\}$ qualify as components of an irreducible tensor associated to the nonunitary irrep $(\frac{1}{2},\frac{1}{2})$ of $\mathrm{sp}'(2,R) \oplus \mathrm{sp}''(2,R)$. This allows the application of a generalized Wigner–Eckart theorem to obtain the matrix elements of the tensor $\mathbf{T}^{(1/2)(1/2)}$ with respect to the states (3.9), namely

$$\langle(ws)\bar\lambda'\bar\lambda''\bar\rho'\bar\rho''|T^{(1/2)(1/2)}_{r'r''}|(ws)\lambda'\lambda''\rho'\rho''\rangle$$
$$= \langle ws\bar\lambda'\bar\lambda''\|\mathbf{T}^{(1/2)(1/2)}\|ws\lambda'\lambda''\rangle$$
$$\times\langle\lambda'\rho';\tfrac{1}{2}r'|\bar\lambda'\bar\rho'\rangle\langle\lambda''\rho'';\tfrac{1}{2}r''|\bar\lambda''\bar\rho''\rangle, \qquad (B3)$$

where on the right-hand side appear a reduced matrix element of the tensor and two Clebsch–Gordan coefficients of the group $\mathrm{Sp}(2,R)$, in that order.

We have used in (B3) a notation which seems the most natural for that particular purpose, the connection of λ',λ'' with w, ν, μ was already given in (3.8a) and (3.8b) and the ρ',ρ'' are related to the labels w, N, M of the states (3.9) by

$$\rho' = \tfrac{1}{2}(w+N+M), \quad \rho'' = \tfrac{1}{2}(w+N-M). \qquad (B4)$$

The Clebsch–Gordan coefficients of $\mathrm{Sp}(2,R)$ needed in (B3) are precisely those studied by Ui in Ref. 22.

There are four distinct reduced matrix elements on the tensor $\mathbf{T}^{(1/2)(1/2)}$. Their values are deduced, as usual, by direct evaluation of (B3) with particular values of ρ',ρ''. Still with the notation of λ',λ'' the reduced matrix elements were obtained as

$$\langle ws,\lambda'+\tfrac{1}{2},\lambda''+\tfrac{1}{2}\|\mathbf{T}^{(1/2)(1/2)}\|ws\lambda'\lambda''\rangle$$
$$= \left[\frac{(\lambda'+\lambda''-w+1)(\lambda'+\lambda''-s-1)(\lambda'+\lambda''+s)(\lambda'+\lambda''+w-2)}{(2\lambda')(2\lambda'')}\right]^{1/2}, \qquad (B5)$$

$$\langle ws,\lambda' - \tfrac{1}{2},\lambda'' - \tfrac{1}{2}\|\mathbf{T}^{(1/2)(1/2)}\|ws\lambda'\lambda''\rangle = \left[\frac{(\lambda'+\lambda''-w)(\lambda'+\lambda''-s-2)(\lambda'+\lambda''+s-1)(\lambda'+\lambda''+w-3)}{(2\lambda'-2)(2\lambda''-2)}\right]^{1/2}, \tag{B6}$$

$$\langle ws,\lambda' + \tfrac{1}{2},\lambda'' - \tfrac{1}{2}\|\mathbf{T}^{(1/2)(1/2)}\|ws\lambda'\lambda''\rangle = -\left[\frac{(\lambda'-\lambda''+s+1)(s-\lambda'+\lambda'')(\lambda'-\lambda''+w-1)(-\lambda'+\lambda''+w-2)}{(2\lambda')(2\lambda''-2)}\right]^{1/2}, \tag{B7}$$

$$\langle ws,\lambda' - \tfrac{1}{2},\lambda'' + \tfrac{1}{2}\|\mathbf{T}^{(1/2)(1/2)}\|ws\lambda'\lambda''\rangle = -\left[\frac{(\lambda'-\lambda''+s)(s-\lambda'+\lambda''+1)(\lambda'-\lambda''+w-2)(-\lambda'+\lambda''+w-1)}{(2\lambda'-2)(2\lambda'')}\right]^{1/2}. \tag{B8}$$

Once we have available the relations (B5)–(B8), the matrix elements of any component of the tensor with respect to arbitrary sp$(4,R)$ states can be obtained from (B3) with the help of the Clebsch–Gordan coefficients tabulated by Ui.[22] By this procedure were obtained (3.20a) and (3.20b).

APPENDIX C: EXPECTATION VALUES OF TWO OPERATORS

The expectation values of the two operators discussed in Sec. VI are

$$\begin{aligned}\langle (ws)\nu\mu NM|2Q_1Q_{-1}|(ws)\nu\mu NM\rangle &= \tfrac{1}{2}(N+W)^2 + \tfrac{1}{2}M(M+2) - \tfrac{1}{2}(\nu+w)(\nu+w-2) - \tfrac{1}{2}\mu^2\\ &+ \frac{\nu(w-s+\nu-2)(w+s+\nu-1)(2w+\nu-3)(N+M-\mu-\nu+2)(N-M+\nu-\mu+2w)}{4(w+\mu+\nu-2)(w+\mu+\nu-1)(w-\mu+\nu-2)(w-\mu+\nu-1)}\\ &+ \frac{(\nu+1)(w-s+\nu-1)(w+s+\nu)(2w+\nu-2)(N+M+\mu+\nu+2w)(N-M+\mu-\nu)}{4(w+\mu+\nu)(w+\mu+\nu-1)(w-\mu+\nu)(w-\mu+\nu-1)}\\ &+ \frac{(\mu+s+1)(s-\mu)(w+\mu-1)(w-\mu-2)(N+M+\mu+\nu+2w)(N-M-\mu+\nu+2w-2)}{4(w+\mu+\nu)(w+\mu+\nu-1)(w-\mu+\nu-2)(w-\mu+\nu-1)}\\ &+ \frac{(\mu+s)(s-\mu+1)(w+\mu-2)(w-\mu-1)(N+M-\mu-\nu+2)(N-M+\mu-\nu)}{4(w+\mu+\nu-2)(w+\mu+\nu-1)(w-\mu+\nu)(w-\mu+\nu-1)}\end{aligned} \tag{C1}$$

and

$$\begin{aligned}\langle (ws)\nu\mu NM|Q_0^2|(ws)\nu\mu NM\rangle &= \tfrac{1}{4}(2N+2w)(2N+2w-1)\\ &+ \frac{\nu(w-s+\nu-2)(w+s+\nu-1)(2w+\nu-3)(N+M-\mu-\nu+2)(N-M+\mu-\nu+2)}{8(w+\mu+\nu-2)(w+\mu+\nu-1)(w-\mu+\nu-2)(w-\mu+\nu-1)}\\ &+ \frac{(\nu+1)(w-s+\nu-1)(w+s+\nu)(2w+\nu-2)(N+M+\mu+\nu+2w)(N-M-\mu+\nu+2w)}{8(w+\mu+\nu)(w+\mu+\nu-1)(w-\mu+\nu)(w-\mu+\nu-1)}\\ &+ \frac{(\mu+s+1)(s-\mu)(w+\mu-1)(w-\mu-2)(N+M+\mu+\nu+2w)(N-M+\mu-\nu+2)}{8(w+\mu+\nu)(w+\mu+\nu-1)(w-\mu+\nu-2)(w-\mu+\nu-1)}\\ &+ \frac{(\mu+s)(s-\mu+1)(w+\mu-2)(w-\mu-1)(N+M-\mu-\nu+2)(N-M-\mu+\nu+2w)}{8(w+\mu+\nu-2)(w+\mu+\nu-1)(w-\mu+\nu)(w-\mu+\nu-1)}.\end{aligned} \tag{C2}$$

APPENDIX D: ESTIMATED VALUE OF ρ^2

The purpose of this Appendix is to outline arguments that allows us to estimate the value of ρ^2 in (6.26). The ρ^2 is correlated to the mean square radius via[20]

$$\langle r'^2\rangle \simeq \frac{2}{A}\sum_{s=1}^{n} x'_s \bar{x}'_s, \tag{D1}$$

where A is the number of nucleons and $A \gg 1$. The prime indicates that the coordinates carry dimensions. The dimensionless coordinates x_s, $\bar{x}_s$ used here are related to x'_s, $\bar{x}'_s$ via

$$x'_s = \sqrt{(\hbar/m\omega)}\, x_s, \quad \bar{x}'_s = \sqrt{(\hbar/m\omega)}\, \bar{x}_s. \tag{D2}$$

We are confronted with the difficulty of which values have to be used in two dimensions for m (mass of the nucleon) and ω (frequency of the oscillator). Though we are in two dimensions, we will use the same values as in three dimensions, i.e.,

$$mc^2 \simeq 1000 \text{ MeV}, \tag{D3a}$$

$$\hbar\omega \simeq 40\,A^{-1/2}. \tag{D3b}$$

The value for $\hbar\omega$ we obtained in a similar way as was done in three dimensions in p. 200 of Ref. 20. The result for the factor in (D3b) turns out to be nearly the same. The only difference lies in the A dependence, which is $A^{-1/3}$ in the three-dimensional space; ρ^2 is defined via

$$\rho^2 = 2 \sum_{s=1}^{n} x_s \bar{x}_s, \tag{D4}$$

where the dimensionless coordinates appear. With (D1) and (D2) we get

$$\langle r'^2 \rangle \simeq (\hbar/m\omega)(1/A)\,\rho^2. \tag{D5}$$

Using the values given in (D3) we get

$$\rho^2 \simeq A^{1/2} \langle r'^2 \rangle. \tag{D6}$$

For $\langle r'^2 \rangle$ we use

$$\langle r'^2 \rangle \simeq \frac{\int_0^{2\pi} d\varphi' \int_0^{R'} r'\,dr'\,r'^2}{\int_0^{2\pi} d\varphi' \int_0^{R'} r'\,dr'} = \frac{1}{2} R'^2. \tag{D7}$$

We integrated over a spherical nucleus. Here R' is the radius of that nucleus. Assuming in two dimensions a similar dependence for R' on A as in the case of three dimensions, namely

$$R' \simeq r_0 A^{1/2} \tag{D8}$$

we finally get for ρ^2,

$$\rho^2 \simeq \tfrac{1}{2} r_0^2 A^{3/2}. \tag{D9}$$

That R is given by (D8) means that the number of particles is proportional to the two-dimensional volume. We put the proportionality factor r_0 in (D8) equal to one fermi as it is similar to the value 1.2–1.4 fermi in the three-dimensional case. Thus we get the formula (6.26).

[1]M. Moshinsky, J. Math. Phys. **25**, 1555 (1984); O. Castaños, E. Chaćon, and M. Moshinsky, *ibid.* **25**, 2815 (1984); M. Moshinsky, M. Nicolescu, and R. T. Sharp, *ibid.* **26**, 2995 (1985).

[2]E. Chacón, P. Hess, and M. Moshinsky, J. Math. Phys. **25**, 1565 (1984).

[3]O. Castaños, E. Chacón, and M. Moshinsky, J. Math. Phys. **25**, 1211 (1984); O. Castaños, A. Frank, E. Chacón, P. Hess, and M. Moshinsky, *ibid.* **23**, 2537 (1982).

[4]O. Castaños, E. Chacón, M. Moshinsky, and C. Quesne, J. Math. Phys. **26**, 2107 (1985); O. Castaños, P. Kramer, and M. Moshinsky, *ibid.* **27**, 924 (1986).

[5]O. Castaños and M. Moshinsky, J. Phys. A: Math. Gen. **20**, 513 (1987).

[6]G. F. Filippov and V. I. Ovcharenko, Yad. Fiz. **30**, 646 (1979); V. C. Vasilevsky, Ya. F. Smirnov, and G. F. Filippov, *ibid.* **32**, 987 (1980); G. F. Filippov, V. I. Ovcharenko, and Ya. F. Smirnov, *Microscopic Theory of Collective Excitations in Nuclei* (Naukova Dumka, Kiev, 1981) (in Russian).

