Nonlinear spinor representations

P. Furlan
International School for Advanced Studies, Trieste, Italy and Dipartimento di Fisica Teorica dell'Università di Trieste, Trieste, Italy and Instituto Nazionale di Fisica Nucleare, Sezione di Trieste, Trieste, Italy

R. Rączka
International School for Advanced Studies, Trieste, Italy and Institute of Nuclear Studies, Nuclear Theory Department, Warsaw, Poland

(Received 8 April 1985; accepted for publication 5 June 1985)

A systematic method for the construction of nonlinear carrier spaces for a class of nonlinear spinor representations of complex and pseudo-orthogonal rotation groups is presented. It is shown that Cartan pure spinors, which satisfy quadratic constraints, are special cases of our construction. A class of new nonlinear spinor representations is discovered, which is particularly interesting in the case of generalized conformal groups $SO(\nu+1,\nu-1)$, $\nu=3,4,\ldots$. The nonlinearity condition considerably diminishes the number of independent spinor components and therefore the corresponding spinor fields are the most natural building blocks for grand unified field theories. The method presented here is universal and can be applied for the construction of new nonlinear representations of other higher-symmetry groups.

I. INTRODUCTION

There is currently a widespread belief that field theories in higher-dimensional space-times may play an important role in understanding four-dimensional space-time quantum field theories.[1] This belief is supported by the fact that for field theories in certain space-times, the anomalies of gauge or gravitational field theories disappear.[2] The many attractive features of higher-dimensional space-times are diminished by the fact that in these cases the number of Dirac fields—when the theory is finally restricted to the four-dimensional space-time—is very high.[3]

In this work we would like to point out that there exist nonlinear spinor representations in higher-dimensional space-times for which the number of independent spinor components is considerably diminished in a natural manner. These spinor field theories resemble nonlinear spinorial σ models with covariant quadratic constraints restricting the number of independent components.[4]

To give a concrete example consider the so-called neutral space-times $\mathbb{R}^{\nu,\nu}$ with the kinematical group $SO(\nu,\nu)$, $\nu=2,3,\ldots$. Let Γ_a, $a=1,\ldots,2\nu$, be $2^\nu\times 2^\nu$ generalized Dirac matrices satisfying the anticommutation relations

$$\{\Gamma_a,\Gamma_b\}=2g_{ab}\mathbf{1},\quad a,b=1,\ldots,2\nu, \tag{1.1}$$

where g_{ab} is the metric tensor of the $\mathbb{R}^{\nu,\nu}$ space-time. Let C be the matrix satisfying the relations

$$C\Gamma_a=(-1)^\nu\Gamma_a^T C,\quad CC^T=\mathbf{1},\quad C^2=(-1)^{\nu(\nu+1)/2}, \tag{1.2}$$

and let $\Gamma_{a_1\cdots a_r}$, $r=1,\ldots,\nu$, be the completely antisymmetrized product of Γ_a matrices. Let ψ be a spinor for $SO(\nu,\nu)$ in the carrier space L^{m_+} of dimension $2^{\nu-1}$ defined by the highest weight $m_+=(\frac{1}{2},\ldots,\frac{1}{2})$ (see Ref. 5). Let $\tilde{\psi}\doteq\psi^T C$. A spinor ψ_P in L^{m_+} is said to be pure if it satisfies the following set of quadratic constraints[6,7]:

$$\tilde{\psi}_P\Gamma_{a_1\cdots a_k}\psi_P=0,\quad \text{for } k=0,1,\ldots,\nu-1. \tag{1.3}$$

It was shown that the constraints (1.3) are covariant and that the number d of independent constraints equals

$$d=2^{\nu-1}-1-\binom{\nu}{2}. \tag{1.4}$$

Hence the number d_ψ of independent pure spinor components is

$$d_\psi=1+\binom{\nu}{2}. \tag{1.5}$$

Formula (1.5) for $\nu=5$ gives 11 components for a pure spinor instead of 16 for an ordinary (semi-)spinor. However for $\nu=10$, formula (1.5) gives 46 independent components versus 512 for an ordinary spinor. This demonstrates the suppressing mechanism of independent spinor components due to the constraints (1.3).

The constraints (1.3) make the carrier space N^{m_+} for pure spinors nonlinear: in fact, if ψ_1 and ψ_2 satisfy the constraints (1.3) then $\psi=\lambda_1\psi_1+\lambda_2\psi_2$ with $\lambda_1,\lambda_2\in\mathbb{C}$ will not satisfy (1.3) in general.

Notice that if a given ψ_P satisfies (1.3) then its group transform $T_g\psi_P$—due to covariance of (1.3)—will also satisfy (1.3). In fact let $D_{a_1\cdots a_r;a_1'\cdots a_r'}(g)$ be the matrix elements of the polyvector representation: then, for $r<\nu$,

$$\begin{aligned}&(\widetilde{T_g\psi_P})\Gamma_{a_1\cdots a_r}(T_g\psi_P)\\&\quad=\tilde{\psi}_P(T_g^{-1}\Gamma_{a_1\cdots a_r}T_g)\psi_P\\&\quad=D_{a_1\cdots a_r;a_1'\cdots a_r'}(g)\tilde{\psi}_P\Gamma_{a_1'\cdots a_r'}\psi_P=0.\end{aligned} \tag{1.6}$$

It is remarkable that the property (1.6) allows us to represent the intrinsic components ψ_α, $\alpha=1,\ldots,1+\binom{\nu}{2}$ of a pure spinor in terms of group parameters of the so-called instability group C. In fact let $H\subset SO(\nu,\nu)$ be the stability subgroup of ψ_P, i.e., for $h\in H$ we have

$$T_h\psi_P=\psi_P. \tag{1.7}$$

Then, due to the Mackey decomposition theorem (see Sec. II), there exists a set $C\subset SO(\nu,\nu)$ such that any $g\in SO(\nu,\nu)$ has the representation

$$g=ch,\quad c\in C,\quad h\in H. \tag{1.8}$$

Then the spinor $\psi=T_g\psi_P$, which is pure by (1.6), can be

 0022-2488/85/123021-12$02.50

written in the form

$$\psi = T_c \psi_P = \psi(c), \tag{1.9}$$

i.e., the elements of the pure spinor space N^{m+} can be labeled by the group elements $c \in C$. We show in Sec. II that C coincides—up to a set of Haar measure zero—with the solvable group

$$\tilde{C} = T^{\binom{\nu}{2}} \circledS T^1, \tag{1.10}$$

where T_2^ν is a $\binom{\nu}{2}$-dimensional Abelian subgroup of SO(ν,ν). Consequently the nonlinear spinor space N^{m+} of pure spinors can be identified with the group space of the solvable group (1.10).The group parameters $\{c_k\}_{k=1}^{1+\binom{\nu}{2}} \in C$ may be considered as the $[1+\binom{\nu}{2}]$-intrinsic coordinates of the pure spinor $\psi(c)$.

The action of T_{g_0} on pure spinors in this representation is determined by the Mackey decomposition (1.8): in fact for any $g_0 \in \mathrm{SO}(\nu,\nu)$ we have

$$T_{g_0}\psi(c) = T_{g_0 c}\psi_P = T_{c_{g_0 c}} T_{h_{g_0 c}} \psi_P = \psi(c_{g_0 c}), \tag{1.11}$$

where the group element $c_{g_0 c}$ is uniquely determined by the formula

$$g_0 c = c_{g_0 c} h_{g_0 c}. \tag{1.12}$$

One may introduce in a natural manner a nonlinear covariant wave equation for spinor fields transforming according to a nonlinear representation.[4,8] In fact following the Cartan construction[9] one may assume that intrinsic spinor components $\{c_k\} \in C$ depend on space-time coordinates $x \in \mathbf{R}^{\nu,\nu}$. In this case the most natural Dirac-like covariant equation

$$\Gamma_a \frac{\partial}{\partial x_a} \psi[\{c_k(x)\}] = 0 \tag{1.13}$$

is a nonlinear equation since the N^{m+}-spinor space to which $\psi[\{c_k(x)\}]$ belongs is nonlinear. Using the representation (1.9) for a pure spinor and (1.10) for the group elements $c \in C$, one reduces (1.13) to a specific system of nonlinear covariant wave equations for $c_k(x)$ functions. The explicit form and the properties of these solutions are considered in our separate work.[8] In this work we limit ourselves to a presentation of the basic results, which allow the reduction of a nonlinear carrier space N^m for a nonlinear spinor representation to the specific homogeneous space G/H, where H is the stability subgroup of a chosen pure spinor ψ_P. The construction of nonlinear N^m spaces underlined above for the SO(ν,ν) group is universal and one carries out this construction for an arbitrary rotation group in two steps.

(1) Find the stability group (1.7) for the chosen spinor representation.

(2) Find the coset space G/H, which, by (1.9), gives a description of the elements in the nonlinear carrier space N^m by means of certain homogeneous space coordinates. It turns out that in the most important cases the coset space G/H may be represented—up to a set of Haar measure zero—with some group space as, e.g., (1.10).

In Sec. II we present a general formalism of nonlinear group representations and we illustrate it in the case of nonlinear spinor representations of the SO(2ν,C) groups.In Sec. III we extend this analysis to SO($2\nu+1$,C) groups. In Sec. IV we present a theory of nonlinear spinor representations for the pseudo-orthogonal groups SO(p,q), $p \geqslant q$, $p+q=2\nu$. The structure of the nonlinear spaces N^m strongly depends on the chosen signature (p,q) of space-time and is rather rich. In the neutral ($\mathbf{R}^{\nu,\nu}$) and the conformal ($\mathbf{R}^{\nu+1,\nu-1}$) cases the N^m spaces can be represented as some specific group spaces. In all other remaining cases the N^m space coincides with specific G/H spaces depending on the space-time signature. The stability subgroup H for the SO($2\nu-h,h$) spinor representations has the general form

$$H = [\mathrm{SU}(\nu-h) \times \mathrm{SL}(h,\mathbf{R})] \circledS R, \tag{1.14}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{h}{2}} + d^{2h(\nu-h)}, \tag{1.15}$$

with $t^{\binom{h}{2}}$ a $\binom{h}{2}$-dimensional Lie algebra and $d^{2h(\nu-h)}$ a $2h(\nu-h)$-dimensional vector space in the so($2\nu-h,h$) Lie algebra.

We see that the stability groups H for even-dimensional pseudo-orthogonal groups have a rather rich structure and run from

$$H = \mathrm{SU}(\nu)$$

for SO(2ν) to

$$H = \mathrm{SL}(\nu,\mathbf{R}) \circledS T^{\binom{\nu}{2}}$$

for the neutral case SO(ν,ν).

The parallel analysis for SO(p,q) groups, $p>q$, $p+q=2\nu+1$, acting in odd-dimensional space-times $\mathbf{R}^{p,q}$ is carried out in Sec. V. The stability groups H for the highest weight spinor ψ_m, $m=(\frac{1}{2},\frac{1}{2},\ldots,\frac{1}{2})$ of SO($2\nu+1-h,h$) groups have also a rich structure and are of the form

$$H = [\mathrm{SU}(\nu-h) \times \mathrm{SL}(h,\mathbf{R})] \circledS R, \tag{1.16}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{h}{2}} \dot{+} d^{2h(\nu-h)+h}, \tag{1.17}$$

with $t^{\binom{h}{2}}$ a $\binom{h}{2}$-dimensional Abelian Lie algebra and $d^{2h(\nu-h)+h}$ a $[2h(\nu-h)+h]$-dimensional vector space in so($2\nu+1-h,h$) Lie algebra.

The nonlinear carrier spinor space N^m may be identified with the quotient space SO($2\nu+1-h,h$)/H, which in turn for $h=\nu$ and $h=\nu-1$ may be represented—up to a set of Haar measure zero—as the group space of a specific solvable group.

Finally, we conclude this work with Sec. VI, where we emphasize the importance of nonlinear spinor representations for the construction of nonlinear relativistic field theories.

Let us note that for an arbitrary group G the set C in the Mackey decomposition (1.8) is homeomorphic with the homogeneous space G/H of the left H cosets. In fact by (1.8) we have

$$gH = c_g h_g = c_g H.$$

Hence every left coset gH can be uniquely characterized by the element c_g in C. Conversely—by virtue of (1.8)—every element $c \in C$ determines uniquely the left coset cH and two

different elements c_1 and c_2 in C determine different cosets c_1H and c_2H in G/H.

We follow in this work the Cartan notation.[6] In the compact SO(n) case in $\mathbb{R}^n$ we take the anticommutation relations of generalized Dirac γ matrices in the form

$$\{\gamma_a,\gamma_b\} = 2\delta_{ab}\mathbf{1}, \quad a,b = 1,...,n. \tag{1.18}$$

It turns out to be convenient in all proofs to use the H basis of Cartan,[6] which is given by the set of $2[n/2]+1$ matrices $\{H_0,H_j,H_{j'}\}, j = 1,...,[n/2]$, satisfying the anticommutation relations

$$\{H_r,H_s\} = 2g_{rs}\mathbf{1}, \quad r,s = 0,1,...,\nu,1',...,\nu', \tag{1.19}$$

with $\nu \equiv [n/2]$, where

$$g = \begin{Vmatrix} 0 & \frac{1}{2}\mathbf{1}_\nu \\ \frac{1}{2}\mathbf{1}_\nu & 0 \end{Vmatrix}, \quad \text{for } n = 2\nu,$$

and

$$g = \begin{Vmatrix} 1 & & 0 \\ & 0 & \frac{1}{2}\mathbf{1}_\nu \\ 0 & & \\ & \frac{1}{2}\mathbf{1}_\nu & 0 \end{Vmatrix}, \quad \text{for } n = 2\nu+1. \tag{1.20}$$

This basis is connected to the γ_a basis by the relations

$$\begin{aligned} \gamma_{2j-1} &= H_j + H_{j'}, \\ \gamma_{2j} &= i(H_{j'} - H_j), \\ \gamma_{2\nu+1} &= H_0. \end{aligned} \tag{1.21}$$

The generators S_{ab} of so(n) for spinor representations are defined by the formula

$$S_{ab} = (-i/4)[\gamma_a,\gamma_b], \tag{1.22}$$

and they satisfy the following commutation relations:

$$[S_{ab},S_{cd}] = i(\delta_{ac}S_{bd} + \delta_{bd}S_{ac} - \delta_{ad}S_{bc} - \delta_{bc}S_{ad}). \tag{1.23}$$

The Clifford algebra units Γ_a and the corresponding generators (1.22) of the so(p,q) Lie algebras are obtained by multiplying $q-\gamma_a$'s by i. The metric tensor for the so($n-h,h$) Lie algebra is taken in the form

$$g_{ab} = \begin{cases} 0, & \text{for } a \neq b, \\ 1, & \text{for } a = b = \begin{cases} 1,2,...,2([n/2]-h), \\ 2l-1, \quad \text{with } l = [n/2]-h+1,...,\{n/2\}, \end{cases} \\ -1, & \text{for } a = b = 2k, \quad \text{with } k = [n/2]-h+1,...,[n/2], \end{cases} \tag{1.24}$$

and the corresponding Γ_a's satisfy the following anticommutation relations:

$$\{\Gamma_a,\Gamma_b\} = 2g_{ab}\mathbf{1}. \tag{1.25}$$

Introducing the generators of so($n-h,h$) in the form

$$X_{ab} = -\tfrac{1}{4}[\Gamma_a,\Gamma_b], \tag{1.26}$$

we have

$$[X_{ab},X_{cd}] = g_{ac}X_{bd} + g_{bd}X_{ac} - g_{ad}X_{bc} - g_{bc}X_{ad}. \tag{1.27}$$

II. NONLINEAR GROUP REPRESENTATIONS

We begin our analysis with a precise definition of a nonlinear group representation. Let G be a topological group and N a nonlinear topological space. We say that the map $g \to T_g$ is a nonlinear representation of G in N if the following conditions are satisfied.

(1) With each $g \in G$ there is associated a transformation $T_g: n \to T_g n$ of N into N.

(2) The identity element e of G is the identity transformation of N.

(3) The mapping $(g,n) \to T_g n$ of $G \times N$ into N is continuous.

(4) For $g_1,g_2 \in G$ and $n \in N$ we have

$$T_{g_1g_2}n = T_{g_1}(T_{g_2}n).$$

The novelty of nonlinear representations consists in the condition that the carrier space N for the representation is nonlinear. The nonlinearity condition for the carrier space makes the analysis and the classification of nonlinear representations very difficult.[10] For instance, the linear unitary irreducible representations of the Poincaré group are all classified.[5] At the same time the classification problem of nonlinear representations of the Poincaré group is equivalent to the classification problem of all possible solutions of all possible nonlinear relativistic wave equations, which is clearly an unsolvable problem. This is the reason why we have so far very few papers on the properties of nonlinear representations.

We shall analyze in this work nonlinear spinor representations. We begin our analysis with an illustration of how a nonlinear carrier space N appears naturally in the case of spinor theory.

Consider first the SO($2\nu,\mathbb{C}$) complex rotation group in the even-dimensional complex "space-time" $\mathbb{C}^{2\nu}$. It is well known that this group possesses two kinds of irreducible spinor representations given by the semispinors $\overset{\text{I}}{\psi}$ and $\overset{\text{II}}{\psi}$ of the first and the second kind, respectively.[5] The linear carrier space L^{m+} for the linear spinor representation T^{m+} of SO($2\nu,\mathbb{C}$) has the dimension $2^{\nu-1}$ (see Ref. 5). It was shown, however, by Cartan[6] that in the carrier space L^{m+} one can introduce the concept of a pure spinor, which provides the carrier space N^{m+} for a nonlinear representation of SO($2\nu,\mathbb{C}$). In fact, let γ_a, $a = 1,...,2\nu$, be the generators of the Clifford algebra for the linear representation T^{m+} of

SO(2ν,C) and let

$$\gamma_{a_1 a_2 \cdots a_r} = \gamma_{[a_1}\gamma_{a_2}\cdots\gamma_{a_r]}, \quad r = 0,1,\ldots,2\nu,$$

by polyvectors in L. Let $C = \gamma_2\gamma_4\cdots\gamma_{2\nu}$. Then the pure spinors are defined as the subset of semispinors, say of the first kind, which satisfy the conditions ($\tilde{\psi} \equiv \psi^T C$)

$$\tilde{\psi}\gamma_{a_1\cdots a_r}\psi = 0, \quad r = 0,1,\ldots,\nu-1. \tag{2.1}$$

Since

$$C\gamma_a = (-1)^\nu \gamma_a^T C, \tag{2.2}$$

the constraints (2.1) have covariant form, i.e., they hold in any reference frame. It is shown in Ref. 6 that the number d of independent constraints given by (2.1) is equal to

$$d = 2^{\nu-1} - 1 - \binom{\nu}{2}. \tag{2.3}$$

Hence the space N of pure spinors has the dimension

$$d_\psi = 1 + \binom{\nu}{2}, \tag{2.4}$$

which, for $\nu > 3$, is smaller than the dimension $2^{\nu-1}$ of the linear spinor representation. Clearly if ψ_1 and ψ_2 satisfy (2.1) then their linear combination

$$\psi = \alpha_1\psi_1 + \alpha_2\psi_2, \quad \alpha_1,\alpha_2 \in \mathbb{C},$$

in general does not satisfy (2.1). Hence the pure spinors form in the $2^{\nu-1}$-dimensional carrier space of the linear spinor representation a $[1 + \binom{\nu}{2}]$-dimensional nonlinear carrier space N^{m_+} for a nonlinear spinor representation.

It is instructive to see how the space N^{m_+} and the nonlinear representation are explicitly realized in the pure spinor case. We show this using the concept of the highest weight spinor ψ_{m_+}. As is well known, the linear spinor representations of SO(2ν,C) are characterized by the highest weights $m_\pm$ given by the formula[5]

$$m_\pm = (\tfrac{1}{2},\tfrac{1}{2},\ldots,\tfrac{1}{2},\pm\tfrac{1}{2}); \tag{2.5}$$

m_+ corresponds to semispinors of the first and m_- of the second kind, respectively. We can choose an explicit representation for the Clifford algebra $\mathbb{C}_{2\nu}$ given by Eq. (1.18), such that the following lemma holds.

Lemma 2.1: The highest weight m_+ semispinor has the form

$$\psi_{m_+} = \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}. \tag{2.6}$$

Proof: Let us label the 2^ν rows and columns of the γ and H matrices *à la* Cartan, i.e., using the completely skew-symmetric set of indices

$$i_1 i_2 \cdots i_p, \quad p = 0,1,\ldots,\nu, \qquad i_1,i_2,\ldots,i_p = 1,2,\ldots,\nu \tag{2.7}$$

(where the index "0" means that no index appears). Then we use the following explicit realization of the Clifford algebra basis [see Ref. 6 Sec. 92(c)]:

$$(H_j)_{i_1 i_2 \cdots i_p}^{\;\;i_1 i_2 \cdots i_p j} = \begin{cases} 0, & \text{for } j \in i_1,i_2,\ldots,i_p, \\ 1, & \text{for } j \notin i_1,i_2,\ldots,i_p, \end{cases} \tag{2.8a}$$

$$(H_{j'})_{i_1 i_2 \cdots i_p}^{\;\;i_1 i_2 \cdots i_p} = \begin{cases} 0, & \text{for } j \in i_1,i_2,\ldots,i_p, \\ 1, & \text{for } j \notin i_1,i_2,\ldots,i_p, \end{cases} \tag{2.8b}$$

$$(H_0)_{i_1 i_2 \cdots i_p}^{\;\;i_1 i_2 \cdots i_p} = (-1)^p, \tag{2.8c}$$

all other matrix elements being zero.

The Cartan subalgebra of so(2ν,C) is given by the ν generators

$$S_{2j-1,2j}, \quad j = 1,2,\ldots,\nu. \tag{2.9}$$

Since only the elements of the first column of S_{ab} count, when S_{ab} acts on ψ_{m_+}, let us investigate the first column's structure of $S_{2j-1,2j}$. From Eqs. (1.21) and (1.22) we have

$$S_{2j-1,2j} = \tfrac{1}{2}(H_j H_{j'} + H_{j'} H_j - H_j H_j - H_{j'} H_{j'}). \tag{2.10}$$

But Eqs. (2.8) give

$$(H_j H_j)_{i_1 i_2 \cdots i_p}^{\;\;0} = (H_{j'} H_{j'})_{i_1 i_2 \cdots i_p}^{\;\;0} = (H_{j'} H_j)_{i_1 i_2 \cdots i_p}^{\;\;0} = 0, \tag{2.11a}$$

$$(H_j H_{j'})_{i_1 i_2} \cdots i_p^{\;\;0} = \delta_{p0}, \tag{2.11b}$$

and then

$$(S_{2j-1,2j})_{i_1 i_2 \cdots i_p}^{\;\;0} = \tfrac{1}{2}\delta_{p0}, \tag{2.12}$$

which proves the lemma. ▼

Let l^+ and l^- be the vector space of raising and lowering operators, respectively, and let h be the Cartan subalgebra of so(2ν,C). Let

$$\text{so}(2\nu,\mathbb{C}) = l^+ \dot{+} h \dot{+} l^- \tag{2.13}$$

be the Cartan decomposition of so(2ν,C). Then the linear envelope of all vectors

$$\prod_{k=1}^{r} l_{i_k}^- \psi_{m_+}, \quad l_{i_k}^- \in l^-, \quad \text{for } r = 0,1,\ldots,\dim T^{m_+},$$

coincides with the linear carrier space of the linear spinor representation T^{m_+} (see Ref. 5). This shows the importance of the highest weight spinor ψ_{m_+} for the linear spinor representation theory.

We show now that ψ_{m_+} is also crucial for the construction of a nonlinear group representation. First we notice that by Eqs. (2.1), (1.21), (2.6), and (2.8) we have

$$\tilde{\psi}_{m_+}\gamma_{a_1\cdots a_r}\psi_{m_+} = 0, \quad r = 0,1,\ldots,\nu-1. \tag{2.14}$$

Hence ψ_{m_+} represents a pure spinor belonging to $N^{m_+} \subset L^{m_+}$. We have the following theorem.

Theorem 2.2: The stability group H of ψ_{m_+} is the connected semidirect product group

$$H = \text{SL}(\nu,\mathbb{C}) \circledS T^{(\nu)}_{(2)}, \tag{2.15}$$

where $T^{(\nu)}_{(2)}$ is an Abelian $\binom{\nu}{2}$-dimensional subgroup of SO(2ν,C).

Proof: We shall first look for the stability subalgebra h of ψ_{m_+}, i.e., for all linearly independent generators S_{ab} in so(2ν,C), which satisfy the condition

$$S_{ab}\psi_{m_+} = 0. \tag{2.16}$$

In order to investigate the action of the generators of so(2ν,C) on ψ_{m_+}, it is necessary (and sufficient) to know all the elements of just the first column of the matrices S_{ab}.

From Eqs. (2.8) we have

$$(H_jH_l)_{i_1\cdots i_p}{}^0 = (H_{j'},H_l)_{i_1\cdots i_p}{}^0 = 0, \tag{2.17a}$$

$$(H_jH_{l'})_{i_1\cdots i_p}{}^0 = \delta_{p0}\delta_{jl}, \tag{2.17b}$$

$$(H_{j'}H_{l'})_{i_1\cdots i_p}{}^0 = \delta_{p2}(\delta_{i_1 l}\delta_{i_2 j} - \delta_{i_1 j}\delta_{i_2 l}), \tag{2.17c}$$

with $j,l = 1,\ldots,\nu$, $p = 0,1,\ldots,\nu$, and $i_1,\ldots,i_p = 1,\ldots,\nu$.

Then, taking into account Eqs. (1.19)–(1.22) and (2.17), we have that the $\binom{2\nu}{2}$ generators S_{ab} of so(2ν,C) can be divided into the two following linearly independent sets: (a) $(3\nu^2 - \nu - 2)/2$ matrices having all elements zero in the first column, i.e.,

$$\binom{\nu}{2} \text{ matrices } \mathscr{T}_{jl}$$

$$= -(i/2)[H_j,H_l]$$

$$= \tfrac{1}{2}\{S_{2j-1,2l-1} - S_{2j,2l} + i(S_{2j-1,2l} + S_{2j,2l-1})\}, \tag{2.18a}$$

$$\nu^2 - 1 \text{ matrices } \mathscr{P}_{jl}$$

$$= -\frac{i}{2}[H_j,H_{l'}] + \delta_{jl}\frac{1}{\nu}\frac{i}{2}\sum_{k=1}^{\nu}[H_k,H_{k'}]$$

$$= \tfrac{1}{2}\{S_{2j-1,2l-1} + S_{2j,2l} + i(S_{2j,2l-1} - S_{2j-1,2l})\}$$

$$+ \frac{i}{\nu}\delta_{jl}\sum_{k=1}^{\nu}S_{2k-1,2k}; \tag{2.18b}$$

(b) $(\nu^2 - \nu + 2)/2$ matrices having some non-null element in the first column, i.e.,

$$\binom{\nu}{2} \text{ matrices } \mathscr{Q}_{jl}$$

$$= -(i/2)[H_{j'},H_{l'}]$$

$$= \tfrac{1}{2}\{S_{2j-1,2l-1} - S_{2j,2l} - i(S_{2j-1,2l} + S_{2j,2l-1})\}, \tag{2.19a}$$

$$1 \text{ matrix } \mathscr{D} = \tfrac{1}{2}[H_\nu,H_{\nu'}] = S_{2\nu-1,2\nu}. \tag{2.19b}$$

We see that only the first set belongs to the stability subalgebra h. The generators (2.18) satisfy the commutation relations

$$[\mathscr{T}_{jl},\mathscr{T}_{mn}] = 0, \tag{2.20a}$$

$$[\mathscr{P}_{jl},\mathscr{P}_{mn}] = i(\delta_{jn}\mathscr{P}_{ml} - \delta_{lm}\mathscr{P}_{jn}), \tag{2.20b}$$

$$[\mathscr{P}_{jl},\mathscr{T}_{mn}] = i(\delta_{ln}\mathscr{T}_{jm} - \delta_{lm}\mathscr{T}_{jn} + (2/\nu)\delta_{jl}\mathscr{T}_{mn}), \tag{2.20c}$$

i.e., the generators $\mathscr{T}_{jl}$ form an Abelian ideal of h, while the generators $\mathscr{P}_{jl}$ satisfy the commutation relations of the simple complex Lie algebra sl(ν, C). Furthermore the generators $\mathscr{P}_{jl}$ generate an automorphism of the Abelian subalgebra. Therefore, if we denote by $t^{\binom{\nu}{2}}$ the $\binom{\nu}{2}$-dimensional Abelian subalgebra, we can say that h is given by

$$h = \mathrm{sl}(\nu,\mathbb{C}) \oplus t^{\binom{\nu}{2}}. \tag{2.21}$$

Using the general connection between Lie algebras and the connected Lie groups given by Theorem 3.3 of Ref. 5 we obtain the assertion of Theorem 2.2. ▼

The crucial tool in understanding the structure of the stability Lie algebra and of the set complementary to it is provided by the Levi–Malcev theorem.[5] Let l be an arbitrary Lie algebra over R or C and let r be its radical, i.e., a maximal solvable ideal. Then the Levi–Malcev theorem states[5] that there exists a semisimple Lie subalgebra s of l such that

$$l = s \oplus r. \tag{2.22}$$

In the case (2.15) $r = t^{\binom{\nu}{2}}$ and $s = \mathrm{sl}(\nu,\mathbb{C})$.

Now let G be a Lie group and H a closed subgroup of G. Then the Mackey decomposition theorem states that there exists a Borel set C in G such that every element $g \in G$ can be uniquely represented in the form[5]

$$g = ch, \quad c \in C, \quad h \in H. \tag{2.23}$$

In general the set C is not a group. It is, however, very interesting that in the case of SO(2ν,C) and H given by (2.15) we have the following theorem.

Theorem 2.3: Let $G = \mathrm{SO}(2\nu,\mathbb{C})$ and let H be the stability subgroup (2.15) of ψ_{m_+}. Then the complementary set C may be represented—up to a set of Haar measure zero—as the connected solvable subgroup

$$\tilde{C} = T^{\binom{\nu}{2}} \circledS T^1. \tag{2.24}$$

Proof: The set of generators in so(2ν,C) complementary to $\mathrm{sl}(\nu,\mathbb{C}) \oplus t^{\binom{\nu}{2}}$ is given by the generators $\mathscr{Q}_{jl}$ and $\mathscr{D}$ of Eqs. (2.19), with the following commutation relations:

$$[\mathscr{Q}_{jl},\mathscr{Q}_{mn}] = 0, \tag{2.25a}$$

$$[\mathscr{D},\mathscr{Q}_{jl}] = \delta_{jv}\mathscr{Q}_{lv} - \delta_{lv}\mathscr{Q}_{jv}. \tag{2.25b}$$

We see that c is a solvable Lie algebra

$$c = t^{\binom{\nu}{2}} \oplus t^1.$$

By Theorem 3.3 in Ref. 5 there is a connected solvable Lie subgroup $\tilde{C}$ of G with the Lie algebra c and

$$\dim \tilde{C} = \dim G - \dim H.$$

On the other hand, for the Mackey set C we have

$$\dim C = \dim G - \dim H.$$

Hence the group space $\tilde{C}$ coincides with the Mackey set C up to a set $C - \tilde{C}$ of the Haar measure zero. Since the group G acts transitively on $\tilde{C}$, the group space $\tilde{C}$ may be also used for a description of the carrier space N^{m_+} for the nonlinear spinor representation. ▼

Theorems 2.2 and 2.3 allow us to give the explicit realization for the N^{m_+}-nonlinear carrier spinor space. In fact, since ψ_{m_+} is a pure spinor, the spinor

$$\psi = T_g\psi_{m_+} \tag{2.26}$$

is also pure. Indeed by (2.2) we have

$$\tilde{\psi}\Gamma_{a_1\cdots a_r}\psi = D_{a_1\cdots a_r;a_1'\cdots a_r'}(g)\tilde{\psi}_{m_+}\Gamma_{a_1'\cdots a_r'}\psi_{m_+} = 0,$$

$$r = 1,\ldots,\nu-1. \tag{2.27}$$

Using the Mackey theorem we have $T_g = T_cT_h$ and by Theorem 2.2 we have

$$\psi = T_c\psi_{m_+} = \psi(c). \tag{2.28}$$

We see, therefore, that the carrier space N^{m_+} can be represented as the connected group space $\tilde{C}$ generated according to Theorem (3.3) of Ref. 5 by the Lie algebra c. By (2.24) the dimension of C is

$$d_c = 1 + \binom{\nu}{2}, \tag{2.29}$$

and coincides with the dimension d_ψ of the pure spinor space given by (2.4). The action of SO(2ν,C) in N^{m_+} is given by the group action dictated by the Mackey decomposition. In fact, by (2.23)

$$g_0 c = c_{g_0 c} h_{g_0 c}. \tag{2.30}$$

Hence

$$T_{g_0}\psi(c) = T_{g_0 c}\psi_{m_+} = \psi(c_{g_0 c}). \tag{2.31}$$

It is evident from this formula that the action of T_g on N^{m_+} is highly nonlinear.

Table I gives the dimensions of the linear and the corresponding nonlinear spinor representations for SO(2ν,C) and the number of constraints.

We see that the dimension of the nonlinear spinor representation grows much more slowly than the dimension of the corresponding linear representation. This fact may be very useful in constructing grand unified theories in higher dimensions, where for the time being we have too many spinor fields.[3]

III. NONLINEAR SPINOR REPRESENTATION OF SO (2ν + 1,C)

The SO(2ν + 1,C) group possesses one kind of linear irreducible fundamental spinor representation T^m, which is determined by the highest weight $m = (\frac{1}{2}, \ldots, \frac{1}{2})$ and has the dimension 2^ν (see Ref. 5). The pure spinor nonlinear carrier space N^m is defined as the set of all ψ's in L^m which satisfy the quadratic constraints

$$\tilde{\psi}\gamma_{a_1 \cdots a_k}\psi = 0, \quad \text{for} \quad k < \nu. \tag{3.1}$$

It is shown in Ref. 6 that the number d of independent constraints (3.1) is given by the formula

$$d = 2^\nu - 1 - \tfrac{1}{2}\nu(\nu+1). \tag{3.2}$$

Hence the dimension d_ψ of the nonlinear pure spinor space N^m is given by the formula

$$d_\psi = 1 + \nu + \binom{\nu}{2}. \tag{3.3}$$

Since the constraints (3.1) are quadratic, the space N^m is evidently nonlinear. Since in addition the constraints (3.1) are given in a covariant form, the nonlinear space N^m is invariant under the action of the group representation $g \to T_{g(\epsilon)} = e^{\epsilon_{ab}\gamma_{ab}}$. Consequently the space N^m represents the carrier space of dimension d_ψ of a nonlinear, SO(2ν + 1,C) representation.

TABLE I. Dimensions of the linear and nonlinear spinor representations for SO(2ν,C) and the number of corresponding constraints.

ν	2	3	4	5	6	10
$2^{\nu-1}$	2	4	8	16	32	512
$1+\binom{\nu}{2}$	2	4	7	11	16	46
$2^{\nu-1}-1-\binom{\nu}{2}$	0	0	1	5	16	466

We now give the explicit realization of N^m space as the coset space of the SO(2ν + 1,C) group over its distinguished subgroup H. We have the following theorem.

Theorem 3.1: The stability group H of the highest weight spinor ψ_m is the group

$$H = \mathrm{SL}(\nu,\mathbb{C}) \circledS R, \tag{3.4}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{\nu}{2}} + e^\nu,$$

with $t^{\binom{\nu}{2}}$ a $\binom{\nu}{2}$-dimensional Abelian subalgebra and e^ν a ν-dimensional vector space in the so(2ν + 1,C) Lie algebra.

Proof: We show first that the highest-weight spinor ψ_m has the form

$$\psi_m = \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}. \tag{3.5}$$

In fact the Cartan subalgebra of so(2ν + 1,C) is given by the same ν generators (2.9) of so(ν,C). Hence from Lemma 2.1 we have that ψ_m, being identical to ψ_{m_+} of Eq. (2.6), still corresponds to the highest weight $m = (\frac{1}{2}, \ldots, \frac{1}{2})$.

Let us then look for the stability subalgebra h of ψ_m. In addition to Eqs. (2.17) we find

$$(H_j H_0)_{i_1 \cdots i_p}{}^0 = (H_0 H_j)_{i_1 \cdots i_p}{}^0 = 0, \tag{3.6a}$$

$$(H_{j'} H_0)_{i_1 \cdots i_p}{}^0 = -(H_0 H_{j'})_{i_1 \cdots i_p}{}^0 = \delta_{p1}\delta_{i_1 j}, \tag{3.6b}$$

$j,l = 1, \ldots, \nu, p = 0,1, \ldots, \nu$, and $i_1, \ldots, i_p = 1, \ldots, \nu$. Then for the so(2$\nu$ + 1,C) Lie algebra, in addition to the generators (2.18) and (2.19) of the so(2ν,C) Lie algebra, we find further

(a) ν matrices $\mathscr{E}_j = -(i/2)[H_j, H_0] = S_{2j-1,2\nu+1} + iS_{2j,2\nu+1}$, (3.7a)

having all elements zero in the first column and

(b) ν matrices $\mathscr{F}_j = -(i/2)[H_{j'}, H_0] = S_{2j-1,2\nu+1} - iS_{2j,2\nu+1}$, (3.7b)

having some non-null element in the first column.

We see that the set (a) given by the $(3\nu^2 + \nu - 2)/2$ matrices $\{\mathscr{T}_{jl}, \mathscr{P}_{jl}, \mathscr{E}_j\}$ of Eqs. (2.18) and (3.7a) form the stability subalgebra h. In addition to the commutation relations (2.20) we have, in fact,

$$[\mathscr{E}_j, \mathscr{E}_l] = 2i\mathscr{T}_{jl}, \tag{3.8a}$$

$$[\mathscr{E}_j, \mathscr{T}_{mm}] = 0, \tag{3.8b}$$

$$[\mathscr{E}_j, \mathscr{P}_{mm}] = i(\delta_{jn}\mathscr{E}_m - (1/\nu)\delta_{mn}\mathscr{E}_j). \tag{3.8c}$$

From Eqs. (2.20) and (3.8) we see that the structure of the stability subalgebra h is just that one corresponding to Eq. (3.4). ▼

The Mackey decomposition states that there exists a Borel set C in SO(2ν + 1,C) such that any $g \in$ SO(2ν + 1,C) can be represented as $c\,h$, $c \in C$, $h \in H$. In the present case we have the following theorem.

Theorem 3.2: The set C may be represented—up to a set of Haar measure zero—as a solvable Lie group whose Lie algebra c has the following structure:

$$c = t^{1} \oplus (t^{(2)} \dot{+} f^{\nu}), \tag{3.9}$$

with $t^{(2)}$ a $\binom{\nu}{2}$-dimensional Abelian Lie algebra and f^{ν} a ν-dimensional vector space in $so(2\nu+1,\mathbb{C})$.

Proof: The set c of generators, complementary in $so(2\nu+1,\mathbb{C})$ to the set of generators giving the stability subalgebra h of ψ_m, is formed by the generators $\mathcal{Q}_{jl}$ and $\mathcal{D}$ of Eqs. (2.19) plus the generators $\mathcal{F}_j$ of Eq. (3.7b). They satisfy the commutation relations (2.25) plus

$$[\mathcal{F}_j, \mathcal{F}_l] = 2i\mathcal{Q}_{jl}, \tag{3.10a}$$

$$[\mathcal{F}_j, \mathcal{Q}_{mn}] = 0, \tag{3.10b}$$

$$[\mathcal{F}_j, \mathcal{D}] = \delta_{j\nu}\mathcal{F}_{\nu}. \tag{3.10c}$$

From Eqs. (2.25) and (3.10) we see that the set c forms a solvable Lie algebra, whose structure is given by Eq. (3.9). Repeating the argument given at the end of the proof of Theorem 2.3 we obtain the assertion of Theorem 3.2. ▼

Now, since ψ_m satisfies the constraints (3.1), it is pure. Since the constraints are convariant, the spinor

$$\psi = T_g\psi_m = T_c\psi_m = \psi(c) \tag{3.11}$$

is also pure. Consequently, the nonlinear carrier space of pure spinors can be identified with the group space C given by Theorem 3.2. By (3.9) the dimension of C is

$$d_C = 1 + \nu + \binom{\nu}{2}$$

and coincides with the dimension d_ψ of the N^m space given by (3.3). Consequently the nonlinear space N^m of pure spinors can be identified with the quotient space G/H, which in the present case may be represented as the group space $\tilde{C}$ given by Theorem 3.2.

Table II compares the dimension of the linear spinor representation, the dimension of the nonlinear one, and the number of independent constraints. We see again that the dimension of the nonlinear spinor representation is much smaller—especially for the higher space-time dimensions—than the corresponding dimension of the linear representation.

IV. NONLINEAR SPINOR REPRESENTATIONS FOR SO(p,q) GROUPS, $p+q=2\nu$

We shall construct now a class of nonlinear spinor representations for pseudo-orthogonal groups $SO(p,q)$, $p+q=2\nu$, $p \geqslant q$. For the sake of simplicity we shall denote this class of groups by the symbol $SO(2\nu-h,h)$, $h=0,1,\ldots,\nu$. We also shall analyze first the case of semispinors of the first kind. Guided by the results of Secs. II and III we shall construct the nonlinear spinor representations for $SO(2\nu-h,h)$ in two steps: (1) determination of the stability subgroup H of the highest weight spinor ψ_{m_+}, and (2) determination of the Mackey complementary set C.

TABLE II. Dimensions of the linear and nonlinear spinor representations for $SO(2\nu+1,\mathbb{C})$ and the number of corresponding constraints.

ν	2	3	4	5	6	10
2^{ν}	4	8	16	32	64	1024
$1+\nu+\binom{\nu}{2}$	4	7	11	16	22	56
$2^{\nu}-1-\nu-\binom{\nu}{2}$	0	1	5	16	42	968

Contrary to the results of Sec. II and III it will turn out in some geometrically distinguished cases that the set C may be represented as the group space of a specific subgroup of $SO(2\nu-h,h)$, but in most cases it is just a homogeneous space. Those nonlinear spinor representations for which C is a group space should be—from the geometrical and the physical point of view—very interesting. We have the following theorem.

Theorem 4.1: Let $G=SO(2\nu-h,h)$, with $h=0,1,\ldots,\nu$. Then the stability group H of the highest weight spinor ψ_{m_+} is the connected group

$$H = [SU(\nu-h)\times SL(h,\mathbb{R})] \circledS R, \tag{4.1}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{(h)} + d^{2h(\nu-h)}, \tag{4.2}$$

with $t^{(h)}$ a $\binom{h}{2}$-dimensional Abelian Lie algebra and $d^{2h(\nu-h)}$ a $2h(\nu-h)$-dimensional vector space in the $so(2\nu-h,h)$ Lie algebra.

Proof: (See Appendix A.)

Theorem 4.1 shows the rich structure of the highest-weight spinor stability subgroups, depending on the signature of the 2ν-dimensional space-time $\mathbb{R}^{2\nu-h,h}$, $h=0,1,\ldots,\nu$. In the case $h=0$ (i.e., Euclidean space-time $\mathbb{R}^{2\nu}$) by (4.1) we have

$$H = SU(\nu). \tag{4.3}$$

In this case the nonlinear carrier spinor space N^{m_+} will coincide with the quotient

$$C = SO(2\nu)/SU(\nu), \tag{4.4}$$

which in turn coincides—up to a phase factor—with the so-called quadric Grassmannians.[11] At the other end, if we set $h=\nu$ (i.e., we consider the so-called neutral space-time $\mathbb{R}^{\nu,\nu}$), then by (4.1) we have

$$H = SL(\nu,\mathbb{R}) \circledS T^{(\nu)}, \tag{4.5}$$

where $T^{(\nu)}$ is an Abelian group of dimension $\binom{\nu}{2}$. The analysis given below shows then that in this case the space N^m of nonlinear spinors may be represented as the group space of the solvable group C given by

$$C = T^{(\nu)} \circledS T_1, \tag{4.6}$$

where $T^{(\nu)}$ is an Abelian $\binom{\nu}{2}$-dimensional subgroup of $SO(\nu,\nu)$. We hope that these two extreme cases illustrate the richness of the considered theory.

Formula (4.1) gives also the dimension of the stability

subgroup and the dimension of the Mackey set C, which coincides with the dimension of the nonlinear carrier space N^m. We present these results in Table III. It follows from Table III that the cases $h = 0$, ν are not obtainable as particular cases of the general h case.

The following theorem shows that some signatures are distinguished.

Theorem 4.2: Let $G = \mathrm{SO}(2\nu\text{–}h,h)$. Then for $h = \nu$ and $\nu - 1$ the carrier space N^{m_+} for the nonlinear spinor representation may be represented—up to a set of Haar measure zero—as the group space

$$\tilde{C} = T^{\binom{\nu}{2}} \circledS T^1, \quad \text{for } h = \nu, \tag{4.7}$$

and

$$\tilde{C} = R, \quad \text{for } h = \nu - 1,$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{\nu-1}{2}} + d^{2\nu}, \tag{4.8}$$

with $t^{\binom{\nu-1}{2}}$ an Abelian algebra of dimension $\binom{\nu-1}{2}$ and $d^{2\nu}$ a 2ν-dimensional vector space in g.

In all remaining cases the carrier space N^{m_+} of the nonlinear spinor representation T^{m_+} coincides with the homogeneous space $\mathrm{SO}(2\nu - h,h)/H$.

Proof: (See Appendix B.)

Theorem 4.2 indicates that the neutral space-time $\mathbb{R}^{\nu,\nu}$ and the conformal space-time $\mathbb{R}^{\nu+1,\nu-1}$ are distinguished. The neutral case was already discussed in Sec. I. The conformal spaces $\mathbb{R}^{\nu+1,\nu-1}$ play the role of a natural generalization of the conformal space $\mathbb{R}^{4,2}$. It was proven recently[12]—using a rather complex extension of the Chevalley theory[13]—that the conformal spaces admit also a kind of nonlinear pure spinor representation. In our formalism these generalized pure spinors appear in an extremely natural manner. In fact the passage from the $\mathrm{so}(\nu,\nu)$ Lie algebra to the $\mathrm{so}(\nu+1,\nu-1)$ Lie algebra is obtained by multiplying by $(-i)$ the element Γ_2 of the Clifford algebra Γ_a, $a = 1,\ldots,2\nu$ of the $\mathbb{R}^{\nu,\nu}$ space. Such an operation will not change the constraints (1.3). Since the highest weight spinor ψ_{m_+} for $\mathrm{SO}(\nu+1,\nu-1)$ has the same form as for $\mathrm{SO}(\nu,\nu)$, the spinor ψ_{m_+} of $\mathrm{SO}(\nu+1,\nu-1)$ will also satisfy (1.3). By covariance arguments the spinor $T_g\psi_{m_+}$ will also satisfy the constraints (1.3). But by Theorems 4.1 and 4.2,

$$T_g\psi_{m_+} = T_c\psi_{m_+} = \psi(c), \tag{4.9}$$

where c is the element of the solvable group given by (4.7). Hence in the conformal case, similarly to the neutral case, the nonlinear carrier space N^{m_+} can be represented as the group space of the solvable group C. The space N^{m_+} has the dimension

$$d_\psi = 1 + \binom{\nu+1}{2}. \tag{4.10}$$

This dimension equals the dimension of the pure spinor nonlinear representation in the neutral $\mathbb{R}^{\nu+1,\nu+1}$ space.

TABLE III. Dimensions of the stability subgroup H and of the Mackey set C for the group $\mathrm{SO}(2\nu - h,h)$.

h	dim H	dim C
0	$\nu^2 - 1$	$\nu^2 - \nu + 1$
h	$\nu^2 - 2 + h(h-1)/2$	$\nu^2 - \nu + 2 - h(h-1)/2$
$\nu - 1$	$\frac{1}{2}\nu(\nu-1) - 1$	$1 + \binom{\nu+1}{2}$
ν	$\frac{1}{2}(3\nu^2 - \nu - 2)$	$1 + \binom{\nu}{2}$

V. NONLINEAR SPINOR REPRESENTATION OF SO(p,q) GROUPS, $p + q = 2\nu + 1$

We shall now construct a class of nonlinear spinor representations for the pseudo-orthogonal groups $\mathrm{SO}(p,q)$, $p + q = 2\nu + 1, p > q$, acting in the odd-dimensional space-time $\mathbb{R}^{p,q}$. As is well known, these groups have only one type of linear fundamental spinor representations, determined by the highest weight $m = (\frac{1}{2},\frac{1}{2},\ldots,\frac{1}{2})$ (ν components). For the sake of simplicity we shall denote this class of groups by the symbol $\mathrm{SO}(2\nu + 1 - h,h)$, $h = 0,1,\ldots,\nu$. We shall state the main results only, which, however, have a more complicated form than in the case of even-dimensional space-times.

Theorem 5.1: Let $G = \mathrm{SO}(2\nu + 1 - h,h)$, with $h = 0,1,\ldots,\nu$. Then the stability group H of the highest weight spinor has the form

$$H = [\mathrm{SU}(\nu\text{–}h) \times \mathrm{SL}(h,\mathbb{R})] \circledS R, \tag{5.1}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{h}{2}} + d^{2h(\nu-h)+h}, \tag{5.2}$$

with $t^{\binom{h}{2}}$ a $\binom{h}{2}$-dimensional Abelian Lie algebra and $d^{2h(\nu-h)+h}$ a $[2h(\nu-h)+h]$-dimensional vector space in the $\mathrm{so}(2\nu + 1\text{–}h,h)$ Lie algebra.

Proof: (See Appendix C.)

Similarly as in the even case, there are distinguished signatures for which the carrier space N^m of nonlinear spinor representations coincides with a group space. In fact we have the following theorem.

Theorem 5.2: Let $G = \mathrm{SO}(2\nu + 1 - h,h)$. Then for $h = \nu$ the carrier space N^m of nonlinear spinor representations may be represented—up to a set of Haar measure zero—as the group space

$$\tilde{C} = R^1 \circledS R, \quad h = \nu, \tag{5.3}$$

where R is a solvable group whose Lie algebra r has the following structure:

$$r = t^{\binom{\nu}{2}} + f^\nu \tag{5.4}$$

with $t^{\binom{\nu}{2}}$ a $\binom{\nu}{2}$-dimensional Abelian Lie algebra and f^ν a ν-dimensional vector space in the $\mathrm{so}(2\nu + 1 - h,h)$ Lie algebra.

For $h = \nu - 1$, N^m coincides with the group space

$$C = \mathrm{SO}(3) \circledS R, \tag{5.5}$$

where R is a solvable group with its Lie algebra r given by

$$r = t^{\binom{\nu}{2}-1} + f^{3\nu-2}, \tag{5.6}$$

with $t^{\binom{\nu}{2}-1}$ a $\binom{\nu}{2}-1$-dimensional Abelian Lie algebra and $f^{3\nu-2}$ a $(3\nu-2)$-dimensional vector space in the so$(2\nu+1-h,h)$ Lie algebra.

In all remaining cases the carrier space N^m coincides with the homogeneous space $SO(2\nu+1-h,h)/H$.

Proof: (See Appendix D.)

Formula (5.1) also determines the dimension of the stability subgroup and the dimension of the Mackey set C which is equal to the dimension of the nonlinear carrier space N^m. We present these results in Table IV.

Note that the cases $h=0$ and ν are not obtainable as particular cases of the general h case.

The nonlinear spinor representations in the $\mathbb{R}^{\nu+1,\nu}$ spaces were considered by Chevalley.[13] It was shown that in the $\mathbb{R}^{\nu+1,\nu}$ spaces there exist pure spinors ψ_P that satisfy the following quadratic constraints:

$$\tilde{\psi}_P \Gamma_{a_1\cdots a_r} \psi_P = 0, \quad \text{for } r = 0,1,...,\nu-1. \tag{5.7}$$

It was shown in Ref. 6 that the number of independent constraints given by (5.7) equals

$$2^\nu - 1 - \nu - \binom{\nu}{2}.$$

Hence the dimension d_ψ of the nonlinear carrier space N^m is

$$d_\psi = 1 + \nu + \binom{\nu}{2}.$$

Using the previous arguments we conclude that

$$\psi_P = T_g \psi_m = T_c \psi_m = \psi(c),$$

where c is an element of the solvable group C given by (5.3). It follows from Table IV that

$$d_C = 1 + \nu + \binom{\nu}{2},$$

i.e., it coincides with the dimension d_ψ determined by the number of independent quadratic constraints.

We see therefore that the nonlinear pure spinor space N^m may be represented as the group space C given by (5.3). The action of the nonlinear representation T_g of $SO(\nu+1,\nu)$ in C is determined by the Mackey decomposition.

The case $h=\nu-1$ is distinguished by the fact that the Mackey set C coincides—up to a set of Haar measure zero—with the solvable group given by (5.5). The construction of N^m nonlinear space is carried out as previously and this space may be represented as the group space of the solvable group given by (5.5).

TABLE IV. Dimensions of the stability subgroup H and of the Mackey set C for the group $SO(2\nu+1-h,h)$.

h	dim H	dim C
0	ν^2-1	$1+\nu+\nu^2$
h	$\nu^2-2+h(h+1)/2$	$2+\nu+\nu^2-h(h+1)/2$
$\nu-1$	$\frac{1}{2}(3\nu^2-\nu-4)$	$1+\binom{\nu+2}{2}$
ν	$\frac{1}{2}(3\nu^2+\nu-2)$	$1+\binom{\nu+1}{2}$

VI. DISCUSSION

We conclude this work with the following remarks.

(1) We have presented a systematic method for the construction of nonlinear spinor representations of complex and pseudo-orthogonal rotation groups. This method consists in finding the stability subgroup representation T_h, $h \in H$, of the highest weight spinor ψ_m and on the realization of the nonlinear carrier space N^m essentially as the quotient set G/H.

It should be stressed that the present method will provide the construction of nonlinear representations of $SO(p,q)$ groups also for other types of representations, in particular for higher spin representations $m=(n_1/2,\ n_2/2...,n_\nu/2)$, $n_i \geqslant 1$, for the tensor representations with $m=(m_1,...,m_\nu)$, $m_i \geqslant m_{i+1}$, m_i-non-negative integers, as well as for spin-tensor representations. In that manner one can associate with any linear $SO(p,q)$ representation determined by the highest weight $m=(m_1,...,m_\nu)$ a nonlinear group representation of much smaller (in general) dimension.

It is also clear that the above construction can be extended to any representation of any group G.

(2) The most interesting applications of nonlinear spinor representations are in field theory and particle physics. In fact, as we showed,[8] due to the nonlinear constraints (1.3) or (3.1) or effectively due to the nonlinearity of the carrier spinor space N^m, the simplest $SO(p,q)$ covariant Dirac-like wave equation

$$\Gamma_a \frac{\partial}{\partial x_a} \psi(x) = 0, \quad a = 1,...,p+q, \tag{6.1}$$

$\psi(x) \in N^m$, represents a nonlinear wave equation. This nonlinearity becomes explicit if we write $\psi(x)$ in terms of the intrinsic components $c_k(x)$, $k=1,2,...,$ dim N^m. In this case (6.1) reduces to a specific system of nonlinear field equations for the intrinsic components $c_k(x), k=1,...,$dim N^m. This is a new field of research which we present in detail.[8]

(3) The pure spinor field theories are attractive for elementary particle model builders since these models contain usually the smallest number of fundamental fields.[14] From the aesthetical point of view the nonlinear spinors that have the smallest number of independent components provide the smallest "building blocks" for an elementary particle model.[7,8] It would be very interesting to develop a canonical formalism for nonlinear spinors and check what particle spectrum for fermions and bosons follows from the nonlinearity of the theory.

From a general quantum field theory point of view nonlinear spinor field theories of the considered kind present a kind of field theory of σ-model type with covariant constraints. Hence one may, in principle, apply a standard Faddeev–Senjanovic method of quantization of such field theories.[15] The only novelty is connected with the fact that quadratic constraints are imposed on spinor components.

(4) The results of the present work might find also some applications in the theory of spontaneously broken quantum gauge field models based on $SO(n)$ or $SO(p,q)$ groups. In fact

the basic problem of such models is to find a stability subgroup of the chosen G-representation for the vacuum, which we solved with full generality for the considered class of groups.

ACKNOWLEDGMENTS

The authors thank Professor P. Budinich for inspiring discussions. One of the authors (R. R.) thanks Dr. S. Giler, Dr. P. Kosinski, Dr. P. Maslanka, and Dr. J. Rembielinski for constructive discussions and for showing their results on spinor theories prior to publication. He also gratefully acknowledges the warm hospitality extended to him by Professor P. Budinich and the administrative staff of the ISAS during his stay in Trieste.

APPENDIX A: PROOF OF THEOREM 4.1

Let us take as a basis for the Clifford algebra $\mathbb{R}_{2\nu-h,h}$ the 2ν ($2^\nu\times 2^\nu$) matrices Γ_m, $m=1,...,2\nu$, satisfying the anticommutation relations (1.25), given explicitly by the relations

$$\Gamma_{2j-1}=H_j+H_{j'}=\gamma_{2j-1}, \qquad j=1,...,\nu-h, \tag{A1a}$$
$$\Gamma_{2j}=i(H_{j'}-H_j)=\gamma_{2j}, \tag{A1b}$$
$$\Gamma_{2l-1}=H_l+H_{l'}=\gamma_{2l-1}, \qquad l=\nu-h+1,...,\nu. \tag{A1c}$$
$$\Gamma_{2l}=H_l-H_{l'}=i\gamma_{2l}, \tag{A1d}$$

It is easy to see that the generators $X_{ab}(a,b=1,...,2\nu)$ defined by Eq. (1.26) in terms of the Γ_a's of Eqs. (A1) satisfy the commutation relations (1.27) of the so($2\nu-h,h$) Lie algebra.

Out of all $\nu(2\nu-1)$ linearly independent generators X_{ab} of so($2\nu-h,h$) we can perform the two following sets of *real* linearly independent combinations: (a) $\nu^2-2+h(h-1)/2$ matrices having all elements zero in the first column, i.e.,

$\binom{\nu-h}{2}$ matrices A_{ij}

$$=-\tfrac{1}{2}\{[H_i,H_{j'}]+[H_{i'},H_j]\}=X_{2i-1,2j-1}+X_{2i,2j}, \tag{A2a}$$

$\binom{\nu-h}{2}+\nu-h-1$ matrices B_{ij}

$$=\frac{i}{2}\{-[H_i,H_{j'}]+[H_{i'},H_j]\}+i\frac{\delta_{ij}}{\nu-h}\sum_{g=1}^{\nu-h}[H_g,H_{g'}]$$
$$=X_{2i-1,2j}-X_{2i,2j-1}-\frac{2\delta_{ij}}{\nu-h}\sum_{g=1}^{\nu-h}X_{2g-1,2g}, \tag{A2b}$$

$\binom{h}{2}$ matrices T_{kl}

$$=-\tfrac{1}{2}[H_k,H_l]$$
$$=\tfrac{1}{2}\{X_{2k-1,2l-1}+X_{2k,2l}+X_{2k-1,2l}+X_{2k,2l-1}\}, \tag{A2c}$$

(h^2-1) matrices Π_{kl}

$$=-\tfrac{1}{2}[H_k,H_{l'}]+\frac{\delta_{kl}}{h}\sum_{m=\nu-h+1}^{\nu}\tfrac{1}{2}[H_m,H_{m'}]$$
$$=\tfrac{1}{2}\{X_{2k-1,2l-1}-X_{2k,2l}-X_{2k-1,2l}+X_{2k,2l-1}\}+\frac{\delta_{kl}}{h}\sum_{m=\nu-h+1}^{\nu}X_{2m-1,2m}, \tag{A2d}$$

$h(\nu-h)$ matrices C_{il}

$$=-\tfrac{1}{4}\{[H_i,H_l]+[H_{i'},H_{l'}]\}$$
$$=\tfrac{1}{2}(X_{2i-1,2l-1}+X_{2i-1,2l}), \tag{A2e}$$

$h(\nu-h)$ matrices D_{il}

$$=(i/4)\{[H_i,H_l]-[H_{i'},H_{l'}]\}$$
$$=\tfrac{1}{2}(X_{2i,2l-1}+X_{2i,2l}); \tag{A2f}$$

(b) $\nu^2-\nu+2-h(h-1)/2$ matrices having some non-null element in the first column, i.e.,

$\binom{\nu-h}{2}$ matrices $\tilde{A}_{ij}$

$$=-\tfrac{1}{2}\{[H_i,H_j]+[H_{i'},H_{j'}]\}=X_{2i-1,2j-1}-X_{2i,2j}, \tag{A3a}$$

$\binom{\nu-h}{2}$ matrices $\tilde{B}_{ij}$

$$=(i/2)\{[H_i,H_j]-[H_{i'},H_{j'}]\}=X_{2i-1,2j}+X_{2i,2j-1}, \tag{A3b}$$

$\binom{h}{2}$ matrices $\tilde{Q}_{kl}$

$$=-\tfrac{1}{2}[H_{k'},H_{l'}]=\tfrac{1}{2}\{X_{2k-1,2l-1}+X_{2k,2l}-X_{2k-1,2l}-X_{2k,2l-1}\}, \tag{A3c}$$

$h(\nu-h)$ matrices $\tilde{C}_{il}$

$$=-\tfrac{1}{4}\{[H_i,H_{l'}]+[H_{i'},H_{l'}]\}$$
$$=\tfrac{1}{2}(X_{2i-1,2l-1}-X_{2i-1,2l}), \tag{A3d}$$

$h(\nu-h)$ matrices $\tilde{D}_{il}$

$$=(i/4)\{[H_i,H_{l'}]-[H_{i'},H_{l'}]\}=\tfrac{1}{2}(X_{2i,2l-1}-X_{2i,2l}), \tag{A3e}$$

1 matrix $\tilde{B}$

$$=-(i/2)[H_{(\nu-h)},H_{(\nu-h)'}]=X_{2(\nu-h)-1,2(\nu-h)}, \tag{A3f}$$

1 matrix $\tilde{D}$

$$=\tfrac{1}{2}[H_\nu,H_{\nu'}]=X_{2\nu-1,2\nu}, \tag{A3g}$$

where $i,j=1,...,\nu-h$ and $k,l=\nu-h+1,...,\nu$.

From Eqs. (2.11) we see that only the first set (a) generates the stability subalgebra h. The generators (A2a)÷(A2f) satisfy the following commutation relations:

$$[A_{ij},A_{fg}]=\delta_{if}A_{jg}+\delta_{jg}A_{if}-\delta_{ig}A_{jf}-\delta_{jf}A_{ig}, \tag{A4a}$$
$$[B_{ij},B_{fg}]=\delta_{if}A_{jg}+\delta_{jg}A_{if}+\delta_{ig}A_{jf}+\delta_{jf}A_{ig}, \tag{A4b}$$
$$[A_{ij},B_{fg}]=\delta_{if}B_{jg}-\delta_{jg}B_{if}+\delta_{ig}B_{jf}-\delta_{jf}B_{ig}, \tag{A4c}$$
$$[\Pi_{kl},\Pi_{mn}]=\delta_{kn}\Pi_{ml}-\delta_{lm}\Pi_{kn}, \tag{A4d}$$
$$[T_{kl},T_{mn}]=0, \tag{A4e}$$

$$[\Pi_{kl},T_{mn}] = \delta_{ln}T_{km} - \delta_{lm}T_{kn} + (2\delta_{kl}/h)T_{mn}, \quad \text{(A4f)}$$

$$[A_{ij},\Pi_{kl}] = [B_{ij},\Pi_{kl}] = [A_{ij},T_{kl}] = [B_{ij},T_{kl}] = 0, \quad \text{(A4g)}$$

$$[C_{ik},C_{jl}] = \delta_{ij}\tfrac{1}{2}T_{kl}, \quad \text{(A4h)}$$

$$[D_{ik},D_{jl}] = \delta_{ij}\tfrac{1}{2}T_{kl}, \quad \text{(A4i)}$$

$$[C_{ik},D_{jl}] = 0, \quad \text{(A4j)}$$

$$[A_{ij},C_{fl}] = \delta_{if}C_{jl} - \delta_{jf}C_{il}, \quad \text{(A4k)}$$

$$[A_{ij},D_{fl}] = \delta_{if}D_{jl} - \delta_{jf}D_{il}, \quad \text{(A4l)}$$

$$[B_{ij},C_{fl}] = \delta_{if}D_{jl} + \delta_{jf}D_{il} - [2\delta_{ij}/(\nu-h)]D_{fl}, \quad \text{(A4m)}$$

$$[B_{ij},D_{fl}] = -\delta_{if}C_{jl} - \delta_{jf}C_{il} + [2\delta_{ij}/(\nu-h)]C_{fl}, \quad \text{(A4n)}$$

$$[C_{il},T_{km}] = [D_{il},T_{km}] = 0, \quad \text{(A4o)}$$

$$[C_{il},\Pi_{km}] = \delta_{lm}C_{ik} - (\delta_{km}/h)C_{il}, \quad \text{(A4p)}$$

$$[D_{il},\Pi_{km}] = \delta_{lm}D_{ik} - (\delta_{km}/h)D_{il}. \quad \text{(A4q)}$$

We see that the $(\nu-h)^2-1$ generators $\{A_{ij},B_{ij}\}$ form a $\mathrm{su}(\nu-h)$ Lie algebra, while the (h^2-1) generators $\{\Pi_{kl}\}$ form a $\mathrm{sl}(h,\mathbb{R})$ Lie algebra and altogether $\{A_{ij},B_{ij},\Pi_{kl}\}$ form the semisimple Lie algebra $\mathrm{su}(\nu-h)\oplus\mathrm{sl}(h,\mathbb{R})$. The remaining generators $\{T_{kl},C_{il},D_{il}\}$ give rise to an Abelian solvable Lie algebra, whose structure is given just by Eq. (4.2). Now using Theorem 3.3 of Ref. 5 we obtain the assertion of Theorem 4.1. ▼

APPENDIX B: PROOF OF THEOREM 4.2

The set of generators, complementary in $\mathrm{so}(2\nu-h,h)$ to the generators giving rise to the stability subalgebra h of ψ_{m_+}, is given by the generators of Eqs. (A3a)÷(A3g) of Appendix A.

Since

$$[\tilde{A}_{ij},\tilde{B}_{ij}] = B_{ii} + B_{jj} - 2B_{(\nu-h),(\nu-h)} + 4\tilde{B}, \quad i\neq j = 1,\ldots,\nu-h, \quad \text{(B1)}$$

we see that for

$$0\leqslant h\leqslant \nu-2 \quad \text{(B2)}$$

the complementary set of generators (A3a)÷(A3g) does *not* form an algebra.

Let us explore the remaining cases $h=\nu$ and $h=\nu-1$.

In the so-called *neutral case* $h=\nu$, the complementary set is given by the $\binom{\nu}{2}$ generators $\tilde{Q}_{kl}$ of Eq. (A3c) and by the generator $\tilde{D}$ of Eq. (A3g). Their commutation relations are given by

$$[\tilde{Q}_{kl},\tilde{Q}_{mn}] = 0, \quad \text{(B3a)}$$

$$[\tilde{D},\tilde{Q}_{kl}] = \delta_{k\nu}\tilde{Q}_{l\nu} - \delta_{l\nu}\tilde{Q}_{k\nu}, \quad \text{(B3b)}$$

with $k,l,m,n = 1,\ldots,\nu$.

We see that in this case the complementary set forms a solvable Lie algebra with the structure given by

$$c = t^{(2)}\oplus t^1, \quad \text{(B4)}$$

where $t^{(2)}$ is an Abelian algebra of dimension $\binom{\nu}{2}$.

In the so-called *conformal case* $h=\nu-1$, instead, the complementary set is given by the

$\binom{\nu-1}{2}$ generators $\tilde{Q}_{kl}$ of Eq. (A3c),

$\nu-1$ generators $\tilde{C}_{1l}$ of Eq. (A3d),

$\nu-1$ generators $\tilde{D}_{1l}$ of Eq. (A3e), (B5)

1 generator $\tilde{B}$ of Eq. (A3f),

1 generator $\tilde{D}$ of Eq. (A3g).

Their commutation relations are given by

$$[\tilde{D},\tilde{Q}_{kl}] = \delta_{k\nu}\tilde{Q}_{l\nu} - \delta_{l\nu}\tilde{Q}_{k\nu}, \quad \text{(B6a)}$$

$$[\tilde{C}_{1k},\tilde{C}_{1l}] = \tfrac{1}{2}\tilde{Q}_{kl}, \quad \text{(B6b)}$$

$$[\tilde{D}_{1k},\tilde{D}_{1l}] = \tfrac{1}{2}\tilde{Q}_{kl}, \quad \text{(B6c)}$$

$$[\tilde{B},\tilde{C}_{1l}] = \tilde{D}_{1l}, \quad \text{(B6d)}$$

$$[\tilde{B},\tilde{D}_{1l}] = -\tilde{C}_{1l}, \quad \text{(B6e)}$$

$$[\tilde{D},\tilde{C}_{1l}] = -\delta_{l\nu}\tilde{C}_{1\nu}, \quad \text{(B6f)}$$

$$[\tilde{D},\tilde{D}_{1l}] = -\delta_{l\nu}\tilde{D}_{1\nu}, \quad \text{(B6g)}$$

$$[\tilde{Q}_{kl},\tilde{Q}_{mn}] = [\tilde{C}_{1k},\tilde{D}_{1l}] = [\tilde{B},\tilde{Q}_{kl}] = [\tilde{B},\tilde{D}] = [\tilde{Q}_{kl},\tilde{C}_{1m}] = [\tilde{Q}_{kl},\tilde{D}_{1m}] = 0, \quad \text{(B6h)}$$

with $k,l,m,n = 2,\ldots,\nu$.

We see that the complementary set forms a solvable Lie algebra r whose structure is that one given by Eq. (4.10). Repeating the argument given at the end of the proof of Theorem 2.3 we obtain the assertion of Theorem 4.2. ▼

APPENDIX C: PROOF OF THEOREM 5.1

Let us take as a basis for the Clifford algebra $\mathbb{R}_{2\nu+1-h,h}$ the $2\nu+1$ $(2^\nu\times 2^\nu)$ matrices Γ_n, $n=1,\ldots,2\nu+1$, satisfying the anticommutation relations (1.25), given explicitly by Eqs. (A1) plus the new relation

$$\Gamma_{2\nu+1} = H_0 = \gamma_{2\nu+1}. \quad \text{(C1)}$$

It is easy to see that the generators X_{ab} $(a,b = 1,\ldots,2\nu+1)$ defined by Eq. (1.26) in terms of the Γ_a's of Eqs. (A1a)÷(A1d) and (C1) satisfy the commutation relations (1.27) of the $\mathrm{so}(2\nu+1-h,h)$ Lie algebra.

Let us now look for the stability subalgebra h of the highest weight spinor ψ_m.

From Eqs. (3.6) we see that for the $\mathrm{so}(2\nu+1-h,h)$ Lie algebra, in addition to the generators (A2a)÷(A2f) and (A3a)÷(A3g) of the $\mathrm{so}(2\nu-h,h)$ Lie algebra, we find further *real* linearly independent combinations of X_{ab} generators, giving

(a) h matrices E_k

$$= -\tfrac{1}{2}[H_k,H_0] = X_{2k-1,2\nu+1} + X_{2k,2\nu+1}, \quad \text{(C2)}$$

having all elements zero in the first column, and

(b) $(\nu-h)$ matrices $\tilde{H}_i$

$$= -\tfrac{1}{4}\{[H_i,H_0] + [H_{i'},H_0]\} = X_{2i-1,2\nu+1}, \quad \text{(C3a)}$$

$(\nu-h)$ matrices $\tilde{K}_i$

$$= (i/4)\{[H_i,H_0] - [H_{i'},H_0]\} = X_{2i,2\nu+1}, \quad \text{(C3b)}$$

h matrices $\tilde{F}_k$

$$= -\tfrac{1}{2}[H_{k'},H_0] = X_{2k-1,2\nu+1} - X_{2k,2\nu+1}, \quad \text{(C3c)}$$

having some non-null element in the first column.

We can easily check that the set given by the $\nu^2 - 2 + h(h+1)/2$ matrices (A2a)÷(A2f) and (C2) form the stability subalgebra h we are looking for. In addition to the commutation relations (A4a)÷(A4q) we have in fact

$$[E_k, E_l] = 2T_{kl}, \tag{C4a}$$

$$[E_k, \Pi_{lm}] = \delta_{km}E_l - (\delta_{lm}/h)E_k, \tag{C4b}$$

$$[E_k, A_{ij}] = [E_k, B_{ij}] = [E_k, T_{lm}] = [E_k, C_{il}] = [E_k, D_{il}] = 0. \tag{C4c}$$

From Eqs. (A4a)÷(A4q) and (C4) we see that the structure of the stability subalgebra h is just as that one of the Lie algebra of (5.1) group. Now using Theorem 3.3 of Ref. 5 we obtain the assertion of Theorem 5.1. ▼

APPENDIX D: PROOF OF THEOREM 5.2

The set c of generators, complementary in $so(2\nu + 1 - h, h)$ to the set of generators giving the stability subalgebra h of ψ_m, is formed by the generators given by Eqs. (A3a)÷(A3g) and (C3).

Since Eq. (B1) still holds, we see that for

$$0 \leqslant h \leqslant \nu - 2, \tag{D1}$$

the complementary set c of generators (A3a)÷(A3g) and (C3) does *not* form an algebra.

Let us analyze the remaining $h = \nu$ and $h = \nu - 1$ cases. In the $h = \nu$ case the complementary set is given by the

$$\begin{pmatrix}\nu\\2\end{pmatrix} \text{ generators } \tilde{Q}_{kl} \text{ of Eq. (A3c)},$$

$$\nu \text{ generators } \tilde{F}_k \text{ of Eq. (C3c)}, \tag{D2}$$

$$1 \text{ generator } \tilde{D} \text{ of Eq. (A3g)},$$

with $k,l = 1, \ldots, \nu$.

Their commutation relations are given by Eqs. (B3) plus

$$[\tilde{F}_k, \tilde{F}_l] = 2\tilde{Q}_{kl}, \tag{D3a}$$

$$[\tilde{F}_k, \tilde{D}] = \delta_{k\nu}\tilde{F}_\nu, \tag{D3b}$$

$$[\tilde{F}_k, \tilde{Q}_{lm}] = 0, \tag{D3c}$$

with $k,l,m = 1, \ldots, \nu$.

We see that in this case the complementary set c forms a solvable Lie algebra with the structure given by

$$c = r^1 \oplus r, \tag{D4}$$

where r is a solvable Lie algebra having the structure described by Eq. (5.4).

In the so-called *conformal case* $h = \nu - 1$, instead, the complementary set c is given by the generators of Eq. (B5) plus

$$\nu - 1 \text{ generators } \tilde{F}_k \text{ of Eq. (C3c)},$$

$$1 \text{ generator } \tilde{H}_1 \text{ of Eq. (C3a)}, \tag{D5}$$

$$1 \text{ generator } \tilde{K}_1 \text{ of Eq. (C3b)},$$

with $k = 2, \ldots, \nu$.

Their commutation relations are given by Eqs.(B6a)÷(B6h) plus

$$[\tilde{H}_1, \tilde{K}_1] = \tilde{B}, \tag{D6a}$$

$$[\tilde{H}_1, \tilde{F}_k] = 2\tilde{C}_{1k}, \tag{D6b}$$

$$[\tilde{K}_1, \tilde{F}_k] = 2\tilde{D}_{1k}, \tag{D6c}$$

$$[\tilde{F}_k, \tilde{F}_l] = 2\tilde{Q}_{kl}, \tag{D6d}$$

$$[\tilde{B}, \tilde{H}_1] = \tilde{K}_1, \tag{D6e}$$

$$[\tilde{K}_1, \tilde{B}] = \tilde{H}_1, \tag{D6f}$$

$$[\tilde{H}_1, \tilde{C}_{1l}] = -\tfrac{1}{2}\tilde{F}_l, \tag{D6g}$$

$$[\tilde{K}_1, \tilde{D}_{1l}] = -\tfrac{1}{2}\tilde{F}_l, \tag{D6h}$$

$$[\tilde{F}_k, \tilde{D}] = \delta_{k\nu}\tilde{F}_\nu, \tag{D6i}$$

$$[\tilde{H}_1, \tilde{Q}_{kl}] = [\tilde{K}_1, \tilde{Q}_{kl}] = [\tilde{H}_1, \tilde{D}] = [\tilde{K}_1, \tilde{D}] = [\tilde{H}_1, \tilde{D}_{1l}] = [\tilde{K}_1, \tilde{C}_{1l}] = [\tilde{F}_k, \tilde{B}] = [\tilde{F}_k, \tilde{Q}_{mn}] = [\tilde{F}_k, \tilde{C}_{1l}] = [\tilde{F}_k, \tilde{D}_{1l}] = 0, \tag{D6j}$$

with $k,l = 2, \ldots, \nu$.

We see from Eqs. (B6a)÷(B6h) and (D6a)÷(D6j) that the complementary set c forms an algebra. Precisely the three generators $\{\tilde{H}_1, \tilde{K}_1, \tilde{B}\}$ satisfy the commutation relations of the simple compact Lie algebra so(3), while the remaining generators $\{\tilde{Q}_{kl}, \tilde{D}, \tilde{F}_k, \tilde{C}_{1l}, \tilde{D}_{1l}\}$ form a solvable subalgebra. Furthermore the so(3) generators produce an automorphism of the solvable subalgebra.

We see therefore that the complementary algebra has just the structure dictated by Eqs. (5.4) and (5.6). Repeating the argument given at the end of the proof of Theorem 2.3 we obtain the assertion of Theorem 5.2. ▼

[1] See, e.g., *Relativity, Groups and Topology*, edited by B. de Witt and R. Stora (North-Holland, Amsterdam, 1984), Vol. II.

[2] See, e.g., G. Moore and P. Nelson, "Aetiology of sigma model anomalies," preprint HUTP-84/A076, Harvard University, 1984.

[3] See, e.g., *Supersymmetry and Supergravity 1982, Proceedings of the School on Supergravity and Supersymmetry*, ICTP, Trieste, September 1982, edited by S. Ferrara, J. C. Taylor, and P. van Nieuwenhuizen (World Scientific, Singapore, 1983).

[4] P. Furlan and R. Rączka, Phys. Lett. B **152**, 75 (1985).

[5] A. O. Barut and R. Rączka, *Theory of Group Representations and Applications* (PWN, Warsaw, 1980).

[6] E. Cartan, *The Theory of Spinors* (Hermann, Paris, 1966).

[7] P. Budinich, "On the possible role of pure conformal spinors in physics," ISAS preprint 36/83/EP, Trieste.

[8] P. Furlan and R. Rączka, "New nonlinear spinor field theories," ISAS preprint 77/84/EP.

[9] See, e.g., M. K. Volkov and V. N. Pervushin, *Essentially Nonlinear Quantum Theories and Dynamical Symmetries in Meson Physics* (Atomizdat, Moscow, 1978) (in Russian).

[10] M. Flato, G. Pinczon, and J. Simon, Ann. Sci. Ecole Norm. Sup. (9) **10**, 405 (1977).

[11] Ian R. Porteous, *Topological Geometry* (Cambridge U. P., Cambridge, 1981).

[12] I. M. Benn and R. W. Tucker, "Pure spinors and real Clifford algebras," preprint University of Lancaster, 1985.

[13] C. Chevalley, *The Algebraic Theory of Spinors* (Columbia U. P., New York, 1954).

[14] E. Fermi and C. N. Yang, Phys. Rev. **76**, 1739 (1949); S. Sakata, Prog. Theor. Phys. **16**, 686 (1956).

[15] See, e.g., L. D. Faddeev and A. A. Slavnov, *Gauge Fields: Introduction to Quantum Theory* (Benjamin/Cummins, Reading, MA, 1980).

On the projective representations of finite Abelian groups. II

M. Saeed-ul-Islam[a)]
International Centre for Theoretical Physics, Trieste, Italy

(Received 5 June 1984; accepted for publication 31 August 1984)

Complete sets of inequivalent irreducible projective representations of $C_m^n = \{w_1,\cdots,w_n; w_i^m = 1, i = 1,...,n; w_iw_j = w_jw_i, i,j = 1,...,n\}$ with respect to a class of factor sets α are determined, where $\alpha(w_i,w_j) = \theta\alpha(w_j,w_i)$, $1 \leqslant i < j \leqslant n$ and θ is a fixed mth root of unity. A single irreducible projective representation of C_m^n for each factor set α is constructed and called the basic projective representation. The rest of the representations are obtained by tensoring the basic projective representations with some ordinary representations of C_m^n. Projective representations of C_m^n are thus parametrized in terms of its ordinary representations.

I. INTRODUCTION

Applications of representation theory of finite Abelian groups in different fields of physics such as solid state physics is well known. More recently, they have been found useful in statistical mechanics (see Rittenberg[1]). Also, projective representations of Abelian groups arise naturally in the study of energy bands in the presence of a magnetic field (see Brown[2]).

II. THEORY

In this paper, we consider C_m^n, the direct product of n copies of a cyclic group C_m of order m and determine its inequivalent irreducible projective representations with respect to a particular class of factor sets.

C_m^n is an Abelian group of order m^n given by

$$C_m^n = \{w_1,...,w_n : w_i^m = 1, \quad i = 1,...,n;\ w_iw_j = w_jw_i,\ i,j = 1,...,n\}.$$

Let α be a factor set of C_m^n (see Morris[3] for definitions and other properties of factor sets and projective representations). The factor set α may be chosen (up to equivalence) in such a way that $\alpha'(w_i,w_j) = \alpha(w_i,w_j)\alpha(w_j,w_i)^{-1}$ is an mth root of unity.

Morris[4] has determined α-regular classes and inequivalent irreducible projective representations (ipr's) of C_m^n with respect to the factor sets in two special cases, when $\alpha'(w_i,w_j) = \alpha(w_i,w_j)\alpha(w_j,w_i)^{-1} = \theta\,(1 \leqslant i < j \leqslant n)$ and either (i) θ is a primitive mth root of unity, or (ii) 2 divides m and θ is a primitive square root of unity.

We consider a more general case when θ is primitive k th root of unity where k is any divisor of m and obtain complete sets of inequivalent ipr's of C_m^n with respect to these factor sets. The results of Morris[4] may be obtained as a particular case by taking $k = 2$ and $k = m$, respectively. The α-regular classes of C_m^n with respect to these factor sets were considered by Saeed-ul-Islam.[5]

If T is a projective representation of C_m^n with factor set α as above over the field of complex numbers and if $T_i = T(w_i)$, $i = 1,...,n$, then it is easy to see that $T_1,...,T_n$ satisfy the following equations:

$$T_i^m = I, \quad i = 1,...,n; \quad T_iT_j = \theta T_jT_i, \quad 1 \leqslant i < j \leqslant n. \tag{1}$$

Conversely, if $T_1,...,T_n$ are any given matrices satisfying Eq. (1) then they generate a projective representation T of C_m^n with factor set α and are given by

$$T(w_1^{a_1}\cdots w_n^{a_n}) = T_1^{a_1}\cdots T_n^{a_n},$$

for all $a_i \in \{0,...,m-1\}$, $i = 1,...,n$.