[7]V. Vanagas, *The Microscopic Nuclear Theory* in *Lecture Notes in Physics* (University of Toronto, Toronto, 1977); V. Vanagas, *The Microscopic Theory of Collective Motions in Nuclei*, in *Group Theory and its Applications in Physics—1980*, edited by T. H. Seligman (AIP, New York, 1980).

[8]L. Weaver, L. C. Biedenharn, and R. Y. Cusson, Ann. Phys. (NY) **77**, 250 (1973); O. L. Weaver, R. Y. Cusson, and L. C. Biedenharn, *ibid.* **102**, 493 (1976).

[9]G. Rosensteel and D. J. Rowe, Ann. Phys. (NY) **96**, 1 (1976); **123**, 36 (1978); **126**, 198 (1980).

[10]J. Deenen and C. Quesne, J. Math. Phys. **23**, 878, 2004 (1982); **25**, 1638, 2354 (1984); J. Phys. A: Math. Gen. **16**, 2095 (1983).

[11]P. Kramer, Ann. Phys. (NY) **141**, 254, 259 (1982); P. Kramer, Z. Papadopolos, and W. Schweizer, *ibid.* **149**, 44 (1983); Nucl. Phys. A **431**, 75 (1984).

[12]K. T. Hecht, J. Phys. A: Math. Gen. **18**, L1003 (1985); K. T. Hecht and Y. Suzuki, Nucl. Phys. A **455**, 315 (1986).

[13]J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **18**, 2259 (1977); M. Moshinsky and P. Winternitz, *ibid.* **21**, 1667 (1980).

[14]A. Arima and F. Iachello, Ann. Phys. (NY) **99**, 253 (1976); **111**, 201 (1978); **123**, 468 (1979).

[15]O. Castaños, E. Chacón, A. Frank, and M. Moshinsky, J. Math. Phys. **20**, 35 (1979).

[16]D. R. Peterson and K. T. Hecht, Nucl. Phys. A **344**, 361 (1980).

[17]A. Ya. Dzublik, V. I. Ovcharenko, A. I. Steshenko, and G. F. Filippov, Yad. Fiz. **15**, 869 (1972) [Sov. J. Nucl. Phys. **15**, 487 (1972)].

[18]W. Zickendraht, J. Math. Phys. **10**, 30 (1969); **12**, 1663 (1971).

[19]Yu. F. Smirnov and G. F. Filippov, Yad. Fiz. **27**, 73 (1978) [Sov. J. Nucl. Phys. **27**, 39 (1978)].

[20]J. M. Eisenberg and W. Greiner, *Nuclear Theory* (North-Holland, Amsterdam, 1975), Vol. 1, pp. 70 and 200.

[21]J. S. Lomont, *Applications of Finite Groups* (Academic, New York, 1959), p. 321.

[22]H. Ui, Ann. Phys. (NY) **49**, 69 (1968).

[23]E. P. Wigner, *Group Theory* (Academic, New York, 1959), p. 325.

[24]A. Bohr and B. Mottelson, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **27**, No. 16 (1953).

[25]J. P. Elliott, Proc. R. Soc. London Ser. A **245**, 128, 562 (1958).

[26]D. J. Rowe, in *Group Theory and its Applications in Physics*, edited by T. H. Seligman (AIP, New York, 1980).

[27]P. O. Hess, M. Seiwert, J. Maruhn, and W. Greiner, Z. Phys. A **296**, 147 (1980); P. O. Hess, J. Maruhn, and W. Greiner, J. Phys. G **67**, 737 (1981).

[28]E. Chacón, M. Moshinsky, and V. Vanagas, J. Math. Phys. **22**, 605 (1981).

[29]P. Park, J. Carvalho, M. Vassanji, and D. J. Rowe, Nucl. Phys. A **414**, 93 (1984); G. Rosensteel, J. P. Draayer, and K. J. Weeks, Nucl. Phys. A **419**, 1 (1984); J. P. Draayer, K. J. Weeks, and G. Rosensteel, Nucl. Phys. A **413**, 215 (1984).

[30]J. N. Ginocchio, M. W. Kirson, Phys. Rev. Lett. **44**, 1744 (1980).

Phenomenological electrodynamics in curvilinear space, with application to Rindler space

I. Brevik
Luftkrigsskolen, N-7000 Trondheim, Norway

(Received 18 December 1986; accepted for publication 29 April 1987)

The phenomenological Maxwell equations and the constitutive relations are derived in curvilinear coordinates endowed with a time-independent, time-orthogonal metric, when there is a uniform, isotropic, and homogeneous medium present. The way of formulating the electromagnetic theory chosen here is basically in agreement with the formulations given by von Laue and Arzeliès. The theory is thereafter applied to the case when the medium is at rest in Rindler space. The fundamental electromagnetic modes (the TE, TM, and TEM modes) are worked out in terms of modified Bessel functions, and the electromagnetic energy-momentum tensor is worked out in the Minkowski picture and shown to possess an interesting analogy with the case when an *inhomogeneous* medium is at rest in an *inertial* frame.

I. INTRODUCTION

Whereas the phenomenological electrodynamic theory of dielectric media in uniform motion has been subject to extensive studies, far less attention has so far been given to the case when the medium is in a state of uniform acceleration. The main purpose of the present work is to study the electrodynamic theory of a medium when it is in uniform *linear* acceleration, that is, when it is at rest in one of the wedges in Rindler space.[1] We assume a medium of the simplest kind: it is taken to possess a constant, (i.e., spatially nonvarying and nondispersive) permittivity ϵ and a constant permeability μ.

Some reasons for undertaking this kind of study are the following.

(1) The electrodynamics of continuous media is a natural generalization of the electrodynamics in vacuum. The theory, if properly constructed, has to be consistent and moreover to be testable experimentally. This point can be more complicated than it looks at first sight. Already within special relativity the phenomenological theory leads to consequences which, although consistent, may appear surprising, for instance the occurrence of negative field energies for a radiation field in a class of inertial systems due to the *spacelike* radiation four-momentum.[2] When developing the phenomenological theory to the extended case of accelerated media we subject it to another consistency test.

(2) Another motivation has its root in quantum field theory. Almost all works published so far on quantum theory in Rindler space (cf., for instance, Refs. 3–16) deal with the simple case of scalar fields, although a few of them, like the work of Candelas and Deutsch,[7b] deal with the electromagnetic field also. The extension of quantum field theory to the case of phenomenological electrodynamics would be an interesting development. However, we will not proceed so far as to present an electromagnetic quantum theory here. Some remarks, based upon the simplification of replacing the electromagnetic field by a "phenomenological" scalar field, will be made at the end of this paper.

(3) An interesting application of this kind of electromagnetism, having been put forward in particular by Lee,[17] is to use it within the context of quantum chromodynamics (QCD). Complicated QCD processes of higher order may be effectively described by a theory that is formally quite similar to electrodynamics; thus it contains color permeability and color permittivity as input parameters. Cross relationships of this kind, linking classical theories to modern ones, make it desirable also to develop the classical theories further.

In the first part of this paper (Secs. II–IV) we give a general formulation of the electromagnetic theory in curvilinear space, when there is a medium present and the metric is time independent and time orthogonal. We thus consider a class of coordinate systems somewhat broader than strictly necessary as regards the specific application to Rindler space. We find it desirable to proceed in this way, because the generalization of the theory is moderate and achieved at low costs and makes the formalism applicable to different spaces, and also because there are different versions of electrodynamic theory in the literature tending to confuse the casual reader unless the foundation of the theory is laid properly.

The second part of the paper deals with the special case when there is a medium at rest in Rindler space. Section V summarizes some of the basic features of that space, and also discusses briefly our idealized assumption about complete filling of the Rindler wedge by the medium in relation to the macroscopic strength of real materials. Section VI derives and solves the Maxwell equations for the three types of fundamental modes: the TE, TM, and TEM modes. Mathematically, modified Bessel functions of imaginary order are the key functions for the TE and TM modes. Solutions are also given for the electromagnetic potentials.

Section VII elaborates upon, and calculates explicitly, the electromagnetic energy-momentum tensor according to the Minkowski picture. It turns out that for the TE and TM modes there is a transport of electromagnetic energy *transversely* to the direction of acceleration, whereas for the TEM mode there is a transport *in* the acceleration direction. Simplified expressions for the velocity of propagation of energy (ray velocity) are obtained at great distances from the horizon. There exists actually a very helpful analogy between the physical system that we study here and the electromagnetic field within an *inhomogeneous* medium at rest in an *inertial* frame. As far as we know, this analogy has not been noted before, and its importance ought therefore to be stressed.

II. MAXWELL'S EQUATIONS

Our treatment rests upon the following assumptions.

(1) The metric tensor $g_{\mu\nu}$ (μ,ν running from 0 to 3) is independent of time, $\partial g_{\mu\nu}/\partial t = 0$. In particular, the spatial distance between two reference points, as measured by standard measuring rods at rest, is independent of time. In the language of Møller, a reference system of this kind is said to be *rigid* (Ref. 18, p. 287).

(2) The metric is *time orthogonal*, i.e.,

$$g_{0i} = 0, \quad i = 1,2,3, \tag{2.1}$$

implying that the vector potential γ_i vanishes (Ref. 18, p. 280). Thus our metric may be written in the form

$$g_{\mu\nu} = \left(\begin{array}{c|c} g_{00} & 0 \\ \hline 0 & g_{ik} \end{array}\right), \tag{2.2}$$

where g_{ik}, because of vanishing γ_i, determines directly the spatial geometry in the reference system. Note that $g^{00} = 1/g_{00}$. The Minkowski metric is according to our conventions

$$\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1). \tag{2.3}$$

A. Four-dimensional formulation

It is convenient to use the Cartan formalism. We choose an arbitrary coordinate basis, i.e., in the language of Misner *et al.*[19] $\omega^\mu = dx^\mu$. The fundamental Faraday two-form is

$$\mathbf{F} = \tfrac{1}{2}F_{\mu\nu}\, dx^\mu \wedge dx^\nu. \tag{2.4}$$

The first half of Maxwell's set of equations is obtained by requiring the exterior derivative of **F** to be zero,

$$d\mathbf{F} = 0. \tag{2.5}$$

Since

$$d\mathbf{F} = \tfrac{1}{2}F_{\mu\nu,\rho}\, dx^\rho \wedge dx^\mu \wedge dx^\nu \tag{2.6}$$

(comma denoting ordinary partial derivative), we can write (2.5) as

$$F_{[\mu\nu,\rho]} = 0, \tag{2.7}$$

with

$$F_{[\mu\nu,\rho]} = \tfrac{1}{3}(F_{\mu\nu,\rho} + F_{\nu\rho,\mu} + F_{\rho\mu,\nu}) \tag{2.8}$$

denoting the antisymmetrized part of $F_{\mu\nu,\rho}$.