The following theorem which gives the number and degrees of the ipr's of C_m^n is proved in Ref. 5.

Theorem 1: Let α be a factor set of C_m^n over C satisfying $\alpha'(w_i,w_j) = \theta\,(1 \leqslant i < j \leqslant n)$, where θ is a primitive k th root of unity, k divides m. Let $m = kl$.

Then we have the following.

(i) (n even) C_m^n has l^n number of α-regular elements given by $w_1^{a_1}\cdots w_n^{a_n}$: $a_i \equiv 0 \pmod k$, $i = 1,...,n$ and therefore l^n number of inequivalent irreducible projective representations of C_m^n each of degree $(k)^{1/2n}$.

(ii) (n odd) C_m^n has kl^n number of α-regular elements of the form $w_1^{a_1}\cdots w_n^{a_n}$: $a_1 \equiv -a_2 \equiv a_3 \equiv \cdots \equiv a_n \pmod k$ and therefore has kl^n number of inequivalent irreducible projective representations of degree $k^{1/2(n-1)}$.

We first construct a set of $k \times k$ matrices, which are used in the construction of the required irreducible projective representations.

If k is odd, let P and Q be the $k \times k$ matrices defined by

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$Q = \begin{pmatrix} 0 & \theta & 0 & 0 & \cdots & 0 \\ 0 & 0 & \theta^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \theta^3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \theta^{k-1} \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

If k is even, let P be defined as above and

[a] Permanent address: Department of Mathematics, Bahauddin Zakariya University, Multan, Pakistan.

 0022-2488/85/123033-03$02.50

$$Q=\begin{pmatrix} 0 & \zeta & 0 & 0 & \cdots & 0 \\ 0 & 0 & \zeta^3 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \zeta^5 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \zeta^{2k-3} \\ \zeta^{2k-1} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

where ζ is the $2k$ th primitive root of unity such that $\zeta^2=\theta$. Then in both cases, it can be readily verified that (see Morris[6])

$$P^k=Q^k=I, \quad PQ=\theta QP.$$

where I is the identity matrix of order $k\times k$. Further, let

$$R=\begin{cases} P^{k-1}Q, & \text{if } k \text{ is odd,} \\ \zeta P^{k-1}Q, & \text{if } k \text{ is even.} \end{cases}$$

It follows easily that in either case,

$$R^k=I, \quad PR=\theta RP, \quad QR=\theta RQ.$$

Now let

$$E_{2i-1}=R\otimes R\otimes\cdots\otimes R\otimes P\otimes I\otimes\cdots\otimes I,$$

$$E_{2i}=R\otimes R\otimes\cdots\otimes R\otimes Q\otimes I\otimes\cdots\otimes I,$$

$i=1,...,\nu=[n/2]$. If n is odd let

$$E_n=R\otimes\cdots\otimes R\otimes R\otimes R\otimes\cdots\otimes R.$$

It follows from the properties of P, Q, and R that $E_1,...,E_n$ satisfy Eq. (1) and therefore generate a projective representation T of C_m^n with factor set α. We call T the basic projective representation of C_m^n with factor set α.

It is further verified that

(i) $E_1^r E_2^{-r} E_3^r\cdots E_{2\nu-1}^r E_{2\nu}^{-r}=\theta^{[-r(r+1)/2]\nu} I$, $r=0,1,...,k-1$.

(ii) No other product of matrices $E_{j_1}^{\mu_{j_1}}\cdots E_{j_t}^{\mu_{j_t}}=\lambda I$, for any nonzero complex number λ, except a reordering of (i) or when $\mu_{j_i}\equiv 0 \pmod k$ in which case $\lambda=1$.

(iii) $E_{j_1}^{\mu_{j_1}}\cdots E_{j_t}^{\mu_{j_t}}$ has nonzero trace if and only if $E_{j_1}^{\mu_{j_1}}\cdots E_{j_t}^{\mu_{j_t}}=\lambda I$ for some nonzero complex number λ.

The following result now follows easily.

Lemma: If χ denotes the projective character of the basic projective representation T of C_m^n as defined above then

(i) n even,

$$\chi(w_1^{a_1}\cdots w_n^{a_n})=(k)^{1/2n},$$

where $a_i\equiv 0 \pmod k$, $i=1,...,n$,

(ii) n odd,

$$\chi(w_{w_1}^{a_1}\cdots w_n^{a_n})=\theta^{-[r(r+1)/2]\nu}(k)^{(1/2)(n-1)},$$

where

$$a_1\equiv -a_2\equiv a_3\equiv\cdots\equiv -a_{2\nu}\equiv a_{2\nu+1}\equiv r \pmod k,$$

$r=0,1,....,k-1$. $\chi(w)=0$ if w is not of the above two types.

If n is even, the number of α-regular classes of C_m^n is equal to l^n (see Saeed-ul-Islam[5]) and χ has value $(k)^{(1/2)n}$ on each of these elements. Therefore,

$$\langle\chi,\chi\rangle=\frac{1}{m^n}\sum_{w\in C_m^n}\chi(w)\bar{\chi}(w)=\frac{1}{m^n}l^n.k^{(1/2)n}.k^{(1/2)n}=1,$$

and hence χ is irreducible in this case. Similarly it can be shown that χ is irreducible if n is odd.

Definition: Let ξ be a primitive mth root of unity such that $\xi^l=\theta$. Define $\theta_{(b_1,...,b_n)}:C_m^n\to C^*$ by $\theta_{(b_1,...,b_n)}(w_1^{a_1}\cdots w_n^{a_n})=\xi^{b_1a_1+\cdots+b_na_n}$, for all $b_i,a_i\in\{0,1,...,m-1\}$, $i=1,...,n$. Then $\theta_{(a_1,...,a_n)}$ is an ordinary irreducible representation of C_m^n and by giving different values to b_i, $i=1,...,n$, we get a complete set of inequivalent ordinary irreducible representations of C_m^n.

We consider a subcollection of these representations given by

$$\{\theta_{(b_1,...,b_n)}:b_i\in\{0,1,...,l-1\},\ i=1,...,n\}$$

and define

$$T_{(b_1,...,b_n)}=\theta_{(b_1,...,b_n)}T,$$

for all $b_i\in\{0,1,...,l-1, i=1,...,n\}$. Clearly, $T_{(b_1,...,b_n)}$ are irreducible projective representations of C_m^n with factor set equal to the factor set of T. If we denote the character of $T_{(b_1,...,b_n)}$ by $\chi_{(b_1,...,b_n)}$ then $\chi_{(b_1,...,b_n)}(w)=\theta_{(b_1,...,b_n)}(w)\cdot\chi(w)$ for all $w\in C_m^n$. Now assume that $\chi_{(b_1,...,b_n)}=\chi_{(b'_1,...,b'_n)}$, where $b_i,b'_1\in\{0,1,...,l-1\}$, $i=1,...,n$. Then we have, in particular,

$$\chi_{(b_1,...,b_n)}(w_i^k)=\chi_{(b'_1,...,b'_n)}(w_i^k), \quad i=1,...,n,$$

i.e.,

$$\theta_{(b_1,...,b_n)}(w_i^k)\chi(w_i^k)\theta_{(b'_1,...,b'_n)}(w_i^k)\chi(w_i^k),$$

i.e.,

$$\theta_{(b_1,...,b_n)}(w_i^k)=\theta_{(b'_1,...,b'_n)}(\omega_i^k), \quad i=1,...,n, \quad \chi(w_i^k)\neq 0.$$

i.e.,

$$\xi^{kb_i}=\xi^{kb'_i},$$

which implies that $k(b_i-b'_1)\equiv 0 \pmod m$, i.e., $b_i-b'_i\equiv 0 \pmod l$. Since $b_i,b'_i<l$, therefore $b_i=b'_i$, $i=1,...,n$. Thus $T_{(b_1,...,b_n)}:b_i\in\{0,1,...,l-1\}$, $i=1,...,n$ are all inequivalent.

If n is odd, define

$$T^{(j)}_{(b_1,...,b_n)}(w)=\xi^{j(\Sigma a_i)}T_{(b_1,...,b_n)}(w),$$

for all $b_1,...,b_n\in\{0,1,...,l-1\}$, $i=1,...,n, j=0,l,...,(k-1)l$, $w=w_i^{a_1}\cdots w_n^{a_n}\in C_m^n$ and also let $\chi^{(j)}_{(b_1,...,b_n)}$ denote the projective character of $T^{(j)}_{(b_1,...,b_n)}$. Considering $w=w_1w_2^{-1}w_3,...,w_{2\nu}^{-1}w_{2\nu+1}$, we get

$$\chi^{(j)}_{(b_1,...,b_n)}(w)=\chi^{(j')}_{(b_1,...,b_n)}(w)$$

$$\Rightarrow\xi^{j(\Sigma a_i)}\chi_{(b_1,...,b_n)}(w)=\xi^{j'(\Sigma a_i)}\chi_{(b_1,...,b_n)}(w)$$

$$\Rightarrow(j-j')\equiv 0 \pmod m.$$

$$\Rightarrow j\equiv j \pmod m \Rightarrow j=j', \text{ because } j,j'\leqslant m-1.$$

Thus if n is odd then $T^{(j)}_{(b_1,...,b_n)}$ are inequivalent for distinct values of j. It is now clear that if n is even, the number of $T_{(b_1,...,b_n)}$ is l^n and if n is odd, the number of $T^{(j)}_{(b_1,...,b_n)}$'s is equal to kl^n which are equal to the number of α-regular classes in the respective cases.

We summarize the above results in the following.

Theorem 2: A complete set of inequivalent ipr's of C_m^n with factor set α is given by (i) n even,

$$\{T_{(b_1,\ldots,b_n)}: b_i \in \{0,1,\ldots,l-1\}, i=1,\ldots,n\},$$

(ii) n odd,

$$\{T^{(j)}_{(b_1,\ldots,b_n)}: b_i \in \{0,1,\ldots,l-1\}, i=1,\ldots,n, \quad j \in \{0,l,\ldots,(k-1)l\}\}.$$

Each of the above representation is of degree k^ν, $\nu = [(n/2)]$.

Corollary 1: (Theorem 1, Ref. 4) If $l=1$, i.e., θ is primitive mth root of unity then (i) n even, C^n_m has only one ipr of degree m, and (ii) n odd, C^n_m has n inequivalent ipr's each of degree n^ν.

Corollary 2: (Theorem 2, Ref. 4) If $\theta = -1$, i.e., $l = m/2$, $k=2$, then (i) n even, C^n_m has $(m/2)^n$ inequivalent ipr's each of degree 2^ν, and (ii) n odd, C^n_m has $2\,(m/2)^n$ inequivalent ipr's each of degree 2^ν.

Remarks: (i) The above lemma and the theorem give an alternative proof of the results proved in Ref. 5.

(ii) The problem of constructing ipr's in the case when $\alpha'(w_i,w_j)$ does not take equal values for different pairs of i,j's will be considered in a subsequent paper. The case $n=3$ has been discussed by Backhouse and Bradley.[7]

ACKNOWLEDGMENTS

The author would 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 finalized. He would like to thank the Government of Pakistan for the award of the travel grant.

He would like to thank Professor Vidossich and Professor Yacoub (visiting professor at I.C.T.P.) for their continuous encouragement and also his colleague Ali Shoukat (visiting scientist) for very useful discussions.

[1] V. Rittenberg, Physica A **114**, 245 (1982).

[2] E. Brown, Solid State Phys. Adv. Research Appl. **22**, 313 (1968).

[3] A. O. Morris, "Projective representations of finite groups," in *"Proceedings of the Conference on Clifford Algebra, its Generalizations and applications,"* Matscience, Madras, 1971 (Institute of Mathematical Science, Madras, 1972), pp. 43–86.

[4] A. O. Morris, J. London. Math. Soc. **7**, 235 (1973).

[5] M. Saeed-ul-Islam, "On the projective representations of finite abelian groups. I," J. Pure Appl. Sci., Fac. Sci., Islamia Univ., Bahawalpur, Pakistan (to appear).

[6] A. O. Morris, Quart. J. Math. Oxford, Ser. (2) **18**, 7 (1967).

[7] N. B. Backhouse and C. J. Bradley, Quart, J. Math. Oxford Ser. (2) **21**, 203 (1970).

The group of gauge transformations as a Schwartz–Lie group

Renzo Cirelli and Alessandro Manià
Dipartimento di Fisica dell'Università, Milano, Italy and Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy

(Received 4 October 1984; accepted for publication 11 June 1985)

The group of gauge transformations of a smooth principal bundle $P(M,G)$ over a not necessarily compact manifold M and with a not necessarily compact structure group G is proved to be a Schwartz–Lie group. Its Lie algebra and exponential map are discussed.

I. INTRODUCTION

In recent years physicists have paid growing attention to "infinite continuous" groups (as they sometimes are called) such as the current group, the group of gauge transformations, the group of volume preserving diffeomorphisms of an oriented manifold X, and the group of symplectic transformations of a symplectic manifold X. Obviously enough, this does not stem from the fact that physicists only recently have recognized the importance of this kind of group for fundamental physics,[1] but from the fact that mathematicians have made progress in endowing them with "smoothness structures," which allow us to treat them to some extent as the usual Lie groups. Frequently these structures are relativized from similar ones for $\mathrm{Map}(X,Y)$, the space of maps from the manifold X to the manifold Y, an object of fundamental relevance in global analysis.[2] However, in the treatment of $\mathrm{Map}(X,Y)$ the compactness of the manifold X, at least, is currently assumed. In application to physics this assumption may be well suited for specific problems but, generally speaking, is clearly reductive. This is particularly evident for the "continuous infinite" groups quoted above since the manifold X is the space-time or the phase space. Many physical objects live in space-time: eventually depending on the mathematical category they are thought to belong to, for some of them to be bounded to live in a compactified space-time does not result in an essential modification of their properties, but for others it results in an artificial enlargement or reduction of their properties. As things are we cannot help but remember the ancient mythological story of Procustes' bed.[3]

Furthermore, often Sobolev space techniques are used jointly with the compactness assumption; these techniques offer the well-known advantages of the Hilbertian structures, but must be paid for with hard limitations and technical complications. And then, the whole setting becomes very unnatural if one tries to remove the compactness assumption.

For these reasons we consider the recent proposal of Michor[4,5] very interesting. He endows $C^\infty(X,Y)$ (the set of C^∞ maps from X to Y) with a natural differentiable manifold structure working in a very natural and simple setting and without assuming compactness for X. Michor's treatment gives $C^\infty(X,Y)$ a topology that is finer than the widely used Whitney C^∞ topology,[6] avoids projective limits[7,8] and Γ-differentiability,[9] and, last, but not least, is meaningful from the physical point of view.

As applications of his treatment, Michor himself has discussed some "infinite continuous" groups of relevant interest for physics, for instance, the group $\mathrm{Diff}X$ and the group of canonical transformations.[4(c),4(d),5] In this paper we want to apply Michor's treatment to the current group G^X and to the group $\mathscr{G}$ of gauge transformations. We will show that these groups can be given a differentiable structure by which they become Schwartz–Lie groups with associated Schwartz–Lie algebras; we will also prove the existence of a nice exponential map for these groups.

In Sec. II we introduce, with a short and simple exposition, Michor's method to give $C^\infty(X,Y)$ a differentiable structure. In Sec. III we study the current group and the group of gauge transformations, viewing them as groups of sections of smooth group fiber bundles. In Sec. IV we discuss the associated Lie algebra and give the exponential map.

II. $C^\infty(X,Y)$ AS A SCHWARTZ MANIFOLD

In this section we endow the set $C^\infty(X,Y)$ of a smooth differentiable structure assuming that X and Y are *ordinary smooth* manifolds, that is Hausdorff, second-countable, and locally compact C^∞ manifolds (hence finite dimensional, paracompact, and metrizable). We follow a procedure that is clearly equivalent to the Michor procedure,[4,5] but, in our opinion, more suitable for physicists' taste. According to this attitude we start with the discussion of the locally convex space, which will play the role of local model of the manifold.

Let $\zeta=(E,M,\pi;\mathbb{R}^m)$ be an ordinary smooth vector bundle over M. We denote by $\mathrm{Sec}_0\,\zeta$ the real linear space of smooth sections with compact support of ζ. We want to endow $\mathrm{Sec}_0\,\zeta$ of a locally convex topology suitable from the point of view of differential calculus.

Given an open subset U of $\mathbb{R}^n$ we denote by $\mathscr{D}(U,\mathbb{R}^m)$ the real linear space of smooth $\mathbb{R}^m$-valued maps on U with compact support [if $m=1$ we shortly write, as usual, $\mathscr{D}(U)$ instead of $\mathscr{D}(U,\mathbb{R})$]. As is well known,[10] in this space a locally convex topology, usually called the Schwartz topology, is induced by the family of seminorms

$$\{q_\vartheta^U\}\quad(\vartheta\in\Theta),$$

where Θ is the collection of all the families $\vartheta\equiv\{\vartheta_p\}$ $(p\in\mathbb{N}^n)$ of continuous real functions ϑ_p on U such that the family $\{\mathrm{supp}\,\vartheta_p\}$ $(p\in\mathbb{N}^n)$ is locally finite and

$$q_\vartheta^U(\sigma):=\max_{p\in\mathbb{N}^n}\max_{x\in U}||\vartheta_p(x)D^p\sigma(x)||_{\mathbb{R}^n},\quad\sigma\in\mathscr{D}(U,\mathbb{R}^m).$$

It is well known from the usual distribution theory that

 0022-2488/85/123036-06$02.50

$\mathscr{D}(U,\mathbb{R}^m)$, endowed by the Schwartz topology induced by the above family of seminorms, is a complete locally convex nuclear space, strict inductive limit of a countable family of separable Fréchet spaces (even dually nuclear and Lindelöf, hence paracompact and normal). Locally convex spaces of this kind are called (NLF)-*spaces* by Michor[4(d)]; we will call them simply *Schwartz spaces*.

If X is an ordinary smooth manifold we can analogously consider the linear space $\mathscr{D}(X,\mathbb{R}^m)$. With a nearly obvious generalization also this space can be made a Schwartz space. Actually, we can take a countable and locally finite atlas

$$\mathscr{A} = \{(U_k,\varphi_k,\mathbb{R}^n)\} \quad (k\in\mathbb{N})$$

for X and introduce the family of seminorms

$$\{q_{k,\vartheta}\} \quad (k\in\mathbb{N}, \quad \vartheta\in\Theta),$$

where Θ is the collection of all the families $\vartheta\equiv\{\vartheta_p\}$ $(p\in\mathbb{N}^n)$ of continuous real functions ϑ_p on X such that the family $\{\operatorname{supp}\vartheta_p\}$ $(p\in\mathbb{N}^n)$ is locally finite and

$$q_{k,\vartheta}(\sigma) = q_{\vartheta_k}^{\varphi_k(U_k)}(\sigma_k), \quad \sigma\in\mathscr{D}(X,\mathbb{R}^m),$$

where σ_k is the local expression of σ and ϑ_k the "local expression" of ϑ.

Now suppose $\zeta = (E,M,\pi;\mathbb{R}^m)$ is an ordinary vector bundle over M. A family of seminorms

$$\{q_{k,\vartheta}\} \quad (k\in\mathbb{N}, \quad \vartheta\in\Theta)$$

can be introduced in the linear space $\operatorname{Sec}_0\zeta$ in an analogous way considering now a countable *locally finite fibered atlas* for ζ, that is, a pair $(\mathscr{A},\zeta\mathscr{A})$, where

$$\mathscr{A}\equiv\{(U_k,\varphi_k,\mathbb{R}^n)\} \quad (k\in\mathbb{N})$$

is an atlas for M, and

$$\zeta\mathscr{A} = \{(\pi^{-1}(U_k),\ \mathbb{R}^m+\mathbb{R}^n,\ (\varphi_k\times\mathrm{id}_{\mathbb{R}^m})\circ\psi_k)\} \quad (k\in\mathbb{N})$$

is an atlas for E, such that (U_k,ψ_k) $(k\in\mathbb{N})$ is locally finite locally trivializing system for ζ.

A simple way to see that $\operatorname{Sec}_0\zeta$ endowed with the topology induced by this family of seminorms is a Schwartz space is the following.

First suppose that ζ is trivial over M, that is, $\zeta = (M\times\mathbb{R}^m, M, pr_M)$; then $\operatorname{Sec}_0\zeta\cong\mathscr{D}(M,\mathbb{R}^m)$. If ζ is not trivial over M, by a very well-known structural theorem[11] there exists a vector bundle $\eta = (E',M,\pi';\mathbb{R}^r)$ over M such that the Whitney sum $\zeta\oplus\eta = (E\oplus E',M,\pi\oplus\pi';\mathbb{R}^{m+r})$ is trivial over M, that is $E\oplus E' = M\times\mathbb{R}^{m+r}$. Therefore $\operatorname{Sec}_0(\zeta\oplus\eta)$ can be given the structure of a Schwartz space as above. Moreover, the decomposition $\operatorname{Sec}_0(\zeta\oplus\eta) = \operatorname{Sec}_0\zeta \oplus \operatorname{Sec}_0\eta$ is also topological, that is, $\operatorname{Sec}_0\zeta$ is a topological direct summand of $\operatorname{Sec}_0(\zeta\oplus\eta)$. Hence $\operatorname{Sec}_0\zeta$ inherits from $\operatorname{Sec}_0(\zeta\oplus\eta)$ the structure of Schwartz space together with the properties given above.[12]

The nice topological properties of the Schwartz spaces are very suitable also from the point of view of differential calculus since nearly all the good definitions of smoothness are equivalent.[13] For the sake of convenience we choose the C_c^∞ smoothness, which can be defined as follows: Let V and W be locally convex vector spaces and U an open subset of V. A map $f:U\to W$ is said to be C_c^1 on U if for every $x\in U$ there exists a linear operator $Df(x):V\to W$ such that

(1) $\lim_{t\to 0}\{[f(x+ty)-f(x)]/t\} = Df(x)y$,

(2) the map $Df:\ U\times V\to W$,
$$(x,y)\rightsquigarrow Df(x,y) := Df(x)y$$
is continuous.

A map $f:U\to W$ is said to be C_c^{k+1} on U $(k\geqslant 1)$ if f is C_c^k on U and D^kf, where $D^kf := D(D^{k-1}f)$, is C_c^1 on U. A map $f:U\to W$ is said to be C_c^∞ on U if it is C_c^k on U for every $k = 1,2\ldots$.

By a *Schwartz manifold* we will mean a manifold in the C_c^∞ sense modeled on a Schwartz space.

The locally convex space candidate for the role of local model at f of the manifold $C^\infty(X,Y)$ is $\operatorname{Sec}_0\zeta_f$, where ζ_f is the pullback $f^*(T_Y)$ of the tangent bundle $T_Y = (TY,Y,p_Y)$ by $f\in C^\infty(X,Y)$. Obviously $\operatorname{Sec}_0\zeta_f$ will be understood as a Schwartz space, as discussed above, endowed with the just defined C_c^∞ smoothness.

An important tool for the definition of charts on $C^\infty(X,Y)$ is the notion of local addition on an ordinary manifold.

Given an ordinary smooth manifold M, let $T_M = (TM,M,p_M)$ be its tangent bundle. A *local addition* on M is a mapping $\tau:TM\to M$ satisfying (1) $(p_M,\tau):TM\to M\times M$ is a diffeomorphism with an open neighborhood of the diagonal in $M\times M$, and (2) $\tau(0_x) = x$, $\forall x\in M$, where 0_x is the zero of T_xM.

The most important property of a local addition τ on M is that, for every $x\in M$, its restriction $\tau_x:T_xM\to M$ is a diffeomorphism with an open neighborhood of x in M. Every ordinary smooth manifold M admits local additions (see Sec. 10.1 of Ref. 5).

Now we are able to introduce an atlas for $C^\infty(X,Y)$. We choose a local addition τ on Y and define, for every fixed $f\in C^\infty(X,Y)$,

$$U_f^\tau := \{g\in C^\infty(X,Y)\,|\,g\sim f,\ g(x)\in\tau_{f(x)}(T_{f(x)}Y),\ \forall x\in X\},$$

where $g\sim f$ means that the set $\{x\in X\,|\,g(x)\neq f(x)\}$ is relatively compact, and

$$\varphi_f^\tau:\ U_f^\tau\to\operatorname{Sec}_0\zeta,$$
$$g\rightsquigarrow\varphi_f^\tau(g), \quad (\varphi_f^\tau(g))(x) := \tau_{f(x)}^{-1}g(x).$$

It is easy to see that φ_f^τ is a bijection; its inverse is

$$\psi_f^\tau:\ \operatorname{Sec}_0\zeta\to U_f,$$
$$s\rightsquigarrow\psi_f^\tau(s) := \tau\circ s.$$

Therefore $(U_f^\tau,\varphi_f^\tau,\operatorname{Sec}_0\zeta_f)$ is a chart at f on $C^\infty(X,Y)$.

Now we can show that

$$\mathscr{A}^\tau := \{(U_f^\tau,\varphi_f^\tau,\operatorname{Sec}_0\zeta_f)\} \quad [f\in C^\infty(X,Y)]$$

is a smooth atlas for $C^\infty(X,Y)$. To do this we must prove that whenever $U_f\cap U_g\neq\varnothing$,

(a) $\varphi_f^\tau(U_f\cap U_g)$ is open in $\operatorname{Sec}_0\zeta$,

(b) $\varphi_g^\tau\cdot X\circ(\varphi_f^\tau)^{-1}:\varphi_f^\tau(U_f\cap U_g)\to\varphi_g^\tau(U_f\cap U_g)$

is a diffeomorphism.

To prove (a) we just recall that the convergence of a sequence $\{s_n\}$ in $\operatorname{Sec}_0\zeta_f$ to $s\in\operatorname{Sec}_0\zeta_f$ implies that there exists a compact subset K of X such that all but a finite number of the s_n's equal s off K and $s_n(x)\to s(x)$ uniformly on K; by a

standard argument this implies that the complement of $\varphi_f^{-1}(U_f \cap U_g)$ in $\mathrm{Sec}_0\,\zeta_f$ is closed.

The proof of (b) is a simple application of the Ω-lemma 8.7 of Michor.[5]

So we can conclude that, for every local addition τ on Y, $\mathscr{A}^\tau$ is a smooth atlas for $C^\infty(X,Y)$.

Again by a simple application of the Ω-lemma it can be shown that the atlases $\mathscr{A}^\tau$ and $\mathscr{A}^{\tau'}$ are equivalent for every local addition τ and τ' on Y. Therefore the C_c^∞-manifold structure introduced in $C^\infty(X,Y)$ does not depend on the choice of the local addition on Y.

We emphasize that, as shown by Michor, $C^\infty(X,Y)$ endowed with this manifold structure is a Hausdorff, paracompact, and normal manifold and admits C_c^∞ partitions of unity. An important point to remark about the tangent space $TC^\infty(X,Y)$ is that, for every $f \in C^\infty(X,Y)$, $T_f C^\infty(X,Y)$ is canonically isomorphic to the local model $\mathrm{Sec}_0\,\zeta_f$ as is shown in Theorem 10.13 of Ref. 5. The map, which identifies $T_f C^\infty(X,Y)$ with $\mathrm{Sec}_0\,\zeta_f$, can be worked out from Lemma 10.15 of Michor[5] and is given by

$$T_f C^\infty(X,Y) \ni [c] \rightsquigarrow \mathrm{ev}[c] \in \mathrm{Sec}_0\,\zeta_f,$$

where $[c]$ is the equivalence class of paths in $C^\infty(X,Y)$ through f, that is $c(0) = f$, and

$$\mathrm{ev}[c](x) := [\mathrm{ev}_x \circ c], \quad \forall x \in X,$$

with $\mathrm{ev}_x : C^\infty(X,Y) \to Y$ the evaluation at x.

In the following we will be particularly interested in $\mathrm{Sec}\,\zeta$, the set of all smooth sections of a given (ordinary) fiber bundle $\zeta = (E,M,\pi)$. It is shown in Michor,[5] Proposition 10.10, that $\mathrm{Sec}\,\zeta$ is a splitting submanifold of $C^\infty(M,E)$ whose local model at s is $\mathrm{Sec}_0(s^*(\mathrm{Ver}(W_s)))$, where $p_s : W \to s(M)$ is a tubular neighborhood of $s(M)$ in E such that $p_s = s \circ \pi \restriction W_s$ and Ver is the functor which to every smooth fiber bundle associates its vertical bundle. We can substitute $s^*(\mathrm{Ver}(E))$ for $s^*(\mathrm{Ver}(W_s))$ if we take into account that

$$T_{s(x)} W_s = T_{s(x)} E, \quad \forall x \in M,$$

obviously and that

$$T_{s(x)} W_s = T_{s(x)}(W_s)_{s(x)} \oplus (T_x s)(T_x M), \quad \forall x \in M,$$

since s is transversal to the fibers of $p_s : W_s \to s(M)$. In fact from these two relations we have

$$\mathrm{Ver}_{s(x)} W_s = \mathrm{Ver}_{s(x)} E, \quad \forall x \in M.$$

Therefore the local model at s for $\mathrm{Sec}\,\zeta$ can be identified with $\mathrm{Sec}_0(s^*(\mathrm{Ver}\,E))$. Since $\mathrm{Sec}\,\zeta$ is a splitting submanifold of $C^\infty(M,E)$, we also obtain that the map ev introduced above gives the canonical identification

$$T_s\,\mathrm{Sec}\,\zeta \simeq \mathrm{Sec}_0(s^*(\mathrm{Ver}\,E)).$$

III. SCHWARTZ–LIE GROUP OF SECTIONS OF A GROUP FIBER BUNDLE

In this section we use the treatment of Sec. II to endow the current group and the group of gauge transformations with the structure of the Schwartz–Lie group.

The (smooth) current group G^M is the group of smooth maps $f : M \to G$ from an ordinary smooth manifold M to an ordinary Lie group G (see Ref. 14). From Sec. II we have immediately that $G^M \equiv C^\infty(M,G)$ is a Schwartz manifold. We have to show that the group structure and the manifold structure are compatible so that we are allowed to say that G^M is a Schwartz–Lie group (that is, loosely speaking, a smooth group modeled on a Schwartz space).

Note that G^M can be identified with the group of (smooth) sections of the *trivial* principal bundle $(M \times G, M, \mathrm{pr}_M; G)$.

The group $\mathscr{G}$ of gauge transformations of a principal bundle $P(M,G)$ over M with structure group G is the set $\mathrm{Sec}\,P[G]$ of (smooth) sections of $P[G]$ with the pointwise defined composition low. Here $P[G]$ means the fiber bundle associated to $P(M,G)$ with typical fiber G and action of G on G given by the adjoint action $b \rightsquigarrow aba^{-1}$. It is well known that $\mathscr{G}$ can be identified with the group of those diffeomorphisms f of the total space P of $P(M,G)$ such that

$$f \circ \pi = \pi, \quad f(pa) = f(p)a, \quad \forall p \in P, \quad \forall a \in G,$$

where $\pi : P \to M$ is the projection of $P(M,G)$ (see Ref. 15).

Again from Sec. II we know that $\mathrm{Sec}\,P[G]$ is a Schwartz manifold and again we are faced with the problem of the compatibility of the two structures.

We can treat the problem in a more general setting that covers both cases. We consider a smooth *group* fiber bundle $\gamma \equiv (E,M,\pi;G)$ with the ordinary Lie group G as typical fiber and make $\mathrm{Sec}\,\gamma$ into a group with pointwise defined composition law. We emphasize that we only assume that M and G are *ordinary* manifolds; in particular we do not assume that M and (or) G are compact as is usually assumed in this context.

Theorem 3.1: Let $\gamma = (E,M,\pi;G)$ be a smooth *group* fiber bundle with the ordinary Lie group G as typical fiber. Then $\mathrm{Sec}\,\gamma$, with the pointwise defined composition law, is a Schwartz–Lie group.

Proof: To prove the smoothness of the composition law

$$\Theta : \mathrm{Sec}\,\gamma \times \mathrm{Sec}\,\gamma \to \mathrm{Sec}\,\gamma$$
$$(s,s') \rightsquigarrow \Theta(s,s') = ss',$$

we first introduce the pullback of the Cartesian product $\gamma \times \gamma$ by the diagonal map $\Delta : M \to M \times M$ and we denote it by $\gamma \times_M \gamma$. The fiber bundle $\gamma \times_M \gamma$ is called the *fibered product* of γ by itself and its total space may canonically be identified with $E \times_M E := \{(u,u') E \times E \mid \pi(u) = \pi(u')\}$ and its fiber over $x \in M$ with the Cartesian product $E_x \times E_x$. Then we define

$$\vartheta : E \times_M E \to E,$$
$$(x,(u,u')) \rightsquigarrow \vartheta(x,(u,u')) := (x,uu'),$$

that is, ϑ is a map which on each fiber $E_x \times E_x$ is the group composition for E_x. Using local triviality it is easy to see that ϑ is a smooth map.

The map Θ is the composition of the canonical identification of $\mathrm{Sec}\,\gamma \times \mathrm{Sec}\,\gamma$ with $\mathrm{Sec}(\gamma \times_M \gamma)$ and the map

$$\mathrm{Comp}_\vartheta : \mathrm{Sec}(\gamma \times_M \gamma) \to \mathrm{Sec}\,\gamma,$$
$$(s,s') \rightsquigarrow \vartheta \circ (s,s'), \quad \vartheta \circ (s,s')(x) = s(x)s'(x).$$

Now, by Ref. 5, Propositions 10.5 and 10.10, the canonical identification is a C_c^∞ map and by Corollary 10.14.1 of Ref. 5, Comp_ϑ too is C_c^∞. Therefore Θ is a C_c^∞ map.

To prove the smoothness of the inverse map Inv on

Sec γ we proceed in an analogous way. We introduce the map

$$\text{inv}: E \to E,$$
$$(x,u) \mapsto \text{inv}(x,u) = (x,u^{-1}),$$

which, using local triviality, is immediately seen to be smooth. Since Inv is the composition with inv, again by Corollary 10.14.1 of Ref. 5 we obtain that Inv is a C_c^∞ map. ///

With essentially similar arguments we can prove the following theorem.

Theorem 3.2: Let $\lambda \equiv (E,M,\pi;L)$ be a smooth *Lie algebra* fiber bundle with the ordinary Lie algebra L as typical fiber. Then $\text{Sec}_0 \lambda$, with pointwise defined composition laws, is a Schwartz–Lie algebra.

It must be remarked that the requirement of compactness for the supports of the sections of λ cannot be removed if we want to obtain a topological vector space; in fact, $\text{Sec}\,\lambda$ is an Abelian Schwartz–Lie group, but the scalar multiplication is not continuous at 0; actually, for $\sigma \in \text{Sec}\,\lambda$, $(1/n)\sigma \to 0$ iff $\sigma \in \text{Sec}_0 \lambda$, so that the open subset $\text{Sec}_0 \lambda$ is clearly the maximal subset of $\text{Sec}\,\lambda$, which is a topological vector space.

Theorem 3.2 suggests a setting for a natural realization of the Lie algebra of the group Sec γ, which is very useful and suitable in applications as, for instance, the possibility of defining an exponential map will show.

We discuss these matters in the following section.

IV. THE LIE ALGEBRA OF THE GROUP Sec γ

Given a C_c^∞ manifold $\mathscr{M}$ modeled on locally convex spaces E_α, let $\mathscr{X}(\mathscr{M})$ be the $\mathscr{S}(\mathscr{M})$ module of the vector fields on $\mathscr{M}$ [$\mathscr{S}(\mathscr{M})$ is the ring of the real-valued C_c^∞ functions on $\mathscr{M}$]. We introduce in this module a Lie bracket in the following way: If A and B are vector fields on $\mathscr{M}$ with local expressions A_α and B_α the Lie bracket $[A,B]$ is the vector field on $\mathscr{M}$ whose local expression is $-A'_\alpha(v)B_\alpha(v) + B'_\alpha(v)A_\alpha(v)$, where $A'_\alpha(v)$ and $B'_\alpha(v)$ are the derivatives of A_α and B_α at $v \in E_\alpha$.

To check that the so-defined $[A,B]$ is actually a vector field on $\mathscr{M}$ is simply a matter of checking transition relations for vector fields.

Analogously we can define the Lie derivative on $\mathscr{S}(\mathscr{M})$: If f is a real-valued C_c^∞ function with local expression f_α and A a vector field on $\mathscr{M}$ with local expression A_α, then the Lie derivative $\mathscr{L}_A f$ is the C_c^∞ function whose local expression is $f'_\alpha(v)A_\alpha(v)$.

It is immediate that the just-defined Lie bracket and Lie derivative satisfy the familiar relations

$$[A,B+C] = [A,B] + [A,C],$$
$$[fA,B] = f[A,B] - (\mathscr{L}_B f)A,$$
$$[A,B] = -[B,A],$$
$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0,$$
$$\mathscr{L}_A(f+g) = \mathscr{L}_A f + \mathscr{L}_A g,$$
$$\mathscr{L}_A(fg) = (\mathscr{L}_A f)g + f(\mathscr{L}_A g),$$
$$\mathscr{L}_{A+B} f = \mathscr{L}_A f + \mathscr{L}_B f,$$
$$\mathscr{L}_{kA} f = k\mathscr{L}_A f, \quad k \in \mathbb{R},$$
$$\mathscr{L}_{[A,B]} = [\mathscr{L}_A, \mathscr{L}_B].$$

From these relations we see that the introduction of the Lie bracket $[\ ,\ ]$ endows $\mathscr{X}(\mathscr{M})$ with the structure of Lie module over $\mathscr{S}(\mathscr{M})$.

If the C_c^∞ manifold $\mathscr{M}$ is a C_c^∞ group $\mathscr{G}$, the invariant (under left translations) vector fields are of paramount importance; they constitute a *Lie algebra over* $\mathbb{R}$, since they are exactly those vector fields which are self-related by every (left) translation. This Lie algebra over $\mathbb{R}$ is called the Lie algebra of $\mathscr{G}$. We will denote it with $L_\mathscr{G}$. As in the case of ordinary Lie groups the map

$$\text{ev}_e: L_\mathscr{G} \to T_e\mathscr{G},$$
$$A \mapsto \text{ev}_e A := A(e),$$

where e is the unit element of $\mathscr{G}$, gives a natural isomorphism of $L_\mathscr{G}$ as linear space with $T_e\mathscr{G}$. This linear isomorphism becomes a Lie algebra isomorphism if we define on $T_e\mathscr{G}$ the Lie bracket

$$[A(e),B(e)] := [A,B](e).$$

As in the case of ordinary Lie groups, it can be easily seen that if $a(t)$ is a path on $\mathscr{G}$ through e belonging to $A(e)$ and $b(t)$ a path through e belonging to $B(e)$, the so-defined $[A(e),B(e)]$ is indeed the equivalence class of paths through e which the path

$$a(\tau)b(\tau)a^{-1}(\tau)b^{-1}(\tau), \quad \tau = (\text{sgn}\, t)\sqrt{|t|},$$

belongs to.

Let now $\mathscr{G}$ be the Schwartz–Lie group Sec γ introduced in Sec. III. The result remarked at the end of Sec. II, i.e., that the tangent space at f to $C^\infty(X,Y)$ can be canonically identified with the local model at f, enables us to give a simple realization of the Lie algebra $L_{\text{Sec}\,\gamma}$ of Sec γ.

Theorem 4.1: Let Sec γ be the Schwartz–Lie group of smooth sections of the smooth group fiber bundle $\gamma = (E,M,\pi;G)$ with the ordinary Lie group G as the typical fiber; let $\text{Sec}_0 \lambda$ be the Schwartz–Lie algebra of smooth sections with compact support of the smooth Lie algebra fiber bundle $\lambda = (F,M,\rho;L)$ with the ordinary Lie algebra L as the typical fiber. If λ is defined by $L = L_G$ and a local trivializing system derived in an obvious way by the one of γ, then the Lie algebra $L_{\text{Sec}\,\gamma}$ is canonically isomorphic with $\text{Sec}_0 \lambda$.

Proof: Since, for every $x \in M$,

$$(e^*(\text{Ver}\,E))_x \simeq (\text{Ver}\,E)_{e(x)} = T_{e(x)}E_{\pi(e(x))}$$
$$= T_{e(x)}E_x \equiv T_{e(x)}G_x \simeq L_{G_x},$$

by the hypothesis $L = L_G$ and local triviality the vector bundles $e^*(\text{Ver}\,E)$ and λ are naturally M isomorphic. Calling l this M isomorphism we must just show that the composition

$$l_* \circ \text{ev}: T_e\,\text{Sec}\,\gamma \to \text{Sec}_0(e^*(\text{Ver}\,E)) \to \text{Sec}_0\,\lambda$$

preserves the Lie bracket. Now, if $\dot{a}(0) \in T_e\,\text{Sec}\,\gamma$, $\dot{b}(0) \in T_e\,\text{Sec}\,\gamma$, and $\dot{c}(0) = [\dot{a}(0),\dot{b}(0)]$, we have

$$((l_* \circ \text{ev})\dot{c}(0))(x)$$
$$= (l_*(\text{ev}\,\dot{c}(0)))(x) = (l \circ \text{ev}\dot{c}(0))(x)$$
$$= l((\text{ev}\dot{c}(0))(x)) = l((\text{ev}_x \circ c)^{\cdot}(0));$$

but

$$(\mathrm{ev}_x \circ c)^{\cdot}(0) = (\mathrm{ev}_x \circ a(\tau) b(\tau) a^{-1}(\tau) b^{-1}(\tau))^{\cdot}(0)$$
$$= ((\mathrm{ev}_x \circ a)(\tau)(\mathrm{ev}_x \circ b)(\tau)(\mathrm{ev}_x \circ a^{-1})(\tau)$$
$$\times (\mathrm{ev}_x \circ b^{-1})(\tau))^{\cdot}(0)$$
$$= [(\mathrm{ev}_x \circ a)^{\cdot}(0), (\mathrm{ev}_x \circ b)^{\cdot}(0)],$$

and therefore

$$((l_* \circ \mathrm{ev})\dot{c}(0))(x) = l([(\mathrm{ev}_x \circ a)^{\cdot}(0), (\mathrm{ev}_x \circ b)^{\cdot}(0)])$$
$$= [l((\mathrm{ev}_x \circ a)^{\cdot}(0)), l((\mathrm{ev}_x \circ b)^{\cdot}(0))]$$
$$= [(l \circ \mathrm{ev}\, \dot{a}(0))(x), (l \circ \mathrm{ev}\, \dot{b}(0))(x)]$$
$$= [(l_*(\mathrm{ev}\, \dot{a}(0))(x), (l_*(\mathrm{ev}\, \dot{b}(0)))(x)]$$
$$= [l_*(\mathrm{ev}\, \dot{a}(0)), l_*(\mathrm{ev}\, \dot{b}(0))](x),$$

that is,

$$(l_* \circ \mathrm{ev})[\dot{a}(0), \dot{b}(0)] = [(l_* \circ \mathrm{ev})\dot{a}(0), (l_* \circ \mathrm{ev})\dot{b}(0)]. \quad ///$$

Coming to the exponential map, one must expect that the possibility of defining a nice exponential map for Schwartz–Lie groups in general will meet with serious difficulties. For instance, for the group $\mathrm{Diff}(M)$ there exists an exponential map whose image generates a dense subgroup of the connected component of the identity,[4(d)] but it is well known that, even in the case of compact M, this exponential map, in general, cannot be subjective on any open neighborhood of the identity. For our Schwartz–Lie groups, Theorem 4.1 enables us to introduce a nice exponential map. In fact, identifying by Theorem 4.1 the Lie algebra of $\mathrm{Sec}\,\gamma$ with $\mathrm{Sec}_0\,\lambda$ we can define

$$\mathrm{Exp}: L_{\mathrm{Sec}\,\gamma} \cong \mathrm{Sec}_0\,\lambda \to \mathrm{Sec}\,\gamma,$$
$$\sigma \rightsquigarrow \mathrm{Exp}(\sigma),$$

where

$$(\mathrm{Exp}\,\sigma)(x) := \exp(\sigma(x))$$

and

$$\exp: F \to E$$

is defined by

$$\exp(v) = \exp_{\rho(v)}(v), \quad v \in F.$$

By local triviality the map exp is smooth and therefore $\mathrm{Exp} = \mathrm{Comp}_{\mathrm{exp}}$ is smooth.

It can be seen very easily that Exp has the familiar properties. The only not evident property is that it is a local diffeomorphism, but this can be established directly in the following way.

We know that, for every $x \in M$, there exists an open neighborhood V_{0_x} in L_{G_x} such that on it $\exp_x$ is a diffeomorphism with an open neighborhood U_{e_x} of e_x in G_x. By the local triviality it is immediately seen that the union of all the U_{e_x} is an open neighborhood of the image of the unit section e. Therefore

$$U_e := \{s \in \mathrm{Sec}\,\gamma \mid s \sim e,\ s(x) \in U_{e(x)},\ \forall x \in M\}$$

is an open neighborhood of e in $\mathrm{Sec}\,\gamma$ and we can define

$$\mathrm{Log}: U_e \to \mathrm{Sec}_0\,\lambda$$
$$s \rightsquigarrow \mathrm{Log}\, s := \log \circ s,$$

where log is defined in an obvious way. Since $\mathrm{Log} = \mathrm{Comp}_{\log}$ we obtain that Log is a smooth map and since

$$(\mathrm{Exp} \circ \mathrm{Log})(s) = \exp \circ \log \circ s = s, \quad s \in U_e,$$

and similarly for $\mathrm{Log} \circ \mathrm{Exp}$, we can conclude that Exp is a local diffeomorphism.

V. CONCLUDING REMARKS

The group of gauge transformations has been studied by Mitter and Viallet[16] (see also Refs. 17–19 and 15) using Sobolev space techniques and assuming the base space M of the principal bundle to be compact. As we have said in the Introduction, in Michor's method Sobolev space techniques are not expected. Though to not dispose of Hilbert space methods may at present cause some trouble (mainly in connection with the lack of an inverse function theorem) we can hope that in the near future hard implicit function theorems in the setting of Schwartz spaces become available. For instance, a workable inverse function theorem on Fréchet spaces admitting smoothing operators, which applies very well to $C^\infty(X, Y)$, even if only when X is compact, is now available.[20]

ACKNOWLEDGMENTS

We gratefully thank P. W. Michor for having drawn our attention to Ref. 20. We wish to thank G. F. Dell'Antonio for stimulating conversations, and M. C. Abbati and P. Cotta-Ramusino for fruitful discussions and useful comments.

[1] W. Pauli, starting a series of lectures on continuous groups in quantum mechanics to the members of the CERN Theoretical Study Division at Copenhagen in September 1955, said, "These lectures will deal with the class of 'finite continuous' groups only; I wish, however, to emphasize that according to my opinion the class of 'infinite continuous' groups will presumably turn out to be of greater importance for physics in future" (mimeographed notes).

[2] J. Eells, "A setting for global analysis," Bull. Am. Math. Soc. **72**, 751 (1966); R. Palais, *Foundations of Global Non-linear Analysis* (Benjamin, New York, 1968); J. Marsden, *Application of Global Analysis in Mathematical Physics* (Publish or Perish, Boston, 1974).

[3] Plutarch, *Parallel lives: Theseus and Romulus* (c.45-125 A.C.).

[4] P. W. Michor, (a) "Manifolds of smooth maps," Cahiers Top. Geom. Diff. **XIX-1**, 47 (1978); (b) "Manifolds of smooth maps II," Cahiers Top. Geom. Diff. **XXI-1**, 63 (1980); (c) "Manifolds of smooth maps III," Cahiers Top. Geom. Diff. **XXI-3**, 325 (1980); (d) "Manifolds of smooth maps IV," Cahiers Top. Geom. Diff. **XXIV-1**, 57 (1984).

[5] P. W. Michor, "Manifolds of differentiable mappings," in *Shiva Mathematics Series*, Vol. 3 (Shiva, Orpington, England, 1980).

[6] M. Golubitsky and V. Guillemin, "Stable mappings and their singularities," in *Springer Graduate Texts in Mathematics*, Vol. 14 (Springer, Berlin, 1973).

[7] D. G. Ebin and J. Marsden, "Groups of diffeomorphisms and the motion of an incompressible fluid," Ann. Math. **92**, 109 (1970).

[8] H. Omori, "Infinite dimensional Lie transformation groups," in *Springer Lecture Notes in Mathematics*, Vol. 427 (Springer, Berlin, 1974).

[9] S. Yamamuro, "Differential calculus in topological linear spaces," in *Springer Lecture Notes in Mathematics*, Vol. 374 (Springer, Berlin, 1974).

[10] J. Horvath, *Topological Vector Spaces and Distributions I* (Addison–Wesley, Reading, MA, 1966).

[11] W. Greub, S. Halperin, and R. Vanstone, *Connections, Curvature and Cohomology. I, II, III* (Academic, New York 1972, 1973, 1976).

[12] H. H. Schäfer, "Topological vector spaces," *Springer Graduate Texts in Mathematics*, Vol. 3 (Springer, Berlin, 1970).

[13] H. H. Keller, "Differential calculus in locally convex spaces," *Springer Lecture Notes in Mathematics*, Vol. 417 (Springer, Berlin, 1974).

[14]H. Araki, "Factorizable representation of current algebra," Publ. RIMS, Kyoto Univ. **5**, 361 (1969–70).
[15]W. Kondracki and J. Rogulski, "On the stratification of the orbit space for the action of automorphisms on connections," Institute of Mathematics, Polish Academy of Sciences Preprint 281, 1983.
[16]P. K. Mitter and C. M. Viallet, "On the bundle of connections and the gauge orbit manifold in the Yang–Mills theory," Commun. Math. Phys. **79**, 457 (1981).
[17]I. M. Singer, "Some remarks on Gribov ambiguity," Commun. Math. Phys. **60**, 7 (1978).
[18]P. K. Mitter, "Geometry of gauge orbits and the Yang–Mills dynamical system," *Lectures at the 1979 Cargèse Summer Institute on "Recent Developments in Gauge Theories"* (Plenum, New York, 1980).
[19]M. Asorey and P. K. Mitter, "On geometry, topology and ϑ-sectors in regularized quantum Yang–Mills theory," Ref. TH. 3424-CERN LPTHE 82/13, 1982.
[20]R. S. Hamilton, "The inverse function theorem of Nash and Moser," Bull. Am. Math. Soc. **1**, 65 (1982).

Lie symmetries of some equations of the Fokker–Planck type

C. C. A. Sastri and K. A. Dunn
Department of Mathematics, Statistics, and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H8

(Received 18 August 1982; accepted for publication 10 July 1985)

The structure of the local Lie groups of symmetries of some partial differential equations of the Fokker–Planck type in one space dimension is investigated. A connection between these groups and the group $SL_2(\mathbb{R})$ is established in the sense that they are all shown to be locally isomorphic to $SL_2(\mathbb{R})A$, where A is the radical. It is conjectured that the groups of Lie symmetries of all Fokker–Planck equations in one space dimension have this structure. The notion of partial invariance, due to Ovsiannikov, is applied to the equations studied. It appears plausible that the class of partially invariant solutions of these equations is larger than the class of invariant solutions although no explicit demonstration of this claim is available at present.

I. INTRODUCTION

The Fokker–Planck equation first arose in kinetic theory,[1,2] where it describes the evolution of the one-particle distribution function of a dilute gas with long-range collisions, such as a Coulomb gas. It can, for instance, be derived[3] from the Boltzmann equation in the limit of large impact parameters. Besides kinetic theory, it occurs in a variety of areas[4–10] such as engineering and biology. In probabilistic literature, it is also called the Kolmogorov forward equation,[11] and describes the evolution of the transition probability density for a diffusion process. In the case of one space variable, to which we shall restrict ourselves here for the sake of simplicity, it can be written in the form

$$\frac{\partial u}{\partial t} = a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial u}{\partial x} + cu, \tag{1}$$

where u is the unknown function, x and t are the space and time coordinates respectively, and a, b, and c are smooth functions of x and t, assumed to be given. We shall further suppose that the processes are homogeneous, which means that a, b, and c depend only on x. While much of what will be said here applies to Eq. (1), three special cases, studied earlier by Bluman and Cole[12] and by Nariboli,[13] among others, will be discussed in some detail: (i) $a = 1$, $b = c = 0$,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}; \tag{2}$$

(ii) $a = \beta x$, $b = 2\beta - \alpha x$, $c = -\alpha$ (α, β constants),

$$\frac{\partial u}{\partial t} = \beta\frac{\partial^2}{\partial x^2}(xu) - \alpha\frac{\partial}{\partial x}(xu); \tag{3}$$

and (iii) $a = \frac{1}{4}x^{1-2p}$, $b = (1-4p)x^{-2p}/4$, $c = p^2x^{-(2p+1)}$ (p a constant such that $2n \equiv 2p + 1 \neq 0$),

$$\frac{\partial u}{\partial t} = \frac{1}{4}\frac{\partial}{\partial x}\left(x^{1-2p}\frac{\partial u}{\partial x} - 2px^{-2p}u\right). \tag{4}$$

Equation (2) describes Brownian motion without drift and Eq. (3) was first introduced by Feller[6] in the study of a problem in population genetics. Equation (4) arises in plasma physics.[9,13] Equation (2) has the same structure as the well-known heat equation, except that in that case u stands for the temperature of a solid. From now on, we shall refer to Eqs. (2), (3), and (4) as the heat equation, the genetics equation, and the plasma physics equation, respectively.

Bluman and Cole[12,14] used the method of Lie symmetries to find the invariant solutions (also called similarity solutions) of the heat equation and Bluman[15] did the same thing for a special case of Eq. (1). Nariboli[13] considered several special cases of Eq. (1) including the genetics and plasma physics equations and found the similarity solutions. Moreover, Bluman[16] has shown that every one-dimensional Fokker–Plank equation with a six-dimensional group of Lie symmetries can be transformed into the heat equation and vice versa. In other words, all Fokker–Planck equations with a six-dimensional Lie group of symmetries form an equivalence class of which the heat equation is the "canonical" member. In view of this result of Bluman's, three of the five special cases of the Fokker–Planck equation considered by Nariboli[13]—the well-known Ornstein–Uhlenbeck equation is one of them—which turn out to have six-dimensional groups of Lie symmetries, belong to the same equivalence class as the heat equation. The remaining two examples—the genetics equation and the plasma physics equation—each have a four-dimensional group of Lie symmetries, which have the same structure. Two natural questions arise at this point. (1) Is there an analog of Bluman's result for Fokker–Planck equations in one space dimension which have groups of Lie symmetries of dimension other than 6, and if so, what is the "canonical" equation for each dimension? (2) What happens to equations in higher dimensions? On the basis of our experience with several cases, it appears very likely that the answer to the first part of the first question is affirmative. Work is still in progress on both questions, and we shall return to them in a later publication. For the present, we shall confine ourselves to identifying the structure of the groups of the equations considered and making some remarks about partially invariant solutions for them.

The heat equation in particular and, in general, second-order linear partial differential equations in two variables have been considered by Ovsiannikov,[17,18] who has revived the group-theoretical study of differential equations in the past two decades and who, specifically, introduced the notion of partial invariance. However, to our knowledge, neither Ovsiannikov nor any of the other authors mentioned

 0022-2488/85/123042-06$02.50

so far has identified the structure of the groups the way we do, i.e., give the Levi decomposition. The remarkable fact is that in each case, the semisimple part is isomorphic to $SL_2(\mathbf{R})$.

The analysis is carried out mainly, but not entirely, by using the methods expounded by Ovsiannikov.[17,18] In doing so, we observe that the basic difference between the method used by Ovsiannikov and that used in Refs. 12–15 (it is the same method in all cases) is the following: Bluman, Cole, and Nariboli consider Eq. (1) (or special cases of it) and look for the infinitesimal generator of the group of transformations acting on the (t, x, u) space which are such that if $(t, x, u) \to (t', x', u')$, then u' satisfies Eq. (1) in t' and x'. Ovsiannikov's method does the same thing, except that instead of working with Eq. (1) he would work with the equivalent first-order system of equations

$$v = u_x, \quad u_t = av_x + bv + cu. \tag{5}$$

So now one looks for the group acting on the (t, x, u, v) space which leaves (5) invariant. Since Eq. (1) and the system (5) are equivalent, one would expect the groups to be the same, and indeed they are. The point, however, is that if one is interested in just the invariant solutions one need consider only the group acting on the (t, x, u) space, while if one is interested in invariant as well as partially invariant solutions, one must consider the group acting on the (t, x, u, v) space. The presence of the extra variable v, the "superfluous" variable in Ovsiannikov's terminology, turns out to be essential.

II. LIE SYMMETRIES

Since the technique of Lie symmetries is well described in the literature,[14,17,18,19] we shall skip the details of the construction of the infinitesimal generators of the groups. Our notation and terminology are those of Refs. 17 and 18 with very slight modifications which are self-explanatory.

The infinitesimal generator of the system (5) is of the form

$$X = \xi \frac{\partial}{\partial t} + \eta \frac{\partial}{\partial x} + \sigma \frac{\partial}{\partial u} + \tau \frac{\partial}{\partial v}, \tag{6}$$

where ξ, η, σ, and τ are all functions of t, x, u, and v, which can be determined using the techniques described in Refs. 17 and 18. It follows from the discussion there that for a linear system such as (5), ξ is a function of t alone, and η a function of t and x only. Moreover, $\sigma = fu + g$ and $\tau = u(\partial f/\partial x) + (\partial g/\partial x) + v(f - (\partial \eta/\partial x))$, where f and g are functions of t and x alone such that

$$\frac{\partial f}{\partial t} - c\frac{d\xi}{dt} - \eta\frac{dc}{dt} - a\frac{\partial^2 f}{\partial x^2} - b\frac{\partial f}{\partial x} = 0, \tag{7}$$

and g satisfies the system (5). Thus the problem of finding the full group of Lie symmetries of (5) involves finding the general solution of (5), which is an impossible task. Like previous authors, we simplify the problem by setting $g = 0$.

A. Heat equation

We get

$$\xi = a_1 + a_2 t + a_3 t^2, \tag{8}$$

$$\eta = a_4 + a_5 t + (x/2)(a_2 + 2a_3 t), \tag{9}$$

$$\sigma = [a_6 - (a_3 t/2) - (a_5 x/2) - (a_3 x^2/4)]u, \tag{10}$$

and

$$\tau = -(u/2)(a_2 + a_3 x) + v[a_6 - (a_2/4) - (3a_3 t/2) - (a_5 x/2) - (a_3 x^2/4)], \tag{11}$$

where a_1, a_2, a_3, a_4, a_5, and a_6 are arbitrary constants. Let $\mathcal{G}$ denote the Lie algebra and G the group. A basis $\{X_i\}$, $1 \le i \le 6$ for $\mathcal{G}$ can be obtained by setting $a_i = 1$ for a fixed i and $a_j = 0$ for all $j \ne i$. One obtains

$$X_1 = \frac{\partial}{\partial t},$$

$$X_2 = t\frac{\partial}{\partial t} + \frac{x}{2}\frac{\partial}{\partial x} - \frac{v}{2}\frac{\partial}{\partial v},$$

$$X_3 = t^2\frac{\partial}{\partial t} + tx\frac{\partial}{\partial x} - \left(\frac{t}{2} + \frac{x^2}{4}\right)u\frac{\partial}{\partial u} - \left(\frac{ux}{2} + \frac{3tv}{2} + \frac{x^2 v}{4}\right)\frac{\partial}{\partial v},$$

$$X_4 = \frac{\partial}{\partial x},$$

$$X_5 = t\frac{\partial}{\partial x} - \frac{xu}{2}\frac{\partial}{\partial u} - \left(\frac{u}{2} + \frac{xv}{2}\right)\frac{\partial}{\partial v},$$

and

$$X_6 = u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}.$$

The table of commutators is as follows:

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	0	X_1	$2X_2 - \frac{1}{2}X_6$	0	X_4	0
X_2		0	X_3	$-\frac{1}{2}X_4$	$\frac{1}{2}X_5$	0
X_3			0	$-X_5$	0	0
X_4				0	$-\frac{1}{2}X_6$	0
X_5					0	0
X_6						0.

Here, X_1 and X_4 correspond to translation in t and x, respectively, while X_6 describes stretching in u and v. The center $\mathcal{Z}$ of $\mathcal{G}$ is the span of X_6. The radical of $\mathcal{G}$ is $\mathcal{A}$, the span of X_4, X_5, and X_6. If $\mathcal{B}$ denotes the quotient algebra $\mathcal{G}/\mathcal{A}$, the Levi decomposition of $\mathcal{G}$ gives

$$\mathcal{G} \cong \mathcal{B} \oplus \mathcal{A}. \tag{12}$$

It can be shown that $\mathcal{B}$ is isomorphic to $s\ell_2(\mathbf{R})$, the algebra of real, 2×2 matrices of trace zero. Let $Z_1 = -X_1$, $Z_2 = 2X_2$, and $Z_3 = X_3$. Then $\{Z_1, Z_2, Z_3\}$ is a basis for $\mathcal{B}$, and

$$[Z_1, Z_2] = 2Z_1, \quad [Z_1, Z_3] = -Z_2, \quad [Z_2, Z_3] = 2Z_3.$$

On the other hand, if

$$W_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad W_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad W_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

then $\{W_1, W_2, W_3\}$ is a basis for $s\ell_2(\mathbf{R})$ and

$$[W_1, W_2] = 2W_1, \quad [W_1, W_3] = -W_2, \quad [W_2, W_3] = 2W_3.$$

Hence $\mathcal{B} \cong s\ell_2(\mathbf{R})$. So if A denotes the Lie group of $\mathcal{A}$, $G \cong SL_2(\mathbf{R})A$, where $SL_2(\mathbf{R})$ is the special linear group.

[Strictly speaking,[18] $\mathcal{G}$ is a quotient algebra of the full Lie algebra of the heat equation. Recall that we constructed $\mathcal{G}$ by setting $g=0$; if g is taken to be any solution of the system (5), the full Lie algebra is the span of $\{X_i\}$, $1\leqslant i\leqslant 6$, and $X_7=g(\partial/\partial u)+(\partial g/\partial x)(\partial/\partial v)$; X_7 corresponds to a translation in u by a solution of (2); and $\mathcal{G}$ is the quotient of the full algebra by the span of X_7, which is an infinite-dimensional ideal.]

B. Genetics equation

We obtain

$$\xi = a_1 e^{\alpha t} + a_2 e^{-\alpha t} + a_3, \tag{13a}$$

$$\eta = \alpha x(a_1 e^{\alpha t} - a_2 e^{-\alpha t}), \tag{13b}$$

$$\sigma = [a_4 - \alpha(a_1 e^{\alpha t} - a_2 e^{-\alpha t}) - (\alpha^2 x/\beta)a_2 e^{-\alpha t}]u, \tag{13c}$$

and

$$\tau = -(\alpha^2/\beta)a_2 e^{-\alpha t}u + [-2\alpha a_1 e^{\alpha t} + 2\alpha a_2 e^{-\alpha t} + a_4 - (\alpha^2 x/\beta)a_2 e^{-\alpha t}]v, \tag{14}$$

where a_1, a_2, a_3, and a_4 are arbitrary. Again, let $\mathcal{G}$ denote the Lie algebra and G the group. A basis $\{X_i\}$, $1\leqslant i\leqslant 4$, for $\mathcal{G}$ can be obtained as before:

$$X_1 = e^{\alpha t}\frac{\partial}{\partial t} + \alpha x e^{\alpha t}\frac{\partial}{\partial x} - \alpha u e^{\alpha t}\frac{\partial}{\partial u} - 2\alpha e^{\alpha t}\frac{\partial}{\partial v},$$

$$X_2 = e^{-\alpha t}\frac{\partial}{\partial t} - \alpha x e^{-\alpha t}\frac{\partial}{\partial x} + \left(\alpha e^{-\alpha t} - \frac{\alpha^2 x}{\beta}e^{-\alpha t}\right)u\frac{\partial}{\partial u} + \left[-\frac{\alpha^2}{\beta}e^{-\alpha t}u + \left(2\alpha e^{-\alpha t} - \frac{\alpha^2 x}{\beta}e^{-\alpha t}\right)v\right]\frac{\partial}{\partial v},$$

$$X_3 = \frac{\partial}{\partial t},$$

and

$$X_4 = u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}.$$

The table of commutators is as follows:

	X_1	X_2	X_3	X_4
X_1	0	$-2\alpha X_3$	$-\alpha X_1$	0
X_2		0	αX_2	0
X_3			0	0
X_4				0.

Here X_3 corresponds to translation in t while X_4 represents stretching in u and v. The center $\mathcal{Z}$ of G is the span of X_4, which is also the radical of $\mathcal{G}$. If $\mathcal{B}$ denotes the quotient algebra $\mathcal{G}/\mathcal{Z}$, the Levi decomposition gives $\mathcal{G}\cong\mathcal{B}\oplus\mathcal{Z}$. Again, $\mathcal{B}$ is isomorphic to $s\ell_2(\mathbb{R})$. Let $Z_1=(1/\alpha)X_1$, $Z_2=-(2/\alpha)X_3$, and $Z_3=-(1/\alpha)X_2$. (These transformations are valid only if $\alpha\neq 0$; but that is a reasonable assumption.[6]) Then $\{Z_1, Z_2, Z_3\}$ is a basis for $\mathcal{B}$, and $[Z_1, Z_2]=2Z_1$, $[Z_1, Z_3]=-Z_2$, and $[Z_2, Z_3]=2Z_3$. Thus $G\cong SL_2(\mathbb{R})Z$, where Z is the Lie group of $\mathcal{Z}$.

C. Plasma physics equation

We obtain

$$\xi = a_1 + a_2 t + a_3 t^2, \tag{15}$$

$$\eta = (a_3 t/n + a_2/2n)x, \tag{16}$$

$$\sigma = [a_4 + (n-1)a_3 t/n - (a_3 x^{2n}/n^2)]u, \tag{17}$$

and

$$\tau = -\frac{2a_3}{n}x^{2n-1}u + v\left[a_4 - \frac{a_2}{2n} + \frac{(n-2)}{n}a_3 t - \frac{a_3 x^{2n}}{n^2}\right], \tag{18}$$

where a_1, a_2, a_3, and a_4 are arbitrary constants. As before, one can obtain a basis $\{X_i\}$, $1\leqslant i\leqslant 4$, for the Lie algebra $\mathcal{G}$:

$$X_1 = \frac{\partial}{\partial t},$$

$$X_2 = t\frac{\partial}{\partial t} + \frac{x}{2n}\frac{\partial}{\partial x} - \frac{v}{2n}\frac{\partial}{\partial v},$$

$$X_3 = t^2\frac{\partial}{\partial t} + \frac{tx}{n}\frac{\partial}{\partial x} + \left[\frac{(n-1)}{n}t - \frac{x^{2n}}{n^2}\right]u\frac{\partial}{\partial u} - \left\{\frac{2x^{2n-1}}{n}u - v\left[\frac{(n-2)}{n}t - \frac{x^{2n}}{n^2}\right]\right\}\frac{\partial}{\partial v},$$

and

$$X_4 = u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}.$$

The table of commutators is as follows:

	X_1	X_2	X_3	X_4
X_1	0	X_1	$2X_2 + \frac{(n-1)}{n}X_4$	0
X_2		0	X_3	0
X_3			0	0
X_4				0

As in the case of the genetics equation, if $\mathcal{G}$ is the Lie algebra and $\mathcal{Z}$ the span of X_4, $\mathcal{Z}$ is the center of $\mathcal{G}$. If $\mathcal{B}$ denotes the quotient algebra $\mathcal{G}/\mathcal{Z}$, then $\mathcal{G}\cong\mathcal{B}\oplus\mathcal{Z}$. As before, $\mathcal{B}$ is isomorphic to $s\ell_2(\mathbb{R})$. Let $Z_1=X_1$, $Z_2=2X_2$, and $Z_3=-X_3$. Then $[Z_1, Z_2]=2Z_1$, $[Z_1, Z_3]=Z_2$, and $[Z_2, Z_3]=2Z_3$, as required.

Remark: As indicated earlier, the symmetry groups of Eq. (1) and the system (5) are the same. For instance, if the terms depending on τ and v are ignored, our X reduces to that of Refs. 12–15. The tables of commutators are the same and hence the groups are the same.

III. CONCLUSIONS

The structure of the groups of Lie symmetries of some Fokker–Planck equations has been studied. While the implications of the connection between Fokker–Planck equations and the group $SL_2(\mathbb{R})$ have yet to be investigated, it is clear that a Bluman-type result[16] would yield a classification of Fokker–Planck equations based on the dimension of the group of Lie symmetries. The examples presented here, as well as some work in progress, strongly indicate the existence of such a result. They also lead us to the conjecture that the group of Lie symmetries of every Fokker–Planck equation in one space dimension has the structure $SL_2(\mathbb{R})A$.

The examples considered here have been studied earlier by Bluman and Cole and by Nariboli, who found the infinitesimal generators and hence the similarity solutions. One might ask what the advantage is in studying the same equa-

tions by the Ovsiannikov method. The answer is that the presence of the extra variable v—the "superfluous variable" is Ovsiannikov's terminology[17]—enables one to look for partially invariant solutions,[17,18] while it is impossible to do so without it. To see this, some notation and terminology are needed, which will be developed here. The reader is referred to the work of Ovsiannikov[17,18] for a detailed discussion.

Let H be a subgroup of G, the group of Lie symmetries of the system (5); H acts on the (t, x, u, v) space. If primes denote transformed quantities (under the action of H), a function I of t, x, u, and v is said to be an invariant of H if $I' = I$, i.e., if $I(t', x', u', v') = I(t, x, u, v)$. A set of functionally independent invariants $I^1, I^2, \ldots, I^\ell$ of H is said to be complete if every invariant I of H can be expressed as a function of $I^1, I^2, \ldots, I^\ell$. A manifold in the (t, x, u, v) space is said to be invariant under H if it is invariant under each element of H. Let $Y = \binom{t}{x}$ and $U = \binom{u}{v}$. Then $U = \phi(Y)$, where ϕ is a vector function, is called an invariant solution (with respect to H) of the system (5) if the manifold $U = \phi(Y)$ is invariant under H. If, on the other hand, the manifold $U = \phi(Y)$ is contained in some manifold invariant under H, U is called a partially invariant solution (with respect to H) of (5). Clearly, invariant solutions are partially invariant, but the converse is not always true.

Let H be generated by the set of infinitesimal operators $\{X_{i_1}, X_{i_2}, \ldots, X_{i_K}\}$, $1 \leqslant K \leqslant r$, where r is the dimension of G. A complete set of invariants of H can be found by solving the first-order, linear partial differential equation $a^j X_j I = 0$, where the Einstein summation convention has been used and where $a^j \in \mathbb{R}$ $(1 \leqslant j \leqslant K)$. Now, in general, there may be more than one invariant manifold containing the manifold $U = \phi(Y)$. Consider the smallest such manifold. It can be described[17] by means of some equations, say q of them, involving $I^1, I^2, \ldots, I^\ell$:

$$\psi^i(I^1, I^2, \ldots, I^\ell) = 0 \quad (i = 1,2,\ldots,q). \tag{19}$$

Let s denote the rank of the matrix $(\partial I^\beta / \partial U^\gamma)$, where $1 \leqslant \beta \leqslant \ell$, $1 \leqslant \gamma \leqslant 2$, and $U = \binom{U^1}{U^2} = \binom{u}{v}$. Let $\rho = \ell - q$. Ovsiannikov calls ρ the rank of the solution $U = \phi(Y)$. A dimensionality argument shows[17,18] that a necessary condition for partially invariant solutions of rank ρ to exist is

$$\min\{\ell - 1, n - m\} \geqslant \rho \geqslant \ell - s,$$

where n is the dimension of the space on which H acts and m the number of dependent variables. In our case, $n = 4$ and $m = 2$. One other condition that must be satisfied for partially invariant solutions to exist is $q < m$. In Ref. 12–15, $m = 1$, so that this is impossible. Before proceeding to give an example, we observe that our computations show only that it is plausible that the special cases of the Fokker–Planck equation considered in this paper have partially invariant solutions that are not invariant. [One reason for seeking such solutions is that in many specific problems,[20] invariant solutions may not have any meaning (i.e., are not admissible) in the context of the problems; in such cases, one could try partially invariant solutions.] Regrettably, we have so far been unable to find an explicit case in which a solution is partially invariant under a subgroup H but not invariant under H or *any other* subgroup of G. Nevertheless, in the expectation that the clue to such a construction might lie in the plausibility argument, it is presented here. It should be observed, however, that our remark about the necessity of m being greater than 1 for partially invariant solutions to exist is valid, no matter whether the partial differential equation considered is a Fokker–Planck equation or some other equation.

Since the computations are lengthy, though straightforward, we shall consider only one example in detail. This concerns the heat equation.

Consider the subgroup generated by X_1 and X_6. To compare our work with that of Refs. 12–15, let us suppress the terms depending on τ or v, so that the group acts on the (t, x, u) space.

Then $X_1 = \partial/\partial t$ and $X_6 = u(\partial/\partial u)$. Let H denote the subgroup generated by X_1 and X_6. Let H_1 denote the subgroup generated by X_1 alone and H_2 that generated by X_6 alone. Let H' be the subgroup of H whose generator is $\alpha X_1 + \beta X_6$, with $\alpha, \beta \in \mathbb{R}$. A complete set of invariants for H_1 is $I^1 = x$, $I^2 = u$. The solutions invariant under H_1 are such that $\psi(x, u) = 0$, or $u = f(x)$, where f must satisfy a differential equation obtained by substituting u in Eq. (2):

$$f''(x) = 0. \tag{20}$$

This gives $u = ax + b$, where a and b are arbitrary. A complete set of invariants for H_2 is $I^1 = t$, $I^2 = x$; there are no invariant solutions. Now consider H'. Invariants can be obtained by solving

$$(\alpha X_1 + \beta X_6) I = 0 \quad (\alpha, \beta \in \mathbb{R}). \tag{21}$$

A complete set is $I^1 = x$, $I^2 = ue^{-\beta t/\alpha}$. The solutions invariant under H' are of the form $u = e^{\beta t/\alpha} f(x)$, where

$$f''(x) - (\beta/\alpha) f(x) = 0. \tag{22}$$

Observe that if $\beta = 0$, H' reduces to H_1 and the solutions given by Eq. (22) coincide with those given by Eq. (20). Observe also that there are no invariant solutions if $\alpha = 0$, since H' would then reduce to H_2. If $\alpha \neq 0$ and $\beta \neq 0$, Eq. (22) leads to the solutions

$$u = e^{\beta t/\alpha}(ae^{\sqrt{\beta/\alpha}x} + be^{-\sqrt{\beta/\alpha}x}), \tag{23}$$

where a and b are arbitrary. Now let us come back to the (t, x, u, v) space. Invariants for H are obtained by solving $(\alpha X_1 + \beta X_6) I = 0$, i.e.,

$$\left(\alpha \frac{\partial}{\partial t} + \beta u \frac{\partial}{\partial u} + \beta v \frac{\partial}{\partial v}\right) I = 0. \tag{24}$$

A complete set is $I^1 = x$, $I^2 = ue^{-\beta t/\alpha}$, and $I^3 = ve^{-\beta t/\alpha}$. Solutions of two ranks are possible. Those of rank one are the ones invariant[17,18] under H:

$$\psi_1(x, ue^{-\beta t/\alpha}, ve^{-\beta t/\alpha}) = 0$$

and

$$\psi_2(x, ue^{-\beta t/\alpha}, ve^{-\beta t/\alpha}) = 0, \tag{25}$$

where ψ_1 and ψ_2 are some functions. Assuming that the implicit function theorem can be applied, we get $ue^{-\beta t/\alpha} = f(x)$, i.e.,

$$u = e^{\beta t/\alpha} f(x).$$

The solutions of rank two are those partially invariant under H:

$$\psi(x,ue^{-\beta t/\alpha}, ve^{-\beta t/\alpha}) = 0. \tag{26}$$

Again assuming that the implicit function theorem can be applied, we get

$$u = e^{\beta t/\alpha} f(\lambda,\mu), \tag{27}$$

where $\lambda = x$ and $\mu = ve^{-\beta t/\alpha}$. Ovsiannikov's algorithm for constructing the partially invariant solutions now consists in deriving a differential equation for $f(\lambda,\mu)$ by using the system of equations (5) and the compatibility conditions $v_{tx} = v_{xt}$ and $f_{\lambda\mu} = f_{\mu\lambda}$ as well as the relations $\lambda_x = 1$, $\lambda_t = 0, \mu_x = v_x e^{-\beta t/\alpha}$, and $\mu_t = (v_t - \beta v/\alpha)e^{-\beta t/\alpha}$. (Here v is what Ovsiannikov calls a superfluous variable.) The computations are straightforward but a little tedious. One obtains (assuming that $f_\mu \neq 0$)

$$\mu^2 f_{\mu\mu} + 2\mu f_\mu f_{\lambda\mu} - 2\mu f_\lambda f_{\mu\mu} - 2f_\lambda f_\mu f_{\lambda\mu} + f_\lambda^2 f_{\mu\mu} + f_\mu^2 f_{\lambda\lambda} + (\beta/\alpha)\mu f_\mu^3 - (\beta/\alpha) f f_\mu^2 = 0. \tag{28}$$

As observed by Bluman and Cole[12] in a similar context, this is not much of a reduction of the problem of solving Eq. (2). Still one can, in principle, obtain solutions partially invariant under H by solving Eq. (28) first and then using that solution, along with the equations $\lambda = x$, and $\mu = ve^{-\beta t/\alpha}$, to solve either of the equations

$$v_x = (e^{\beta t/\alpha}/f_\mu)(ve^{-\beta t/\alpha} - f_\lambda), \tag{29}$$

$$v_t = v\left(\frac{1}{f_\mu^2} + \frac{\beta}{\alpha}\right) - \frac{e^{\beta t/\alpha}}{f_\mu^2}\left(f_\lambda + \frac{\beta}{\alpha} f f_\mu\right). \tag{30}$$

Let us consider some special cases now. Let $f_\lambda = 0$. Then Eq. (28) reduces to

$$\mu^2 f''(\mu) + (\beta/\alpha)\mu f'^3(\mu) - (\beta/\alpha) f f'^2(\mu) = 0. \tag{31}$$

If we further assume that $\beta = 0$, we obtain a subclass of solutions partially invariant under H_1:

$$\mu f''(\mu) = 0, \tag{32}$$

i.e.,

$$f(\mu) = a\mu + b, \tag{33}$$

where a and b are arbitrary.

Equation (29) now gives $v_x = v/a$, which leads to the solution $v = e^{x/a} g(t)$. It can be shown that this, together with Eq. (33), gives the three-parameter family of solutions

$$u(t,x) = ace^{(ax+t)/a^2} + b, \tag{34}$$

where a, b, and c are arbitrary. Observe that these solutions, while not invariant under H_1 since u is a function of t as well as of x, are actually invariant under the transformations

$$\begin{aligned} t' &= t + k_1, \\ x' &= x + k_2, \\ u' &= ue^{(k_1 + ak_2)/a^2} + b(1 - e^{(k_1 + ak_2)/a^2}), \end{aligned} \tag{35}$$

where k_1 and k_2 are arbitrary. These transformations describe simultaneous translations in t, x, and u and stretching in u. It is easy to check that they form a group. However, that group is not a subgroup of the six-dimensional group G found earlier unless $b = 0$. This is because $b \neq 0$ implies a translation in u, which is represented by the operator $X_7 = g(t,x)(\partial/\partial u)$; and X_7, together with $\{X_i\}$, $1 \leqslant i \leqslant 6$, spans an infinite-dimensional Lie algebra. Thus, if $b \neq 0$, the solutions given by Eq. (34) are partially invariant under H_1 but invariant under the subgroup (of the full group) generated by X_1, X_4, X_6, and X_7. If $b = 0$, they are invariant under the subgroup generated by X_1, X_4, and X_6, which *is* a subgroup of G. It is plausible, though difficult to show explicitly, that the general solution of Eq. (28) would lead to solutions of the heat equation which are not invariant under any subgroup of the full, infinite-dimensional group of Lie symmetries of Eq. (2).

The analysis given for the subgroup generated by X_1 and X_6 applies to some other subgroups of G as well, for instance the subgroup generated by X_2 and X_6. Similar things can be done for the genetics and plasma physics equations also.

Note added in proof: The condition that f satisfies Eq. (28) is necessary but not sufficient for Eq. (2) to have solutions of the type (27). This point, which will be pursued further in a later publication, was made by Professor D. R. K. S. Rao, to whom we are grateful.

ACKNOWLEDGMENTS

We are very grateful to the referee, whose criticism of earlier versions of the paper has led to a substantial clarification and revision of the work. We also thank him/her for pointing out Ref. 16. In addition, one of us (C.C.A.S.) takes pleasure in thanking Professor S. Berman, Professor W. H. Hui, Professor R. Manohar, Professor R. V. Moody, and Professor G. Tenti for helpful discussions at various stages. Professor Moody and Professor Tenti were particularly helpful in improving his understanding of the subject.

[1] A. Fokker, Ann. Phys. Leipzig **43**, 312 (1914).
[2] M. Planck, Sitzungsber. Preuss. Acad. Wiss. Phys. Math. Kl. **1917**, 324.
[3] R. L. Liboff, *Introduction to the Theory of Kinetic Equations* (Wiley, New York, 1969).
[4] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[5] J. D. Cowan, "Stochastic Models of Neuroelectric Activity", in *Statistical Mechanics: New Concepts, New Problems*, edited by S. A. Rice, K. F. Freed, and J. C. Light (Univ. of Chicago, Chicago, 1972).
[6] W. Feller, "Diffusion Processes in Genetics", in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman (Univ. of California, Berkeley, 1951).
[7] D. Lidwig, "Stochastic Population Theories", in *Lecture Notes in Biomathematics*, No. 3 (Springer, New York, 1974).
[8] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper and Row, New York, 1970).
[9] M. A. Lieberman and A. J. Lichtenberg, Plasma Phys. **15**, 125 (1973).
[10] L. S. Ornstein and G. E. Uhlenbeck, Phys. Rev. **36**, 823 (1930).
[11] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I (Wiley, New York, 1968),3rd ed.
[12] G. W. Bluman and J. D. Cole, J. Math. Mech. (now Indiana Univ. Math. J.) **18**, 1025 (1969).
[13] G. A. Nariboli, Stochastic Processes Appl. **5**, 157 (1977).
[14] G. W. Bluman and J. D. Cole, "Similarity Methods for Differential Equations," in *Applied Mathematical Sciences*, No. 13 (Springer, New York, 1974).
[15] G. W. Bluman, Int. J. Non-Linear Mech. **6**, 143 (1971).
[16] G. W. Bluman, SIAM J. Applied Math. **39**, 2, 238 (1980).
[17] L. V. Ovsiannikov, *Group Properties of Differential Equations*, English translation by G. W. Bluman (1967), available (perhaps) from the Dept. of

Applied Math. at Caltech; original Russian edition (Siberian section of the Academy of Sciences, USSR, Novosibirsk, 1962).
[18]L. V. Ovsiannikov, *Group Analysis of Differential Equations*, English translation edited by W. F. Ames (Academic, New York, 1982).
[19]J. M. Hill, *Solution of Differential Equations by means of One-Parameter Groups*, No. 63 in the series *Research Notes in Mathematics* (Pitman, Boston, 1982).
[20]G. Tenti and W. H. Hui, J. Math. Phys. **19**, 774 (1978).