Proceeding to the second half of Maxwell's equations we first introduce a new two-form describing the response of the dielectric medium to the applied electromagnetic field:

$$\mathbf{G} = \tfrac{1}{2}G_{\mu\nu}\, dx^\mu \wedge dx^\nu. \tag{2.9}$$

Its dual $^*\mathbf{G}$, which may be called the Maxwell two-form, is

$$^*\mathbf{G} = \tfrac{1}{2}\, {}^*G_{\mu\nu}\, dx^\mu \wedge dx^\nu = -\tfrac{1}{4}\epsilon_{\mu\nu\alpha\beta}G^{\alpha\beta}\, dx^\mu \wedge dx^\nu. \tag{2.10}$$

Here $\epsilon_{\mu\nu\alpha\beta}$ is the completely antisymmetric pseudotensor of rank 4:

$$\epsilon_{\mu\nu\alpha\beta} = (-g)^{1/2}\delta_{\mu\nu\alpha\beta}, \quad g = \det(g_{\mu\nu}), \tag{2.11}$$

and $\delta_{\mu\nu\alpha\beta}$ is the completely antisymmetric Levi-Civita symbol, with

$$\delta_{0123} = 1. \tag{2.12}$$

The contravariant pseudotensor $\epsilon^{\mu\nu\alpha\beta}$ is

$$\epsilon^{\mu\nu\alpha\beta} = -(-g)^{-1/2}\delta_{\mu\nu\alpha\beta}. \tag{2.13}$$

We are now in a position to write down the second half of Maxwell's equations, assuming for completeness that there may be also free charges and currents in the medium. Introducing the charge-current one-form

$$\mathbf{J} = J_\mu\, dx^\mu, \tag{2.14}$$

whose dual is the three-form

$$^*\mathbf{J} = -(1/3!)J^\mu\epsilon_{\mu\alpha\beta\gamma}\, dx^\alpha \wedge dx^\beta \wedge dx^\gamma, \tag{2.15}$$

we can express the remaining Maxwell's equations compactly as

$$d\,^*\mathbf{G} = {}^*\mathbf{J} \tag{2.16}$$

($c = 1$). In component form, (2.16) means

$$(-g)^{-1/2}[(-g)^{1/2}G^{\mu\nu}]_{,\nu} = J^\mu. \tag{2.17}$$

In this four-dimensional formulation of the theory, all information about the medium's influence upon the fields is hidden in the induction tensor $G_{\mu\nu}$. To make the influence from the medium explicit, we now turn to the three-dimensional formulation.

B. Three-dimensional formulation

We aim at a construction of the theory in (1 + 3) space. This means that, for a given value of the coordinate time t, we have to introduce three-dimensional vectors, tensors, and differential forms.

First, introduce the electric field one-form

$$\mathbf{E} = E_i\, dx^i \tag{2.18}$$

and the magnetic flux density two-form

$$\mathbf{B} = \tfrac{1}{2}B_{ik}\, dx^i \wedge dx^k, \tag{2.19}$$

whereby the four-dimensional two-form **F** introduced in (2.4) may be decomposed as

$$\mathbf{F} = \mathbf{E} \wedge dx^0 + \mathbf{B}. \tag{2.20}$$

In three-space, the form $^*\mathbf{B}$ dual to **B** is a one-form; its components are

$$^*B_i = \tfrac{1}{2}\epsilon_{ikl}B^{kl}, \tag{2.21}$$

where ϵ_{ikl} is the completely antisymmetric pseudotensor of rank 3:

$$\epsilon_{ikl} = \gamma^{1/2}\delta_{ikl}, \quad \gamma = \det(g_{ik}), \tag{2.22}$$

and δ_{ikl} is the completely antisymmetric Levi-Civita symbol, with $\delta_{123} = 1$. Since there are no gravitational potentials, spatial indices are raised and lowered by means of g_{ik} directly.

The contravariant components corresponding to the $^*B_i$ are, when the star in front of the symbol (the Hodge operator) is omitted,

$$B^i = \tfrac{1}{2}\epsilon^{ikl}B_{kl}, \quad \epsilon^{ikl} = \gamma^{-1/2}\delta_{ikl}. \tag{2.23}$$

The B^i are the components of a three-dimensional *axial vector*, $\mathbf{B} = B^i\mathbf{e}_i$, where the basis vectors $\mathbf{e}_i$, in a coordinate basis as assumed here, are $\mathbf{e}_i = \partial/\partial x^i$. The axial vector property of **B** follows from the appearance of the *pseudo*tensor ϵ^{ikl} in (2.23).

We can now express the two-form **F** in terms of three-dimensional quantities

$$F_{0i} = -E_i, \quad F_{ik} = \epsilon_{ikl} B^l. \tag{2.24}$$

Now introduce the three-dimensional field components D^i, H_i in the two-form **G** in a similar way. Expression (2.17) makes it natural to start from the tensor *density* (of weight 1)

$$\mathcal{G}^{\mu\nu} = (-g)^{1/2} G^{\mu\nu}, \tag{2.25}$$

the zeroth row of which is associated with the components D^i of the displacement vector **D**:

$$\mathcal{G}^{0i} = \gamma^{1/2} D^i, \tag{2.26}$$

whereas the spatial components are expressed as

$$\mathcal{G}^{ik} = \gamma^{1/2} H^{ik}, \tag{2.27}$$

so that the H^{ik} constitute a three-dimensional *tensor* (not tensor density). The lowered components H_{ik} constitute a two-form

$$\mathbf{H} = \tfrac{1}{2} H_{ik}\, dx^i \wedge dx^k, \tag{2.28}$$

whose dual is a one-form, ***H**. The components of the latter are, when we again omit the Hodge operator,

$$H_i = \tfrac{1}{2} \epsilon_{ikl} H^{kl}. \tag{2.29}$$

For reference purposes it is useful also to write the following equations:

$$\mathcal{G}^{ik} = \delta_{ikl} H_l, \tag{2.30}$$

$$^*G_{0i} = -H_i, \quad ^*G_{ik} = -\epsilon_{ikl} D^l. \tag{2.31}$$

Care has been exerted here to choose the definitions such that $^*G_{\mu\nu}$ follows from $F_{\mu\nu}$ if one makes the following replacements:

$$E_i \to H_i, \quad B^i \to -D^i. \tag{2.32}$$

For the case of an electromagnetic field in *vacuum* in an inertial system, the duality transformations are usually written as $\mathbf{E} \to \mathbf{H}$, $\mathbf{B} \to -\mathbf{D}$. These transformations are thus strictly analogous to (2.32).

Let us now write down Maxwell's equations in three-dimensional form. Defining curl and divergence operators by

$$\begin{aligned} (\text{curl}\,\mathbf{E})^i &= \epsilon^{ikl} \partial_k E_l, \\ \text{div}\,\mathbf{B} &= \gamma^{-1/2} \partial_i (\gamma^{1/2} B^i), \end{aligned} \tag{2.33}$$

we can write (2.7) as

$$(\text{curl}\,\mathbf{E})^i = -\partial_0 B^i, \quad \text{div}\,\mathbf{B} = 0. \tag{2.34}$$

The remaining Maxwell equations, (2.17), are treated similarly. Decomposing the four-current density J^μ as

$$J^\mu = (-g_{00})^{-1/2} (\rho, \mathbf{j}), \tag{2.35}$$

we get

$$(\text{curl}\,\mathbf{H})^i = j^i + \partial_0 D^i, \quad \text{div}\,\mathbf{D} = \rho. \tag{2.36}$$

It ought to be noted that the present theory assigns "mediumlike" properties to the vacuum. Assuming no material medium to be present $(G_{\mu\nu} = F_{\mu\nu})$ we get

$$D_i = (-g_{00})^{-1/2} E_i, \quad B_i = (-g_{00})^{-1/2} H_i, \tag{2.37}$$

showing that the vacuum behaves as if it were a medium endowed with a permittivity $(-g_{00})^{-1/2}$ and an equally large permeability.

Our theory above is in agreement with the classical works of von Laue,[20] Arzeliès,[21] and others, although there are differences in notation. Compare also the more recent work of Van Bladel.[22] For the case of a field in vacuum, we are in agreement with Møller (Ref. 18, Sec. 10.9) and Landau and Lifshitz.[23] There is essentially agreement also with Post.[24]

The formulation of Schmutzer[25] is, however, different, a characteristic feature of his theory being that the medium-like properties of vacuum, as expressed in (2.37) above, are no longer present. At first sight this may appear to be an attractive feature. However, the price one has to pay for formulating the theory in this way is that the four-dimensional Maxwell equations, (2.7) and (2.17), can no longer be expressed in three-dimensional form such as in (2.34) and (2.36) under maintenance of the curl and divergence operators *as they are naturally defined in (2.33)*. This appears to us to be a drawback of the theory.

Another way of formulating the three-dimensional theory is also worth mentioning: it consists in expressing the four-dimensional quantities $F_{\mu\nu}$, $\mathcal{G}^{\mu\nu}$ in terms of three-dimensional fields *exactly as if Minkowskian coordinates were used*. In the special case of no matter being present, the gravitational field acts itself as some kind of "medium" having in general *nondiagonal* permittivity and permeability. This way of formulation has been advocated by Skrotskii[26] and Plebanski[27] (assuming no medium), and by Volkov and Kiselev,[28] who introduced a medium. Further references along these lines are Refs. 29–31.

In conclusion, the various ways of formulating the three-dimensional theory appear to be internally consistent, although in our opinion the most natural alternative is the von Laue–Arzeliès version presented above.

III. CONSTITUTIVE RELATIONS

We shall make use of the same method introduced in an earlier work[32] dealing with inertial frames. The method makes use of projection operators, which are able to map the electromagnetic theory in a *medium* onto the electromagnetic theory in a *vacuum*. In the present problem, it is sufficient to introduce the following operator:

$$O^{\mu\nu} = g^{\mu\nu} - \kappa u^\mu u^\nu, \quad \kappa = n^2 - 1. \tag{3.1}$$

Here u^μ is the four-velocity of the medium, and $n = (\epsilon\mu)^{1/2}$ is the refractive index.

Using (3.1) we can write the constitutive relations compactly as

$$G^{\mu\nu} = \mu^{-1} O^{\mu\alpha} O^{\nu\beta} F_{\alpha\beta}. \tag{3.2}$$

Alternatively, we can write them out as

$$G^{\mu\nu} = \mu^{-1} [F^{\mu\nu} - \kappa (F^\mu u^\nu - F^\nu u^\mu)], \tag{3.3}$$

with $F^\mu = F^{\mu\nu} u_\nu$. The derivatives of u^μ are thus not permitted to occur; cf. also the discussion by Anderson and Ryon[33] on this point.

In the rest system of matter, where

$$u^i = 0, \quad u^0 = (-g_{00})^{-1/2}, \quad u_0 = -(-g_{00})^{1/2}, \tag{3.4}$$

Eqs. (3.3) imply, in view of (2.24)–(2.30),

$$D_i = \epsilon(-g_{00})^{-1/2}E_i, \quad B_i = \mu(-g_{00})^{-1/2}H_i, \tag{3.5}$$

which are in agreement with (2.37) if there is no material medium present.

It ought to be stressed that the particular way of formulating the constitutive relations shown in (3.2) or (3.3) is quite useful, since it solves for the induction tensor $G^{\mu\nu}$ explicitly. Often one will find in the literature these relations given in a way that determines $G^{\mu\nu}$ only implicitly. According to our knowledge the formulation (3.3) was first given by Jauch and Watson.[34] Its significance for a Lagrangian theory of the electromagnetic field in media was immediately recognized by Novobátzky,[35] and it became in the following years extensively employed by the Hungarian group.[36–39] Other authors also made use of it, in various contexts, including the generalization to the case of anisotropic media.[25,32,33,40,41]

IV. ELECTROMAGNETIC POTENTIALS

We can introduce the constitutive relations (3.2) into the second half of Maxwell's equations (2.17) and thereafter introduce the electromagnetic four-potential A_μ by the equation

$$F_{\mu\nu} = A_{\nu;\mu} - A_{\mu;\nu} \tag{4.1}$$

(semicolon denoting covariant derivative), to obtain equations for A_μ.

What form does the *subsidiary condition* take under the present circumstances? The most natural expression for it, in Lorentz gauge, seems to be

$$O^{\mu\nu}A_{\mu;\nu} = 0, \tag{4.2}$$

since this is a covariant equation reducing to the conventional Lorentz condition in an inertial frame. We accordingly adopt (4.2); this is in accordance also with Mo.[41]

Let us hereafter make the simplifying assumption that the curvature tensor is equal to zero. Covariant derivatives accordingly commute. Some calculation then yields the following wave equations for the potential components:

$$(O^{\alpha\beta}A_{\mu;\beta})_{;\alpha} = -\mu(O^{-1})_{\mu\nu}J^\nu + O^{\alpha\beta}_{\ ;\beta}A_{\alpha;\mu} - O^{\alpha\beta}_{\ ;\mu}A_{\beta;\alpha} + (O^{-1})_{\mu\alpha}O^{\alpha\beta}_{\ ;\gamma}O^{\gamma\nu}F_{\beta\nu}, \tag{4.3}$$

where it is understood that $F_{\beta\nu}$ means $(A_{\nu;\beta} - A_{\beta;\nu})$. These complicated equations are, in general, coupled. Equations (4.3) together with (4.2) are equivalent to Maxwell's equations.