The infinite coupling limit of perturbative expansions from a variational extrapolation method

B. Bonnier, M. Hontebeyrie, and E. H. Ticembal[a)]
Laboratoire de Physique Théorique,[b)] Université de Bordeaux 1, Rue du Solarium, 33170 Gradignan, France

(Received 13 March 1985; accepted for publication 7 June 1985)

A method for extrapolating perturbative power series to infinity is described. It is a Borel partial resummation stabilized by a variational parameter. Two kinds of series relative to the anharmonic oscillators $|x|^k$, $k>0$, are extrapolated in order to illustrate the effectiveness of the method: the Rayleigh–Schrödinger series, on the one hand, which, after extrapolation, provides the strong coupling expansion of the energy levels, and their lattice expansion, on the other hand, from which is extracted the continuum limit.

I. INTRODUCTION

The extrapolation of a physical quantity to its value at infinite argument from the knowledge of its perturbative series at the origin is needed in various circumstances, especially in computing the continuum limit of lattice strong coupling expansions. The problem can be generally stated as follows.

One considers a physical quantity $\Sigma(z)$ as a function of some variable z, which may be, for example, a coupling constant or an inverse lattice spacing. When this variable goes to infinity, the behavior of $\Sigma(z)$ is assumed to be of the form

$$\Sigma(z)\simeq z^{-\alpha}\sum_p A_p z^{-\beta p},\quad p\geqslant 0, \tag{1.1}$$

where the coefficients A_p are unknown: it is the aim of the extrapolation to obtain their value in terms of the perturbative expansion of $\Sigma(z)$ at the origin

$$\Sigma(z)\simeq \Sigma_N(z)\equiv\sum_n a_n z^n,\quad 0\leqslant n\leqslant N. \tag{1.2}$$

To be precise, the leading asymptotic index α is assumed also to be known, since it is always provided by dimensional or scaling arguments. As for the remaining β, it may or may not be given *a priori*, and in the latter case it has to be fixed by self-consistency. In most situations these indices fall in the ranges $-1\leqslant\alpha\leqslant 1$, $0<\beta\leqslant 2$.

There exists a rigorous approach to such a standard problem, generally restricted to quantum mechanics, and various practical solutions, such as the well-established Padé analysis. We shall adopt here the practitioner's point of view and in order to enlarge the panoply of numerical tools we present a simple method, which we have found to be particularly efficient in various cases, some of which are given below. As usual, the efficiency criteria are the consistency of the full available sequence of extrapolated values (i), which have to appear precociously (ii), and this pattern must remain true where other approaches fail (iii).

Our method is, roughly speaking, a Borel-like resummation stabilized by a variational parameter. It is described in Sec. II, and applied in Secs. III and IV to two very different extrapolations, although both concern the energy levels of the Hamiltonian $H(k)$:

$$H(k)\equiv\tfrac{1}{2}p^2+g|x|^k,\quad k>0. \tag{1.3}$$

This has provided for a long time a canonical testing ground.[1] In Sec. III, the (true) strong coupling expansion of the levels is derived from their weak coupling expansion (thus $z\equiv g$, the coupling constant). By itself this result is useful, since generally the extrapolants work only at finite coupling, the main difficulty being to resum the Rayleigh–Schrödinger (asymptotic) series. In Sec. IV we derive the continuum limit of the strong coupling expansion of the ground state. In this case the variable z is the inverse lattice coupling and the limit we seek is given by A_p, $p=0$, in the expansion (1.1).

All these examples are considered according to various values of the anharmonicity parameter k in the range $k>0$, and some of the associated series are known to be very difficult,[2,3] even "impossible,"[2] to extrapolate. We feel that the present method greatly improves the situation in this particular framework and, as it is not restricted to quantum mechanics, our hope is that it can facilitate the analysis of various lattice expansions in field theory.

II. CONSTRUCTION OF THE EXTRAPOLANTS

Taking into account the set of input parameters previously defined, i.e., $\{a_n, 0\leqslant n\leqslant N, \alpha, \beta\}$, we first construct $N+1$ polynomials $P_n(\lambda)$ according to

$$P_n(\lambda)=\sum_l a_l\Gamma^{-1}(n-l+1)\Gamma^{-1}((\alpha+l)/\beta)\lambda^{n-l},$$
$$0\leqslant l\leqslant n\leqslant N. \tag{2.1}$$

It will be clear in the following that λ, which we treat for the moment as a variable, is for the extrapolation a variational parameter and as such, it will be fixed at some positive value in terms of the input parameters. It is useful to observe that

$$\frac{dP_n}{d\lambda}(\lambda)=P_{n-1}(\lambda),\quad P_n(0)=\Gamma^{-1}\left(\frac{\alpha+n}{\beta}\right)a_n, \tag{2.2}$$

$$\frac{1}{n}\sum_{i=1}^{n}\lambda_i^n=\lambda_1^1=-\frac{a_1}{a_0}\Gamma\left(\frac{\alpha}{\beta}\right)\Gamma^{-1}\left(\frac{\alpha+1}{\beta}\right),\quad 1\leqslant n\leqslant N. \tag{2.3}$$

[a)] Stagiaire de Recherche du Gouvernement Algérien.
[b)] Unité de Recherche associée au Centre National de la Recherche Scientifique.

 0022-2488/85/123048-05$02.50

The relation (2.3), where λ_i^n denotes the ith root of $P_n(\lambda)$, has a simple geometrical interpretation: the first root λ_1^1 is the common barycenter of all the n-polygons constructed with the n roots of $P_n(\lambda)$.

As a next step we consider the one-dimensional integrals $I_l(z,\lambda)$ involving the indices α and β previously defined:

$$I_l(z,\lambda)=\beta z^l\int_0^\infty u^{\alpha+l-1}e^{-\lambda z u-u^\beta}du, \quad l\geqslant 0. \tag{2.4}$$

From this definition it is obvious that

$$I_l(z=0^+,\lambda)=z^l\Gamma((\alpha+l)/\beta)+O(z^{l+1},\lambda), \tag{2.5}$$

$$I_l(z,\lambda=0)=z^l\Gamma((\alpha+l)/\beta)),$$

$$\frac{\partial I_l}{\partial\lambda}(z,\lambda)=-I_{l+1}(z,\lambda). \tag{2.6}$$

On the other hand, the change of variable $v=\lambda z u$ in definition (2.4) gives an equivalent representation,

$$I_l(z,\lambda)=\beta z^{-\alpha}\lambda^{-\alpha-l}\int_0^\infty v^{\alpha+l-1}e^{-v-(v/\lambda z)^\beta}dv, \tag{2.7}$$

which displays the $z=+\infty$ behavior of $I_l(z,\lambda)$ for $\lambda>0$

$$I_l(z,\lambda)=z^{-\alpha}\sum_p \omega_p^l(\lambda)z^{-\beta p}, \quad p\geqslant 0, \tag{2.8}$$

with

$$\omega_p^l(\lambda)=(-1)^p\beta\lambda^{-\alpha-l-\beta p}\Gamma^{-1}(p+1)\Gamma(\alpha+l+\beta p). \tag{2.9}$$

Thus, comparing the relations (1.1) and (1.2) and (2.8) and (2.5) shows that $\Sigma(z)$ and $I_l(z,\lambda)$, $l\geqslant 0$, have the same kind of formal expansions at $z=+\infty$ and $z=0^+$. Our method of extrapolation can be interpreted as choosing these integrals as a "basis" in which to expand $\Sigma(z)$. The coefficients of this expansion are fixed from the knowledge of the $z=0$ perturbation series $\Sigma_N(z)$ and, as we shall see, turn out to be the polynomials $P_n(\lambda)$. Our nth-order extrapolant $\sigma_n(z,\lambda)$ then takes the form

$$\sigma_n(z,\lambda)=\sum_l P_l(\lambda)I_l(z,\lambda), \quad 0\leqslant l\leqslant n\leqslant N. \tag{2.10}$$

In order to prove that $\sigma_n(z,\lambda)$ and $\Sigma(z)$ match at $z=0^+$, we observe that, from relations (2.2) and (2.6),

$$\frac{\partial\sigma_n}{\partial\lambda}(z,\lambda)=P_n(\lambda)I_{n+1}(z,\lambda), \tag{2.11}$$

in such a way that

$$\sigma_n(z,\lambda)=\sigma_n(z,0)+\int_0^\lambda P_n(\mu)I_{n+1}(z,\mu)d\mu, \tag{2.12}$$

for all z allowing the μ integration (which excludes $z=+\infty$ due to the end point singularity $\mu=0$). Since, from relations (2.2) and (2.6),

$$\sigma_n(z,0)=\sum_l a_l z^l, \quad 0\leqslant l\leqslant n, \tag{2.13}$$

the representation (2.12) indicates that $\sigma_n(z,\lambda)$ and $\Sigma(z)$ have the same expansion (1.2) at $z=0^+$ up to order n, $0\leqslant n\leqslant N$, independently of λ ($\lambda>0$). The $z=+\infty$ extrapolation is then provided by definition (2.10), where $I_l(z,\lambda)$ is replaced by its asymptotic expansion [(2.8) and (2.9)]. Identifying the result with the form (1.1) gives the extrapolated values of A_p:

$$A_p\sim A_p^n, \quad p\geqslant 0,$$

$$A_p^n=\sum_l \omega_p^l(\lambda)P_l(\lambda) \quad 0\leqslant l\leqslant n\leqslant N. \tag{2.14}$$

We must now comment on the λ dependence of the extrapolants A_p^n, which follows from Eq. (2.11):

$$\frac{dA_p^n}{d\lambda}=\omega_p^{n+1}(\lambda)P_n(\lambda). \tag{2.15}$$

It indicates that for any $p\geqslant 0$ [or any $z>0$ in relation (2.11)] the roots λ_i^n of $P_n(\lambda)$ stabilize the nth-order extrapolants A_p^n. As n increases, these roots (or their real parts) spread over a region R_n increasing in magnitude but always containing λ_1^1, as indicated by relation (2.3). Moreover, at least in the examples we have investigated, the variation of $A_p^n(\lambda)$, when λ describes R_n, is limited to small oscillations: this can be interpreted as an indication of convergence towards a λ-independent value, as it should. Yet at each order of the extrapolation an optimal value $\bar{\lambda}_n$ can be chosen inside R_n, as an illustration of "Stevenson's principle."[4] However, taking advantage of the central position of λ_1^1, which we have already emphasized (and as $\lambda_1^1>0$,) we make in what follows the simple choice $\lambda=\lambda_1^1$, independent of the order. This completes the definition of our sequence of extrapolants in terms of the input parameters.

It is clear from the above derivation that we remain at the level of formal series manipulations. Some comments can be made, however, in order to justify our choice for the "basis" $I_l(z,\lambda)$. First, these functions (with $\lambda=1$) already appear in perturbative expansions of multidimensional functional integrals.[5] Second, and this gives some hints about the convergence of the extrapolation, they realize a particular Borel resummation of $\Sigma_N(z)$, which can be seen in the following way. The Borel transform $\rho(u)$ of $\Sigma_N(z)$ with indices (α,β) is defined as

$$\rho(u)=\sum_n a_n u^n\Gamma^{-1}\left(\frac{\alpha+n}{\beta}\right), \quad n\geqslant 0, \tag{2.16}$$

in such a way that

$$\Sigma(z)=\beta\int_0^\infty e^{-y^\beta}y^{\alpha-1}\rho(zy)dy. \tag{2.17}$$

On the other hand, it can be checked from their definition that the polynomials $P_n(\lambda)$ are generated by $e^{\lambda u}\rho(u)$, i.e.,

$$e^{\lambda u}\rho(u)=\sum_{n\geqslant 0}u^n P_n(\lambda) \quad u\simeq 0. \tag{2.18}$$

Thus, if λ is such that the relation (2.18) extends to $u\geqslant 0$, the continuation it implies for $\rho(u)$ can be put into the integral (2.17), which is then at order n nothing else than $\sigma_n(z,\lambda)$ [from definition (2.10) with $I_l(z,\lambda)$ given in definition (2.4)].

III. FROM WEAK TO STRONG COUPLING: THE ANHARMONIC OSCILLATOR EXAMPLE

As a typical illustration of the previous method we consider the strong coupling regime ($g\to+\infty$) of the L th energy level $E_L(g)$, $L\geqslant 0$, of the quantum mechanical Hamiltonian H_k given in definition (1.3). It is known from scaling[1] that they behave according to a convergent expansion of the form (1.1), where

TABLE I. First-order approximant $A_0^{\frac{1}{2}}$ of the leading coefficient A_0 of the strong coupling expansion for the ground state of the Hamiltonian H_k, for various values of k. The exact results A_0 are taken from Ref. 2.

k	$\frac{1}{2}$	$\frac{2}{3}$	1	$\frac{3}{2}$	4	6	8	10
$A_0^{\frac{1}{2}}$	0.918 22	0.874 18	0.806 88	0.743 54	0.666 73	0.680 30	0.707 90	0.739 98
A_0	0.922 45	0.878 94	0.808 61	0.743 88	0.667 98	0.680 70	0.704 05	0.728 48 $<A_0$ $<0.729\ 60$

$$\alpha = -2/(2+k), \quad \beta = -2\alpha, \tag{3.1}$$

i.e.,

$$E_L(g) = g^{2/(2+k)} \sum_p A_p g^{-4p/(2+k)}, \quad p \geqslant 0. \tag{3.2}$$

Our aim is to express the coefficients A_p $p \geqslant 0$, in terms of (α, β) and the perturbative expansion of $E_L(g)$ at $g = 0^+$:

$$E_L(g) = \sum_p a_p g^p, \quad 0 < p < N. \tag{3.3}$$

This approach is thus an improvement of earlier ones[6] undertaken in the same spirit, since it seems there does not exist a specific algorithm to generate the coefficients A_p.

We first consider the ground state $E_0(g)$, whose Rayleigh–Schrödinger series (3.3) begins with

$$a_0 = \tfrac{1}{2}, \quad a_1 = 2^{1-k}\Gamma(k)\Gamma^{-1}(k/2). \tag{3.4}$$

This allows us to compute the root $\lambda_1^{\frac{1}{2}}$ defined in Eq. (2.3),

$$\lambda_1^{\frac{1}{2}} = 2^{3-k}\sqrt{\pi}\Gamma(k)\Gamma^{-1}(k/2)\Gamma^{-1}(k/4), \tag{3.5}$$

which is positive, as required (but unbounded as k goes to infinity). It is instructive to write explicitly the first-order approximant $A_p^{\frac{1}{2}}$ of A_p from Eq. (2.14), i.e.,

$$A_p^{\frac{1}{2}} = \frac{(-1)^{p+1}(\lambda_1^{\frac{1}{2}})^{(2-4p)/(2+k)}}{\sqrt{\pi}(2+k)} \times \Gamma^{-1}(p+1)\Gamma\left(\frac{4p-2}{2+k}\right), \tag{3.6}$$

which gives, for the first values of p ($p = 0, 1, 2$)

$$\begin{aligned} A_0^{\frac{1}{2}} &= (1/2\sqrt{\pi})(\lambda_1^{\frac{1}{2}})^{2/(2+k)}\Gamma(k/(k+2)), \\ A_1^{\frac{1}{2}} &= [1/\sqrt{\pi}(k+2)]\ (\lambda_1^{\frac{1}{2}})^{-2/(2+k)}\Gamma(2/(k+2)), \\ A_2^{\frac{1}{2}} &= [-1/\sqrt{\pi}(k+2)]\,(\lambda_1^{\frac{1}{2}})^{-6/(k+2)}\Gamma(6/(k+2)). \end{aligned} \tag{3.7}$$

It can be verified in the soluble example of the harmonic oscillator $k = 2$ that these expressions are then the exact ones. In this case the approximants are independent of the order n since $\lambda_1^{\frac{1}{2}}$ turns out to be the multiple common root of all the polynomials $P_n(\lambda)$. We give in Table I some values of the leading coefficient $A_0^{\frac{1}{2}}$ as a function of the anharmonicity parameter k and later in Table III (first line) the values of $A_1^{\frac{1}{2}}$ and $A_2^{\frac{1}{2}}$ when $k \leqslant 4$. Since $\lambda_1^{\frac{1}{2}}$ is exponentially divergent as $k \to +\infty$, the agreement of the values implied by Eqs. (3.7) with the true ones, which is perfect at $k = 2$, must be lost when k increases. Despite the fact that only a_0 and a_1 were used as input, the agreement is seen to be rather good over a large range of anharmonicity.

This feature persists for higher levels, as we show now in the case of the x^4 anharmonic oscillator, as a standard example. The first terms of the series (3.3) in such a case are

$$a_0 = L + \tfrac{1}{2}, \quad a_1 = \tfrac{3}{2}a_0^2 + \tfrac{3}{8}, \tag{3.8}$$

and at first order the leading behavior of $E_L(g)$ for $g \to +\infty$ is found to be

$$E_L(g) \sim g^{1/3}(L+\tfrac{1}{2})^{4/3}C(L), \tag{3.9}$$

with

$$C(L) = (3/\pi)^{1/3}\Gamma(\tfrac{2}{3})[1 + (2L+1)^{-2}]. \tag{3.10}$$

The numerical variations of $C(L)$ with L, up to $L \to +\infty$, shown in Table II, indicate a general agreement with the exact values within 3% at most.

We turn now to the investigation of the higher-order corrections. From the last comment at the end of the previous section, these are expected to be small, even convergent, at fixed $\lambda = \lambda_1^{\frac{1}{2}}$, when the asymptotic indices (α, β) ensure Borel resummation of the Rayleigh–Schrödinger series. We thus compare these indices with the leading exponential behavior of the coefficients a_n, when n goes to infinity, which is known to be[7]

$$a_n \sim (-1)^n\Gamma(n(k/2-1)+\tfrac{1}{2}), \quad L = 0. \tag{3.11}$$

The convergence can be expected when

$$1/\beta \geqslant k/2 - 1, \quad \text{i.e, } k \leqslant 3. \tag{3.12}$$

We have not studied examples ruled out by the criteria (3.12) but we think that a higher order can be stabilized by relaxing the constraint $\lambda = \lambda_1^{\frac{1}{2}}$, as is obviously necessary when $k \to +\infty$. As the typical anharmonic oscillator x^4 lays in the admissible range (3.12), we have computed the sequence of approximants to A_0, A_1, and A_2 for its ground state. These are listed in Table III up to the order $N \sim 15$ and the results obviously support our conjecture. It is also interesting from a practical point of view to observe that high precision is reached within the first few orders.

We thus think that the method, at least when the anharmonicity is not too strong, provides us with a systematic,

TABLE II. First-order approximant $C(L)$ of the L th level of the anharmonic oscillator H_4, as given by Eq. (3.10), for various values of L. The exact values are taken from Ref. 9.

L	0	1	2	3	4	5	$+\infty$
$C(L)$	1.6800	1.3811	1.3510	1.3425	1.3389	1.3371	1.3345
Exact	1.6832	1.3940	1.3842	1.3804	1.3789	1.3781	1.3765

TABLE III. Numerical values at increasing order n of the extrapolants of the three first coefficients A_0, A_1, and A_2 of the strong coupling expansion (1.1) in the case of the ground state of the anharmonic oscillator H_4. The values for A_0 and A_1 are taken from Ref. 9, and from Ref. 10 for A_2.

n	A_0^n	A_1^n	$-10^2\times A_2^n$
1	0.666 730 450	0.144 323 759	0.884 194
2	0.667 449 228	0.144 012 579	0.875 615
3	0.667 883 970	0.143 749 081	0.866 275
4	0.667 962 767	0.143 689 382	0.863 736
5	0.667 986 291	0.143 668 319	0.862 702
6	0.667 987 601	0.143 669 783	0.862 628
7	0.667 987 692	0.143 666 874	0.862 621
8	0.667 986 823	0.143 667 968	0.862 695
9	0.667 986 578	0.143 668 303	0.862 718
10	0.667 986 335	0.143 668 660	0.862 746
11	0.667 986 304	0.143 668 709	0.862 750
12	0.667 986 253	0.143 668 794	0.862 758
13	0.667 986 264	0.143 668 776	0.862 756
14	0.667 986 250	0.143 668 800	0.862 758
Exact value	0.667 986 259	0.143 67	~0.863

simple, and analytic algorithm for going from weak to strong coupling expansions.

IV. EXTRAPOLATION OF LATTICE SERIES

As a different kind of application of our method, we compute the continuum limit of a family of lattice strong coupling expansions. The series we have chosen gives in that limit the ground state $E_0(g)$ of the Hamiltonians H_k previously considered, according to

$$E_0(g) = g^{2/(2+k)}\epsilon(k), \tag{4.1}$$

$$\epsilon(k) = [(k+2)/2k]\lim_{t=+\infty}\{t^{k/(k+2)}\Sigma(t)\}. \tag{4.2}$$

In the relation (4.2), $\Sigma(t)$ denotes a perturbative series corresponding to the expansion in the inverse coupling of the functional integral associated to H_k, within the regularization provided by a one-dimensional lattice of spacing a. This expansion reads

$$\Sigma_N(t) = \Sigma a_n t^n, \quad a_0 = 1, \quad a_1 = -2\Gamma(3/k)\Gamma^{-1}(1/k), \tag{4.3}$$

where t is a dimensionless variable

$$t = a^{-(2+k)/k}g^{-2/k}, \quad \text{i.e., } a \to 0 \Rightarrow t \sim +\infty. \tag{4.4}$$

A detailed derivation of these relations can be found in Refs. 2 and 3.

Our objective here is to compute $\epsilon(k)$ from the knowledge of $\Sigma_N(t)$, $n \lesssim N \sim 20$, since the limit (4.2) implies that $\Sigma(t)$ behaves at $t = +\infty$ according to the form (1,1) where the index α is

$$\alpha = k/(k+2). \tag{4.5}$$

As a first technical difficulty, absent in the previous example, it thus appears that the remaining index β is unspecified. The other difficulty is that $\Sigma_N(t)$ extrapolates badly, according to previous analyses, for some specific values of the anharmonicity. In fact the behavior of the input series $\Sigma_N(t)$ presents the following characteristics.[3,8]

(i) For $k < 1$ the series is asymptotic:

$$a_n \sim (-1)^n\Gamma(n(2/k-1)) \quad \text{as } n \to +\infty. \tag{4.6}$$

It is Borel summable in the interval $\frac{2}{3} < k < 1$ and when $k < \frac{2}{3}$ there is no successful extrapolation of the series (the Padé approximants converge to a wrong answer[2]).

TABLE IV. Range of values found for β and ϵ_N by applying the criteria (4.8) at order N for various choices of the anharmonicity parameter k. Here, ϵ_N is the Nth-order approximant of $\epsilon(k)$ defined in Eq. (4.2) for the ground state of H_k. The exact values are taken from Ref. 2.

N	k	β	ϵ_N	Exact value
5		0.34 – 0.41	0.857 – 0.912	
10	$\frac{1}{2}$	0.32 – 0.37	0.882 – 0.949	0.922 45
15		0.31 – 0.35	0.890 – 0.916	
20		0.30 – 0.34	0.928 – 0.949	
5		0.47 – 0.51	0.846 – 0.865	
10	$\frac{2}{3}$	0.45 – 0.50	0.865 – 0.872	0.878 94
15		0.40 – 0.47	0.863 – 0.892	
20		0.41 – 0.46	0.875 – 0.886	
5		0.63 – 0.68	0.795 – 0.805	
10	1	0.62 – 0.69	0.799 – 0.809	0.808 61
15		0.55 – 0.68	0.803 – 0.816	
20		0.60 – 0.63	0.808 – 0.811	
5		1.0 – 1.5	0.653 – 0.676	
10	4	1.0 – 1.5	0.663 – 0.680	0.667 98
15		1.0 – 1.5	0.651 – 0.671	
20		1.28 – 1.33	0.6680 – 0.6685	
5		0.75 – 4.8	0.70 – 0.95	
10	∞	0.75 – 2.7	0.78 – 1.3	1.233 7
15		0.9 – 1.85	0.83 – 0.96	
20		0.95 – 1.65	0.91 – 1.2	

(ii) For $1 \leqslant k < 2$ the series remain asymptotic, the large orders being unknown. Various extrapolations work.

(iii) For $k \geqslant 2$ the series has a finite radius of convergence. As k increases it becomes more difficult to extract the continuum limit, especially at $k = +\infty$ (the square well).

Taking these facts into account, we want to apply our method to some characteristic values of k, i.e., $k = \frac{1}{2}, \frac{2}{3}, 1, 4, +\infty$. We want also to stick to our simplifying choice $\lambda = \lambda_1^1$, which is admissible, since here

$$\lambda_1^1 = \frac{1}{2}\Gamma\left(\frac{1}{k}\right)\Gamma^{-1}\left(\frac{3}{k}\right)\Gamma\left(\frac{k}{\beta(k+2)}\right) \times \Gamma^{-1}\left(\frac{1}{\beta}\left(1+\frac{k}{k+2}\right)\right), \tag{4.7}$$

i.e., λ_1^1 is positive ($\beta > 0$) and bounded for all k. The results ϵ_p we obtain in that way for $\epsilon(k)$ at order p are thus β dependent and this freedom has been restricted by self-consistency, as we look for the minimum in β of the function $\Delta_n(\beta)$ defined by

$$\Delta_n(\beta) = \sum_p \left(\frac{\epsilon_p - \epsilon_{p-1}}{\epsilon_p}\right)^2, \quad \sup(1, n-10) \leqslant p \leqslant n. \tag{4.8}$$

In fact we allow for $\Delta_n(\beta)$ a variation of one order of magnitude around its minimum: this gives an "admissible" range of values for β, and the corresponding extrapolants ϵ_p. These results are listed in Table IV for the chosen values of k, and they display a significant improvement over previous approaches,[2,3] especially as $k \to 0$. We emphasize that the harmonic oscillator case is exactly treated with the value $\beta = 1$, and that our estimate of β from the criteria (4.8) is consistent with the large-order behavior given in Eq. (4.6), which indicates that

$$1/\beta \geqslant 2/k - 1,$$

i.e.,

$$\beta \leqslant \beta_c(k) = k/(2-k), \quad k < 1, \quad \beta_c(\tfrac{1}{2}) = \tfrac{1}{3}. \tag{4.9}$$

Our main conclusion is to observe that satisfactory results are obtained with a reasonable amount of perturbative terms ($N \sim 10$) on the whole range of values of k, in spite of the very different kinds of behavior of the extrapolated series. This indicates that such a method may be successfully applied to lattice expansions in field theory.

[1] B. Simon in *Cargese Lectures in Physics* (Gordon and Breach, New York, 1970), p. 383; A. Martin, *ibid.*, p. 415.
[2] C. M. Bender, L. R. Mead, and L. M. Simmons, Jr., Phys. Rev. D **24**, 2674 (1981).
[3] B. Bonnier, Nucl. Phys. B **243**, 253 (1984).
[4] P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981).
[5] B. S.Shaverdyan and A. G. Ushveridze, Phys. Lett. B **123**, 316 (1982).
[6] G. Parisi, Phys. Lett. B. **69**, 329 (1977); B. Bonnier, Phys. Lett. B **78**, 107 (1978).
[7] E. Brezin, J. C. Le Guillou, and J. Zinn-Justin, Phys. Rev. D **15**, 1544 (1977).
[8] C. M. Bender, L. R. Mead, and L. M. Simmons, Jr., Phys. Rev. D **28**, 936 (1983).
[9] F. T. Hioe and E. W. Montroll, J. Math. Phys. **16**, 1945 (1975).
[10] M. A. Smondyrev, J.O.I.N.T preprint, Dubna, 1984.

Dynamical group chains and integrity bases

R. Gilmore
Department of Physics and Atmospheric Science, Drexel University, Philadelphia, Pennsylvania 19104

J. P. Draayer
Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803

(Received 27 March 1985; accepted for publication 18 June 1985)

An algorithm for constructing a Hamiltonian from the generators of a dynamical group G, which is invariant under the operations of a symmetry group $H \subset G$, is presented. In practice, this algorithm is subject to a large number of simplifications. It is sufficient to construct an integrity basis of H scalars in terms of which all H scalars can be expressed as polynomial functions. In many instances the integrity basis exists in 1–1 correspondence with the Casimir operators for a group–subgroup lattice based on the pair $H \subset G$. When this is so the theory embodies natural symmetry limits and analytic results for observables can be given. Examples of the application of the algorithm are given for the dynamical group SU(2) with symmetry subgroups C_3 and U(1) and for SU(N) $\supset$ SO(3), $N = 3$, 4, and 6.

I. INTRODUCTION

The physical properties of many systems are determined by a dynamical group G. Often G is a finite-dimensional Lie group, whose Lie algebra $\mathbf{G}$ is spanned by a set of operators $\mathbf{X}_i$, $i = 1,2,\ldots,n = \dim(\mathbf{G})$, the infinitesimal generators of G. The dynamical properties of the system are then defined by specifying a Hamiltonian. The Hamiltonian is a function of the basis vectors of $\mathbf{G}$, $\mathbf{H} = \mathbf{H}(\mathbf{X})$. This function can be expressed as a graded and symmetrized power series in the $\mathbf{X}_i$,

$$\mathbf{H}(\mathbf{X}) = A^{(0)}\mathbf{I} + A^{(1)}_i \mathbf{X}_i + (1/2!) A^{(2)}_{ij} \mathbf{X}_i \mathbf{X}_j + (1/3!) A^{(3)}_{ijk} \mathbf{X}_i \mathbf{X}_j \mathbf{X}_k + \cdots . \tag{1.1}$$

The terms homogeneous of degree d in the $\mathbf{X}_i$ may be taken as symmetric under permutation of the operators. If they are not symmetric, they may be written as the sum of a symmetric part and an antisymmetric part. The degree of the antisymmetric part can be reduced by one using the commutation relations of $\mathbf{G}$, $[\mathbf{X}_i, \mathbf{X}_j] = C^k_{ij} \mathbf{X}_k$. As a result, there are $(n + d - 1)!/(n - 1)!d!$ operators which are homogeneous of degree d in the n generators $\mathbf{X}_i$ and fully symmetric under the action of the permutation group P_d. These operators span a linear vector space of operators, $\mathbf{U}^d(\mathbf{G})$. The direct sum of these operator spaces is called the universal enveloping algebra of $\mathbf{G}$, $\mathbf{U}(\mathbf{G}) = \Sigma^{\infty}_{d=0} \oplus \mathbf{U}^d(\mathbf{G})$ (see Ref. 1).

If the physical system possesses a symmetry group $H \subset G$, then the Hamiltonian must be invariant under the action of H. This severely restricts the form of the expansion (1.1), since the Hamiltonian must now be a sum of the H scalars of $U(\mathbf{G})$; that is, operators which transform under the identity representation $\gamma^e(H)$ of H. The determination of the H scalars in $\mathbf{U}(\mathbf{G})$ is a classic group theory problem resolved by a simple algorithm.

Algorithm:

(1) Determine $\Gamma^d(G)$, the representation of G carried by $\mathbf{U}^d(\mathbf{G})$.

(2) Determine the number of times $\gamma^e(H)$ occurs in $\Gamma^d(G)$ under the restriction of G to H.

(3) The basis operators for each $\gamma^e(H)$ are the H scalars in $\mathbf{U}^d(\mathbf{G})$.

If, in addition, the Hamiltonian is to display time reversal invariance and be Hermitian, the expansion (1.1) must be further restricted to accommodate the requirements of these order-2 discrete operations.

If G_i is a subgroup of the dynamical group G containing the symmetry group ($H \subseteq G_i \subseteq G$), the invariant operators of G_i (Casimir invariants) are H scalars in $\mathbf{U}(\mathbf{G})$. If a set of subgroups G_i (including H and G) exists with the property that the number of H scalars of degree d is equal to the number of distinct products of degree d of the Casimir operators of the G_i, then the search for H scalars can be resolved by the construction of a group–subgroup lattice. When this is possible, major simplifications follow since the Casimir operators and their spectra are known for the simple Lie groups. In particular, the existence of such group–subgroup chains produces a theory that embodies natural symmetry limits and has analytic expressions for observables.

II. GUIDELINES AND SIMPLIFICATIONS

The algorithm presented above is much easier to state than to implement. However, in many cases of physical interest (e.g., restriction to boson Hilbert spaces) a number of simplifications occur, making it unnecessary to deal with the full group theoretical machinery implicit in steps (1)–(3). For this reason we devote this section to a number of guidelines for using the algorithm, to the conditions under which simplifications occur, and to methods available for implementing the simplifications in an economical way.

(1) The representation $\Gamma^d(G)$ of G on $\mathbf{U}^d(\mathbf{G})$ is generally reducible. It is then useful to carry out the reduction

$$\Gamma^d(G) = \Sigma \oplus \Gamma^{\lambda}(G) \tag{2.1}$$

into a direct sum of irreducible representations $\Gamma^{\lambda}(G)$, since a great deal is known about the irreducible representations of Lie groups. In particular, standard tools can be invoked to effect the reduction of $\Gamma^{\lambda}(G)$ to $\gamma^e(H)$.

(2) The $\mathbf{X}_i$ span the adjoint representation of $\mathbf{G}$. The symmetric tensor product $\mathbf{X}_i \otimes \mathbf{X}_j \otimes \cdots \otimes \mathbf{X}_k$ can then be computed, and its irreducible content deduced by applying a

 0022-2488/85/123053-15$02.50

symmetric plethysm to the adjoint representation of G (see Ref. 2).

(3) When U(G) acts in an invariant Hilbert space, which carries an irreducible representation $\Gamma^{\mu}(G)$, the number of inequivalent irreducible tensors of type (λ) [i.e., tensors that transform irreducibly under $\Gamma^{\lambda}(G)$] that occur in U(G) is equal to the number of times (μ) occurs in the decomposition of $(\lambda) \otimes (\mu)$ (see Ref. 3).

(4) If two sets T^{λ_1} and T^{λ_2} of irreducible tensor operators of degrees d_1 and $d_2 > d_1$ are equivalent within each invariant Hilbert space, then $\lambda_1 = \lambda_2$ and T^{λ_2} can be obtained by multiplying T^{λ_1} by a function of the Casimir operators of G of degree $d_2 - d_1$, $T^{\lambda_2} = F(\mathbf{C})T^{\lambda_1}$.

(5) When $U(\mathbf{G})$ acts on an invariant Hilbert space which carries an irreducible representation $\Gamma^{\mu}(G)$, where μ is a Young pattern of arbitrary size and row lengths (generic or nondegenerate case), the number of inequivalent tensors of type (λ) that occur in U(G) is equal to the number of zero weights that occur in the representation $\Gamma^{\lambda}(G)$ [e.g., 2, 3, 5 for the adjoint representations of SU(3), SU(4), SU(6), respectively].[4]

(6) Under restriction of the action of U(G) from generic to degenerate classes of representations, characterized by having one or more Young partition row lengths equal to 0 or more than one of equal length (i.e., $[k]$ *or* $[1^k]$, for boson or fermion representations of $SU(N)$, $N > 2$), the inequivalent tensor content of U(G) can be reduced in three ways: (i) some operators may have only zero matrix elements within the restricted class of Hilbert spaces under consideration; (ii) inequivalent tensors of the same type in U(G) may become linearly dependent on restriction to a special class of Hilbert spaces; and (iii) functionally independent operators (e.g., the Casimir operators) may exhibit functional dependences on restriction to degenerate Hilbert spaces.

(7) When U(G) acts on nongeneric or degenerate representations of type $\Gamma^{\mu}(G)$, the number of inequivalent irreducible tensors of type (λ) that occur can be determined by using a simple algorithm.

(i) Determine the highest weight in (μ).

(ii) Determine the largest subgroup $K \subset G$ which leaves this highest weight invariant up to a phase factor.

(iii) Determine the number of times the identity representation of K, $\gamma^{e}(K)$, is contained in the restriction of $\Gamma^{(\lambda)}(G)$ to K.

This number is equal to the number of inequivalent tensors of type (λ) which occur in U(G), and these operators are the basis vectors on which the $\gamma^{e}(K)$ act. This algorithm is an implementation of remark (3), above, for degenerate representations. For generic representations, K is equal to the Cartan subgroup, so remark (5) is a special case of this result.

(8) As d increases, the number of G tensors in $U^{d}(\mathbf{G})$ increases rapidly. It then becomes useful to apply the same methods to the description of irreducible tensor operators as have been applied to a description of irreducible representations. In the latter case, a complete set of fundamental irreducible representations is introduced, with the property that every irreducible representation can be obtained as a "stretched" (fully symmetric) product formed from members of this fundamental set.[3] A semisimple group or algebra of rank l has exactly l fundamental irreducible representations. Similarly, in the case of tensor operators, a complete set of fundamental irreducible tensor operators is introduced, with the property that every irreducible tensor operator can be obtained as a "stretched" tensor product formed from members of this set.[5] The set of fundamental tensor operators is finite; a folk theorem says that the number of non-Casimir fundamental H-scalar tensors is twice the number of missing labels in the reduction of representations of the dynamical group G to the symmetry subgroup H (see Ref. 6). The number of missing labels depends on the class of representations of G under consideration and decreases as the degeneracy increases. The complete set of fundamental tensor operators is called an integrity basis for U(G).

(9) It is possible to construct generating functions for the irreducible tensor content of $U^{d}(\mathbf{G})$ in terms of the integrity basis for U(G).

(10) The simplifications brought about by the introduction of generating functions in the integrity basis encounters a minor problem; namely, not all possible products of tensors in the intregity basis may be independent. The generating function must take account of this.

(11) The generating function for the irreducible G-tensor content of $U^{d}(\mathbf{G})$ can be transformed into a generating function for the H-scalar content in $U^{d}(\mathbf{G})$ using the known branching rules of $G \downarrow H$.

(12) The generating functions and branching rules for Lie subgroups of Lie groups generally do not incorporate the discrete operations under which the Hamiltonian must remain invariant. The two most important discrete transformations for our purposes are space reflection P and time reversal T. Both of these discrete operations are of order two. As a result, under each discrete operation the Lie algebra of G splits into two parts corresponding to the positive and negative eigenvalues ($+1$ and -1) of these discrete transformations. Only H scalars that are positive eigenvectors of P and T can occur in the expansion (1.1).

(13) Hermitian conjunction is an operation on representations of an algebra G rather than on G itself. The Hamiltonian must be invariant under Hermitian conjugation. When G is compact, the adjoint representation [see remark (2)] of the algebra G is Hermitian. The H-scalar operators that occur in $U^{d}(\mathbf{G})$ are then either self-adjoint or occur in Hermitian adjoint pairs, since $U^{d}(\mathbf{G})$ is a symmetric d th-order tensor product of self-adjoint operators. Each self-adjoint operator in U(G) introduces one real parameter while each Hermitian adjoint pair of operators introduces one complex parameter.

In the following four sections we illustrate the use of the algorithm presented in Sec. I as well as the mechanisms and procedures listed in this section, for four dynamical groups of physical importance. These are SU(2) (point group tensor harmonics, Morse oscillator, and isospin), SU(3) (Elliott model), SU(4) (vibron model), and SU(6) (interacting boson model).

III. DYNAMICAL GROUP SU(2)

In this section we consider SU(2) as a dynamical group in three manifestations. In the first case the symmetry sub-

group H is taken to be the discrete group C_3. This makes contact with previous work in ligand field theory. In the second case the symmetry subgroup is taken to be U(1). The SU(2) $\supset$ U(1) structure is the simplest case in which G and H are both Lie groups. This situation is encountered, for example, with the Morse oscillator[7,8] and in isospin theories.[9] In the third case the dynamical group G is taken to be the direct product group SU(2) $\otimes$ SU(2) with the symmetry subgroup H again taken to be U(1). This example serves as a prototype for semisimple Lie groups containing a common Lie subgroup.

A. SU(2) $\supset C_3$

The Lie algebra SU(2) of the Lie group SU(2) is spanned by the three basis operators $\mathbf{J}_i$, which may be taken in either the Cartesian ($i = x,y,z$) or spherical ($i = +,3,-$) basis. The universal enveloping algebra U[SU(2)] has the form

$$F(J) = A^{(0)}I + A^{(1)}_i J_i + (1/2!)\,A^{(2)}_{ij}\,\mathbf{J}_i\mathbf{J}_j + (1/3!)\,A^{(3)}_{ijk}\,\mathbf{J}_i\,\mathbf{J}_j\,\mathbf{J}_k + \cdots . \tag{3.1}$$

The series coefficients $A^{(d)}_{i_1\cdots i_d}$ are invariant under permutation of the indices. There are $\frac{1}{2}(d+1)(d+2)$ independent homogeneous symmetric polynomials of degree d in the generators $\mathbf{J}_i$. These polynomials carry a representation Γ^d of SU(2), which is reducible for $d>1$ [cf. Sec. II, remark (1)]. The irreducible constant of Γ^d is D^L, $L = d,d-2,\ldots,1$ or 0 [cf. Sec. II, remark (2)]. For any rank L, only one inequivalent tensor of rank L occurs, $Y^L_M(\mathbf{J})$ [cf. Sec. II, remark (3)]. The terms in U^d[SU(2)] homogeneous of degree d in the $\mathbf{J}_i$ can be written [cf. Sec. II, remark (4)] as follows:

$$A^{(d)}_{i_1\cdots i_d}\,\mathbf{J}_{i_1}\,\mathbf{J}_{i_2}\cdots\mathbf{J}_{i_d} = \sum A^{(d)}_{NLM}\,(\mathbf{J}\cdot\mathbf{J})^N\,Y^L_M(\mathbf{J}), \tag{3.2}$$

where the sum extends over all non-negative values of N and L such that $2N+L=d$ and $-L\leqslant M\leqslant +L$. The spherical tensor operators $Y^L_M(\mathbf{J})$ are obtained from the corresponding spherical functions $Y^L_M(\theta,\phi)$ by the substitutions $(\sin\theta)\,e^{\pm i\theta}\to\mathbf{J}_\pm$ and $\cos\theta\to\mathbf{J}_3$.

The spherical tensor operator $Y^L_M(\mathbf{J})$ can be constructed as a stretched, or symmetrized, L th-order tensor product based on $Y^1_M(\mathbf{J})$. The integrity basis for U[SU(2)] therefore consists of the two operators, $\mathbf{J}\cdot\mathbf{J}$ and $Y^1(\mathbf{J})$ [cf. Sec. II, remark (8)]. The generating function for the irreducible SU(2) tensor content of the SU(2) enveloping algebra is

$$G[D;\mathrm{SU(2)}] = \frac{1}{(1-D^2\mathbf{J}\cdot\mathbf{J})(1-D\,\mathbf{J})} \tag{3.3a}$$

$$= \sum m^{(d)}_{NL}\,D^d(\mathbf{J}\cdot\mathbf{J})^N(\mathbf{J})^L. \tag{3.3b}$$

Here $m^{(d)}_{NL}$ is the multiplicity of occurrence of the term $(\mathbf{J}\cdot\mathbf{J})^N\,Y^L(\mathbf{J})$ in the d th-order symmetrized tensor product of the basis operators $\mathbf{J}_i$ with themselves [cf. Sec. II, remark (9)].

The generating function (3.3) may be converted to a generating function for scalars in the symmetry subgroup as follows [cf. Sec. II, remark (11)]. The generating function is first expanded as a power series in D, the C_3 invariant operators in the expansion are retained while all other tensor operators are projected to zero, and the expansion is then finally resummed. The only C_3 scalar operators in U[SU(2)] are $\mathbf{J}\cdot\mathbf{J}$ and $Y^L_M(\mathbf{J})$, with M a multiple of 3 and $L\geqslant|M|$. An integrity basis for this set of operators is $\mathbf{J}\cdot\mathbf{J}$, $Y^1_0(\mathbf{J})\sim\mathbf{J}_3$, and $Y^3_{\pm 3}(\mathbf{J})$. The latter operators may be replaced by $\mathbf{J}^3_\pm$. The generating function for C_3 scalar operators in U[SU(2)] is thus

$$G(D;\mathrm{SU(2)}\supset\mathbf{C}_3) = \frac{1}{(1-D^2\mathbf{J}\cdot\mathbf{J})(1-D\,\mathbf{J}_3)} \times\left\{\frac{1}{1-D^3\mathbf{J}^3_+}+\frac{D^3\mathbf{J}^3_-}{1-D^3\mathbf{J}^3_-}\right\} \tag{3.4a}$$

$$= \sum m^{(d)}_{Nn_3n_+n_-}\,D^d(\mathbf{J}\cdot\mathbf{J})^N\mathbf{J}_3^{n_3}(\mathbf{J}^3_+)^{n_+}(\mathbf{J}^3_-)^{n_-}. \tag{3.4b}$$

Products of the form $\mathbf{J}^3_+\,\mathbf{J}^3_-$ are not functionally independent of the other members of the integrity basis. This product can be expressed in terms of $\mathbf{J}\cdot\mathbf{J}$ and $\mathbf{J}_3^2$. The structure of the terms within the curly brackets { } in (3.4a) ensures that only functionally independent terms are retained in the series expansion of the generating function [cf. Sec. II, remark (10)].

From (3.3) it follows that the Hamiltonian for a system with dynamical group SU(2) and symmetry subgroup C_3 has the form

$$\mathbf{H}(\mathbf{J}) = \sum A^{(d)}_{N,n_3,M}(\mathbf{J}\cdot\mathbf{J})^N(\mathbf{J}_3)^{n_3}(\mathbf{J}^3_\pm)^{|M|}, \tag{3.5}$$

with $\mathbf{J}_+$ for $M>0$, $\mathbf{J}_-$ for $M<0$, and M a multiple of 3. The Hermiticity requirement on the Hamiltonian places the following reality restrictions on the coefficients: $A^{(d)}_{N,n_3,M} = A^{(d)*}_{N,n_3,-M}$.

Point group tensor harmonics have been used to construct Hamiltonians for systems with crystal and ligand field symmetries,[10] but such constructions have not been coupled with the use of integrity bases and generating functions.

B. SU(2) $\supset$ U(1)

In this case the enveloping algebra remains unchanged. However, the subgroup scalar operators change because the symmetry subgroup is different. So whereas results (3.1)–(3.3) still apply, (3.4) and (3.5) must be replaced by

$$G[D;\mathrm{SU(2)}\supset\mathbf{U(1)}] = \frac{1}{(1-D^2\mathbf{J}\cdot\mathbf{J})(1-D\,\mathbf{J}_3)} \tag{3.6a}$$

$$= \sum m^{(d)}_{Nn_3}(\mathbf{J}\cdot\mathbf{J})^N\,\mathbf{J}_3^{n_3}, \tag{3.6b}$$

$$\mathbf{H}(\mathbf{J}) = \sum A^{(d)}_{Nn_3}(\mathbf{J}\cdot\mathbf{J})^N J_3^{n_3}, \tag{3.7}$$

with all coefficients in (3.7) real. The integrity basis for U(1) scalars is smaller than that for C_3 scalars since U(1) is larger than C_3: U(1) $\supset$ C_3. As a result the generating function (3.6) and expansion (3.7) are simpler than (3.4) and (3.5).

Quite often the Hamiltonian is computed within a single invariant subspace of G. In such a space the values of the dynamical group Casimir operators are fixed and can be neglected. For example, in a fixed-J space the Hamiltonian for SU(2) $\supset$ U(1) systems assumes the simple form

$$\mathbf{H}(\mathbf{J}) = \sum A_L \, Y_0^L(\mathbf{J}). \tag{3.8}$$

Such a Hamiltonian has been used to describe the Morse oscillator.[7] Up to terms of degree 2, the Hamiltonian is

$$\begin{aligned}\mathbf{H}_{\text{Morse}} &= A_0 + A_1 \, Y_0^1(\mathbf{J}) + A_2 Y_0^2(\mathbf{J}) \\ &= B_0 + B_1 \, \mathbf{J}_3 + B_2 \mathbf{J}_3^2.\end{aligned} \tag{3.9}$$

C. SU(2) ⊗ SU(2) ⊃ U(1)

The Lie algebra for this direct product group is the direct sum algebra spanned by the six generators $\mathbf{J}_i(\alpha)$, $\alpha = 1,2$; $i = +,3,-$. The generators of the two SU(2) subgroups ($\alpha = 1,2$) commute. The symmetry subgroup U(1) is generated by $\mathbf{J}_3(1) + \mathbf{J}_3(2)$.

The generating function for the irreducible tensor content of the SU(2) ⊗ SU(2) enveloping algebra is

$$\begin{aligned}&G[D;\mathbf{SU(2)} \otimes \mathbf{SU(2)}] \\ &\quad = \{[1 - D^2\mathbf{J}(1)\cdot\mathbf{J}(1)][1 - D^2\mathbf{J}(2)\cdot\mathbf{J}(2)] \\ &\qquad \times [1 - D(\mathbf{J}(1) + \mathbf{J}(2))]\}^{-1} \\ &\quad = \sum m^{(d)}_{N_1N_2L_1L_2} \, D^d \, [\mathbf{J}(1)\cdot\mathbf{J}(1)]^{N_1} \, [\mathbf{J}(2)\cdot\mathbf{J}(2)]^{N_2} \\ &\qquad \times [\mathbf{J}(1)]^{L_1} [\mathbf{J}(2)]^{L_2},\end{aligned} \tag{3.10}$$

where $d = 2N_1 + 2N_2 + L_1 + L_2$. From this, the generating function for U(1) scalars in U[SU(2) ⊗ SU(2)] can be constructed. The result is

$$\begin{aligned}&G[D;\mathbf{SU(2)} \otimes \mathbf{SU(2)} \supset \mathbf{U(1)}] \\ &\quad = \prod_{\alpha=1}^{2} G_\alpha[D;\mathbf{SU(2)} \supset \mathbf{U(1)}] \\ &\qquad \times G_{12}[D;\mathbf{SU(2)} \otimes \mathbf{SU(2)} \supset \mathbf{U(1)}],\end{aligned} \tag{3.11}$$

where each $G_\alpha[D;\mathbf{SU(2)} \supset \mathbf{U(1)}]$ has the form (3.6a) and

$$\begin{aligned}&G_{12}[D;\mathbf{SU(2)} \otimes \mathbf{SU(2)} \supset \mathbf{U(1)}] \\ &\quad = \left\{\frac{1}{1 - D^2\mathbf{J}_+(1)\mathbf{J}_-(2)} + \frac{D^2\mathbf{J}_-(1)\mathbf{J}_+(2)}{1 - D^2\mathbf{J}_-(1)\mathbf{J}_+(2)}\right\}.\end{aligned} \tag{3.12}$$

The integrity basis for U(1) scalars in the enveloping algebra of SU(2) ⊗ SU(2) consists of $\mathbf{J}(\alpha)\cdot\mathbf{J}(\alpha)$, $\mathbf{J}_3(\alpha)$ ($\alpha = 1,2$), $\mathbf{J}_+(1)\mathbf{J}_-(2)$ and $\mathbf{J}_-(1)\mathbf{J}_+(2)$, with the understanding that cross terms involving the last two terms $\mathbf{J}_+(1)\mathbf{J}_-(2)$ and $\mathbf{J}_-(1)\mathbf{J}_+(2)$ are not functionally independent of the remaining terms.

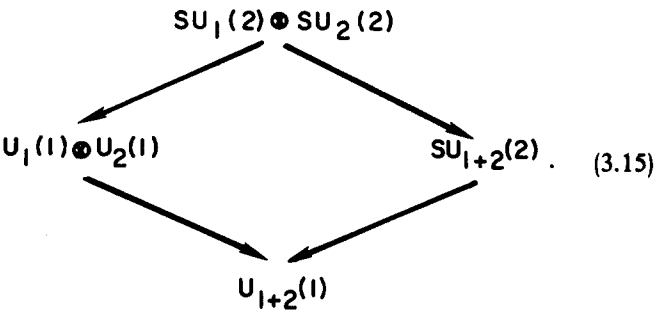
The most general Hamiltonian for a system with dynamical group→symmetry subgroup structures SU(2) ⊗ SU(2) ⊃ U(1) is

$$\begin{aligned}&\mathbf{H}[\mathbf{J}(1),\mathbf{J}(2)] \\ &\quad = \sum A^{(d)}_{N_1N_2n_3n_3'k_3k_3'} \, [\mathbf{J}(1)\cdot\mathbf{J}(1)]^{N_1} \, [\mathbf{J}(2)\cdot\mathbf{J}(2)]^{N_2} \\ &\qquad \times \mathbf{J}_3(1)^{n_3}\mathbf{J}_3(2)^{n_3'}[\mathbf{J}_+(1)\mathbf{J}_-(2)]^{k_3}[\mathbf{J}_-(1)\mathbf{J}_+(2)]^{k_3'},\end{aligned} \tag{3.13}$$

where either $k_3 = 0$ or $k_3' = 0$. The reality conditions required by the Hermiticity of H are

$$A^{(d)}_{N_1N_2n_3n_3'k_3k_3'} = A^{(d)*}_{N_1N_2n_3n_3'k_3'k_3}. \tag{3.14}$$

In an effort to reduce the number of parameters in the Hamiltonian (3.13), a unitary transformation may be applied. The most general suitable transformation is $\exp\{i[\phi_1\mathbf{J}_3(1) + \phi_2\mathbf{J}_3(2)]\}$. At best this can transform two complex parameters to real values. Since the number of complex parameters in (3.13) of degree $d = 1,2,3,4,\ldots$ is $0,1,2,6,\ldots$, only Hamiltonians of maximum degree 2 can be expressed in terms of the operators $\mathbf{J}(\alpha)\cdot\mathbf{J}(\alpha)$, $\mathbf{J}_3(\alpha)$, and $\mathbf{J}_+(1)\mathbf{J}_-(2) + \mathbf{J}_-(1)\mathbf{J}_+(2) = 2[\mathbf{J}(1)\cdot\mathbf{J}(2) - \mathbf{J}_3(1)\mathbf{J}_3(2)]$. These five operators exist in 1–1 correspondence with the invariant operators in the following chains of dynamical subgroups:



(3.15)

As a result, the integrity basis in (3.13) can be replaced by invariant operators from dynamical group chains only for Hamiltonians of maximum degree 2 and then only because of the possibility of performing a unitary transformation. These group chains have been used to describe the spectra of triatomic molecules.[11]

IV. SU(3) ⊃ SO(3)

In this section we consider SU(3) as the dynamical group and SO(3) as its symmetry subgroup. This is the underlying structure of the Elliott model.[12,13] It also occurs as a subalgebra in the vibron and interacting boson models (Secs. V and VI, ahead). The embedding of SO(3) in SU(3) is defined by the irreducibility of the defining three-dimensional representation of SU(3) on restriction to the subgroup. Two cases are considered, the generic case in which U[SU(3)] acts on an invariant Hilbert space of arbitrary symmetry type, and the degenerate case in which only symmetric (bosonlike) Hilbert spaces are considered.

A. Generic case

The tensor content of the enveloping algebra of SU(3) is obtained by reducing the symmetric tensor product of the adjoint (octet) representation, $[f] = [21] = (1,1)$. This has been done for low degree by using the method of S functions, or plethysms.[14] The number of inequivalent tensors of type (p,q), of dimensions $\frac{1}{2}(p+1)(q+1)(p+q+2)$, that occur in U[SU(3)] is equal to $1 + \min(p,q)$ if $p + 2q = 0 \pmod 3$ and zero otherwise [cf. Sec. II, remark (5)]. Thus there are two inequivalent octets, three 27-plets (2,2), one each of the decouplets (3,0) and (0,3), etc.

In describing tensor types, we adopt the Cartan notation (μ) rather than the Young partition notation $[f]$. In this notation

$$\mu_i = f_i - f_{i+1} = 2(\alpha_i, M^h)/(\alpha_i, \alpha_i). \tag{4.1}$$

Here f_i is the length of the ith row in the Young partition and μ_i is the overhang between the ith and $(i+1)$st row. Further, M^h is the highest weight in the representation space

$[M^h = \mathbf{f}$ for $U(N), SO(N), USp(N)]$, and the α_i are the highest weights of the l fundamental irreducible representations of the simple Lie group. The integer μ_i is the number of times the ith fundamental irreducible representation occurs in the stretched tensor product from which the representation with highest weight M^h is constructed. By adopting this notation, we bring the construction of representations and of irreducible tensorial sets into close analogy with one another [cf. Sec. II, remark (8)].

The generating function for the irreducible tensor content in the SU(3) enveloping algebra is[5]

$$\begin{aligned} G[D;\mathbf{SU(3)}] &= [(1-D^2\mathbf{C}^2)(1-D^3\mathbf{C}^3) \\ &\times(1-D\,\mathbf{U}_1\mathbf{U}_2)(1-D^2\mathbf{U}_1\mathbf{U}_2)]^{-1} \\ &\times\left\{\frac{1}{1-D^3\mathbf{U}_1^3}+\frac{D^3\mathbf{U}_2^3}{1-D^3\mathbf{U}_2^3}\right\} \qquad (4.2) \\ &= \sum m_{pq}^{(d)} D^d \mathbf{U}_1^p \mathbf{U}_2^q, \qquad (4.3) \end{aligned}$$

where $m_{pq}^{(d)}$ is the multiplicity of occurrence of tensors of type (p,q) in U^d [SU(3)]. Here $\mathbf{C}^2$ and $\mathbf{C}^3$ are the second and third Casimir operators of SU(3), $D\,\mathbf{U}_1\mathbf{U}_2$ represents symbolically the adjoint representation, $D^2\mathbf{U}_1\mathbf{U}_2$ represents the octet, which occurs in the symmetrized square of the adjoint representation (Gell-Mann's d coupling[15]), and $D^3\mathbf{U}_1^3$ and $D\,\mathbf{U}_2^3$ represent the 10 and $\overline{10}$, or (3,0) and (0,3), representations. Only five of these tensor operators are functionally independent; in particular, the cube of $D^2\mathbf{U}_1\mathbf{U}_2$ is a polynomial in the remaining operators.[5]

The SO(3) scalar content of U[SU(3)] can be obtained as described in Sec. III. The generating function (4.2) is expanded, the SO(3) scalar operators within each SU(3) tensor are projected out, and the expansion is resummed. A SU(3) tensor of type (p,q) contains a SO(3) scalar exactly once if p and q are both even, zero otherwise. The operators $\mathbf{C}^2,\mathbf{C}^3$ are SO(3) scalars. The representations (2,2) obtained from the two octets each contain one SO(3) scalar

$$\begin{aligned} &(D\,\mathbf{U}_1\mathbf{U}_2)\times(D\,\mathbf{U}_1\mathbf{U}_2)\to D^2\mathbf{U}_1^2\mathbf{U}_2^2 \supset \mathbf{J}\cdot\mathbf{J}, \\ &(D\,\mathbf{U}_1\mathbf{U}_2)\times(D^2\mathbf{U}_1\mathbf{U}_2)\to D^3\mathbf{U}_1^2\mathbf{U}_2^2 \supset \mathbf{X}_3, \\ &(D^2\mathbf{U}_1\mathbf{U}_2)\times(D^2\mathbf{U}_1\mathbf{U}_2)\to D^4\mathbf{U}_1^2\mathbf{U}_2^2 \supset \mathbf{X}_4. \qquad (4.4) \end{aligned}$$

An additional independent SO(3) scalar can be formed from the square of the 10 or the $\overline{10}$

$$\begin{aligned} &(D^3\mathbf{U}_1^3)^2\to D^6\mathbf{U}_1^6 \\ &\qquad\qquad \supset \mathbf{X}_6\,. \qquad (4.5) \\ &(D^3\mathbf{U}_2^3)^2\to D^6\mathbf{U}_2^6 \end{aligned}$$

The generating function for SO(3) scalars in U[SU(3)] is[16]

$$\begin{aligned} G[D;\mathbf{SU(3)} \supset \mathbf{SO(3)}] &= [(1-D^2\mathbf{C}^2)(1-D^3\mathbf{C}^3)(1-D^2\mathbf{J}\cdot\mathbf{J})]^{-1} \\ &\times\frac{(1+D^6\mathbf{X}_6)}{(1-D^3\mathbf{X}_3)(1-D^4\mathbf{X}_4)}\,. \qquad (4.6) \end{aligned}$$

The square of the operator $\mathbf{X}_6$ is a polynomial in the remaining five members of the integrity basis.

The effect of discrete transformations [cf. Sec. II, remark (12)] on the Hamiltonian H [SU(3) $\supset$ SO(3)] is determined by considering the effect of P and T on the generators $\mathbf{X}_i$ of SU(3). All eight generators are invariant under P, which therefore imposes no additional constraints on the Hamiltonian. The effect of the time reversal operator T is nontrivial. The generators of SU(3) can be written as an $L=1$ spherical tensor (L) and an $L=2$ spherical tensor (Q), with eigenvalues $(-1)^L$ under T:

$$\mathbf{SU(3)} = \mathbf{L}+\mathbf{Q} \overset{T}{\to} (-1)\mathbf{L}+(+1)\mathbf{Q}. \qquad (4.7)$$

The six operators which appear in (4.6) can be expressed in terms of these spherical tensor operators as follows[16]:

$$\begin{aligned} &\mathbf{C}^2 = (\tfrac{3}{4})^2(\mathbf{L}\cdot\mathbf{L}+2\mathbf{Q}\cdot\mathbf{Q}), \\ &\mathbf{C}^3 = \mathbf{L}\cdot\mathbf{Q}\cdot\mathbf{L}-\tfrac{4}{3}\mathbf{Q}\cdot\mathbf{Q}\cdot\mathbf{Q}, \\ &\mathbf{J}\cdot\mathbf{J} = \mathbf{L}\cdot\mathbf{L}, \qquad (4.8) \\ &\mathbf{X}_3 = \mathbf{L}\cdot\mathbf{Q}\cdot\mathbf{L}, \\ &\mathbf{X}_4 = \mathbf{L}_a\mathbf{Q}_{ab}\mathbf{Q}_{bc}\mathbf{L}_c, \\ &\mathbf{X}_6 = \epsilon_{abc}\mathbf{Q}_{bd}\mathbf{Q}_{ce}\mathbf{Q}_{ef}\mathbf{L}_a\mathbf{L}_d\mathbf{L}_f. \end{aligned}$$

Of these, only the operator $\mathbf{X}_6$ goes into its negative under time reversal ($T\mathbf{X}_6T^{-1} = -\mathbf{X}_6$). As a result, if it appears in the Hamiltonian it must be multiplied by an imaginary number. The remaining five operators are positive under time reversal and so each must be multiplied by a real parameter when appearing in a time reversal invariant Hamiltonian.

The Hermiticity requirement imposes no additional reality conditions on the coefficients $A^{(d)}_{i_1\cdots i_d}$ in the Hamiltonian. The first five SO(3) scalars in (4.8) are self-adjoint and $\mathbf{X}_6$ is anti-Hermitian. As a result, any term in the Hamiltonian containing the operator $(\mathbf{X}_6)^{n_6}$ must be multiplied by a real or imaginary coefficient, depending on whether n_6 is zero or one. No other values of n_6 are possible by (4.6).

A generating function for the *number* of SO(3) scalars, $n^{(d)}$, of degree d is

$$G(D) = \frac{1+D^6}{(1-D^2)^2(1-D^3)^2(1-D^4)} = \sum n^{(d)}D^d. \qquad (4.9)$$

To terms of degree $d=4$ a SU(3) Hamiltonian with SO(3) symmetry contains $2+2+4=8$ terms: $\mathbf{J}^2$, $\mathbf{C}^2$; $\mathbf{X}_3$, $\mathbf{C}^3$; $(\mathbf{J}^2)^2$, $\mathbf{J}^2\mathbf{C}^2$, $(\mathbf{C}^2)^2$, $\mathbf{X}_4$. Acting within a single SU(3) representation, four of these terms become redundant. The resulting four-parameter Hamiltonian has been studied extensively over a wide class of rotational nuclei.[17–20]

Less the operator $\mathbf{X}_6$ and the Casimir invariants for the dynamical group SU(3) and the symmetry subgroup SO(3), there are two SO(3) scalars. This is twice the number of missing labels in the reduction of SU(3) to SO(3), in agreement with the missing label folk theorem.

B. Degenerate representations

Degenerate representations are constructed from fewer than the full complement of fundamental representations possessed by a semisimple Lie group of rank l. As a result, one or more of their Cartan representation labels μ_i vanish. For SU(3) the degenerate representations have labels $(p,0)$ or

$(0,q)$. They are symmetric representations based on the 3 or $\bar{3}$ fundamental irreducible representation.

Under reduction from generic to degenerate representations, simplifications occur in the enveloping algebra. These simplifications are of three types [cf. Sec. II, remark (6)].[4]

(i) The Casimir operators are no longer independent. In the present case $\mathbf{C}^3$ is a function of $\mathbf{C}^2$.

(ii) Some tensor operators have only vanishing matrix elements within degenerate representations. For example, tensor operators of type $(p,0)$ or $(0,q)$ have only vanishing matrix elements within degenerate representations [cf. Sec. II, remark (3)].

(iii) Linear dependences occur among formerly independent operators of the same tensor type when restricted to particular classes of degenerate representations. For example, the two octets $D\,\mathbf{U}_1\mathbf{U}_2$ and $D^2\mathbf{U}_1\mathbf{U}_2$ become proportional within the representation $(p,0)$ or $(0,q)$. This can be seen, for example, by computing their expectation values within the coherent state representation.[21,22]

The generating function for SU(3) tensors in U[SU(3)] acting on degenerate representations is[4]

$$G(D;\mathbf{U}_1\mathbf{U}_2) = [(1 - D^2\mathbf{C}^2)(1 - D\,\mathbf{U}_1\mathbf{U}_2)]^{-1}. \tag{4.10}$$

From this, the generating function for SO(3) scalars is easily derived as

$$G[D;\mathbf{SU(3)} \supset \mathbf{SO(3)}] = [(1 - D^2\mathbf{C}^2)(1 - D^2\mathbf{J}^2)]^{-1}. \tag{4.11}$$

Within a single representation $\mathbf{C}^2$ can be neglected. The Hamiltonian then becomes simply a function of the total angular momentum operator $\mathbf{J}^2$. Invariance under time reversal and Hermitian conjugation requires that the Hamiltonian be a *real* function of $\mathbf{J}^2$. The absence of additional SO(3) scalars besides the Casimir operators of SU(3) and SO(3) is consistent with the missing label folk theorem.[6]

The problem of replacing an integrity basis by the invariant operators of a group–subgroup lattice has a clear-cut solution when the dynamical group is SU(3) and the symmetry group is SO(3). In this case there is only one lattice,

$$\begin{array}{c}\mathbf{SU(3)}\\ \Big\downarrow\\ \mathbf{SO(3)}\end{array} . \tag{4.12}$$

In the generic case, this replacement is possible only for terms up to degree $d = 2$. For $d = 3$ the group lattice provides three scalars while the integrity basis requires four. For $d = 4$ the group lattice fails to provide yet another scalar $(\mathbf{X}_4)$. In the case of degenerate representations, the Casimir operators of SU(3) and SO(3) coincide precisely with the integrity basis. As a result, the integrity basis can be replaced by the group–subgroup lattice.

V. SU(4) ⊃ SO(3)

Diatomic[23] and triatomic[24] molecules have recently been described in terms of the dynamical group SU(4). The basic bond excitations that are important are assumed to have quantum numbers $J^\pi = 1^-, 0^+$. These are created by operators $b_\mu^+ = (\pi_m^+, \sigma^+)$, $m = +1,0,-1$. The dynamical group associated with each bond is generated by the bilinear number conserving operator products $b_\mu^+ b_\nu$. These 16 opeators generate U(4). Removing the first-order Casimir invariant $\Sigma_\mu\, b_\mu^+ b_\mu$ yields 15 generators of the simple Lie group SU(4). This group acts in a Hilbert space, which carries a completely symmetric, or boson, representation of SU(4) with representation labels $(N,0,0)$.

In the event that several (covalent) bonds are present in a molecule, the dynamical group is assumed to be a direct product of SU(4) dynamial groups, one for each bond. That is, $G = \mathrm{SU}_1(4) \otimes \mathrm{SU}_2(4) \otimes \cdots \otimes \mathrm{SU}_k(4)$, where k is the number of bonds. The $15k$ generators of G are $b_\mu^+(i)b_\nu(i) - \frac{1}{4}[\,\Sigma_\gamma\, b_\gamma^+(i)b_\gamma(i)\,]\delta_{\mu\nu}$, where $1 \leqslant \mu, \nu \leqslant 4$, and $1 \leqslant i \leqslant k$. The dynamical subgroup for the ith bond, $\mathrm{SU}_i(4)$, acts in a bosonic Hilbert space, which carries a completely symmetric representation $(N_i,0,0)$ of SU(4). The total Hilbert space is the direct product of each of these completely symmetric subspaces, one for each bond.

The semisimple direct product dynamical group G contains a "collective" or "diagonal" subgroup $\mathrm{SU}_D(4)$, with generators $B_\mu^+ B_\nu - \frac{1}{4}[\,\Sigma_\gamma\, B_\gamma^+ B_\gamma\,]\delta_{\mu\nu}$, where B_μ^+ is the collective operator $B_\mu^+ = \Sigma_{i=1}^k\, b_\mu^+(i)$. Under restriction of G to $\mathrm{SU}_D(4)$, the Hilbert space on which G acts reduces to a direct sum over SU(4) invariant subspaces with Young partitions containing $N_D = \Sigma\, N_i$ boxes and no more than k or four rows, whichever is smaller.

The transformation properties of the generators of U(4) under the proper rotation group SO(3), the order-2 group symmetries P (space reflection), T (time reversal), and Hermitian conjugation, are determined as follows. The U(4) generators are expressed as spherical tensor operators, using vector coupling coefficients where necessary:

$$\begin{aligned} &T_M^{(L)} = (\pi^+\tilde{\pi})_M^{(L)} = \sum_{m,(m')} \langle 1,m;1,m'|L,M\rangle\, \pi_m^+\, \tilde{\pi}_{m'},\\ &S = \sigma^+\tilde{\sigma},\\ &P_+ = \pi^+\tilde{\sigma} + \sigma^+\tilde{\pi}, \qquad P_- = i(\pi^+\tilde{\sigma} - \sigma^+\tilde{\pi}). \end{aligned} \tag{5.1}$$

The operators $\tilde{\pi},\tilde{\sigma}$ are related to π,σ by

$$\tilde{\pi}_m = (-)^m \pi_{-m}, \quad \tilde{\sigma} = \sigma. \tag{5.2}$$

The operators $\tilde{\pi}$, rather than π, transform under $D^1[\mathrm{SO(3)}]$. The operators $T^{(0)}$ and S are rotational scalars, $T^{(1)}$ and $P_\pm$ are vectors, and $T^{(2)}$ is a spherical tensor of rank 2. Since $P(\pi,\sigma)P^{-1} = (-\pi, +\sigma)$, $T^{(L)}$ and S are even and $P_\pm$ are odd under space inversion. And because $T(\pi,\sigma)T^{-1} = (\tilde{\pi},\tilde{\sigma})$, $T^{(1)}$ and P_- are odd under time reversal and the remaining operators are even. All irreducible tensorial sets are Hermitian in the sense they obey $O_M^{L+} = (-1)^M O^L{}_{-M}$. Notice that the choice (5.2) for the transformation properties of the π operators under time reversal differs from the standard transformation $O_M^L = (-1)^{L-M} O^L{}_{-M}$ for operators based on angular momentum (odd under time reversal). This choice was made because the π operators are associated with dipole displacements, which are even under time reversal.

The effect of the order-2 discrete operations, P,T is summarized in Table I. They effect a decomposition of the Lie algebra G into two subspaces

$$G = S_e + S_o, \tag{5.3}$$

where S_e is even and S_o is odd under the discrete operation.

TABLE I. Effect of discrete operations on generators of U(4).

	S_e:Even	S_o:Odd	Subgroup	Quotient
P→Space inversion	$T^{(0)},T^{(1)},T^{(2)},S$	P_+,P_-	U(3) ⊗ U(1)	SU(4)/U(3)
T→Time reversal	$T^{(0)},T^{(2)},P_+,S$	$T^{(1)},P_-$	SO(4)	U(4)/SO(4)

For the parity operation, S_e forms a subalgebra in G, while for time reversal the odd subspace S_o forms a subalgebra in G. The reason is as follows: The group elements in U(4) are of the form $\exp(i\mathbf{X})$, where $\mathbf{X}\in\mathbf{G}$. Under time reversal

$$T[\exp(i\mathbf{X})]T^{-1}=\exp[T(i\mathbf{X})T^{-1}]=\exp[-i(T\mathbf{X}T^{-1})]. \tag{5.4}$$

The operation T leaves invariant the subset of elements in G generated by the X's, which obey $T\mathbf{X}T^{-1}=-\mathbf{X}$. A similar argument holds for Hermitian conjugation.

A. Diatomic molecules

For molecules with a single bond, the dynamical group SU(4) acts only within fully symmetric representations. This is a very degenerate class of representations. Accordingly, the generating function for the SU(4) irreducible tensor content of U[SU(4)] is relatively simple,[4]

$$G[D;\mathbf{SU(4)}]=[(1-D^2\mathbf{C}^2)(1-D\,\mathbf{U}_1\mathbf{U}_3)]^{-1}\,. \tag{5.5}$$

Here $D\,\mathbf{U}_1\mathbf{U}_3$ represents the adjoint representation of SU(4) [i.e., the generators of SU(4)] and $D^2\mathbf{C}^2$ represents the second-order Casimir operator of SU(4). Since this is constant within any representation and only one representation is considered for diatomic systems, this part of the generating function will be ignored in the remainder of this subsection.

Next, it is necessary to determine the SO(3) content of the SU(4) irreducible tensors in the enveloping algebra. This can be obtained by determining the U(3) content of the SU(4) tensors (μ_1,μ_2,μ_3) that occur and then determining the SO(3) scalars (excluding pseudoscalars) in these U(3) representations. As the U(3) subalgebra $(T^{(0)},T^{(1)},T^{(2)})$ is even under space inversion, each U(3) irreducible representation in (μ_1,μ_2,μ_3) of SU(4) carries a good parity label. We are interested in the SO(3) scalars in positive parity representations of SU(3).

The adjoint representation (1,0,1) of SU(4) carries one 0^+ representation of SO(3), with basis operator $n_\pi=(\pi^+\tilde{\pi})^{(0)}$. The representation (2,0,2) of SU(4) [i.e., $(D\,\mathbf{U}_1\mathbf{U}_3)^2$] contains five positive parity representations of U(3). Four of these contain SO(3) scalars. Two of these scalars are self-adjoint and there is one Hermitian adjoint pair. The two self-adjoint representations of SU(3) are (0,0) and (2,2); they contain SO(3) scalars $(n_\pi)^2$ and $[(\pi^+\tilde{\pi})^{(2)}(\pi^+\tilde{\pi})^{(2)}]^{(0)}$. The Hermitian adjoint pair (2,0) and (0,2) contain the SO(3) scalars $[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}$ and $[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}$.

More generally, when $N=2K$ or $2K+1$ the representation $(N,0,N)$ of SU(4) contains $K+1$ positive parity self-adjoint SO(3) scalars and $K(K+1)/2$ pairs of Hermitian adjoint operators. The generating function for SO(3) scalars in U[SU(4)] is

$$\begin{aligned}G[D;&\mathbf{SU(4)}\supset\mathbf{SO(3)}]\\&=((1-Dn_\pi)\{1-D^2[(\pi^+\tilde{\pi})^{(2)}(\pi^+\tilde{\pi})^{(2)}]^{(0)}\})^{-1}\\&\times\left\{\frac{1}{1-D^2[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}}\right.\\&\left.+\frac{D^2[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}}{1-D^2[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}}\right\}.\end{aligned} \tag{5.6}$$

The integrity basis for SO(3) scalars in the SU(4) enveloping algebra consists of the four operators n_π, $[(\pi^+\tilde{\pi})^{(2)}(\pi^+\tilde{\pi})^{(2)}]^{(0)}$, $[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}$, and $[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}$.

The operators n_π and $[(\pi^+\tilde{\pi})^{(2)}(\pi^+\tilde{\pi})^{(2)}]^{(0)}$ are invariant under the three discrete operations: P, T, and † (Hermitian conjugation). The remaining two operators are even under P. The effect of time reversal and Hermitian conjugation is

$$\begin{aligned}&T\{\alpha[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}+\beta[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}\}T^{-1}\\&\quad=\alpha^*[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})^{(0)}]+\beta^*[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)},\end{aligned} \tag{5.7a}$$

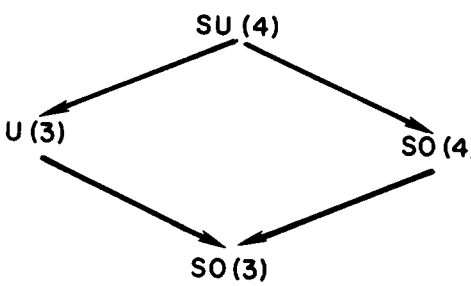
$$\begin{aligned}&\{\alpha[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]+\beta[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^0\}^\dagger\\&\quad=\beta^*[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}+\alpha^*[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}.\end{aligned} \tag{5.7b}$$

The effect of these two operations is to impose additional reality restrictions on the coefficients that appear in an expansion of the form (1.1). In particular, the coefficients of $[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}$ and $[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}$ must be real and equal. This condition causes modification of the generating function (5.6):

$$\begin{aligned}G[D;&\mathbf{SU(4)}\supset\mathbf{SO(3)}\otimes\mathbf{P}\otimes\mathbf{T}\otimes\dagger]\\&=((1-Dn_\pi)\{1-D^2[(\pi^+\tilde{\pi})^{(2)}(\pi^+\tilde{\pi})^{(2)}]^{(0)}\})^{-1}\\&\times(1-D^2\{[(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)}+[(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)}\})^{-1}.\end{aligned} \tag{5.8}$$

The integrity basis contains one first-degree operator and two second-degree operators. To second order, the Hamiltonian is a linear combination of the four operators $T^{(0)}$, $[T^{(L)}T^{(L)}]^{(0)}$ with $L=0$ and 2, and $[P_+P_+]^{(0)}$ with real coefficients.

These three operators may be replaced by Casimir operators for the group lattice



(5.9)

The generators for the subgroups are

$$\begin{array}{lccc} \mathrm{U}(3): & T^{(0)} & T^{(1)} & T^{(2)} \\ \mathrm{SO}(4): & & T^{(1)} & P_+ . \\ \mathrm{SO}(3): & & T^{(1)} & \end{array}$$

The generators for SO(4) are chosen as $T^{(1)}$ and P_+ rather than $T^{(1)}$ and P_- (which generate the time-reversed subgroup $\mathrm{SO}_T(4)$, see Table I) in order to yield wave functions that are real. The operator n_π is the first-order Casimir operator for U(3). Only two of the three operators $[(\pi^+\tilde{\pi})^{(L)}(\pi^+\tilde{\pi})^{(L)}]^{(0)}$, $L = 0,1,2$, are independent within fully symmetric representations. This is a consequence of the existence of only $K+1$ independent self-adjoint operators of degree $d = 2K$. The operator $n_\pi = \mathbf{C}^1[\mathrm{U}(3)](L = 0 \to n_\pi^2)$ is already included as an element in an integrity basis of Casimir operators. We may choose either of the other two operators ($L = 1,2$) as an additional member of the integrity basis. The operator with $L = 1$ is the second-order Casimir operator of the rotational subgroup SO(3) with generators $(\pi^+\tilde{\pi})^1_m$. Finally, we may make the identification

$$\begin{aligned} &\mathbf{C}^2[\mathrm{SO}(4)] - \mathbf{C}^2[\mathrm{SO}(3)] \\ &\quad = P_+P_+ \\ &\quad = [(\pi^+\tilde{\sigma})(\pi^+\tilde{\sigma})]^{(0)} + [(\sigma^+\tilde{\pi})(\sigma^+\tilde{\pi})]^{(0)} + 2n_\sigma n_\pi + 4. \end{aligned} \tag{5.10}$$

The generating function for even parity, time reversal invariant, Hermitian SO(3) scalars in the SU(4) enveloping algebra, in terms of the integrity basis of Casimir operators, is therefore

$$\begin{aligned} &G[D;\mathrm{SU}(4) \supset \mathrm{SO}(3) \otimes \mathbf{P} \otimes \mathbf{T} \otimes \dagger] \\ &\quad = (1 - D^2\mathbf{C}^2[\mathrm{SU}(4)])^{-1} \\ &\qquad \times \{(1 - D\mathbf{C}^1[\mathrm{U}(3)])(1 - D^2\mathbf{C}^2[\mathrm{SO}(4)]) \\ &\qquad \times (1 - D^2\mathbf{C}^2[\mathrm{SO}(3)])\}^{-1}. \end{aligned} \tag{5.11}$$

Thus, in the case of fully symmetric representations of SU(4), the Casimir operators for a group lattice are sufficient to construct all rotational scalars of any order.

B. Triatomic molecules

For molecules with two bonds, the dynamical group is $\mathrm{SU}_1(4) \otimes \mathrm{SU}_2(4)$. The $\mathrm{SU}_\alpha(4)$, $\alpha = 1,2$, act in Hilbert spaces H_1 and H_2, which carry fully symmetric representations $(N_1,0,0)$ and $(N_2,0,0)$ respectively. The direct product Hilbert space is reducible with respect to the diagonal subgroup $\mathrm{SU}_D(4)$. The irreducible content of this Hilbert space consists of representations with Young partitions $[f_1 = N_1 + N_2 - K, \quad f_2 = K, \quad f_3 = 0], \quad K = 0,1,2,\ldots, \min(N_1,N_2)$.

The irreducible tensor content in $\mathrm{U}[\mathrm{SU}_1(4) \otimes \mathrm{SU}_2(4)]$ is constructed from the generating function for the $\mathrm{SU}_D(4)$ irreducible tensor content in $\mathrm{U}[\mathrm{SU}_D(4)]$ acting on the degenerate representations in the class $[f_1,f_2,0]$. This generating function is[4]

$$\begin{aligned} &G[D;\mathrm{SU}(4)] \\ &\quad = [\underline{(1 - D^2\mathbf{C}^2)}(1 - D^3\mathbf{C}^3)\underline{(1 - D\mathbf{U}_1\mathbf{U}_3)}(1 - D^2\mathbf{U}_1\mathbf{U}_3)]^{-1} \\ &\qquad \times \frac{1}{(1 - D^2\mathbf{U}_2^2)}\left\{\frac{1}{(1 - D^3\mathbf{U}_1^2\mathbf{U}_2)} + \frac{D^3\mathbf{U}_2\mathbf{U}_3^2}{(1 - D^3\mathbf{U}_2\mathbf{U}_3^2)}\right\}. \end{aligned} \tag{5.12}$$

Two of these factors (underlined) occur in the generating function introduced in (5.5). The SO(3) scalars, which are derived from those two terms have a structure which has already been determined. The remaining terms in (5.12) give rise to SO(3) scalars not previously considered.