In the special case of vanishing covariant derivatives, $u^\mu_{;\nu} = 0$, (4.3) become simplified considerably since they decouple. We then have

$$(O^{\alpha\beta}A_{\mu;\beta})_{;\alpha} = -\mu(O^{-1})_{\mu\nu}J^\nu. \tag{4.4}$$

The potential equations derived by Mo [Ref. 41, Eq. (4.10a)] presuppose that $u^\mu_{;\nu} = 0$. This is actually a delicate point: in a reference frame where the matter is at rest one might perhaps expect that $u^\mu_{;\nu}$ would be zero. In general, this does *not* hold true, not even in the simple Rindler space, as we shall show below. Therefore, if one wishes to describe the electromagnetic field in Rindler space by the potential wave equations, one has to return to the full set of equations (4.3).

V. GENERAL PROPERTIES OF THE RINDLER SPACE

From now on we shall specialize to the case of uniform linear acceleration. Assume that the frame K (the Rindler frame) is uniformly accelerated along the x axis with respect to the inertial background space, called I. We first observe that, as regards the relative motion between observer and medium, there is actually a variety of possibilities to choose from here: case I, the observer can be inertial, and the medium accelerated; case II, the observer can be accelerated, and the medium inertial; and case III, the observer and the medium can be coaccelerated.

In this paper we will consider only case III. This implies the simplifying feature that all spatial velocity components of the medium vanish in the comoving Rindler frame. On the other hand, it is a complicating factor in the analysis that the medium, even if assumed to fill the entire Rindler wedge, fills only a *part* of the global Minkowski space. There exists accordingly no simple homogeneous medium filling the inertial background space I.

Before embarking upon the specific electromagnetic theory, it is convenient to summarize some basic properties of the Rindler space. In the frame I we represent the Minkowski coordinates by capital letters, $X^\mu = (cT,X,Y,Z)$. (For physical reasons it is desirable in this section to keep c as a dimensional quantity.) The Rindler coordinates in the frame K will be donoted by $x^\mu = (ct,x,y,z)$, where t is the "global" coordinate time within the Rindler wedge. The relations between X^μ and x^μ are

$$cT = \sigma x \sinh(at/c), \quad X = \sigma x \cosh(at/c), \quad Y = y, \quad Z = z, \tag{5.1}$$

implying $X^2 - c^2T^2 = x^2$. We have introduced a constant a having the dimension of an acceleration, and also a parameter σ to distinguish between the two wedges: $\sigma = +1$ refers to the right wedge (R) and $\sigma = -1$ refers to the left (L). The situation is shown in Fig. 1. Since $cT/X = \tanh(at/c)$, a straight line through the origin of slope less than unity in magnitude singles out a spacelike surface corresponding to constant coordinate time $t \in \langle -\infty, \infty \rangle$. With the conventions adopted in (5.1), this holds true for both regions R and L. Our conventions are in accordance with those of Birrell and Davies,[14] Sciama *et al.*,[8] and Takagi.[13] As seen from Fig. 1 there are two other regions of Minkowski space also, viz., the future region F and the past region P. These regions are not described by the formulas (5.1). In the following we will consider the right region R only, thus $\sigma = +1$. Its past horizon is the line $cT = -X, X > 0$, corresponding to $t = -\infty$, whereas the future region is $cT = X, X > 0$, corresponding to $t = \infty$. If, as will be assumed here, the material medium occupies the whole Rindler wedge, then the left end of it is travelling with the velocity of light: it moves towards the origin for $T < 0$ and outwards from the origin for $T > 0$. At $T = 0$ the left end experiences an infinite acceleration (in

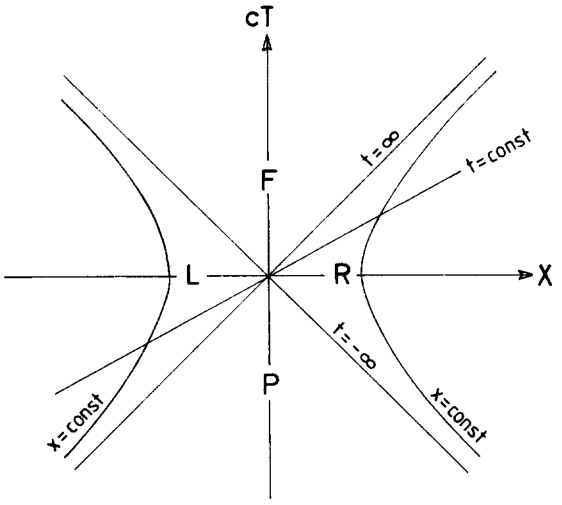


FIG. 1. Rindler coordinates in Minkowski space.

pictorial terms, there is an exchange of "light platforms").

An arbitrary fixed point $x = \text{const}$ in the medium travels with velocity

$$\frac{dX}{dT} = c \tanh\left(\frac{at}{c}\right) = cT\left(T^2 + \frac{x^2}{c^2}\right)^{-1/2}. \quad (5.2)$$

As seen from the Minkowski frame the medium is thus not rigid: for a given T the magnitude of the velocity of a material point is greater the closer the point is located to the origin. From (5.2) we find that the nonvanishing component of the four-velocity U^μ of the medium in the frame I are

$$U^0 = c \cosh(at/c), \quad U^1 = c \sinh(at/c). \quad (5.3)$$

At the instant $T = t = 0$, all points in the medium are at rest in I, as well as in K.

From (5.1) we construct the line element

$$ds^2 = -(a^2x^2/c^2)dt^2 + dx^2 + dy^2 + dz^2, \quad (5.4)$$

from which it follows that

$$g_{00} = -a^2x^2/c^4, \quad g_{0i} = 0, \quad g_{ik} = \delta_{ik}. \quad (5.5)$$

The spatial metric is thus Cartesian. All deviations from Minkowski space are described by the metric component g_{00}. From (5.4) it follows that the proper time τ for a material point at rest in K is

$$\tau = (ax/c^2)t. \quad (5.6)$$

The four-acceleration of such a point, as viewed from I, is

$$A^\mu = \frac{dU^\mu}{d\tau} = \frac{c^2}{x}\left(\sinh\frac{at}{c}, \cosh\frac{at}{c}, 0, 0\right), \quad (5.7)$$

leading to the invariant, proper acceleration

$$A = (A^\mu A_\mu)^{1/2} = c^2/x. \quad (5.8)$$

Note that this is independent of which value is assigned to the acceleration parameter a.

As mentioned, we will assume that the medium occupies the whole Rindler wedge, i.e., extends down to $x = 0$. Conceptually, there is nothing against making such an assumption. For a physical material, however, strong stresses will be produced near the horizon, and the material breaks. It is of interest to estimate the position of the breaking point. Let us consider maraging steel, one of the strongest construction materials, as an example. Its yield strength is[42] 2 GPa $= 2\times 10^{10}$ dyn/cm^2. If a stress of this magnitude acts across one end (the base area) of a coin-shaped sample, of thickness 1 mm we obtain, since the density is about 8 g/cm^3, the acceleration of the sample to be about 3×10^{10} cm/sec^2. Assume next that the left end of the medium is located at $x = x_0$ (not to be confused with the covariant time coordinate), and require the acceleration of this end, which according to (5.8) is c^2/x_0, to be equal to 3×10^{10} cm/sec^2. This yields $x_0 \cong 3\times 10^{10}$ cm. If the medium approaches essentially closer to the origin than this, it will be torn apart. Thus the strength of real macroscopic materials puts severe limits on how close we can approach the horizon. If we instead consider *microscopic* systems, the conditions can be more extreme. For instance, in the SLAC linac[10a] one obtains an electric field of 7 MV/m; for an electron this yields an acceleration of about 10^{20} cm/sec^2. If some kind of "medium" were able to withstand such an acceleration, then the left end would come as close as $x_0 = 9$ cm to the horizon, before breaking.

We put henceforth $a = 1$, $c = 1$. The following relations, in the frame K, are useful:

$$\Gamma^1_{00} = x, \quad \Gamma^0_{01} = \Gamma^0_{10} = 1/x,$$
$$u^0 = 1/x, \quad u^i = 0, \quad u^1_{;0} = 1. \quad (5.9)$$

All other Christoffel components, and covariant derivatives of the velocity, are equal to zero in this rest frame of the medium. The nonvanishing $u^1_{;0}$, in particular, implies that the simplified version (4.4) of the potential wave equation cannot be used; we have to resort to the full wave equation (4.3).

VI. THE FUNDAMENTAL ELECTROMAGNETIC MODES

A. Maxwell's equations

We consider in this section a pure radiation field in K. Our purpose is to derive the expressions for the fundamental electromagnetic modes. Rather than dealing with potentials, we choose to start directly from the Maxwell equations. We let the modes vary with time as $e^{-i\omega t}$. Since the spatial geometry in K is Cartesian, we do not have to distinguish between covariant and contravariant components of the vector fields. The Maxwell equations become

$$\text{curl}\,\mathbf{E} = i\omega\mathbf{B}, \quad \text{div}\,\mathbf{B} = 0, \quad (6.1a)$$

$$\text{curl}\,\mathbf{H} = -i\omega\mathbf{D}, \quad \text{div}\,\mathbf{D} = 0, \quad (6.1b)$$

where the operators curl and div have their usual Cartesian meaning. The constitutive relations are, in accordance with (3.5),

$$\mathbf{D} = (\epsilon/x)\mathbf{E}, \quad \mathbf{B} = (\mu/x)\mathbf{H}. \quad (6.2)$$

These equations in fact are exactly of the same form as for an electromagnetic field in an *inhomogeneous medium at rest in an inertial frame*, with the following effective permittivity and permeability:

$$\epsilon_{\text{eff}} = \epsilon/x, \quad \mu_{\text{eff}} = \mu/x. \quad (6.3)$$

This analogy is most helpful, when interpreting the theory in K.

Eliminating **D** and **B** from (6.1) and (6.2) we obtain the wave equations for **E** and **H**. The wave equation for **E** can be written as

$$\mathrm{curl}(x\,\mathrm{curl}\,\mathbf{E}) = (\epsilon\mu\omega^2/x)\mathbf{E}, \qquad (6.4a)$$

which is equivalent to

$$\nabla^2\mathbf{E} - \mathrm{grad}\,\mathrm{div}\,\mathbf{E} - (1/x)\mathbf{e}_x \times \mathrm{curl}\,\mathbf{E} + (\epsilon\mu\omega^2/x^2)\mathbf{E} = 0, \qquad (6.4b)$$

where $\mathbf{e}_x$ is the unit vector in the x direction. The equation for **H** is identical.

We consider a wave propagating in the xz plane. All quantities are independent of y, and the uniformity of the distribution of fields in the z direction means that the fields vary with z through a factor e^{ikz}, where k is a constant.[43] If $k = 0$, the fields depend only on x, and (borrowing terminology from the theory of ordinary inhomogeneous media[43]) we call the wave a *normal* wave. If $k \neq 0$, we call it an *oblique* wave. In the latter case two independent cases of polarization must be distinguished. We shall now introduce the concepts of TE and TM modes.

B. TE mode

Assume that **E** is directed along the y axis, i.e., transverse to the plane of polarization. Writing for simplicity E instead of E_y, we obtain from (6.4b)

$$\frac{d^2E}{dx^2} + \frac{1}{x}\frac{dE}{dx} + \left(\frac{\epsilon\mu\omega^2}{x^2} - k^2\right)E = 0, \qquad (6.5)$$

which is the governing equation for the TE mode. In the derivation of (6.5), we first took into account the z dependence mentioned above, and thereafter omitted irrelevant factors, so that E in (6.5) depends on x only.