For concreteness, suppose one of the SO(3) scalars of degree $d = 2$ constructed in Sec. V A above for $\mathrm{SU}(4) \supset \mathrm{SO}(3)$ has the structure $B_{ij}\mathbf{X}_i\mathbf{X}_j$. Then a set of scalars for $\mathrm{SU}_1(4) \otimes \mathrm{SU}_2(4) \supset \mathrm{SO}(3)$ is

$$\begin{aligned} &B_{ij}[\mathbf{X}_i(1) + \mathbf{X}_i(2)][\mathbf{X}_j(1) + \mathbf{X}_j(2)] \\ &\quad = B_{ij}\mathbf{X}_i(1)\mathbf{X}_j(1) + B_{ij}\mathbf{X}_i(2)\mathbf{X}_j(2) + B_{ij}\{\mathbf{X}_i(1),\mathbf{X}_j(2)\}. \end{aligned} \tag{5.13}$$

The first two terms are the SO(3) scalars constructed for each bond separately. The third term, in which the bracket $\{\ ,\ \}$ indicates symmetrization with respect to the bond indices, is new and represents a rotationally invariant bond–bond interaction. This construction is perhaps most familiar for the SU(2) case, where for $\mathbf{J} = \mathbf{j}_1 + \mathbf{j}_2$ one has that $\mathbf{J}^2 = \mathbf{j}_1^2 + \mathbf{j}_2^2 + 2\mathbf{j}_1\cdot\mathbf{j}_2$.

In general, it is possible to construct operators invariant under the subgroup $H \subset G$ by constructing Casimir covariant operators of the direct product group $G_1 \otimes G_2$ (G_1,G_2 isomorphic to G). This construction proceeds as follows. If $\mathbf{C}^2 = B_{ij}\,\mathbf{X}_i\mathbf{X}_j$ is a Casimir invariant for G, one can construct the Casimir covariant operator $\mathbf{C}^2_{cc} = B_{ij}\{\mathbf{X}_i(1),\mathbf{X}_j(2)\}$ for $G_1 \otimes G_2$. This operator is not an invariant of G_1 or G_2 separately, but is an invariant for the direct product group. The construction of Casimir covariants extends easily to higher-order Casimir operators and larger numbers of isomorphic subgroups.

The interbond interactions of degree $d = 2$, which are constructed from the SO(3) scalars arising from the underlined terms in (5.12), can be written down immediately from the operators in (5.8) using the result (5.13):

$$\begin{aligned} &\{(\pi^+\tilde{\pi})^{(0)}_1,(\pi^+,\tilde{\pi})^{(0)}_2\}, \\ &\{(\pi^+\tilde{\pi})^{(2)}_1,(\pi^+,\tilde{\pi})^{(2)}_2\}^{(0)}, \\ &\{(\pi^+\tilde{\sigma})_1,(\pi^+\tilde{\sigma})_2\}^{(0)} + \{(\sigma^+\tilde{\pi})_1,(\sigma^+\tilde{\pi})_2\}^{(0)}. \end{aligned} \tag{5.14a}$$

The additional SO(3) scalars, which cannot be constructed from the operators in (5.6), are obtained from the remaining terms of the generating function (5.12). The additional SO(3) scalars of degree $d = 2$ are

$$\begin{aligned} &D^2\mathbf{U}_1\mathbf{U}_3 \to \{(\pi^+\tilde{\pi})^{(1)}_1,(\pi^+,\tilde{\pi})^{(1)}_2\}^{(0)}, \\ &D^2\mathbf{U}_2^2 \to \{(\pi^+\tilde{\sigma})_1,(\sigma^+\tilde{\pi})_2\}^{(0)} \quad \text{and} \quad \{(\sigma^+\tilde{\pi})_1,(\pi^+\tilde{\sigma})_2\}^{(0)}. \end{aligned} \tag{5.14b}$$

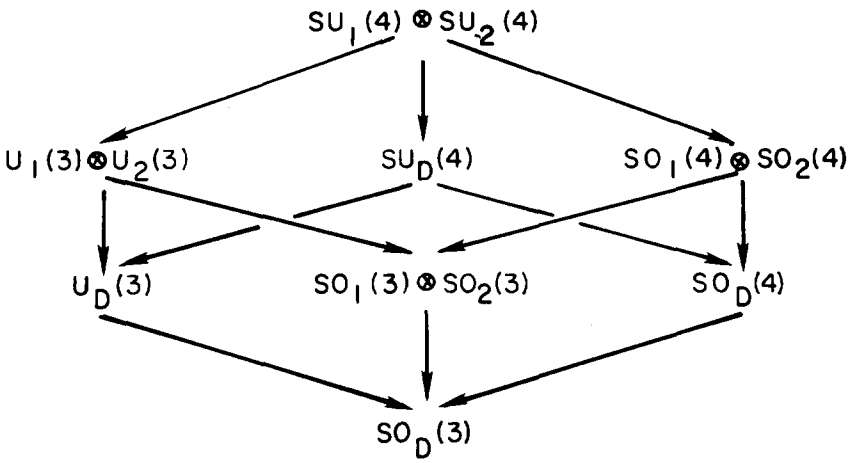
The discrete operations of time reversal and Hermitian conjugation impose constraints on the coefficients of the two terms arising from $D^2\mathbf{U}_2^2$. The independent SO(3) scalars involving both the π's and σ's (5.14a) and (5.14b) may be conveniently rewritten in terms of the $P_\pm$ operators of (5.1). The five interbond interactions of degree $d = 2$ are then

$$\begin{aligned} &\{T^{(L)}_1, T^{(L)}_2\}^{(0)}, \quad \text{with } L = 0,1,2, \\ &\{P_\sigma(1), P_\sigma(2)\}^{(0)}, \quad \sigma = +,-. \end{aligned} \tag{5.14c}$$

The total number of terms of degree up to 2 in the Hamiltonian with dynamical and symmetry groups $\mathrm{SU}(4) \otimes \mathrm{SU}(4) \supset \mathrm{SO}(3)$ is 13. Each bond contributes one

linear and three quadratic terms, while the bond–bond interaction introduces five additional terms. These operators do not form an integrity basis of SO(3) scalars in U[SU(4) ⊗ SU(4)] since new tensors occur for degree 3 (e.g., D^3C^3, $D^3U_1^2U_2$, and $D^3U_2U_3^2$).

It is possible to replace the SO(3) scalars in SU(4) ⊗ SU(4) up to degree 2 by the Casimir operators belonging to a group–subgroup lattice. A number of lattices are possible. We present one lattice which is particularly well matched to the SO(3) scalar operators given in (5.14):



(5.15)

In this group–subgroup lattice there are three types of subgroup restrictions:

$$\mathbf{G}_1 \oplus \mathbf{G}_2 \supset \mathbf{H}_1 \oplus \mathbf{H}_2, \tag{5.16a}$$

$$\mathbf{G} \oplus \mathbf{G}_2 \supset \mathbf{G}_D, \tag{5.16b}$$

$$\mathbf{G}_D \supset \mathbf{H}_D. \tag{5.16c}$$

Here $\mathbf{G}_i$ is a Lie algebra of operators for bond i, $\mathbf{H}_i$ is a subalgebra, and $(\mathbf{H}_D)\mathbf{G}_D$ is a diagonal (sub)algebra whose basis vectors are direct sum operators of the form $\mathbf{X}_i(D) = \mathbf{X}_i(1) + \mathbf{X}_i(2)$, where $\mathbf{X}_i(\alpha)$ is a boson number-conserving operator for bond α.

The group to subgroup restrictions of the form (5.16a) generate SO(3) scalars for each bond separately. The operators so generated have the form (5.8) or with the identification $1 \leftrightarrow 2$, (5.14a). The group to diagonal-subgroup restrictions of the form (5.16b) generate SO(3) scalars describing bond–bond interactions. For example, the difference between the second-order Casimir operator of $U_D(3)$ and those of $U_1(3)$ and $U_2(3)$ is a sum of cross terms given in (5.14c):

$$\mathbf{C}^2[\mathrm{U}_D(3)] - \{\mathbf{C}^2[\mathrm{U}_1(3)] + \mathbf{C}^2[\mathrm{U}_2(3)]\}$$

$$= \sum_{L=0,1,2} \{T_1^{(L)}, T_2^{(L)}\}^{(0)}. \tag{5.17}$$

[Note that with $i \neq j$ all three of the operators $\{T_i^{(L)}, T_j^{(L)}\}^{(0)}$ are independent while if $i = j$ only two are. This can be seen by expressing these operators as linear combinations of the operators $\{(\pi_i^+\ \pi_j^+)^{(L)}, (\tilde{\pi}_i\ \tilde{\pi}_j)^{(L)}\}^{(0)}$ using Racah recoupling techniques and noting that the operator with $L = 1$ vanishes by symmetry if $i = j$.] And finally, the diagonal group to diagonal subgroup restrictions of the form (5.14c) are useful for expressing the bond–bond scalar interactions as linear combinations of Casimir operators:

$$\mathbf{C}^2[\mathrm{SO}_D(4)] - \mathbf{C}^2[\mathrm{SO}_D(3)]$$

$$- \sum_{\alpha=1}^{2} \{\mathbf{C}^2[\mathrm{SO}_\alpha(4)] - \mathbf{C}^2[\mathrm{SO}_\alpha(3)]\}$$

$$= \{P_+(1), P_+(2)\}^{(0)}, \tag{5.18a}$$

$$\mathbf{C}^2[\mathrm{SU}_D(4)] - \mathbf{C}^2[\mathrm{U}_D(3)]$$

$$- \sum_{\alpha=1}^{2} \{\mathbf{C}^2[\mathrm{SU}_\alpha(4)] - \mathbf{C}^2[\mathrm{U}_\alpha(3)]\}$$

$$= \tfrac{1}{2}\{P_+(1), P_+(2)\}^{(0)} + \tfrac{1}{2}\{P_-(1), P_-(2)\}^{(0)}. \tag{5.18b}$$

Additional chains in the group–subgroup lattice (5.15) are possible whenever a mapping M of the Lie algebra G into itself exists whose square is the identity (involutive automorphism).[25] The Lie algebra G then splits into eigenspaces of M with eigenvalues $+1$ and -1,

$$\mathbf{G} = \mathbf{K} \oplus \mathbf{P} \overset{MGM^{-1}}{\rightarrow} \mathbf{K} \ominus \mathbf{P}, \tag{5.19}$$

restriction (5.16a) can be replaced by an alternative restriction

$$\mathbf{G}_1 \oplus \mathbf{G}_2 = (\mathbf{K}_1 \oplus \mathbf{P}_1) \oplus (\mathbf{K}_2 \oplus \mathbf{P}_2)$$

$$\rightarrow \mathbf{G}_D^* = (\mathbf{K}_1 \oplus \mathbf{K}_2) \oplus (\mathbf{P}_1 \ominus \mathbf{P}_2). \tag{5.20}$$

This means that if the $\mathbf{X}_i(\alpha)$ are even and the $\mathbf{Y}_j(\alpha)$ are odd under M ($\alpha = 1,2$), the generators $\mathbf{X}_i^*(D)$ and $\mathbf{Y}_j^*(D)$ are

$$\mathbf{X}_i^*(D) = \mathbf{X}_i(1) + \mathbf{X}_i(2), \quad \mathbf{Y}_i^*(D) = \mathbf{Y}_j(1) - \mathbf{Y}_j(2). \tag{5.21}$$

Each of the four group–subgroup reductions contained in (5.9) obeys this automorphism property. The operators obtained by including group–subgroup chains formed in this way are not independent of those occurring in the chains already present in (5.15).

For symmetric triatomic molecules (e.g., CO_2) it might be expected that the quantum numbers N_1 and N_2 describing the two bonds are equal. (In fact, spectral data have been

fitted with $N_1 \simeq 2N_2$.[22] In the case that $N_1 = N_2$, one of the representations that occurs in $SU_D(4) \supset SU_1(4) \otimes SU_2(4)$ is of the degenerate type $(\mu_1 = 0, \mu_2 = N, 0)$. Within this representation the generating function for SU(4) tensor operators reduces to[4]

$$G[D;SU_1(4) \otimes SU_2(4)] = [(1 - D^2\mathbf{C}^2)(1 - D^3\mathbf{C}^3)(1 - D\,\mathbf{U}_1\mathbf{U}_3)(1 - D^2\mathbf{U}_2^2)]^{-1}. \tag{5.22}$$

In this class of degenerate representations, the tensor operator $D^2\mathbf{U}_1\mathbf{U}_3$ is proportional to $D\,\mathbf{U}_1\mathbf{U}_3$. As a result, the three SO(3) scalars $\{(\pi^+\tilde{\pi})_1^{(L)}, (\pi^+\tilde{\pi})_2^{(L)}\}^{(0)}$ become linearly dependent, and the number of independent parameters in the Hamiltonian (to second degree) is reduced by 1.

C. Polyatomic molecules (Poly > 3)

Molecules with three interatomic bonds must have four or more atoms, or must have three atoms in a cyclic configuration. The dynamical group for the vibron model of such molecules is $SU_1(4) \otimes SU_2(4) \otimes \cdots \otimes SU_k(4)$. In the reduction to the diagonal group $SU_D(4)$, generic representations, (μ_1, μ_2, μ_3) with the μ_α all nonzero, will typically occur. The generating function for the tensor content of U[SU(4)] on such representations is known.[5] Although it is fairly complicated, terms of degree less than 3 that occur in this generating function are identical to those that occur in the generating function for the degenerate class of representations $(\mu_1, \mu_2, 0)$. Therefore, the Hamiltonian describing a polyatomic system has the same structure as the Hamiltonian describing a two-bond triatomic system to second degree. It may be constructed as follows.

(1) Include terms of the form shown in (5.8) for each bond, $i = 1,2,\ldots,k$.

(2) For each atom sharing bonds i and j, include terms in the Hamiltonian of the form (5.14) with the substitution $(1,2) \rightarrow (i,j)$.

In going from the generic representation (μ_1, μ_2, μ_3) to the degenerate representations $(\mu_1, \mu_2, 0)$, $(\mu_1, 0, \mu_3)$, $(0, \mu_2, \mu_3)$, no simplifications occur through second degree. The simplification that occurs for the slightly more degenerate representation $(0, \mu_2, 0)$ has been described at the end of Sec. V B. The most degenerate representations $(\mu_1, 0, 0)$ and $(0, 0, \mu_3)$ have the same degenerate enveloping algebra, described in Sec. V A.

VI. SU(6) ⊃ SO(3)

Nuclei exhibiting low-lying collective excited states have been described in terms of the dynamical group SU(6) (see Refs. 26–30). The basic excitations of importance are assumed to have quantum numbers $J^\pi = 0^+, 2^+$. These are created by operators $b_\mu^+ = (s^+, d_m^+)$, $m = +2, +1, 0, -1, -2$. The dynamical group is generated by the bilinear number conserving operators $b_\mu^+ b_\nu$. These operators generate U(6). Removing the first-order Casimir invariant $\Sigma_\mu b_\mu^+ b_\mu = s^+ s + d^+ d = N$, where N has the interpretation of half the number of valence nucleons (or nucleon holes), yields thirty-five generators of the simple Lie group SU(6). This group acts in a Hilbert space that carries a completely symmetric, or boson, representation $(N, \dot{0})$ of SU(6).

In the event that a distinction is to be made between protons and neutrons, collective operators $b_\mu^+(i)$, $i = \pi, \nu$, are introduced.[28–30] From these operators, 72 number conserving, charge conserving operators can be constructed, $b_\mu^+(i) b_\nu(i)$. Removal of the invariants $s_\pi^+ s_\pi + d_\pi^+ d_\pi = N_\pi$ and $s_\nu^+ s_\nu + d_\nu^+ d_\nu = N_\nu$ leads to the direct product dynamical group $SU_\pi(6) \otimes SU_\nu(6)$. This group acts on the direct product of two symmetric Hilbert spaces, $(N_\pi, \dot{0}) \otimes (N_\nu, \dot{0})$. Under restriction to the diagonal subgroup $SU_D(6) \subset SU_\pi(6) \otimes SU_\nu(6)$, this Hilbert space reduces to a direct sum of Hilbert spaces characterized by Young partitions $(N_\pi + N_\nu - K, K, 0)$, $K = 0,1,\ldots,\min(N_\pi, N_\nu)$.

In both cases the symmetry subgroup is SO(3). For the model in which proton–neutron equivalence is assumed (IBM-1), only the most degenerate class of representations, $[N, \dot{0}] = (N, \dot{0})$, occurs. For the model in which protons and neutrons are considered to be distinguishable (IBM-2), the next most degenerate representations of the type $[f_1, f_2, 0]$ must also be considered.

It is convenient to express the generators of SU(6) as spherical tensor operators as follows:

$$T_M^{(L)} = (d^+ \tilde{d})_M^{(L)} = \sum_{m,(m')} \langle 2,m;2,m'|L,M\rangle\, d_m^+ \tilde{d}_{m'}, \tag{6.1}$$

$$D_+ = d^+ \tilde{s} + s^+ \tilde{d}, \quad D_- = i(d^+ \tilde{s} - s^+ \tilde{d}).$$

An additional SO(3) scalar, $S = s^+ s$, is included among the generators of U(6). Here

$$\tilde{s} = s, \quad \tilde{d}_m = (-)^m d_{-m}, \tag{6.2}$$

where $\sim$ denotes the time reversal operation. All these tensor operators are Hermitian in the sense that $O_M^{L+} = (-1)^M O_{-M}^L$. The 15 operators $T^{(L)}$ $(L = 1,3)$, D_- are negative under time reversal. The latter generate a subgroup $SO_T(6)$ of SU(6).

A. IBM-1

The generating function for the SU(6) tensor content of U[SU(6)] acting within the most degenerate class of representation $(N, \dot{0})$ is[31]

$$G[D;SU(6)] = [(1 - D^2\mathbf{C}^2)(1 - D\,\mathbf{U}_1\mathbf{U}_5)]^{-1}. \tag{6.3}$$

Here $D\,\mathbf{U}_1\mathbf{U}_5$ represents the 35-dimensional adjoint representation of SU(6) and $D^2\mathbf{C}^2$ is its second-order Casimir operator. Since the Hilbert space for a nucleus is taken to be a single SU(6) invariant subspace, the Casimir operator may be dropped from the generating function.

Next, it is necessary to determine the SO(3) content of each SU(6) irreducible tensor operator in the enveloping algebra. The only SU(6) tensors that occur are $(K, 0^3, K)$, of degree K. The adjoint representation $(K = 1)$ contains one SO(3) scalar, $n_d = (d^+ \tilde{d})^{(0)}$, the d-boson number operator. The next representation $(K = 2)$ contains seven SO(3) scalars, of which three are self-adjoint and there are two Hermitian adjoint pairs. These operators, together with their group–subgroup parentage, are summarized in Table II.

TABLE II. SO(3) scalars in the SU(6) enveloping algebra, up to second degree, and the classification of their group theoretical parentage. Representation labels are given in terms of Young partitions.

SU(6)	SU(5)	SO(5)	Form
$[21^4]$	[0]	(00)	$(d^+\tilde{d})^{(0)}$
$[42^4]$	[0]	(00)	$\sum_{L=0}^{4}[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}$
	$[42^3]$	(00)	Linear combinations of forms:
		(22)	$[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}$; $L=0,\ldots,4$
	[2]	(00)	$[(d^+\tilde{s})^{(2)}(d^+\tilde{s})^{(2)}]^{(0)}$
	$[2^4]$	(00)	$[s^+\tilde{d})^{(2)}(s^+\tilde{d})^{(2)}]^{(0)}$
	$[31^3]$	(30)	$[(d^+d^+)^{(2)}(\tilde{d}\tilde{s})^{(2)}]^{(0)}$
	$[32^3]$	(30)	$[(s^+d^+)^{(2)}(\tilde{d}\tilde{d})^{(2)}]^{(0)}$

The number of missing labels in the reduction of symmetric representations of SU(6) to the subgroup SO(3) is 3. The missing label folk theorem suggests that there should therefore be six SO(3) scalars in the SU(6) ⊃ SO(3) integrity basis for symmetric representations, in addition to the second-order Casimir operators of SU(6) and SO(3). Seven independent SO(3) scalars are listed in Table II. Note that $[(d^+\tilde{d})^{(0)}(d^+\tilde{d})^{(0)}]^{(0)} = [(d^+\tilde{d})^{(0)}]^2$. The operator $C^2(SU(6))$ is not listed. Also, only three of the five operators $[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}$ $(L=0,1,2,3,4)$ are independent. The operator with $L=1$ is the Casimir operator of SO(3). As a result, the folk theorem is valid in this case.

The fact that only three of the five operators $[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}$ are independent can be seen in two ways. Standard tables show that the representations [0] and $[42^3]$ of U(5) contain one and two SO(3) scalars, respectively. Alternatively, these five operators may be obtained from the ordered operators $[(d^+d^+)^{(L)}(\tilde{d}\tilde{d})^{(L)}]^{(0)}$ using Racah recoupling techniques. Two of these ordered operators vanish by the symmetry of the 3*j* symbols.

The generating function for SO(3) scalars of arbitrary degree in the enveloping algebra of SU(6) is

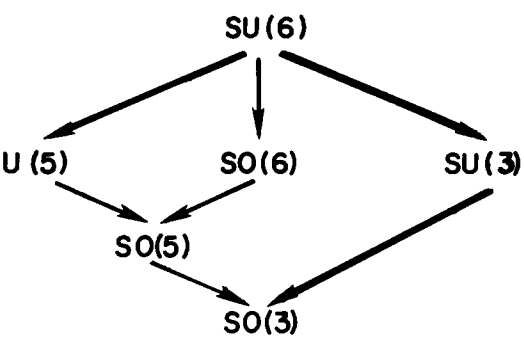
$$
\begin{aligned}
&G[D;\mathrm{SU(6)} \supset \mathrm{SO(3)}] \\
&\quad = F_0^{(1)}(D)F_2^{(2)}(D)F_4^{(2)}(D)\tilde{F}_1^{(2)}(D)\tilde{F}_2^{(2)}(D), \qquad (6.4)\\
&F_0^{(1)} = [1 - D(d^+\tilde{d})^{(0)}]^{-1},\\
&F_L^{(2)} = \{1 - D^2[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}\}^{-1},\\
&\tilde{F}_1^{(2)} = \{1 - D^2[(d^+\tilde{d})^{(2)}(d^+\tilde{s})^{(2)}]^{(0)}\}^{-1}\\
&\qquad + \frac{D^2[(s^+\tilde{d})^{(2)}(d^+\tilde{d})^{(2)}]^{(0)}}{1 - D^2[(s^+\tilde{d})^{(2)}(d^+\tilde{d})^{(2)}]^{(0)}}, \qquad (6.5)\\
&F_2^{(2)} = \{1 - D^2[(d^+\tilde{s})^{(2)}(d^+\tilde{s})^{(2)}]^{(0)}\}^{-1}\\
&\qquad + \frac{D^2[(s^+\tilde{d})^{(2)}(s^+\tilde{d})^{(2)}]^{(0)}}{1 - D^2[(s^+\tilde{d})^{(2)}(s^+\tilde{d})^{(2)}]^{(0)}}.
\end{aligned}
$$

To second order the Hamiltonian with dynamical group–symmetry subgroup structure SU(6) ⊃ SO(3) is

$$
\begin{aligned}
\mathbf{H} = {} & \epsilon_d n_d + \sum_{L=0,2,4} A_L^{(2)}[(d^+\tilde{d})^{(L)}(d^+\tilde{d})^{(L)}]^{(0)}\\
& + \alpha[(d^+\tilde{s})^{(2)}(d^+\tilde{s})^{(2)}]^{(0)} + \alpha'[(s^+\tilde{d})^{(2)}(s^+\tilde{d})^{(2)}]^{(0)}\\
& + \beta[(d^+\tilde{s})^{(2)}(d^+\tilde{s})^{(2)}]^{(0)} + \beta'[(s^+\tilde{d})^{(2)}(s^+\tilde{d})^{(2)}]^{(0)}. \qquad (6.6)
\end{aligned}
$$

Invariance under time reversal requires all parameters to be real. Invariance under Hermitian conjugation requires, in addition, that $\alpha=\alpha', \beta=\beta'$. As a result, there are six independent real parameters in (6.6) corresponding to the linear SO(3) scalar $n_d = T^{(0)}$ and five quadratic SO(3) scalars $[T^{(L)}T^{(L)}]^{(0)}$, $L=0,2,4$, $[T^{(2)}D_+]^{(0)}$, and $[D_+D_+]^{(0)}$.

The linear operator $T^{(0)}$ and the four quadratic operators $[T^{(L)}T^{(L)}]^{(0)}$ with $L=2,4$, $[T^{(2)}D_+]^{(0)}$, and $[D_+D_+]^{(0)}$ provide an integrity basis for Hermitian, time reversal invariant SO(3) scalars in fully symmetric representations $[N,0]$ of SU(6). These operators can be related to the Casimir operators for the group lattice



(6.7)

The generators for the subgroups are

U(5):	$T^{(0)}$	$T^{(1)}$	$T^{(2)}$	$T^{(3)}$	$T^{(4)}$
SO(6):		$T^{(1)}$		$T^{(3)}$	D_+
SO(5):		$T^{(1)}$		$T^{(3)}$	
SU(3):		$T^{(1)}$			$D_+ + \chi T^{(2)}$ with $\chi^2 = \frac{35}{4}$
SO(3):		$T^{(1)}$			

Two different SO(6) subgroups may be chosen, with generators

$$\begin{aligned} &\mathrm{SO}(6): \quad T^{(1)}, T^{(3)}, D_+, \\ &\mathrm{SO}_T(6): \quad T^{(1)}, T^{(3)}, D_-. \end{aligned} \tag{6.8}$$

All generators of $\mathrm{SO}_T(6)$ are negative under time reversal. The subgroup SO(6) is chosen as a matter of convenience. The quadrupole operator D_+ is able to generate states from the vacuum $[(s^+)^\mu |0\rangle]$ with respect to which all Hamiltonian matrix elements are real. Relationships between Casimir operators of the groups in (6.7) are

$$\begin{aligned} \mathrm{C}^1[\mathrm{U}(5)] &= T^{(0)} = n_d, \\ \mathrm{C}^2[\mathrm{U}(5)] - \mathrm{C}^2[\mathrm{SO}(5)] &= \sum_{L=0,2,4} [T^{(L)} T^{(L)}]^{(0)}, \\ \mathrm{C}^2[\mathrm{SO}(6)] - \mathrm{C}^2[\mathrm{SO}(5)] &= [D_+ D_+]^{(0)}, \\ \mathrm{C}^2[\mathrm{SO}(5)] - \mathrm{C}^2[\mathrm{SO}(3)] &= [T^{(3)} T^{(3)}]^{(0)}, \\ \mathrm{C}^2[\mathrm{SU}(3)] - \mathrm{C}^2[\mathrm{SO}(3)] &= [(D_+ + \chi T^{(2)})(D_+ + \chi T^{(2)})]^{(0)}, \quad \text{with } \chi^2 = \tfrac{35}{4}, \\ \mathrm{C}^2[\mathrm{SO}(3)] &= [T^{(1)} T^{(1)}]^{(0)}. \end{aligned} \tag{6.9}$$

B. IBM-2

When a distinction is to be made between protons and neutrons, the dynamical group is $\mathrm{SU}_\pi(6) \otimes \mathrm{SU}_\nu(6)$ (see Refs. 28–30). These groups act in Hilbert spaces H_π and H_ν, which carry fully symmetric representations $(N_\pi,0)$ of $\mathrm{SU}_\pi(6)$ and $(N_\nu,0)$ of $\mathrm{SU}_\nu(6)$. The direct product Hilbert space $H_\pi \otimes H_\nu$ is reducible with respect to the diagonal subgroup $\mathrm{SU}_D(6)$. The irreducible content of this Hilbert space consists of representations with Young partitions $[f_1 = N_1 + N_2 - K, f_2 = K, 0]$, $K = 0,1,2,\ldots, \min(N_\pi, N_\nu)$.

In addition to the SO(3) scalars, which can be constructed from the generators of $\mathrm{SU}_\pi(6)$ alone or of $\mathrm{SU}_\nu(6)$ alone (Table II), there are scalars which are "cross terms." These describe specifically the proton–neutron interaction, and first occur for $d = 2$. Many of these operators can be constructed from the second-degree operators given (6.8). The structure of all SO(3) scalars can be determined by constructing the generating function for the $\mathrm{SU}_D(6)$ tensor content in U[SU(6)] acting on degenerate representations in the class $(\mu_1, \mu_2, 0)$, and then constructing the SO(3) scalar generating function. The tensor generating function is

$$\begin{aligned} &G[D;\mathrm{SU}(6) \supset \mathrm{SO}(3)] \\ &= [(1 - D^2\mathrm{C}^2)(1 - D^3\mathrm{C}^3)(1 - D\,\mathrm{U}_1\mathrm{U}_5)(1 - D^2\mathrm{U}_1\mathrm{U}_5)(1 - D^2\mathrm{U}_2\mathrm{U}_4)]^{-1} \\ &\times \left\{ \frac{1}{(1 - D^3\mathrm{U}_1^2\,\mathrm{U}_4)} + \frac{D^3\mathrm{U}_2\mathrm{U}_5^2}{(1 - D^3\mathrm{U}_2\mathrm{U}_5^2)} \right\}. \end{aligned} \tag{6.10}$$

Once again, the Casimir operators are now explicitly included in the generating function because several irreducible representations occur in the reduction of $H_\pi \otimes H_\nu$. To second order, two new tensor operators besides C^2 occur, $D^2\mathrm{U}_1\mathrm{U}_5$ and $D^2\mathrm{U}_2\mathrm{U}_4$. The operators C^2, $D^2\mathrm{U}_1\mathrm{U}_5$, and $D^2\mathrm{U}_2\mathrm{U}_4$ each contribute one self-adjoint SO(3) scalar.

The degree-2 SO(3) scalars describing the proton–neutron interaction, which are derived from the term $(1 - D\,\mathrm{U}_1\mathrm{U}_5)^{-1}$, and whose structure can be inferred from the scalars in (6.8), are

$$\begin{aligned} &\{T_\pi^{(L)}, T_\nu^{(L)}\}^{(0)}, \quad \text{with } L = 0,2,4, \\ &\{T_\pi^{(2)}, D_{+\nu}\}^{(0)}, \quad \{D_{+\pi}, D_{+\nu}\}^{(0)}. \end{aligned} \tag{6.11a}$$

In these expressions the brackets { , } mean that the operators within should be symmetrized with respect to the interchange of proton–neutron indices.

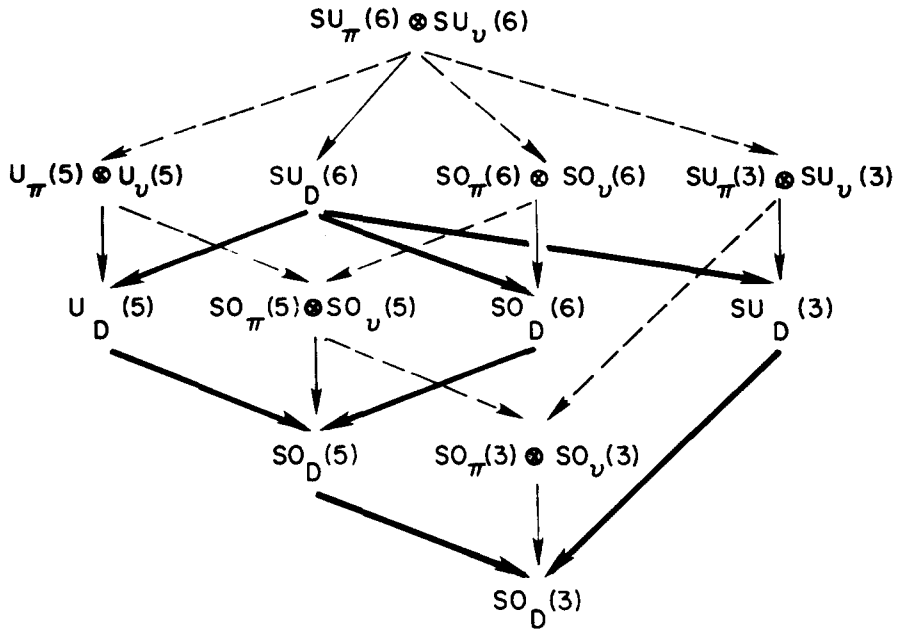
The three additional self-adjoint operators of degree $d = 2$, which are derived from the additional terms $(1 - D^2\mathrm{C}^2)^{-1}$, $(1 - D^2\mathrm{U}_1\mathrm{U}_5)^{-1}$, and $(1 - D^2\mathrm{U}_2\mathrm{U}_4)^{-1}$, are

$$\{(d^+ \tilde{d})_\pi^{(L)}, (d^+ \tilde{d})_\nu^{(L)}\}^{(0)}, \quad \text{with } L = 1,3, \tag{6.11b}$$

$$(d_\pi^+ \tilde{d}_\nu)^{(0)} (d_\nu^+ \tilde{d}_\pi)^{(0)} \text{ or equivalently } \{D_{-\pi}, D_{-\nu}\}^{(0)}.$$

The total number of terms up to degree $d = 2$ in the Hamiltonian with dynamical and symmetry groups $\mathrm{SU}_\pi(6) \times \mathrm{SU}_\nu(6) \supset \mathrm{SO}(3)$ is 20. Of these, six describe the protons alone and six the neutrons. The proton–neutron interaction is described by eight terms. Five of these are obtained directly from the scalar operators in $\mathrm{SU}(6) \supset \mathrm{SO}(3)$ as Casimir covariants using the construction of (5.13). The remaining three terms, given in (6.11b), are sometimes called Majorana terms.

It is possible to replace the SO(3) scalars in $\mathrm{SU}_\pi(6) \times \mathrm{SU}_\nu(6)$ up to degree 2 by the Casimir operators belonging to a suitably chosen group lattice. As in the case $\mathrm{SU}(4) \otimes \mathrm{SU}(4) \supset \mathrm{SO}(3)$, this group lattice can be chosen to reflect the construction of the proton–neutron interaction terms from those of the simpler model with proton–neutron equivalence [cf. (5.15)]:



(6.12)

Once again, this lattice provides a clear separation of the SO(3) scalars describing protons alone or neutrons alone from those describing the proton–neutron interaction. The former are derived as Casimir operators of direct product groups with the latter obtained from diagonal subgroups. To guide the eye, in (6.12) heavy solid lines are used to indicate diagonal group to diagonal subgroup connections, broken lines for direct product group to direct product subgroup connections, and regular solid lines to indicate the direct product to diagonal subgroup connections. This lattice does not describe an integrity basis for terms of degree $d \geqslant 3$.

Many additional groups can be placed in the group lattice with $SU_\pi(6) \otimes SU_\nu(6)$ as the dynamical group and SO(3) as the symmetry subgroup. First, "off-diagonal cross-product terms" such as $U_\pi(5) \times SO_\nu(6)$ can occur. Second, D^* subgroups can occur by the process described in (5.20). And finally, subgroups can be constructed by combining two groups, G_i and G_i', in an out of phase way. For example, the quadrupole tensor for the Lie algebra SU(3) for the proton and neutron subgroups of the dynamical groups are

$$Q_\pi = [d^+\tilde{s} + s^+\tilde{d} + \chi(d^+\tilde{d})^{(2)}]_\pi, \quad \text{with } \chi_\pi^2 = \tfrac{35}{4}, \tag{6.13}$$

$$Q_\nu = [d^+\tilde{s} + s^+\tilde{d} + \chi(d^+\tilde{d})^{(2)}]_\nu, \quad \text{with } \chi_\nu^2 = \tfrac{35}{4}.$$

When the two factors χ_π and χ_ν are chosen with the same sign the standard diagonal subgroup SU(3) results. When they are chosen with the opposite sign the subgroup called SU*(3) results. In this group–subgroup reduction the representations of $SU_\pi(3)$ are combined with the conjugate representations of $SU_\nu(3)$ and then reduced.

VII. SUMMARY AND CONCLUSIONS

An algorithm has been presented for constructing a Hamiltonian from the generators of a dynamical group G, which is invariant under the operations of a symmetry subgroup $H \subset G$. Such an algorithm is necessary to determine when the H scalars on which the Hamiltonian depends can be replaced by the Casimir operators for the members in a group–subgroup lattice.

In principle, the algorithm is simple; it is certainly straightforward to state: The terms homogeneous of degree d in the universal enveloping algebra of the Lie algebra of G, $U^d(G)$, carry a representation of G, which in general is reducible. Under reduction to the subgroup $H \subset G$ the number of times the identity representation, $\gamma^e(H)$, occurs in $U^d(G)$ can be determined. The basis vector operators for these identity representations are the H scalars in $U^d(G)$ from which the Hamiltonian can be constructed.

As the implementation of this algorithm is usually somewhat less than straightforward, and since there are simplifications that reduce the complexity of the results in many cases of physical interest, a number of guidelines for using and simplifying the algorithm have been presented. The major part of these simplifications are effected by introducing generating functions for the irreducible tensor content of $U^d(G)$. For several cases of physical interest these functions have been constructed explicitly. Furthermore, it is shown how these functions can be used in turn to construct generating functions for the H scalars in $U^d(G)$.

Generating functions are generally available only for connected Lie subgroups of Lie groups. The effects of discrete operations (such as space inversion, time reversal, and Hermitian conjugation), which must also leave the Hamiltonian invariant, must be studied separately. These symmetries impose additional reality constraints on the parameters which appear in the Hamiltonian. Thus, they may reduce the number of H scalars which can appear in the integrity basis for the H scalars in the universal enveloping algebra of the dynamical group G from which Hamiltonians can be constructed.

The use of this algorithm was illustrated first for the dynamical group SU(2). In the case that $H \subset SU(2)$ is finite, the algorithm leads to the construction of the point group tensor harmonics as linear combinations of the spherical ten-

sor operators. These have been used in the field of solid state physics to describe ligand and crystal fields.

The algorithm was then used to construct the integrity basis for SU(3) ⊃ SO(3), SU(4) ⊃ SO(3), and SU(6) ⊃ SO(3). In addition, another algorithm was presented for constructing the integrity basis for direct product dynamical groups. The problem of replacing the integrity basis by the Casimir operators belonging to a group–subgroup lattice was also considered for the three cases SU(N) ⊃ SO(3), $N = 3,4$, and 6.

Compact simple Lie groups have been used to illustrate the algorithm stated in Sec. I. However, the algorithm can be applied to noncompact and nonsemisimple groups as well: The universal enveloping algebra, U(G), is structurally the same whether the group is simple, semisimple, nonsemisimple, compact, or noncompact.[32] Each subspace $\mathrm{U}^d(\mathbf{G})$ is finite dimensional if the algebra is finite dimensional. Its reduction into G-invariant tensor operators proceeds without change. When G is noncompact, the irreducible representations so obtained are finite-dimensional but nonunitary. The H scalars in these finite-dimensional representations can be obtained as before. The only change brought about by the use of a noncompact dynamical group G is in the reality restrictions on the complex parameters, which appear in the Hamiltonian **H**, brought about by the requirement that **H** be Hermitian. If G' is a noncompact dynamical group obtained by analytic continuation of the compact group G, then straightforward analytic continuation of the computation of the integrity basis of $G \supset H$ can be used to construct the integrity basis for $G' \supset H$. This remark extends to the group–subgroup lattice as well.

It would be useful to have a criterion for determining completeness of an integrity basis. The generating function for the G tensor content of U(G), and the extension of this generating function to one for the H scalars in U(G) suffices. However, construction of the latter from the former may be difficult. A more easily and directly applicable criterion, such as the missing label folk theorem, would be useful. This result holds for a number of the applications considered, and fails to hold for several applications also considered above. Something like it should be true, but we cannot recommend it as a test for completeness of an integrity basis.

The dependences, which occur among inequivalent sets of tensor operators with identical transformation properties on restriction from one class of representations to a more degenerate class of representations, can be studied using as a tool the coherent states associated with the dynamical group G. The close relation between coherent states and independent tensor operators can be seen as follows. Generalized coherent states are defined by specifying (i) a Lie group G, (ii) a Hilbert space in which G acts (in this case, a class of representations), and (iii) an extremal state within this class of representations. An immediate output of the coherent state construction procedure is a stability subgroup (in this case, K). The structures used in the construction of coherent states are exactly those used in determining the number of inequivalent tensors of a given type that act within a class of representations [cf. Sec. II, remark (6)]. The expectation value of these operators with respect to the coherent states provides a system of functions in which linear dependences show up clearly on restriction to a more degenerate class of representations.

In this work we have not investigated the problem of constructing the most general group–sub group lattice based on the dynamical group–symmetry subgroup pair $G \supset H$. This is the counterpart of the problem of determining the integrity basis for the same pair of groups. The objective of such a study would be the development of an algorithm for constructing a "universal enveloping" group–subgroup lattice, which could be compared with the integrity basis and from which a minimal complete group–subgroup lattice could be isolated and put in 1–1 correspondence with the elements in the integrity basis.

Note added in proof: The concept of integrity basis developed in the context of invariant theory was studied extensively in the past century by Molien, Grace, Young, and others. See Weyl[33] for extensive references. The use of integrity bases in physical applications is implicit in the work of many authors. Explicit use of the integrity basis appears in the work of Judd,[34] Smith and Wybourne,[35] Wybourne,[36] Killingbeck,[37] and McClellan.[38] We are indebted to Professor Wybourne for providing this brief outline of the use of the integrity basis concept in physical applications.

ACKNOWLEDGMENTS

We thank Professor F. Iachello for stressing the importance of this problem, Professor R. T. Sharp for suggesting the use of degenerate generating functions, and Professor J. P. Elliott, Professor K. T. Hecht, Professor G. Rosensteel, and Professor O. Scholten for many useful remarks. We also thank Professor D. J. Rowe and B. Castel for their hospitality at the Summer Institute of Theoretical Physics at Queen's University, Kingston, Ontario, Canada, where this work was initiated.

Support from the National Science Foundation under Grants No. PHY84-41891 and No. PHY84-09662 is also acknowledged.

[1] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[2] D. E. Littlewood, Philos. Trans. R. Soc. London Ser. A **239**, 305 (1944).
[3] H. Weyl, Math. Z. **24**, 277, 328 (1926); *The Classical Groups* (Princeton U.P., Princeton, NJ, 1946).
[4] Y. Giroux, M. Couture, and R. T. Sharp, J. Phys. A **17**, 715 (1984).
[5] M. Couture and R. T. Sharp, J. Phys. A **13**, 1925 (1980).
[6] A. Peccia and R. T. Sharp, J. Math. Phys. **17**, 1313 (1976).
[7] M. Berrondo and A. Palma, J. Phys. A **13**, 773 (1980).
[8] Y. Alhassid, F. Iachello, and F. Gürsey, Chem. Phys. Lett. **99**, 27 (1983).
[9] J. B. French, in *Isospin in Nuclear Physics*, edited by D. H. Wilkinson (North-Holland, Amsterdam, 1969).
[10] E. R. Callen and H. B. Callen, Phys. Rev. **129**, 578 (1963).
[11] O. S. van Roosmalen, R. D. Levine, and A. E. L. Dieperink, Chem. Phys. Lett. **101**, 512 (1983).
[12] J. P. Elliott, Proc. R. Soc. London Ser. A **245**, 128 (1958).
[13] J. P. Draayer and G. Rosensteel, Nucl. Phys. A **439**, 61 (1985).
[14] P. H. Butler and B. G. Wybourne, Atomic Data **3**, 133 (1971).
[15] M. Gell-Mann, Cal. Tech. Report CTSL-20, 1961 (unpublished) [reprinted in M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (Benjamin, New York, 1964)].
[16] B. R. Judd, W. Miller, Jr., J. Patera, and P. Winternitz, J. Math. Phys. **15**, 1787 (1974).
[17] J. P. Draayer and K. J. Weeks, Phys. Rev. Lett. **51**, 1422 (1983).

[18]J. P. Draayer, K. J. Weeks, and G. Rosensteel, Nucl. Phys. A **413**, 215 (1984).
[19]G. Rosensteel, J. P. Draayer, and K. J. Weeks, Nucl. Phys. A **419**, 1 (1984).
[20]J. P. Draayer and K. J. Weeks, Ann. Phys. **156**, 91 (1984).
[21]F. T. Arecchi, R.Gilmore, and D. M. Kim, Lett. Nuovo Cimento **6**, 219 (1973).
[22]R. Gilmore, Rev. Mex. Fis. **23**, 143 (1974).
[23]F. Iachello and R. D. Levine, J. Chem. Phys. **77**, 3046 (1982).
[24]O. S.van Roosmalen, F. Iachello, R. D. Levine, and A. E. L. Dieperink, J. Chem. Phys. **79**, 2515 (1983).
[25]R. Gilmore, *Lie Groups, Lie Algebras, and Some of Their Applications* (Wiley–Interscience, New York, 1974).
[26]A. Arima and F. Iachello, Phys. Rev. Lett. **35**, 1069 (1975).
[27]A. Arima and F. Iachello, Ann. Phys. **99**, 253 (1976); **111**, 201 (1978); **123**, 468 (1979).
[28]A. Arima, T. Otsuka, F. Iachello, and I. Talmi, Phys. Lett. B **66**, 205 (1977).
[29]T. Otsuka, A. Arima, F. Iachello, and I. Talmi, Phys. Lett. B **76**, 139 (1978).
[30]T. Otsuka, A. Arima, and F. Iachello, Nucl. Phys. A **309**, 1 (1978).
[31]R. C. King, J. Patera, and R. T. Sharp, J. Phys. A **15**, 1143 (1982).
[32]G. Rosensteel and J. P. Draayer, Nucl. Phys. A **436**, 445 (1985).
[33]H. Weyl, *The Classical Groups* (Princeton U.P., Princeton, NJ, 1946).
[34]B. R. Judd, Phys. Rev. **141**, 4 (1966).
[35]P. R. Smith and B. G. Wybourne, J. Math. Phys. **8**, 2434 (1967).
[36]B. G.Wybourne, *Symmetry Principles in Atomic Spectroscopy* (Wiley, New York, 1970).
[37]J. Killingbeck, J. Phys. C **5**, 2497 (1972).
[38]A. G. McClellan, J. Phys. C **1**, 3326 (1974); **9**, 939, 4083 (1976); **11**, 4665 (1978); **12**, 5411 (1978); A. G. McClellan, *The Classical Thermodynamics of Deformable Materials* (Cambridge U.P., London, 1980).

The Green's function for a finite linear chain

Ronald Bass
Department of Physics and Astronomy, University of Kansas, Lawrence, Kansas 66045

(Received 28 May 1985; accepted for publication 25 June 1985)

A new expression for the Green's function of a finite-length one-dimensional harmonic lattice with nearest-neighbor interactions is reported. Simple closed expressions in terms of Chebyshev polynomials are developed for periodic, fixed, and free end boundary conditions.

I. THE GENERAL GREEN'S FUNCTION

The finite-length one-dimensional single-component harmonic lattice with nearest-neighbor interactions is a well-studied standard model whose natural frequencies and normal modes are well known for a variety of boundary conditions.[1]

However, the real space representation of the Green's function for this model has only been developed as a determinantal expression for periodic boundary conditions. We shall show here that the Green's function for this model can readily be derived as a simple closed expression in terms of Chebyshev polynomials for a wide range of different boundary conditions.

The equations of motion for the N atom lattice may be written in dimensionless units as

$$2zu(n) - u(n+1) - u(n-1) = 0, \tag{1}$$

where $u(n)$ is the displacement of the nth particle from its equilibrium position, n lies in the range $1 \leqslant n \leqslant N$, and $z = 1 - \omega^2$, where ω is the dimensionless frequency.

This model is usually constrained by one of four sets of boundary conditions expressed as constraints on $u(0)$ and $u(N+1)$: (i) periodic boundary conditions, where $u(N+n) \equiv u(n)$; (ii) both ends clamped, so that $u(0) = u(N+1) = 0$; (iii) both ends free, so that $u(0) = u(1)$ and $u(N) = u(N+1)$; and (iv) one end free and the other clamped, for example $u(0) = u(1)$ and $u(N+1) = 0$.

The Green's function[2] for the model is the solution to the inhomogeneous equation

$$zG(n,m) - \tfrac{1}{2}G(n-1,m) - \tfrac{1}{2}G(n+1,m) = \delta_{n,m}, \tag{2}$$

for all $N \geqslant n, m \geqslant 0$. When $n \neq m$, we note that the difference equation (2) is the recurrence relation for Chebyshev polynomials,[3] and that consequently the solution to Eq. (2) must be a sum of Chebyshev polynomials. As the solution must also obey the time-reversal requirement that $G(n,m) = G(m,n)$, the polynomials with the appropriate symmetry to be included are $T_{n-m}(z)$, $T_{n+m}(z)$, $U_{|n-m|-1}(z)$, and $U_{n+m-1}(z)$.

Direct substitution into Eq. (2) reveals that $-U_{|n-m|-1}(z)$ is a particular solution to the equation, while the other three polynomials are solutions to the related homogeneous equation (1). Hence the Green's function may be written in general form as

$$\begin{aligned} G(n,m) = {} & A(z)T_{n-m}(z) + B(z)T_{n+m}(z) \\ & + C(z)U_{n+m-1}(z) - U_{|n-m|-1}(Z), \end{aligned} \tag{3}$$

where $A(z)$, $B(z)$, and $C(z)$ are determined by the boundary conditions.

II. APPLYING THE BOUNDARY CONDITIONS

The solutions under the four sets of boundary conditions may now be found as follows.

(*i*) *Periodic boundary conditions:* The Green's function depends only on the relative position $n-m$ so that $B(z) = C(z) = 0$, and $A(z)$ is the site diagonal Green's function, $G(n,n)$. For an infinite chain it is known that $G(n,n) = (z^2-1)^{-1/2}$ so that

$$G(n,m) = T_{n-m}(z)/\sqrt{z^2-1} - U_{|n-m|-1}(z). \tag{4}$$

For a finite lattice, periodic boundary conditions can be written as $G(N,0) = G(0,0)$ so that

$$A(z) = U_{N-1}(z)/[T_N(z) - 1]. \tag{5}$$

The Green's function for N particles subject to periodic boundary conditions is then

$$\begin{aligned} G(n,m) &= G_N(n-m) \\ &= \frac{T_{n-m}(z)U_{N-1}(z)}{T_N(z)-1} - U_{|n-m|-1}(z). \end{aligned} \tag{6}$$

(*ii*) *Both ends clamped:* The boundary conditions are $G(0,m) = G(N+1,m) = 0$, regardless of the value of $m \geqslant 0$. The first of these conditions is

$$G(0,m) = 0 = (A(z) + B(z))T_m(z) + (C(z) - 1)U_{m-1}(z).$$

The only solution valid for all m is when $B(z) = -A(z)$ and $C(z) = 1$. Then

$$\begin{aligned} G(N+1,m) = 0 = {} & A(z)(T_{N+1-m}(z) - T_{N+1+m}(z)) \\ & + U_{N+m}(z) - U_{N-m}(z) \\ = {} & \{-2A(z)(z^2-1)U_N(z) \\ & + T_{N+1}(z)\}U_{m-1}(z), \end{aligned}$$

so that

$$A(z) = \frac{T_{N+1}(z)}{(z^2-1)U_N(z)} = \frac{U_{2N+1}(z)}{T_{2N+2}-1}.$$

Hence the Green's function for the N atom chain with clamped ends is

$$\begin{aligned} G(n,m) &= (T_{n-m}(z) - T_{n+m}(z))\frac{U_{2N+1}(z)}{T_{2N+2}(z)-1} \\ &\quad - U_{|n-m|-1}(z) + U_{n+m-1}(z) \\ &= G_{2N+2}(n-m) - G_{2N+2}(n+m), \end{aligned} \tag{7}$$

where $G_N(m)$ was defined in Eq. (6).

 0022-2488/85/123068-02$02.50

(*iii*) *Both ends free:* The boundary conditions when both ends are free are designed to ensure that the force on each end of the chain is zero. This is ensured by setting $G(0, m) = G(1, m)$ and $G(N+1, m) = G(N, m)$. The first of these conditions can be written as

$$\begin{aligned} G(0, m) - G(1, m) = 0 = & A(z)(T_m(z) - T_{m-1}(z)) \\ & + B(z)(T_m(z) - T_{m+1}(z)) \\ & + C(z)(U_{m-1}(z) - U_m(z)) \\ & - (U_{m-1}(z) - U_{m-2}(z)). \end{aligned}$$

Using the identities for Chebyshev polynomials,[3] $T_m(z) = zT_{m+1}(z) - (z^2-1)U_m(z)$ and $U_{m-1}(z) = zU_m(z) - T_{m+1}(z)$, we find that the solution, valid for all m, is $B(z) = 1 + zA(z)$ and $C(z) = (1-z^2)A(z) - z$. Substituting these into Eq. (3), we find the Green's function to be

$$\begin{aligned} G(n, m) = & A(z)\{T_{n-m}(z) + T_{n+m}(z)\} \\ & - U_{|n-m|-1}(z) - U_{n+m-1}(z). \end{aligned} \tag{8}$$

The value of $A(z)$ is then found from the boundary condition at the other end of the chain. Here $G(N+1, m) = G(N, m)$, so that

$$\begin{aligned} 0 = & A(z)\{T_{N-m}(z) + T_{N+m-1}(z) \\ & - T_{N+1-m}(z) - T_{N+m}(z)\} \\ & - U_{N-m-1}(z) - U_{N-m-2}(z) \\ & + U_{N-m}(z) + U_{N+m-1}(z) \\ = & \{2A(z)(z^2-1)U_{N-1}(z) - 2T_N(z)\} \\ & \times \{U_{m-2}(z) - U_{m-1}(z)\} \end{aligned}$$

so that

$$A(z) = \frac{T_N(z)}{(z^2-1)U_{N-1}(z)} = \frac{U_{2N-1}(z)}{T_{2N}(z) - 1}$$

and

$$\begin{aligned} G(n, m) = & \frac{U_{2N-1}(z)}{T_{2N}(z) - 1}(T_{n-m}(z) + T_{n+m-1}(z)) \\ & - U_{|n-m|-1}(z) - U_{n+m-2}(z) \\ = & G_{2N}(n-m) + G_{2N}(n+m-1). \end{aligned} \tag{9}$$

(*iv*) *One end clamped, the other free:* The Green's function that obeys the free boundary condition $G(0, m) = G(1, m)$ is given by Eq. (8). The clamped boundary condition on the far end, $G(N+1, m) = 0$, determines the value of $A(z)$ through

$$\begin{aligned} G(N+1, m) = 0 = & A(z)\{T_{N+1-m}(z) + T_{N+m}(z)\} \\ & - U_{N-m}(z) - U_{N+m-1}(z), \end{aligned}$$

so that

$$A(z) = U_{2N}(z)/[T_{2N+1}(z) + 1],$$

and the Green's function is

$$\begin{aligned} G(n, m) = & [U_{2N}(z)/T_{2N+1}(z) + 1]\{T_{n-m}(z) \\ & + T_{n+m-1}(z)\} - U_{|n-m|-1}(z) \\ & - U_{n+m-2}(z). \end{aligned} \tag{10}$$

Consequently, we have found that the Green's function for a finite-length one-component one-dimensional lattice with nearest-neighbor spring constants may be expressed in simple closed form in terms of Chebyshev polynomials, the exact form being determined by the applicable boundary conditions.

[1] See, for example, *Mathematical Methods in One Dimension*, edited by E. H. Lieb and D. C. Mattis (Academic, New York, 1966).

[2] E. N. Economu, *Green's Functions in Quantum Physics* (Springer, New York, 1979).

[3] See Chap. 22 in *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Natl. Bur. Stand., Washington, D. C., 1968).

Integrable Hamiltonian systems with velocity-dependent potentials

B. Dorizzi and B. Grammaticos
Département de Mathématiques (MTI), Centre National d'Etude de Télécommunications, F-92131, Issy-les-Moulineaux, France

A. Ramani
Centre de Physique Théorique, Ecole Polytechnique F-91128, Palaiseau, France

P. Winternitz
Centre de recherche mathématiques, Université de Montréal, C. P. 6128, Montréal, Québec, Canada H3C 3J7

(Received 27 September 1984; accepted for publication 6 June 1985)

The integrability of a two-dimensional Hamiltonian in which the potential depends explicitly on the momenta is investigated. Hamiltonians of this kind are encountered in the description of the motion of a particle in a magnetic field. Two integrable classes of potentials are identified and the second integral of motion is constructed for each of them. The singularity analysis of the equations of motion is also performed, confirming once more the relation between the (weak) Painlevé property and integrability.

I. INTRODUCTION

The purpose of this article is to study two-dimensional Hamiltonian systems corresponding to the motion of a particle in a plane under the influence of a scalar potential and a transverse magnetic field. The Hamiltonian is postulated to have the form

$$H = \tfrac{1}{2}(p_x^2 + p_y^2) + A(x,y)p_x + B(x,y)p_y + W(x,y), \qquad (1.1)$$

where A, B, and W are, at this stage, arbitrary functions of the coordinates x and y, and p_x and p_y are the momenta canonically conjugate to x and y. We are interested in establishing the conditions on the scalar and vector potentials W and $\mathbf{A} = (A,B)$, under which the system becomes integrable, i.e., a second integral of motion exists.

The problem of finding integrable dynamical systems of physical interest is a difficult one. Research in this direction has been actively pursued during the last few years for a variety of reasons. These include the good physical properties of integrable systems, namely the regular behavior of trajectories and the related predictability of the behavior of the system over long periods of time. In particular, integrable systems with well-behaved integrals of motion will not exhibit chaotic behavior. A knowledge of these sytems, on the other hand, helps in the study of "neighboring" nonintegrable systems, e.g., in the investigation of the onset of chaos.

Even for Hamiltonian systems not too many results are known. The Toda and Calogero systems[1-3] are among the rare examples of integrable N-particle systems on a line with pair interactions.

For a two-dimensional Hamiltonian system the problem should, in principle, be a simpler one. In order to ensure integrability, all that is needed is for a second constant of motion, independent of the Hamiltonian, to exist. Even in this case few general results are known, but the combined use of different techniques has made it possible to identify quite a few two-dimensional integrable systems of interest, and in some cases whole classes of such systems.

A straightforward method, due to Bertrand,[4] involves a direct search for additional integrals of motion, making the assumption that these constants are polynomials in the velocities (or momenta). The coefficients of the polynomial are functions of the coordinates and are obtained by solving certain, in general nonlinear, partial differential equations. This direct method is most powerful in the simpler cases, when the additional integrals of motion are first- or second-order polynomials in the momenta. It has been generalized to the case of quantum mechanical integrable systems,[5,6] and applied to the case of two- and three-dimensional nonrelativistic one-particle Hamiltonian systems with velocity-independent forces.[5-7] In this case the existence of second-order integer polynomial integrals of motion has been related to the separation of variables in the corresponding Hamilton–Jacobi or Schrödinger equation.

The drawbacks of the direct method are quite obvious. It cannot be applied to find nonpolynomial integrals of motion, which are known to exist and to be of interest.[8] Moreover, even in the polynomial case, the method becomes extremely cumbersome for polynomials of order higher than 2 in the momenta. More significantly, there is no guarantee that we will be able to solve the differential equations involved in the reconstruction of the coefficients in the polynomial.

Other methods are hence needed to identify integrable systems or at least to pinpoint candidates for integrability. A powerful tool for this purpose is the Painlevé criterion,[9] based on the study of the singularity structure of the solutions in the complex time plane. According to this criterion, a system is a candidate for integrability, if it does not allow any moving critical points, i.e., if the only singularities (in an arbitrary solution) that depend on the initial conditions are poles. The application of the Painlevé criterion has led to the identification of several two-dimensional integrable Hamiltonian systems.

As mentioned above, Hamiltonian systems of the type

$$H = \tfrac{1}{2}(p_x^2 + p_y^2) + V(x,y), \qquad (1.2)$$

 0022-2488/85/123070-10$02.50

with a second integral of motion that is quadratic in the momenta, have been completely classified.[5,7] The integral of motion in this case has been shown to be the sum of a second-order element in the enveloping algebra of the Lie algebra of motions of the plane e(2) [or its complexification e(2,C) if complex Hamiltonians are considered] and a function $h(x,y)$, related to the potential $V(x,y)$. A classification of quadratic integrals of motion then amounts to a classification of such second-order elements in the enveloping algebra into orbits under the action of the corresponding Euclidean Lie group E(2) [or E(2,C)]. In the real case four types of such orbits exist, corresponding to potentials allowing the separation of variables in the Hamilton–Jacobi (or Schrödinger) equation in Cartesian, parabolic, polar, and elliptic coordinates, respectively. In the complex case, three more orbits, and correspondingly, three more integrable types of potentials exist.[5,7]

Some partial results also exist[10–14] for integrals of motion of the order 3,4, or 6. Here group theory is somewhat less helpful, since only the highest-order terms in the polynomial lie in the enveloping algebra of e(2).

An interesting and different approach has been adopted by Hall.[12] He generalizes the concept of integrability by allowing the second integral of motion to depend explicitly on the energy E. Such an integral will have different values on each energy surface, may have different properties on different subsets of energy surfaces, and may indeed only exist on certain subsets. Such "configurational invariants," in Hall's terminology, are of considerable interest in both classical and quantum mechanics. Indeed, in quantum mechanics they could be used to analyze accidental degeneracy of certain energy levels, rather than of the entire energy spectrum. We shall not go into this interesting question in the present article.

In this paper we concentrate on the Hamiltonian (1.1) and restrict ourselves to systems that allow integrals of motion that are first- or second-order polynomials in the velocities (or momenta). Indeed, the integral of motion will be written as

$$C = g_0\dot{x}^2 + g_1\dot{x}\dot{y} + g_2\dot{y}^2 + k_0\dot{x} + k_1\dot{y} + h, \tag{1.3}$$

where g_i, k_i, and h are functions of the coordinates x and y only. By assumption, they are thus independent of time t and energy E.

In Sec. II we first analyze the linear case, i.e., $g_0 = g_1 = g_2 = 0$ in (1.3). As usual, such an invariant leads to potentials with purely geometric symmetries. The quadratic case (at least one of the g_i's nonvanishing) is more interesting. As in the case of velocity-independent potentials (1.2), we are led to consider several cases, corresponding to different types of separable coordinate sytems. A major complication arises in the present case: the linear terms $k_0(x,y)$ and $k_1(x,y)$ are present and cannot, as opposed to the case of velocity-independent potentials, be set equal to zero as a consequence of time reversal invariance. The g_i terms of (1.3) will again lie in the enveloping algebra of e(2), not, however, the linear terms.

In Sec. III we obtain the most general Hamiltonian (1.1) allowing a "Cartesian" type integral of motion. The motion of a particle in several special cases of such Hamiltonian systems is investigated in Sec. IV: we present some bounded trajectories and special Poincaré sections. Section V deals with the Painlevé analysis of the general system found in Sec. III: the potentials and the integral of motion in general involve Weierstrass elliptic functions and hence have interesting periodicity properties. Some conclusions are presented in Sec. VI. The Appendix deals with a "degenerate" case of the Cartesian type of integral of motion, occurring only for complex Hamiltonians and reflecting the existence of isotropic (zero-length) momenta in the complex case.

II. FORMULATION OF THE PROBLEM

The equations of motion for the system with Hamiltonian (1.1) are

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p_x} = p_x + A, \\ \dot{y} &= \frac{\partial H}{\partial p_y} = p_y + B, \\ \dot{p}_x &= -\frac{\partial H}{\partial x} = -W_x - A_x p_x - B_x p_y, \\ \dot{p}_y &= -\frac{\partial H}{\partial y} = -W_y - A_y p_x - B_y p_y, \end{aligned} \tag{2.1}$$

or, after eliminating the momenta,

$$\begin{aligned} \ddot{x} &= -W_x + AA_x + BB_x + \dot{y}(A_y - B_x), \\ \ddot{y} &= -W_y + AA_y + BB_y - \dot{x}(A_y - B_x). \end{aligned} \tag{2.2}$$

Putting

$$\Omega \equiv A_y - B_x, \quad V = W - \tfrac{1}{2}(A^2 + B^2), \tag{2.3}$$

we simplify the equations of motion to

$$\begin{aligned} \ddot{x} &= -V_x + \Omega\dot{y}, \\ \ddot{y} &= -V_y - \Omega\dot{x}, \end{aligned} \tag{2.4}$$

and the Hamiltonian to

$$H = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) + V(x,y). \tag{2.5}$$

Once Ω and V are determined we can use the definitions (2.3) to reconstruct the scalar potential W and the vector $\mathbf{A} = (A,B)$, up to a gauge transformation

$$\begin{pmatrix} V \\ \mathbf{A} \end{pmatrix} \to \begin{pmatrix} V + (\mathbf{A},\nabla\phi) + \frac{1}{2}(\nabla\phi)^2 \\ \mathbf{A} + \nabla\phi \end{pmatrix}, \tag{2.6}$$

where $\phi(x,y)$ is an arbitrary function.

We will consider the Hamiltonian system in the form (2.4) and (2.5) and investigate the conditions under which the system admits a second integral of motion C, independent of the Hamiltonian and linear or quadratic in the velocities.

A. Linear integral of motion

Let us first consider the linear case, i.e.,

$$C = f_0(x,y)\dot{x} + f_1(x,y)\dot{y} + h(x,y), \tag{2.7}$$

where again f_i and h are independent of time and energy. Following Bertrand's method we impose $dC/dt = 0$, use the equations of motion (2.4), and equate the coefficients of each term in $\dot{x}$ and $\dot{y}$ to zero. We obtain

$$f_{0x} = 0, \quad f_{1y} = 0, \quad f_{0y} + f_{1x} = 0, \tag{2.8}$$

$$-f_1\Omega + h_x = 0, \quad f_0\Omega + h_y = 0, \quad f_0V_x + f_1V_y = 0. \tag{2.9}$$

Thus

$$f_0 = \alpha y + \beta, \quad f_1 = -\alpha x + \gamma. \tag{2.10}$$

The compatibility condition $h_{xy} = h_{yx}$ for (2.9) gives

$$f_0\Omega_x + f_1\Omega_y = 0.$$

For $\alpha = 0$ we find

$$V = V(\xi), \quad \Omega = \Omega(\xi), \quad h = h(\xi), \qquad \xi = \gamma x - \beta y, \quad \dot{h}(\xi) = \Omega(\xi). \tag{2.11}$$

For $\alpha \neq 0$ an adequate translation in x and y permits the choice $\beta = \gamma = 0$ so that, putting $\alpha = 1$, we have

$$V = V(\rho), \quad \Omega = \Omega(\rho), \quad h = h(\rho), \qquad \rho = \sqrt{x^2 + y^2}, \quad \dot{h}(\rho) = -\rho\Omega. \tag{2.12}$$

We see that a linear integral of motion (2.7) exists if and only if both V and Ω are invariant either under translations in some direction, or under rotations about some point (which we have translated to the origin).

Hall[12] has studied the less restrictive case of a linear constant depending explicitly on a fixed value of the energy. In that context he found extended families of solutions.

B. Quadratic integral of motion

Let us now consider a quadratic constant of motion C in the form (1.3). The condition $dC/dt = 0$ directly implies

$$g_{0x} = 0, \quad g_{2y} = 0, \quad g_{0y} + g_{1x} = 0, \quad g_{1y} + g_{2x} = 0, \tag{2.13}$$

so that

$$g_0 = \alpha y^2 - \beta y + \delta, \qquad g_1 = -2\alpha xy + \beta x - \gamma y + \xi, \qquad g_2 = \alpha x^2 + \gamma x + \zeta, \tag{2.14}$$

and the constant C reduces to

$$C = \alpha(x\dot{y} - y\dot{x})^2 + (x\dot{y} - y\dot{x})(\beta\dot{x} + \gamma\dot{y}) + \delta\dot{x}^2 + \zeta\dot{y}^2 + \xi\dot{x}\dot{y} + k_0(x,y)\dot{x} + k_1(x,y)\dot{y} + h(x,y). \tag{2.15}$$

It was shown in Ref. 12 that a weaker requirement, namely, that C be a configurational invariant that may depend on the energy E, implies that $F \equiv \frac{1}{2}(g_0 - g_2) + (i/2)g_1$ should be an analytic function of $z \equiv x + iy$. It is a simple matter to verify that (2.15) satisfies this requirement. Since α, β, γ, δ, ζ, and ξ are constants, the quadratic part of C can be interpreted as a second-order element in the enveloping algebra of the Lie algebra e(2), with a basis consisting of two translations $\dot{x}$ and $\dot{y}$ and the rotation $(x\dot{y} - y\dot{x})$. Performing a Euclidean transformation of the x,y plane (in general a complex one) including reflections of x and y, and if necessary taking linear combinations of C and H, we can reduce C to one of the following seven cases[5,7]: (1) the Cartesian case, $\alpha = \beta = \gamma = 0, \xi/(\delta - \zeta) \neq \epsilon i, \epsilon = \pm 1$,

$$C_1 = \dot{x}^2 + k_0\dot{x} + k_1\dot{y} + h; \tag{2.16}$$

(2) the degenerate Cartesian case, $\alpha = \beta = \gamma = 0$, $\xi/(\delta - \zeta) = \epsilon i$,

$$C_2 = \dot{x}(\dot{x} + i\dot{y}) + k_0\dot{x} + k_1\dot{y} + h; \tag{2.17}$$

(3) the parabolic case, $\alpha = 0, \beta^2 + \gamma^2 \neq 0$,

$$C_3 = \dot{x}(x\dot{y} - y\dot{x}) + k_0\dot{x} + k_1\dot{y} + h; \tag{2.18}$$

(4) the degenerate parabolic case, $\alpha = 0, \gamma = \epsilon i\beta \neq 0$,

$$C_4 = (\dot{x} + i\dot{y})(x\dot{y} - y\dot{x}) + k_0\dot{x} + k_1\dot{y} + h. \tag{2.19}$$

In the case $\alpha \neq 0$, define

$$\lambda_1 = \frac{\beta\gamma}{2\alpha} - \xi, \quad \lambda_2 = \frac{\beta^2 - \gamma^2}{4\alpha} + \zeta - \delta, \tag{2.20}$$

$$\sigma = (1/2\alpha)(\lambda_1^2 + \lambda_2^2)^{1/2}. \tag{2.21}$$

We then have (5) the spherical case, $\alpha \neq 0, \lambda_1 = \lambda_2 = 0$,

$$C_5 = (x\dot{y} - y\dot{x})^2 - k_0\dot{x} + k_1\dot{y} + h; \tag{2.22}$$

(6) the elliptic case, $\alpha \neq 0, (\lambda_1, \lambda_2) \neq (0,0), \lambda_1 \neq \pm i\lambda_2$,

$$C_6 = (x\dot{y} - y\dot{x})^2 + \sigma(\dot{x}^2 - \dot{y}^2) + k_0\dot{x} + k_1\dot{y} + h \tag{2.23}$$

(σ is the focal distance); and (7) the degenerate elliptic case, $\alpha \neq 0, \lambda_1 = \pm i\lambda_2 \neq 0$,

$$C_7 = (x\dot{y} - y\dot{x})^2 + (\dot{x} + i\dot{y})^2 + k_0\dot{x} + k_1\dot{y} + h. \tag{2.24}$$

In addition to (2.13) the condition $dC/dt = 0$ implies the following system of differential equations:

$$\begin{aligned} &k_{0x} - g_1\Omega = 0, \quad k_{1y} + g_1\Omega = 0, \\ &2g_0\Omega - 2g_2\Omega + k_{0y} + k_{1x} = 0, \\ &-2g_0V_x - g_1V_y - k_1\Omega + h_x = 0, \\ &-g_1V_x - 2g_2V_y + k_0\Omega + h_y = 0, \\ &k_0V_x + k_1V_y = 0, \end{aligned} \tag{2.25}$$

with g_i as in (2.14).

In general these equations are not easy to solve. The polynomials g_i should be chosen in one of the "standard" forms implied by the expressions $C_1, \ldots, C_7$ above and each case must be considered separately. In this article we restrict ourselves to the Cartesian and degenerate Cartesian cases, i.e., we put $\alpha = \beta = \gamma$ in (2.14), so that g_0, g_1, and g_2 in the integral of motion (1.3) are constants.

III. THE CARTESIAN CASE

A. Derivation of basic equations

Consider the case $\alpha = \beta = \gamma = 0$, $\xi/(\delta - \zeta) \neq \pm i$ in Eqs. (2.14). Performing an appropriate rotation and linear combination with H, we reduce C to the form C_1 of (2.16):

$$C = \tfrac{1}{2}\dot{x}^2 + k_0\dot{x} + k_1\dot{y} + h. \tag{3.1}$$

Equations (2.25), following from the condition $dC/dt = 0$, reduce to

$$\Omega + k_{0y} + k_{1x} = 0, \quad k_{0x} = 0, \quad k_{1y} = 0, \tag{3.2}$$

$$k_0\Omega + h_y = 0, \quad -k_1\Omega + h_x - V_x = 0, \tag{3.3}$$

$$k_0V_x + k_1V_y = 0. \tag{3.4}$$

The case of a purely scalar potential, considered earlier[5-7] corresponds to $k_0 = k_1 = 0, \Omega = 0, V(x,y) = V_1(x) + V_2(y)$. From now on we assume that at least one of the quantities k_0 or k_1 is not identically zero.

Equations (3.2) can be immediately solved and imply

$$k_0 = -g_y(y), \quad k_1 = -f_x(x), \quad \Omega = f_{xx}(x) + g_{yy}(y). \tag{3.5}$$

The compatibility condition $h_{xy} = h_{yx}$ for the two equations (3.3) is

$$V_{xy} + k_1\Omega_y + k_0\Omega_x = 0,$$

which can be integrated to yield

$$V(x,y) = f(x)g_{yy}(y) + g(y)f_{xx}(x) + u(x) + v(y), \tag{3.6}$$

where $u(x), f(x), v(y)$, and $g(y)$ are functions of one variable as indicated. Equation (3.4) reduces to

$$g_y[gf_{xxx} + f_x g_{yy} + u_x] + f_x[fg_{yyy} + g_y f_{xx} + v_y] = 0. \tag{3.7}$$

Defining

$$F(x) = u + f_x^2/2, \quad G(y) = v + g_y^2/2, \tag{3.8}$$

and performing some simple manipulations, we transform (3.7) to the form

$$\frac{G_y}{g_y} + \frac{F_x}{f_x} + g\frac{f_{xxx}}{f_x} + f\frac{g_{yyy}}{g_y} = 0. \tag{3.9}$$

Taking the mixed derivative $\partial^2/\partial x\,\partial y$ of (3.9) and separating variables we find

$$\frac{1}{f_x}\frac{d}{dx}\frac{f_{xxx}}{f_x} = -\frac{1}{g_y}\frac{d}{dy}\frac{g_{yyy}}{g_y} = 2\alpha. \tag{3.10}$$

Integrating (3.10) we obtain the following equations for the functions $f(x)$ and $g(y)$:

$$\begin{aligned} f_{xx} &= \alpha f^2 + \beta f + \gamma, \\ g_{yy} &= -\alpha g^2 + \delta g + \zeta, \end{aligned} \tag{3.11}$$

where $\alpha, \beta, \gamma, \delta$, and ζ are constants.

Returning to Eqs. (3.3) and (3.4) and their consequences, we can express the potentials and all other relevant quantities in terms of the functions $f(x)$ and $g(y)$:

$$\begin{aligned} \Omega &= \alpha(f^2 - g^2) + \beta f + \delta g + \gamma + \zeta, \\ V &= \frac{\alpha}{3}(g-f)^3 - \frac{\beta+\delta}{2}(g-f)^2 + (\gamma + \kappa - \zeta)(g-f), \\ k_0 &= -g_y, \quad k_1 = -f_x, \\ h &= -(\alpha/3)(g^3 + 2f^3 - 3gf^2) + \beta(fg - f^2) \\ &\quad + \tfrac{1}{2}\delta(g^2 - f^2) + \gamma(g - 2f) + \zeta g - \kappa f. \end{aligned} \tag{3.12}$$

All the Greek letters represent arbitrary constants. The functions f and g introduced in (3.5) are defined up to an arbitrary additive constant.

The problem of constructing the integrable Hamiltonian H, the constant C and the field Ω has thus been reduced to solving Eqs. (3.11). Let us now examine these equations.

B. Analysis and solution of basic equations

(1a) $\alpha=\beta=\delta=0$, $\gamma\zeta \neq 0$: In this case f and g are second-order polynomials and we can, after a possible translation, put

$$V = \tfrac{1}{2}(Ax^2 + By^2), \quad \Omega = \Omega_0, \tag{3.13}$$

where A, B, and Ω_0 are constants.

This case describes a harmonic oscillator in a homogeneous magnetic field and the equations of motion can be solved explicitly. (See Sec. IV.)

(1b) $\alpha=\beta=\delta=0$, $\zeta=0$, $\gamma \neq 0$ (or $\gamma=0$, $\zeta \neq 0$): In this case we have

$$V = \tfrac{1}{2}Ax^2 + By, \quad \Omega = \Omega_0 \tag{3.14}$$

(A, B, and Ω_0 are constants).

(2a) $\alpha=0$, $\beta\delta \neq 0$: Adding constants to f and g we can eliminate the constant terms in (3.11) to obtain

$$f_{xx} = \beta f, \quad g_{yy} = \delta g. \tag{3.15}$$

Depending on the signs of β and δ we find that f and g are either trigonometric or hyperbolic functions.

(2b) $\alpha=0$, $\delta=0$, $\beta \neq 0$ (or $\delta \neq 0$, $\beta=0$): We find that f is a trigonometric or hyperbolic function and g a quadratic polynomial (or vice versa).

(3) $\alpha \neq 0$: Adding appropriate constants to f and g we reduce (3.11) to

$$f_{xx} = \alpha f^2 + \gamma, \quad g_{yy} = -\alpha g^2 + \zeta. \tag{3.16}$$

Putting

$$\begin{aligned} f &= 2\sqrt{3\epsilon_1\gamma/\alpha}\,F((\epsilon_1 a\gamma/3)^{1/4}x), \\ g &= -2\sqrt{-3\epsilon_2\zeta/\alpha}\,F((-\epsilon_2\alpha\zeta/3)^{1/4}y), \end{aligned} \tag{3.17}$$

where $\epsilon_i = \pm 1$ and $F(z)$ satisfies

$$F'' = 6F^2 + \epsilon/2, \quad \epsilon = \pm 1, \tag{3.18}$$

we express both $f(x)$ and $g(y)$ in terms of the Weierstrass elliptic function

$$F(z) = P(z - k; \epsilon, h), \tag{3.19}$$

where k and h are arbitrary constants. The function $P(z; g_2, g_3)$ is, in general, a doubly periodic function of the complex variable z, analytic in z except for an infinite number of second-order poles (e.g., at $z = 0$). In the limit when one or both of the periods becomes infinite, the Weierstrass elliptic function is expressed in terms of elementary functions (trigonometric, hyperbolic, or inverse powers). Notice that, e.g., for $\gamma = 0$, $\zeta = 0$, in (3.16) a particular solution is

$$f(x) = 6/\alpha x^2, \quad g(y) = -6/\alpha y^2. \tag{3.20}$$

Many other special cases can be extracted from Eq. (3.11).

IV. EXAMPLES OF TRAJECTORIES AND POINCARÉ SECTIONS

In this section we shall study numerically some examples of Hamiltonian systems of the type (2.4) and (2.5), both integrable and nonintegrable ones.

Let us first consider the integrable case (3.13), i.e., $\alpha = \beta = \delta = 0$ in Eqs. (3.11). Thus, we have $V = \frac{1}{2}(Ax^2 + By^2)$, $\Omega = \text{const}$, and the equations of motion can be integrated analytically. Equations (2.4) in this case lead to a fourth-order differential equation for x and y separately and, e.g., for x we obtain

$$x^{(4)} + (A + B + \Omega^2)\ddot{x} + ABx = 0. \tag{4.1}$$

The solutions of the characteristic equation are

$$\begin{aligned} r_j^2 &= \{-(A + B + \Omega^2) \\ &\quad + \epsilon_j[(A + B + \Omega^2)^2 - 4AB]^{1/2}/2, \\ \epsilon_j &= \pm 1, \quad j = 1,2. \end{aligned} \tag{4.2}$$

Assuming that all four roots $\pm r_1$, $\pm r_2$ are different, we obtain

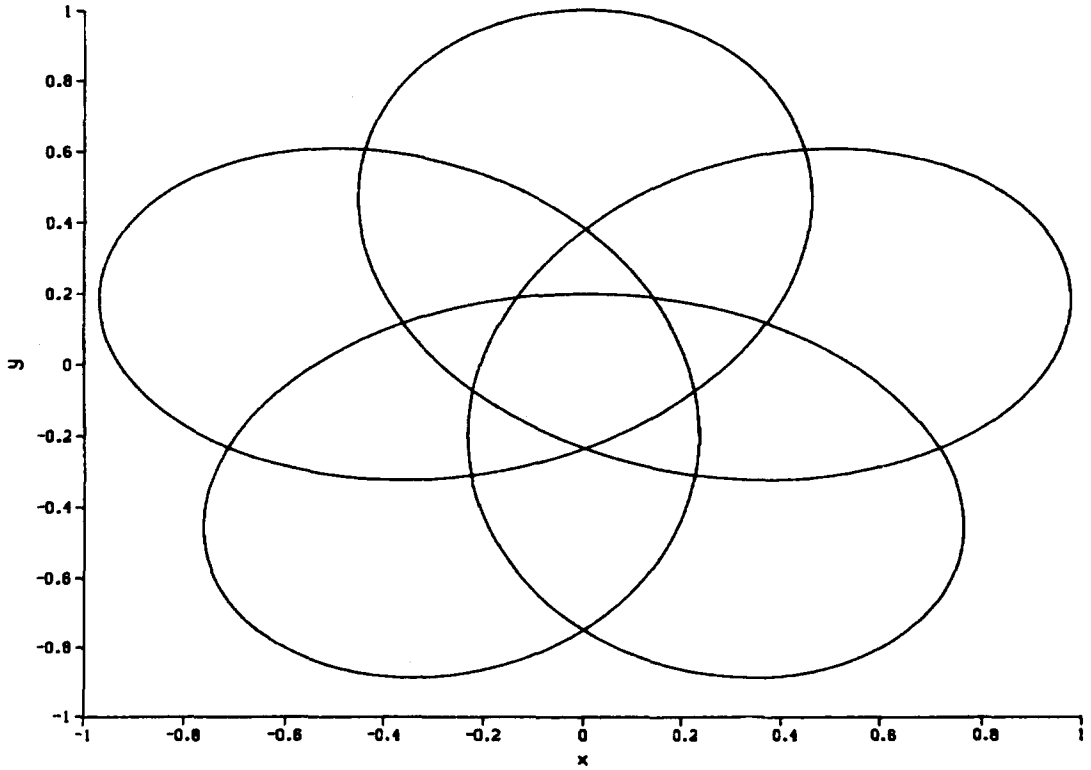


FIG. 1. A periodic trajectory for the potential $V = (x^2 + y^2)/2$ and constant magnetic field $\Omega = \frac{3}{2}$. The period is $T = 4\pi$.

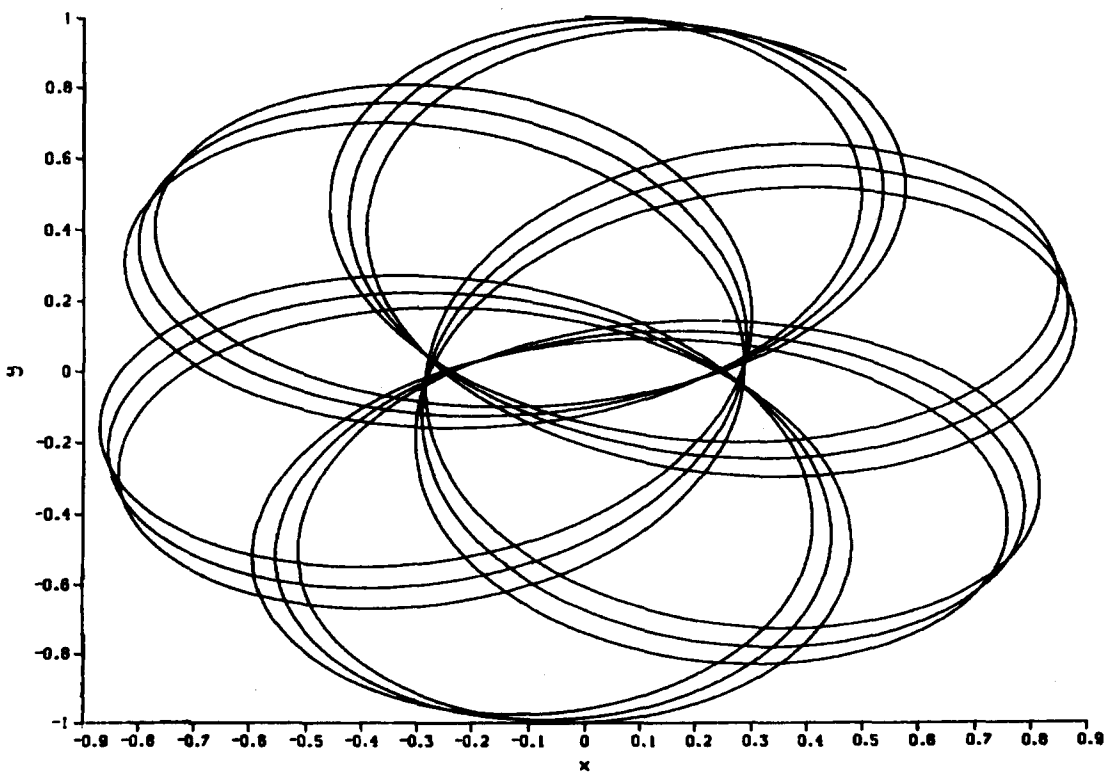


FIG. 2. A nonperiodic bounded trajectory for the integrable Hamiltonian with $V = x^2/2 + y^2/4$ and $\Omega = \frac{3}{2}$. The frequencies in this case are not comeasurable.

$$x(t) = ae^{r_1 t} + be^{-r_1 t} + ce^{r_2 t} + de^{-r_2 t},$$
$$y(t) = (1/\Omega r_1 r_2)[r_2(A + r_1^2)(ae^{r_1 t} - be^{-r_1 t}) + r_1(A + r_2^2)(ce^{r_2 t} - de^{-r_2 t})], \quad (4.3)$$

where $a,\ldots,d$ are arbitrary constants. We see that the motion is bounded (and quasiperiodic, i.e., restricted to a torus in phase space) if r_1 and r_2 are pure imaginary. This happens if

$$(A + B + \Omega^2)^2 > 4AB > 0 \quad \text{and} \quad A + B + \Omega^2 > 0. \quad (4.4)$$

The trajectories are actually periodic in configuration space if r_1/r_2 is a rational number.

In Fig. 1 we present a trajectory for the case $A = B = 1$, $\Omega = \frac{3}{2}$. Conditions (4.4) are satisfied and we have $r_1 = 2i$, $r_2 = i/2$, so the trajectories are periodic with period $T = 4\pi$. An integrable bounded, but nonperiodic trajectory is shown in Fig. 2, where we have chosen $A = 1$, $B = \frac{1}{2}$, and $\Omega = \frac{3}{2}$.

A very useful tool in studies of two-dimensional Hamiltonian systems are the Poincaré sections, defined as follows. Consider a Hamiltonian system with Hamiltonian $H(x,y,p_x,p_y)$. The phase space is of dimension 4, but since the total energy $H = E$ is an integral of motion, the trajectories in phase space are restricted to a three-dimensional manifold, so that, e.g., p_x can be expressed in terms of p_y, x, and y. A Poincaré section is obtained by intersecting this three-dimensional fixed energy manifold by a hyperplane, e.g., the hyperplane $x = 0$. If the system is integrable, the Poincaré sections are regular, since they are sections of a torus. For chaotic systems, on the other hand, points are scattered on the Poincaré sections in an irregular manner.

Figure 3 represents such a regular Poincaré section for the integrable system discussed above with $A = 1$, $B = \frac{1}{2}$, and $\Omega = \frac{3}{2}$ and initial position $(x_0,y_0) = (1,0)$.

For comparison we also have studied some nonintegrable Hamiltonian systems. We again consider the potential $V = \frac{1}{2}x^2 + \frac{1}{4}y^2$, but introduce a nonconstant magnetic field

$$\Omega = \tfrac{3}{2} + cx, \quad c = \text{const} \quad (4.5)$$

(the only integrable case corresponds to $c = 0$). For small values of c ($c < 1$) the system remains close to the integrable one and the lack of integrability is not visible in the numerical studies. For $c = \frac{3}{2}$ a trajectory is shown in Fig. 4 and it does not seem to differ qualitatively from the nonperiodic "integrable" trajectory of Fig. 2. The fact that the system is not integrable manifests itself quite clearly on the Poincaré section of Fig. 5.

Finally, in Fig. 6, we show a trajectory for a different integrable system, namely,

$$V = (\cosh \alpha x + \cosh \beta y)^2, \quad \Omega = \alpha^2 \cosh \alpha x - \beta^2 \cosh \beta y, \quad (4.6)$$

corresponding to the case when Eqs. (3.11) are linear. The trajectories in this case remain bounded.

The trajectories for trigonometric-type potentials, or the doubly periodic ones involving Weierstrass elliptic functions, do not, unfortunately, remain bounded, so there is little point in presenting them on figures. In the case of the elliptic functions this is due to the fact that the quadratic terms in the two equations (3.11) have opposite signs, so either the f or g contribution to the potential will be repulsive and the other one attractive [see, e.g., (3.20)].

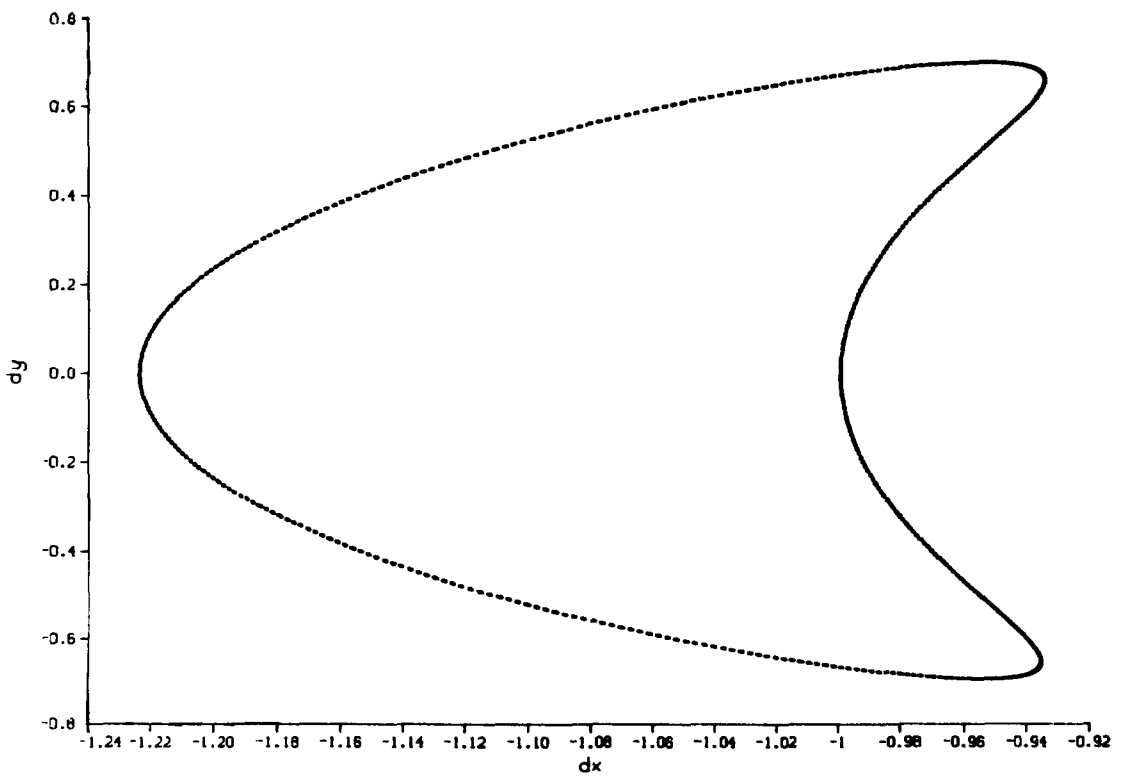


FIG. 3. A Poincaré section for the integrable system with $V = x^2/2 + y^2/4$ and $\Omega = \frac{3}{2}$.

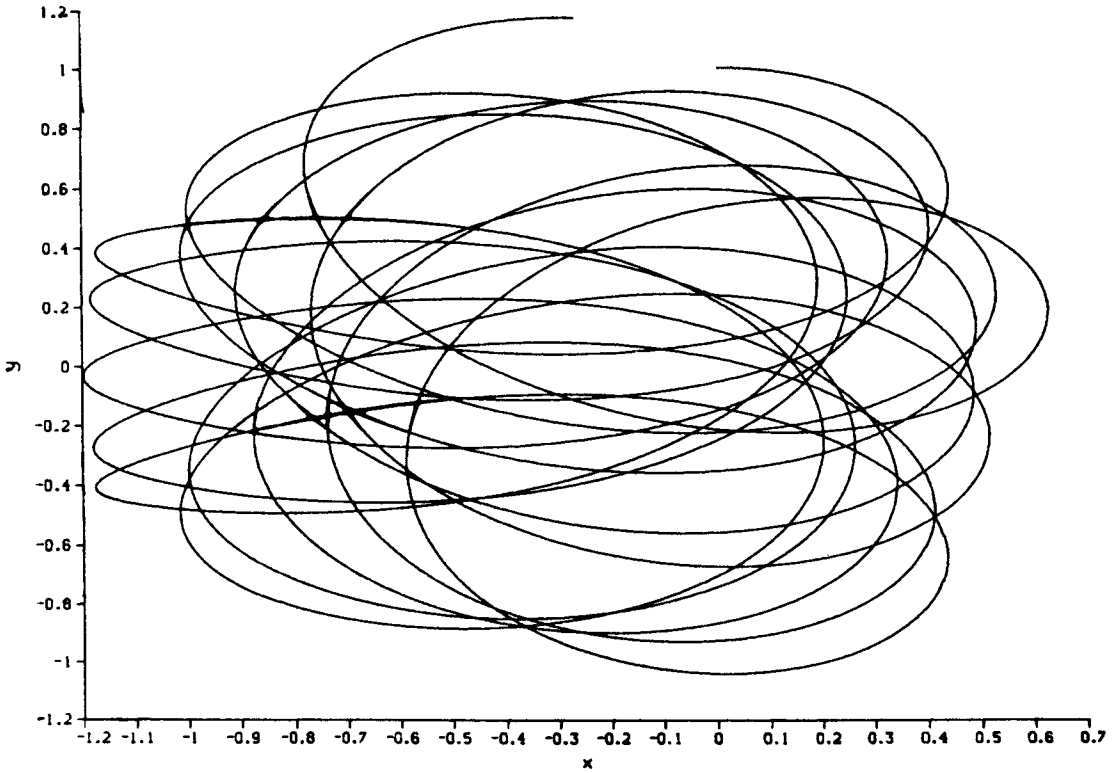


FIG. 4. A bounded nonperiodic trajectory for a perturbed nonintegrable system with $V = x^2/2 + y^2/4$ and magnetic field $\Omega = 3(1 + x)/2$.

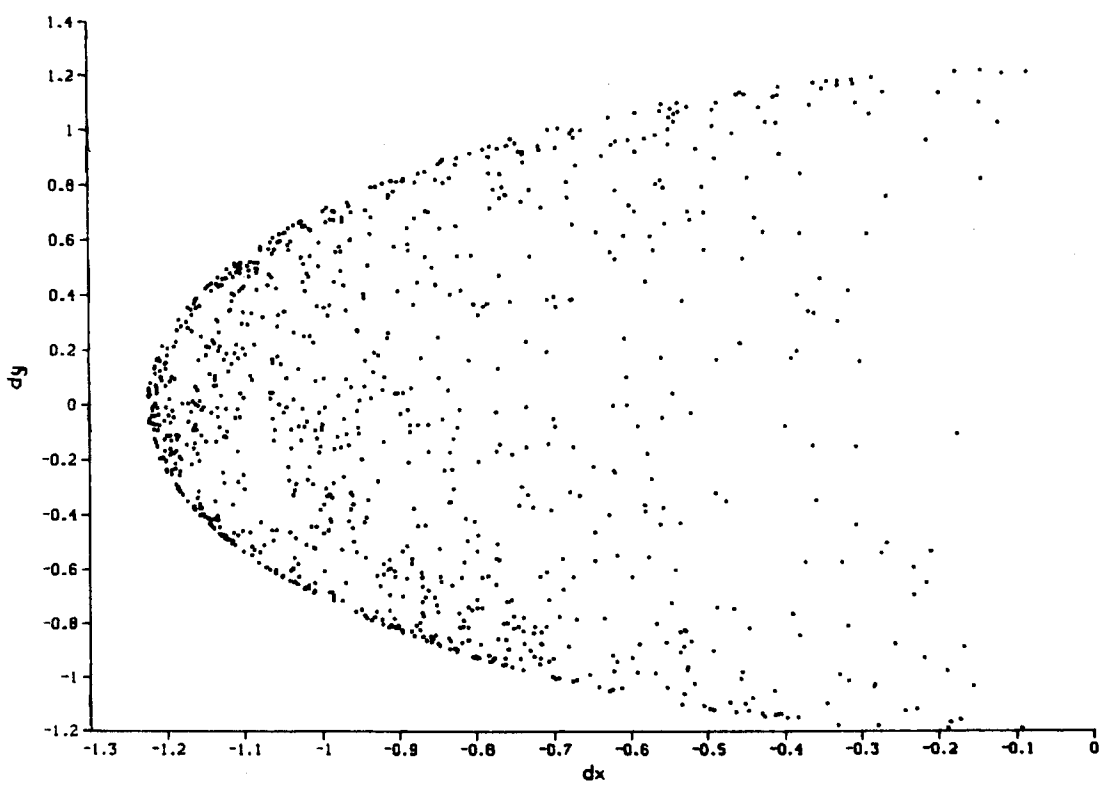


FIG. 5. A Poincaré section for the same nonintegrable system as in Fig. 4.

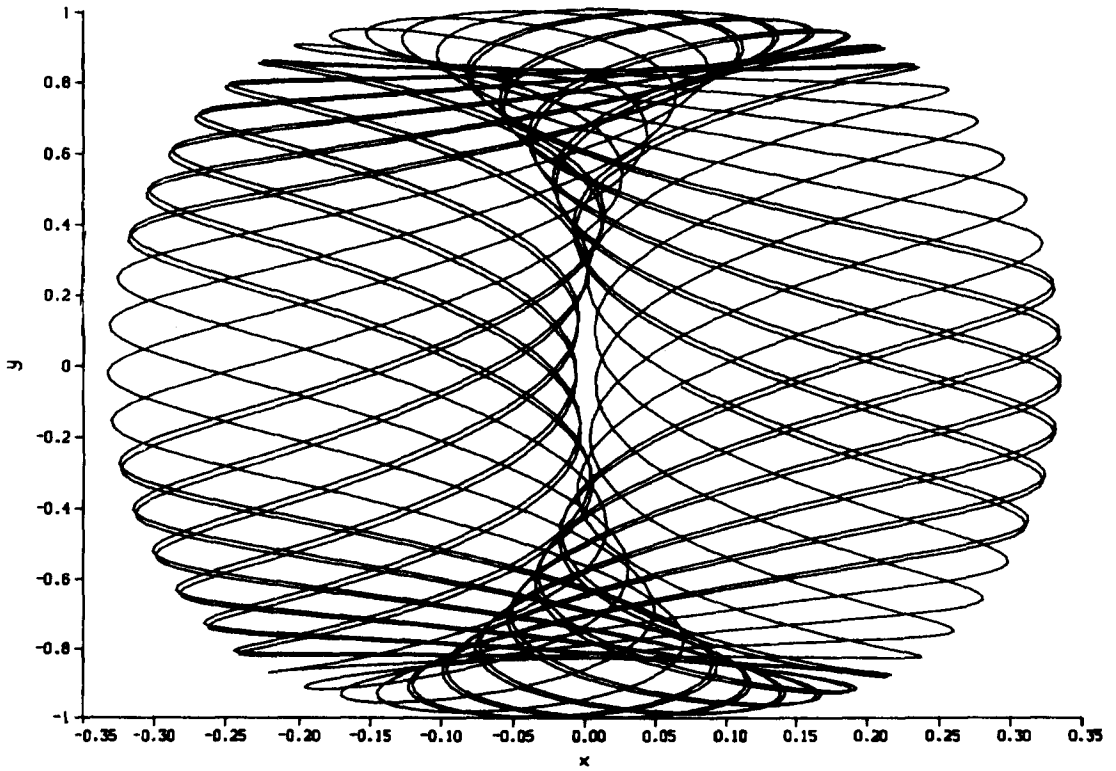


FIG. 6. A trajectory for the integrable system with $V=(\cosh \alpha x+\cosh \beta y)^2$ and $\Omega=\alpha^2 \cosh \alpha x-\beta^2 \cosh \beta y$.

V. PAINLEVÉ ANALYSIS OF THE CARTESIAN CASE

We recall that the equations of motion are

$$\ddot{x}=-V_x+\Omega \dot{y}, \quad \ddot{y}=-V_y-\Omega \dot{x}, \tag{5.1}$$

where Ω and V are expressed in terms of two elliptic functions $f(x)$ and $g(y)$ by the following steps.

If f and g satisfy

$$f_{xx}=\alpha f^2+\beta f+\gamma, \quad g_{yy}=-\alpha g^2+\delta g+\zeta, \tag{5.2}$$

then $\Omega=f_{xx}+g_{yy}$ and

$$V=(\alpha/3)(g-f)^3-[(\beta+\delta)/2](g-f)^2+\kappa(g-f)+\nu,$$

where κ and ν are arbitrary constants. In the following, α will be taken equal to 3.

The singularity analysis is somewhat unusual in this case, because when x and y go to infinity, the elliptic functions do not go to any limit. This prevents x and y from going to infinity in a finite time. Rather, the singularities at finite times occur when x and y go to poles, respectively, of f and g. Indeed, when x goes to a pole of f at a finite time t_0, y cannot go to a regular point of g but must go to a pole as well.

Let x_0 be a pole of f and y_0 a pole of g. From Eqs. (5.2) one can see that f behaves as

$$f \sim 2/(x-x_0)^2,$$

while g behaves as

$$g \sim -2/(y-y_0)^2.$$

Balancing the leading powers in (5.1), we see that $x-x_0$ and $y-y_0$ must behave as $z^{1/4}$, where $z=t-t_0$. Note that all the terms in Eq. (5.1) are dominant and behave as $z^{-7/4}$. As usual, t_0 is one of the free parameters of the expansion but x_0 and y_0 are fixed as they must be the locations of poles of the *given* functions f and g. From there on, the Painlevé analysis follows as usual: We write

$$x-x_0 \sim \alpha z^{1/4}, \quad y-y_0 \sim \beta z^{1/4}.$$

Balancing the coefficients of the leading powers, we find only one relation between α and β, namely,

$$\alpha^3 \beta^3=16(\alpha^2+\beta^2).$$

The resonances are found to be at $-1, 0, \frac{1}{2}$, and $\frac{3}{2}$. As usual, -1 is related to the arbitrariness of t_0 and 0 reflects the freedom of α (or β), since we have only one relationship between α and β. The other resonances are half-integers, therefore the expansion does not have the full Painlevé property. Still the expansion may be of "weak Painlevé" type[15] because a power $z^{1/2}$ in the expansion is a "natural" one since the leading behavior goes as $z^{1/4}$. In order to check for the resonance condition, one must first expand f and g to order 3 in $(x-x_0)$ and $(y-y_0)$, respectively, then substitute expansions of $(x-x_0)$ and $(y-y_0)$ in terms of z. Checking the resonance conditions (especially at order $\frac{3}{2}$) is a very tedious task, which was accomplished using the REDUCE formal language. Both resonance conditions were actually found to be satisfied and no logarithm enters in the expansion, which is thus indeed of the "weak Painlevé type," as expected for integrability.

VI. CONCLUSIONS

The main result of this paper can be summarized as follows.

(1) We have shown that a nonrelativistic Hamiltonian system with the Hamiltonian (1.1), involving both scalar and vector potentials, admits a quadratic integral of motion C, only if the terms in C that are quadratic in the velocities are actually second-order operators in the enveloping algebra of the Euclidean Lie algebra with the basis $\{\dot{x}, \dot{y}, x\dot{y} - \dot{x}y\}$.

(2) We have completely analyzed the case of "Cartesian" and "degenerate Cartesian" integrals of motion, i.e., the case when the coefficients g_i of the second-order terms in the constant C are independent of the coordinates. The most general integrable system of the Cartesian type is given by (3.12), where $f(x)$ and $g(y)$ are solutions of the elliptic-function-type ordinary differential equations (3.11). In general, $f(x)$ and $g(y)$ are doubly periodic functions of the (complex) variables x and y, respectively.

(3) We have investigated periodic and nonperiodic bounded trajectories and also Poincaré sections for several of the obtained integrable systems and compared them with some nonintegrable ones.

(4) We have shown that, in general, the obtained integrable systems manifest the "weak Painlevé property," i.e., the solutions may have moving poles and rational branch points, not, however, logarithmic ones. The case of a pure scalar potential is an exception. Then $V(x,y) = F(x) + G(y)$ and $\Omega = k_0 = k_1 = 0$. The functions $F(x)$ and $G(y)$ are arbitrary, the equations of motion separate in Cartesian coordinates, and the singularity structure of the solutions can be arbitrary [depending on the form of $F(x)$ and $G(y)$].

Several comments are in order here.

(i) In general the vector and scalar potentials will be doubly periodic in the complex x and y planes, though of course one or both periods may, in special cases, be infinite. In the general case the motion is unbounded. Periodic potentials are of obvious interest in crystallography and solid state physics and we plan to return to the case of integrable Hamiltonian systems with potentials expressed in terms of Weierstrass elliptic functions.

(ii) The case when a magnetic field is present is much more difficult to treat than the case of a purely scalar potential, when the existence of an integral of motion quadratic in the velocities is tantamount to the separability of the Hamiltonian in one of several coordinate systems. This is the reason why we have restricted ourselves to two of seven classes of quadratic integrals of motion. The other classes are presently under study.

ACKNOWLEDGMENTS

The authors thank R. Padjen and V. Papageorgiou for many helpful discussions.

One of the authors (P. W.) is indebted to DICET for financial support and to R. Padjen for hospitality at the Centre National d'Etude de Télecommunications (CNET), Paris where a large part of this study was performed. He (P. W.) also was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the *Fonds FCAC pour l'aide et le soutien à la recherche du Gouvernement du Québec.*

APPENDIX: THE DEGENERATE CARTESIAN CASE

As mentioned in Sec. II, if the variables x and y, as well as the potentials, are allowed to be complex, isotropic vectors exist in velocity space and they lead to "degenerate" orbits of second-order integrals of motion. One such integral is the "degenerate Cartesian integral"

$$C = \dot{x}(\dot{x} + i\dot{y}) + \tilde{k}_0\,\dot{x} + \tilde{k}_1\,\dot{y} + h. \tag{A1}$$

[See (2.17) and Ref. 7.]

In the coordinates

$$z = x + iy, \quad \bar{z} = x - iy,$$

the constant C reduces to

$$C = \tfrac{1}{2}(\dot{z}^2 + \dot{z}\dot{\bar{z}}) + k_0\,\dot{z} + k_1\,\dot{\bar{z}} + h, \tag{A2}$$

and the equations of the motion are

$$\ddot{z} = -2V_{\bar{z}} - i\Omega\dot{z}, \quad \ddot{\bar{z}} = 2V_z + i\Omega\dot{\bar{z}}. \tag{A3}$$

Putting

$$\frac{dC}{dt} = 0,$$

using the equations of motion, and equating the coefficients of each term to zero, we obtain

$$k_{0,z} - i\Omega = 0, \quad k_{0,\bar{z}} + k_{1,z} = 0, \quad k_{1,\bar{z}} = 0, \tag{A4}$$

$$h_z - 2V_{\bar{z}} - V_z - ik_0\,\Omega = 0, \quad h_{\bar{z}} - V_{\bar{z}} + ik_1\,\Omega = 0, \tag{A5}$$

$$k_0 V_{\bar{z}} = k_1 V_z = 0. \tag{A6}$$

We can easily integrate for k_0, k_1, and Ω:

$$k_1 = \phi_1(z), \quad k_0 = -\phi_1'(z)\bar{z} + \phi_0(z), \tag{A7}$$

$$i\Omega = \phi_0'(z) - \phi_1''(z)\bar{z}. \tag{A8}$$

The compatibility condition $h_{z\bar{z}} = h_{\bar{z}z}$ for Eqs. (A5) implies

$$2V_{\bar{z}\bar{z}} = (\phi_0\phi_1'' - \phi_1\phi_0'') + \bar{z}(\phi_1\phi_1''' - \phi_1'\phi_1''), \tag{A9}$$

leading to

$$2V = u_0(z) + \bar{z}u_1(z) + (\bar{z}^2/2)(\phi_0\phi_1'' - \phi_1\phi_0'') + (\bar{z}^3/6)(\phi_1\phi_1''' - \phi_1'\phi_1''). \tag{A10}$$

Equation (A6) now reduces to

$$(-\phi_1'\bar{z} + \phi_0)(u_1 + \bar{z}A(z) + (\bar{z}^2/2)B(z)) + \phi_1(u_0' + \bar{z}u_1' + (\bar{z}^2/2)A' + (\bar{z}^3/6)B'(z)) = 0,$$

where

$$A(z) = \phi_0\phi_1'' - \phi_1\phi_0'', \quad B(z) = \phi_1\phi_1''' - \phi_1'\phi_1''.$$

Equating to zero the coefficient of each monomial in $\bar{z}$, we obtain

$$\begin{aligned} &-\tfrac{1}{2}\phi_1' B(z) + (\phi_1/6)B'(z) = 0, \\ &-\phi_1' A(z) + (\phi_0/2)B(z) + (\phi_1/2)A'(z) = 0, \\ &-\phi_1' u_1 + \phi_0 A(z) + u_1'\phi_1 = 0, \\ &u_1\phi_0 + u_0'\phi_1 = 0. \end{aligned} \tag{A11}$$

From the first equation we immediately deduce

$$B(z) = k\phi_1^3,$$

which yields an equation for ϕ_1 only:

$$\phi_1\phi_1''' - \phi_1'\phi_1'' = k\phi_1^3. \tag{A12}$$

Once ϕ_1 is known, we obtain a linear third-order equation for ϕ_0. The two last equations allow the calculation of u_1 and u_0 in terms of ϕ_0 and ϕ_1.

Let us now integrate Eq. (A12). Multiplying by ϕ_1''/ϕ_1^3 we obtain

$$\phi_1''\phi_1'''/\phi_1^2 - \phi_1'\phi_1''^2/\phi_1^3 = k\phi_1'',$$

that is,

$$(\phi_1''^2/\phi_1^2)' = 2k\phi_1'', \quad \phi_1''^2 = 2k(\phi_1'\phi_1^2 + l\phi_1^2),$$

if $k \neq 0$, we set

$$u'' = \phi_1' + l, \tag{A13}$$

thus $u' = \phi_1 + lz + m$,

$$\begin{aligned} &\phi_1 = u' - lz - m, \\ &u'''^2 = (2ku'')(u' - lz - m)^2, \\ &u'''/\sqrt{u''} = \sqrt{2k}(u' - lz - m), \\ &u'' = \rho(u - lz^2 - mz - n)^2, \quad \rho = \text{const}, \end{aligned} \tag{A14}$$

that is, u is an elliptic function and ϕ_1 is defined in terms of u by (A13).

[1]H. Toda, J. Phys. Soc. Jpn. **23**, 501 (1967).
[2]F. Calogero, J. Math. Phys. **12**, 419 (1971).
[3]F. Calogero, "Integrable dynamical systems and related mathematical results," in *Nonlinear Phenomena, Lecture Notes in Physics*, Vol. 189 (Springer, Berlin, 1983).
[4]J. Bertrand, J. Math. (i) **XVII**, 121 (1852).
[5]I. Friš, V. Mandrosov, J. Smorodinsky, M. Uhlíř, and P. Winternitz, Phys. Lett. **16**, 354 (1965); P. Winternitz, Ya. A. Smorodinskiĭ, M. Uhlíř, and I. Friš, Yad. Fiz. **4**, 625 (1966) [Sov. J. Nucl. Phys. **4**, 444 (1967)].
[6]A. Makarov, J. Smorodinsky, Kh. Valiev, and P. Winternitz, Nuovo Cimento **52** A, 1061 (1967).
[7]B. Dorizzi, B. Grammaticos, and A. Ramani, J. Math. Phys. **24**, 2282 (1983).
[8]J. Hietarinta, Phys. Rev. Lett. **52**, 1057 (1984).
[9]M. J. Ablowitz, A. Ramani, and H. Segur, Lett. Nuovo Cimento **23**, 333 (1978).
[10]C. R. Holt, J. Math. Phys. **23**, 1037 (1982).
[11]B. Grammaticos, B. Dorizzi, and A. Ramani, J. Math. Phys. **24**, 2289 (1983).
[12]L. Hall, Physica D **8**, 90 (1983).
[13]V. I. Inozemtsev, J. Phys. A **17**, 815 (1984).
[14]B. Dorizzi, B. Grammaticos, and A. Ramani, J. Math. Phys. **25**, 481 (1984).
[15]A. Ramani, B. Dorizzi, and B. Grammaticos, Phys. Rev. Lett. **49**, 1539 (1982).

Characteristic functional structure of infinitesimal symmetry mappings of classical dynamical systems. I. Velocity-dependent mappings of second-order differential equations

Gerald H. Katzin
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202
Jack Levine
Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205

(Received 14 May 1985; accepted for publication 18 June 1985)

The primary purpose of this paper is to show that infinitesimal velocity-dependent symmetry mappings [(a) $\bar{x}^i = x^i + \delta x^i$, $\delta x^i \equiv \xi^i(\dot{x},x,t)\delta a$ with associated change in path parameter (b) $\bar{t} = t + \delta t$, $\delta t \equiv \xi^0(\dot{x},x,t)$] of classical (including relativistic) particle systems (c) $E^i(\ddot{x},\dot{x},x,t) = 0$ are expressible in a form with a characteristic functional structure which is the same for all dynamical systems (c) and is manifestly dependent upon constants of motion of the system. In this characteristic form the symmetry mappings are determined by (d) $\xi^i = Z^i(\dot{x},x,t) + \dot{x}^i\xi^0, \xi^0$ arbitrary; the functions Z^i appearing in (d) have the form (e) $Z^i = B^A g^i_A(C^1,...,C^r;t)$, $0<r<2n$, $A = 1,...,2n$, where the B^A are arbitrary constants of motion and the C's appearing in the functions g^i_A are specified constants of motion. A procedure is given to determine the g^i_A. For Lagrangian systems it follows that velocity-dependent Noether mappings are a subclass of the above-mentioned general symmetry mappings of the form (a)–(e). An analysis of velocity-dependent Noether mapping theory is included in order to compare for Lagrangian systems the procedure for obtaining the characteristic form (e) of the general mappings with the procedure for obtaining the well-known formula (f) $Z^i_N = H^{ij}(\dot{x},x,t)\partial Z/\partial\dot{x}^j$ ($Z \equiv$ constant of motion), characteristic of velocity-dependent Noether mappings. It is shown how any *given* velocity-dependent symmetry mapping function $Z^i(\dot{x},x,t)$ (including Noether mappings) can be expressed in the form (e). A collection of variational formulas and identities is derived in order to develop from first principles the velocity-dependent symmetry mapping theory. Throughout, comparisons are made between velocity-dependent and velocity-independent symmetry theory.

I. INTRODUCTION

An infinitesimal mapping which maps the solution set of a system of differential equations into itself is said to be a symmetry mapping of the system of equations. In a series of earlier papers[1–14] the authors have investigated various relationships between infinitesimal symmetry mappings of classical (including relativistic) particle dynamical systems[15]

$$E^i(\ddot{x}^1,...,\ddot{x}^n,\dot{x}^1,...,\dot{x}^n,x^1,...,x^n,t) \equiv E^i(\ddot{x},\dot{x},x,t) = 0, \quad i = 1,...,n, \tag{1.1}$$

and their constants of motion. (For further references see Sarlet and Cantrijn.[16]) The infinitesimal mappings considered in those papers were of the form

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i \equiv \xi^i(x,t)\delta a, \tag{1.2}$$

$$\bar{t} = t + \delta t, \quad \delta t \equiv \xi^0(x,t)\delta a. \tag{1.3}$$

It is to be noted that the functions ξ^i (1.2) which determined the point mappings and the functions ξ^0 (1.3) which determined the associated change in path parameter t were assumed to be functions of only x^i and t. We shall now refer to such mappings as *velocity-independent* in order to distinguish them from *velocity-dependent* mappings determined by mapping functions $\xi^i(\dot{x},x,t)$, $\xi^0(\dot{x},x,t)$ to be considered in this paper.

The primary purpose of this paper is to show that general infinitesimal velocity-dependent symmetry mappings of classical particle systems are always expressible in a form with a characteristic functional structure which is the same for all second-order dynamical systems (1.1) and is manifestly dependent upon constants of motion of the system. In an accompanying paper (paper II of this series)[17] the characteristic functional structure of velocity-independent symmetry mappings is obtained for systems of first-order differential equations. In future papers knowledge of the characteristic symmetry structure will be applied to linear dynamical systems and to dynamical systems with cyclic coordinates. Additional relationships between velocity-dependent symmetry mappings of general dynamical systems and constants of motion will also be considered in later papers.

In this present paper we start from first principles and develop (from the local point of view) the variational procedures for determining general velocity-dependent symmetry mappings of dynamical systems (1.1), with emphasis on Lagrangian systems.

For Lagrangian systems velocity-dependent Noether symmetry mappings are a subclass of the *general* velocity-dependent symmetry mappings. Such Noether mappings are known to be expressible in a form with a characteristic functional structure which is the same for all Lagrangian systems and dependent upon constants of motion of the system. It therefore follows that this known Noether structure must be expressible in the form of the above-mentioned general structure we have obtained, and this will be shown. A clever mathematical technique used by Candotti, Palmieri, and Vi-

 0022-2488/85/123080-20$02.50

tale[18,19] in their analysis of the inverse Noether theorem is essential to our derivation of the characteristic functional structure for general symmetry mappings. For this reason, and to make it convenient to compare the methods of deriving the characteristic functional structures associated with these two respective classes of symmetries, we include an extensive presentation of Noether theory as a prerequisite to our analysis of general symmetry mappings.

Throughout this paper, where it is of interest, we shall make comparisons between velocity-dependent and velocity-independent symmetry properties.

In the formulation and analysis of infinitesimal velocity-dependent symmetry mappings of dynamical systems we have found that certain basic variational formulas and identities are frequently used in a number of different derivations. Therefore, before proceeding with any dynamical symmetry analysis *per se* we shall first develop in Sec. II a collection of useful fundamental variational formulas and identities. These will be derived in a systematic manner from a few basic variational definitions. Not only will these formulas expedite many derivations in this paper, but we believe they will be useful for future reference. For completeness many of the formulas will be given in a more general form than actually required in this paper.

In Sec. III we apply the variational formulas developed in Sec. II to formulate several important variational identities which involve the Lagrangian operator. Two of these identities clearly show the distinction in functional structure of the variation of the Lagrangian operator based upon infinitesimal velocity-dependent and velocity-independent mappings. A third identity gives a decomposition of the variation of the Lagrangian operator which will be used to show that the variation in path parameter δt may be arbitrarily chosen in the formulation of velocity-dependent symmetry mappings.

In Sec. IV the identities developed in the previous two sections are employed to formulate conditions in order that infinitesimal velocity-dependent mappings define general symmetry mappings of an n-dimensional Lagrangian dynamical system. These conditions are in the form of a system of n partial differential equations in the $n+1$ unknown mapping functions $\xi^i(\dot{x},x,t)$, $\xi^0(\dot{x},x,t)$. By a decomposition of the mapping functions $\xi^i(=Z^i+\dot{x}^i\xi^0)$ the symmetry equations reduce to a system of n equations in the n auxiliary mapping functions $Z^i(\dot{x},x,t)$. The absence of $\xi^0(\dot{x},x,t)$ from these latter equations indicates that in a velocity-dependent symmetry mapping the function $\xi^0(\dot{x},x,t)$ may be chosen arbitrarily. It then follows for each Z^i solution that a choice for ξ^0 will lead to functions ξ^i which together with ξ^0 determine a velocity-dependent symmetry mapping. The symmetry conditions are shown to be invariant with respect to the "Lagrangian gauge change" $L\to\bar{L}\equiv L+df(x,t)/dt$.

In Sec. V a presentation of velocity-dependent Noether theory is given. We define Noether mappings at the *differential* level (as opposed to a formulation based upon the variation of the action integral) in a manner which is applicable to both velocity-dependent and independent mappings. For velocity-dependent mappings this definition leads to conditions which must hold identically in the $\ddot{x}^i$ and thereby avoids the situation in which every arbitrary mapping is a Noether mapping.

By means of elementary analysis we obtain a general solution of the Noether mapping condition for velocity-dependent mappings. This familiar Noether mapping solution has a characteristic functional structure which is the same for all Lagrangian systems. In this form each velocity-dependent Noether mapping is manifestly functionally dependent upon a constant of motion of the dynamical system. The mapping function ξ^0 is arbitrary. Due to this arbitrariness, for any given constant of motion there is an associated class of velocity-dependent Noether mappings. These Noether mapping solutions are shown to satisfy the general velocity-dependent symmetry mapping condition for Lagrangian dynamical systems (derived in Sec. IV), thereby establishing at the differential level that velocity-dependent Noether mappings are *symmetry* mappings (as to be expected).

For any given velocity-dependent Noether symmetry mapping (described above) the Noether identity (given in Sec. III) shows that the well-known concomitant Noether constant of motion is independent of the arbitrary mapping function ξ^0. Moreover, this concomitant Noether constant of motion is the constant of motion which appears in the given Noether mapping. By an appropriate choice of the arbitrary function ξ^0 this latter property is the inverse Noether theorem as developed by Candotti, Palmieri, and Vitale.[18,19]

Section V also contains an analysis of the invariance of velocity-dependent Noether mapping theory with respect to the Lagrangian gauge change of Sec. IV.

In Sec. VI we develop a method for obtaining the characteristic functional structure of the velocity-dependent mappings which are solutions of the symmetry equations for the class of dynamical systems (1.1). The structure of symmetry mappings for Lagrangian systems is included as a subcase. For the dynamical system (1.1), as in the case of the Lagrangian systems, a decomposition in the mapping function $\xi^i(=Z^i+\dot{x}^i\xi^0)$ reduces the symmetry condition to a system of n second-order partial differential equations in the n unknown auxiliary mapping functions $Z^i(\dot{x},x,t)$. Here again the absence of the mapping function ξ^0 in the reduced system of equations indicates ξ^0 may be arbitrarily chosen. However, except for the simplest dynamical systems this reduced system of partial differential equations remains formidable.

As a prerequisite to deriving the characteristic functional structure of the solution of this reduced system of symmetry equations we examine the reciprocity between any complete finite solution of the dynamical system (which has $2n$ essential integration constants) and its *associated* set of $2n$ functionally independent constants of motion. A consequence of this reciprocity is that the constants of integration of any complete finite solution are also the respective values which its associated constants of motion assume on dynamical paths.

In principle, by means of any assumed finite complete solution of the dynamical equations we show how the system of n *partial* differential symmetry equations in the n $Z^i(\dot{x},x,t)$ may be reduced to a system of n second-order, *linear*, homogeneous *ordinary* differential equations in n unknowns $z^i(t)$.

Any number of the above-mentioned $2n$ integration constants of the finite solution of the dynamical equation may appear as parameters in this affiliated system of ordinary linear differential equations. The general form of the solution of such a system of ordinary linear equations is known to be a linear combination of $2n$ fundamental solution functions. We thus conclude that when evaluated on dynamical trajectories every solution of the symmetry equations must be of this general form.

In the solution of the affiliated system of linear equations we replace those constants of integration of the dynamical equations, which now act as parameters, by their associated (by means of the reciprocity mentioned above) constants of motion. In addition the $2n$ constants of integration which originate in the solution of these affiliated equations are replaced by arbitrary constants of motion of the dynamical equations. As a consequence of these replacements the solution of the affiliated system of ordinary linear equations is converted into a set of n functions of $\dot{x}^i, x^i$, and t. We then prove these n functions will be a solution $Z^i(\dot{x},x,t)$ of the original partial differential symmetry equations. Such Z^i solutions will be manifestly dependent upon the constants of motion of the dynamical system and will have a characteristic functional structure which is the same for all dynamical systems. Moreover, we prove that *every* solution of the system of partial differential symmetry equations is expressible in this characteristic form. Thus velocity-dependent Noether symmetry mappings, which have their own (familiar) characteristic functional structure (as discussed in Sec. V), can also be reexpressed in this general characteristic form.

Finally, we show in detail how any given velocity-dependent symmetry solution $Z^i(\dot{x},x,t)$ may be expressed in a form which has the characteristic functional structure of the above-described general symmetry solution.

In Sec. VII we illustrate the theory developed in the preceding sections by determining the symmetries of a one-dimensional nonlinear dynamical system.

II. BASIC VARIATIONAL FORMULAS AND IDENTITIES

Let x^i be the coordinates of a point P in a local region of an n-dimensional space. A *general* curve γ passing through P is expressed in terms of a path parameter t as $x^i = x^i(t)$. At P the curve γ has a tangent vector dx^i/dt.

We consider infinitesimal point mappings with associated change in parameter defined by

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i \equiv \xi^i\left(\frac{dx}{dt}, x, t\right)\delta a, \tag{2.1}$$

$$\bar{t} = t + \delta t, \quad \delta t \equiv \xi^0\left(\frac{dx}{dt}, x, t\right)\delta a, \tag{2.2}$$

with inverse

$$x^i = \bar{x}^i - \delta\bar{x}^i, \quad \delta\bar{x}^i \equiv \xi^i\left(\frac{d\bar{x}}{d\bar{t}}, \bar{x}, \bar{t}\right)\delta a, \tag{2.3}$$

$$t = \bar{t} - \delta\bar{t}, \quad \delta\bar{t} \equiv \xi^0\left(\frac{d\bar{x}}{d\bar{t}}, \bar{x}, \bar{t}\right)\delta a. \tag{2.4}$$

Based upon the mappings (2.1) and (2.2) we define to within first order in δa

$$\delta\frac{d^\alpha x^i}{dt^\alpha} \equiv \frac{d^\alpha \bar{x}^i}{d\bar{t}^\alpha} - \frac{d^\alpha x^i}{dt^\alpha}, \quad \alpha = 1,2,\ldots . \tag{2.5}$$

We shall subsequently express (2.5) explicitly in terms of the basic mapping functions ξ^i, ξ^0 and their derivatives. At this point we shall first establish a recursion relation between[20] $\delta(d^\alpha x^i/dt^\alpha)$ and $\delta(d^{\alpha-1}x^i/dt^{\alpha-1})$. We therefore express (2.5) in the form

$$\delta\frac{d^\alpha x^i}{dt^\alpha} = \left(\frac{d}{dt}\frac{d^{\alpha-1}\bar{x}^i}{d\bar{t}^{\alpha-1}}\right)\frac{dt}{d\bar{t}} - \frac{d}{dt}\frac{d^{\alpha-1}x^i}{dt^{\alpha-1}}, \quad \alpha = 1,2,\ldots . \tag{2.6}$$

From (2.4) it follows that to first order in δa we have

$$\frac{dt}{d\bar{t}} = 1 - \frac{d\delta t}{dt}. \tag{2.7}$$

Hence (2.6) may be expressed in the form

$$\delta\frac{d^\alpha x^i}{dt^\alpha} = \frac{d}{dt}\left(\frac{d^{\alpha-1}\bar{x}^i}{d\bar{t}^{\alpha-1}} - \frac{d^{\alpha-1}x^i}{dt^{\alpha-1}}\right) - \left(\frac{d}{dt}\frac{d^{\alpha-1}\bar{x}^i}{d\bar{t}^{\alpha-1}}\right)\frac{d\delta t}{dt}, \quad \alpha = 1,2,\ldots , \tag{2.8}$$

which to first order in δa reduces to the desired recursion relation[20]

$$\delta\frac{d^\alpha x^i}{dt^\alpha} = \frac{d}{dt}\left(\delta\frac{d^{\alpha-1}x^i}{dt^{\alpha-1}}\right) - \frac{d^\alpha x^i}{dt^\alpha}\frac{d\delta t}{dt}, \quad \alpha = 1,2,\ldots . \tag{2.9}$$

Next we introduce a decomposition of the infinitesimal mapping (2.1) by expressing δx^i in the form

$$\delta x^i = \Delta x^i + \frac{dx^i}{dt}\delta t, \tag{2.10}$$

$$\Delta x^i \equiv Z^i\left(\frac{dx}{dt}, x, t\right)\delta a. \tag{2.11}$$

It follows from (2.1), (2.10), and (2.11) that

$$\xi^i\left(\frac{dx}{dt}, x, t\right) = Z^i\left(\frac{dx}{dt}, x, t\right) + \frac{dx^i}{dt}\xi^0\left(\frac{dx}{dt}, x, t\right). \tag{2.12}$$

Remark 2.1: We note that if the mapping functions ξ^i and ξ^0 are assumed to be velocity independent [i.e., $\xi^i(x,t)$, $\xi^0(x,t)$] then Z^i must be at most linear in dx^i/dt, since for this case we have from (2.12) that $Z^i = Z^i_0 \equiv \xi^i(x,t) - (dx^i/dt)\,\xi^0(x,t)$.[21] ■

As we shall demonstrate, the decomposition (2.12) in the infinitesimal mapping will expedite many calculations. More importantly, it will be shown in a straightforward fashion by means of (2.12) that when velocity-dependent point mappings of the type (2.1) are employed in the analysis of dynamical symmetries then the δt occurring in (2.2) may be arbitrarily chosen.[22]

Remark 2.2: When $\delta t = 0$ a mapping of the type (2.1) and (2.2) reduces by (2.10) to the form

$$\bar{x}^i = x^i + \Delta x^i, \tag{2.13}$$

$$\bar{t} = t, \quad (\delta t \equiv 0). \tag{2.14}$$

Thus we may consider the Δ variation to be a special case of the δ variation for which $\bar{t} \equiv t$. It follows, therefore, that formulas based upon the δ variation [associated with the map-

ping (2.1) and (2.2)] will reduce to formulas based upon the Δ variation [associated with the mapping (2.13) and (2.14)] by setting $\delta t = 0$ in the former. ■

We find from (2.9) and Remark 2.2 that

$$\Delta \frac{d^\alpha x^i}{dt^\alpha} = \frac{d}{dt}\left(\Delta \frac{d^{\alpha-1} x^i}{dt^{\alpha-1}}\right), \quad \alpha = 1,2,\ldots . \tag{2.15}$$

From (2.15) we immediately find for $\alpha = 1$ that

$$\Delta \frac{dx^i}{dt} = \frac{d\Delta x^i}{dt}. \tag{2.16}$$

It now follows from (2.15) and (2.16) that[20]

$$\Delta \frac{d^\alpha x^i}{dt^\alpha} = \frac{d^\alpha \Delta x^i}{dt^\alpha}, \quad \alpha = 0,1,\ldots . \tag{2.17}$$

We next determine the relationship between $\delta(d^\alpha x^i/dt^\alpha)$ and $\Delta(d^\alpha x^i/dt^\alpha)$. From (2.10) we have

$$\frac{d\delta x^i}{dt} = \frac{d\Delta x^i}{dt} + \frac{d^2 x^i}{dt^2}\delta t + \frac{dx^i}{dt}\frac{d\delta t}{dt}. \tag{2.18}$$

By use of (2.18) we may express (2.9) for the case $\alpha = 1$ in the form

$$\delta \frac{dx^i}{dt} = \frac{d\Delta x^i}{dt} + \frac{d^2 x^i}{dt^2}\delta t. \tag{2.19}$$

From (2.19) we obtain

$$\frac{d}{dt}\left(\delta \frac{dx^i}{dt}\right) = \frac{d^2 \Delta x^i}{dt^2} + \frac{d^3 x^i}{dt^3}\delta t + \frac{d^2 x^i}{dt^2}\frac{d\delta t}{dt}. \tag{2.20}$$

By (2.20) we may express (2.9) for the case $\alpha = 2$ in the form

$$\delta \frac{d^2 x^i}{dt^2} = \frac{d^2 \Delta x^i}{dt^2} + \frac{d^3 x^i}{dt^3}\delta t. \tag{2.21}$$

If we continue in a similar manner we find[23]

$$\delta \frac{d^\alpha x^i}{dt^\alpha} = \frac{d^\alpha \Delta x^i}{dt^\alpha} + \frac{d^{\alpha+1} x^i}{dt^{\alpha+1}}\delta t, \quad \alpha = 0,1,\ldots . \tag{2.22}$$

By means of (2.17) and (2.22) we obtain the desired relationship

$$\delta \frac{d^\alpha x^i}{dt^\alpha} = \Delta \frac{d^\alpha x^i}{dt^\alpha} + \frac{d^{\alpha+1} x^i}{dt^{\alpha+1}}\delta t, \quad \alpha = 0,1,\ldots . \tag{2.23}$$

It is now a straightforward procedure to express $\delta(d^\alpha x^i/dt^\alpha)$, $\alpha = 1,2,\ldots$, as a function of the basic deformations δx^i, δt, and appropriate derivatives by use of (2.10) and (2.22). We thus obtain

$$\delta \frac{d^\alpha x^i}{dt^\alpha} = \frac{d^\alpha \delta x^i}{dt^\alpha} - \frac{d^\alpha}{dt^\alpha}\left(\frac{dx^i}{dt}\delta t\right) + \frac{d^{\alpha+1} x^i}{dt^{\alpha+1}}\delta t, \quad \alpha = 1,2,\ldots . \tag{2.24}$$

Equation (2.24) can be rewritten in the expanded form

$$\delta \frac{d^\alpha x^i}{dt^\alpha} = \frac{d^\alpha \delta x^i}{dt^\alpha} - \sum_{r=0}^{\alpha}\left[\frac{\alpha!}{r!(\alpha-r)!}\frac{d^{\alpha+1-r} x^i}{dt^{\alpha+1-r}}\frac{d^r \delta t}{dt^r}\right] + \frac{d^{\alpha+1} x^i}{dt^{\alpha+1}}\delta t, \quad \alpha = 1,2,\ldots . \tag{2.25}$$

We evaluate (2.25) for the cases $\alpha = 1,2,3$:

$$(\alpha = 1) \quad \delta \frac{dx^i}{dt} = \frac{d\delta x^i}{dt} - \frac{dx^i}{dt}\frac{d\delta t}{dt}, \tag{2.26}$$

$$(\alpha = 2) \quad \delta \frac{d^2 x^i}{dt^2} = \frac{d^2 \delta x^i}{dt^2} - 2\frac{d^2 x^i}{dt^2}\frac{d\delta t}{dt} - \frac{dx^i}{dt}\frac{d^2 \delta t}{dt^2}, \tag{2.27}$$

$$(\alpha = 3) \quad \delta \frac{d^3 x^i}{dt^3} = \frac{d^3 \delta x^i}{dt^3} - 3\frac{d^3 x^i}{dt^3}\frac{d\delta t}{dt} - 3\frac{d^2 x^i}{dt^2}\frac{d^2 \delta t}{dt^2} - \frac{dx^i}{dt}\frac{d^3 \delta t}{dt^3}. \tag{2.28}$$

It is now convenient to introduce the notation

$$x^i_\alpha \equiv \frac{d^\alpha x^i}{dt^\alpha}, \ \alpha = 0,1,\ldots; \quad x^i_0 \equiv x^i. \tag{2.29}$$

The variational formulas (2.1), (2.2), and (2.25) allow us to define the δ variation of a function $G(x^i_N, x^i_{N-1},\ldots, x^i_1 x^i_0, t) \equiv G(x_N, x_{N-1},\ldots, x_1, x_0, t)$ to be[24]

$$\delta G(x_N, x_{N-1}, x_1, x_0, t) \equiv \frac{\partial G}{\partial x^i_\alpha}\delta x^i_\alpha + \frac{\partial G}{\partial t}\delta t, \quad \alpha = 0,\ldots,N. \tag{2.30}$$

In a similar manner the variational formulas (2.13), (2.14), and (2.25) allow us to define the Δ variation of G to be

$$\Delta G(x_N, x_{N-1},\ldots, x_1, x_0, t) \equiv \frac{\partial G}{\partial x^i_\alpha}\Delta x^i_\alpha, \quad \alpha = 0,\ldots,N. \tag{2.31}$$

By means of (2.22) and (2.30) we obtain

$$\delta G = \frac{\partial G}{\partial x^i_\alpha}\frac{d^\alpha \Delta x^i}{dt^\alpha} + \left(\frac{\partial G}{\partial x^i_\alpha}x^i_{\alpha+1} + \frac{\partial G}{\partial t}\right)\delta t, \quad \alpha = 0,\ldots,N. \tag{2.32}$$

With the use of (2.17), (2.31), and the definition

$$\frac{dG}{dt} \equiv \frac{\partial G}{\partial x^i_\alpha}x^i_{\alpha+1} + \frac{\partial G}{\partial t}, \quad \alpha = 0,\ldots,N, \tag{2.33}$$

we express (2.32) in the form

$$\delta G = \Delta G + \frac{dG}{dt}\delta t. \tag{2.34}$$

From (2.30) with $G \equiv UV$ it immediately follows that

$$\delta(UV) = (\delta U)V + U\,\delta V. \tag{2.35}$$

Similarly from (2.31) we find

$$\Delta(UV) = (\Delta U)V + U\,\Delta V. \tag{2.36}$$

It follows from (2.31) and (2.33) by direct expansion [and use of (2.15)] that

$$\frac{d}{dt}\Delta G = \Delta \frac{dG}{dt}. \tag{2.37}$$

We use (2.34) and form

$$\frac{d}{dt}\delta G = \frac{d}{dt}\Delta G + \frac{d^2 G}{dt^2}\delta t + \frac{dG}{dt}\frac{d\delta t}{dt}. \tag{2.38}$$

Again we employ (2.34) with G replaced by dG/dt to obtain

$$\delta \frac{dG}{dt} = \Delta \frac{dG}{dt} + \frac{d^2 G}{dt^2}\delta t. \tag{2.39}$$

It now follows from (2.37)–(2.39) that

$$\frac{d}{dt}\delta G = \delta \frac{dG}{dt} + \frac{dG}{dt}\frac{d\delta t}{dt}. \tag{2.40}$$

By use of (2.17) and (2.31) a straightforward calculation shows that for $G \equiv G(x_N, x_{N-1},..., x_1, x_0, t)$ we have

$$\Delta \frac{\partial G}{\partial x_\alpha^i} - \frac{\partial \Delta G}{\partial x_\alpha^i} = -\frac{\partial G}{\partial x_\beta^j} \frac{\partial}{\partial x_\alpha^i}\left(\frac{d^\beta \Delta x^j}{dt^\beta}\right), \quad 0 \leqslant \alpha \leqslant N, \quad \beta = 0,1,...,N. \tag{2.41}$$

$$\Delta \frac{\partial G}{\partial t} - \frac{\partial \Delta G}{\partial t} = -\frac{\partial G}{\partial x_\beta^j} \frac{\partial}{\partial t}\left(\frac{d^\beta \Delta x^j}{dt^\beta}\right), \quad \beta = 0,1,...,N. \tag{2.42}$$

III. GENERAL AND VARIATIONAL IDENTITIES INVOLVING THE LAGRANGIAN OPERATOR Λ_i

The basic variational formulas developed in the previous section will now be used to derive several important variational identities which involve the Lagrangian operator

$$\Lambda_i \equiv \frac{d}{dt}\frac{\partial}{\partial \dot{x}^i} - \frac{\partial}{\partial x^i}. \tag{3.1}$$

We shall, however, first give some general identities which will be freely used in this section and elsewhere in the paper. Since these general identities are easily verified by direct expansion we omit all derivational details.

By use of the total derivative operator (2.33) we find for a function $G = G(x_N, x_{N-1},..., x_0, t)$ that

$$\frac{\partial}{\partial t}\frac{dG}{dt} - \frac{d}{dt}\frac{\partial G}{\partial t} \equiv 0, \tag{3.2}$$

$$\frac{\partial}{\partial x^i}\frac{dG}{dt} - \frac{d}{dt}\frac{\partial G}{\partial x^i} \equiv 0, \tag{3.3}$$

$$\frac{\partial}{\partial x_\alpha^i}\frac{dG}{dt} - \frac{d}{dt}\frac{\partial G}{\partial x_\alpha^i} \equiv \frac{\partial G}{\partial x_{\alpha-1}^i}, \quad \alpha = 1,2,..., N. \tag{3.4}$$

By use of (2.33) and (3.1)–(3.4) we find for the above G that

$$\Lambda_i\left(\frac{dG}{dt}\right) \equiv \frac{d^2}{dt^2}\frac{\partial G}{\partial \dot{x}^i}, \tag{3.5}$$

$$\Lambda_i\left(\frac{\partial G}{\partial t}\right) = \frac{\partial}{\partial t}\Lambda_i(G), \tag{3.6}$$

$$\Lambda_i\left(\frac{\partial G}{\partial x^j}\right) = \frac{\partial}{\partial x^j}\Lambda_i(G). \tag{3.7}$$

For a function $L(\dot{x}^i, x^i, t)$ we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i - L\right) + \frac{\partial L}{\partial t} \equiv \dot{x}^i \Lambda_i(L). \tag{3.8}$$

Based upon mappings (2.1) and (2.2) with the use of (2.26), (3.1), and (3.8) it is easily verified for $L = L(\dot{x}, x, t)$ that

$$\delta L + L\frac{d\delta t}{dt} \equiv \frac{d}{dt}\left[\frac{\partial L}{\partial \dot{x}^i}\delta x^i - \left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i - L\right)\delta t\right] - (\delta x^i - \dot{x}^i \delta t)\Lambda_i(L). \tag{3.9}$$

If in (3.9) we eliminate δL and δx^i by use of (2.34) and (2.10), respectively, we obtain

$$\Delta L \equiv \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\Delta x^i\right) - \Delta x^i \Lambda_i(L). \tag{3.10}$$

Equations (3.9) and (3.10) are well known. We shall refer to either as the *Noether identity*.

We now derive several identities which involve $\delta \Lambda_i(L)$ and $\Delta \Lambda_i(L)$. We consider these identities to be of fundamental importance in the analysis of symmetries of Lagrangian dynamical systems. The first identity follows immediately from (2.34) when we replace $G(\ddot{x}, \dot{x}, x, t)$ by $\Lambda_i[L(\dot{x}, x, t)]$.

Identity 3.1:

$$\delta \Lambda_i(L) \equiv \Delta \Lambda_i(L) + \delta t \frac{d}{dt}\Lambda_i(L). \tag{3.11}$$

■

By use of (2.31) [with $G \equiv \Lambda_i(L)$], (2.11), and (2.17) we find

$$\Delta \Lambda_i(L) = (H_{ij}\ddot{Z}^j + \tilde{J}_{ij}\dot{Z}^j + \tilde{K}_{ij}Z^j)\delta a, \tag{3.12}$$

where

$$H_{ij}(\dot{x}, x, t) \equiv \frac{\partial}{\partial \dot{x}^j}\Lambda_i(L), \tag{3.13}$$

$$\tilde{J}_{ij}(\ddot{x}, \dot{x}, x, t) \equiv \frac{\partial}{\partial \dot{x}^j}\Lambda_i(L), \tag{3.14}$$

$$\tilde{K}_{ij}(\ddot{x}, \dot{x}, x, t) \equiv \frac{\partial}{\partial x^j}\Lambda_i(L). \tag{3.15}$$

From (3.1) we find that H_{ij} defined by (3.13) is given by

$$H_{ij} = \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} = H_{ji}. \tag{3.16}$$

From (3.1) with the use of (3.4) (with G replaced by $\partial L/\partial \dot{x}^i$) the function $\tilde{J}_{ij}$ defined by (3.14) can be expressed in the form

$$\tilde{J}_{ij} = \dot{H}_{ij} + \Omega_{ij}, \tag{3.17}$$

where

$$\Omega_{ij} \equiv \frac{\partial^2 L}{\partial \dot{x}^i \partial x^j} - \frac{\partial^2 L}{\partial \dot{x}^j \partial x^i} = -\Omega_{ji}. \tag{3.18}$$

We next obtain an alternative form for $\Delta \Lambda_i(L)$ of (3.12). By means of (3.1) and the identity (2.37) (with $G = L$) we find

$$\Delta \Lambda_i(L) = \frac{d}{dt}\Delta\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \Delta\left(\frac{\partial L}{\partial x^i}\right). \tag{3.19}$$

Use of (2.41) (with $G = L$ and the choices $\alpha = 0$ and $\alpha = 1$) allows us to express (3.19) in the form [recall from (2.11) that $\Delta x^j \equiv Z^j \delta a$]

$$\Delta \Lambda_i(L) \equiv \Lambda_i(\Delta L) + \left[-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \dot{Z}^j}{\partial \dot{x}^i} + \frac{\partial L}{\partial x^j}\frac{\partial Z^j}{\partial \dot{x}^i}\right) + \frac{\partial L}{\partial \dot{x}^j}\frac{\partial \dot{Z}^j}{\partial x^i} + \frac{\partial L}{\partial x^j}\frac{\partial Z^j}{\partial x^i}\right]\delta a. \tag{3.20}$$

By use of (3.3) and (3.4) (with G replaced by Z^j) we express (3.20) in the form

$$\Delta \Lambda_i(L) \equiv \Lambda_i(\Delta L) + \left[-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{d}{dt}\frac{\partial Z^j}{\partial \dot{x}^i} + \frac{\partial L}{\partial \dot{x}^j}\frac{\partial Z^j}{\partial x^i} + \frac{\partial L}{\partial x^j}\frac{\partial Z^j}{\partial \dot{x}^i}\right)\right] + \frac{\partial L}{\partial \dot{x}^j}\frac{d}{dt}\frac{\partial Z^j}{\partial x^i} + \frac{\partial L}{\partial x^j}\frac{\partial Z^j}{\partial x^i}\Bigg]\delta a. \tag{3.21}$$

Equation (3.1) allows us to write

$$\frac{\partial L}{\partial \dot{x}^j}\frac{d}{dt}\frac{\partial Z^j}{\partial \dot{x}^i}=\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial Z^j}{\partial \dot{x}^i}\right)-\frac{\partial Z^j}{\partial \dot{x}^i}\Lambda_j(L)-\frac{\partial L}{\partial x^j}\frac{\partial Z^j}{\partial \dot{x}^i}. \tag{3.22}$$

With (3.22) used in (3.21) we obtain the above-mentioned alternative form for $\Delta\Lambda_i(L)$ of (3.12).

Identity 3.2:

$$\Delta\Lambda_i(L)\equiv\Lambda_i(\Delta L)-\frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \Delta x^j}{\partial \dot{x}^i}\right)+\Lambda_i(\Delta x^j)\Lambda_j(L)+\frac{\partial \Delta x^j}{\partial \dot{x}^i}\frac{d}{dt}\Lambda_j(L). \tag{3.23}$$

■

It is of interest to express $\delta\Lambda_i(L)$ of (3.11) explicitly in terms of the basic functions δx^i and δt. To achieve this we first express the right side of (3.23) in this form. By use of (2.34) (with $G=L$) we may express $\Lambda_i(\Delta L)$ in the form

$$\Lambda_i(\Delta L)=\Lambda_i\left[\delta L+L\frac{d}{dt}\delta t-\frac{d}{dt}(L\,\delta t)\right]. \tag{3.24}$$

If we now use (3.5) (with $G=L\,\delta t$) we may write (3.24) in the form

$$\Lambda_i(\Delta L)=\Lambda_i\left(\delta L+L\frac{d}{dt}\delta t\right)-\frac{d^2}{dt^2}\left[\frac{\partial}{\partial \dot{x}^i}(L\delta t)\right]. \tag{3.25}$$

By use of (3.25) and (2.10) we express (3.23) in the form

$$\Delta\Lambda_i(L)=\Lambda_i\left(\delta L+L\frac{d}{dt}\delta t\right)-\frac{d^2}{dt^2}\left[\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \delta x^j}{\partial \dot{x}^i}-\left(\frac{\partial L}{\partial \dot{x}^j}\dot{x}^j-L\right)\frac{\partial \delta t}{\partial \dot{x}^i}\right]+\Lambda_i(\delta x^j-\dot{x}^j\delta t)\Lambda_j(L)+\left(\frac{\partial \delta x^j}{\partial \dot{x}^i}-\delta^j_i\delta t-\dot{x}^j\frac{\partial \delta t}{\partial \dot{x}^i}\right)\frac{d}{dt}\Lambda_i(L). \tag{3.26}$$

Employing (3.26) we express (3.11) in the desired form which shows the explicit dependence upon δx^i and δt and thereby obtain the following identity.

Identity 3.3:

$$\delta\Lambda_i(L)=\Lambda_i\left(\delta L+L\frac{d}{dt}\delta t\right)-\frac{d^2}{dt^2}\left[\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \delta x^j}{\partial \dot{x}^i}-\left(\frac{\partial L}{\partial \dot{x}^j}\dot{x}^j-L\right)\frac{\partial \delta t}{\partial \dot{x}^i}\right]+\Lambda_i(\delta x^j-\dot{x}^j\delta t)\Lambda_j(L)+\left(\frac{\partial \delta x^j}{\partial \dot{x}^i}-\dot{x}^j\frac{\partial \delta t}{\partial \dot{x}^i}\right)\frac{d}{dt}\Lambda_j(L). \tag{3.27}$$

■

Inspection of (3.27) readily shows how the structural form of $\delta\Lambda_i(L)$ depends upon the velocity dependence of the mapping functions δx^i and δt.

IV. SYMMETRIES OF LAGRANGIAN DYNAMICAL SYSTEMS

Identities developed in the previous two sections will now be employed in the formulation and analysis of infinitesimal symmetry mappings of n-dimensional, $n>1$, Lagrangian dynamical systems

$$\Lambda_i(L)=0,\quad L=L(\dot{x},x,t). \tag{4.1}$$

Expansion of (4.1) by means of (3.1) leads to

$$\Lambda_i(L)=H_{ij}\,\ddot{x}^j-R_i=0, \tag{4.2}$$

where $H_{ij}(\dot{x},x,t)$ is given by (3.16), and

$$R_i(\dot{x},x,t)\equiv-\left[\frac{\partial^2 L}{\partial \dot{x}^i\,\partial x^j}\dot{x}^j+\frac{\partial^2 L}{\partial \dot{x}^i\,\partial t}-\frac{\partial L}{\partial x^i}\right]. \tag{4.3}$$

It is assumed that

$$\det H_{ij}\neq 0, \tag{4.4}$$

so that we may define

$$H^{ij}\equiv\frac{\operatorname{cof} H_{ji}}{\det H_{ij}}=H^{ji}. \tag{4.5}$$

From (4.5) it follows that

$$H^{ij}H_{jk}=\delta^i_k. \tag{4.6}$$

By use of (4.6) and (4.2) we obtain ($R^j\equiv H^{ji}R_i$)

$$\ddot{x}^j\doteq R^j(\dot{x},x,t). \tag{4.7}$$

Remark 4.1: The "$\doteq$" notation: Before discussing the concept of symmetry mappings of a dynamical system we first discuss the idea of "dynamical functional composition." By means of the dynamical equation (4.7) and its total derivatives with respect to t we may eliminate from any function [for notation refer to (2.29)] $G(x^i_N, x^i_{N-1},\ldots,x^i_1,x^i_0,t)$, $N\geqslant 2$, all derivatives of the coordinates which are of order higher than dx^i/dt to obtain a function of $dx^i/dt, x^i, t$. We indicate this particular type of functional composition by the notation "$\doteq$." For example, if from the dynamical equation we obtain $\ddot{x}=R(\dot{x},x,t)$ then $G(\ddot{x},\dot{x},x,t)\doteq G[R(\dot{x},x,t),\dot{x},x,t]\equiv F(\dot{x},x,t)$.

The appearance of the "$\doteq$" sign in place of the usual "$=$" sign in an equation indicates the above dynamical functional composition is to be used to express all functions appearing in the equation in terms of the "coordinates" $dx^i/dt, x^i, t$. For example, if the function defined by the operation [refer to (2.11) and (2.31)] $\Delta G(\ddot{x},\dot{x},x,t)$ appears in an equation which employs the "$\doteq$" notation, then all derivatives of the coordinates of order greater than dx^i/dt are to be eliminated from the various partial derivatives of G and also from the expanded expressions $\dot{Z}^i(\dot{x},x,t)$ and $\ddot{Z}^i(\dot{x},x,t)$. It is understood equations so obtained by this procedure of dynamical functional composition are to hold only on dynamical trajectories. We will retain the "$\doteq$" notation in such equations as a reminder that these equations are to hold for dynamical trajectories. ■

With the above-described notational scheme in mind we turn now to the formulation of the condition for a symmetry mapping of Lagrange's equation. An infinitesimal mapping [(2.1), (2.2)] which maps the set of all solution curves of (4.1) into itself is customarily said to be a symmetry mapping of the Lagrangian dynamical system (4.1). Such mappings are determined by the condition[25]

$$\delta\Lambda_i(L)\doteq 0. \tag{4.8}$$

Mappings of the form (2.1) and (2.2) may be classified with regard to the assumed *explicit* dependence of the mapping functions ξ^i, ξ^0 upon the $\dot{x}^i$ variables. There are four

possibilities to be considered. However, in the analysis of the symmetry mapping condition (4.8) we find it convenient to define only two main types as follows.

(I) *Velocity-dependent point mappings*:

$$\text{(a)}\ \xi^i = \xi^i(\dot{x},x,t), \quad \xi^0 = \xi^0(\dot{x},x,t), \tag{4.9}$$

$$\text{(b)}\ \xi^i = \xi^i(\dot{x},x,t), \quad \xi^0 = \xi^0(x,t). \tag{4.10}$$

(II) *Velocity-independent point mappings*:

$$\text{(a)}\ \xi^i = \xi^i(x,t), \quad \xi^0 = \xi^0(\dot{x},x,t), \tag{4.11}$$

$$\text{(b)}\ \xi^i = \xi^i(x,t), \quad \xi^0 = \xi^0(x,t). \tag{4.12}$$

To see how the form of the symmetry condition (4.8) is dependent upon the assumed velocity dependence of ξ^i, ξ^0 we use (3.27) (Identity 3.3) to express (4.8) in the form

$$\delta\Lambda_i(L) \doteq \Lambda_i\left(\delta L + L\frac{d}{dt}\delta t\right) - \frac{d^2}{dt^2}\left[\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \delta x^j}{\partial \dot{x}^i} - \left(\frac{\partial L}{\partial \dot{x}^j}\dot{x}^j - L\right)\frac{\partial \delta t}{\partial \dot{x}^i}\right] \doteq 0. \tag{4.13}$$

Inspection of the symmetry condition in the form (4.13) shows that if ξ^i and ξ^0 are both independent of $\dot{x}^i$ [type (II)(b)] then it reduces to the form

$$\text{(II)(b)}\quad \delta\Lambda_i(L) \doteq \Lambda_i\left(\delta L + \frac{d}{dt}\delta t\right) \doteq 0. \tag{4.14}$$

This form of the symmetry condition for completely velocity-independent mappings was previously obtained by Katzin and Levine.[26] Since in (4.14) neither ξ^i nor ξ^0 contain $\dot{x}^i$, the expansion of (4.14) leads to n equations which must hold identically in $\dot{x}^i$. Consequently, this generally results in an overdetermined system of equations to be solved for $\xi^i(x,t)$, $\xi^0(x,t)$. Hence not every dynamical system will admit symmetry solutions of the type (II)(b). Symmetries of this type have been extensively treated in the literature.

Inspection of (4.13) shows that for type (II)(a) solutions, for which $\xi^i = \xi^i(x,t)$, $\xi^0 = \xi^0(\dot{x},x,t)$, the symmetry condition reduces to

$$\text{(II)(a)}\quad \delta\Lambda_i(L) \doteq \Lambda_i\left(\delta L + L\frac{d\delta t}{dt}\right) + \frac{d^2}{dt^2}\left[\left(\frac{\partial L}{\partial \dot{x}^j}\dot{x}^j - L\right)\frac{\partial \delta t}{\partial \dot{x}^i}\right] \doteq 0. \tag{4.15}$$

We are unaware of any systematic analysis of symmetries of this type.

For both types (I)(a) and (I)(b) of velocity-dependent point mappings, when (4.8) is formally expanded [refer to (2.30)] and use is made of (4.7) to eliminate all time derivatives of x^i of order higher than $\dot{x}^i$, we are led to a system of n homogeneous, second-order, linear partial differential equations to be solved for the $(n+1)$ symmetry mapping functions $[\xi^i(\dot{x},x,t), \xi^0(\dot{x},x,t)]$ or $[\xi^i(\dot{x},x,t), \xi^0(x,t)]$.

This system of differential equations may be simplified by changing n of the $(n+1)$ dependent variables from $\xi^i(\dot{x},x,t)$ to $Z^i(\dot{x},x,t)$ by the transformation

$$\xi^i(\dot{x},x,t) = Z^i(\dot{x},x,t) + \dot{x}^i\xi^0. \tag{2.12'}$$

We are easily led to this simplifying transformation by inspection of (3.11) (Identity 3.1) which indicates the effect of such a variable change on $\delta\Lambda_i(L)$. Since $d\Lambda_i(L)/dt \doteq 0$, we have from (3.11) that the transformation (2.12) expresses the symmetry condition (4.8) in the form

$$\text{(I)(a), (I)(b)}\quad \Delta\Lambda_i(L) \doteq 0. \tag{4.16}$$

From the definition of the Δ operator (2.31) it is apparent that the $Z^i(\dot{x},x,t)$ are the only unknown functions occurring in (4.16). [The detailed expansion of (4.16) is given in Theorem 4.2 below.]

Remark 4.2: It is of significance that the elimination of ξ^i from (4.8) by the transformation (2.12) also resulted in the simultaneous elimination of ξ^0. Hence for the case of velocity-dependent point mappings [type (I)] the solutions $Z^i(\dot{x},x,t)$ of (4.16) can be immediately used in (2.12) to obtain the functions $\xi^i(\dot{x},x,t)$ with the functions $\xi^0(\dot{x},x,t)$ arbitrarily chosen. Thus for a given $Z^i(\dot{x},x,t)$ solution any choice of $\xi^0(\dot{x},x,t)$ will lead to an associated $\xi^i(\dot{x},x,t)$ such that ξ^i, ξ^0 will satisfy (4.8), and hence define a velocity-dependent symmetry mapping [(2.1) and (2.2)]. ■

For the present purpose of comparison of (4.16) and (4.13), and for later convenience in the analysis of Noether mappings we briefly digress and employ (3.23) (Identity 3.2) to express (4.16) in the form

$$\text{(I)(a), (I)(b)}\quad \Delta\Lambda_i(L) \doteq \Lambda_i(\Delta L) - \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \Delta x^j}{\partial \dot{x}^i}\right) \doteq 0. \tag{4.17}$$

We may summarize the above results in the following theorem.

Theorem 4.1: An infinitesimal velocity-dependent mapping

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i \equiv \xi^i(\dot{x},x,t)\delta a, \tag{2.1'}$$

$$\bar{t} = t + \delta t, \quad \delta t \equiv \xi^0(\dot{x},x,t)\delta a, \tag{2.2'}$$

will define a symmetry mapping of a Lagrangian dynamical system

$$\Lambda_i(L) = 0, \quad L = L(\dot{x},x,t), \tag{4.1'}$$

if the mapping functions $\delta x^i, \delta t$ are solutions of the symmetry condition

$$\delta\Lambda_i(L) \doteq 0. \tag{4.8'}$$

The symmetry condition (4.8') may be expressed in the equivalent form

$$\Lambda_i\left(\delta L + L\frac{d\delta t}{dt}\right) - \frac{d^2}{dt^2}\left[\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \delta x^j}{\partial \dot{x}^i} - \left(\frac{\partial L}{\partial \dot{x}^j}\dot{x}^j - L\right)\frac{\partial \delta t}{\partial \dot{x}^i}\right] \doteq 0. \tag{4.13'}$$

The velocity-dependent mapping functions $\delta x^i, \delta t$ which are solutions of (4.8') are given by

$$\delta x^i = \Delta x^i + \dot{x}^i\delta t, \tag{2.10'}$$

$$\Delta x^i \equiv Z^i(\dot{x},x,t)\delta a, \tag{2.11'}$$

$$\delta t = \text{arbitrary}, \tag{4.18}$$

where Z^i is a solution of the system of partial differential equations determined by either of the following equivalent conditions:

$$\Delta\Lambda_i(L) \doteq 0, \tag{4.16'}$$

or

$$\Lambda_i(\Delta L) - \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial \Delta x^j}{\partial \dot{x}^i}\right) \stackrel{\circ}{=} 0. \tag{4.17'}$$

■

To obtain the partial differential equations for $Z^i(\dot{x},x,t)$, referred to in Theorem 4.1, we first make use of (3.12)–(3.15) to express the symmetry condition (4.16) in the form

$$H_{ij}\ddot{Z}^j + J_{ij}\dot{Z}^j + K_{ij}Z^j \stackrel{\circ}{=} 0, \tag{4.19}$$

where in (4.19) the functions $J_{ij}(\dot{x},x,t)$ and $K_{ij}(\dot{x},x,t)$ result from the elimination of any $\ddot{x}^i$ dependence from $\tilde{J}_{ij}(\ddot{x},\dot{x},x,t)$ (3.14) and $\tilde{K}_{ij}(\ddot{x},\dot{x},x,t)$ (3.15) by use of (4.7) in that

$$\tilde{J}_{ij}(\ddot{x},\dot{x},x,t) \stackrel{\circ}{=} \tilde{J}_{ij}[R(\dot{x},x,t),\dot{x},x,t] \equiv J_{ij}(\dot{x},x,t), \tag{4.20}$$

$$\tilde{K}_{ij}(\ddot{x},\dot{x},x,t) \stackrel{\circ}{=} \tilde{K}_{ij}[R(\dot{x},x,t),\dot{x},x,t] \equiv K_{ij}(\dot{x},x,t). \tag{4.21}$$

Formal expansion of (4.19) with use of (4.7) leads to a system of n homogeneous, second-order, linear, partial differential equations to be solved for the n functions $Z^i(\dot{x},x,t)$ only.

We may therefore state the following theorem.

Theorem 4.2: For a Lagrangian dynamical system

$$\Lambda_i(L) \equiv H_{ij}(\dot{x},x,t)\ddot{x}^j - R_i(\dot{x},x,t) = 0, \tag{4.2'}$$

based upon the Lagrangian $L = L(\dot{x},x,t)$, where

$$H_{ij}(\dot{x},x,t) \equiv \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j}, \tag{3.16'}$$

$$\det H_{ij} \neq 0, \tag{4.4}$$

$$R_i(\dot{x},x,t) \equiv -\left[\frac{\partial^2 L}{\partial \dot{x}^i \partial x^j}\dot{x}^j + \frac{\partial^2 L}{\partial \dot{x}^i \partial t} - \frac{\partial L}{\partial x^i}\right], \tag{4.3}$$

the symmetry condition

$$\Delta\Lambda_i(L) \stackrel{\circ}{=} 0 \tag{4.16'}$$

may be formally expressed in the form

$$H_{ij}\ddot{Z}^j + J_{ij}\dot{Z}^j + K_{ij}Z^j \stackrel{\circ}{=} 0, \tag{4.19}$$

where

$$H_{ij}(\dot{x},x,t) \equiv \frac{\partial \Lambda_i(L)}{\partial \ddot{x}^j}, \tag{3.13'}$$

$$J_{ij}(\dot{x},x,t) \stackrel{\circ}{\equiv} \frac{\partial \Lambda_i(L)}{\partial \dot{x}^j} \tag{4.22}$$

$$K_{ij}(\dot{x},x,t) \stackrel{\circ}{\equiv} \frac{\partial \Lambda_i(L)}{\partial x^j}. \tag{4.23}$$

For the dynamical system (4.2') the symmetry condition (4.16') [or (4.19')] is equivalent to the following system of partial differential equations for the symmetry mapping function $Z^i(\dot{x},x,t)$:

$$A_{ij}^{ab}\frac{\partial^2 Z^j}{\partial \dot{x}^a \partial \dot{x}^b} + B_{ij}^{ab}\frac{\partial^2 Z^j}{\partial \dot{x}^a \partial x^b} + C_{ij}^{ab}\frac{\partial^2 Z^j}{\partial x^a \partial x^b} + D_{ij}^{a}\frac{\partial^2 Z^j}{\partial \dot{x}^a \partial t} + E_{ij}^{a}\frac{\partial^2 Z^j}{\partial x^a \partial t} + H_{ij}\frac{\partial^2 Z^j}{\partial t \partial t} + G_{ij}^{a}\frac{\partial Z^j}{\partial \dot{x}^a} + F_{ij}^{a}\frac{\partial Z^j}{\partial x^a} + J_{ij}\frac{\partial Z^j}{\partial t} + K_{ij}Z^j \stackrel{\circ}{=} 0, \tag{4.24}$$

where

$$A_{ij}^{ab}(\dot{x},x,t) \equiv H_{ij}R^a R^b, \tag{4.25}$$

$$B_{ij}^{ab}(\dot{x},x,t) \equiv 2H_{ij}R^a\dot{x}^b, \tag{4.26}$$

$$C_{ij}^{ab}(\dot{x},x,t) \equiv H_{ij}\dot{x}^a\dot{x}^b, \tag{4.27}$$

$$D_{ij}^{a}(\dot{x},x,t) \equiv 2H_{ij}R^a, \tag{4.28}$$

$$E_{ij}^{a}(\dot{x},x,t) \equiv 2H_{ij}\dot{x}^a, \tag{4.29}$$

$$G_{ij}^{a}(\dot{x},x,t) \equiv H_{ij}\left(\frac{\partial R^a}{\partial \dot{x}^b}R^b + \frac{\partial R^a}{\partial x^b}\dot{x}^b + \frac{\partial R^a}{\partial t}\right) + J_{ij}R^a, \tag{4.30}$$

$$F_{ij}^{a}(\dot{x},x,t) \equiv H_{ij}R^a + J_{ij}\dot{x}^a, \tag{4.31}$$

$$J_{ij}(\dot{x},x,t) \equiv \frac{\partial H_{ia}}{\partial \dot{x}^j}R^a - \frac{\partial R_i}{\partial \dot{x}^j}, \tag{4.32}$$

$$K_{ij}(\dot{x},x,t) \equiv \frac{\partial H_{ia}}{\partial x^j}R^a - \frac{\partial R_i}{\partial x^j}. \tag{4.33}$$

The function $R^a(\dot{x},x,t)$ appearing in (4.25)–(4.33) are defined by $R^a = H^{ab}R_b$ with R_b given by (4.3'), and where H^{ij} is defined by $H^{ij}H_{jk} = \delta^i_k$. ■

The fact that the system of partial differential equations (4.24) can be expressed formally in the condensed form (4.19) will be utilized in Sec. VI to develop a procedure for solving these equations for the vectors $Z^i(\dot{x},x,t)$.