Independent solutions of (6.5) are[44] (when $k \neq 0$) the modified Bessel functions of imaginary order, $I_{\pm i\alpha}(kx)$ and $K_{i\alpha}(kx)$, where we have introduced

$$\alpha = (\epsilon\mu)^{1/2}\omega = n\omega. \qquad (6.6)$$

We take ω to be positive, so that α becomes positive as well. Since $I_{\pm i\alpha}\ (kx \to \infty) \to \infty$, we reject these functions as solutions. Introducing the symbol

$$u = kx, \qquad (6.7)$$

we can then write the TE mode as follows, including again the z and t factors,

$$E = E_0 K_{i\alpha}(u) e^{i(-\omega t + kz)}, \qquad (6.8)$$

where E_0 is a constant. Recall that this expression holds in the region

$$x \in [0,\infty\rangle, \quad y,z \in \langle -\infty,\infty\rangle, \qquad (6.9)$$

assumed that $k \neq 0$, $\omega > 0$.

It is useful to observe that $K_{i\alpha}(u)$ is a real quantity. This follows from formula 9.6.24 in Ref. 44:

$$K_\nu(u) = \int_0^\infty e^{-u\cosh t}\cosh(\nu t)dt. \qquad (6.10)$$

This expression shows that K_ν is real when ν is real or purely imaginary, thus including $\nu = i\alpha$ with α real.

The magnetic field is most easily determined by calculating the components of **B** from the first equation in (6.1a):

$$B_x = -(k/\omega)E, \quad B_y = 0,$$
$$B_z = (kE_0/i\omega)K'_{i\alpha}(u)e^{i(-\omega t + kz)}, \qquad (6.11)$$

where prime means differentiation with respect to the argument. The remaining fields **D** and **H** now follow from (6.2).

There are two limiting cases of interest here. First, near the horizon, $x \to 0$, all (the nonvanishing) components of **E** and **B** diverge; the fields "pile up" near the horizon. In this region also the effective permittivity and permeability go to infinity.

Next, the case of large u is of interest, since the fields can then be expressed in terms of elementary functions. Namely, in this limit we have, when we include terms of order $1/u$,

$$K_{i\alpha}(u) = (\pi/2u)^{1/2}e^{-u}[1 - (4\alpha^2 + 1)/8u], \qquad (6.12a)$$

$$K'_{i\alpha}(u) = -(\pi/2u)^{1/2}e^{-u}[1 - (4\alpha^2 - 3)/8u], \qquad (6.12b)$$

see Ref. 44, formulas 9.7.2 and 9.7.4. Therefore, if we omit the $O(1/u)$ terms we find that $B_z/B_x \to i$ when $u \to \infty$, which implies that B_z lies $\frac{1}{2}\pi$ ahead of B_x in phase.

We have based this analysis upon Maxwell's equations. Alternatively we might have started from the potential wave equations (4.3). The following relations are useful here:

$$O^{\mu\nu}_{\ ;\alpha} = -(\kappa/x)(g^{\mu}_{1}g^{\nu}_{0} + g^{\nu}_{1}g^{\mu}_{0})g^{0}_{\alpha}. \qquad (6.13)$$

Some calculation shows that the wave equation for the component $A \equiv A_y$ is identical to (6.5) above. It is consistent to put all other potential components equal to zero, and so from $E = -\partial_0 A$ we have

$$A = E/i\omega, \qquad (6.14)$$

where E is given by (6.8). The subsidiary condition (4.2), and all remaining potential equations, are satisfied automatically.

C. TM mode

Assume now that **H** is transverse and directed along the y axis. The equation for $H \equiv H_y$ is the same as (6.5), and so we obtain immediately for the TM mode

$$H = H_0 K_{i\alpha}(u)e^{i(-\omega t + kz)}, \qquad (6.15)$$

in analogy to (6.8), H_0 being a constant.

The natural pair of field variables is now **H** and **D**. From the first equation in (6.1b) we determine the components of **D**,

$$D_x = (k/\omega)H, \quad D_y = 0,$$
$$D_z = (ikH_0/\omega)K'_{i\alpha}(u)e^{i(-\omega t + kz)}. \qquad (6.16)$$

The remaining fields follow from (6.2).

Near the horizon, all field components in (6.15) and (6.16) diverge. Again, for large u the approximation (6.12) is useful.

It is possible to introduce the electromagnetic potential in full analogy to (6.14): the ansatz

$$A_\mu = (A_0, \mathbf{A}) = (0, A_x, 0, A_z), \qquad (6.17a)$$

with

$$A_x = E_x/i\omega, \quad A_z = E_z/i\omega, \qquad (6.17b)$$

leads to correct expressions for the fields **E** and **B**. The ansatz is thereby justified.

The TM case is more complicated than the TE case as far as the electromagnetic potential equations are concerned. The reason is that the expressions for A_μ, as given by (6.17), do not satisfy the subsidiary condition and the potential wave equations *separately*, with the particular choice (4.2) for the subsidiary condition. The general lesson to be drawn from this is that when dealing with the electromagnetic field in curvilinear coordinates, the best method is to start from the Maxwell equations directly. There is no natural, simplifying choice for the subsidiary condition, and the use of electromagnetic potentials may in fact easily make the mathematical analysis more complicated.

D. TEM mode

An interesting special case occurs when the wave is a normal one, propagating along the x axis. Then all fields are transverse, i.e, they are lying in the yz plane. The constant k is in this case equal to zero, so that the variable $u = kx$ can no longer be used. Instead, we return to the basic equation (6.5), which becomes

$$\frac{d^2E}{dx^2} + \frac{1}{x}\frac{dE}{dx} + \frac{\alpha^2}{x^2}E = 0. \qquad (6.18)$$

The independent solutions are simply $x^{\pm i\alpha}$. It is convenient to express the solutions in terms of a new variable ξ, defined by $x = e^\xi$, $\xi \in \langle -\infty, \infty\rangle$. Choosing the right-moving solution, we can write the fundamental TEM mode as

$$\mathbf{E} = \mathbf{E}_0 e^{i(-\omega t + \alpha\xi)}, \quad \mathbf{H} = (\epsilon/\mu)^{1/2}\mathbf{e}_x \times \mathbf{E}. \qquad (6.19)$$

It means in (t,ξ) coordinates a *plane* wave; the amplitudes of **E** and **H** do not change with position at all. In this respect the TEM mode contrasts the TE and TM modes.

VII. THE ELECTROMAGNETIC ENERGY-MOMENTUM TENSOR

The situation under study in this paper is very suitable for making an explicit test of the electromagnetic energy-momentum conservation equations in curvilinear coordinates. In general relativity, it is known that the construction of the energy-momentum complex for a gravitational system is by no means trivial; the problem being essentially the transfer of the covariant divergence of the energy-momentum tensor into an ordinary divergence of a (not unique) energy-momentum complex. These matters have in particular been discussed by Møller;[18] cf. also the recent paper by Kovacs.[45]

Since the electromagnetic field in matter is a nonclosed system, we cannot expect beforehand that the electromagnetic four-force density f_μ is zero, not even for a pure radiation field. Let us for completeness assume that there are extraneous charges and currents present, described by a four-current density J^μ. If $S_\mu{}^\nu$ is the electromagnetic energy-momentum tensor, the differential conservation equations are

$$-S_{\mu;\nu}^{\ \nu} = f_\mu. \qquad (7.1)$$

It is known that electromagnetic phenomena in a medium in inertial space, at least for high frequencies, are best described in terms of Minkowski's energy-momentum tensor (there is no relationship to Minkowski *coordinates* here!). See, for instance, Refs. 18, 40, and 46. In fact, we are not aware of any experiment in optics that cannot be explained by means of this tensor in a straightforward way. It is natural, therefore, to base the analysis upon the Minkowski tensor also in the case of curvilinear coordinates. Consequently, $S_\mu{}^\nu$ is given by

$$S_\mu{}^\nu = F_{\mu\alpha}G^{\nu\alpha} - \tfrac{1}{4}g_\mu^\nu F_{\alpha\beta}G^{\alpha\beta}, \qquad (7.2)$$

which corresponds to

$$f_\mu = F_{\mu\nu}J^\nu + \tfrac{1}{4}(F_{\alpha\beta}G^{\alpha\beta}_{\ \ ;\mu} - F_{\alpha\beta;\mu}G^{\alpha\beta}). \qquad (7.3)$$

Here the first term gives the force on the extraneous charges and currents, whereas the second term gives the force on the *medium*.

Now specialize to Rindler space. Introduce the electromagnetic energy density h, the Poynting vector **S**, the momentum density **g**, and the Maxwell stress tensor $t_i{}^k$. Explicitly

$$\begin{aligned} &S_0{}^0 = -h/x, \quad S_0{}^i = -S^i/x, \\ &S_i{}^0 = g_i/x, \quad S_i{}^k = -t_i{}^k/x, \end{aligned} \qquad (7.4)$$

and so (7.2) leads to

$$\begin{aligned} &h = \tfrac{1}{2}(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B}), \quad S^i = (\mathbf{E}\times\mathbf{H})^i, \\ &g_i = (\mathbf{D}\times\mathbf{B})_i, \\ &t_i{}^k = E_iD^k + H_iB^k - \tfrac{1}{2}\delta_i^k(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B}). \end{aligned} \qquad (7.5)$$

These expressions, in fact, are formally identical to those one obtains for an electromagnetic field in an ordinary dielectric medium.

The four-force density f_μ in terms of the three-fields can be found from (7.3). We shall not write down the expressions for its components here. It is, however, of great interest to recognize that the differential conservation equations can be written in the form

$$\begin{aligned} &\partial_0 h + \operatorname{div}\mathbf{S} = -\mathbf{E}\cdot\mathbf{j}, \\ &-\partial_0 g_i + t_{i,k}^{\ k} = \rho E_i + (\mathbf{j}\times\mathbf{B})_i \\ &\qquad - \tfrac{1}{2}E^2\,\partial_i\epsilon_{\text{eff}} - \tfrac{1}{2}H^2\,\partial_i\mu_{\text{eff}}, \end{aligned} \qquad (7.6)$$

cf. (6.3). These are *formally identical to those holding for an inhomogeneous medium at rest with permittivity ϵ_{eff} and permeability μ_{eff}* in an inertial system, when the Minkowski tensor is being used. In fact, this is exactly what we would expect, on basis of the correspondence (7.5) with the Minkowski theory. This analogy ought in our opinion to be of value also in related problems, such as in the examination of the subtle emission and absorption phenomena that take place when a detector is at rest in an accelerated system.

Before closing this section let us define also the propagation velocity of electromagnetic energy:

$$\mathbf{u}^* = \mathbf{S}/h. \qquad (7.7)$$

Using (7.5) we can calculate $\mathbf{u}^*$ explicitly for the various modes. For the TE mode we obtain

$$\mathbf{u}^* = (0,0,u^{*z}), \qquad (7.8a)$$

$$u^{*z} = \frac{2\omega}{k}\left\{1 + \left[\frac{K'_{i\alpha}(u)}{K_{i\alpha}(u)}\right]^2 + \frac{\alpha^2}{u^2}\right\}. \qquad (7.8b)$$

The expression for the TM mode is identical. For these modes there is thus an electromagnetic energy transport *parallel* to the horizon. Making use of the large u expansion (6.12) the expression (7.8b) can be simplified. If $u \to \infty$, the velocity of the energy becomes simply ω/k. This is the same expression as for a plane wave of wavenumber k propagating in the z direction.

Considering finally the TEM mode, we find that there is an energy transport *in the x direction*. We obtain

$$\mathbf{u}^* = (1/n_{\text{eff}})\mathbf{e}_x, \tag{7.9a}$$

with

$$n_{\text{eff}} = n/x = (\epsilon\mu)^{1/2}/x, \tag{7.9b}$$

in analogy to (6.3). The physical system behaves as an inhomogeneous medium of refractive index n_{eff}.