Since $Z^i = 0$ is a solution of (4.24) [or (4.19)], it follows from Theorem 4.1 that we may state the following corollary.

Corollary 4.1.1: The velocity-dependent infinitesimal mappings defined by

$$\delta x^i = \dot{x}^i\,\delta t, \tag{4.34}$$

$$\delta t = \text{arbitrary} \tag{4.35}$$

will be symmetry mappings in that they satisfy the symmetry condition (4.8). Such mappings will map points of each trajectory into points of the same trajectory. ■

Let us denote the symmetry equation (4.24) for $Z^i(\dot{x},x,t)$ by

$$S^i[Z(\dot{x},x,t)] \stackrel{\circ}{=} 0. \tag{4.36}$$

By inspection of (4.2) and (4.24)–(4.33), it follows that if

$$\frac{\partial \Lambda_i(L)}{\partial t} = 0, \tag{4.37}$$

then the coefficients appearing in the differential equation (4.24) [or equivalently (4.36)] will be independent of t. For this case we have

$$\frac{\partial}{\partial t}S^i[Z(\dot{x},x,t)] = S^i\left[\frac{\partial Z(\dot{x},x,t)}{\partial t}\right], \tag{4.38}$$

which leads to the following corollary to Theorem 4.2.

Corollary 4.2.1: If $Z^i(\dot{x},x,t)$ is a solution of the symmetry equation (4.24) associated with a Lagrangian dynamical system (4.2) for which

$$\frac{\partial \Lambda_i(L)}{\partial t} = 0, \tag{4.37'}$$

then $\partial Z^i(\dot{x},x,t)/\partial t$ will also be a solution of the symmetry equation. ■

A result similar to Corollary 4.2.1 was previously obtained for the case of velocity-independent symmetry mappings.[3]

Remark 4.3: From (3.5) with G replaced by $f(x,t)$ we obtain the well-known result that

$$\Lambda_i\left[L+\frac{df(x,t)}{dt}\right]=\Lambda_i(L). \tag{4.39}$$

Therefore it is immediately obvious that if we define

$$\overline{L}(\dot{x},x,t)\equiv L(\dot{x},x,t)+\frac{df(x,t)}{dt}, \tag{4.40}$$

then the symmetry condition (4.16) [for $Z^i(\dot{x},x,t)$] will be invariant with respect to the Lagrangian gauge change

$$L\rightarrow\overline{L}=L+df(x,t)/dt. \tag{4.41}$$

Consequently the set of infinitesimal velocity-dependent symmetry mappings (2.1) and (2.2) of a Lagrangian dynamical system as described in Theorem 4.1 will be invariant with respect to this gauge change. A similar result was found for the case of velocity-independent symmetry mappings.[8] ■

Thus far we have found that the problem of obtaining infinitesimal velocity-dependent symmetry mappings determined by the functions $\xi^i(\dot{x},x,t)$, $\xi^0(\dot{x},x,t)$ can be reduced to that of solving the system of partial differential equations (4.24) for the functions $Z^i(\dot{x},x,t)$. For all but the simplest dynamical systems this is a formidable task. However, as we shall prove (in Sec. VI), *every* solution of the above-mentioned system of partial differential symmetry equations can be expressed in a form which has a characteristic functional structure which is independent of the specific dynamical system being considered.

There exists a major subclass of solutions of the symmetry equation (4.24) which is based upon the work of Noether.These known velocity-dependent Noether symmetry solutions $Z^i(\dot{x},x,t)$ (discussed in Sec. V) also have a characteristic functional structure which is the same for all dynamical systems; however, in standard form the Noether functional structure differs from the functional structure of the general solution mentioned in the paragraph above. It is not apparent how the functional structure of the Noether subclass is reconcilable with the functional structure of the general solution—but this will be shown in Sec. VI.

As is known, there is a direct relation between the Noether subclass of symmetries and constants of motion. We shall show there is also a relation between general velocity-dependent symmetries and constants of motion, although more involved than in the Noether case.

To fully appreciate how velocity-dependent Noether symmetries are related to the above-mentioned general velocity-dependent symmetries and to compare the two methods used for obtaining these symmetries we first discuss Noether theory in detail (in the next section) and then give a detailed discussion of general symmetries in Sec. VI.

V. NOETHER MAPPINGS

An infinitesimal mapping (2.1) and (2.2) is said to define a Noether mapping if there exists a function $\tau(\dot{x},x,t)$ such that the condition

$$\delta L+L\frac{d\delta t}{dt}=-\frac{d\tau}{dt}\delta a \tag{5.1}$$

is satisfied. It is to be noted that in this condensed form the Noether mapping condition is the same for both velocity-dependent and velocity-independent cases.[27]

Use of (1.1), (1.2), (2.26), (2.27), and (2.30) allows us to expand (5.1) to the form[28]

$$\begin{aligned}&\left[\frac{\partial L}{\partial\dot{x}^i}\frac{\partial\xi^i}{\partial\dot{x}^j}-\left(\frac{\partial L}{\partial\dot{x}^i}\dot{x}^i-L\right)\frac{\partial\xi^0}{\partial\dot{x}^j}+\frac{\partial\tau}{\partial\dot{x}^j}\right]\ddot{x}^j\\&\quad+\frac{\partial L}{\partial\dot{x}^i}\left[\left(\frac{\partial\xi^i}{\partial x^j}\dot{x}^j+\frac{\partial\xi^i}{\partial t}\right)-\dot{x}^i\left(\frac{\partial\xi^0}{\partial x^j}\dot{x}^j+\frac{\partial\xi^0}{\partial t}\right)\right]\\&\quad+\frac{\partial L}{\partial x^i}\xi^i+\frac{\partial L}{\partial t}\xi^0+L\left(\frac{\partial\xi^0}{\partial x^j}\dot{x}^j+\frac{\partial\xi^0}{\partial t}\right)\\&\quad+\frac{\partial\tau}{\partial x^j}\dot{x}^j+\frac{\partial\tau}{\partial t}=0.\end{aligned} \tag{5.2}$$

The well-known classical case of Noether mappings is the one for which the functions ξ^i, ξ^0, and τ are all velocity independent, i.e., $\xi^i(x,t)$, $\xi^0(x,t)$, $\tau(x,t)$. For this classical case the Noether mapping condition (5.2) must hold identically in the $\dot{x}^i$. In general this leads to a set of overdetermined equations in the unknowns ξ^i, ξ^0, τ, and hence there may exist Lagrangians for which there are no Noether mappings of this classical type.

It is readily shown at the differential level that any classical Noether mapping determined by (5.1) will be a symmetry mapping. To see this we first note from (3.5) with the choice $G=\tau(x,t)$ that

$$\Lambda_i\left[\frac{d\tau(x,t)}{dt}\right]\equiv 0. \tag{5.3}$$

As a consequence of (5.3) it follows for classical Noether mapping solutions of (5.1) that

$$\Lambda_i\left[\delta L+L\frac{d\delta t}{dt}\right]\equiv 0. \tag{5.4}$$

Hence the symmetry condition (4.14) is satisfied by classical Noether mappings.[8]

We now consider solutions of (5.1) which determine type (I) velocity-dependent point mappings [(4.9) or (4.10)], where the associated function τ may or may not contain $\dot{x}^i$. We shall refer to such solutions as velocity-dependent Noether mappings. As is known, such velocity-dependent Noether mappings are also symmetry mappings. In the latter part of this section this will be shown at the differential level without reference to the invariance of the action integral.

By means of (2.34) (with $G=L$) we may express (5.1) in terms of the mapping functions $Z^i(\dot{x},x,t)$ [refer to (2.12)]. We thereby obtain the Noether mapping condition in the alternative form

$$\Delta L=-\frac{d\tau^*}{dt}\delta a, \tag{5.5}$$

where

$$\tau^*(\dot{x},x,t)\equiv L\xi^0+\tau. \tag{5.6}$$

Expansion of (5.5) [refer to (2.31) and (2.11)] gives

$$\begin{aligned}&\left(\frac{\partial L}{\partial\dot{x}^i}\frac{\partial Z^i}{\partial\dot{x}^j}+\frac{\partial\tau^*}{\partial\dot{x}^j}\right)\ddot{x}^j+\frac{\partial L}{\partial\dot{x}^i}\left(\frac{\partial Z^i}{\partial x^j}\dot{x}^j+\frac{\partial Z^i}{\partial t}\right)\\&\quad+\frac{\partial L}{\partial x^i}Z^i+\frac{\partial\tau^*}{\partial x^i}\dot{x}^i+\frac{\partial\tau^*}{\partial t}=0.\end{aligned} \tag{5.7}$$

Note that the elimination of ξ^i from the (velocity-dependent) Noether mapping condition (5.1) by means of (2.12) has again led to an equation in which ξ^0 does not appear and hence may be considered as arbitrary. [See Remark 4.2 following (4.16).[29]]

For the case of velocity-dependent functions $\tau(\dot{x},x,t)$ it is important to stress that in the Noether mapping condition (5.1) and subsequently in (5.5) the " $=$ " sign is used rather than the " $\mathring{=}$ " sign. To understand the reason for this we note first that for dynamical paths (denoted by the " $\mathring{=}$ " sign) the Noether identity (3.10) reduces to

$$\Delta L \mathring{=} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} Z^i\right) \delta a. \tag{5.8}$$

Hence if " $\mathring{=}$ " were used in the Noether mapping condition (5.1), then by use of (5.8) it would follow that (5.5) could be expressed in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} Z^i + \tau^*\right) \mathring{=} 0. \tag{5.9}$$

Solutions to (5.9) are given by

$$Z^i(\dot{x},x,t) = \text{arbitrary}, \tag{5.10}$$

$$\tau^* = M(\dot{x},x,t) - \frac{\partial L}{\partial \dot{x}^i} Z^i, \tag{5.11}$$

where

$$\frac{dM}{dt} \mathring{=} 0, \tag{5.12}$$

i.e., $M(\dot{x},x,t)$ is any constant of motion of the dynamical system (4.1).

If in (2.12) we consider $\xi^i(\dot{x},x,t)$ as arbitrary in addition to the above-mentioned arbitrariness in $\xi^0(\dot{x},x,t)$, then the determined value of $Z^i(\dot{x},x,t)$ will satisfy condition (5.10). It then follows from (5.6) and (5.11) that a value of $\tau(\dot{x},x,t)$ will exist corresponding to each such arbitrarily chosen ξ^i, ξ^0. The above analysis implies that if " $\mathring{=}$ " were used in (5.1), *any* infinitesimal velocity-dependent mapping (2.1) and (2.2) would be a Noether mapping. To avoid this situation the Noether mapping condition (5.1) has been formulated with the " $=$ " sign.[30] Consequently (5.7) must hold identically in the $\ddot{x}^i$.[31] In contrast it is readily seen for the case of velocity-independent Noether mapping $[\xi^i(x,t),\xi^0(x,t),\tau(x,t)]$ that $\dot{x}^i$ is the highest-order derivative of a coordinate which appears in (5.2). Hence for the velocity-independent case it is immaterial whether or not " $=$ " or " $\mathring{=}$ " is used in the Noether mapping condition (5.1).

We now continue with the analysis of velocity-dependent Noether mappings determined by (5.7). We observe that if a set of functions $Z^i(\dot{x},x,t)$, $\tau^*(\dot{x},x,t)$ is a solution to (5.7), then it follows from (2.12), (5.6), and the above-mentioned arbitrariness in $\xi^0(\dot{x},x,t)$ that a solution to the Noether mapping condition (5.1) [or (5.2)] will be given by the functions

$$\xi^i(\dot{x},x,t) = Z^i(\dot{x},x,t) + \dot{x}^i\xi^0(\dot{x},x,t), \tag{5.13}$$

$$\xi^0(\dot{x},x,t) = \text{arbitrary}, \tag{5.14}$$

$$\tau(\dot{x},x,t) = \tau^*(\dot{x},x,t) - L\xi^0(\dot{x},x,t). \tag{5.15}$$

[Note, however, for the case of velocity-independent Noether mappings, it follows from the remarks in the paragraph immediately after (5.2) that the function $\xi^0(x,t)$ will in general not be arbitrary.]

We next determine the form of the functions $Z^i(\dot{x},x,t)$ and $\tau^*(\dot{x},x,t)$ appearing in (5.13) and (5.15) by obtaining the solution to (5.7). As mentioned above (5.7) must hold identically in $\ddot{x}^i$. This leads to the following equations in the unknown functions $Z^i(\dot{x},x,t)$ and $\tau^*(\dot{x},x,t)$:

$$\frac{\partial L}{\partial \dot{x}^i}\frac{\partial Z^i}{\partial \dot{x}^j} + \frac{\partial \tau^*}{\partial \dot{x}^j} = 0, \tag{5.16}$$

$$\frac{\partial L}{\partial \dot{x}^i}\left(\frac{\partial Z^i}{\partial x^j}\dot{x}^j + \frac{\partial Z^i}{\partial t}\right) + \frac{\partial L}{\partial x^i} Z^i + \frac{\partial \tau^*}{\partial x^i}\dot{x}^i + \frac{\partial \tau^*}{\partial t} = 0. \tag{5.17}$$

Solutions to (5.16) and (5.17) can be obtained in terms of the function $L(\dot{x},x,t)$, provided $\det H_{ij} \neq 0$ [see (3.16) for the definition of H_{ij}]. Sarlet and Centrijn[16] obtained solutions to (5.16) and (5.17) in their analysis of symmetries of Lagrangian dynamical systems (based upon the Cartan one-form) which led to constants of motion of the Noether type.[32] Their solutions, however, were determined using the succinct methods of the modern techniques of calculus on manifolds. Palmieri and Vitale[18] in their analysis of the inverse Noether problem stated without detailed proof a special solution to essentially (5.16) and (5.17) for the case where $L = L(\dot{x},x)$ and where the mapping functions are assumed to have no explicit dependence on t. Saletan and Cromer[33] in their textbook treatment of an inverse Noether theorem [for systems with $L(\dot{x},x,t)$] followed the method of Palmieri and Vitale[18] and essentially verified a solution to (5.16) and (5.17). Candotti, Palmieri, and Vitale[19] gave a detailed proof of an inverse Noether theorem for the case $L(\dot{x},x,t)$ with the assumption of velocity and time-dependent mapping functions ξ^i, ξ^0. However, they based their work upon the condition $\tau = 0$ and hence did not obtain a general solution to (5.1) [or (5.16) and (5.17)].

Due to the importance of Noether mappings as a subclass of the more general symmetry mappings to be considered in Sec. VI we shall solve in detail by elementary methods the Noether mapping conditions (5.16) and (5.17). Of particular interest to us in the derivation of this Noether solution is a technique employed by Candotti, Palmieri, and Vitale,[19] since we shall apply a similar technique in an analysis of velocity-dependent general symmetry mappings. This detailed derivation of the solution of the Noether mapping conditions will also provide a unified notation so that the reader may readily compare the method used to obtain the Noether subclass of velocity-dependent symmetry mappings with the method used to obtain the velocity-dependent general symmetry mappings (treated in the next section).

From (5.16) the integrability conditions on τ^* take the form

$$H_{ij}\frac{\partial Z^i}{\partial x^k} - H_{ik}\frac{\partial Z^i}{\partial \dot{x}^j} = 0. \tag{5.18}$$

We define

$$Z_i \equiv H_{ij} Z^j, \tag{5.19}$$

and by (4.5) and (4.6) find

$$Z^i = H^{ij} Z_j. \tag{5.20}$$

By means of (5.20) we may express (5.18) in the form [use being made of a formula for $\partial H^{ic}/\partial \dot{x}^k$ obtained from (4.6), and the property $\partial H_{jb}/\partial \dot{x}^k \equiv \partial H_{kb}/\partial \dot{x}^j$, which follows from the definition (3.16)]

$$\frac{\partial Z_j}{\partial \dot{x}^k} - \frac{\partial Z_k}{\partial \dot{x}^j} = 0. \tag{5.21}$$

The solution to (5.21) is

$$Z_j = \frac{\partial Z}{\partial \dot{x}^j}, \quad Z(\dot{x},x,t) \text{ arbitrary.} \tag{5.22}$$

From (5.22) and (5.20) we find that the integrability condition for τ^* (5.18), obtained from (5.16), requires Z^i to have the form

$$Z^i = H^{ij}\frac{\partial Z}{\partial \dot{x}^j}, \quad Z(\dot{x},x,t) \text{ arbitrary.} \tag{5.23}$$

We now proceed with the integration of (5.16) by first rewriting it in the form [with the use of (3.16)]

$$\frac{\partial \tau^*}{\partial \dot{x}^j} = -\frac{\partial}{\partial \dot{x}^j}\left(\frac{\partial L}{\partial \dot{x}^i}Z^i\right) + H_{ij}Z^i. \tag{5.24}$$

If Z^i in (5.24) is replaced by use of (5.23), the resulting equation may be readily integrated to give

$$\tau^* = -H^{im}\frac{\partial L}{\partial \dot{x}^i}\frac{\partial Z}{\partial \dot{x}^m} + Z. \tag{5.25}$$

Since $Z(\dot{x},x,t)$ is arbitrary there has been no loss of generality in dropping an arbitrary additive integration function of x^i and t from (5.25). The integration of (5.16) is now complete; it has led to Z^i with the form (5.23) and τ^* with the form (5.25).

Consider next the integration of the remaining condition (5.17). Use of Z^i (5.23), τ^* (5.25), along with (4.3) allows us to express (5.17) in the form

$$H^{im}R_i\frac{\partial Z}{\partial \dot{x}^m} + \frac{\partial Z}{\partial x^m}\dot{x}^m + \frac{\partial Z}{\partial t} = 0. \tag{5.26}$$

From the decomposition of $\Lambda_i(L)$ given in (4.2), along with the use of (4.6), we may eliminate R_i from (5.26) to obtain

$$\frac{dZ}{dt} - H^{im}\frac{\partial Z}{\partial \dot{x}^m}\Lambda_i(L) = 0. \tag{5.27}$$

If we evaluate (5.27) for a dynamical path we find

$$\frac{dZ}{dt} \doteq 0. \tag{5.28}$$

Hence the unknown function $Z(\dot{x},x,t)$ appearing in (5.26) must necessarily be a constant of motion of the dynamical system (4.1).

It is next shown that a sufficient condition for a function $Z(\dot{x},x,t)$ to be a solution of (5.26) is that it be a constant of motion of the dynamical system (4.1). Assume then that a function $Z(\dot{x},x,t)$ is a constant of motion of the dynamical system (4.1) so that (5.28) holds. We thus have upon expansion of (5.28)

$$\frac{\partial Z}{\partial \dot{x}^m}\ddot{x}^m + \frac{\partial Z}{\partial x^m}\dot{x}^m + \frac{\partial Z}{\partial t} \doteq 0. \tag{5.29}$$

Use of the dynamical equation in the form (4.7) allows (5.29) to be written in the form

$$F(\dot{x},x,t) \equiv H^{im}R_i\frac{\partial Z}{\partial \dot{x}^m} + \frac{\partial Z}{\partial x^m}\dot{x}^m + \frac{\partial Z}{\partial t} \doteq 0. \tag{5.30}$$

Since the function $F(\dot{x},x,t)$ vanishes at every point of every dynamical path it must vanish identically in the space with variables $(\dot{x}^i,x^i,t)$.[34] Thus if $Z(\dot{x},x,t)$ is a constant of motion of the dynamical system (4.1) it will identically satisfy (5.26); this establishes the above-mentioned sufficiency condition.

Equations (5.16) and (5.17) have now been solved for $\tau^*(\dot{x},x,t)$ and $Z^i(\dot{x},x,t)$ given by (5.23) and (5.25), wherein $Z(\dot{x},x,t)$ must be a constant of motion of the dynamical system (4.1). It therefore follows that the solution to the alternative form of the velocity-dependent Noether mapping condition (5.5) has been obtained. These functions for τ^* and Z^i when used in (5.13) and (5.15) lead to (ξ^i,ξ^0,τ), the general solution of the velocity-dependent Noether mapping condition (5.1):

$$\xi^i(\dot{x},x,t) = H^{ij}\frac{\partial Z}{\partial \dot{x}^j}\dot{x}^i\xi^0, \tag{5.31}$$

$$\xi^0(\dot{x},x,t) = \text{arbitrary}, \tag{5.32}$$

$$\tau(\dot{x},x,t) = -H^{im}\frac{\partial L}{\partial \dot{x}^i}\frac{\partial Z}{\partial \dot{x}^m} + Z - L\xi^0, \tag{5.33}$$

where $Z(\dot{x},x,t)$ is a constant of motion of the dynamical system (4.1) with Lagrangian $L(\dot{x},x,t)$.

Remark 5.1: The Noether mapping functions given by Palmieri and Vitale[18] can be obtained as a special case of (5.31)–(5.33) with the assumption that $L = L(\dot{x},x)$, $Z = Z(\dot{x},x)$, and $\xi^0 = 0$.

The Noether mapping functions given by Candotti, Palmieri, and Vitale[19] can be obtained from (5.31)–(5.33) with the assumption that $\tau = 0$. In this case ξ^0 is no longer arbitrary and is determined by (5.33).

The Noether mapping functions obtained by Sarlet and Cantrijn[16] in their search for generalized symmetry mappings with a concomitant constant of motion of the Noether type were essentially the same as (5.31)–(5.33). ■

We now establish at the *differential level* that the above-obtained velocity-dependent Noether mapping functions (5.31)–(5.33) determine symmetry mappings of the dynamical equation (4.1). To show that an infinitesimal velocity-dependent mapping (2.1) and (2.2) defined by mapping functions ξ^i, ξ^0 is a symmetry mapping it is sufficient to show that the associated mapping function Z^i [refer to (2.12)] satisfies (4.17). Since the Z^i and τ^* associated with the Noether mapping functions ξ^i, τ [refer to (5.13), (5.15), (5.31), and (5.33)] satisfy (5.5) it follows for Noether mappings that (4.17) immediately takes the form

$$\Delta\Lambda_i(L) \doteq -\left[\Lambda_i\left(\frac{d\tau^*}{dt}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \dot{x}^j}\frac{\partial Z^j}{\partial \dot{x}^i}\right)\right]\delta a. \tag{5.34}$$

Use of the identity (3.5) and the definition (3.16) allows us to express (5.34) in the form

$$\Delta\Lambda_i(L) \doteq -\frac{d^2}{dt^2}\left[\frac{\partial}{\partial \dot{x}^i}\left(\tau^* + \frac{\partial L}{\partial \dot{x}^j}Z^j\right) - H_{ij}Z^j\right]\delta a. \tag{5.35}$$

The term in square brackets in (5.35) vanishes identically upon substitution of Z^i and τ^* given by (5.23) and (5.25).

Hence every infinitesimal velocity-dependent Noether mapping (2.1) and (2.2) defined by (5.31)–(5.33) is a symmetry mapping of the dynamical system (4.1). We note that this method for proving that Noether mappings are symmetries has bypassed any reference to the invariance of the action integral.

It is next shown that the above Noether solution (5.31)–(5.33) leads in a natural manner to the inverse Noether theorem.

As is well known, from the Noether identity (3.9) and the Noether mapping condition (5.1) it follows that corresponding to every velocity-dependent or independent Noether mapping (ξ^i,ξ^0,τ) there exists a Noether constant of motion I_N of the dynamical system (4.1), where

$$I_N \equiv \frac{\partial L}{\partial \dot{x}^i}\xi^i - \left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i - L\right)\xi^0 + \tau. \tag{5.36}$$

For the case of velocity-dependent Noether mappings it is found by use of (5.13)–(5.15) that I_N reduces to the form

$$I_N = \frac{\partial L}{\partial \dot{x}^i} Z^i + \tau^*. \tag{5.37}$$

This latter form for I_N shows it is independent of the arbitrary function ξ^0. Moreover, if the Noether mapping functions (5.31)–(5.33) (which incorporate the detailed form of Z^i and τ^*) are used in (5.36), we find

$$I_N = Z, \tag{5.38}$$

i.e., the Noether constant of motion I_N is exactly the constant of motion Z which appears in the Noether mapping functions. This relation (5.38) between I_N and the arbitrarily chosen constant of motion Z, which determines the Noether symmetry mapping functions Z^i,τ^*, essentially forms the basis of what is generally referred to as the inverse Noether theorem, that is, corresponding to every constant of motion Z there will exist velocity-dependent Noether symmetry mapping functions (Z^i,τ^*) whose concomitant Noether constant of motion is Z. However, as a consequence of the arbitrariness of the mapping function ξ^0, an infinite number of velocity-dependent Noether symmetry mappings (ξ^i,ξ^0,τ) [(5.31)–(5.33)] may be associated with each given (Noether) constant of motion (5.38). Thus even a Noether constant of motion originally obtained as a concomitant of a velocity-independent Noether mapping could also be considered as a concomitant of an infinite number of velocity-dependent Noether mappings.

We summarize the velocity-dependent Noether mapping theory in the theorem that follows.

Theorem 5.1: (Noether theory): An infinitesimal velocity-dependent mapping

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i = \xi^i(\dot{x},x,t)\delta a, \tag{2.1'}$$

$$\bar{t} = t + \delta t, \quad \delta t = \xi^0(\dot{x},x,t)\delta a \tag{2.2'}$$

is said to be a velocity-dependent Noether mapping of a Lagrangian dynamical system

$$\Lambda_i(L) = 0, \quad L = L(\dot{x},x,t) \tag{4.1'}$$

$[H_{ij} \equiv \partial^2 L/\partial \dot{x}^i \partial \dot{x}^j, \det H_{ij} \neq 0]$, if and only if there exist functions $\xi^i(\dot{x},x,t)$, $\xi^0(\dot{x},x,t)$, and $\tau(\dot{x},x,t)$ which satisfy the Noether mapping condition

$$\delta L + L\frac{d\delta t}{dt} = -\frac{d\tau}{dt}\delta a. \tag{5.1'}$$

The change in variables

$$\xi^i(\dot{x},x,t) = Z^i(\dot{x},x,t) + \dot{x}^i\xi^0(\dot{x},x,t), \tag{2.12'}$$

$$\tau(\dot{x},x,t) = \tau^*(\dot{x},x,t) - L(\dot{x},x,t)\xi^0(\dot{x},x,t) \tag{5.6'}$$

reduces the Noether mapping condition (5.1') to the form

$$\Delta L = -\frac{d\tau^*}{dt}\delta a, \tag{5.5'}$$

in which ξ^0 does not appear. The solution to (5.5') is expressible in the form

$$Z^i = H^{ij}\frac{\partial Z}{\partial \dot{x}^j}, \tag{5.39}$$

$$\tau^* = -H^{ij}\frac{\partial L}{\partial \dot{x}^i}\frac{\partial Z}{\partial \dot{x}^j} + Z, \tag{5.40}$$

where $H^{ij}H_{jk} = \delta^i_k$, and $Z(\dot{x},x,t)$ is an arbitrary constant of motion of the dynamical system (4.1'). Hence the solution to the Noether mapping condition (5.1') is expressible in the form

$$\xi^i(\dot{x},x,t) = H^{ij}\frac{\partial Z}{\partial \dot{x}^j} + \dot{x}^i\xi^0, \tag{5.31'}$$

$$\xi^0(\dot{x},x,t) = \text{arbitrary}, \tag{5.32'}$$

$$\tau(\dot{x},x,t) = H^{ij}\frac{\partial L}{\partial \dot{x}^i}\frac{\partial Z}{\partial \dot{x}^j} + Z - L\xi^0, \tag{5.33'}$$

and the most general infinitesimal velocity-dependent Noether mapping (2.1') and (2.2') is determined by the Noether mapping functions ξ^i (5.31') and ξ^0 (5.32'). All velocity-dependent Noether mappings are symmetry mappings in that the Noether mapping functions ξ^i (5.31') and ξ^0 (5.32') will identically satisfy the symmetry condition

$$\delta\Lambda_i(L) \doteq 0. \tag{4.8}$$

Alternatively, the Noether mapping function Z^i (5.39) will identically satisfy the symmetry condition

$$\Delta\Lambda_i(L) \doteq 0. \tag{4.16'}$$

Associated with each infinitesimal velocity-dependent Noether symmetry mapping there will exist a concomitant Noether constant of motion

$$I_N = \frac{\partial L}{\partial \dot{x}^i}\xi^i - \left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i - L\right)\xi^0 + \tau, \tag{5.36'}$$

which can be expressed in the equivalent forms

$$I_N = \frac{\partial L}{\partial \dot{x}^i}Z^i + \tau^*, \tag{5.37'}$$

or

$$I_N = Z. \tag{5.38'}$$

■

Remark 5.2: From the form of the velocity-dependent Noether symmetry solution Z^i (5.39) we may regard the problem of obtaining solutions of the Noether symmetry condition (5.5) as being reducible to that of obtaining constants of motion of the associated Lagrangian dynamical system (4.1). ■

Remark 5.3: Infinitesimal velocity-dependent Noether symmetry mapping solutions Z^i and τ^* may be obtained as

explicit functions of $\dot{x}^i, x^i$, and t so that their functional structures in terms of the constant of motion Z [refer to (5.39) and (5.40)] are not immediately apparent. However, one may readily express such known $Z^i(\dot{x},x,t)$ and $\tau^*(\dot{x},x,t)$ in the respective forms (5.39) and (5.40) by first employing the Z^i and τ^* in (5.37) to obtain I_N, which by (5.38) is the required constant of motion $Z(\dot{x},x,t)$ needed to express Z^i and τ^* in the desired forms. This is one of the points illustrated by Example I given in Sec. VII. ■

Remark 5.4: Noether mappings (refer to Theorem 5.1) are a subclass of general symmetry mappings (refer to Theorem 4.1). It therefore follows (refer to Remark 4.3) from the invariance of the symmetry condition (4.16) with respect to a Lagrangian gauge transformation (4.40) that the set of solutions $\{Z^i(\dot{x},x,t)\}$ of the Noether symmetry condition (5.5) must be Lagrangian gauge invariant. However, it remains of interest to examine the effect of Lagrangian gauge transformations on the Noether symmetry condition (5.5) in order to determine how τ^* is affected and then to use the results to investigate the behavior of the Noether constant of motion I_N (5.37) with respect to this gauge transformation.

For completeness it is of interest to not assume the above-mentioned invariance of the Noether mapping functions Z^i but to investigate how such invariance arises from the direct analysis of the effect of a Lagrangian gauge change on the Noether mapping condition (5.5). We therefore now assume that the Noether mapping condition (5.5) is based upon the Lagrangian $\overline{L}(\dot{x},x,t)$ (4.40). We thereby obtain

$$\overline{\Delta}\,\overline{L}(\dot{x},x,t) = -\frac{d\overline{\tau}^*(\dot{x},x,t)}{dt}\,\delta a \tag{5.41}$$

($\overline{\Delta}$ indicates variation with respect to $\overline{Z}^i$), which is to be solved for $\overline{Z}^i(\dot{x},x,t)$ and $\overline{\tau}^*(\dot{x},x,t)$. By means of (4.40) and (5.5) it follows by the use of (2.37) [with $G = f(x,t)$] that (5.41) may be expressed in the form

$$\overline{\Delta}L = -\frac{d\tau^\dagger}{dt}\,\delta a, \tag{5.42}$$

where

$$\tau^\dagger\,\delta\alpha \equiv \overline{\Delta}f(x,t) + \overline{\tau}^*\,\delta a. \tag{5.43}$$

Equation (5.42) is identical in form to (5.5); it therefore follows from Theorem 5.1 that the solution to (5.42) is given by

$$\overline{Z}^i = H^{ij}\frac{\partial\overline{Z}}{\partial\dot{x}^i}, \tag{5.44}$$

$$\tau^\dagger = -H^{ij}\frac{\partial L}{\partial\dot{x}^i}\frac{\partial\overline{Z}^i}{\partial\dot{x}^j} + \overline{Z}, \tag{5.45}$$

where $\overline{Z}(\dot{x},x,t)$ is an arbitrary constant of motion. The set of arbitrary constants of motion $\{\overline{Z}\}$ appearing in (5.44) and (5.45) is the same as the set of arbitrary constants of motion $\{Z\}$ which appears in (5.39) and (5.40). Hence there is no loss in generality in assuming $\overline{Z} = Z$.

It therefore follows that

$$\overline{Z}^i = Z^i, \tag{5.46}$$

which verifies the previous observation that Noether symmetry solutions are invariant with respect to the Lagrangian gauge transformation (4.40).

By use of (5.43)–(5.46) we find

$$\overline{\tau}^* = \tau^* - \frac{\partial f}{\partial x^i}Z^i, \tag{5.47}$$

which shows how the function τ^* is affected by the Lagrangian gauge change (4.40).

We next examine the effect of the Lagrangian gauge change (4.40) on the functional structure of a Noether constant of motion I_N (5.37). To do this we formulate the Noether constant of motion (5.37) in terms of $\overline{L}, \overline{Z}^i$, and $\overline{\tau}^*$ to obtain

$$\overline{I}_N = \frac{\partial\overline{L}}{\partial\dot{x}^i}\overline{Z}^i + \overline{\tau}^*. \tag{5.48}$$

Upon use of (4.40), (5.46), (5,47), and (5.37) we find that (5.48) takes the form

$$\overline{I}_N = I_N. \tag{5.49}$$

Thus the Lagrangian gauge transformation (4.40) leaves the functional structure of a Noether constant of motion invariant because the induced gauge change in τ^* compensates for the gauge change in L (with Z^i remaining invariant).

Since in the Noether solution (5.31)–(5.33) the function $\xi^0(\dot{x},x,t)$ is arbitrary, there is no loss in generality in considering it unchanged with respect to the Lagrangian gauge change (4.40). We then find under this gauge change that

$$\overline{\xi}^i = \xi^i, \tag{5.50}$$

$$\overline{\tau} = \tau - \frac{\partial f}{\partial x^i}\xi^i - \frac{\partial f}{\partial t}\xi^0. \tag{5.51}$$

For a discussion of velocity-independent Noether theory with respect to Lagrangian gauge transformations refer to Katzin and Levine.[8] ■

Remark 5.5: For Lagrangian systems for which the Lagrangian has no explicit t dependence so that $\partial L(\dot{x},x)/\partial t = 0$, it follows from (5.5) [by use of (3.2)] that if $Z^i(\dot{x},x,t)$ and $\tau^*(\dot{x},x,t)$ satisfy (5.5) then so will $\partial Z^i/\partial t$ and $\partial\tau^*/\partial t$. Hence if Z^i defines a Noether mapping of such a dynamical system, then $\partial Z^i/\partial t$ will also be a Noether mapping of the system, that is, $\partial Z^i/\partial t$ remains within the subclass of Noether mappings (refer to Corollary 4.1). A similar result was proved for velocity-independent Noether mapping.[9] ■

Remark 5.6: For Lagrangian systems for which $\partial L/\partial t = 0$ it follows from (5.39) and Remark 5.5 that for such dynamical systems if $M(\dot{x},x,t)$ is a constant of motion so too will be $\partial M(\dot{x},x,t)/\partial t$. This well-known result is generally proved by other means.[35] ■

Remark 5.7: Based upon the Noether mapping condition (5.1) we have discussed Noether mappings for the classical velocity-independent case $[\xi^i(x,t),\xi^0(x,t),\tau(x,t)]$ and for the most general velocity dependence case $[\xi^i(\dot{x},x,t), \xi^0(\dot{x},x,t),\tau(\dot{x},x,t)]$. It would be of further interest to systematically investigate Noether mappings with regard to the assumed explicit velocity dependence or independence of each of the functions ξ^i, ξ^0, τ. To do so one must consider the two choices $\tau(\dot{x},x,t)$ and $\tau(x,t)$ for each of the four types of ξ^i, ξ^0 considered in Sec. IV [refer to (4.9)–(4.12)]. By analysis of the $(n+1)$ equations obtained from (5.2) (which must hold identically in the $\ddot{x}^i$) it can be shown for unconstrained Lagrangian systems that of the eight possible situations, the two cases $[\xi^i(x,t),\xi^0(\dot{x},x,t),\tau(x,t)]$ and $[\xi^i(x,t),\xi^0(x,t),\tau(\dot{x},x,t)]$

cannot exist. We have not attempted a literature search with regard to this detailed classification system. ■

For a Lagrangian dynamical system (4.1) the Noether symmetry condition (5.5) defines a subclass of symmetry solutions of the general symmetry condition (4.16). These velocity-dependent Noether symmetry solutions Z^i were shown to be expressible in a form that has a characteristic functional structure which is the same for all (Lagrangian) dynamical systems and dependent upon the constants of motion of the system. We are thus led to the problem of determining if the general velocity-dependent symmetry solutions Z^i of (4.16), that is, solutions which are either non-Noether or Noether symmetries, can also be expressible in a form which has a characteristic functional structure which is the same for all dynamical systems and is dependent upon the constants of motion of the dynamical system. This problem will be considered in the next section.

VI. A CHARACTERISTIC FUNCTIONAL STRUCTURE OF ALL VELOCITY-DEPENDENT SYMMETRY MAPPINGS

In this section it is shown that every velocity-dependent solution $Z^i(\dot{x},x,t)$ (Noether or non-Noether type) of the symmetry condition (4.16) is expressible in a characteristic form which in principle is derivable if a set of $2n$ functionally independent constants of motion of the dynamical system is known. For generality in the derivation to follow we shall consider dynamical equations which are more general than Lagrange's equation (4.1).

Consider then a system of dynamical equations of the form

$$E_i(\ddot{x},\dot{x},x,t) \equiv H_{ij}(\dot{x},x,t)\ddot{x}^j - F_i(\dot{x},x,t) = 0, \tag{6.1}$$

in which $\det H_{ij} \neq 0$ so that the $\ddot{x}^j$ terms can be expressed in the form

$$\ddot{x}^i \stackrel{\circ}{=} F^i(\dot{x},x,t), \tag{6.2}$$

where

$$F^i \equiv H^{ij}F_j, \quad H^{ij}H_{jk} = \delta^i_k. \tag{6.3}$$

Dynamical equations (6.1) include Lagrange equations of the form (4.1) for the choice $E_i(\ddot{x},\dot{x},x,t) \equiv \Lambda_i(L)$ [refer to (4.2) and (4.3)], in which case $F_i(\dot{x},x,t)$ of (6.1) reduces to $R_i(\dot{x},x,t)$ of (4.3), and H_{ij} of (6.1) reduces to (3.13) [which may then be expressed in the form (3.16)].

Let $C^A(\dot{x},x,t)$, $A = 1,...,2n$, be a set of $2n$ functionally independent constants of motion of the dynamical system (6.1) so that on a dynamical path we may write

$$C^A(\dot{x},x,t) \stackrel{t}{=} c^A, \quad c^A \equiv \text{const}, \tag{6.4}$$

where in general the constants c^A will vary from path to path.

Remark 6.1: The "$\stackrel{t}{=}$" notation: the notation "$\stackrel{t}{=}$" indicates that any function [refer to (2.9)] $G(x^i_N, x^i_{N-1},..., x^i_1, x^i, t)$ is to be expressed completely in terms of the path parameter t by means of the finite equations of the dynamical paths. ■

The $2n$ equations (6.4) may be inverted to express x^i and $\dot{x}^i$ in the form

$$x^i = \phi^i(c^1,c^2,...,c^{2n};t) \equiv \phi^i(c,t), \tag{6.5}$$

$$\dot{x}^i = \psi^i(c^1,c^2,...,c^{2n};t) \equiv \psi^i(c,t) = \frac{\partial \phi^i(c,t)}{\partial t}. \tag{6.6}$$

The set of equations (6.5) is a complete solution to the dynamical equations (6.1) [or (6.2)]. Conversely, for a complete solution (6.5) we could solve Eqs. (6.5) and (6.6) for the constants c^A and thereby obtain the $2n$ constants of motion $C^A(\dot{x},x,t)$ (6.4). This reciprocity between a complete solution of the dynamical equations and its associated set of $2n$ functionally independent constants of motion is essential to the proof to follow.

As with Lagrange's equations (4.1), an infinitesimal velocity-dependent mapping (2.1) and (2.2) which maps the set of all solution curves of (6.1) into itself will define a symmetry mapping of this more general dynamical system. Such mappings are determined by the condition

$$\delta E_i \stackrel{\circ}{=} 0. \tag{6.7}$$

Use of (2.34) with G replaced by E_i shows (in a similar manner to that of Sec. IV) that such symmetry mappings are determined by velocity-dependent mapping functions $\delta x^i, \delta t$ of the form

$$\delta x^i = \Delta x^i + \dot{x}^i\,\delta t, \tag{6.8}$$

$$\delta t = \text{arbitrary}. \tag{6.9}$$

In (6.8), $\Delta x^i [\equiv Z^i(\dot{x},x,t)\delta a]$ is determined by the set of partial differential equations obtained by the formal expansion of the symmetry conditions

$$\frac{\Delta E_i}{\delta a} \stackrel{\circ}{=} H_{ij}\ddot{Z}^j + J_{ij}\dot{Z}^j + K_{ij}Z^j \stackrel{\circ}{=} 0, \tag{6.10}$$

$$H_{ij}(\dot{x},x,t) \stackrel{\circ}{=} \frac{\partial E_i}{\partial \ddot{x}^j}, \quad J_{ij}(\dot{x},x,t) \stackrel{\circ}{=} \frac{\partial E_i}{\partial \dot{x}^j},$$

$$K_{ij}(\dot{x},x,t) \stackrel{\circ}{=} \frac{\partial E_i}{\partial x^j}. \tag{6.11}$$

The expanded form of (6.10) is given by (4.24)–(4.33), wherein for applicability to the more general dynamical system (6.1) currently being considered, we now employ H_{ij}, J_{ij}, K_{ij} as defined by (6.11), and R^i is now replaced by F^i as defined in (6.2).

In the definitions (6.11) and in the formal expansion of $\ddot{Z}^j$ and $\dot{Z}^j$ occurring in (6.10) all $\ddot{x}^i$ and $\dddot{x}^i$ terms are to be expressed as functions of $\dot{x}^i, x^i, t$ by means of the dynamical equation (6.2).

It should be noted that if $E_i \equiv \Lambda_i(L)$, [see (3.1)] then J_{ij}, K_{ij} given by (6.11) will reduce to the respective J_{ij}, K_{ij} based upon the Lagrange equaton (4.2) [refer to (3.14), (3.15), (4.20), and (4.21)], and the functions H_{ij} of (6.11) will reduce to (3.16).

It is convenient to transvect (6.10) with H^{ik} and express the system of equations in the form

$$\ddot{Z}^i + J^i_j\dot{Z}^j + K^i_j Z_j \stackrel{\circ}{=} 0, \tag{6.12}$$

where

$$J^i_j(\dot{x},x,t) \equiv H^{im}J_{mj}, \quad K^i_j(\dot{x},x,t) \equiv H^{im}K_{mj}. \tag{6.13}$$

We first derive a necessary condition that every solution $Z^i(\dot{x},x,t)$ of the set of partial differential symmetry equations

obtained by the expansion of (6.12) must satisfy. If (6.12) is evaluated on the dynamical trajectories of (6.1) by means of (6.5) and (6.6) we obtain an associated set of ordinary differential equations

$$\ddot{z}^i(c,t) + j^i{}_m(c,t)\dot{z}^m(c,t) + k^i{}_m(c,t)z^m(c,t) \stackrel{t}{=} 0, \tag{6.14}$$

where

$$Z^i(\dot{x},x,t) \stackrel{t}{=} Z^i[\psi(c,t),\phi(c,t),t] \equiv z^i(c,t), \tag{6.15}$$

$$J^i{}_m(\dot{x},x,t) \stackrel{t}{=} J^i{}_m[\psi(c,t),\phi(c,t),t] \equiv j^i{}_m(c,t), \tag{6.16}$$

$$K^i{}_m(\dot{x},x,t) \stackrel{t}{=} K^i{}_m[\psi(c,t),\phi(c,t),t] \equiv k^i{}_m(c,t). \tag{6.17}$$

We note that the associated system of equations (6.14) which must be satisfied by $z^i(c,t)$ (6.15) is always a system of *linear* equations irrespective of whether the original dynamical system (6.1) is linear or nonlinear. Solutions of (6.14) are expressible in the form

$$z^i(c,t) = b^A g^i_A(c,t), \tag{6.18}$$

where the b^A, $A = 1,...,2n$ are arbitrary constants.

We may thus state the following theorem.

Theorem 6.1. When evaluated on trajectories

$$x^i = \phi^i(c,t) \tag{6.5'}$$

of the dynamical system

$$H_{ij}(\dot{x},x,t)\ddot{x}^j - F_i(\dot{x},x,t) = 0, \tag{6.1'}$$

every solution $Z^i(\dot{x},x,t)$ of the symmetry (partial differential) equations obtained by the formal expansion of

$$\ddot{Z}^i + J^i_j\,\dot{Z}^j + K^i_j\,Z^j \stackrel{\cdot}{=} 0 \tag{6.12'}$$

must be expressible in the form

$$Z^i(\dot{x},x,t) \stackrel{t}{=} z^i(c,t) = b^A g^i_A(c,t),$$

$$b^A \equiv \text{arbitrary const}, \quad A = 1,...,2n, \tag{6.19}$$

where the $z^i(c,t)$ are solutions of the associated ordinary differential equations

$$\ddot{z}^i(c,t) + j^i{}_m(c,t)\dot{z}^m(c,t) + k^i{}_m(c,t)z^m(c,t) \stackrel{t}{=} 0 \tag{6.14'}$$

obtained by evaluating the symmetry equations (6.12') on the trajectories (6.5'). The $2n$ vectors $g^i_A(c,t)$ appearing in (6.19) constitute a fundamental solution set of (6.14'). ■

Since on a trajectory every solution $Z^i(\dot{x},x,t)$ of (6.12) has the form (6.19) we are led to assume (to be proved below) that the functions $Z^{*i}(\dot{x},x,t)$ defined by

$$Z^{*i}(\dot{x},x,t) \equiv B^{*A}(\dot{x},x,t)g^i_A[C^1(\dot{x},x,t),...,C^{2n}(\dot{x},x,t),t]$$
$$\equiv B^{*A}g^i_A(C,t) \tag{6.20}$$

will be a solution of the (partial differential) symmetry equation (6.12). In (6.20) the $B^{*A}(\dot{x},x,t)$ are arbitrarily chosen constants of motion. Hence, they can always be expressed as functions of the $2n$ functionally independent constants of motion $C^A(\dot{x},x,t)$ (6.4) which are associated with the inversion of the dynamical solution (6.5). We may thus write

$$B^{*A}(\dot{x},x,t) = b^{*A}[C(\dot{x},x,t)] \stackrel{t}{=} b^{*A}(c),$$

$$b^{*A}(c) = \text{const}, \quad A = 1,...,2n. \tag{6.21}$$

In (6.20) the functions $g^i_A[C(\dot{x},x,t),t]$ are obtained by replacing the constants c^A appearing in the $g^i_A(c,t)$ of (6.18) by the respective above-mentioned constants of motion C^A by means of (6.4).

It will now be verified that the assumed solution $Z^{*i}(\dot{x},x,t)$ (6.20) satisfies the (partial differential) symmetry equations (6.12). To do so it is notationally convenient to first define

$$Q^i(Z) \equiv \ddot{Z}^i + J^i_j\,\dot{Z}^j + K^i_j\,Z^j, \tag{6.22}$$

so that it is required to prove $Q^i(Z^*) \stackrel{\cdot}{=} 0$.

We first form from (6.20)

$$\dot{Z}^{*i} = \dot{B}^{*A}g^i_A + B^{*A}\left(\frac{\partial g^i_A}{\partial C^B}\dot{C}^B + \frac{\partial g^i_A}{\partial t}\right). \tag{6.23}$$

Since the $B^{*A}(\dot{x},x,t)$ and $C^A(\dot{x},x,t)$ are constants of motion we have by use of (6.2) that $\dot{B}^{*A} \stackrel{\cdot}{=} 0$, $\dot{C}^A \stackrel{\cdot}{=} 0$, and we thereby obtain from (6.23) that

$$\dot{Z}^{*i} \stackrel{\cdot}{=} B^{*A}\frac{\partial g^i_A}{\partial t}. \tag{6.24}$$

In a similar manner we find

$$\ddot{Z}^i \stackrel{\cdot}{=} B^{*A}\frac{\partial^2 g^i_A}{\partial t\,\partial t}. \tag{6.25}$$

Use of (6.20), (6.24), (6.25), and (6.22) leads to

$$Q^i(Z^*) \stackrel{\cdot}{=} B^{*A}\left[\frac{\partial^2 g^i_A(C,t)}{\partial t\,\partial t} + J^i{}_m \times \frac{\partial g^m_A(C,t)}{\partial t} + K^i{}_m\,g^m_A(C,t)\right]. \tag{6.26}$$

It is to be noted as a consequence of the " $\stackrel{\cdot}{=}$ " in (6.26) that $Q^i(Z^*)$ is now a function of $\dot{x}^i$, x^i, and t, so that we may write

$$Q^i(Z^*) \stackrel{\cdot}{=} P^i(\dot{x},x,t), \tag{6.27}$$

where

$$P^i(\dot{x},x,t) \equiv B^{*A}\left[\frac{\partial^2 g^i_A(C,t)}{\partial t\,\partial t} + J^i{}_m\frac{\partial g^m_A(C,t)}{\partial t} + K^i{}_m\,g^m_A(C,t)\right]. \tag{6.28}$$

We note that in the derivatives $\partial g^i_A(C,t)/\partial t$ and $\partial^2 g^i_A(C,t)/\partial t\,\partial t$ which appear in (6.28) the functions $C^A(\dot{x},x,t)$ are held fixed. Therefore if we evaluate the expression $P^i(\dot{x},x,t)$ (6.28) on the dynamical paths by means of (6.5), (6.6), (6.15)–(6.17), and (6.21) we may write the resulting expression in the form

$$P^i(\dot{x},x,t) \stackrel{t}{=} b^{*A}(c)[\ddot{g}^i_A(c,t) + j^i{}_m(c,t)\dot{g}^m_A(c,t) + k^i{}_m(c,t)g^m_A(c,t)]. \tag{6.29}$$

The functions $g^i_A(c,t)$ in (6.29) are solutions of (6.14) for *all* values of c^A. It follows that $P^i(\dot{x},x,t) \stackrel{t}{=} 0$. Since the functions $P^i(\dot{x},x,t)$ vanish at each point of every trajectory they vanish

identically in the variables $\dot{x}^i$, x^i, and t.[36] From (6.27) we thereby conclude that $Q^i(Z^*) \overset{\circ}{=} 0$. Hence the symmetry condition (6.12) is satisfied by $Z^{*i}(\dot{x},x,t)$ (6.20).

It is next proved that every solution $Z^i(\dot{x},x,t)$ of the symmetry equation (6.12) is expressible in the form (6.20). To do this we assume on the contrary the existence of a solution $Z^{**i}(\dot{x},x,t)$ which is not expressible in the form (6.20). By Theorem 6.1, there will exist constants $b^A = b^{**A}(c)$ so that on dynamical trajectories [refer to (6.19)]

$$Z^{**i}(\dot{x},x,t) \overset{t}{=} z^{**i}(c,t) \overset{t}{=} b^{**A}(c)\, g^i_A(c,t). \tag{6.30}$$

By means of (6.20) and the assumed solution Z^{**i} we form $W^i(\dot{x},x,t) \equiv Z^{**i}(\dot{x},x,t) - Z^{*i}(\dot{x},x,t)$ to obtain

$$W^i(\dot{x},x,t) = Z^{**i}(\dot{x},x,t) - B^{*A}(\dot{x},x,t)\, g^i_A[C(\dot{x},x,t),t]. \tag{6.31}$$

Evaluating $W^i(\dot{x},x,t)$ on dynamical trajectories by use of (6.4)–(6.6), (6.21), and (6.30), we obtain

$$W^i(\dot{x},x,t) \overset{t}{=} b^{**A}(c)\, g^i_A(c,t) - b^{*A}(c)\, g^i_A(c,t). \tag{6.32}$$

Since in (6.31) the B^{*A} [associated with the Z^* solution (6.20)] may be arbitrarily chosen constants of motion, we may always pick them so that $B^{*A}(\dot{x},x,t) \equiv b^{*A}[C(\dot{x},x,t)]$. For this choice it follows [refer to (6.21)] that $b^{**A}[C(\dot{x},x,t] \overset{t}{=} b^{**A}(c)$ and hence by (6.32) $W^i(\dot{x},x,t) \overset{t}{=} 0$ (for all c). Since $W^i(\dot{x},x,t)$ vanishes at every point of every trajectory it vanishes identically in the variables $\dot{x}^i$, x^i, and t. Therefore, by (6.31)

$$Z^{**i}(\dot{x},x,t) = B^{*A}(\dot{x},x,t)\, g^i[C(\dot{x},x,t),t], \tag{6.33}$$

which contradicts our assumption that Z^{**i} was not expressible in the form (6.20). We are thus able to state the following theorem.

Theorem 6.2: Consider a system of dynamical equations

$$E_i \equiv H_{ij}\, \ddot{x}^j - F_i(\dot{x},x,t) = 0, \quad \det H_{ij} \neq 0, \tag{6.1''}$$

with a complete solution for the trajectories given by

$$x^i = \phi^i(c^1,\ldots,c^{2n},t) \equiv \phi^i(c,t),$$
$$c^A \equiv \text{const}, \quad A = 1,\ldots,2n, \tag{6.5''}$$

so that

$$\dot{x}^i = \frac{\partial \phi^i(c,t)}{\partial t}. \tag{6.6'}$$

Let

$$C^A(\dot{x},x,t) \overset{t}{=} c^A \tag{6.4'}$$

be the specific set of $2n$ functionally independent constants of motion obtained by inversion of (6.5'') and (6.6') for the constants c^A.

A solution $Z^i(\dot{x},x,t)$ of the (partial differential) symmetry equations

$$\frac{\Delta E_i}{\delta a} \overset{\circ}{=} H_{ij}\, \ddot{Z}^j + J_{ij}\, \dot{Z}^j + K_{ij}\, Z^j \overset{\circ}{=} 0, \tag{6.10}$$

where

$$H_{ij}(\dot{x},x,t) \overset{\circ}{=} \frac{\partial E_i}{\partial \ddot{x}^j}, \quad J_{ij}(\dot{x},x,t) \overset{\circ}{=} \frac{\partial E_i}{\partial \dot{x}^j},$$
$$K_{ij}(\dot{x},x,t) \overset{\circ}{=} \frac{\partial E_i}{\partial x^j}, \tag{6.11}$$

is

$$Z^i(\dot{x},x,t) = B^A(\dot{x},x,t)\, g^i_A[C^1(\dot{x},x,t),\ldots,C^{2n}(\dot{x},x,t),t]$$
$$\equiv B^A g^i_A(C,t); \tag{6.20'}$$

the $B^A(\dot{x},x,t)$ are arbitrary constants of motion of (6.1'') and hence may be regarded as arbitrary functions of the $2n$ constants of motion $C^A(\dot{x},x,t)$ (6.4'), and the functions $g^i_A[C(\dot{x},x,t),t]$ are obtained by replacing the constants c^A appearing in the functions $g^i_A(c,t)$ of Theorem 6.1 by the respective constants of motion $C^A(\dot{x},x,t)$ by means of (6.4'). Moreover, every solution $Z^i(\dot{x},x,t)$ of the symmetry equation (6.10) is expressible in the form (6.20') by a suitable choice of the constants of motion $B^A(\dot{x},x,t)$. ■

Remark 6.2: With $Z(\dot{x},x,t)$ given by (6.20'), symmetry mappings for a dynamical system (6.1) will be determined by (6.8) and (6.9). ■

Remark 6.3: Since in (6.20) we may choose the arbitrary constants of motion B^A to be absolute constants, it follows that the $2n$ vectors

$$Z^i_A(\dot{x},x,t) \equiv g^i_A[C(\dot{x},x,t),t], \quad A = 1,\ldots,2n \tag{6.34}$$

will be solutions of the symmetry condition (6.10), and hence may be regarded as a "basis" for all symmetry solutions $Z^i(\dot{x},x,t)$. ■

Remark 6.4: Had we initially chosen a different set of $2n$ functionally independent constants of motion (6.4), e.g.,

$$\overline{C}^A(\dot{x},x,t) \overset{t}{=} \overline{c}^A, \tag{6.35}$$

then the associated solution (6.5) (obtained by inversion) of the dynamical equations would correspondingly change to

$$x^i = \overline{\phi}^i(\overline{c}^A,t). \tag{6.36}$$

This change would permeate through the whole procedure and in general results in a different set of fundamental solution functions $\overline{g}^i_A(\overline{c},t)$ (refer to Theorem 6.1). Consequently, we would obtain a different set of solution functions $\overline{g}^i_A[\overline{C}(\dot{x},x,t),t]$ which would also constitute a basis for $Z^i(\dot{x},x,t)$ solutions of the symmetry equation (6.12). It would then follow that

$$Z^i(\dot{x},x,t) = B^A[C(\dot{x},x,t)]\, g^i_A[C(\dot{x},x,t),t]$$
$$= \overline{B}^A[\overline{C}(\dot{x},x,t)]\, \overline{g}^i_A[\overline{C}(\dot{x},x,t),t]. \tag{6.37}$$

■

Since Noether symmetry solutions are a subclass of the solutions $Z^i(\dot{x},x,t)$ of the symmetry condition (4.16) for a Lagrangian dynamical system (4.2), and since such Lagrangian dynamical systems are a subclass of dynamical systems (6.1), we may state the following corollary to Theorem 6.2.

Corollary 6.2.1: Every velocity-dependent Noether symmetry solution $Z^i(\dot{x},x,t)$ (5.39) (as described in Theorem 5.1) of a Lagrangian dynamical system (as described in Theorem 4.1) is expressible in the form (6.20) [as described in Theorem (6.2)]. ■

With reference to Theorem 6.2 let $\overline{Z}^i(\dot{x},x,t)$ be any given solution to the symmetry equation (6.10) [or equivalently (6.12)]. We now show how to choose the arbitrary constants of motion $B^A(\dot{x},x,t)$ so that $\overline{Z}^i$ is expressible in the characteristic form (6.20), that is,

$$\overline{Z}^i(\dot{x},x,t) = B^A(\dot{x},x,t)\, g^i_A[C(\dot{x},x,t),t]. \tag{6.38}$$

From (6.38) we form (recall $\dot{B}^A \doteq 0, \dot{C}^A \doteq 0$)

$$\dot{\overline{Z}}^i \doteq B^A \dot{g}^i_A(C,t) \doteq B^A \frac{\partial g^i_A(C,t)}{\partial t}. \tag{6.39}$$

The functions $g^i_A(c,t)$ ($c^A = \text{const}$) define a complete set of solutions to the ordinary differential equation (6.14); it thus follows that[37]

$$W(c,t) \equiv \begin{vmatrix} g^i_A(c,t) \\ \hline \dot{g}^i_A(c,t) \end{vmatrix} \neq 0, \tag{6.40}$$

where $A = 1,\ldots,2n$ denotes columns and $i = 1,\ldots,n$ denotes rows in the partitioned determinant. Since the constants of motion $C^A(\dot{x},x,t)$ appearing in the functions $g^i_A[C(\dot{x},x,t),t]$ act as constants in (6.39), it follows from (6.40) that

$$W(C,t) \doteq \begin{vmatrix} g^i_A(C,t) \\ \hline \dot{g}^i_A(C,t) \end{vmatrix} \neq 0. \tag{6.41}$$

Therefore the $2n$ linear equations (6.38) and (6.39) may be solved algebraically for the $2n$ constants of motion B^A by Cramer's rule.

We are thus led to the following corollary to Theorem 6.2.

Corollary 6.2.2: For a given solution $\overline{Z}^i(\dot{x},x,t)$ of the symmetry condition (6.10) to be expressed in the generic form (6.20) with respect to the basis functions $g^i_A[C(\dot{x},x,t),t]$ the arbitrary constants of motion $B^A(\dot{x},x,t)$ must be chosen to have the form

$$\overline{B}^A(\dot{x},x,t) \doteq \frac{\Delta_A}{W}, \tag{6.42}$$

where

$$\Delta_A \doteq \begin{vmatrix} g^i_1 \cdots g^i_{A-1} & \overline{Z}^i & g^i_{A+1} \cdots g^i_{2n} \\ \hline \dot{g}^i_1 \cdots \dot{g}^i_{A-1} & \dot{\overline{Z}}^i & \dot{g}^i_{A+1} \cdots \dot{g}^i_{2n} \end{vmatrix} \tag{6.43}$$

and

$$W \doteq \begin{vmatrix} g^i_A \\ \hline \dot{g}^i_A \end{vmatrix}. \tag{6.41'}$$

In the partitioned determinants Δ_A and W, the indices $i = 1,\ldots,n$ denote rows and the indices $A = 1,\ldots,2n$ denote columns. In (6.41') and (6.43) the constants of motion $C^A(\dot{x},x,t)$ appearing in the functions $g^i_A(C,t)$ behave as constants in that $\dot{g}^i_A(C,t) \doteq \partial g^i_A(C,t)/\partial t$. ■

It is of interest to demonstrate that the $\overline{B}^A(\dot{x},x,t)$ (6.42) actually satisfy the conditions $\dot{\overline{B}}^A \doteq 0$, required of constants of motion. With reference to (6.42) we form

$$\dot{\overline{B}}^A \doteq \dot{\Delta}_A/W - \Delta_A \dot{W}/W^2. \tag{6.44}$$

We first form $\dot{W}$ by differentiation of the $2n \times 2n$ determinant W (6.41). Since differentiation of row i ($i = 1,2,\ldots,n$) results in a row whose elements are identical to those of row $i+n$, we obtain a sum of only n determinants which arise from differentiation of the last n rows of W. Hence

$$\dot{W} \doteq \sum_{i=1}^{n} \dot{W}_i, \tag{6.45}$$

where

$$\dot{W}_i \doteq \begin{vmatrix} g^1_1 & \cdots & g^1_{2n} \\ \vdots & & \vdots \\ g^n_1 & \cdots & g^n_{2n} \\ \hline \dot{g}^1_1 & \cdots & \dot{g}^1_{2n} \\ \vdots & & \vdots \\ \dot{g}^{i-1}_1 & \cdots & \dot{g}^{i-1}_{2n} \\ \ddot{g}^i_1 & \cdots & \ddot{g}^i_{2n} \\ \dot{g}^{i+1}_1 & \cdots & \dot{g}^{i+1}_{2n} \\ \vdots & & \vdots \\ \dot{g}^n_1 & \cdots & \dot{g}^n_{2n} \end{vmatrix}. \tag{6.46}$$

Expand $\dot{W}_i$ on row $i+n$ to obtain

$$\dot{W}_i \doteq \sum_{A=1}^{2n} \ddot{g}^i_A \operatorname{cof} \ddot{g}^i_A. \tag{6.47}$$

Remark 6.5: Note that the appearance of a free index "i" on the left side of an equation implies the index i is also free on the right side of the equation. ■

Comparison of the $\dot{W}_i$ determinants (6.46) with the W determinant (6.41) shows that

$$(\operatorname{cof} \ddot{g}^i_A \text{ of } \dot{W}_i) \equiv (\operatorname{cof} \dot{g}^i_A \text{ of } W). \tag{6.48}$$

Hence (6.47) may be written as

$$\dot{W}_i \doteq \sum_{A=1}^{2n} \ddot{g}^i_A (\operatorname{cof} \dot{g}^i_A \text{ of } W). \tag{6.49}$$

With reference to Remark 6.3 and (6.34) therein we have from (6.12) that

$$\ddot{g}^i_A \doteq -(J^i{}_j\, \dot{g}^j_A + K^i{}_j\, g^j_A). \tag{6.50}$$

Upon substitution of (6.50) in (6.49) the resulting equation may be expressed in the form

$$\begin{aligned} \dot{W}_i \doteq & -\sum_{\substack{j=1 \\ j\neq i}}^{n} J^i{}_j \sum_{A=1}^{2n} [\dot{g}^j_A (\operatorname{cof} \dot{g}^i_A \text{ of } W)] \\ & - J^i{}_i \sum_{A=1}^{2n} [\dot{g}^i_A (\operatorname{cof} \dot{g}^i_A \text{ of } W)] \\ & - \sum_{j=1}^{n} K^i{}_j \sum_{A=1}^{2n} [g^j_A (\operatorname{cof} \dot{g}^i_A \text{ of } W)]. \end{aligned} \tag{6.51}$$

From the theory of determinants[38] we have for $\Delta \equiv |a_{ij}|$ and $A_{ij} \equiv \operatorname{cof} a_{ij}$ that

$$\sum_{k=1}^{n} a_{ik} A_{jk} = \delta^i_j \Delta. \tag{6.52}$$

With reference to (6.41) it follows from (6.52) that (6.51) reduces to

$$\dot{W}_i \doteq -J^i{}_i W, \tag{6.53}$$

and hence (6.45) becomes[39]

$$\dot{W} \doteq -\left(\sum_{i=1}^{n} J^i{}_i\right) W. \tag{6.54}$$

The functions Δ_A, which also appear in (6.44), will now be calculated. By the same argument preceding (6.45) it is found that

$$\dot{\Delta}_A \doteq \sum_{i=1}^{n} \dot{\Delta}_{Ai}, \tag{6.55}$$

where

$$\dot{\Delta}_{Ai} \doteq \begin{vmatrix} g_1^1 & \cdots & g_{A-1}^1 & \overline{Z}^1 & g_{A+1}^1 & \cdots & g_{2n}^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ g_1^n & \cdots & g_{A-1}^n & \overline{Z}^n & g_{A+1}^n & \cdots & g_{2n}^n \\ \hline \dot{g}_1^1 & \cdots & \dot{g}_{A-1}^1 & \dot{\overline{Z}}^1 & \dot{g}_{A+1}^1 & \cdots & \dot{g}_{2n}^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \dot{g}_1^{i-1} & \cdots & \dot{g}_{A-1}^{i-1} & \dot{\overline{Z}}^{i-1} & \dot{g}_{A+1}^{i-1} & \cdots & \dot{g}_{2n}^{i-1} \\ \ddot{g}_1^i & \cdots & \ddot{g}_{A-1}^i & \ddot{\overline{Z}}^i & \ddot{g}_{A+1}^i & \cdots & \ddot{g}_{2n}^i \\ \dot{g}_1^{i+1} & \cdots & \dot{g}_{A-1}^{i+1} & \dot{\overline{Z}}^{i+1} & \dot{g}_{A+1}^{i+1} & \cdots & \dot{g}_{2n}^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \dot{g}_1^n & \cdots & \dot{g}_{A-1}^n & \dot{\overline{Z}}^n & \dot{g}_{A+1}^n & \cdots & \dot{g}_{2n}^n \end{vmatrix}. \tag{6.56}$$

$$\dot{\Delta}_{Ai} \doteq \sum_{\substack{B=1 \\ B \neq A}}^{2n} \ddot{g}_B^i \operatorname{cof} \ddot{g}_B^i + \overline{Z}^i \operatorname{cof} \overline{Z}^i. \tag{6.57}$$

By means of (6.12), (6.50), and relations similar to (6.48), Eq. (6.57) can be expressed in the form

$$\begin{aligned} \dot{\Delta}_{Ai} \doteq & -\sum_{j=1}^{n} J^i{}_j \Bigg[\sum_{\substack{B=1 \\ B \neq A}}^{2n} \dot{g}_B^j (\operatorname{cof} \dot{g}_B^j \text{ of } \Delta_B) \\ & + \overline{Z}^j (\operatorname{cof} \overline{Z}^i \text{ of } \Delta_B) \Bigg] \\ & - \sum_{j=1}^{n} K^i{}_j \Bigg[\sum_{\substack{B=1 \\ B \neq A}}^{2n} g_B^j (\operatorname{cof} \dot{g}_B^j \text{ of } \Delta_B) \\ & + \overline{Z}^j (\operatorname{cof} \overline{Z}^i \text{ of } \Delta_B) \Bigg]. \end{aligned} \tag{6.58}$$

With reference to (6.43) and (6.52), Eq. (6.58) reduces to

$$\dot{\Delta}_{Ai} \doteq -J^i{}_i \Delta_A, \tag{6.59}$$

and hence (6.54) takes the form

$$\dot{\Delta}_A \doteq -\left(\sum_{i=1}^{n} J^i{}_i \right) \Delta_A. \tag{6.60}$$

By means of (6.44), (6.54), and (6.60) we find $\overline{B}^A \doteq 0$, which verifies that $\overline{B}^A$ (6.42) is a constant of motion.

As a consequence of the form of the symmetry solution (6.20) we obtain a third corollary to Theorem 6.2 whose proof is immediate.

Corollary 6.2.3: If $Z^i(\dot{x},x,t)$ is a solution of the symmetry equation (6.12), then so will be

$$\tilde{Z}^i(\dot{x},x,t) \equiv A(\dot{x},x,t) Z^i(\dot{x},x,t), \tag{6.61}$$

where $A(\dot{x},x,t)$ is an arbitrary constant of motion of the dynamical system (6.1). ■

Remark 6.6: Katzin and Levine[7] had previously obtained a result similar to Corollary 6.2.3 for the $2n$-dimensional symmetry vector of Hamilton's equations expressed in sympletic notation. More recently, Sarlet and Cantrijn[16] gave another proof for this using the concepts of calculus on manifolds. ■

VII. EXAMPLES

To illustrate various aspects of the theory developed in the preceding sections we now give two examples.

Example I: Consider the one-dimensional dynamical system defined by the Lagrangian

$$L = e^{-\dot{x}} - x. \tag{7.1}$$

It follows from (3.1), (6.1), and (7.1) that

$$E_1 = \Lambda_1(L) = e^{-\dot{x}} \ddot{x} + 1 = 0, \tag{7.2}$$

and hence we have [refer to (6.2)]

$$\ddot{x} \doteq -e^{\dot{x}}. \tag{7.3}$$

The dynamical equation (7.3) may be solved to obtain [refer to (6.5)]

$$x = (t + c^1)[1 - \ln(t + c^1)] + c^2, \tag{7.4}$$

from which it follows that

$$\dot{x} = -\ln(t + c^1). \tag{7.5}$$

By solving (7.4) and (7.5) for the constants c^1 and c^2 we obtain two functionally independent constants of motion C^1 and C^2, where

$$C^1(\dot{x},x,t) = e^{-\dot{x}} - t \doteq c^1, \tag{7.6}$$

$$C^2(\dot{x},x,t) = -e^{-\dot{x}}(\dot{x} + 1) + x \doteq c^2. \tag{7.7}$$

By use of (7.2) and (7.3) we find from (3.13)–(3.15), (4.20), and (4.21) [or alternatively from (6.11)] that

$$H_{11} \doteq e^{-\dot{x}}, \quad J_{11} \doteq 1, \quad K_{11} \doteq 0, \tag{7.8}$$

from which it follows that

$$H^{11} = H_{11}^{-1} = e^{\dot{x}}. \tag{7.9}$$

It now follows that the symmetry condition (4.19) [or (6.10)] takes the form

$$e^{-\dot{x}} \ddot{Z}^1 + \dot{Z}^1 \doteq 0. \tag{7.10}$$

Formal expansion of (7.10) [including elimination of all dot derivatives of x or order higher than $\dot{x}$ as required by the "$\doteq$" notation] will lead to the partial differential symmetry equation for the symmetry vector $Z^1(\dot{x},x,t)$ [refer to (4.24)].

We transvect (7.10) by means of H^{11} (7.9). The resulting equation may be evaluated on the dynamical trajectory by means of (7.4) to obtain the associated ordinary differential equation [refer to (6.14)–(6.17)]

$$(t + c^1)\ddot{z}^1 + \dot{z}^1 \doteq 0. \tag{7.11}$$

The solution to (7.11) is

$$z^1 = b^1 \ln(t + c^1) + b^2. \tag{7.12}$$

Hence corresponding to (6.18) we have by inspection of (7.12) that

$$g_1^1 = \ln(t + c^1), \quad g_2^1 = 1. \tag{7.13}$$

By the procedure described in Theorem 6.2 we may now construct the solution $Z^1(\dot{x},x,t)$ to the partial differential equation (4.24) obtained from the expansion of the symmetry condition (7.10). Hence from (7.12) the solution Z^1 as given by (6.20) takes the form

$$Z^1(\dot{x},x,t) = B^1 \ln[t + C^1(\dot{x},x,t)] + B^2, \tag{7.14}$$

where $B^1(\dot{x},x,t)$ and $B^2(\dot{x},x,t)$ are arbitrary constants of motion of the dynamical system (7.3) and hence may be regarded as arbitrary functions of the constant of motion C^1 (7.6) and C^2 (7.7). We now make use of (7.6) in (7.14) and find the *general* solution of the partial differential symmetry equation obtained by expansion of the symmetry condition (7.10) to be

$$Z^1(\dot{x},x,t) = -B^1\dot{x} + B^2. \tag{7.15}$$

It is of interest to note that by appropriate choices of the arbitrary constants of motion B^1 and B^2, Noether and non-Noether symmetry solutions may be obtained from the general solution (7.15). For example, the choice $B^1 = C^1$ (7.6) and $B^2 = C^2$ (7.7) reduces (7.15) to a Noether solution as described in Theorem 5.1 (refer also to Corollary 6.2.1) in that the resulting symmetry vector

$$Z^1 = -C^1\dot{x} + C^2 \tag{7.16}$$

may be expressed in the Noether form (5.39)

$$Z^1 = e^{\dot{x}} \frac{\partial}{\partial \dot{x}} (-C^1C^2), \tag{7.17}$$

where with reference to (5.37) and (5.38) the Noether constant of motion $Z = I_N = -C^1C^2$, and with reference to (5.40) the Noether function $\tau^* = -\dot{x}e^{-2\dot{x}} - te^{-\dot{x}} + xt$.

If in the general symmetry solution (7.15) we choose $B^1 = C^1$ (7.6) and $B^2 = 0$ it is readily shown that the resulting symmetry vector

$$Z^1 = -(e^{-\dot{x}} - t)\dot{x} \tag{7.18}$$

is a non-Noether solution.

Example II: We again use the procedure developed in the proceding section to obtain the general velocity-dependent mappng of the one-dimensional system

$$E_1 \equiv x\ddot{x} - \dot{x}^2 = 0. \tag{7.19}$$

It is easily verified that the solution to the dynamical equation (7.19) is

$$x = c^1e^{c^2t}, \quad c^1,c^2 \text{ const.} \tag{7.20}$$

From (7.20) we find

$$\dot{x} = c^1c^2e^{c^2t}. \tag{7.21}$$

Equations (7.20) and (7.21) can be solved for c^1 and c^2. This inversion procedure leads to two functionally independent constants of motion

$$C^1(\dot{x},x,t) \equiv xe^{-\dot{x}t/x} \overset{t}{=} c^1, \tag{7.22}$$

$$C^2(x,x,t) \equiv \dot{x}/x \overset{t}{=} c^2. \tag{7.23}$$

With reference to (6.11) and (7.19) we find

$$H_{11} = x, \quad J_{11} = -2\dot{x}, \quad K_{11} = \dot{x}^2/x. \tag{7.24}$$

It then follows from (7.24) that the symmetry condition (6.10) takes the form

$$\Delta E_1/\delta a \doteq x\ddot{Z} - 2\dot{x}\dot{Z} + (\dot{x}^2/x)Z \doteq 0. \tag{7.25}$$

Expansion of (7.25) with use of the dynamical equation (7.19) to eliminate $\ddot{x}$ and $\dddot{x}$ gives the partial differential equation which determines the symmetry mapping function $Z(\dot{x},x,t)$ [refer to the two paragraphs immediately following (6.11)]

$$\frac{\dot{x}^4}{x}\frac{\partial^2 Z}{\partial \dot{x}\,\partial \dot{x}} + 2\dot{x}^3\frac{\partial^2 Z}{\partial \dot{x}\,\partial x} + x\dot{x}^2\frac{\partial^2 Z}{\partial x\,\partial x} + 2\dot{x}^2\frac{\partial^2 Z}{\partial \dot{x}\,\partial t} + 2x\dot{x}\frac{\partial^2 Z}{\partial x\,\partial t} + x\frac{\partial^2 Z}{\partial t\,\partial t} - \frac{\dot{x}^3}{x}\frac{\partial Z}{\partial \dot{x}} - \dot{x}^2\frac{\partial Z}{\partial t} - 2\dot{x}\frac{\partial Z}{\partial t} + \frac{\dot{x}^2}{x}Z \doteq 0. \tag{7.26}$$

To solve (7.26) we follow the procedures described in Theorems 6.1 and 6.2. We evaluate (7.25) on the trajectories (7.20) to obtain (after simplificaton) the associated ordinary differential equation

$$\ddot{z} - 2c^2\dot{z} + (c^2)^2z \overset{t}{=} 0, \tag{7.27}$$

which has as its solution

$$z = b^1e^{c^2t} + b^2te^{c^2t}, \quad b^1,b^2 = \text{const.} \tag{7.28}$$

[From (7.28) the functions $g_A^i(c,t)$ appearing in (6.19) have the form

$$g_1^1(c,t) \equiv e^{c^2t}, \quad g_2^1(c,t) = te^{c^2t}.] \tag{7.29}$$

The solution $Z(\dot{x},x,t)$ of (7.26) is constructed from (7.28) by replacing the constants b^1,b^2 by arbitrary constants of motion $B^1(\dot{x},x,t)$, $B^2(\dot{x},x,t)$ of the dynamical system (7.19), and by replacing the constant c^2 by the constant of motion $C^2(\dot{x},x,t)$ (7.23). As result of this procedure we obtain

$$Z(\dot{x},x,t) = B^1(\dot{x},x,t)e^{\dot{x}t/x} + B^2(\dot{x},x,t)te^{\dot{x}t/x}. \tag{7.30}$$

It is easily verified that $Z(\dot{x},x,t)$ (7.30) will satisfy (7.26) identically in $\dot{x}$, x, and t if use is made of the fact that the constants of motion $B^A(\dot{x},x,t)$, $A = 1,2$, satisfy the conditions

$$\dot{B}^A \doteq \frac{\partial B^A}{\partial \dot{x}}\frac{\dot{x}^2}{x} + \frac{\partial B^A}{\partial x}\dot{x} + \frac{\partial B^A}{\partial t} \doteq 0. \tag{7.31}$$

[1]G. H. Katzin and J. Levine, J. Math. Phys. **9**, 8 (1968).
[2]G. H. Katzin and J. Levine, Colloq. Math. (Wroclaw, Poland) **26**, 21 (1972).
[3]G. H. Katzin, J. Math. Phys. **14**, 1213 (1973).
[4]G. H. Katzin, Lett. Nuovo Cimento **7**, 213 (1973).
[5]G. H. Katzin and J. Levine, J. Math. Phys. **14**, 1886 (1973).
[6]G. H. Katzin and J. Levine, J. Math. Phys. **15**, 1460 (1974).
[7]G. H. Katzin and J. Levine, J. Math. Phys. **16**, 548 (1975).
[8]G. H. Katzin and J. Levine, J. Math. Phys. **17**, 1345 (1976).
[9]G. H. Katzin, J. Levine, and R. N. Sane, J. Math. Phys. **18**, 424 (1977).
[10]G. H. Katzin and J. Levine, J. Math. Phys. **18**, 1267 (1977).
[11]G. H. Katzin and J. Levine, Tensor **34**, 179 (1980).
[12]G. H. Katzin and J. Levine, J. Math. Phys. **22**, 1878 (1981).
[13]G. H. Katzin and J. Levine, J. Math. Phys. **23**, 552 (1982).
[14]G. H. Katzin and J. Levine, J. Math. Phys. **24**, 1761 (1983).
[15]When it is convenient we shall employ the standard dot notation to indicate total derivative with respect to t. Hence $\dot{F} \equiv dF/dt$, $\ddot{F} \equiv d^2F/dt^2$.
[16]W. Sarlet and F. Cantrijn, SIAM Rev. **23**, 467 (1981).
[17]G. H. Katzin and J. Levine, J. Math. Phys. **26**, 3100 (1985).
[18]C. Palmieri and B. Vitale, Nuovo Cimento A **66**, 299 (1970).
[19]E. Candotti, C. Palmieri, and B. Vitale, Am. J. Phys. **40**, 424 (1972).

[20]For any function $F(t, x^i, dx^i/dt, d^2x^i/dt^2, \cdots)$ the notation $d^0F/dt^0 \equiv F$, For example, $d^0x^i/dt^0 \equiv x^i$.

[21]For additional discussion refer to Ref. 16.

[22]In the formulation of dynamical symmetries of Hamilton's equations in phase space one is led in a natural manner to introduce such a decomposition to simplify the system of differential equations which determine the mapping functions ξ^i, ξ^0. See Sec. II of Ref. 7.

[23]Note for the choice $\alpha = 0$ in (2.22) we recover the defining relationship (2.10).

[24]The Einstein summation notation is applied to both the coordinate indices (lowercase Latin) and the derivative indices (lowercase Greek) so that

$$\frac{\partial G}{\partial x^i_\alpha}\delta x^i_\alpha \equiv \sum_{\alpha=0}^{N}\sum_{i=1}^{n}\frac{\partial G}{\partial x^i_\alpha}\delta x^i_\alpha.$$

We shall indicate the range of Greek indices in each equation, since in some equations they start with "0" and in others they start with "1."

[25]Other procedures for formulating infintesimal symmetries exist. For example, one could require that the condition

$$\delta\Lambda_i(L) \equiv \sigma^j_i\Lambda_j(L) + \mu^j_i\, d\Lambda_j(L)/dt$$

be satisfied identically in the variables $d^2x^i/dt^2, d^3x^i/dt^3$ for some functions $\sigma^j_i(\dot{x},x,t), \mu^j_i(\dot{x},x,t)$.

[26]See Eq. (3.10) of Ref. 8.

[27]See, for example, Ref. 8, Eq. (3.1) for the velocity-independent case.

[28]D. S. Djukic, Int. J. Non-Linear Mech. **8**, 479 (1973), has obtained this equation from a variational integral formulation.

[29]This is to be expected since, as will be shown below, the class of velocity-dependent Noether mapping is a subclass of *symmetry* mappings and hence satisfies (4.16).

[30]M. Lutzky, J. Phys. A: Math. Gen. **12**, 973 (1979), has formulated velocity-dependent Noether mappings for the " $\doteq$ " case but avoids the trivial situation by assuming τ is indepdent of $\dot{x}^i$.

[31]This viewpoint is in agreement with that expressed in Refs. 16 and 29.

[32]Constants of motion of the Noether type are of the form (5.36) or (5.37).

[33]E. J. Saletan and A. H. Cromer, *Theoretical Mehcanics* (Wiley, New York, 1971), p. 83.

[34]This argument concerning the vanishing identically of $F(\dot{x},x,t)$ appears in Ref. 19. It will also be essential in the analysis of general symmetries which we shall consider in Sec. VI.

[35]C. W. Kilmister, *Hamiltonian Dynamics, Mathematical Physics Series* (Am. Elsevier, New York, 1965), Sec. 4.3.

[36]This argument is similar to that used in Ref. 19.

[37]R. P. Agnew, *Differential Equations* (McGraw–Hill, New York, 1960), p. 466.

[38]T. Muir, *A Treatise on the Theory of Determinants*, revised by William H. Metzler (Dover, New York, 1960), p. 63.

[39]This result is essentially Abel's theorem.

Characteristic functional structure of infinitesimal symmetry mappings of classical dynamical systems. II. Mappings of first-order differential equations

Gerald H. Katzin
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202
Jack Levine
Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205

(Received 14 May 1985; accepted for publication 18 June 1985)

Infinitesimal velocity-dependent symmetry mappings of second-order dynamical systems (a) $E^i(\ddot{x}, \dot{x}, x, t) \equiv \ddot{x}^i - F^i(\dot{x}, x, t) = 0$, $i = 1,..., n$, were studied in considerable detail in a previous paper [J. Math. Phys. **26**, 3080 (1985), the first of this series]. Among the results developed in that paper was a procedure for determining the characteristic functional structure of symmetry mappings for such second-order systems. In this present companion paper it is shown that a similar procedure may be used to obtain the characteristic functional structure of infinitesimal symmetry mappings (b) $\bar{y}^I = y^I + \delta y^I$, $\delta y^I \equiv \eta^I(y, t)\delta a$; (c) $\bar{t} = t + \delta t$, $\delta t \equiv \eta^0(y, t)\delta a$, for systems of first-order differential equations (d) $E^I(\dot{y}, y, t) \equiv \dot{y}^I - \lambda^I(y, t) = 0$, $I = 1,..., N$. This characteristic structure is the same for all first-order systems (d) and is explicitly dependent upon constants of motion of the system. For the special case in which (d) is a system of $N = 2n$ equations derived from a system of n second-order equations (a) it is shown how the respective symmetry equations based upon these two equivalent dynamical descriptions are related and how their symmetry solutions are correlated. Two examples are given.

I. INTRODUCTION

In the first paper of this series[1] (denoted by I) an analysis of infinitesimal *velocity-dependent* symmetry mappings of classical (including relativistic) particle dynamical systems described by second-order differential equations

$$E^i(\ddot{x}^1,..., \ddot{x}^n, \dot{x}^1,..., \dot{x}^n, x^1,... ,x^n, t) \equiv E^i(\ddot{x}, \dot{x}, x, t) = 0, \quad i = 1,..., n, \tag{1.1}$$

was presented. For such dynamical equations (1.1) it was shown that symmetry mappings

$$\bar{x}^i = x^i + \delta x^i, \quad \delta x^i \equiv \xi^i(\dot{x}, x, t)\delta a, \tag{1.2}$$

with associated change in path parameter

$$\bar{t} = t + \delta t, \quad \delta t \equiv \xi^0(\dot{x}, x, t)\delta a, \tag{1.3}$$

were expressible in a form with a characteristic functional structure which was the same for all dynamical equations (1.1), and which was manifestly dependent upon constants of motion of the dynamical system. In this characteristic form the (velocity-dependent) symmetry mapping functions were given by

$$\xi^i(\dot{x}, x, t) = Z^i(\dot{x}, x, t) + x^i\xi^0(\dot{x}, x, t), \tag{1.4}$$

$$\xi^0(\dot{x}, x, t) = \text{arbitrary}; \tag{1.5}$$

the functions $Z^i(\dot{x}, x, t)$ in (1.4) had the form

$$Z^i(\dot{x}, x, t) = B^A(\dot{x}, x, t)g^i_A[C^1(\dot{x}, x, t),...,C^r(\dot{x}, x, t), t], \quad 0 \leqslant r \leqslant 2n, \quad A = 1,...,2n, \tag{1.6}$$

where the B^A were arbitrary constants of motion, and the C's appearing in the functions g^i_A were specific constants of motion. A procedure was given to determine the g^i_A.

With slight modifications the method developed in paper I for obtaining the characteristic structure of *velocity-dependent* symmetry mappings of second-order differential equations is also applicable for determining the characteristic structure of *velocity-independent* symmetry mappings of systems of first-order differential equations. Such first-order systems of differential equations may arise, for example, from a $2n$-dimensional description of an n-dimensional dynamical system, as in the case of Hamilton's equations.

Due to the similarity in the derivations of the characteristic structures of symmetry mappings of first- and second-order systems of differential equations we give in this paper only a brief sketch of the derivation for first-order equations.

For the special case in which a system of $2n$ first-order dynamical equations is derived from a system of n second-order dynamical equations it is shown how the respective symmetry equations, based upon these two equivalent descriptions of a dynamical system, are related and how their symmetry solutions are correlated. Two examples are given.

II. STRUCTURE OF SYMMETRY MAPPINGS

Consider a system of first-order differential equations which is expressible in the form

$$E^I(\dot{y}^A, y^A, t) \equiv E^I(\dot{y}, y, t) \equiv \dot{y}^I - \lambda^I(y, t) = 0, \quad I, A = 1,..., N. \tag{2.1}$$

A complete solution of Eqs. (2.1) is denoted by

$$y^I = \Phi^I(\gamma^1,..., \gamma^N, t) \equiv \Phi^I(\gamma, t), \quad \gamma^I \equiv \text{const.} \tag{2.2}$$

Inversion of the solution (2.2) leads to the N functionally independent constants of motion

$$\Gamma^I(y, t) \equiv \Gamma^I(y^{i+n}, y^i, t) \overset{t}{=} \gamma^I, \quad i = 1,..., n. \tag{2.3}$$

Remark 2.1: The notation $\overset{t}{=}$ indicates that the finite solution (2.2) is used to express a function completely in terms of the path parameters t. ■

The reciprocity between a complete solution and the constants of motion obtained by its inversion will be essential

 0022-2488/85/123100-05$02.50

in the determination of the characteristic functional structure of symmetry mappings of first-order systems, as was the case for second-order systems.

In our symmetry analysis of second-order differential equations treated in paper I we assumed the highest-order derivatives of the coordinates appearing in the mapping functions ξ^i, ξ^0 to be one less than the order of the differential equations. As a consequence the mapping functions were taken to be velocity dependent. In our symmetry analysis of first-order differential equations to be treated in this present paper we make a similar assumption and thereby take the mapping functions η^I, η^0 to be velocity independent. We thus assume δ variations to be based on point mappings

$$\bar{y}^I = y^I + \delta y^I, \quad \delta y^I \equiv \eta^I(y, t)\delta a, \tag{2.4}$$

with associated change in path parameter

$$\bar{t} = t + \delta t, \quad \delta t \equiv \eta^0(y, t)\delta a. \tag{2.5}$$

Based on such velocity-independent mappings (2.4) and (2.5) the symmetry mapping condition for first-order systems (2.1) is also formally written in the form[2]

$$\delta E^I(\dot{y}, t, t) \doteq 0. \tag{2.6}$$

Remark 2.2: The " $\doteq$ " notation for first-order differential equations: Conceptually the " $\doteq$ " notation is essentially the same for first-order systems of equations as it is for second-order systems (refer to paper I, Remark 4.1). However, for first-order systems (2.1) the " $\doteq$ " notation implies that the differential equation (2.1) is to be used to eliminate *all* dot derivatives of the coordinates from expressions or equations. ■

With reference to [I-(2.26)] and [I-(2.30)] (this notation used to refer to equations of paper I), we now find for the differential equation (2.1) that the symmetry condition (2.6) leads to

$$\dot{\eta}^I - \dot{y}^I\dot{\eta}^0 - \frac{\partial \lambda^I}{\partial y^J}\eta^J - \frac{\partial \lambda^I}{\partial t}\eta^0 \doteq 0. \tag{2.7}$$

For first-order systems a modified version of the decomposition [I-(2.12)] is found to be useful.[3] We therefore write

$$\eta^I(y, t) = U^I(y, t) + \lambda^I(y, t)\eta^0(y, t). \tag{2.8}$$

By use of (2.1) and (2.8) the symmetry condition (2.7) may be expressed in the form

$$\dot{U}^I + K^I{}_J U^J = 0, \tag{2.9}$$

where

$$K^I{}_J(y, t) \equiv -\frac{\partial \lambda^I}{\partial y^J}. \tag{2.10}$$

The function $\eta^0(y, t)$ does not appear in (2.9) and therefore may be taken to be arbitrary, as was the case for $\xi^0(\dot{x}, x, t)$ in the analysis of velocity-dependent symmetry mappings of second-order differential equations. However, the arbitrariness of $\eta^0(y, t)$ for velocity-independent symmetry mappings of first-order systems is in contrast to the nonarbitrariness of $\xi^0(x, t)$ for velocity-independent symmetry mappings of second-order systems.

From this point on the procedure for obtaining the characteristic functional structure of the $U^I(y, t)$ solutions of (2.9) is essentially the same as that for determining the characteristic structure of the $Z^i(\dot{x}, x, t)$ symmetry functions of second-order systems. We therefore omit the details of the analysis which leads to the following theorem about symmetry mappings of first-order differential equations.

Theorem 2.1: Consider a system of first-order differential equations expressible in the form

$$E^I \equiv \dot{y}^I - \lambda^I(y, t) = 0, \quad I = 1,...,N, \tag{2.1'}$$

with a complete solution given by

$$y^I = \Phi^I(\gamma^1,..., \gamma^N, t) \equiv \Phi^I(\gamma, t), \quad \gamma^I = \text{const.} \tag{2.2'}$$

Let

$$\Gamma^I(y, t) \overset{t}{=} \gamma^I \tag{2.3'}$$

be the specific set of N functionally independent constants of motion obtained by inversion of (2.2') for the constants γ^I.

An infinitesimal mapping

$$\bar{y}^I = y^I + \delta y^I, \quad \delta y^I \equiv \eta^I(y, t)\delta a, \tag{2.4'}$$

$$\bar{t} = t + \delta t, \quad \delta t \equiv \eta^0(y, t)\delta a, \tag{2.5'}$$

will define a symmetry mapping of the system (2.1') if the mapping functions are solutions of the symmetry condition

$$\delta E^I \doteq 0. \tag{2.6'}$$

The mapping functions $\eta^I(y, t), \eta^0(y, t)$ which are determined by (2.6') are expressible in the form

$$\eta^I(y, t) = U^I(y, t) + \lambda^I(y, t)\eta^0(y, t), \tag{2.8'}$$

$$\eta^0(y, t) = \text{arbitrary}, \tag{2.11}$$

where the functions $U^I(y, t)$ are solutions of the partial differential equations

$$\frac{\partial U^I}{\partial y^J}\lambda^J + \frac{\partial U^I}{\partial t} - \frac{\partial \lambda^I}{\partial y^J}U^J \doteq 0, \tag{2.12}$$

obtained by the formal expansion of the auxiliary symmetry condition

$$\dot{U}^I + K^I{}_J(y, t)U^J \doteq 0, \tag{2.9'}$$

where

$$K^I{}_J(y, t) = -\frac{\partial \lambda^I(y, t)}{\partial y^J}. \tag{2.10'}$$

Evaluation of (2.9') on the solution curves (2.2') of (2.1') gives the associated ordinary differential equations (refer to paper I, Theorem 6.1)

$$\dot{u}^I(\gamma, t) + k^I{}_J(\gamma, t)u^J(\gamma, t) \overset{t}{=} 0, \tag{2.13}$$

where

$$U^I(y, t) \overset{t}{=} u^I(\gamma, t), \quad K^I{}_J(y, t) \overset{t}{=} k^I{}_J(\gamma, t). \tag{2.14}$$

A complete solution to (2.13) is

$$u^I(\gamma, t) = \beta^J G^I_J(\gamma, t), \quad \beta^J = \text{const.} \tag{2.15}$$

Every solution $U^I(y, t)$ of the partial differential equations (2.12) is expressible in the form (refer to paper I, Theorem 6.2)

$$U^I(y, t) = B^J(y, t)G^I_J[\Gamma(y, t), t], \tag{2.16}$$

where the $B^J(y, t)$ are arbitrary constants of motion of (2.1')

and therefore are expressible as functions of the constants of motion $\Gamma^I(y,t)$ (2.3'), and the functions $G^I_J[\Gamma(y,t),t]$ are obtained by replacing the constants γ^I appearing in the functions $G^I_J(\gamma,t)$ of (2.15) by the respective constants of motion $\Gamma^I(y,t)$ of (2.3'). ■

Corollary 2.1.1: Since in (2.16) the $B^J(y,t)$ are arbitrary constants of motion it follows that each of the N vectors $U^I_{(J)} \equiv G^I_J$, where I denotes component and J denotes vector, are symmetry vectors. ■

Corollary 2.1.2: If $U^I(y,t)$ is a solution of the partial differential symmetry equation (2.12), then so is

$$\tilde{U}^I(y,t) \equiv A(y,t)U^I(y,t), \tag{2.17}$$

where $A(y,t)$ is a constant of motion of the differential equation (2.1) ■

Remark 2.3: Results similar to Corollary 2.1.2 were obtained for second-order systems of differential equations (refer to paper I, Corollary 6.2.3), and also for first-order differential equations of Hamilton form (refer to Theorem 5.1 of Ref. 3). ■

Corollary 2.1.3: If $U^I(y,t)$ is a solution of the partial differential symmetry equations (2.12) associated with an autonomous first-order system of differential equations (2.1) [in that $\lambda^I = \lambda^I(y)$], then $\partial U^I(y,t)/\partial t$ will also be a solution of (2.12). ■

Remark 2.4: A similar result was found for velocity-dependent symmetry solutions $Z^i(\dot{x},x,t)$ in the analysis of autonomous second-order differential equations (refer to paper I, Corollary 4.2). A similar result also holds for velocity-independent symmetries of second-order systems.[4] ■

Corollary 2.1.4: For autonomous differential equations (2.1) the functions $\partial G^I_J[\Gamma(y,t),t]/\partial t$ can be expressed in the form

$$\frac{\partial G^I_J[\Gamma(y,t),t]}{\partial t} = B^K_J[\Gamma(y,t)]G^I_K[\Gamma(y,t),t], \tag{2.18}$$

for appropriately chosen constants of motion $B^K_J[\Gamma(y,t)]$. In (2.18) the indicated partial differentiation with respect to t is to also include the t appearing in the argument $\Gamma(y,t)$ (keeping y fixed). ■

III. AN ILLUSTRATION OF THEOREM 2.1

The procedure described in Theorem 2.1 will be illustrated by determining the symmetry mappings of the one-dimensional first-order differential equation[5]

$$E(\dot{y},y,t) \equiv \dot{y} - (2y/t)\ln y = 0. \tag{3.1}$$

It is readily verified that the solution to (3.1) is

$$y = e^{\gamma t^2}, \quad \gamma = \text{const}, \tag{3.2}$$

and that inversion of (3.2) for γ gives the constant of motion

$$\Gamma(y,t) = (1/t^2)\ln y \overset{t}{=} \gamma. \tag{3.3}$$

Comparison of (3.1) with (2.1) shows that

$$\lambda(y,t) = (2y/t)\ln y. \tag{3.4}$$

With reference to (2.10), it follows from (3.4) that

$$K(y,t) = -(2/t)(1+\ln y). \tag{3.5}$$

Hence (2.9) takes the form

$$\dot{U} - (2/t)(1+\ln y)U \doteq 0. \tag{3.6}$$

The expanded form of (3.6) obtained by use of (3.4) in (2.17) gives

$$\left[\frac{2y}{t}\ln y\right]\frac{\partial U}{\partial y} + \frac{\partial U}{\partial t} - \frac{2}{t}(1+\ln y)U \doteq 0, \tag{3.7}$$

which is the partial differential equation to be solved for the auxiliary symmetry mapping function $U(y,t)$.

To obtain the solution of (3.7) by the method summarized in Theorem 2.1 the finite solution curves (3.2) are used in (3.5) to obtain [refer to (2.14)]

$$k(\gamma,t) = -2(\gamma t + 1/t). \tag{3.8}$$

With the use of (3.8) the associated ordinary differential equation (2.13) is found to be

$$\dot{u} - 2(\gamma t + 1/t)u \overset{t}{=} 0. \tag{3.9}$$

The solution of (3.9) is

$$u(\gamma,t) = \beta t^2 e^{\gamma t^2}, \quad \beta = \text{arbitrary const}. \tag{3.10}$$

The solution of the partial differential equation (3.7) is constructed from (3.10) by replacing the constant γ with the constant of motion $\Gamma(y,t)$ by means of (3.3) and by replacing the integration constant β with an arbitrary constant of motion $B(y,t)$ [any function of $\Gamma(y,t)$]. We thereby obtain [refer to (2.16)]

$$U(y,t) = B(y,t)t^2 y. \tag{3.11}$$

It can be readily verified that the auxiliary symmetry function $U(y,t)$ given by (3.11) is a solution of (3.7), if use is made of the condition

$$\left[\frac{2y}{t}\ln y\right]\frac{\partial B}{\partial y} + \frac{\partial B}{\partial t} \doteq 0, \tag{3.12}$$

which is a consequence of $B(y,t)$ being a constant of motion of (3.1).

It therefore follows from (2.4), (2.5), (2.8), (2.11), (3.4), and (3.11) that the differential equation (3.1) admits the infinitesimal symmetry mapping

$$\bar{y} = y + [B(y,t)t^2 y + [(2y/t)\ln y]\eta^0(y,t)]\delta a, \tag{3.13}$$

$$\bar{t} = t + \eta^0(y,t)\delta a, \quad \eta^0(y,t) \text{ arbitrary}. \tag{3.14}$$

Each choice of the constant of motion B and the mapping function η^0 will determine a symmetry mapping.

IV. CORRELATIONS BETWEEN SYMMETRY MAPPINGS OF RELATED SYSTEMS OF FIRST- AND SECOND-ORDER DIFFERENTIAL EQUATIONS

Consider a system of n second-order differential equations expressible in the form

$$E^i(\ddot{x},\dot{x},x,t) \equiv \ddot{x}^i - F^i(\dot{x},x,t) = 0, \quad i = 1,\ldots,n. \tag{4.1}$$

In terms of $N = 2n$ variables y^I, $I = 1,\ldots,2n$, defined by

$$y^i \equiv x^i, \tag{4.2}$$

$$y^{i+n} \equiv \dot{x}^i, \tag{4.3}$$

the n second-order equations (4.1) may be expressed as $2n$ first-order equations

$$\dot{y}^I - \lambda^I(y,t) = 0, \quad I = 1,\ldots,N, \tag{4.4}$$

where

$$\lambda^i(y,t) \equiv y^{i+n}, \quad i = 1,\ldots,n, \tag{4.5}$$

$$\lambda^{i+n}(y,t) \equiv F^i(y^{j+n}, y^j, t) = F^i(\dot{x}^j, x^j, t). \tag{4.6}$$

Theorem 2.1 describes the characteristic functional structure of the symmetry mapping of first-order systems of the type (4.4). In paper I the characteristic functional structure of symmetry mappings of second-order systems of the type (4.1) was discussed. Since (4.1) and (4.4) are two descriptions of the same dynamical system it is of interest to show how the respective symmetry conditions based upon these two descriptions are related and to correlate their solutions.

With use of (2.10) the expansion of the symmetry conditions (2.9) for a general system of $N = 2n$ first-order equations (2.1) with the range $I, J = 1,\ldots,2n$ leads to the equations

$$\dot{U}^i - \frac{\partial\lambda^i(y,t)}{\partial y^j} U^j - \frac{\partial\lambda^i(y,t)}{\partial y^{j+n}} U^{j+n} \doteq 0, \tag{4.7}$$

$$\dot{U}^{i+n} - \frac{\partial\lambda^{i+n}(y,t)}{\partial y^j} U^j - \frac{\partial\lambda^{i+n}(y,t)}{\partial y^{j+n}} U^{j+n} \doteq 0. \tag{4.8}$$

When the $N = 2n$ first-order equations (2.1) are derived from the system of n second-order equations (4.1), the $\lambda^I(y)$ of (2.1) take the form (4.5) and (4.6) and the symmetry conditions (4.7) and (4.8) for $U^I(y,t)$ reduce, respectively, to

$$\dot{U}^i - U^{i+n} \doteq 0, \tag{4.9}$$

$$\dot{U}^{i+n} - \frac{\partial F^i(y,t)}{\partial y^j} U^j - \frac{\partial F^i(y,t)}{\partial y^{j+n}} U^{j+n} \doteq 0. \tag{4.10}$$

With reference to [I-(6.11)] and [I-(6.12)] the symmetry conditions which determine the velocity-dependent symmetry mapping functions $Z^i(\dot{x}, x, t)$ for the second-order dynamical equations (4.1) take the form

$$\ddot{Z}^i - \frac{\partial F^i(\dot{x},x,t)}{\partial \dot{x}^j} \dot{Z}^j - \frac{\partial F^i(\dot{x},x,t)}{\partial x^j} Z^j \doteq 0. \tag{4.11}$$

We now verify that the $2n$ first-order symmetry equations (4.9) and (4.10) for $U^I(y,t)$ are equivalent to the n second-order symmetry equations (4.11) for $Z^i(\dot{x}, x, t)$. From (4.2), (4.3), and the basic definitions [I-(2.11)], and [I-(2.16)] we have

$$U^i(y,t) = Z^i(\dot{x}, x, t), \tag{4.12}$$

$$U^{i+n}(y,t) \doteq \dot{Z}^i(\dot{x}, x, t), \tag{4.13}$$

where the equality indicated in (4.13) requires the use of " $\doteq$ " in order to eliminate the $\ddot{x}^i$ which occurs in the expansion of $\dot{Z}^i$. By means of (4.2), (4.3), (4.6), (4.12), and (4.13), it is found that (4.9) transforms into an equation which is identically zero and (4.10) transforms into (4.11). In a similar manner by use of (4.12) and (4.13) the second-order system (4.11) may be transformed into the first-order system (4.9) and (4.10).

We show finally the correlation between the characteristic functional structure of the n functions $Z^i(\dot{x}, x, t)$ which are solutions of the second-order symmetry equations (4.11) and the characteristic functional structure of the $2n$ functions $U^I(y,t)$ which are solutions of the equivalent first-order symmetry equations (4.9) and (4.10). With reference to [I-(6.5)], [I-(6.6)], and (2.2) we obtain from (4.2) and (4.3) the following relations based upon the finite solutions of the n second-order dynamical equations [I-(6.2)] and the finite solutions of the equivalent $2n$ first-order dynamical equations (4.4)–(4.6):

$$\phi^i(c,t) = \Phi^i(\gamma, t), \tag{4.14}$$

$$\frac{\partial\phi^i(c,t)}{\partial t} = \Phi^{i+n}(\gamma, t), \tag{4.15}$$

$$c^A = \gamma^A. \tag{4.16}$$

It then follows from (4.2), (4.3), and (4.14)–(4.16) that the constants of motion $C^A(\dot{x}, x, t)$ [I-(6.4)] (obtained from the finite solution of the n second-order dynamical equation [I-(6.2)] by inversion of [I-(6.5)] and [I-(6.6)]) and the constants of motion $\Gamma^A(y,t)$ (2.3) [obtained from the finite solutions of the $2n$ equivalent first-order dynamical equations (4.4)–(4.6) by inversion of (2.2)] are equal, that is,

$$C^A(\dot{x}^i, x^i, t) = \Gamma^A(y^{i+n}, y^i, t). \tag{4.17}$$

From (4.2), (4.3), (4.12), (4.13), and (4.17), along with [I-(6.20′)] contained in Theorem 6.2 of paper I, and (2.16) contained in Theorem 2.1 of the present paper, it follows that

$$G^i_A[\Gamma(y,t), t] = g^i_A[C(\dot{x}, x, t)], \tag{4.18}$$

$$G^{i+n}_A[\Gamma(y,t), t] = \frac{\partial g^i_A}{\partial t}[C(\dot{x}, x, t), t]. \tag{4.19}$$

[The functions C^A are considered constant with respect to the partial differentiation indicated in (4.19).] Hence the fundamental solution functions G^I_A and g^I_A which determine the respective characteristic functional structure of the symmetry solutions associated with the equivalent first- and second-order dynamical equations are correlated by (4.18) and (4.19).

V. ILLUSTRATION OF DISCUSSION IN SEC. IV

To illustrate the correlation between symmetries of equivalent first- and second-order systems of differential equations we continue with the study of a dynamical system which in Sec. VII of paper I was used to illustrate symmetry theory of second-order differential equations.

Consider then the dynamical equation [I-(7.19)]

$$E_1 = x\ddot{x} - \dot{x}^2 = 0. \tag{5.1}$$

The associated symmetry condition for the dynamical equation (5.1) is given by [I-(7.25)]

$$\ddot{Z} - (2\dot{x}/x)\dot{Z} + (\dot{x}/x)^2 Z \doteq 0, \tag{5.2}$$

with symmetry solution

$$Z(\dot{x}, x, t) = B^1(\dot{x}, x, t) g^1_1(C, t) + B^2(\dot{x}, x, t) g^1_2(C, t), \tag{5.3}$$

where the B's are arbitrary functions of the constants of motion [I-(7.22)] and [I-(7.23)]

$$C^1(\dot{x}, x, t) = xe^{-\dot{x}t/x}, \quad C^2(\dot{x}, x, t) = \dot{x}/x, \tag{5.4}$$

and where the functions [I-(7.29)]

$$g^1_1(C, t) = e^{C^2 t}, \quad g^1_2 = te^{C^2 t} \tag{5.5}$$

define a fundamental solution set.

In terms of the variables y^1, y^2 defined by

$$y^1 = x, \quad y^2 = \dot{x}, \tag{5.6}$$

the one-dimensional second-order differential equation (5.1) is converted to the equivalent two-dimensional first-order differential equations [refer to (4.4)–(4.6)]

$$\dot{y}^1 - y^2 = 0, \tag{5.7}$$

$$\dot{y}^2 - (y^2)^2/y^1 = 0. \tag{5.8}$$

For the dynamical equations (5.7) and (5.8) we find that the symmetry conditions (4.9) and (4.10) expand to the partial differential symmetry equations

$$y^2 \frac{\partial U^1}{\partial y^1} + \frac{(y^2)^2}{y^1}\frac{\partial U^1}{\partial y^2} + \frac{\partial U^1}{\partial t} - U^2 \doteq 0, \tag{5.9}$$

$$y^2 \frac{\partial U^2}{\partial y^1} + \frac{(y^2)^2}{y^1}\frac{\partial U^2}{\partial y^2} + \frac{\partial U^2}{\partial t} + \frac{(y^2)^2}{y^1} U^1 - 2\frac{y^2}{y^1} \doteq 0. \tag{5.10}$$

Based upon the fundamental solution functions g^i_A (5.5) of the second-order partial differential symmetry equation resulting from the expansion of (5.2) we construct the associated fundamental solution functions G^I_A for the equivalent first-order partial differential symmetry equations (5.9) and (5.10) by means of the correlation (4.17)–(4.19) [with use of (5.4) and (5.6)–(5.8)] to obtain

$$G^1_1(\Gamma, t) = e^{\Gamma^2 t} = e^{y^2 t/y^1}, \tag{5.11}$$

$$G^2_1(\Gamma, t) = \Gamma^2 e^{\Gamma^2 t} = (y^2/y^1)e^{y^2 t/y^1}, \tag{5.12}$$

$$G^1_2(\Gamma, t) = te^{\Gamma^2 t} = te^{y^2 t/y^1}, \tag{5.13}$$

$$G^2_2(\Gamma, t) = (1 + y^2 t/y^1)e^{y^2 t/y^1}. \tag{5.14}$$

The fundamental solution functions (5.11)–(5.14) may be used in (2.16) to obtain the general solution to the first-order partial differential symmetry equations (5.9) and (5.10). Since the B^A in (2.16) are arbitrary constants of motion it follows that

$$\mathbf{U}_{(1)} \equiv (U^1_{(1)}, U^1_{(2)}) \equiv (G^1_1, G^2_1) \tag{5.15}$$

and

$$\mathbf{U}_{(2)} \equiv (U^1_{(2)}, U^2_{(2)}) \equiv (G^1_2, G^2_2), \tag{5.16}$$

where the G^I_A are given by (5.11)–(5.14), will each be solutions of (5.9) and (5.10), as may be readily verified.

[1]G. H. Katzin and J. Levine, J. Math. Phys. **26**, 3080 (1985).
[2]Refer to Sec. II in Ref. 1 for basic variational definitions.
[3]G. H. Katzin and J. Levine, J. Math. Phys. **16**, 548 (1975).
[4]G. H. Katzin, J. Levine, and R. N. Sane, J. Math. Phys. **18**, 424 (1977).
[5]Since $N = 1$, we suppress all indices.

Exact reduced density matrices for a model problem

Leon Cohen[a)]
Center for Advanced Studies, University of New Mexico, Albuquerque, New Mexico 87131

Chongmoon Lee
Hunter College of the City University, New York, New York 10021

(Received 27 March 1985; accepted for publication 10 July 1985)

The reduced density matrices of arbitrary order for the boson problem of N particles, each attracted harmonically to a central point and interacting with each other harmonically, are analytically calculated.

I. INTRODUCTION

Although the applications of reduced density matrices have been quite extensive there are very few examples for which they have been calculated exactly. The advantage of having exact results is that they can be used to study and test different ideas relating to their properties and applications. We consider here an N-body boson problem for which we have been able to calculate exactly the reduced density matrices of arbitrary order.

The system we consider consists of N bosons of unit mass, each one harmonically attracted to a "nucleus" with spring constant ω^2 and interacting with each other with a harmonic force whose spring constant is γ^2. Various aspects regarding this model problem have been previously developed.[1-4] The Hamiltonian is

$$H=\tfrac{1}{2}\sum_{i=1}^{N}(-\nabla_i^2+\omega^2 r_i^2)\pm\frac{1}{2}\gamma^2\sum_{i<j}^{N}r_{ij}^2, \tag{1.1}$$

where the minus sign signifies mutual repulsion and the plus sign attraction. Schrödinger's equation

$$H\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)=E\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) \tag{1.2}$$

can be decoupled by making the following coordinate transformation[5,6]:

$$\begin{aligned}\mathbf{Q}_k&=\frac{1}{\sqrt{k(k+1)}}\sum_{i=1}^{k}(\mathbf{r}_{k+1}-\mathbf{r}_i),\quad 1\leqslant k\leqslant N-1,\\ \mathbf{Q}_N&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{r}_i,\end{aligned} \tag{1.3}$$

in which case the Hamiltonian becomes

$$H=\frac{1}{2}\sum_{k=1}^{N-1}(P_k^2+\delta_N^2Q_k^2)+\frac{1}{2}P_N^2+\frac{1}{2}\omega^2Q_N^2, \tag{1.4}$$

with

$$P_k=(1/i)\nabla_{Q_k}, \tag{1.5}$$

$$\delta_N^2=\omega^2\pm N\gamma^2. \tag{1.6}$$

This is the Hamiltonian for N independent harmonic oscillators in the coordinates Q_i, the first $N-1$ having spring constant δ_N^2 and the last one ω^2. The ground state wave function is therefore

$$\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)=\left(\frac{\delta_N}{\pi}\right)^{3/4(N-1)}\left(\frac{\omega}{\pi}\right)^{3/4}e^{-1/2\delta_N\sum_{k=1}^{N-1}Q_k^2-\frac{1}{2}\omega Q_N^2}, \tag{1.7}$$

and the corresponding energy is

$$E=\tfrac{3}{2}(N-1)\delta_N+\tfrac{3}{2}\omega. \tag{1.8}$$

We now express the wave function in terms of the spatial coordinates. Using Eq. (1.3) one may show that

$$Q_k^2=r_{k+1}^2-\frac{1}{k+1}\sum_{i=1}^{k+1}\sum_{j=1}^{k+1}\mathbf{r}_i\cdot\mathbf{r}_j+\frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{k}\mathbf{r}_i\cdot\mathbf{r}_j, \tag{1.9}$$

$$\sum_{k=1}^{N-1}Q_k^2=(1-\frac{1}{N})\sum_{k=1}^{N}r_k^2-\frac{2}{N}\sum_{i<j}^{N}\mathbf{r}_i\cdot\mathbf{r}_j, \tag{1.10}$$

$$Q_N^2=\frac{1}{N}\sum_{i=1}^{N}r_i^2+\frac{2}{N}\sum_{i<j}^{N}\mathbf{r}_i\cdot\mathbf{r}_j. \tag{1.11}$$

Substituting Eqs. (1.9)–(1.11) into Eq. (1.17) the ground state wave function is expressed in terms of the space coordinates as

$$\begin{aligned}\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)=&\left(\frac{\delta_N}{\pi}\right)^{(3/4)(N-1)}\left(\frac{\omega}{\pi}\right)^{3/4}\\ &\times\exp\left\{-\frac{1}{2N}((N-1)\delta_N+\omega)\right.\\ &\left.\times\sum_{i=1}^{N}r_i^2-\frac{1}{N}(\omega-\delta_N)\sum_{i<j}^{N}\mathbf{r}_i\cdot\mathbf{r}_j\right\}.\end{aligned} \tag{1.12}$$

II. EXACT REDUCED DENSITY MATRICES

We now derive the exact reduced density matrices for the following class of wave functions:

$$\psi(r_1,r_2,\ldots,r_N)=K_N\exp\left(-2A\sum_{i=1}^{N}r_i^2-4B\sum_{i<j}^{N}\mathbf{r}_i\cdot\mathbf{r}_j\right), \tag{2.1}$$

where A and B are constants and K_N is a normalizing factor. The wave function given by Eq. (1.12) is a special case. The full N-body density matrix is

$$\begin{aligned}&\rho_N(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N;\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_N')\\ &\quad=K_N^2\exp\left\{-2A\sum_{i=1}^{N}(r_i^2+r_i'^2)\right.\\ &\qquad\left.-4B\sum_{i<j}^{N}(\mathbf{r}_i\cdot\mathbf{r}_j+\mathbf{r}_i'\cdot\mathbf{r}_j')\right\}.\end{aligned} \tag{2.2}$$

For convenience we define the following function:

$$\begin{aligned}&g_N(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N;\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_N';A,B,C_N)\\ &\quad=\exp\left\{-2A\sum_{i=1}^{N}(r_i^2+r_i'^2)\right.\\ &\qquad\left.-4B\sum_{i<j}^{N}(\mathbf{r}_i\cdot\mathbf{r}_j+\mathbf{r}_i'\cdot\mathbf{r}_j')-C_NR_N^2\right\},\end{aligned} \tag{2.3}$$

[a)] Permanent address: Hunter College of the City University, New York, New York, 10021.

 0022-2488/85/123105-04$02.50

where C_N is a constant and

$$R_N = \sum_{i=1}^{N} (\mathbf{r}_i + \mathbf{r}_i') . \tag{2.4}$$

The factor $-C_N R$ has been inserted for reasons which will become clear below. Eventually we will take C_N to be zero.

Now, consider integration with respect to the N th particle. Using the fact that

$$\int e^{-ar^2 - br\cdot J}\, d\mathbf{r} = \left(\frac{\pi}{a}\right)^{3/2} e^{b^2 J^2/4a} \tag{2.5}$$

and

$$R_N^2 = 4r_N^2 + 4\mathbf{r}_N\cdot\mathbf{R}_{N-1} + R_{N-1}^2 , \tag{2.6}$$

we have

$$\begin{aligned}
\int g_N\, d\mathbf{r}_N = \exp\Big\{ & -2A\sum_{i=1}^{N-1}(r_i^2 + r_i'^2) \\
& - 4B\sum_{i<j}^{N-1}(\mathbf{r}_i\cdot\mathbf{r}_j + \mathbf{r}_i'\cdot\mathbf{r}_j') - C_N R_{N-1}^2 \Big\} \\
& \times \int \exp\left(-4(A + C_N)r_N^2\right. \\
& \left. - 4(B + C_N)\mathbf{r}_N\cdot\mathbf{R}_{N-1}\right) d\mathbf{r}_N \\
= {} & \tfrac{1}{8}\pi^{3/2}/(A + C_N)^{3/2} \\
& \times g_{N-1}(\mathbf{r}_1,\ldots,\mathbf{r}_{N-1};\mathbf{r}_1',\ldots,\mathbf{r}_{N-1}';A,B,C_N) \\
& \times \exp\{[(B + C_N)^2/(A + C_N)]R_{N-1}^2\} \\
= {} & \tfrac{1}{8}\pi^{3/2}/(A + C_N)^{3/2} \\
& \times g_{N-1}(\mathbf{r}_1,\ldots,\mathbf{r}_{N-1};\mathbf{r}_1',\ldots,\mathbf{r}_{N-1}'; \\
& A,\ B, C_{N-1}) ,
\end{aligned} \tag{2.7}$$

where we have taken

$$C_{N-1} = C_N - (B + C_N)^2/(A + C_N) . \tag{2.8}$$

Hence, integrating the $N - (s + 1)$ coordinates out of Eq. (2.3) yields

$$\begin{aligned}
\int g_N\, d\mathbf{r}_{s+1}\cdots d\mathbf{r}_N & \\
= {} & \left[\frac{\pi^{N-s}}{4^{N-s}\Pi_{j=s+1}^{N}(A + C_j)}\right]^{3/2} \\
& \times g_s(\mathbf{r}_1,\ldots,\mathbf{r}_s;\mathbf{r}_i',\ldots,\mathbf{r}_s';A,B,C_s) ,
\end{aligned} \tag{2.9}$$

where in general

$$C_i = C_{i+1} - (B + C_{i+1})^2/(A + C_{i+1}) . \tag{2.10}$$

To complete the solution we must find C_i explicitly and evaluate the product that appears in the denominator of Eq. (2.9). Equation (2.10) is a nonlinear difference equation which can be transformed into a linear difference equation by making the transformation

$$C_i = -A + (B - A)(Z_i/Z_{i+1}) . \tag{2.11}$$

Substituting this into Eq. (2.10) results in

$$Z_{i+2} + 2Z_{i+1} + Z_i = 0 , \tag{2.12}$$

which is linear and can be solved by standard methods. Taking

$$Z_i = m^i , \tag{2.13}$$

where m is to be determined, results in the requirement that

$$m^{i+2} + 2m^{i+1} + m^i = 0 \tag{2.14}$$

or

$$(m + 1)^2 = 0 . \tag{2.15}$$

As both roots are equal to -1 the general solution is

$$Z_i = (-1)^i(\eta_1 + i\eta_2) , \tag{2.16}$$

where η_1 and η_2 are arbitrary constants. Substituting this solution into Eq. (2.11) results in

$$C_i = -A + (A - B)(1 + i\eta)/[1 + (i + 1)\eta] , \tag{2.17}$$

where

$$\eta = \eta_2/\eta_1 . \tag{2.18}$$

We see therefore that there is only one arbitrary constant, η, to be solved for. We express it in terms of C_N:

$$C_N = -A + (A - B)(1 + \eta N)/[1 + \eta(N + 1)] . \tag{2.19}$$

For our case $C_N = 0$, which results in

$$\eta = -B/(A + NB) . \tag{2.20}$$

We now evaluate the product that appears in the numerator in Eq. (2.9):

$$\begin{aligned}
\prod_{i=s+1}^{N}(A + C_i) &= \prod_{i=s+1}^{N}\left\{A - B - \frac{(A - B)\eta}{1 + (i + 1)\eta}\right\} \\
&= (A - B)^{N-s}\prod_{i=s+1}^{N}\frac{1 + i\eta}{1 + (i + 1)\eta} \\
&= (A - B)^{N-s}\frac{\Pi_{i=s+1}^{N}(1 + \eta i)}{\Pi_{i=s+1}^{N}(1 + (i + 1)\eta)} \\
&= (A - B)^{N-s}\frac{1 + (s + 1)\eta}{1 + (N + 1)\eta} \\
&= (A - B)^{N-s}\frac{A + (N - s - 1)B}{A - B} .
\end{aligned} \tag{2.21}$$

By taking $s = 0$ for the first factor in Eq. (2.9) and using Eq. (2.21) we obtain the normalizing factor

$$K_N^2 = \{4(A - B)/\pi\}^{3N/2}\{[A + (N - 1)B]/[A - B]\}^{3/2} . \tag{2.22}$$

We now specialize to the wave function given by Eq. (1.12). Comparing Eq. (2.1) with Eq. (1.12) we take

$$A = (1/4N)\{(N - 1)\delta_N + \omega\} , \tag{2.23}$$

$$B = (1/4N)(\omega - \delta_N) , \tag{2.24}$$

from which

$$A - B = \tfrac{1}{4}\delta_N , \tag{2.25}$$

$$A + B(N - 1) = \tfrac{1}{4}\omega , \tag{2.26}$$

$$A + B(N - s - 1) = (1/4N)\{(N - s)\omega + s\delta_N\} . \tag{2.27}$$

Also, the factor in Eq. (2.9) is

$$(\delta_N/\pi)^{(3/2)s}\{N\omega/[(N - s)\omega + s\delta_N]\}^{3/2} . \tag{2.28}$$

The reduced density matrix of order s is defined by

$$\begin{aligned}
\rho_s(\mathbf{r}_1\cdots\mathbf{r}_s;\mathbf{r}_i'\cdots\mathbf{r}_s') = {} & \binom{N}{s}\int \psi^*(\mathbf{r}_1'\cdots\mathbf{r}_s';\mathbf{r}_{s+1}\cdots\mathbf{r}_N) \\
& \times \psi(\mathbf{r}_1\cdots\mathbf{r}_s,\mathbf{r}_{s+1}\cdots\mathbf{r}_N)\, d\mathbf{r}_{s+1}\cdots d\mathbf{r}_N .
\end{aligned} \tag{2.29}$$

Using Eqs. (2.24)–(2.29) we finally have the reduced density matrices

$$\rho_s = \binom{N}{s}\left(\frac{\delta_N}{\pi}\right)^{(3/2)s}\left\{\frac{N\omega}{(N-s)\omega+s\delta_N}\right\}^{3/2} \times g_s\left(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_s;\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_s';\right.$$
$$\left.\frac{1}{4N}[(N-1)\delta_N+\omega],\frac{1}{4N}(\omega-\delta_N),C_s\right), \tag{2.30}$$

where for convenience we rewrite some of the previous expressions in terms of the physical parameters:

$$g_s = \exp\left\{-\frac{1}{2N}(\omega+(N-1)\delta_N)\sum_{i=1}^{s}(\mathbf{r}_i^2+\mathbf{r}_i'^2) - \frac{1}{N}\sum_{i<j}^{s}(\mathbf{r}_i\cdot\mathbf{r}_j+\mathbf{r}_i'\cdot\mathbf{r}_j') - C_s R_s^2\right\}, \tag{2.31}$$

where

$$\mathbf{R}_s = \sum_{i=1}^{s}(\mathbf{r}_i+\mathbf{r}_i'), \tag{2.32}$$

$$C_s = -\frac{1}{4N}\frac{(N-s)(\omega-\delta_N)^2}{(N-s)\omega+s\delta_N}. \tag{2.33}$$

We write out the cases $s=1$ and $s=2$

$$\rho_1(\mathbf{r}_1,\mathbf{r}_i') = N\left\{\frac{\delta_N N\omega/\pi}{(N-1)\omega+\delta_N}\right\}^{3/2} e^{-a_1(\mathbf{r}_1^2+\mathbf{r}_i'^2)+a_2\mathbf{r}_1\cdot\mathbf{r}_i'}, \tag{2.34}$$

$$\rho_2(r_1,r_2;r_1',r_2') = \frac{N(N-1)}{2}\left\{\frac{N\omega\delta_N^2/\pi^2}{(N-2)\omega+2\delta_N}\right\}^{3/2} \times \exp\{-b_1(r_1^2+r_1'^2+r_2^2+r_2'^2) - b_2(\mathbf{r}_1\cdot\mathbf{r}_2+\mathbf{r}_1'\cdot\mathbf{r}_2') + b_3(\mathbf{r}_1\cdot\mathbf{r}_2'+\mathbf{r}_1'\cdot\mathbf{r}_2'+\mathbf{r}_1\cdot\mathbf{r}_1'+\mathbf{r}_2\cdot\mathbf{r}_2')\}, \tag{2.35}$$

where

$$a_1 = \frac{1}{4N}\frac{(N-1)(\omega^2+\delta_N^2)+2(N^2-N+1)\omega\delta_N}{(N-1)\omega+\delta_N}, \tag{2.36}$$

$$a_2 = \frac{1}{2N}\frac{(N-1)(\omega-\delta_N)^2}{(N-1)\omega+\delta_N}, \tag{2.37}$$

$$b_1 = \frac{1}{4N}\frac{(N-2)\omega^2+(3N-2)\delta_N^2+2(N^2-2N+2)\omega\delta_N}{(N-2)\omega+2\delta_N}, \tag{2.38}$$

$$b_2 = \frac{1}{2N}\frac{(N-2)\omega^2-(N+2)\delta_N^2+4\omega\delta_N}{(N-2)\omega+2\delta_N}, \tag{2.39}$$

$$b_3 = \frac{1}{2N}\frac{(N-2)(\omega-\delta_N)^2}{(N-2)\omega+2\delta_N}. \tag{2.40}$$

The reduced density matrix can be expressed in different forms:

$$\rho_s(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_s;\mathbf{r}_1',\mathbf{r}_2',\ldots,\mathbf{r}_s') = \exp\left\{-\frac{1}{N}(\omega+(N-1)\delta_N)\left(\sum_{i=1}^{s}\frac{\mathbf{r}_i-\mathbf{r}_i'}{2}\right)^2 + 2\delta_N\sum_{i<j}^{s}\frac{\mathbf{r}_i-\mathbf{r}_i'}{2}\cdot\frac{\mathbf{r}_j-\mathbf{r}_j'}{2}\right\} \times\rho_s\left(\frac{\mathbf{r}_1+\mathbf{r}_1'}{2},\frac{\mathbf{r}_2+\mathbf{r}_2'}{2},\ldots,\frac{\mathbf{r}_s+\mathbf{r}_s'}{2}\right), \tag{2.41}$$

where the ρ appearing on the right-hand side is the s particle density, that is

$$\rho_s(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_s) = \rho_s(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_s;\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_s). \tag{2.42}$$

Also the reduced density matrix of order s can be expressed in terms of the one particle density:

$$\rho_s(\mathbf{r}_1,\ldots,\mathbf{r}_s;\mathbf{r}_1,\ldots,\mathbf{r}_s) = \frac{(N-1)!}{(N-s)!s!}\left(\frac{\delta_N}{\pi}\right)^{(3/2)(s-1)}\left(\frac{(N-1)\omega+\delta_N}{(N-s)\omega+s\delta_N}\right)^{3/2} \times\exp\left\{-\frac{1}{N}(\omega+(N-1)\delta_N)\left(\sum_{i=1}^{s}\frac{\mathbf{r}_i-\mathbf{r}_i'}{2}\right)^2 + \delta_N\sum_{i<j}^{s}(\mathbf{r}_i\cdot\mathbf{r}_j+\mathbf{r}_i'\cdot\mathbf{r}_j')\right\} \times\rho_1\left(\sum_{i=1}^{s}\frac{\mathbf{r}_i+\mathbf{r}_i'}{2}\right). \tag{2.43}$$

The parts of the total energy can be readily expressed as

$$\left\langle -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2\right\rangle = \frac{3}{4}\{\omega+(N-1)\delta_N\}, \tag{2.44}$$

$$\left\langle \frac{1}{2}\omega^2\sum_{i=1}^{N}r_1^2\right\rangle = \frac{3}{4}\frac{\omega}{\delta_N}((N-1)\omega+\delta_N), \tag{2.45}$$

$$\left\langle \frac{1}{2}\gamma^2\sum_{i<j}^{N}r_{ij}^2\right\rangle = \frac{3}{4}\frac{\gamma^2}{\delta_N}N(N-1). \tag{2.46}$$

Also, the kinetic energy can be expressed as an integral involving the one body density,

$$\left\langle -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2\right\rangle = \int\left\{C+\frac{1}{2}\frac{N\omega\delta_N}{(N-1)\omega+\delta_N}\ln\rho_1(\mathbf{r},\mathbf{r})\right\}\rho_1\,d\mathbf{r}, \tag{2.47}$$

where C is a constant given by

$$C = \frac{3}{4N}\frac{(N-1)\omega^2+(N-1)\delta_N^2+(2N^2-2N+2)\omega\delta_N}{(N-1)\omega+\delta_N} - \frac{1}{2}\frac{N\omega\delta_N}{(N-1)\omega+\delta_N} \times\ln\left[N\left(\frac{\delta_N}{\pi}\right)^{3/2}\left\{\frac{N\omega}{(N-1)\omega+\delta_N}\right\}^{3/2}\right]. \tag{2.48}$$

III. CONCLUSION

Equation (2.30) gives the reduced density matrices of arbitrary order. In conclusion we show that the Hartree solution for this problem can also be obtained exactly. Taking

$$\psi^H(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) = \prod_{i=1}^{N}\varphi(\mathbf{r}_i), \tag{3.1}$$

the Hartree equation is

$$\epsilon\varphi(r) = \left\{-\frac{1}{2}\nabla^2+\frac{1}{2}\omega^2r^2\pm\frac{1}{2}\sum_{k=2}^{N}\gamma^2 \times\int\varphi^*(\mathbf{r}_k)(\mathbf{r}_k-\mathbf{r})^2\varphi(\mathbf{r}_k)d\mathbf{r}_k\right\}\varphi(\mathbf{r}). \tag{3.2}$$

Each of the terms in the summation is identical and evaluates to

$$\int \varphi^*(\mathbf{r}_k)(\mathbf{r}_k - \mathbf{r})^2 \varphi(\mathbf{r}_k) d\mathbf{r}_k$$
$$= r^2 + \int \varphi^*(\mathbf{r}_k) r_k^2 \varphi(\mathbf{r}_k) d\mathbf{r}_k$$
$$+ 2\int \varphi^*(\mathbf{r}_k) \mathbf{r}\cdot\mathbf{r}_k \varphi(\mathbf{r}_k) d\mathbf{r}_k$$
$$= r^2 + \int \varphi^*(\mathbf{r}_k) r_k^2 \varphi(\mathbf{r}_k) d\mathbf{r}_k , \quad (3.3)$$

where the last term in the middle step is zero due to symmetry. Equation (3.2) then becomes

$$\epsilon\varphi(\mathbf{r}) = \{ -\tfrac{1}{2}\nabla^2 + \tfrac{1}{2}\delta_{N-1}^2 r^2 \pm \tfrac{1}{2}(N-1)I\gamma^2 \}\varphi(\mathbf{r}) , \quad (3.4)$$

or

$$(-\tfrac{1}{2}\nabla^2 + \tfrac{1}{2}\delta_{N-1}^2 r^2)\varphi(\mathbf{r}) = (\epsilon \mp \tfrac{1}{2}(N-1)\gamma^2 I)\varphi(\mathbf{r}) , \quad (3.5)$$

where we have set

$$I = \int \varphi^*(\mathbf{r}) r^2 \varphi(\mathbf{r}) d\mathbf{r} . \quad (3.6)$$

Equation (3.5) is identical to the standard one particle harmonic oscillator equation. The ground state solution is

$$\varphi(r) = (\delta_{N-1}/\pi)^{3/4} e^{-(1/2)\delta_{N-1} r^2} , \quad (3.7)$$

with energy

$$\epsilon \pm \tfrac{1}{2}(N-1)\gamma^2 I = \tfrac{3}{2}\delta_{N-1} . \quad (3.8)$$

To complete the solution we must evaluate I. Substituting Eq. (3.7) into Eq. (3.6),

$$I = \left(\frac{\delta_{N-1}}{\pi}\right)^{3/2} 4\pi \int_0^\infty r^4 e^{-\delta_{N-1} r^2} dr = 3/2\ \delta_{N-1} . \quad (3.9)$$

Using Eq. (3.8) the orbital energies are therefore

$$\epsilon = \tfrac{3}{2}\delta_{N-1} \mp \tfrac{3}{4}(N-1)\gamma^2/\delta_{N-1} . \quad (3.10)$$

The total energy of the system is

$$E_H = \sum_{i=1}^N \epsilon - \sum_{i=1}^N \sum_{i<j}^N \int\int \{ \mp \tfrac{1}{2}\gamma^2 (\mathbf{r}_i - \mathbf{r}_j)^2 \}$$
$$\times |\varphi(r_i)|^2 |\varphi(r_j)|^2 d\mathbf{r}_i\, d\mathbf{r}_j$$
$$= N\epsilon \pm \tfrac{1}{4}N(N-1)\gamma^2$$
$$\times \int\int (\mathbf{r}_i - \mathbf{r}_j)^2 |\varphi(\mathbf{r}_i)|^2 |\varphi(\mathbf{r}_j)|^2 d\mathbf{r}_i\, d\mathbf{r}_j$$
$$= \tfrac{3}{2}N\delta_{N-1} . \quad (3.11)$$

The Hartree N-body wave function is

$$\psi^H(\mathbf{r}_1, \mathbf{r}_2, .., \mathbf{r}_N) = \left(\frac{\delta_{N-1}}{\pi}\right)^{(3/4)N} \exp\left(-\frac{1}{2}\delta_{N-1} \sum_{i=1}^N r_i^2 \right) , \quad (3.12)$$

from which the reduced density matrices may be calculated straightforwardly as

$$\rho_s^H(\mathbf{r}_1, ..., \mathbf{r}_s; \mathbf{r}_1', ..., \mathbf{r}_s')$$
$$= \binom{N}{s} \left(\frac{\delta_{N-1}}{\pi}\right)^{(3/2)s} \exp\left(-\frac{1}{2}\delta_{N-1} \sum_{i=1}^s (r_i^2 + r_i'^2) \right) . \quad (3.13)$$

Note added in proof: After the acceptance of this paper, the referee noticed and brought to our attention the work of S. Pruski, J. Mackowiak, and O. Missuno dealing with this model [Rep. Math. Phys. **1**, 309 (1971); **3**, 227, 241 (1972)]. Sage has also considered this model [Theoret. Chim. Acta **19**, 179 (1970)]. These authors have obtained many interesting results and have given expressions for the reduced density matrices. We thank the referee for bringing this to our attention.

ACKNOWLEDGMENT

This research was supported in part by a grant from the City University of New York, PS–CUNY Research Award Program.

[1] J. M. Levy-Leblond, Phys. Lett. A **26**, 540 (1968).
[2] E. R. Davidson, Phys. Rev. A **1**, 30 (1970).
[3] E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry* (Academic, New York, 1976).
[4] P. Phillips and E. R. Davidson, Int. J. Quantum Chem. **23**, 185 (1983).
[5] J. O. Hirschfelder and J. S. Dahler, **42**, 363 (1956).
[6] D. Jepsen and J. O. Hischfelder, Proc. Natl. Acad. Sci. USA **45**, 249 (1959).

On the spectra of SO(3) scalars in the enveloping algebra of SU(3)

H. De Meyer,[a] G. Vanden Berghe, and J. Van der Jeugt[b]
Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit-Gent, Krijgslaan 281, gebouw S9, B9000 Gent, Belgium

(Received 18 March 1985; accepted for publication 4 June 1985)

Formulas are given that make it possible to calculate the eigenvalues of the two independent SO(3) scalars O_l^0 and Q_l^0 in the SU(3) enveloping algebra.

I. INTRODUCTION

It is well known that the degeneracy problem for SO(3) states in SU(3) representations can be solved by the construction of the orthonormal eigenfunctions of the SU(3) Casimirs I_2 and I_3, the SO(3) Casimir L^2, the SO(3) generator l_0, and one additional Hermitian operator X in the enveloping algebra of SU(3) (see Refs. 1 and 2). The choice for X can be restricted in the sense that only two algebraically independent SO(3) scalars exist, one of third order and one of fourth order. All other SO(3) scalars are polynomials in these two independent ones, I_2, I_3, and L^2. Many alternative definitions and notations have been proposed for the two scalars mentioned and these are summarized by Partensky and Quesne.[3] In the present paper we treat the so-called scalar shift operators O_l^0 and Q_l^0, which were introduced by Hughes.[4] On account of relations that we previously established between products of shift operators associated to both scalars, we have been able to derive expressions for certain O_l^0 and Q_l^0 eigenvalues in closed form.[5,6]

It is the aim of the present paper to establish new simple formulas by which O_l^0 and Q_l^0 eigenvalue expressions can be easily deduced. We thereby overcome the calculational restrictions induced by the application of the shift operator formalism. The essential point in the present approach to the problem is the introduction of matrix elements in a particular nonorthogonal SO(3) basis, namely the Elliott basis,[7] which nuclear physicists are very familiar with. In fact, the results that we obtain in the present paper are very important in the context of a new extension of the nuclear interacting boson model[8] in the rotational SU(3) limit, which we have developed very recently.[9]

II. SPECTRUM OF THE O_l^0 SHIFT OPERATOR

The SO(3) scalar operator O_l^0 is a polynomial in the SU(3) generators that is homogeneous and quadratic in the principal SO(3) subalgebra generators l_0, $l_\pm$ and that is linear with respect to the remaining generators q_μ ($\mu = -2, -1, 0, 1, 2$), the components of a five-dimensional SO(3) tensor representation. This polynomial form of O_l^0 together with the equivalent polynomial expressions of the associated shift operators O_l^k ($k = -2, -1, 1, 2$), which have previously been derived,[4,5] allow us to obtain the following relationship between certain O_l^0- and q_0-matrix elements, respectively, in a SO(3) basis:

$$\langle l, m'=0, \alpha' | O_l^0 | l, m=0, \alpha\rangle = \sqrt{6}(2l+3)(2l-1)\langle l, m'=0, \alpha' | q_0 | l, m=0, \alpha\rangle. \tag{2.1}$$

Herein α and α' denote appropriate values of an additional label that distinguishes between states with the same l value, which usually occur in the SO(3) decomposition of a SU(3) (λ, μ) representation. Furthermore, it should be clear that the relation (2.1) only holds for the elements that are diagonal in the SO(3) representation label l.

In order to calculate the O_l^0 spectrum it is extremely useful to select as a particular SO(3) basis the so-called Elliott basis[7] consisting of states $|l, m, K\rangle$, where K is the additional label. The decomposition of a SU(3) representation (λ, μ) in Elliott states $|l, m, K\rangle$ is dictated by the following formulas prescribing the values which the state labels can take:

$$\begin{aligned} &K = \min(\lambda, \mu), \min(\lambda, \mu) - 2, \ldots, 0 \text{ or } 1, \\ &l = K, K+1, \ldots, K + \max(\lambda, \mu), \quad \text{if } K > 0, \\ &l = \max(\lambda, \mu), \max(\lambda, \mu) - 2, \ldots, 0 \text{ or } 1, \quad \text{if } K = 0, \\ &m = -l, -l+1, \ldots, l-1, l. \end{aligned} \tag{2.2}$$

It is well known that the Elliott basis is not an orthogonal basis,[7] but this fact is definitely not prohibitive to the calculation of O_l^0 eigenvalues. Moreover, such a calculation is facilitated by a formula, which has been derived by Elliott[7] and which gives an expression for the action of the q_0 generator upon basis states, namely,

$$\begin{aligned} q_0|l, m, K\rangle = &\sum_{l'}\left(\frac{2l+1}{2l'+1}\right)\frac{\langle l\,2\,m\,0|l'\,m\rangle}{c(K,l)}[\langle l\,2\,K\,0|l'\,K\rangle \\ &\times c(K, l')(\mu + 2\lambda + 3 + l'(l'+1)/2 \\ &- l(l+1)/2)|l', m, K\rangle \\ &+ \sum_{\pm}\langle l\,2\,K\ \pm 2|l'\,K \pm 2\rangle \\ &\times(3(\mu \mp K)(\mu \pm K + 2)/2)^{1/2} \\ &\times c(K \pm 2, l')|l', m, K \pm 2\rangle] \quad (\lambda \geqslant \mu). \end{aligned} \tag{2.3}$$

The coefficients $c(K, l)$ can be regarded as normalization coefficients. Hence, they can be absorbed by a redefinition of states as follows:

$$|l, m, K\rangle\rangle = c(K, l)|l, m, K\rangle. \tag{2.4}$$

If $K < 2$ then the second term of (2.3) involves a state with negative K. In Elliott's scheme $|l, m, -K\rangle$ is defined to be identical with $|l, m, K\rangle$, which is consistent with the choice

$$c(-K, l) = (-1)^{l+\lambda+\mu}c(K, l). \tag{2.5}$$

[a] Senior research associate of the National Fund for Scientific Research (N.F.W.O.) (Belgium).
[b] Research assistant of the N.F.W.O. (Belgium).

 0022-2488/85/123109-03$02.50

It follows that

$$|l, m, -K\rangle\rangle = (-1)^{l+\lambda+\mu}|l, m, K\rangle\rangle. \tag{2.6}$$

By combining all the foregoing results it is clear that we can express the action of O_l^0 as

$$O_l^0|l, m=0, K\rangle\rangle = \sum_{K'=K, K\pm 2} \Omega_{K'K}|l, m=0, K'\rangle\rangle. \tag{2.7}$$

Next we substitute into (2.3) the algebraic closed expressions of the Clebsch–Gordan coefficients. Then the matrix elements $\Omega_{K'K}$ are brought into the following simple form:

$$\begin{aligned}\Omega_{KK} &= \sqrt{6}(2\lambda+\mu+3)[l(l+1)-3K^2],\\ \Omega_{K\pm 2K} &= -3[3(\mu\mp K)(\mu\pm K+2)(l\pm K+2)\\ &\quad\times(l\pm K+1)(l\mp K)(l\mp K-1)/2]^{1/2}.\end{aligned} \tag{2.8}$$

The calculation of the O_l^0 eigenvalues therefore necessitates the diagonalization of a tridiagonal matrix of which the dimension coincides with the degree of l degeneracy.

As an example, let us consider the case of even $\lambda \geq 2$ and $\mu = 2$. From (2.2) we learn that $K=0$ for $l=0$, that $K=2$ for $l=3, 5, ..., \lambda-1, \lambda+1, \lambda+2$, whereas for $l=2, 3, ..., \lambda$ either $K=0$ or $K=2$. Hence, for $l=3, 5, ..., \lambda+1, \lambda+2$ the O_l^0 eigenvalue is determined by $\sqrt{6}(2\lambda+5)(l-3)(l+4)$. Also $O_0^0|0,0,0,\rangle\rangle = 0$. For $l=2, 4, ..., \lambda$ we need to diagonalize the matrix

$$\begin{pmatrix}\Omega_{00} & \Omega_{20}+\Omega_{-20}\\ \Omega_{02} & \Omega_{22}\end{pmatrix}.$$

The corresponding O_l^0 eigenvalues are $\sqrt{6}\{(2\lambda+5)(l-2)(l+3) \pm 6[l(l+1)(l-2)(l+3)+(2\lambda+5)^2]^{1/2}\}$.

These and other results are in complete agreement with the eigenvalues obtained previously by different methods.[1-5]

III. SPECTRUM OF THE Q_l^0 SHIFT OPERATOR

The SO(3) scalar shift operator Q_l^0 is quadratic in the q's and also quadratic in the l generators. Its l-diagonal matrix elements can be related to the matrix elements of coupled q generators. Indeed, on account of the explicit form of Q_l^0 and of the associated higher-order shift operators Q_l^k $(k=-2,-1,1,2)$, one can prove that[6]

$$\begin{aligned}&\langle l, m'=0, \alpha'|Q_l^0|l, m=0, \alpha\rangle\\ &\quad= -\sqrt{14}(2l+3)(2l-1)\\ &\qquad\times\langle l, m'=0, \alpha'|[q\times q]_0^2|l, m=0, \alpha\rangle\\ &\qquad-81l(l+1),\end{aligned} \tag{3.1}$$

whereby

$$[q\times q]_0^2 = \sum_\sigma \langle 2\,2\,\sigma\,-\sigma|2\,0\rangle q_\sigma q_{-\sigma}. \tag{3.2}$$

In (2.3) we have presented Elliott's formula expressing the action of q_0 on Elliott basis states. It is straightforward to deduce from it analogous formulas for the action of the other q generators on the same states. Applying such a formula twice one then arrives at

$$\begin{aligned}[q\times q]_0^2|l, 0, K\rangle\rangle &= \sum_\sigma \sum_{l'} \sum_{l''} \left(\frac{2l+1}{2l''+1}\right)\langle 2\,2\,\sigma\,-\sigma|2\,0\rangle\,\langle l\,2\,0\,-\sigma|l'\,-\sigma\rangle\\ &\times\langle l'\,2\,-\sigma\,\sigma|l''\,0\rangle\Big\{[\langle l\,2\,K\,0|l'\,K\rangle(2\lambda+\mu+3+l'(l'+1)/2-l(l+1)/2)]\\ &\times\Big[\langle l'\,2\,K\,0|l''\,K\rangle(2\lambda+\mu+3+l''(l''+1)/2-l'(l'+1)/2)|l'', 0, K\rangle\rangle\\ &+\sum_\pm\langle l'\,2\,K\,\pm 2|l''\,K\pm 2\rangle(3(\mu\mp K)(\mu\pm K+2)/2)^{1/2}|l'', 0, K\pm 2\rangle\rangle\Big]\\ &+\langle l\,2\,K\,2|l'\,K+2\rangle(3(\mu-K)(\mu+K+2)/2)^{1/2}\Big[\langle l'\,2\,K+2\,0|l''\,K+2\rangle\\ &\times(2\lambda+\mu+3+l''(l''+1)/2-l'(l'+1)/2)|l'', 0, K+2\rangle\rangle\\ &+\sum_\pm\langle l'\,2\,K+2\,\pm 2|l''\,K+2\pm 2\rangle(3(\mu\mp K\mp 2)(\mu+2\pm K\pm 2)/2)^{1/2}|l'', 0, K+2\pm 2\rangle\rangle\Big]\\ &+\langle l\,2\,K\,-2|l'\,K-2\rangle(3(\mu+K)(\mu-K+2)/2)^{1/2}\Big[\langle l'\,2\,K-2\,0|l''\,K-2\rangle\\ &\times(2\lambda+\mu+3+l''(l''+1)/2-l'(l'+1)/2)|l'', 0, L-2\rangle\rangle\\ &+\sum_\pm\langle l'\,2\,K-2\,\pm 2|l''\,K-2\pm 2\rangle(3(\mu\mp K\pm 2)(\mu+2\pm K\mp 2)/2)^{1/2}|l'', 0, K-2\pm 2\rangle\rangle\Big]\Big\}.\end{aligned} \tag{3.3}$$

If one further defines matrix elements $\Lambda_{K'K}$ by

$$Q_l^0|l, m=0, K\rangle\rangle = \sum_{K'=K, K\pm 2, K\pm 4} \Lambda_{K'K}|l, m=0, K'\rangle\rangle, \tag{3.4}$$

then these can be explicitly calculated by substituting into (3.3) the closed expressions for the Clebsch–Gordan coefficients, which we obtain by standard techniques.[10] By doing so it turns out that $\Lambda_{K\pm 4, K}$ is zero except for $K=\mp 2$, in which case we can, on account of the phase convention (2.6), reabsorb the contribution into the Λ_{KK} element. In fact, this property is not that remarkable if one realizes that the q dependence of Q_l^0 is contained into a SO(3) tensor of rank 2.

The final simplified forms of the $\Lambda_{K'K}$ elements, obtained after lengthy calculations, are

$$
\begin{aligned}
\Lambda_{KK} &= 2(2\lambda+\mu+3)^2[l(l+1)-3K^2] \\
&\quad -18K^4+6K^2[5l(l+1)-3] \\
&\quad -12l^2(l+1)^2-72l(l+1) \\
&\quad -3(\mu-K)(\mu+K+2)[l(l+1)-3K^2] \\
&\quad -3(\mu+K)(\mu-K+2)[l(l+1)-3K^2], \\
\Lambda_{K\pm 2K} &= 6[(\mu\mp K)(\mu\pm K+2)(l\pm K+2)(l\pm K+1) \\
&\quad \times(l\mp K)(l\mp K-1)]^{1/2}(2\lambda+\mu\mp 3K), \\
\Lambda_{K\pm 4K} &= 0.
\end{aligned}
\tag{3.5}
$$

Considering again the case of even $\lambda \geqslant 2$ and $\mu = 2$ we readily obtain that the Q_0^0 eigenvalue is zero, that for $l = 3, 5, ..., \lambda+1, \lambda+2$ it is given by the formula $2(l-3)(l+4)(4\lambda^2+20\lambda-12)-(12l^4+24l^3-86l^2-98l+960)$, whereas for $l = 2, 4, ..., \lambda$ the diagonalization of the two-dimensional Λ matrix leads to Q_l^0 eigenvalues expressed as

$$
\begin{aligned}
&2[(l-2)(l+3)(2\lambda+5)^2-6(l^2+l+1)(l^2+l+3)] \\
&\quad \pm 12[(2\lambda+5)^4+2(l^4+2l^3-9l^2-10l+3)(2\lambda+5)^2 \\
&\quad +9(2l+1)^2]^{1/2}.
\end{aligned}
$$

These results, which to our knowledge have never been obtained so far by any other method, are of particular interest to the extension of the interacting boson model for nuclei in the rotational limit, which we actually propose.[9]

[1] B. R. Judd, W. Miller, Jr., J. Patera, and P. Winternitz, J. Math. Phys. **15**, 1787 (1974), and references therein.

[2] M. Moshinsky, J. Patera, R. T. Sharp, and P. Winternitz, Ann. Phys. **95**, 139 (1975), and references therein.

[3] A. Partensky and C. Quesne, J. Math. Phys. **20**, 2014 (1979).

[4] J. W. B. Hughes, J. Phys. A **6**, 48, 281 (1973).

[5] H. E. De Meyer, G. Vanden Berghe, and J. W. B. Hughes, J. Math. Phys. **22**, 2360, 2366 (1981).

[6] J. Van der Jeugt, H. E. De Meyer, and G. Vanden Berghe, J. Math. Phys. **24**, 1025 (1983).

[7] J. P. Elliott, Proc. R. Soc. London Ser. A **245**, 128, 562 (1958).

[8] F. Iachello, in *Nuclear Structure*, edited by K. Abrahams, K. Allaart, and A. E. L. Dieperink (Plenum, New York, 1980).

[9] G. Vanden Berghe, H. E. De Meyer, and P. Van Isacker, "Symmetry-conserving higher-order interaction terms in the interacting boson model," Phys. Rev. C (in press).

[10] A. P. Yutsis, I. B. Levinson, and V. V. Vanagas, *Mathematical Apparatus of the Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962).

On Komar integrals in asymptotically anti-de Sitter space-times

Anne Magnon[a)]
Physics Department, Syracuse University, Syracuse, New York 13210 and Département de Mathématiques, Université de Clermont-Ferrand, 63170 Aubière, France

(Received 23 April 1985; accepted for publication 25 June 1985)

Recently, boundary conditions governing the asymptotic behavior of the gravitational field in the presence of a negative cosmological constant have been introduced using Penrose's conformal techniques. The subsequent analysis has led to expressions of conserved quantities (associated with asymptotic symmetries) involving asymptotic Weyl curvature. On the other hand, if the underlying space-time is equipped with isometries, a generalization of the Komar integral which incorporates the cosmological constant is also available. Thus, in the presence of an isometry, one is faced with two apparently unrelated definitions. It is shown that these definitions agree. This coherence supports the choice of boundary conditions for asymptotically anti-de Sitter space-times and reinforces the definitions of conserved quantities.

I. INTRODUCTION

The current observed value of the cosmological constant Λ is very small (<0.003 eV). On the theoretical side, however, no "explanation" is available for this experimental result. In fact, many models—particularly in supergravity—naturally predict a very large value for Λ. This dichotomy between theory and observation has prompted several detailed investigations of Einstein's equation with cosmological constants. It was hoped, for example, that an analysis of stability of the ground state might show that theories with a nonzero value of Λ have intrinsic instabilities and are therefore not realized in nature.

The central idea in these investigations is that, in the presence of a cosmological constant, de Sitter space (if $\Lambda>0$) or anti-de Sitter space (if $\Lambda<0$) replace Minkowski space as the ground state of the theory and that physically interesting states are represented by space-times which are asymptotically de Sitter or anti-de Sitter. As in the asymptotically Minkowskian context, one expects that the asymptotic symmetry groups for such space-times would be the isometry group of the de Sitter space [O(4,1)] or of anti-de Sitter space [O(2,3)]. One is therefore led to the problem of introducing precise definitions of the boundary conditions which capture these ideas and of investigating the structure that results from these conditions. Now, since the space-like sections in the de Sitter space are compact, the asymptotically de Sitter space-times admit only timelike infinity. The analysis of conserved quantities in these space-times is therefore physically uninteresting. In the anti-de Sitter case, on the other hand, the spacelike sections are noncompact and, as in the asymptotically Minkowskian case, there is a rich asymptotic structure. Therefore, much of the literature is focused on asymptotically anti-de Sitter spaces. In this paper, we shall restrict ourselves to this case. Our purpose here is to show that, in the presence of isometries, the conserved quantities defined at infinity using the asymptotic Weyl curvature reduce to (certain multiples of) the appropriately generalized Komar integrals. This reduction strengthens one's faith in the choice of boundary conditions as well as in the definitions of conserved quantities thereby putting many of the previous analyses on a sounder footing.

Section II is devoted to preliminaries. We recall[1,2] the definitions of asymptotically anti-de Sitter spaces, first in terms of a conformal completion (Sec. II A), then in terms of the behavior of metric components in suitable charts, and list some of the consequences of these boundary conditions (Sec. II B). Recently, the two sets of conditions have been shown to be equivalent.[3] In Sec. III we display the generalization of the Komar integral which incorporates the presence of a nonzero cosmological constant. Section IV is devoted to axisymmetric space-times, Sec. V to static space-times, and Sec. VI to stationary space-times. In all cases, we show that the conserved quantity associated with the asymptotic symmetry corresponding to the isometry, defined in terms of the asymptotic Weyl curvature, is a multiple of the generalized Komar integral. More precisely, we show that the generalized Komar integral yields the correct value for the "angular momentum" but $\frac{1}{2}$ the correct value of "energy." *A priori*, one would not have expected such a simple relationship to hold. For, whereas the conserved quantity at infinity depends sensitively on the choice of the boundary conditions, the generalized Komar integral can be evaluated anywhere in the interior (outside sources). In five-dimensional Kaluza–Klein theories, for example, the boundary conditions are such that the Komar integral for "energy" can vanish identically even when the asymptotic "energy" is nonzero. The fact that we can obtain the same relation as in the asymptotically Minkowskian case reinforces the belief that the boundary conditions used here are natural generalizations of the standard asymptotically flat ones.

II. BOUNDARY CONDITIONS FOR ASYMPTOTICALLY ANTI-DE SITTER SPACE-TIMES

A. Covariant formulation

An analysis of the structure of asymptotically anti-de Sitter space-times based on Penrose's conformal treatment[1] has been proposed.[2] We shall summarize it briefly. (We assume that all manifolds and fields are C^∞.)

[a)] Détachée du Ministère des Relations Extérieures, Paris, France.

 0022-2488/85/123112-06$02.50

Definition 1: A space-time (M, g_{ab}) will be said to be weakly asymptotically anti-de Sitter if there exists a manifold $\hat{M}$ with boundary $\partial\hat{M}$, equipped with a metric $\hat{g}_{ab}$ and a diffeomorphism from M onto $\hat{M} - \partial\hat{M}$ such that (i) there exists a function Ω on $\hat{M}$ such that $\hat{g}_{ab} = \Omega^2 g_{ab}$ on M, (ii) $I \equiv \partial\hat{M}$ is topologically $S^2 \times R$, and on I, $\Omega = 0$, and (iii) g_{ab} satisfies $R_{ab} - \frac{1}{2} R\, g_{ab} = 8\pi T_{ab} - \Lambda g_{ab}$, with $\Lambda < 0$, where $\Omega^{-3} T_a{}^b$ admits a smooth limit to I.

Remarks: (i) Schwarszchild anti-de Sitter and Kerr anti-de Sitter space-times satisfy Definition 1 if one sets $\Omega = 1/r$. (ii) Let $\hat{n}_a = \nabla_a \Omega$. It is easy to check[2] that $\hat{n}^a \hat{n}_a$ reduces to $-\Lambda/3$ on I. Hence I is a time-like cylinder.[1] When $\Lambda = 0$, the above definition reduces to that of asymptotic flatness at null infinity. (iii) Given a conformal completion $(\hat{M}, \hat{g}_{ab})$ which satisfies Definition 1, $(\hat{M}\, \tilde{g}_{ab})$ also satisfies it if and only if $\tilde{g}_{ab} = \omega^2 \hat{g}_{ab}$, where ω is a smooth nowhere vanishing function on $\hat{M}$. Using this conformal freedom, it is easy to set $\nabla_a \hat{n}_b \hat{=} 0$, where from now on, $\hat{=}$ will denote "equals at points of I to." It follows that the conformal freedom is restricted by $\Omega \to \omega\Omega$, where $\pounds_n \omega \hat{=} 0$. Note that, unlike in the asymptotically Minkowskian context, and since $\hat{n}^a$ is a spacelike normal to I, this last condition does not restrict ω to be a function of θ and φ only, on I. It has been proven in Ref. 2 that

$$D_a \hat{E}^{ab} \hat{=} -\tfrac{1}{2} \lim_{\to I} \Omega^{-4} \hat{T}_a{}^c \hat{n}^a \hat{q}^b{}_c, \tag{2.1}$$

where $\hat{E}_{ab} = (-3/\Lambda)\Omega^{-1} \hat{C}_{ambn} \hat{n}^m \hat{n}^n$ is the electric part of the asymptotic Weyl curvature and D is the intrinsic derivative operator on $(I, \hat{q}_{ab})$ induced by $\hat{g}_{ab}$. (iv) In the asymptotically Minkowskian context, the space of generators of I, being topologically S^2, admits a unique conformal class of metrics. Since I is ruled by null generators (the integral curves of $\hat{n}_a$) the pullback of these metrics yield, on I, a unique class of degenerate conformal metrics. In the present case, since $\hat{n}^a$ is not tangential to I, I does not have a fiber bundle structure and $S^2 \times R$ admits many conformal structures; hence the asymptotic symmetry group is the diffeomorphism group of I. The reduction of this group can be obtained by strengthening the asymptotic conditions. It is natural to require that the three-manifold I admits, as its conformal group, the anti-de Sitter group. This is achieved by imposing that the Bach tensor

$$\hat{\beta}_{abc} \hat{=} \hat{D}_{[a} (\hat{R}_{b]c} - \tfrac{1}{4} \hat{R} \hat{q}_{b]c}) \tag{2.2}$$

of $\hat{q}_{ab}$ vanishes, i.e., that I be conformally flat. This leads to the following definition.

Definition 2: A space-time (M, g_{ab}) will be said to be asymptotically anti-de Sitter if, in addition to Definition 1, one has the condition that I be conformally flat [or equivalently that the conformal group of $(I, \hat{q}_{ab})$ is O(2,3) or the anti-de Sitter group (or covering group)]. For this class of space-times, there are[2] ten "conserved quantities"

$$Q_{\xi}(C) = -(8\pi)^{-1} \left(\frac{-3}{\Lambda}\right)^{1/2} \oint_C \hat{E}_{ab} \hat{\xi}^a\, dS^b, \tag{2.3}$$

where ξ^a is a conformal Killing field on $(I, \hat{q}_{ab})$ and C is a cross section of I. The fluxes F_{ξ} of these quantities are completely determined by matter terms:

$$F_{\xi}(\Delta) = (16\pi)^{-1} \int_{\Delta} \lim (\Omega^{-4} \hat{T}_a{}^c \underset{\sim}{n}{}^a \hat{q}_c{}^b) \hat{\xi}_b\, d^3 I, \tag{2.4}$$

where Δ is a three-dimensional region of I, and $\underset{\sim}{n}{}^a$ is the unit normal to I with respect to $\hat{g}_{ab}$.

B. A reformulation using asymptotic charts

The above boundary conditions can be reformulated using admissible charts in the neighborhood of infinity. In Sec. II A, a conformal rescaling of the physical metric was introduced,

$$d\hat{s}^2 = \Omega^2\, ds^2, \tag{2.5}$$

which brings I, the surface at infinity defined by $\Omega = 0$, to a finite distance. The available conformal freedom enables the choice of a chart in which the induced metric on I is given by

$$d\hat{s}^2_I = -dt^2 + d\theta^2 + \sin^2\theta\, d\varphi^2 \equiv \bar{g}_{ab}\, dy^a\, dy^b. \tag{2.6}$$

This chart is then extended off I in such a way that in the coordinates $\Omega, t, \theta, \varphi$, the metric is given by

$$d\hat{s}^2 = d\Omega^2 + \hat{g}_{ab}(\Omega, y) dy^a\, dy^b. \tag{2.7}$$

Analytic solutions of Einstein's equation in the neighborhood of I have been investigated in Ref. 3. The result is the following. Let

$$\hat{g}_{ab}(\Omega, y) = \sum_{n>0} \frac{1}{n!} \hat{g}_{ab}{}^{(n)}(y) \Omega^n. \tag{2.8}$$

It is shown that

$$\begin{aligned} &\hat{g}^{(0)}_{tt} = -1, \quad \hat{g}^{(0)}_{\theta\theta} = 1, \qquad \hat{g}^{(0)}_{\varphi\varphi} = \sin^2\theta, \\ &g^{(1)}_{ab} = 0, \\ &\hat{g}^{(2)}_{tt} = -1, \quad \hat{g}^{(2)}_{\theta\theta} = -1, \quad \hat{g}^{(2)}_{\varphi\varphi} = -\sin^2\theta. \end{aligned} \tag{2.9}$$

The other three components being 0. Furthermore, there are as many analytic solutions to the Einstein's equation as there are coefficients $\hat{g}^{(3)}_{ab}(y) = \hat{E}_{ab}$, which are traceless $(\hat{E}^{ab} \hat{g}^{(0)}_{ab} = 0)$ and transverse $(D_b \hat{E}^b{}_a = 0)$ fields on I. In other words, $\hat{E}_{ab}$, the electric component of the asymptotic Weyl curvature, governs the existence of solutions to the field equation. A further transformation of the radial coordinate,

$$r = \Omega^{-1} - \tfrac{1}{4}\Omega, \tag{2.10}$$

gives the physical metric ds^2 as a deviation from the anti-de Sitter background (with metric ds^2_0) compatible with the chart on I:

$$ds^2 = ds^2_0 + h_{\lambda\mu}\, dx^{\lambda}\, dx^{\mu}. \tag{2.11}$$

These results will be useful for the next sections. In these admissible charts it has been checked that (i) The Kerr anti-de Sitter metric is a prototype of these asymptotically anti-de Sitter space-times, (ii) The boundary conditions are invariant under the anti-de Sitter group O(2,3), and (iii) The charge integrals, generators of O(2,3), are finite. If the admissible chart is that in which the anti-de Sitter metric reads

$$ds^2_0 = -[1 + r^2/R^2] dt^2 + [1 + r^2/R^2]^{-1} dr^2 + r^2 d\omega^2, \tag{2.12}$$

where R, the radius of curvature, is related to the cosmological constant Λ by $R = (3/-\Lambda)^{1/2}$, and $d\omega^2$ is the usual spherical element in coordinates θ, φ, it can be shown that if one considers deviations $h_{\lambda\mu}$, more precisely metrics ds^2

$= ds_0^2 + h_{\lambda\mu}\, dx^\lambda\, dx^\mu$ which behave asymptotically as r^{-1}, the invariance under the O(2,3) group is not fulfilled. The appropriate falloff requires that $h_{r\mu}$ should be of the order r^{-4} and h_{rr} of the order r^{-5}, the other components behaving as r^{-1}.