VIII. CONCLUSIONS AND FINAL REMARKS

(1) It ought to be stressed that we have based the calculation upon a model where the medium is isotropic, homogeneous, and nondispersive, filling the Rindler wedge completely. Recall also that only case III from the listing in Sec. V has been discussed. With these underlying assumptions, the formalism, as we have seen, becomes quite tractable. Perhaps the most noteworthy result is the circumstance that there exists a direct analogy between the electromagnetic theory in the frame K and the theory in an ordinary inhomogeneous dielectric medium at rest.

(2) We have discussed the classical theory only. As far as we know, no quantum treatment of the phenomenological electromagnetic field has so far been given. It is worthwhile to point out in this context, without going into great detail, that it is quite straightforward to construct the quantum field theory for the following related simplified situation: The electromagnetic field is replaced by a fictitious *scalar* field, which is assumed to respond to the presence of the medium in terms of a refractive index, in essentially the same way as an electromagnetic field does. We find it very reasonable to expect that a scalar theory of this kind is able to reflect many of the essential features of a complete electromagnetic theory.

Let us therefore outline the main properties of this scalar theory. The basic field equation for the scalar field Φ must be the following expression:

$$\nabla^2\Phi + (1/x)\partial_x\Phi + (\alpha^2/x^2)\Phi = 0, \tag{8.1}$$

with $\alpha^2 = \epsilon\mu\omega^2$ as before, and $\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$. [The reason for adopting (8.1) as governing equation becomes clear from analogy, if we recall (6.4b) or (6.5).] Assuming uniform distribution of the scalar field in the transverse y and z directions, we see that the fundamental modes will have to vary in these directions through the factor $e^{i\mathbf{k}\cdot\mathbf{y}}$, the transverse vectors $\mathbf{y}$ and $\mathbf{k}$ being defined as

$$\mathbf{y} = (y,z), \quad \mathbf{k} = (k_2,k_3), \quad k = (k_2^2 + k_3^2)^{1/2}, \tag{8.2}$$

with $y,z \in \langle -\infty,\infty\rangle$, $k_2,k_3 \in \langle -\infty,\infty\rangle$, $k \in [0,\infty\rangle$. Requiring the boundary condition at infinity to be $\Phi \to 0$, and introducing again the variable $u = kx$ in analogy to (6.7), we can write the expression for the fundamental mode $\phi_{\omega\mathbf{k}}$ of (8.1) as

$$\phi_{\omega\mathbf{k}} = (1/2\pi^2)(\sinh \pi\alpha)^{1/2} K_{i\alpha}(u) e^{i(-\omega t + \mathbf{k}\cdot\mathbf{y})}. \tag{8.3}$$

The modes are orthonormalized,

$$(\phi_{\omega\mathbf{k}},\phi_{\omega'\mathbf{k}'}) = \delta(\omega - \omega')\delta(\mathbf{k} - \mathbf{k}'), \tag{8.4}$$

in accordance with the following general expression for the inner product:

$$(\phi_1,\phi_2) = -i\int_\Sigma O^{\mu\nu}[\phi_1^* \overleftrightarrow{\nabla}_\nu \phi_2] d^3\Sigma_\mu, \tag{8.5}$$

in standard notation. An expansion of Φ into fundamental modes,

$$\Phi = \int_0^\infty d\omega \int_{-\infty}^\infty d^2k\, [a_{\omega\mathbf{k}}\phi_{\omega\mathbf{k}} + a_{\omega\mathbf{k}}^\dagger \phi_{\omega\mathbf{k}}^*], \tag{8.6}$$

leads to the following nonvanishing commutation rules:

$$[a_{\omega\mathbf{k}},a_{\omega'\mathbf{k}'}^\dagger] = \delta(\omega - \omega')\delta(\mathbf{k} - \mathbf{k}'). \tag{8.7}$$

[1]For a general introduction to Rindler space, see W. Rindler, Am. J. Phys. **34**, 1174 (1966); *Essential Relativity* (Springer, New York, 1977), 2nd ed. See also, J. D. Hamilton, Am. J. Phys. **46**, 83 (1978).
[2]See, for instance, I. Brevik, Phys. Rep. **52**, 133 (1979).
[3]S. A. Fulling, Phys. Rev. D **7**, 2850 (1973).
[4]B. S. DeWitt, Phys. Rep. **19**, 295 (1975).
[5]P. C. W. Davies, J. Phys. A **8**, 609 (1975).
[6]W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
[7](a) P. Candelas and D. G. Raine, J. Math. Phys. **17**, 2101 (1976); (b) P. Candelas and D. Deutsch, Proc. R. Soc. London Ser. A **354**, 79 (1977).
[8]D. W. Sciama, P. Candelas, and D. Deutsch, Adv. Phys. **30**, 327 (1981).
[9]H. K. Flø, Cand. Real Thesis, University of Oslo, 1981.
[10]J. S. Bell and J. Leinaas, (a) Nucl. Phys. B **212**, 131 (1983); (b) J. S. Bell, R. J. Hughes, and J. M. Leinaas, Z. Phys. C **28**, 75 (1985).
[11]W. G. Unruh and R. W. Wald, Phys. Rev. D **25**, 942 (1982); **29**, 1047 (1984).
[12]K. Hinton, P. C. W. Davies, and J. Pfautsch, Phys. Lett. B **120**, 88 (1983).
[13]S. Takagi, Prog. Theor. Phys. **72**, 505 (1984); Phys. Lett. B **148**, 116 (1984); Prog. Theor. Phys. **74**, 1219 (1985).
[14]N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge U. P., Cambridge, 1982).
[15]R. J. Hughes, Ann. Phys. (NY) **162**, 1 (1985).
[16]N. Sanchez, Phys. Lett. A **112**, 133 (1985).
[17]T. D. Lee, *Particle Physics and Introduction to Field Theory* (Harwood, New York, 1981).
[18]C. Møller, *The Theory of Relativity* (Oxford U. P., London, 1972), 2nd ed.
[19]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[20]M. von Laue, *Die Relativitätstheorie* (Vieweg, Braunschweig, 1923), 2. Band, § 15.
[21]H. Arzeliès, *Milieux Conducteurs ou Polarisables en Mouvement* (Gauthier–Villars, Paris, 1959), Chap. VI.
[22]J. Van Bladel, *Relativity and Engineering* (Springer, Berlin, 1984), Chap. 9.
[23]L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1975), § 90.
[24]E. J. Post, *Formal Structure of Electromagnetism* (North-Holland, Amsterdam, 1962).
[25]E. Schmutzer, *Relativistische Physik* (Teubner, Leipzig, 1968).
[26]G. V. Skrotskii, Dok. Akad. Nauk SSSR **114**, 73 (1957) [Sov. Phys. Dokl. **2**, 226 (1957)].
[27]J. Plebanski, Phys. Rev. **118**, 1396 (1960).
[28]A. M. Volkov and V. A. Kiselev, Zh. Eksp. Teor. Fiz. **57**, 1353 (1969) [Sov. Phys. JETP **30**, 733 (1970)].
[29]A. M. Volkov, A. A. Izmest'ev, and G. V. Skrotskii, Zh. Eksp. Teor. Fiz. **59**, 1254 (1970) [Sov. Phys. JETP **32**, 686 (1971)].
[30]B. Mashhoon, Phys. Rev. D **8**, 4297 (1973).
[31]K. Tanaka, J. Appl. Phys. **49**, 4311 (1978); Phys. Rev. A **25**, 385 (1982).
[32]I. Brevik and B. Lautrup, Mat. Fys. Medd. Dan. Vid. Selsk. **38**, 1 (1970).
[33]J. L. Anderson and J. W. Ryon, Phys. Rev. **181**, 1765 (1969).
[34]J. M. Jauch and K. M. Watson, Phys. Rev. **74**, 950, 1485 (1948); **75**, 1249 (1949).
[35]K. F. Novobátzky, Acta Phys. Hung. **1**, 25 (1949).

[36]G. Marx, Acta Phys. Hung. **2**, 67 (1952); **3**, 75 (1953); Bull. Acad. Pol. Sci. **4**, 29 (1956).
[37]G. Marx and K. Nagy, Acta Phys. Hung. **4**, 297 (1955).
[38]G. Marx and G. Györgyi, Ann. Phys. (Leipzig) **16**, 241 (1955).
[39]K. Nagy, Acta Phys. Hung. **5**, 95 (1955).
[40]I. Brevik, Mat. Fys. Medd. Dan. Vid. Selsk. **37**, 13 (1970).
[41]T. C. Mo, J. Math. Phys. **11**, 2589 (1970).
[42]F. Herlach, Europhys. News **17**, 29 (1986).
[43]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984), 2nd ed. Section 88 discusses waves in inhomogeneous media.
[44]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Natl. Bur. Stand., Washington, D.C., 1964) (reprinted by Dover, New York).
[45]D. Kovacs, Gen. Relativ. Gravit. **17**, 927 (1985).
[46]I. Brevik, Mat. Fys. Medd. Dan. Vid. Selsk. **37**, 11 (1970); Phys. Rev. B **33**, 1058 (1986).

Relaxation of the Dicke maser model

D. Goderis,[a] A. Verbeure, and P. Vets[a]
Instituut voor Theoretische Fysica, Universiteit Leuven, B-3030 Leuven, Belgium

(Received 4 February 1987; accepted for publication 20 May 1987)

For the Dicke maser model the approach to equilibrium for a quantum version of a detailed balance evolution is studied rigorously. The exponential decay with a relaxation time of the order $|T-T_c|^{-1}$ for $T\leqslant T_c$ is proved. A complete spectral resolution of the evolution is obtained.

I. INTRODUCTION

In 1963, Glauber proposed a stochastic Ising model in which the spins change their state randomly with time according to a continuous Markov process.[1] The model turned out to be successful in various ways. In particular, it turned out to be an efficient tool to investigate some dynamical aspects of the Ising model near the critical point.[2,3]

Recently a generalization to quantum systems of Glauber-type dynamics was proposed.[4,5] In a rigorous mathematical scheme a quantum mechanical Markov process satisfying the detailed balance condition is defined. These detailed balance evolutions have been used already to investigate the critical behavior of the free boson gas.[6] A result about the critical slowing down was obtained, which coincides with the result obtained from a nonlinear treatment, based on the weak coupling limit.[7]

In this paper we investigate the critical behavior of the Dicke maser model.[8–10] As in Ref. 10 we consider the model described by the Hamiltonian

$$H_N=\sum_{k=-N}^{N}(a_k^*a_k+\epsilon\sigma_k^+\sigma_k^-)+\frac{\lambda}{2N+1}\sum_{k,l=-N}^{N}(a_k^*\sigma_l^-+a_k\sigma_l^+),\qquad(1)$$

where the $a_k^{(*)}$ are the boson creation and annihilation operators and the $\sigma_k^\pm$ the Pauli matrices.

We also make the restriction $0<\epsilon\leqslant\lambda^2$, yielding the existence of a phase transition. Our main result is an exact calculation of the critical exponent governing the approach to equilibrium in this genuinely quantum mechanical model. It is shown that the relaxation time diverges as $|T-T_c|^{-1}$ below the critical temperature. We do not obtain a polynomial behavior like for the free Bose gas.[6] Above the critical temperature we have always an energy gap. Our work must be compared with the work of Hepp and Lieb.[11] In both cases one studies an irreversible behavior of the Dicke maser model. Technically however, our master equation is derived from the condition of detailed balance, whereas they obtain a Markovian evolution by using the so-called "singular reservoirs." They obtain a bifurcation phenomenon. We have a linear evolution equation.

In Sec. II we take this occasion to give a mathematically rigorous and detailed description of the equilibrium states of the Dicke maser model, with a particular emphasis on the explicit construction of the condensed phase below the critical temperature. In Sec. III we define our Markovian dynamics and give a complete spectral resolution of the generator of the process, showing the above mentioned phenomenon of critical slowing down.

II. MATHEMATICAL DESCRIPTION OF THE MODEL AND ITS EQUILIBRIUM STATES

A. Algebra of observables

Consider the C^*-tensor product $\mathscr{C}=\mathscr{B}\otimes\mathscr{A}$, where $\mathscr{B}$ is the boson algebra and $\mathscr{A}$ the atomic algebra; $\mathscr{A}$ is the usual quasilocal spin-$\frac{1}{2}$ lattice algebra

$$\mathscr{A}=\overline{\cup_N(\otimes_{i=-N}^{N}(M_2)_i)},$$

where $(M_2)_i$ is a copy of the complex 2×2 matrices M_2 at the site $i\in\mathbb{Z}$; M_2 is generated by the Pauli matrices $\sigma^\pm$ satisfying $\sigma^{+2}=0$, $\sigma^+\sigma^-+\sigma^-\sigma^+=1$.

For the boson algebra $\mathscr{B}$ we take the CCR-C^*-algebra $\overline{\Delta(H,\sigma)}$ constructed on a symplectic space (H,σ). It is generated by the set of Weyl operators $\{W(\phi)\,|\,\phi\in H\}$ satisfying

$$W(\phi)W(\psi)=W(\phi+\psi)e^{-(i/2)\sigma(\phi,\psi)},$$
$$W(\phi)^*=W(-\phi).$$

Now we discuss the sympletic space (H,σ), which is relevant for our purposes.

Denote

$$L^2(\mathbb{Z})=\left\{f\,\middle|\,f\colon\mathbb{Z}\to C;\ \sum_{n\in\mathbb{Z}}|f(n)|^2<\infty\right\}$$

and $\mathscr{S}$ the subspace of $L^2(\mathbb{Z})$ defined by

$$\mathscr{S}=\left\{f\in L^2(\mathbb{Z})\,\middle|\,\lim_{|n|\to\infty}|n|^pf(n)=0,\ \forall\,p\in\mathbb{N}\right\},$$

$L^2(\mathbb{Z})$ is a Hilbert space with scalar product

$$\langle f,g\rangle_2=\sum_{n\in\mathbb{Z}}\bar f(n)g(n),\quad f,g\in L^2(\mathbb{Z}).$$

Define the Fourier transform $\mathscr{F}\colon\mathscr{S}\to L^2([0,2\pi])$ by

$$f\to(\mathscr{F}f)(k)=\hat f(k)=\sum_{n\in\mathbb{Z}}e^{-ikn}f(n).$$

Define the Hilbert space H as the completion of $\mathscr{S}$ with respect to the scalar product

$$\langle f,g\rangle=\langle f,g\rangle_2+\bar{\hat f}(0)\hat g(0);\ f,g\in\mathscr{S}.$$

Note that H is isomorphic to $L^2(\mathbb{Z})\oplus\mathbb{C}$.

[a] Onderzoeker I.I.K.W., Belgium.

 0022-2488/87/092250-04$02.50

The imbedding of $\mathscr{S}$ into $H\simeq L^2(\mathbf{Z})\oplus\mathbf{C}$ is given by

$$f\in\mathscr{S}\to\phi=(f,\hat{f}(0))\in H$$

and as an orthonormal basis of H we consider

$$\{(e_n,0),(0,1)\,|\,n\in\mathbf{Z}\},$$

where $e_n\in\mathscr{S}$ is the function defined by $e_n(m)=\delta_{nm}$. Remark that $\hat{e}_n(0)=1$ and that in the L^2 sense of H,

$$(0,1)=\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}(e_n,1). \qquad (2)$$

The symplectic form σ on H is now defined: for $\phi=(f,\delta_1)$, $\psi=(g,\delta_2)$,

$$\sigma(\phi,\psi)=\operatorname{Im}\langle f,g\rangle_2.$$

It is clear that σ is degenerated, hence the C^*-algebra $\mathscr{B}=\overline{\Delta(H,\sigma)}$ (Ref. 12) has a nontrivial center generated by the set

$$\{W((0,\gamma))\,|\,\gamma\in\mathbf{C}\}.$$

Let ω be a regular state[13] of $\mathscr{B}$, then there exists in the GNS representation of ω creation and annihilation operators $a^\pm(\phi)$; $\phi\in H$, satisfying the following commutation relations: for $(f,\gamma),(g,\delta)\in H$,

$$W((f,\gamma))=\exp(i/\sqrt{2})(a^+((f,\gamma))+a^-((f,\gamma)))$$

and

$$[a^-((f,\gamma)),a^+((g,\delta))]=\langle f,g\rangle_2.$$

Denote

$$a^\pm((e_n,1))=a_n^\pm;\ n\in\mathbf{Z},$$
$$a^\pm((0,1))=\alpha^\pm,$$

then (2) implies that

$$\underset{N\to\infty}{\text{s-lim}}\frac{1}{2N+1}\sum_{n=-N}^{N}a_n=\alpha. \qquad (3)$$

Finally we need the one-parameter group of gauge transformations $\{\tau_\theta\,|\,\theta\in[0,2\pi)\}$:

$$\tau_\theta a^\pm(\phi)=e^{\pm i\theta}a^\pm(\phi),\quad \tau_\theta\sigma_k^\pm=e^{\pm i\theta}\sigma_k^\pm.$$

B. The equilibrium states

The model shows a phase transition, i.e, there exists a critical temperature T_c determined by

$$\beta_c=(2/\epsilon)th^{-1}(\epsilon/\lambda^2)$$

such that for $\beta<\beta_c$ one has the normal phase state given by a product state ω_n with a density matrix per lattice site $k\in\mathbf{Z}$ (Ref. 10),

$$\rho_k^n=\frac{\exp[-\beta(a_k^+a_k+\epsilon\sigma_k^+\sigma_k^-)]}{\operatorname{tr}\exp[-\beta(a_k^+a_k+\epsilon\sigma_k^+\sigma_k^-)]}.$$

One computes

$$\begin{aligned}&\omega_n(W(f,\gamma))=\exp[-\tfrac14\coth(\beta/2)\langle f,f\rangle_2],\\&\omega_n(a^+(f,\gamma_1)a(g,\gamma_2))=[1/(e^\beta-1)]\,\langle g,f\rangle_2,\\&\omega_n(\alpha)=0.\end{aligned} \qquad (4)$$

For $\beta>\beta_c$ one has also the condensed phase described by a convex set K_β of equilibrium states with extremal elements $\{\omega_{c,\theta}\,|\,\theta\in[0,2\pi)\}$, where $\omega_{c,\theta}$ is a product state with a density matrix per lattice site $k\in\mathbf{Z}$ (Ref. 10),

$$\rho_{k,\theta}^c=\frac{\exp\{-\beta\,[a_k^+a_k+\epsilon\sigma_k^+\sigma_k^-+re^{i\theta}(a_k-\lambda\sigma_k^-)+re^{-i\theta}(a_k^+-\lambda\sigma_k^+)]\}}{\operatorname{tr}\exp\{-\beta\,[a_k^+a_k+\epsilon\sigma_k^+\sigma_k^-+re^{i\theta}(a_k-\lambda\sigma_k^-)+re^{-i\theta}(a_k^+-\lambda\sigma_k^+)]\}},$$

where r is the order parameter, determined by the gap equation

$$2\sigma=th\beta\lambda^2\sigma\neq0;\qquad r^2=\lambda^2\sigma^2-\epsilon^2/4\lambda^2. \qquad (5)$$

One checks that

$$\omega_{c,\theta}\cdot\tau_\phi=\omega_{c,(\theta+\phi)\bmod 2\pi}$$

and we denote

$$\omega_c=\frac{1}{2\pi}\int_0^{2\pi}d\theta\,\omega_{c,\theta}$$

the unique gauge invariant equilibrium state. One computes

$$\begin{aligned}\omega_{c,\theta}(W((f,\gamma)))&=\exp[-\tfrac14\coth(\beta/2)\langle f,f\rangle_2\\&\quad-(i/\sqrt{2})(\bar{\gamma}re^{-i\theta}+\text{h.c.})],\\\omega_{c,\theta}(a^+(f,\gamma_1)a(g,\gamma_2))&=[1/(e^\beta-1)]\langle g,f\rangle_2+\bar{\gamma}_2\gamma_1r^2,\\\omega_{c,\theta}(\alpha)&=-re^{-i\theta}.\end{aligned} \qquad (6)$$

C. Time evolution

All equilibrium states are symmetric for arbitrary $\mathbf{Z}$-lattice point permutations and as the model is of the mean field type (see Refs. 14 and 15) one has an effective evolution $\{\alpha_t\,|\,t\in\mathbf{R}\}$ on the von Neumann algebra $\mathscr{C}''$ induced by each of the equilibrium states yielding

$$\alpha_t(a^\pm((f,\gamma)))=a^\pm((e^{it}f,\gamma)).$$

III. DETAILED BALANCE EVOLUTIONS AND APPROACH TO EQUILIBRIUM

Return to equilibrium has recently been studied in some models by means of quantum stochastic evolutions satisfying the condition of detailed balance.[5,6]

Here we consider another model and in this section we introduce the evolutions on the level of the von Neumann algebra $\mathscr{C}''$ induced by an equilibrium state of the model. Hence let ω be any equilibrium state, a linear map γ with a dense domain $\mathscr{D}$ of $\mathscr{C}''$ into itself satisfies the detailed balance condition if, for all $x,y\in\mathscr{D}$,

$$\omega(x\gamma(y))=\omega(\gamma(x)y).$$

Following Refs. 4 and 5 we consider the following linear dissipative maps satisfying this condition:

$$\begin{aligned}L_x^F(\cdot)=\int dt\,F(t)\mathscr{M}_s(\alpha_t(x)[\cdot,\alpha_{t+s}(x)]\\+[\alpha_s(x),\cdot]\alpha_{t+s}(x)),\end{aligned} \qquad (7)$$

where $\mathscr{M}_s$ stands for the mean over s,

$$\mathscr{M}_s = \lim_{a\to\infty} \frac{1}{2a} \int_{-a}^{a} \cdot ds,$$

and where F is any complex analytic function satisfying

(i) $F \in L^1(\mathbb{R}, dt)$,

(ii) $\hat{F} > 0$,

(iii) $F(-t + i\beta) = F(t)$,

and $x = x^*$ a self-adjoint operator, α_t is the effective time evolution in the state ω.

Now there are many possible choices for the function F as well as for the observable x. Here we consider two types of states; a normal state ω_n and a condensed phase state ω_c.

For each of them we make a particular choice for the observable x, given by

$$x_n = a^+(\psi) + a(\psi),$$

$$x_c = \alpha a^+(\psi) + \alpha^+ a(\psi),$$

where $\psi = (g,0) \in H$; and $g \in \mathscr{S}$, such that $\hat{g}(0) = 0$.

This choice of x implies that the maps L_x^F are so-called quasifree maps, i.e., their dual action on the state space maps generalized free states into generalized free states. The choice of x_c can also be motivated on physical grounds. Indeed, it is the minimal gauge-invariant observable coupling the condensed mode with an excited one.

In the following we denote shortly L_n^F (resp. L_c^F) for $L_{x_n}^F$ (resp. $L_{x_c}^F$). A tedious but straightforward computation yields the following lemma.

Lemma 3.1: For any $\phi \in H$, then

$$L_{n,c}^F(W(\phi)) = \left[(i/\sqrt{2})\left(a(T_{n,c}^g \phi) + a^+(T_{n,c}^g \phi)\right) + \chi_{n,c}^g(\phi) \right] W(\phi),$$

where the $T_{n,c}^g$ are linear operators on H,

$$T_{n,c}^g(\phi) = -(\epsilon_{n,c}^g / \langle \psi, \psi \rangle) \langle \psi, \phi \rangle \psi, \quad \phi \in H,$$

and

$$\epsilon_n^g = \hat{F}(1)(1 - e^{-\beta}) \langle g, g \rangle_2, \quad \epsilon_n^g > 0,$$

$$\epsilon_c^g = \hat{F}(1)(1 - e^{-\beta}) \langle g, g \rangle_2 r^2, \quad \epsilon_c^g \geqslant 0,$$

$$\chi_n^g(\phi) = -\tfrac{1}{2}\hat{F}(1) |\langle \phi, \psi \rangle|^2 (1 + e^{-\beta}),$$

$$\chi_c^g(\phi) = -\tfrac{1}{2}\hat{F}(1) |\langle \phi, \psi \rangle|^2 (1 + e^{-\beta}) r^2. \quad ∎$$

It is clear from the definition of the operators $T_{n,c}^g$ that they are rank-1 operators projecting on the vector ψ. Let $H = H_1 \oplus H_2$, with H_1 being the subspace generated by ψ.

By identification

$$W(\phi_1 \oplus \phi_2) = W(\phi_1) \otimes W(\phi_2), \quad \phi_1 \in H_1, \quad \phi_2 \in H_2,$$

and

$$\overline{\Delta(H,\sigma)} = \overline{\Delta(H_1,\sigma)} \otimes \overline{\Delta(H_2,\sigma)}.$$

From formulas (4) and (6) one has

$$\omega_{n,c}(x_1 \otimes x_2) = \omega_{n,c}(x_1) \omega_{n,c}(x_2)$$

for $x_1 \in \overline{\Delta(H_1,\sigma)}$, $x_2 \in \overline{\Delta(H_2,\sigma)}$, i.e., the $\omega_{n,c}$ are product states with respect to this decomposition, and

$$L_{n,c}^F x_2 = 0 \quad \text{for all } x_2 \in \overline{\Delta(H_2,\sigma)}.$$

Another computation yields the following lemma.

Lemma 3.2: For $\phi \in H_1$ and $m, m' \in \mathbb{N}_0$,

$$L_{n,c}^F(a^+(\phi)^m a(\phi)^{m'}) = -\epsilon_{n,c}^g (m + m') a^+(\phi)^m a(\phi)^{m'} + 2\epsilon_{n,c}^g m m' \omega_{n,c}(a^+(\phi) a(\phi)) a^+(\phi)^{m-1} a(\phi)^{m'-1},$$

$$L_{n,c}^F(a(\phi)^m a^+(\phi)^{m'}) = -\epsilon_{n,c}^g (m + m') a(\phi)^m a^+(\phi)^{m'} + 2\epsilon_{n,c}^g m m' \omega_{n,c}(a(\phi) a^+(\phi)) a(\phi)^{m-1} a^+(\phi)^{m'-1}.$$

If m or m' equals zero, the second term is absent. ∎

Now we consider the action of the operators $L_{n,c}^F$ on the representation space induced by the equilibrium states $\omega_{n,c}$. Actually, we consider the operators $\tilde{L}_{n,c}^F$ defined by

$$\tilde{L}_{n,c}^F(x \Omega_{n,c}) = L_{n,c}^F(x) \Omega_{n,c}, \tag{8}$$

where the $\Omega_{n,c}$ are the cyclic vectors of the states $\omega_{n,c}$ and x runs through a dense set of operators on the representation spaces. The detailed balance condition yields that the operator $\tilde{L}_{n,c}^F$ can be defined as a self-adjoint, negative operator. We have the following result about $\sigma(\tilde{L}_{n,c}^F)$ = spectrum of $\tilde{L}_{n,c}^F$.

Theorem 3.3:

$$\sigma(\tilde{L}_{n,c}^F) = \{ -p\epsilon_{n,c}^g \mid p \in \mathbb{N} \}.$$

Proof: From Lemma 3.1 it is clear that all vectors of the type $x_2 \Omega_{n,c}$, $x_2 \in \overline{\Delta(H_2,\sigma)}$ belong to the zero eigenvalue of $\tilde{L}_{n,c}^F$. Therefore it remains to consider vectors generated by $\{x_1 \Omega_{n,c} \mid x_1 \in \Delta(H,\sigma)\}$. Denote by $\mathscr{H}_{nc}^1$ the subspace generated by this set. It is also generated by the vectors of the type

$$a^+(\phi)^m a(\phi)^{m'} \Omega_{n,c},$$

$$a(\phi)^m a^+(\phi)^{m'} \Omega_{n,c},$$

$\phi \in H_1$, $m, m' \in \mathbb{N}$. In turn $\mathscr{H}_{nc}^1$ is generated by the vectors

$$A^{m,m'} \Omega_{n,c} \quad \text{and} \quad B^{m,m'} \Omega_{n,c},$$

$m, m' \in \mathbb{N}$, where

$$A^{m,m'} = \sum_{k=0}^{p} c_k(m,m') a^+(\phi)^{m-k} a(\phi)^{m'-k},$$

$$B^{m,m'} = \sum_{k=0}^{p} d_k(m,m') a(\phi)^{m-k} a^+(\phi)^{m'-k},$$

with $p = \min(m,m')$, and

$$c_k(m,m') = \frac{(-1)^k [\omega_{n,c}(a^+(\phi) a(\phi))]^k m! m'!}{k!(m-k)!(m'-k)!},$$

$$d_k(m,m') = \frac{(-1)^k [\omega_{n,c}(a(\phi) a^+(\phi))]^k m! m'!}{k!(m-k)!(m'-k)!}.$$

From Lemma 3.2 it follows that

$$L_{n,c}^F A^{m,m'} = -(m + m') \epsilon_{n,c}^g A^{m,m'},$$

$$L_{n,c}^F B^{m,m'} = -(m + m') \epsilon_{n,c}^g B^{m,m'},$$

and the theorem follows. ∎

This theorem makes clear that the spectrum of $\tilde{L}_{n,c}^F$ is completely determined by the number $\epsilon_{n,c}^g$. This constant is called the energy gap. For all $\beta < \beta_c$ this constant is strictly larger than zero (see Lemma 3.1) but for $\beta > \beta_c$ (i.e., $T < T_c$) $\epsilon_c^g \to 0$ as $T \to T_c$. This is the germ of the phenomenon of critical slowing down, which we discuss now.

Using Theorem 3.3 it is clear that the map $\tilde{L}_c^F$ is exponentiable in order to yield the dynamical semigroup evolution

$$\gamma_t^F = \exp t\tilde{L}_c^F, \quad t \geqslant 0.$$

The energy gap ϵ_c^g for fixed g and F determines clearly the rate of convergence to equilibrium of any locally perturbed state, i.e., for any local observables A one has

$$\omega_c(A^*\gamma_t^F A) - \omega_c(A^*A) \leqslant e^{-t\epsilon_c^g}\omega_c(A^*A)$$

and the lifetime $\tau(T)$ of the locally perturbed state

$$B \to \omega_c(A^*BA)/\omega_c(A^*A)$$

is therefore given by $(\epsilon_c^g)^{-1}$. From the gap equation (5) and the definition of the critical temperature T_c it readily follows that ϵ_c^g tends to zero as $T \to T_c$ like $T - T_c$. Therefore the lifetime

$$\tau(T) \simeq 1/(T - T_c)^\eta \quad \text{with } \eta = 1.$$

Hence we obtained a mathematically rigorous proof of the phenomenon of critical slowing down with critical exponent equal to unity, in complete agreement with the result for mean fields of classical lattice systems and the Glauber dynamics. This result enhances our statement that our model (7) is the quantum mechanical version of the classical Glauber dynamics.

[1]R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[2]M. Suzuki and R. Kubo, J. Phys. Soc. Jpn. **24**, 51 (1968).

[3]E. Stoll, K. Binder, and T. Schneider, Phys. Rev. B **8**, 3266 (1973).

[4]J. Quaegebeur, G. Stragier, and A. Verbeure, Ann. Inst. H. Poincaré **41**, 25 (1984).

[5]G. Stragier, Ph. D. thesis, Leuven, 1984.

[6]J. Quaegebeur and A. Verbeure, Lett. Math. Phys. **9**, 93 (1985).

[7]E. Buffet, Ph. de Smedt, and J. V. Pulé, Ann. Phys. (NY) **155**, 269 (1984).

[8]R. H. Dicke, Phys. Rev. **93**, 99 (1954).

[9]K. Hepp and E. H. Lieb, Ann. Phys. (NY) **76**, 360 (1973).

[10]M. Fannes, P. N. M. Sisson, A. Verbeure, and J. C. Wolfe, Ann. Phys. (NY) **98**, 38 (1976).

[11]K. Hepp and E. H. Lieb, Helv. Phys. Acta **46**, 573 (1974).

[12]J. Manuceau, M. Sirugue, D. Testard, and A. Verbeure, Commun. Math. Phys. **32**, 231 (1973).

[13]O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics II* (Springer,Berlin, 1979).

[14]E. Størmer, J. Funct. Anal. **3**, 48 (1969).

[15]M. Fannes, H. Spohn, and A. Verbeure, J. Math. Phys. **21**, 355 (1980).

Erratum: Regular subalgebras of Lie superalgebras and extended Dynkin diagrams [J. Math. Phys. 28, 292 (1987)]

J. Van der Jeugt[a)]
Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit Gent, Krijgslaan 281-S9, B9000 Gent, Belgium

(Received 10 March 1987; accepted for publication 25 March 1987)

We refer to the paper in the title as I. Our classification of regular subalgebras of simple classical Lie superalgebras was based on Kac's list of nonequivalent simple root systems.[1] After the publication of I, it was pointed out to me that the results of Kac are not complete as far as the exceptional Lie superalgebras $G(3)$ and $F(4)$ are concerned.[2] In view of this, our paper should be corrected as follows.

At the end of Sec. VII one should add: Besides the simple root systems (7.3) and (7.9) there exist two more nonequivalent systems, namely,

$$\pi_1 = \{-\delta - \epsilon_1, \delta - \epsilon_3, \epsilon_3 - \epsilon_2\},$$

$$\pi_2 = \{\epsilon_3 - \delta, \delta - \epsilon_2, \epsilon_2\}.$$

These last two choices do not give rise to any new regular subalgebras of $G(3)$.

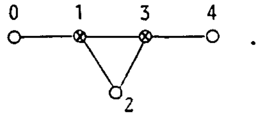
At the end of Sec. VIII one should add: Besides the four root systems given by Kac, there are still two nonequivalent simple root systems for $F(4)$, namely,

$$\pi = \{-\delta + \tfrac{1}{2}(\epsilon_1 + \epsilon_2 - \epsilon_3), \epsilon_3, \delta + \tfrac{1}{2}(-\epsilon_1 + \epsilon_2 - \epsilon_3), \epsilon_1 - \epsilon_2\},$$

$$\pi' = \{2\delta, -\delta + \tfrac{1}{2}(-\epsilon_1 + \epsilon_2 + \epsilon_3), \epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3\}.$$

The system π can be extended by $\gamma = -\epsilon_1 - \epsilon_2$, and the extended Cartan matrix and Dynkin diagram are given by

$$\begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 2 & 0 & 1 & -2 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & -2 & 1 & 0 & 2 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}, \quad \tau = \{1,3\},$$



Deleting node 2 gives rise to the Dynkin diagram and root system of the Lie superalgebra $A(3,0)$. This regular subalgebra of $F(4)$ was not found by means of any other simple root system. Note that π' does not lead to any regular subalgebras not obtained before.

In conclusion, Table I of I should be extended by the regular subalgebra $A(3,0)$ of $F(4)$.

[a)] Senior research assistant N.F.W.O. (Belgium).

[1] V. G. Kac, Adv. Math. 26, 8 (1977).
[2] J. Van de Leur, Ph.D. dissertation, University of Utrecht, 1986 (unpublished).

Erratum: Quantum motion on a half-line connected to a plane [J. Math. Phys. 28, 386 (1987)]

P. Exner and P. Šeba
Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, 141980 Dubna, USSR

(Received 3 February 1987; accepted for publication 11 February 1987)

The vectors (10) should have the same norm. Hence the first one of the relations (11) must be replaced

$$f_1(x) := 8^{1/4} \exp(\bar{\epsilon} x);$$

then the additional multiplicative factor $8^{1/4}$ will appear in (16c) and (30b), and $8^{-1/4}$ will appear in (16d), (17b)–(17d), and (30c).

 0022-2488/87/092254-01$02.50