III. GENERALIZATION OF KOMAR'S INTEGRALS TO SPACE-TIMES WITH A NONZERO COSMOLOGICAL CONSTANT

Recall that in the case $\Lambda = 0$, if the space-time (M, g_{ab}) describing the gravitational field admits a Killing field ξ^a, the Komar integral[4]

$$Q_k(\xi) = \frac{1}{16\pi}\int_{S_2} \epsilon_{abcd}\nabla^c\xi^d\, dS^{ab} \tag{3.1}$$

represents a conserved quantity: This integral does not depend on the particular choice of the two-sphere S_2 surrounding the matter sources if these sources have compact support. [Here ϵ_{abcd} and ∇_a are, respectively, the alternating tensor and the derivative operator on (M, g_{ab}).] If the matter sources are allowed to go to infinity, I, with the falloff described in Sec. II, $Q_k(\xi)$ will be evaluated on a particular S_2, now a cross section of I. However, for simplicity we can restrict ourselves to an isolated system and consider matter with compact spatial support. If ξ^a is a stationary Killing field, Q represents the total energy of the system, while if ξ^a is an axial Killing field, it has the interpretation of the component of the total angular momentum along the responding axis.

We now show how the Komar integral can be extended to space-times with a nonvanishing cosmological constant. In this case Einstein's equation is given by

$$R_{ab} - \tfrac{1}{2}Rg_{ab} = 8\pi T_{ab} - \Lambda g_{ab}. \tag{3.2}$$

Let the space-time (M, g_{ab}) furthermore admit a Killing vector field ξ^a. It follows from the affine colineation equation $\nabla_a\nabla_b\xi_c = R^{\ m}{}_{abc}\xi_m$, and the field equation

$$R_{ab} = 8\pi T_{ab} - 4\pi T g_{ab} + \Lambda g_{ab}, \tag{3.3}$$

that $\nabla^a\nabla_a\xi_b = 8\pi\xi^m[-T_{mb} + \frac{1}{2}g_{mb}T] - \Lambda\xi_b$. Hence the current $J_b = \nabla^a\nabla_a\xi_b$ is conserved: $\nabla^b J_b = 0$. A straightforward calculation making use of Stokes' theorem implies the following equality:

$$\frac{1}{16\pi}\int_{\partial\Sigma}(\nabla_a\xi_b)\epsilon^{ab}{}_{cd}\, dS^{cd} = \int_\Sigma \xi^m\left[-T_{mb} + \frac{1}{2}g_{mb}T\right]d\Sigma^b - \frac{\Lambda}{8\pi}\int_\Sigma \xi_b\, d\Sigma^b, \tag{3.4}$$

where Σ is a spacelike three-surface which intersects the matter tube, $\partial\Sigma$ being outside the matter. In this equality, each integral is independent of a particular choice of Σ which spans a given $\partial\Sigma$ lying outside sources. The quantity

$$Q_{k_\Lambda}(\xi) = \frac{1}{16\pi}\int_{\partial\Sigma}(\nabla_a\xi_b)\epsilon^{ab}{}_{cd}\, dS^{cd} + \frac{\Lambda}{8\pi}\int_\Sigma \xi_b\, d\Sigma^b \tag{3.5}$$

is to be considered as a generalization of the Komar integral. One can imagine using the Kerr–anti-de Sitter space-time[5] to evaluate the integral in (3.5) explicitly. Results of such an evaluation were reported in an independent analysis[6] of the generalized Komar integrals. However, in this case, the presence of a horizon requires some care since formula (3.5) has been derived for solutions with matter contents, and since the integral on the right now involves a volume term which is absent in the more familiar $\Lambda = 0$ case.

Hence, when an isometry ξ^a is available on a space-time with nonvanishing Λ, one is faced with two apparently unrelated definitions of the corresponding conserved quantity. The first is provided by the charge integral involving the asymptotic Weyl curvature,[2] the second is $Q_{k_\Lambda}(\xi)$. Since in the generalized Komar integral, the two-sphere S_2 surrounding the matter sources can be located anywhere inside the space-time, one might have expected that it would be local and insensitive to the asymptotic behavior of the gravitational field, i.e., unrelated to the integral at I. If this were the case, the notion of energy would have been ambiguous and the stability arguments[7] based on the behavior of conserved quantities would have inherited these problems. Fortunately, as we shall see in the next sections, the two definitions do agree.

IV. AXISYMMETRIC SPACE-TIMES

In this section, we shall assume that the space-time (M, g_{ab}) is asymptotically anti-de Sitter and admits a rotational Killing vector field ξ^a. The orbits of ξ^a are topologically S^1 lying on a family of nested two-spheres S_2. Since these two spheres generate a surface Σ, which is the Cauchy surface for some diamond-shaped region of the compactified space-time,[1,8] $\int_\Sigma \xi_b\, d\Sigma^b = 0$, and the Komar integral reduces to the surface integral:

$$Q_{k_\Lambda}(\xi) = \lim_{S_2\to I}\frac{1}{16\pi}\int_{S_2}(\nabla_a\xi_b)\epsilon^{ab}{}_{cd}\, dS^{cd} = \lim_{S_2\to I}\frac{1}{16\pi}\int_{S_2}(\nabla_a\xi_b)dS^{*ab}. \tag{4.1}$$

We want to relate $Q_{k_\Lambda}(\xi)$ to the corresponding conserved quantity at I:

$$Q_\xi(C) = -(8\pi)^{-1}\left(\frac{-3}{\Lambda}\right)^{1/2}\oint_C \hat{E}_{ab}\hat{\xi}^a\, d\hat{S}^b, \tag{4.2}$$

where C is the cross section of I corresponding to the boundary $\partial\Sigma$ of Σ, and $\hat{E}_{ab}$ is the electric component of the rescaled Weyl tensor, as described in Sec. II. Let $(\hat{M}, \hat{g}_{ab} = \Omega^2 g_{ab})$ be a conformal completion of (M, g_{ab}) satisfying Definition 2, such that $\pounds_\xi\Omega = 0$. Since $\xi_b = \Omega^{-2}\hat{\xi}_b$ and $\hat{\xi}_b\, dS^{*ab} = 0$, one has

$$16\pi Q_{k_\Lambda}(\xi) = \lim_{\Omega\to 0}\oint_{S_2(\Omega)}\Omega^{-2}(\hat{\nabla}_a\hat{\xi}_b)\hat{\epsilon}^{ab}{}_{cd}\, dS^{cd} = \lim_{\Omega\to 0}\frac{1}{\Omega^2}\oint_{S_2(\Omega)}(\hat{\nabla}_a\hat{\xi}_b)\hat{\epsilon}^{ab}{}_{cd}\, dS^{cd} = \lim_{\Omega\to 0} A(\Omega)/\Omega^2,$$

say, where $S_2(\Omega) = \Sigma\cap\{\Omega = \text{const}\}$. On the other hand, this last integral is equal to

$$16\pi\int_\Sigma \xi^m\left[-T_{mb}+\frac{1}{2}g_{mb}T\right]d\Sigma^b, \tag{4.3}$$

which is finite provided $\Omega^{-3}\hat{T}_a{}^b$ admits a smooth limit to I. As a result, A must vanish when $\Omega\to 0$, and one can make use of l'Hôpital's rule. Let $\hat{\eta}^m$ denote the vector field inducing the unit normal at I, and such that $\hat{\nabla}_{(a}\hat{\eta}_{b)}\hat{=}0$. Thus $\hat{\eta}^m\hat{\nabla}_m\hat{\eta}^a\hat{=}0$. Let furthermore $\hat{\eta}^{[a}\hat{t}^{\,b]}$ be the two-flat orthogonal to the two-spheres $S_2(\Omega)$ ($\hat{t}^{\,a}$ can be chosen such that $\pounds_\eta\hat{t}\hat{=}0$). We shall consider a neighborhood of I sufficiently small so that $A=A(\Omega)$ converges uniformly to zero as Ω goes to zero. We thus have

$$\lim_{\Omega\to 0}A/\Omega^2=\lim_{\Omega\to 0}(\hat{\eta}^m\hat{\nabla}_m A)(2\Omega\hat{\eta}^m\hat{\nabla}_m\Omega)^{-1}$$

$$=(-\Lambda/3)^{-1/2}\lim_{\Omega\to 0}(1/2\Omega)\,\hat{\eta}^m\hat{\nabla}_m A,$$

where

$$\lim_{\Omega\to 0}\hat{\eta}^m\hat{\nabla}_m A=\lim_{\Omega\to 0}\oint_{S_2(\Omega)}(\hat{\eta}^m\hat{\nabla}_m\hat{\nabla}_a\hat{\xi}_b)dS^{*ab}.$$

We have used the equality $\hat{\eta}^m\hat{\nabla}_m(\hat{t}^{\,[a}\hat{\eta}^{\,b]})\hat{=}0$. Hence $\lim_{\Omega\to 0}A/\Omega^2$ is equal to

$$\lim_{\Omega\to 0}\oint_{S_2(\Omega)}\hat{\eta}^m\hat{R}^{\,p}{}_{mab}\hat{\xi}_p\left[2\Omega\left(-\frac{\Lambda}{3}\right)^{1/2}\right]^{-1}dS^{*ab}, \tag{4.4}$$

where we have used the fact that $\pounds_\xi\Omega=0$ implies that $\hat{\xi}^{\,a}$ is a Killing vector field for $\hat{g}_{ab}$ and thus satisfies the affine colineation equation $\hat{\nabla}_a\hat{\nabla}_b\hat{\xi}_c=\hat{R}_{cba}{}^p\hat{\xi}_p$. On the other hand, one has

$$\hat{R}_{abcd}=\hat{C}_{abcd}+\hat{g}_{a[c}\hat{S}_{d]b}-\hat{g}_{b[c}\hat{S}_{d]a},$$

where $\hat{S}_{ab}=(\hat{R}_{ab}-\frac{1}{6}\hat{R}\hat{g}_{ab})$. Hence (4.4) splits into three integrals:

$$A_1=\lim_{\Omega\to 0}\oint_{S_2(\Omega)}\hat{\eta}^b\hat{C}_{abcd}\hat{\xi}^{\,a}\left[2\Omega\left(-\frac{\Lambda}{3}\right)^{1/2}\right]^{-1}dS^{*cd}, \tag{4.5}$$

$$A_2=\lim_{\Omega\to 0}\oint_{S_2(\Omega)}\hat{\eta}^b\hat{g}_{a[c}\hat{S}_{d]b}\hat{\xi}^{\,a}\left[2\Omega\left(-\frac{\Lambda}{3}\right)^{1/2}\right]^{-1}dS^{*cd}, \tag{4.6}$$

$$A_3=\lim_{\Omega\to 0}-\oint_{S_2(\Omega)}\hat{\eta}^b\hat{g}_{b[c}\hat{S}_{d]a}\hat{\xi}^{\,a}\left[2\Omega\left(-\frac{\Lambda}{3}\right)^{1/2}\right]^{-1}dS^{*cd}. \tag{4.7}$$

The definition of $\hat{E}_{ab}$, the electric part of the asymptotic Weyl curvature, implies that $A_1=-(-3/\Lambda)^{1/2}\times\oint_c\hat{E}_{ab}\hat{\xi}^{\,a}\,dS^b$. The integral A_2 vanishes due to the fact that $\hat{\xi}\cdot\hat{t}=0$ and $\hat{\xi}\cdot\hat{\eta}=0$. Finally A_3 can be expressed using

$$\hat{S}_{ab}=S_{ab}-2\Omega^{-1}\hat{\nabla}_a\hat{\nabla}_b\Omega+\hat{g}_{ab}\Omega^{-2}\hat{\nabla}^m\Omega\hat{\nabla}_m\Omega.$$

Note first that since, by field equations, S_{ab} is a multiple of g_{ab} and since the vectors $\hat{\xi},\hat{\eta},\hat{t}$ are orthogonal, the only terms which contribute to the integral in A_3 are terms in $\hat{\nabla}_c\hat{\nabla}_a\Omega$ and $\hat{\nabla}_d\hat{\nabla}_a\Omega$, respectively. A straightforward calculation reduces A_3 to

$$\lim_{\Omega\to 0}\frac{1}{2}\oint_{S_2(\Omega)}\Omega^{-2}(\hat{\nabla}_a\hat{\xi}_b)\epsilon^{ab}{}_{mn}\,dS^{mn},$$

i.e., to $8\pi\,Q_k(\xi)$. The result follows immediately

$$Q_{k_\Lambda}(\xi)=Q_\xi(C).$$

Hence we have the following theorem.

Theorem: If an asymptotically anti-de Sitter space-time (in which matter sources are spatially compact) admits a rotational Killing vector field ξ^a, the Komar integral $Q_{k_\Lambda}(\xi)$ and the charge integral $Q_\xi(C)$ at I are related via $Q_{k_\Lambda}(\xi)=Q_\xi(C)$, where C is any cross section of I.

V. STATIC SPACE-TIMES

We shall now assume that the space-time (M,g_{ab}) is static, i.e., is equipped with an everywhere timelike hypersurface orthogonal Killing vector field ξ^a, and is asymptotically anti-de Sitter, with conformal completion $(\hat{M},\hat{g}_{ab})$. The static foliation will be denoted by $\mathcal{S}_t$. On the same manifold M, an anti-de Sitter metric $\mathring{g}_{ab}$ can be introduced with the same static hypersurfaces $\mathcal{S}_t$ and hypersurface orthogonal Killing vector field $\mathring{\xi}^{\,a}=\xi^a$. In addition we shall require that $(M,\mathring{g}_{ab})$ and (M,g_{ab}) belong to the same O(2,3) invariant-equivalence class.[9] According to Sec. II, this is equivalent to requiring that, in the usual anti-de Sitter chart (t,r,θ,φ) associated with $\mathring{g}_{ab}$, the components of $g_{ab}-\mathring{g}_{ab}$ are of order r^{-1} with the exception of $g_{tr}-\mathring{g}_{tr}$, $g_{r\theta}-\mathring{g}_{r\theta}$, and $g_{r\varphi}-\mathring{g}_{r\varphi}$, which are of order r^{-4}, and $g_{rr}-\mathring{g}_{rr}$, which is of order r^{-5}.

The generalized Komar integrals associated with $\xi^a\equiv\mathring{\xi}^{\,a}$ in (M,g_{ab}) and $(M,\mathring{g}_{ab})$ are, respectively,

$$Q_{k_\Lambda}(\xi)=\lim_{\partial\Sigma\to I}\frac{1}{16\pi}\int_{\partial\Sigma}(\nabla_a\xi_b)\epsilon^{ab}{}_{cd}\,dS^{cd}+\frac{\Lambda}{8\pi}\int_\Sigma\xi_b\,d\Sigma^b \tag{5.1a}$$

and

$$Q_{k_\Lambda}(\mathring{\xi})=\lim_{\partial\Sigma\to I}\frac{1}{16\pi}\int_{\partial\Sigma}(\mathring{\nabla}_a\mathring{\xi}_b)\mathring{\epsilon}^{ab}{}_{cd}\,dS^{cd}+\frac{\Lambda}{8\pi}\int_\Sigma\mathring{\xi}_b\,d\Sigma^b, \tag{5.1b}$$

where $\Sigma(\Sigma\subset\mathcal{S}_t$, for some value of $t)$ is a compact volume spanning $\partial\Sigma$, a two-sphere surrounding the matter sources. Note that, in the expression of Q_{k_Λ}, the volume integral as well as the surface integral, are both infinite in the limit. The first step of our analysis will consist in the elimination of the volume integral in $Q_{k_\Lambda}(\xi)$. In these volume integrals, the integrands are, respectively, $\xi^{\,b}\epsilon_{bijk}\equiv F_{ijk}$ and $\mathring{\xi}^{\,b}\mathring{\epsilon}_{bijk}\equiv\mathring{F}_{ijk}$, which, being nondegenerate three-forms on Σ, are proportional to one another. Thus, $F_{abc}=\alpha(r,\theta,\varphi)\mathring{F}_{abc}$ for some nowhere vanishing, smooth function α. Now, we can perform a diffeomorphism ψ on $(M,\mathring{g}_{ab})$ such that (a) the boundary conditions [i.e., the O(2,3) equivalence class of asymptotically anti-de Sitter space-times] are preserved, (b) the vector field $\xi^a\equiv\mathring{\xi}^{\,a}$ is mapped to itself, and (c) $(M,\psi(\mathring{g}_{ab})=\mathring{\mathring{g}}_{ab})$ is an anti-de Sitter space-time with $F_{abc}=\mathring{F}_{abc}$. (Recall that under any diffeomorphism ψ an anti-de Sitter space-time is transformed into another anti-de Sitter space-time since $0=\psi[R_{ab}-\frac{1}{2}Rg_{ab}+\Lambda g_{ab}]=\psi(R_{ab})-\frac{1}{2}\psi(R)\times\psi(g_{ab})+\Lambda\psi(g_{ab})\equiv\tilde{R}_{ab}-\frac{1}{2}\tilde{R}\tilde{g}_{ab}+\Lambda\tilde{g}_{ab})$. Thus we have a new anti-de Sitter background $(M,\mathring{\mathring{g}}_{ab})$ for which the volume integral in $Q_{k_\Lambda}(\xi)$ coincides with that in $Q_{k_\Lambda}(\mathring{\xi})$. Subtracting $Q_{k_\Lambda}(\mathring{\xi})$ from $Q_{k_\Lambda}(\xi)$ and using the fact that all generalized Komar integrals vanish identically in any anti-de Sitter

space, we have

$$Q_{k_\Lambda}(\xi) = Q_{k_\Lambda}(\xi) - Q_{k_\Lambda}(\mathring{\xi})$$
$$= \frac{1}{16\pi} \lim_{\partial\Sigma \to I} \int_{\partial\Sigma} [(\nabla_a \xi_b)\epsilon^{ab}{}_{cd} - (\mathring{\nabla}_a \mathring{\xi}_b)\mathring{\epsilon}^{ab}{}_{cd}]\, dS^{cd}. \tag{5.2}$$

From now on, we shall drop the tilde on quantities associated with $\mathring{g}_{ab}$.

Let $g_{ab}\xi^a\xi^b = -\lambda$ (resp. $\mathring{g}_{ab}\mathring{\xi}^a\mathring{\xi}^b = -\lambda_0$) and $\hat{g}_{ab} = \Omega^2 g_{ab}$ (resp. $\mathring{\hat{g}}_{ab} = \mathring{\Omega}^2 \mathring{g}_{ab}$), where, for further convenience, Ω (resp. $\mathring{\Omega}$) will be chosen equal to $\lambda^{-1/2}$ (resp. $\mathring{\lambda}^{-1/2}$). Then, $\hat{\xi}^a \equiv \xi^a$ (resp. $\mathring{\hat{\xi}}^a \equiv \mathring{\xi}^a$) is a Killing field also of $\hat{g}_{ab}$ (resp. $\mathring{\hat{g}}_{ab}$). Furthermore, it has unit norm: $\hat{\lambda} = \hat{g}_{ab}\hat{\xi}^a\hat{\xi}^b = -1$ (and $\mathring{\hat{\lambda}} = -1$). Therefore, using the fact that $\hat{\xi}^a \equiv \mathring{\hat{\xi}}^a$ is hypersurface orthogonal, we have $\hat{\nabla}_a\hat{\xi}_b = \hat{\lambda}^{-1} \times \hat{\xi}_{[a}\hat{\nabla}_{b]}\hat{\lambda} = 0$ and similarly $\mathring{\hat{\nabla}}_a\mathring{\hat{\xi}}_b = 0$. Now, since $\nabla_{[a}\xi_{b]} = \hat{\nabla}_{[a}\Omega^{-2}\hat{\xi}_{b]}$, the expression $Q_{k_\Lambda}(\xi)$ in (5.2) reduces to

$$Q_{k_\Lambda}(\xi) = -\frac{1}{8\pi} \lim_{\Omega \to 0} \Omega^{-3} \int_{\partial\Sigma(\Omega)} [(\hat{\nabla}_a\Omega)\hat{\xi}_b\hat{\epsilon}^{ab}{}_{cd} - \Omega^3\mathring{\Omega}^{-3}(\mathring{\hat{\nabla}}_a\mathring{\Omega})\mathring{\hat{\xi}}_b\mathring{\hat{\epsilon}}^{ab}{}_{cd}]\, dS^{cd}, \tag{5.3}$$

where $\partial\Sigma(\Omega) = S_t \cap \{\Omega = \text{const}\}$. Since we have assumed that the matter has compact support, $Q_{k_\Lambda}(\xi)$ is finite. This implies that, although the integrand of the integral in (5.3) might not go to zero when $\Omega \to 0$, the integral itself does so, and the limit can be evaluated using l'Hôpital's rule. Let $\hat{\eta}^a$ denote the vector field inducing the unit normal to I, and let $h_{ab} = g_{ab} + \lambda^{-1}\xi_a\xi_b$. From now on (since the conformal freedom allows it) we shall assume that $\hat{\nabla}_a\hat{\eta}_b \hat{=} 0$. The action of l'Hôpital's rule on the expression (5.3) yields

$$Q_{k_\Lambda}(\xi) = -\frac{1}{8\pi} \lim_{\Omega \to 0} \int_{\partial\Sigma(\Omega)} \left\{ \left[3\Omega^2 \left(-\frac{\Lambda}{3} \right)^{1/2} \right]^{-1} \times \hat{\eta}^m\hat{\eta}^a(\hat{\nabla}_m\hat{\nabla}_a\Omega) - \left[3\mathring{\Omega}^2 \left(-\frac{\Lambda}{3} \right)^{1/2} \right]^{-1} \mathring{\hat{\eta}}^m\mathring{\hat{\eta}}^a \times (\mathring{\hat{\nabla}}_m\mathring{\hat{\nabla}}_a\mathring{\Omega}) \right\} \left(-\frac{\Lambda}{3} \right)^{1/2} d^2S. \tag{5.4}$$

Let us evaluate the integrand in (5.4). A straightforward calculation gives us the Ricci curvature of (S_t, h_{ab}) in terms of the norm $\lambda \equiv \Omega^{-2}$ and the twist ω_a of the Killing field ξ^a:

$$\mathscr{R}_{ab} = \Lambda h_{ab} + (1/2\lambda)\, h_a{}^p h_b{}^q \nabla_p\nabla_q\lambda - (1/4\lambda^2) D_a\lambda D_b\lambda + (1/2\lambda^2)(\omega_a\omega_b - h_{ab}\omega^m\omega_m), \tag{5.5}$$

where D is the derivative operator compatible with h_{ab}. (The twist terms, of course, vanish in the static case now under consideration. They will be important in the stationary case, to be discussed in Sec. VI.) Hence, we have

$$\hat{\eta}^a\hat{\eta}^b\nabla_a\nabla_b\Omega \equiv \hat{\eta}^a\hat{\eta}^b\nabla_a\nabla_b\lambda^{-1/2} = \lambda^{-1/2}[(1/2\lambda^2)D_a\lambda D_b\lambda + \Lambda h_{ab} - \mathscr{R}_{ab}]\hat{\eta}^a\hat{\eta}^b. \tag{5.6}$$

The idea now is to relate the right side of (5.6) to the electric part of the Weyl tensor. Recall, first, the identity which holds on any time symmetric slice $\mathscr{S}_t$:

$$\mathscr{E}_{ab} \equiv C_{ambn}\zeta^m\zeta^n = \mathscr{R}_{ab} - \tfrac{1}{2}\{h_a{}^m h_b{}^n L_{mn} + h_{ab}L_{pq}h^{pq}\}, \tag{5.7}$$

where ζ^m is unit normal to $\mathscr{S}_t$ and where $L_{ab} = R_{ab} - \frac{1}{6}Rg_{ab}$. ($\mathscr{E}_{ab}$ is the electric part of the Weyl tensor of the physical metric g_{ab}, *relative to* $\mathscr{S}_t$.) Using the field equations outside sources, we have

$$\mathscr{E}_{ab} = \mathscr{R}_{ab} - (2\Lambda/3)h_{ab}, \tag{5.8}$$

hence

$$\hat{\eta}^a\hat{\eta}^b\mathscr{R}_{ab} = \Omega(\Omega^{-1}C_{ambn})\hat{\zeta}^m\hat{\zeta}^n\hat{\eta}^a\hat{\eta}^b + \hat{\eta}^a\hat{\eta}^b((2\Lambda/3)h_{ab}); \tag{5.9}$$

here $(\Omega^{-1}C_{ambn}\,\hat{\eta}^a\hat{\eta}^b)\hat{\zeta}^m\hat{\zeta}^n = E_{mn}\hat{\zeta}^m\hat{\zeta}^n$, where $\hat{E}_{mn}$ is the electric component of the Weyl tensor with respect to the unit normal $\hat{\eta}$ to I, as introduced in (Ref. 2) [$\hat{\zeta}^m$ denotes the unit normal to $(S, \hat{h}_{ab})$]. One immediately deduces from (5.6) and (5.9) that $\hat{\eta}^m\hat{\eta}^n\nabla_m\nabla_n\lambda^{-1/2} = -\Omega^2\hat{E}_{ab}\hat{\zeta}^a\hat{\zeta}^b + \Omega^{-1}(\Lambda/3)\hat{h}_{ab}\hat{\eta}^a\hat{\eta}^b + 2\Omega^{-1}\hat{\eta}^a\hat{\eta}^b\hat{\nabla}_a\Omega\hat{\nabla}_b\Omega$. A straightforward calculation using $\hat{\nabla}_a k_b = \nabla_a k_b + C_{ab}{}^m k_m$, where $C_{ab}{}^m = -\Omega^{-1}[2\delta^m{}_{(a}\hat{\nabla}_{b)}\Omega - \hat{g}_{ab}\hat{\nabla}^m\Omega]$, finally gives us

$$\hat{\eta}^m\hat{\eta}^a\hat{\nabla}_m\hat{\nabla}_a\lambda^{-1/2} = -\Omega^2\hat{E}_{mn}\hat{\zeta}^m\hat{\zeta}^n + \left(\frac{\Lambda}{3}\right)\left(\frac{\hat{h}_{ab}}{\Omega}\right)\hat{\eta}^a\hat{\eta}^b + \frac{1}{\Omega}(\hat{D}^b\Omega)(D_b\Omega), \tag{5.10}$$

where $\hat{D}$ is the derivative operator compatible with $\hat{h}_{ab}$. This last equality enables us to transform the expression (5.4) into

$$Q_{k_\Lambda}(\xi) = -\frac{1}{8\pi}\left(-\frac{3}{\Lambda}\right)^{1/2}\frac{1}{3}\lim_{\Omega\to 0}\oint_{\partial\Sigma(\Omega)}\left\{\left[-\hat{E}_{mn}\hat{\zeta}^m\hat{\zeta}^n + \frac{\Lambda}{3}\Omega^{-3}\hat{h}_{ab}\hat{\eta}^a\hat{\eta}^b + \Omega^{-3}\hat{D}^b\Omega D_b\Omega\right] - \left[-\hat{\hat{E}}_{mn}\hat{\hat{\zeta}}^m\hat{\hat{\zeta}}^n + \frac{\Lambda}{3}\mathring{\Omega}^{-3}\hat{\hat{h}}_{ab}\hat{\hat{\eta}}^a\hat{\hat{\eta}}^b + \mathring{\Omega}^{-3}\hat{\hat{D}}^b\Omega\hat{\hat{D}}_b\Omega\right]\right\} dV_\Sigma. \tag{5.11}$$

(Actually $\hat{\hat{E}}_{mn} = 0$ since anti-de Sitter space is conformally flat.) We shall now be concerned with the integrand in (5.11). Let us first evaluate the quantity $Q_1 = (\Lambda/3)[\Omega^{-3} - \mathring{\Omega}^{-3}]$. Recall that we have chosen $\Omega = \lambda^{-1/2} = (g_{00})^{-1/2} = (\mathring{g}_{00} + (h_{tt}/r)(\theta,\varphi) + O(1/r^2))^{-1/2}$. A straightforward expansion of $Q_1 = (\Lambda/3)[(\mathring{g}_{00} + h_{tt}/r + O(1/r^2))^{3/2} - (\mathring{g}_{00})^{3/2}]$ with respect to r^{-1} provides us with the following result:

$$Q_1 = 3(-\Lambda/3)^{3/2}(h_{tt}/-2) + O(1/r). \tag{5.12}$$

Let us next evaluate the quantity $Q_2 = R - R_0 \equiv \Omega^{-3}(\hat{D}^b\Omega)D_b\Omega - (\mathring{\Omega})^{-3}(\mathring{\hat{D}}^b\mathring{\Omega})(D_b\mathring{\Omega})$. Recall first that $g^{rr} = \mathring{g}^{rr} + r^{-1}h^{rr} + O(r^{-2})$, $g^{\theta\theta} = \mathring{g}^{\theta\theta} + r^{-1}h^{\theta\theta} + O(r^{-2})$, $g^{\varphi\varphi} = \mathring{g}^{\varphi\varphi} + r^{-1}h^{\varphi\varphi} + O(r^{-2})$. A somewhat long but direct calculation provides us with the leading terms in the expression

$$R = \frac{1}{4}\lambda^{-1/2}D_r\left(\mathring{g}_{00} + \frac{h_{tt}}{r}\right)^2\left(\mathring{g}^{rr} + \frac{h^{rr}}{r}\right) + O\left(\frac{1}{r}\right), \tag{5.13}$$

where it must be noticed that terms in $h^{\theta\theta}$ and $h^{\psi\psi}$ and cross terms will not contribute to the surface integral due to their falloff on the domain of integration. The result is the following:

$$Q_2 \equiv R - R_0 = (-\Lambda/3)^{3/2}(h_{tt}/-2) + O(1/r). \tag{5.14}$$

Remark: note that in the case of the Schwarschild anti-de Sitter metric $g_{00} = 1 - 2M/r - (\Lambda/3)r^2$, and therefore $h_{tt} = -2M$.

Finally, the radius of the cross section at I being $-3/\Lambda$,

$$Q_{k_\Lambda}(\xi) = -1/8\pi(-3/\Lambda)^{1/2} \times \frac{1}{3}\lim_{\Omega\to 0}\int_{\partial\Sigma(\Omega)}\left\{-\hat{E}_{mn}\hat{\xi}^m\hat{\xi}^n + 4\left(-\frac{\Lambda}{3}\right)^{1/2}\frac{h_{tt}}{-2}\right\}dV_\Sigma. \tag{5.15}$$

If we return to the notations introduced in (Ref. 3, Appendix D),

$$-\tfrac{1}{2}h_{tt} = \bar{g}_{00}^{(3)} = \hat{E}_{00} = \hat{E}_{mn}\hat{\xi}^m\hat{\xi}^n,$$

the result follows immediately:

$$Q_{k_\Lambda}(\xi) = \tfrac{1}{2}Q_C(\xi),$$

where C is the cross section induced on I by the static slice Σ. We thus have the following theorem.

Theorem: If an asymptotically anti-de Sitter space-time (in which matter sources are spatially compact) is equipped with a static Killing vector field ξ^a, the corresponding generalized Komar integral $Q_{k_\Lambda}(\xi)$ is related to the charge integral $Q_C(\xi)$ at I, via $Q_{k_\Lambda}(\xi) = \frac{1}{2}Q_C(\xi)$.

VI. STATIONARY SPACE-TIMES

Let us now suppose that the space-time (M, g_{ab}) is stationary. The situation is then analogous to that in the static case; the only difference is that the Killing vector ξ^a is not hypersurface orthogonal.

Recall that the generalized Komar integral associated with ξ^a in (M, g_{ab}) is given by

$$Q_{k_\Lambda}(\xi) = \lim_{\partial\Sigma\to I}\left\{\frac{1}{16\pi}\int_{\partial\Sigma}(\nabla_a\xi_b)\epsilon^{ab}{}_{cd}\,dS^{cd} + \frac{\Lambda}{8\pi}\int_\Sigma \xi^a\epsilon_{abcd}\,d\Sigma^{bcd}\right\}, \tag{6.1}$$

where $\nabla_a\xi_b = \lambda^{-1}\xi_{[b}\nabla_{a]}\lambda + \frac{1}{2}\lambda^{-1}\epsilon_{abcd}\xi^c\omega^d$, λ and ω^a being, respectively, the norm and the twist of ξ^a in (M, g_{ab}).

Let us focus on the integrands in (6.1). The integrand in the volume term, $\xi^a\epsilon_{abcd}$ satisfies $\pounds_\xi(\xi^a\epsilon_{abcd}) = 0$ and is orthogonal to ξ^a. It can be therefore naturally identified with a three-form on the three-manifold $\mathcal{O}$, the manifold of orbits of ξ. In the surface integral, $\lambda^{-1}\xi_b(\nabla_a\lambda)\epsilon^{ab}{}_{cd}$ is also a two-form on $\mathcal{O}$. We shall now show that the term involving the twist of ξ does not contribute to the surface integral in $Q_{k_\Lambda}(\xi)$, so that the calculation can be performed entirely on $\mathcal{O}$. In the anti-de Sitter chart (t, r, θ, φ) available in the neighborhood of I, $\int_{\partial\Sigma}\xi_a\omega_b\,dS^{ab} = \int_{\partial\Sigma}\xi_c\omega_d\theta^{[c}\varphi^{d]}\,d\theta\,d\varphi$, where θ^c and φ^d are dual to the one-forms $d\theta$ and $d\varphi$, respectively. Using the asymptotic behavior near I, displayed in Sec. II, we have

$$\xi\cdot\varphi\,(=g_{0\varphi})\sim r^{-1};\quad \xi\cdot\theta\,(=g_{0\theta})\sim r^{-1};$$

$$\omega_a\theta^a = \hat{\epsilon}_{\theta 0}{}^{\varphi r}(\partial_\varphi g_{0r} - \partial_r g_{0\varphi})\sim f_1(\theta,\varphi);$$

and

$$\omega_a\varphi^a = \hat{\epsilon}_{\varphi 0}{}^{\theta r}(\partial_\theta g_{0r} - \partial_r g_{0\theta})\sim f_2(\theta,\varphi).$$

[It is straightforward to check that $\hat{\epsilon}_{\theta 0}{}^{\varphi r}\sim r^2 f(\theta,\varphi)$.] As a result, the twist term does not contribute to the surface integral in $Q_{k_\Lambda}(\xi)$; the only nonzero contribution to $Q_{k_\Lambda}(\xi)$ comes from integrals on forms on the manifold of orbits.

We can therefore proceed as in the case of static space-times, replacing the static slice by the manifold $\mathcal{O}$ of orbits of ξ^a in the calculation. The availability of diffeomorphisms on this manifold enables us to choose "data"—a three-metric $\mathring{q}_{ab}$ and scalar field $\mathring{\lambda}$—on $\mathcal{O}$ which, when evolved give rise to an anti-de Sitter solution $(\mathring{M}, \mathring{g}_{ab}, \mathring{\xi}^a)$ such that $\int_\Sigma \xi_b\epsilon^b{}_{ijk}\,d\Sigma^{ijk}$ equals $\int_\Sigma \mathring{\xi}_b\mathring{\epsilon}^b{}_{ijk}\,d\Sigma^{ijk}$. The evaluation of $Q_{k_C}(\xi)$ is then obtained as in the static case with the following modifications: (i) It is straightforward to show that the twist terms in (5.5) do not contribute to the surface integral; and the terms involving the extrinsic curvature in (5.8) and (5.9) do not contribute because, due to the falloff of π_{ab} (studied in Ref. 3), all these terms are at least of order r^{-2}, except π^{rr} which cancels in (5.9). Hence the resulting expression falls off like r^{-4}. We thus have the following theorem.

Theorem: If an asymptotically anti-de Sitter space-time (in which matter sources are spatially compact) is equipped with a stationary Killing vector field ξ^a, the corresponding generalized Komar integral $Q_{k_\Lambda}(\xi)$ is related to the charge integral $Q_C(\xi)$ at I via $Q_{k_\Lambda}(\xi) = \frac{1}{2}Q_C(\xi)$, where C is any cross section of I lifting the boundary $\partial\Sigma$ of the manifold of orbits of ξ.

ACKNOWLEDGMENTS

It is a pleasure to thank the Syracuse Relativity Group for its support, A. Komar and R. Penrose for suggestions, and A. Ashtekar for fruitful discussions.

[1] R. Penrose, "Structure of Space-Time," in *Battelle Rencontres 1967, Lectures in Mathematics and Physics*, edited by C. DeWitt and J. Wheeler (Benjamin, New York, 1968).
[2] A. Ashtekar and A. Magnon, edited by Class. Quantum Gravity **1**, L39 (1984).
[3] M. Henneaux and C. Teitelboim, "Asymptotically anti-de Sitter spaces," Commun. Math. Phys. **98**, 391 (1985).
[4] A. Komar, Phys. Rev. **113**, 934 (1959).
[5] B. Carter, Commun. Math. Phys. **10**, 280 (1968); M. Demianski, Act. Astron. **23**, 211 (1973).
[6] S. L. Bazanski, *Proceedings of the 10th International Conference on General Relativity and Gravitation*, Padova, 1983, edited by B. Bertotti, F. de Felice, and A. Pascolini, p. 24.
[7] L. F. Abbott and S. Deser, Nucl. Phys. B 195 (1982).
[8] S. W. Hawking and C. F. R. Ellis, *The Large Scale Structure of Space-Time*, (Cambridge U.P., Cambridge, 1973), p. 132.
[9] Anti-de Sitter space-time as embedded in Minkowski space M_5 with signature $(-,-,+,+,+)$ defines an O(2,3) asymptotic group which does not include the supertranslations of M_5. Hence there exist infinitely many O(2,3) invariant equivalence classes of asymptotically anti-de Sitter space-times.

Some exact inhomogeneous solutions of Einstein's equations with symmetries on the hypersurfaces $t = \text{const}$

F. Argüeso and J. L. Sanz
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Santander, 39005 Santander, Spain

(Received 11 February 1985; accepted for publication 24 May 1985)

The solution of Einstein's field equations is studied for a metric written in the form $(\delta \neq \gamma) ds^2 = -\alpha^2(t,r,\theta,\varphi)dt^2 + e^{2\beta(t,r)}dr^2 + e^{2\gamma(t,r)}d\theta^2 + e^{2\delta(t,r)}M^2(\theta)d\varphi^2$. A perfect fluid, which flows orthogonally to the hypersurfaces $t = \text{const}$, is considered as matter content. These hypersurfaces admit a translational Killing vector, which will not be, in general, a Killing vector of the whole space-time. All the possible solutions are obtained when α depends on the variable φ. These solutions represent either a perfect fluid without expansion or vacuum with a cosmological constant $\Lambda \lesseqgtr 0$. Some particular inhomogeneous solutions are obtained for α independently of φ. These solutions are physical, the fluid obeys an equation of state $p = \rho$ (stiff matter), and the space-time admits, apparently, only a group G_2 of isometries. A vacuum family is also obtained in this case.

I. INTRODUCTION

Collins[1] has proposed the idea of "intrinsic symmetries" as an alternative approach in order to obtain inhomogeneous cosmologies. He and other authors[2–4] have used this technique, applying it to different space-times. In the present paper, we start from a metric with one translational symmetry on the hypersurfaces $t = \text{const}$. This metric can be written as $(\delta \neq \gamma)$

$$ds^2 = -\alpha^2(t,r,\theta,\varphi)dt^2 + e^{2\beta(t,r)}dr^2 + e^{2\gamma(t,r)}d\theta^2 + e^{2\delta(t,r)}M^2(\theta)d\varphi^2. \tag{1}$$

The form of this metric is very similar to one previously analyzed by Martinez and Sanz.[4] The difference is that we assume $\delta \neq \gamma$, whereas they only studied the case $\delta = \gamma$, which included the spherical, planar, and hyperbolic intrinsic symmetries.

The only Killing vector admitted by the spatial part of the metric (1) is ∂_φ. This Killing vector constitutes the only "intrinsic symmetry" of the metric; in the case $\partial_\varphi \alpha = 0$, this vector will be a Killing vector of the whole space-time.

We shall assume that the source of the gravitational field is a perfect fluid whose flowlines are orthogonal to the hypersurfaces $t = \text{const}$—we shall analyze the vacuum case, too. This fluid will not be, in general, geodesic and its shear will be different from zero, whereas its vorticity will be always equal to zero. This can be easily computed from the metric (1), taking into account the definitions of the kinematical quantities.[5]

In order to solve Einstein's field equations we use the ADM equations[6] as they have been written by York.[7] This enables us to separate the equations into constraint equations on the hypersurfaces and evolution equations.

In this formalism, a general metric reads[7]

$$ds^2 = -(\alpha^2 - \beta^i\beta_i)dt^2 + 2\beta_i\,dt\,dx^i + \gamma_{ij}\,dx^i\,dx^j \quad (i,j,\ldots = 1,2,3), \tag{2}$$

where γ_{ij} is the metric of the hypersurfaces $t = \text{const}$, β_i is called the shift vector and moves the spatial coordinates as the data are evolved from one slice to the next, and α is called the lapse function and measures the orthogonal proper time between neighboring slices $\alpha\delta t$.

Following the notation established in the reference mentioned above[7] we write the corresponding equations for our metric in Sec. II. First, we impose the condition $\partial_\varphi \alpha \neq 0$ and obtain the general solution for this case. We calculate the density, pressure, and expansion scalar, finding that for all the solutions the density is constant and the expansion vanishes. In several cases the fluid obeys a nonphysical equation of state $p = -\rho$. However, these solutions can be interpreted as vacuum solutions with a cosmological constant $\Lambda = \rho = -p$. In Sec. III we study the equations when $\partial_\varphi \alpha = 0$, $\partial_\theta \alpha \neq 0$. Now, the space-time will admit at least one Killing vector ∂_φ. We obtain some particular solutions with a group G_2 of isometries acting on two-dimensional orbits. The two Killing vectors are orthogonal and the corresponding space-time belongs to the class Bii in the classification scheme developed by Wainwright[8] for inhomogeneous cosmologies. Moreover, the equation of state satisfied by the fluid is $p = \rho$ (stiff matter) and one of the solutions is a subcase of one found by Wainwright *et al.*[9] Some vacuum solutions are also obtained in this case.

II. BASIC EQUATIONS. CASE $\partial_\varphi \alpha \neq 0$

In this section, we will write and solve Einstein's field equations for a metric written in the form (1), using for it the ADM formalism. The general equations of this formalism can be written in the following way:

$$2\tau = R + (\operatorname{tr} K)^2 - K^{ij}K_{ij}, \tag{3}$$

$$j^i = D_j(K^{ij} - \gamma^{ij}\operatorname{tr} K), \tag{4}$$

$$\partial_t \gamma_{ij} = -2\alpha K_{ij} + D_i\beta_j + D_j\beta_i, \tag{5}$$

$$\partial_t K_{ij} = -D_iD_j\alpha + (\mathcal{L}_\beta K)_{ij} + \alpha[R_{ij} - 2K_{il}K^l{}_j + K_{ij}\operatorname{tr} K - S_{ij} + \tfrac{1}{2}\gamma_{ij}(\operatorname{tr} S - \tau)]. \tag{6}$$

We use the corresponding definitions of the quantities appearing in these equations as given in Ref. 7. Equations (3)

 0022-2488/85/123118-07$02.50

and (4) are the constraint equations and (5) and (6) are the evolution equations. Besides, we can use the following two equations, which emerge from the conservation law of the energy-momentum tensor:

$$\partial_t \tau + \alpha D_j j^i = \alpha(S^{ij}K_{ij} + \tau \operatorname{tr} K) - 2j^i D_i \alpha + \beta^i D_i \tau, \tag{7}$$

$$\partial_t j^i + \alpha D_j S^{ij} = \alpha(2K^{ij}j_j + j^i \operatorname{tr} K) - S^{ij}D_j\alpha - \tau D^i\alpha + \mathcal{L}_\beta j^i. \tag{8}$$

For the metric (1), if we calculate the quantities that appear in these equations, we can write the equivalent set

$$\rho = -M_{\theta\theta}M^{-1}e^{-2\gamma} + e^{-2\beta}(-\gamma'' - \delta'' + \gamma'\beta' + \delta'\beta' - \delta'\gamma' - \gamma'^2 - \delta'^2) + K_1K_2 + K_1K_3 + K_2K_3, \tag{9}$$

$$(K_2 + K_3)' + \gamma'(K_2 - K_1) + \delta'(K_3 - K_1) = 0, \tag{10}$$

$$(K_1 + K_3)_\theta + M_\theta M^{-1}(K_3 - K_2) = 0, \tag{11}$$

$$(K_1 + K_2)_\varphi = 0, \tag{12}$$

$$\dot\beta = -\alpha K_1, \tag{13}$$

$$\dot\gamma = -\alpha K_2, \tag{14}$$

$$\dot\delta = -\alpha K_3, \tag{15}$$

$$\alpha'_\theta - \gamma'\alpha_\theta - \alpha M_\theta M^{-1}(\gamma' - \delta') = 0, \tag{16}$$

$$\alpha'_\varphi - \delta'\alpha_\varphi = 0, \tag{17}$$

$$\alpha_{\theta\varphi} - M_\theta M^{-1}\alpha_\varphi = 0, \tag{18}$$

$$\dot K_1 = e^{-2\beta}(-\alpha'' + \beta'\alpha' + \alpha R_{11}) + \alpha K_1 \operatorname{tr} K + \tfrac{1}{2}\alpha(p - \rho), \tag{19}$$

$$\dot K_2 = e^{-2\gamma}(-\alpha_{\theta\theta} - e^{2\gamma - 2\beta}\gamma'\alpha' + \alpha R_{22}) + \alpha K_2 \operatorname{tr} K + \tfrac{1}{2}\alpha(p - \rho), \tag{20}$$

$$\dot K_3 = e^{2\delta}M^{-2}(-\alpha_{\varphi\varphi} - e^{2\delta - 2\beta}M^2\delta'\alpha' - e^{2\delta - 2\gamma}MM_\theta\alpha_\theta + \alpha R_{33}) + \alpha K_3 \operatorname{tr} K + \tfrac{1}{2}\alpha(p - \rho), \tag{21}$$

where $(' \equiv \partial_r, {}_\theta \equiv \partial_\theta, {}_\varphi \equiv \partial_\varphi, \dot{} \equiv \partial_t)$, p is the pressure and ρ is the density of the fluid, $K_i \equiv K^i{}_i$ are the mixed components of the extrinsic curvature—which is diagonal, R_{ij} is the three-Ricci tensor of the hypersurfaces, and R is the curvature ($R = \gamma^{ij}R_{ij}$). A direct calculation leads to the values

$$-R_{11} = \gamma'' + \gamma'^2 + \delta'' + \delta'^2 - \beta'(\gamma' + \delta'), \tag{22a}$$

$$-R_{22} = e^{2\gamma - 2\beta}(\gamma'' + \gamma'^2 + \gamma'\delta' - \gamma'\beta') + M_{\theta\theta}M^{-1}, \tag{22b}$$

$$-R_{33} = e^{2\delta - 2\beta}M^2(\delta'' + \delta'^2 + \delta'\gamma' - \delta'\beta') + e^{2\delta - 2\gamma}MM_{\theta\theta}, \tag{22c}$$

$$R_{12} = (\gamma' - \delta')M_\theta M^{-1}, \quad R_{13} = R_{23} = 0. \tag{22d}$$

Finally, the conservation equations can be written as follows:

$$\dot\rho = \alpha(\rho + p)\operatorname{tr} K, \tag{23}$$

$$\alpha\,\partial_j p = -(\rho + p)\partial_j\alpha. \tag{24}$$

If we calculate the kinematical quantities, we obtain, taking into account that the fluid is orthogonal to the slices $t = \text{const}$, that the vorticity is identically zero and the expansion scalar (θ), shear ($\sigma_{\alpha\beta}$), and acceleration (ξ^α) are given by

$$\theta = -\operatorname{tr} K, \tag{25}$$

$$\sigma^0{}_0 = \sigma^0{}_i = 0, \quad \sigma^i{}_i = -(K_i + \theta/3), \tag{26}$$

$$\xi^0 = 0, \quad \xi^i = \alpha^{-1}\gamma^{ij}\partial_j\alpha. \tag{27}$$

So, the fluid is irrotational; it will be geodesic, according to (27), if and only if the lapse function only depends on time, and it will be shear-free provided that $K_1 = K_2 = K_3 = -\frac{1}{3}\theta$ as we can easily see from Eqs. (25) and (26). This implies, taking into account the expression of the extrinsic curvature in terms of the metric, that $\dot\beta = \dot\gamma = \dot\delta$.

These results are general for a diagonal metric, with the only assumption of a matter content, which is a perfect fluid flowing orthogonally to the slices $t = \text{const}$. Therefore, a solution of this type will be a Friedmann–Robertson–Walker model if and only if $\alpha(t)$ and $\dot\beta = \dot\gamma = \dot\delta$.

Now, we will begin to solve Eqs. (9)–(24). First, we only consider the case $\alpha_\varphi \neq 0$, if we take into account Eqs. (12)–(14): $K_1 = -K_2$ and $\dot\beta = -\dot\gamma$. So, $\beta = -\gamma + F(r)$ and, with a suitable change of frame, redefining the coordinate r, we find $\beta = -\gamma$.

If we turn our attention to Eqs. (17) and (18) and integrate them, we obtain the general expression for the lapse function

$$\alpha = \phi(t,\varphi)Me^{\delta} + H(t,r,\theta). \tag{28}$$

Substituting this value into Eq. (16) and taking the derivative with respect to φ, we find

$$(\delta' - \gamma')\phi_\varphi M_\theta = 0. \tag{29}$$

Therefore, two possibilities arise from this equation, either $\delta' = -\beta'$ or $M_\theta = 0$. In the second case, the hypersurfaces will admit two "intrinsic" Killing vectors ∂_θ and ∂_φ. We shall consider separately the two cases A ($M_\theta \neq 0$, $\delta' = -\beta'$) and B ($M_\theta = 0$).

Case A ($M_\theta \neq 0$, $\delta' = -\beta'$): If we develop Eq. (16), we obtain the following expression for $H(t,r,\theta)$:

$$H = L(t,\theta)e^{-\beta} + A(t,r). \tag{30}$$

If we take into account Eq. (11), this leads to two different subcases: $A_1(\dot\delta = -\dot\beta = \dot\gamma)$ and $A_2(\alpha^{-1}\alpha_\theta = M^{-1}M_\theta)$. The subcase A_1 has been treated exhaustively by other authors,[4] so we will only consider A_2. The general solution for α corresponding to A_2 will be

$$\alpha = \phi(t,\varphi)Me^{-\beta}. \tag{31}$$

We know that $\delta' = -\beta'$, so δ can be expressed as

$$\delta = -\beta + G(t) \tag{32}$$

and the general metric can be written as

$$ds^2 = -\phi^2(t,\varphi)M^2e^{-2\beta}dt^2 + e^{2\beta}dr^2 + e^{-2\beta}(d\theta^2 + e^{2G}M^2d\varphi^2). \tag{33}$$

It is very easy to conclude that Eq. (10) is equivalent to

$$\dot\beta' = \dot\beta\beta'. \tag{34}$$

On the other hand, introducing the expression (33) for the metric into Eqs. (9) and (19)–(21) and substituting the values of the extrinsic curvature given by (13)–(15), and the value of the three-Ricci tensor, we obtain a set of four equations (two of them can be considered as definitions of the density and pressure). If we subtract Eqs. (19) and (20) we find

$$(-\dot\beta\alpha^{-1})^{\cdot} = \tfrac{1}{2}e^{-2\beta}(-\alpha'' - \alpha\beta'^2 + \alpha\beta'') + \tfrac{1}{2}e^{2\beta}(\alpha_{\theta\theta} + \alpha M^{-1}M_{\theta\theta}) + \alpha^{-1}\dot\beta(\dot G - \dot\beta), \tag{35}$$

and substituting Eq. (31) into the previous equation and taking the derivatives with respect to θ, r, and φ, we obtain

$$(M^{-2}\phi^{-1}\dot\phi\dot\beta)'_{\theta\varphi} = 0, \tag{36}$$

which implies necessarily one of the two possibilities $\dot\beta' = 0$ or $(\phi^{-1}\dot\phi)_{\varphi} = 0$. The first one leads—see Eq. (34)—to $\dot\beta = 0$ or $\beta' = 0$, but if $\beta' = 0$, we can easily prove, interchanging the coordinates r and θ, that $M_\theta = 0$ (this corresponds to case B, to be studied later) so $\dot\beta = 0$ and then Eq. (35) is equivalent to

$$M_{\theta\theta} = kM, \quad e^{-4\beta}(\beta'' - \beta'^2) = -k, \quad k = \text{const.} \tag{37}$$

The other possibility $(\phi^{-1}\dot\phi)_{\varphi} = 0$ leads to $\dot\beta = 0$, after a tedious calculation taking into account Eqs. (34) and (35). Therefore, it is a particular case of the previous one.

Integrating Eqs. (37), we obtain the following values for M and β (after rescaling the variables θ and r):

$$M = \begin{cases} \sin\theta, & k = -1, \\ \theta, & k = 0, \\ \sinh\theta, & k = +1, \end{cases} \tag{38a}$$

$$k = 0, \quad e^{-2\beta} = A^2 r^2, \tag{38b}$$

$$k \neq 0, \quad \begin{cases} A \neq 0, & e^{-2\beta} = Ar^2 + kA^{-1}, \quad k = \pm 1, \\ A = 0, & e^{-2\beta} = 2r, \quad k = -1, \end{cases} \tag{38c}$$

where A is a constant which appears in the integration. Now, we subtract Eqs. (19) and (21). A straightforward calculation leads to the following relation between ϕ and G:

$$\phi^{-1}\dot\phi\dot G - \ddot G - \dot G^2 = -(\phi^2 + \phi\phi_{\varphi\varphi}e^{-2G}). \tag{39}$$

Our next step is to calculate the density and pressure to establish if the solutions are physical. The final result for the density, see Eq. (9), is

$$k = 0, \quad \rho = -3A^2, \tag{40}$$

$$k = 0, \quad A \neq 0, \quad \rho = -3A, \tag{41}$$

$$k \neq 0, \quad A = 0, \quad \rho = 0. \tag{42}$$

Taking into account Eq. (23), since $\dot\rho = 0$, we find that either $p = -\rho$ or $\dot\delta = 0$. This last possibility leads to $\delta = \gamma$ (the case studied in Ref. 4). Therefore, we deduce that the only possible equation of state satisfied by the fluid will be $p = -\rho$.

However, we can reinterpret it as a vacuum solution with a cosmological constant $\Lambda = \rho = -p$. From Eq. (38), we see that when $A < 0$ and $k = 1$, $e^{-2\beta} < 0$, which is impossible. Then, in the case $\Lambda > 0$, the only solution will read after an appropriate change of coordinates,

$$ds^2 = 3\Lambda^{-1}\{-\phi^2\sin^2\theta(1-r^2)dt^2 + (1-r^2)^{-1}dr^2 + (1-r^2)(d\theta^2 + e^{2G}\sin^2\theta\, d\varphi^2)\}. \tag{43}$$

When $\Lambda < 0$, we have three possible solutions:

$$k = 0, \quad ds^2 = -\phi^2\theta^2 r^2 dt^2 + 3(|\Lambda|r^2)^{-1}dr^2 + r^2(d\theta^2 + e^{2G}\theta^2 d\varphi^2), \tag{44}$$

$$k = \pm 1, \quad ds^2 = 3|\Lambda|^{-1}\{-\phi^2 M_k^2(r^2 + k)dt^2 + (r^2 + k)^{-1}dr^2 + (r^2 + k)(d\theta^2 + e^{2G}M_k^2\, d\varphi^2). \tag{45}$$

For $\Lambda = 0$, we obtain a vacuum solution without a cosmological constant. By calculating the Riemann tensor,[7] we prove that this solution is the Minkowski space-time. In all the cases ϕ and G are related by Eq. (39).

In conclusion, we have surveyed exhaustively case A, obtaining the general solution, which is a new—as far as we know—vacuum solution with cosmological constant, depending on two arbitrary functions ϕ and G related by Eq. (39).

Case B($M_\theta = 0$): Now, we are going to consider the case $M_\theta = 0$. Taking into account Eqs. (16)–(18), we can write the lapse function as

$$\alpha = f(t,r) + g(t,\theta)e^{\gamma} + h(t,\varphi)e^{\delta}, \tag{46}$$

where $\beta = -\gamma$. If we regard Eq. (11) and develop it, taking into account Eqs. (13) and (15), we find

$$(\dot\gamma - \dot\delta)\alpha_\theta = 0. \tag{47}$$

Therefore, either $\alpha_\theta = 0$ or $\dot\gamma = \dot\delta$. If we now use Eq. (10), substituting into it the value of α given by (46) and taking the derivative with respect to φ, this equation leads to

$$(\dot\delta + \dot\gamma)' = -2\dot\gamma\gamma'. \tag{48}$$

Reintroducing this expression into Eq. (10), we obtain

$$(\dot\delta + \dot\gamma)(\alpha' - \delta'\alpha) = 0. \tag{49}$$

Therefore, we have to deal with four different subcases [see Eqs. (47) and (49)]:

$$\mathrm{B}_1\text{: } \alpha_\theta \neq 0, \quad \dot\delta + \dot\gamma = 0 \quad (\Rightarrow \gamma(r),\ \delta(r),\ K_1 = K_2 = K_3 = 0), \tag{50}$$

$$\mathrm{B}_2\text{: } \alpha_\theta \neq 0, \quad \dot\delta + \dot\gamma \neq 0 \quad (\Rightarrow \alpha' = \delta'\alpha,\ \dot\gamma = \dot\delta,\ \gamma' = \delta',\ f' = \delta' f), \tag{51}$$

$$\mathrm{B}_3\text{: } \alpha_\theta = 0, \quad \dot\delta + \dot\gamma \neq 0 \ (\Rightarrow g = 0,\ \alpha' = \delta'\alpha,\ f' = \delta' f), \tag{52}$$

$$\mathrm{B}_4\text{: } \alpha_\theta = 0, \quad \dot\delta + \dot\gamma = 0 \ (\Rightarrow g = 0). \tag{53}$$

Subcase B_1($\alpha_\theta \neq 0$, $\dot\delta + \dot\gamma = 0$): By writing Eqs. (19)–(21) and combining them in an appropriate way, we find

$$\alpha(p - \rho) + 2e^{2\gamma}(-\alpha'' - \gamma'\alpha' + \alpha R_{11}) = 0, \tag{54}$$

$$\alpha_{\theta\theta} - \alpha R_{22} + e^{4\gamma}(\alpha R_{11} - \alpha'') = 0, \tag{55}$$

$$e^{2\delta - 2\gamma}(\alpha_{\theta\theta} + e^{4\gamma}\alpha'\gamma' - \alpha R_{22}) = \alpha_{\varphi\varphi} + e^{2\delta + 2\gamma}\alpha'\delta' - \alpha R_{33}, \tag{56}$$

and we have another equation as definition of the density [Eq. (9)],

$$\rho = -e^{2\gamma}(\gamma'' + 2\gamma'^2 + \delta'' + \delta'^2 + 2\gamma'\delta'). \tag{57}$$

Now, if we calculate the third derivative of Eq. (56) with respect to the variable φ, it is deduced that

$$h_{\varphi}^{-1}h_{\varphi\varphi\varphi} = e^{2\delta + 2\gamma}(2\gamma'^2 + \gamma'' - \delta'' - 2\delta'^2) = a, \quad h_{\varphi\varphi} = ah + B(t), \tag{58}$$

where a is a constant and $B(t)$ is an arbitrary function of its

argument. Doing the same with respect to θ, we find a similar relation for g,

$$g_\theta^{-1} g_{\theta\theta\theta} = -e^{4\gamma}[(\delta' - \gamma')^2 + ae^{-2\delta-2\gamma}] = b,$$
$$g_{\theta\theta} = bg + D(t), \tag{59}$$

where b is a constant and $D(t)$ is an arbitrary function of its argument. If we substitute Eqs. (58) and (59) into (56), this equation can be written as

$$(\gamma' - \delta')(f' - \delta' f) + fae^{-2\delta-2\gamma} = Be^{-\delta-2\gamma} - De^{-3\gamma}. \tag{60}$$

Equation (55) can be reduced to the following set:

$$\delta'' + \delta'^2 = 0, \quad be^{-4\gamma} = \gamma'' + \gamma'^2, \quad f'' = De^{-3\gamma}. \tag{61}$$

From Eqs. (58), (59), and (61), we obtain

$$b = \gamma'(\delta' - \gamma')e^{4\gamma}. \tag{62}$$

Now, we will consider different subcases depending on the value of the constant b: $b \neq 0$.

Integrating Eqs. (61) for δ and γ, we find

$$e^\delta = cr + d, \quad \begin{matrix} e^{2\gamma} = kr^2 + bk^{-1}, & k \neq 0, \\ e^{2\gamma} = 2r|b|^{1/2}, & b < 0, \end{matrix} \tag{63}$$

where c, d, and k are arbitrary constants.

Let us consider the case $k \neq 0$. Substituting γ and δ given by the previous formulas into Eq. (62), one obtains that $d = 0$.

For $b < 0$, an appropriate change of coordinates leads to the following expression for the metric:

$$ds^2 = -(f + ge^\gamma + he^\delta)dt^2 + k^{-1}(r^2 - 1)^{-1} dr^2 + k(r^2 - 1)d\theta^2 + r^2 d\varphi^2. \tag{64}$$

Moreover, we obtain from Eqs. (58) and (59) that $a = k$, $b = -k^2$, so h and g must satisfy the differential equations

$$h_{\varphi\varphi} = kh + B(t), \quad g_{\theta\theta} = -k^2 g + D(t). \tag{65}$$

Integrating these equations and the corresponding ones for f [Eqs. (60) and (61)]:

$$k < 0: \quad \alpha = [R\cos(|k|^{1/2}\varphi) + S\sin(|k|^{1/2}\varphi)]e^\delta + (T\cos k\theta + Q\sin k\theta)e^\gamma + W, \tag{66a}$$

$$k > 0: \quad \alpha = (Re^{|k|^{1/2}\varphi} + Se^{-|k|^{1/2}\varphi})e^\delta + (T\cos k\theta + Q\sin k\theta)e^\gamma + W. \tag{66b}$$

All the arbitrary functions appearing in these expressions are only time dependent. For, $b > 0$, the metric can be written as

$$ds^2 = -\alpha^2 dt^2 + k^{-1}(r^2 + 1)^{-1} dr^2 + k(r^2 + 1)d\theta^2 + r^2 d\varphi^2. \tag{67}$$

But in this case, k must be always positive, and the only possible value for α will be

$$\alpha = [N\cos(k^{1/2}\varphi) + O\sin(k^{1/2}\varphi)]e^\delta + (Pe^{k\theta} + Ue^{-k\theta})e^\gamma + W, \tag{68}$$

taking into account that $a = -k$ and $b = k^2$ from Eqs. (58) and (59). Calculating the density and pressure according to Eqs. (57) and (54) for the two cases $b \gtrless 0$, we find

$$\rho = -3k, \quad p = -(\rho + 2Wk\alpha^{-1}). \tag{69}$$

So, the solution will be physical when $k < 0$. Therefore, the only physical solution will be given by [see Eqs. (64) and (66a) and write $k = -A^2$, $A > 0$]

$$ds^2 = A^{-2}\{-\alpha^2 dt^2 + (1 - r^2)^{-1} dr^2 + (1 - r^2)d\theta^2 + r^2 d\varphi^2\}, \tag{70a}$$

$$\alpha = r(B\sin\varphi + C\cos\varphi) + (1 - r^2)^{1/2}(D\sin\theta + E\cos\theta) + F, \tag{70b}$$

$$\rho = 3A^2, \quad p = -\rho + 2A^2 F\alpha^{-1}. \tag{70c}$$

The density is always constant and the only kinematical quantity different from zero is the acceleration [see Eqs. (25)–(27)].

We could obtain a vacuum solution with a cosmological constant provided that $F \equiv 0$. All the solutions with $F \equiv 0$, could be interpreted in this way since the fluid will obey an equation of state $p = -\rho$ in this case and defining $\Lambda = \rho = -p$, we introduce the cosmological constant in Einstein's field equations.

Other results can be obtained assuming that $k = 0$. In this case, the solution for γ is $e^{2\gamma} = 2|b|^{1/2}r$. We can find the value of α, integrating the relations for f, g, and h and the final result is ($k \equiv |b|^{1/2}$)

$$\alpha = \tfrac{1}{2}B\varphi^2 + L\varphi + [R\cos(k\theta) + T\sin(k\theta)]e^\gamma + Bk^{-1}r + W, \tag{71}$$

and the values of the density and pressure are

$$\rho = 0, \tag{72}$$

$$p = 2B\alpha^{-1}. \tag{73}$$

For $B \neq 0$, we have an unphysical solution since $\rho = 0, p \neq 0$. Therefore, B must be equal to zero and then the solution is Minkowski space-time as can be easily proven by a direct calculation of the Riemann tensor.

For $b = 0$, we obtain a solution with the same characteristics that appear in the solution studied before. In particular the density and the pressure have the structure (72), (73). In the same way, the vacuum solution can be proven to be Minkowski space-time.

Subcase B_2 ($\alpha_\theta \neq 0$, $\dot\delta + \dot\gamma \neq 0$): In this subcase $\delta' = \gamma'$ and $\dot\delta = \dot\gamma$, so $\delta = \gamma$ and we do not consider it since it only gives solutions that can be found in Ref. 4.

Subcase B_3 ($\alpha_\theta = 0$, $\dot\delta + \dot\gamma \neq 0$): In this subcase ($\alpha_\theta = 0$, $\dot\delta + \dot\gamma \neq 0$) we can write the lapse function as $\alpha = he^\delta$, because $g = 0$ and $f' = f\delta'$, subtracting Eqs. (19) and (20), and combining Eqs. (19)–(21) algebraically, we obtain a couple of equations which read [we shall consider Eq. (19) as the definition of the pressure]

$$\dot\gamma\dot h h^{-1} - \ddot\gamma = e^{2\delta+2\gamma}h^2(\delta'' + \delta'^2), \tag{74}$$

$$\ddot\delta - \dot\delta\dot h h^{-1} = e^{2\delta+2\gamma}h^2(\delta'^2 - \gamma'' - 2\gamma'^2) + h_{\varphi\varphi}h. \tag{75}$$

A tedious manipulation of Eqs. (48), (74), and (75) leads to the simple relations

$$\dot\gamma\gamma' = 0, \quad \dot\delta' = 0. \tag{76}$$

Therefore $\dot\gamma = 0$ or $\gamma' = 0$. The second possibility with $\gamma \neq$ const leads to an unphysical solution ($\rho < 0$) and the first one allows us to integrate easily the function δ, through Eqs. (74)–(76),

$$e^\delta = A(t)r + B(t), \tag{77}$$

where A and B are arbitrary functions.

For $A(t)\neq 0$, the variable r can be redefined and without loss of generality we can take $B\equiv 0$ (we use the property $\dot{B}A=\dot{A}B$ deduced from $\dot{\delta}'=0$).

By using Eq. (75), taking the derivative with respect to the variable r, we find the following expression for γ:

$$e^{2\gamma}=ar^{-1}+br^2+c, \tag{78}$$

where a, b, c are constants. Then h and A will be related by Eq. (75):

$$\ddot{A}A^{-1}-\dot{A}^2A^{-2}-\dot{A}A^{-1}h^{-1}\dot{h}=A^2ch^2+hh_{\varphi\varphi}. \tag{79}$$

The density will be given by

$$\rho=-e^{2\gamma}(\ddot{\gamma}+2\gamma'^2+2\gamma'\delta'). \tag{80}$$

Developing this expression, we easily prove that the density is constant, $\rho=-3b$, and using Eq. (19) we obtain an equation of state for the fluid of the form $p=-\rho$. This equation is unphysical and can be again interpreted as a vacuum solution with $\Lambda\neq 0$. The metric can be written as

$$ds^2=-h^2A^2r^2\,dt^2+(ar^{-1}+br^2+c)^{-1}\,dr^2+(ar^{-1}+br^2+c)d\theta^2+A^2r^2\,d\varphi^2. \tag{81}$$

If $\rho=0$, the constant b vanishes and we obtain a vacuum solution with $\Lambda=0$. If we compute the Riemann tensor, we find that it is different from zero if $a\neq 0$, i.e., we have found a family of nonflat vacuum solutions. It admits, at least, one Killing vector ∂_θ as can be seen by simple inspection.

Summing up, two new families (as far as we know) of vacuum solutions emerge: the metric given by Eq. (81), being $A(t)$, an arbitrary function, and $h(t,\varphi)$, satisfying the differential equation expressed by Eq. (79). For $b\neq 0$ the cosmological Λ term needs to be incorporated.

Next, we shall briefly discuss the case where $A(t)\equiv 0$. The integration of Eq. (75) leads to

$$e^{2\gamma}=ar^2+br+c,\quad hh_{\varphi\varphi}-h^{-1}\dot{h}\dot{\delta}=\ddot{\delta}+ae^{2\delta}, \tag{82}$$

where a, b, and c are constants. Now, a direct calculation of ρ and p through Eqs. (9) and (19) gives $\rho=-a$ and $p=-\rho$. The particular case $a\equiv 0$ is flat space-time and for $a\neq 0$, we can reinterpret the solution as corresponding to Einstein's equations with a Λ term.

Subcase B_4 ($\alpha_\theta=0$, $\dot{\delta}+\dot{\gamma}=0$): Now, we can write the lapse function as

$$\alpha=f(t,r)+h(t,\varphi)e^{\delta}. \tag{83}$$

According to Eq. (48) two possibilities emerge: $\dot{\gamma}=0$ or $\gamma'=0$. If $\gamma'=0$, it is easily proven that this generates a solution with $\delta=\gamma$, so we will only consider the subcase $\dot{\gamma}=0$ ($\Rightarrow\dot{\delta}=0$).

Developing Eqs. (19)–(21) and combining them in an appropriate way, we find

$$he^{\delta}(\delta''+\delta'^2)=-\tfrac{1}{2}[f''+f(\delta''+\delta'^2)], \tag{84}$$

$$-e^{2\gamma+\delta}h(\delta''+\gamma''+2\gamma'^2)+e^{-\delta}h_{\varphi\varphi}+e^{2\gamma}[-f''+(\delta'-\gamma')f'-f(\gamma''+2\gamma'^2-\delta'\gamma')]=0; \tag{85}$$

integrating the first equation we obtain

$$\delta=\ln(dr+e),\quad f=Ar+B, \tag{86}$$

where d and e are constants and $A(t)$ and $B(t)$ are arbitrary functions of their argument.

For $d\neq 0$, the variable r can be redefined and without loss of generality we can take $e=0$ and $d=1$, i.e., $e^{\delta}=r$.

Taking the derivative with respect to φ and r in Eq. (85), we can solve for γ as

$$e^{2\gamma}=ar^{-1}+br^2+c, \tag{87}$$

where a, b, and c are arbitrary constants. We realize that when $B=0$ this solution will be a particular case of solution (81) with $A=1$, because we can reexpress α as $\bar{h}(t,\varphi)e^{\delta}$ with $\bar{h}=A+h$.

If we analyze the case $B\neq 0$, we find that it has the same structure as the metrics found before in the subcase B_1 with $\rho=\text{const}$, $p=-(\rho+2Wk\alpha^{-1})$, and $g_\theta=0$.

For $d\equiv 0$ the only vacuum solution is obtained if we have flat space-time and for nonvacuum we obtain nonphysical solutions with $\rho=0$ and $p\neq 0$.

In conclusion, we have obtained all the possible solutions for a metric written in the form (1) provided that $\alpha_\varphi\neq 0$ and $\delta\neq\gamma$. We have found a family of solutions representing a perfect fluid without expansion, irrotational and shear-free, satisfying a nonbarotropic equation of state. This solution cannot be considered as a cosmological model, since the natural physical interpretation of perfect-fluid solutions with zero expansion is as models of perfect fluids in equilibrium.

We have also found several families of vacuum solutions with a cosmological constant $\Lambda\neq 0$ and a non-Minkowskian vacuum solution with $\Lambda=0$.

If $\alpha_\varphi=0$, $\alpha_\theta\neq 0$, and $M_\theta=0$, interchanging θ and φ, we easily see that this case is equivalent to the one (case B) considered before. So, we have found the general solution for the metric (1) with $M_\theta=0$, α_θ, or $\alpha_\varphi\neq 0$ and $\delta\neq\gamma$. This type of metric admits two "intrinsic Killing vectors" on the hypersurfaces $t=\text{const}$, but these vectors will not be Killing on the whole space-time, due to the θ or φ dependence of α.

III. PARTICULAR SOLUTIONS WITH $\alpha_\varphi\equiv 0$

In this section, we will study the case $\alpha_\varphi=0$, $\alpha_\theta\neq 0$, and $M_\theta\neq 0$. In the general case, the equations are too complicated and no general solution is obtained. We will only obtain some particular solutions, assuming that γ and δ are only time dependent. This assumption simplifies the equations, allowing us to integrate them.

Equation (16) is written now as $\alpha'_\theta=0$. On the other hand, Eq. (10) is equivalent to $\alpha'(\dot{\gamma}+\dot{\delta})=0$. It is very easy to prove, taking into account Eq. (11), that the possibility $\dot{\gamma}+\dot{\delta}=0$ leads, after a redefinition of the coordinates, to a particular case of the metrics studied in Sec. II. Therefore, we shall consider that $\alpha'=0$ and the integration of Eq. (11) leads to

$$\alpha=G(t)M(\theta)^{F(t)},\quad F(t)=(\dot{\delta}-\dot{\gamma})(\dot{\delta}+\dot{\beta})^{-1},\ \beta(t), \tag{88}$$

after a redefinition of the coordinate r.

The only remaining equations to be solved are Eqs. (9) and (19)–(21). The difference between Eqs. (19) and (20) and (19) and (21) can be written as

$$\ddot{\gamma}-\ddot{\beta}+(\dot{\gamma}-\dot{\beta})(\dot{\beta}+\dot{\gamma}+\dot{\delta}-\dot{F}\ln M-\dot{G}G^{-1})$$
$$=G^2e^{-2\gamma}M^{2F-2}[(1+F)MM_{\theta\theta}+F(F-1)M_\theta^2], \tag{89}$$

$$\ddot{\delta}-\ddot{\beta}+(\dot{\delta}-\dot{\beta})(\dot{\beta}+\dot{\gamma}+\dot{\delta}-\dot{F}\ln M-\dot{G}G^{-1})$$
$$=G^2e^{-2\gamma}M^{2F-2}(MM_{\theta\theta}+FM_\theta^2), \tag{90}$$

and if we subtract Eq. (89) minus $(1+F)$ times Eq. (90) and take two derivatives with respect to the variable θ and one derivative with respect to t, we obtain

$$c^{-1}=F=\text{const}\neq 0, \quad M=\theta^c, \tag{91}$$

after a redefinition of the variable θ. Moreover, Eq. (88) can be rewritten as

$$\alpha=G(t)\theta, \quad c(\dot{\delta}-\dot{\gamma})=\dot{\delta}+\dot{\beta}. \tag{92}$$

The particular case $c=1$ leads either to an unphysical solution ($\rho<0$) or to a flat space-time.

For $c\neq 1$ it is more convenient to combine algebraically Eqs. (89) and (90), taking into account Eq. (92), in the following way:

$$\ddot{\beta}+\dot{\beta}(\dot{\beta}+\dot{\gamma}+\dot{\delta}-\dot{G}G^{-1})=0, \tag{93}$$
$$\ddot{\gamma}+\ddot{\gamma}(\dot{\beta}+\dot{\gamma}+\dot{\delta}-\dot{G}G^{-1})=e^{-2\gamma}G^2c(c-1). \tag{94}$$

We will consider two different cases. The first one appears for $\dot{\beta}=0$ and we can always take, without loss of generality, $\gamma=t$, because $\dot{\gamma}\neq 0$ [see Eq. (92)]. This last equation can be integrated in the form $\delta=c(c-1)^{-1}t$ and from Eq. (94) we obtain the general solution for G,

$$G=[(c-1)^2e^{-2t}+ke^{(2-4c)(c-1)^{-1}t}]^{-1/2}, \quad k=\text{const}. \tag{95}$$

For $k\neq 0$, calculating the density and pressure [from Eqs. (9) and (19)], interchanging the variables $r\leftrightarrow\theta$, and doing a time translation, we find

$$\rho=e^{(2-4c)(c-1)^{-1}t}r^{-2}, \quad p=\rho, \tag{96}$$

so, the fluid satisfies a stiff equation of state and the metric reads

$$ds^2=-[(c-1)^2e^{-2t}+(1-c^{-1})e^{(2-4c)(c-1)^{-1}t}]^{-1}r^2\,dt^2$$
$$+e^{2t}\,dr^2+d\theta^2+e^{2(1-c^{-1})^{-1}t}r^{2c}\,d\varphi^2. \tag{97}$$

For $k=0$, we obtain a vacuum solution that is Minkowski space-time. The metric (97) admits two orthogonal Killing vectors ∂_θ and ∂_φ. The new solution (as far as we know) is inhomogeneous as regards the definition given by Wainwright[8] and it belongs to the class Bii apparently in his classification scheme.

Now, let us consider the case $\dot{\beta}\neq 0$. We can always choose $\beta=t$ and integrating Eqs. (92)–(94) the following is obtained:

$$G=ae^{[(2c-1)\gamma+ct](c-1)^{-1}}, \tag{98}$$
$$\delta=(c\gamma+t)(c-1)^{-1}, \tag{99}$$
$$\dot{y}^2-4a^2c^2e^y=k, \quad y\equiv 2c(\gamma+t)(c-1)^{-1}, \tag{100}$$

where a and k are arbitrary constants.

The case $k=0$ will not be considered because it leads to an unphysical solution ($\rho<0$).

On the one hand, for $k\neq 0$, Eq. (19) implies a stiff equation of state, i.e.,

$$p=\rho. \tag{101}$$

On the other hand, for $k>0$ the integration of Eq. (100) leads to the metric—after redefining the variables (t,θ,φ)—

$$ds^2=-c^{-2}[t(t-1)^{-2}]^{2-c^{-1}}t^{s-2}\theta^2dt^2$$
$$+[t(t-1)^{-2}]^{1-c^{-1}}t^s\,d\theta^2$$
$$+t^{-s}\,dr^2+[t(t-1)^{-2}]t^s\theta^{2c}\,d\varphi^2, \tag{102}$$

where s is an arbitrary constant. A direct calculation of ρ, through Eq. (9), gives

$$\rho=\tfrac{1}{4}c(c-1-cs^2)t^{-s}[t(t-1)^{-2}]^{c^{-1}-2}\theta^{-2}. \tag{103}$$

It is interesting to remark that a single change $r\leftrightarrow\theta$ and a redefinition of the variables $(r\to e^r,\ t\to[\coth(\tfrac{1}{2}ct)]^2)$ allows us to write the solution (102) in a form which is a particular case of a family found by Wainwright *et al.*[9] ($n=\pm 1$, $m=\mp 2s$, $\beta=0$, $c=-2q$, and $\tfrac{1}{4}m^2+\tfrac{1}{2}\alpha^2=1-c^{-1}$).

For $k<0$ the integration of Eq. (100) leads to (after a redefinition of the variables)

$$ds^2=-c^{-2}e^{2st}(\cos t)^{2c^{-1}-4}\theta^2\,dt^2+e^{-2st}\,dr^2$$
$$+e^{2st}(\cos t)^{2c^{-1}-2}d\theta^2+e^{2st}(\cos t)^{-2}\theta^{2c}\,d\varphi^2, \tag{104}$$

where s is an arbitrary constant. A direct calculation of ρ [see Eq. (9)] gives

$$\rho=\theta^{-2}(\cos t)^{4-2c^{-1}}e^{-st}c(1-c-cs^2). \tag{105}$$

The particular relation $s^2=c^{-1}-1$ corresponds to a nonflat vacuum solution with the metric given by [see Eq. (104)]

$$ds^2=-(1+s^2)^2e^{2st}(\cos t)^{2s^2-2}\theta^2\,dt^2+e^{-2st}\,dr^2$$
$$+e^{2st}(\cos t)^{2s^2}\,d\theta^2+e^{2st}(\cos t)^{-2}\theta^{2(1+s^2)^{-1}}d\varphi^2, \tag{106}$$

where s is an arbitrary constant.

For $s^2<c^{-1}-1$ the metric (104) represents a new (as far as we know) physical solution of Einstein's equations, apparently inhomogeneous, corresponding to stiff matter.

Taking into account Eqs. (25)–(27), we can easily calculate the kinematical quantities corresponding to the new metrics (97) and (104). These quantities read, respectively,

$$\theta=(2c-1)(c-1)^{-1}\alpha^{-1},$$
$$\sigma=3^{-1/2}[c(c-1)+1]^{1/2}(2c-1)^{-1}\alpha^{-1}, \quad \xi=(e^tr)^{-1}, \tag{107}$$

$$\theta=[s+(2c-1)c^{-1}\tan t\,]\alpha^{-1},$$
$$\xi=e^{-st}(\cos^2 t)^{1-c^{-1}}r^{-1}, \tag{108}$$
$$\sigma=[4s^2+(4-2c^{-1})s\tan t+(c-1+c^{-1})\tan^2 t\,]^{1/2}\alpha^{-1}.$$

We comment that the relative distortion σ/θ is a constant for the metric (97) according to (107).

IV. CONCLUSIONS

We have surveyed the solutions of Einstein's field equations for a metric written in the form (1), which admits one translational Killing vector on the hypersurfaces $t=\text{const}$ and we have assumed a perfect fluid as matter content.

The general solution is obtained for the case $\alpha_\varphi\neq 0$ and several different families emerge. These families can repre-

sent either a perfect fluid without expansion and shear-free [Eq. (70)] or a vacuum with a cosmological constant Λ [Eqs. (43)–(45), (81), and (82)]. In the former case the fluid obeys a nonbarotropic equation of state and satisfies the standard energy condition $\rho > 0$. All the solutions are apparently inhomogeneous.

We have found some particular solutions in the case $\alpha_\varphi = 0$, $\alpha_\theta \neq 0$ [Eqs. (97), (102), (105), and (106)]. All the solutions are apparently inhomogeneous (i.e., admit a group G_2 of isometries) and the fluid satisfies in this case a stiff matter equation of state with expansion and shear different from zero. Equation (106) represents a vacuum solution. We comment that the metric given by Eq. (102) is, in fact, a particular case of a family of inhomogeneous solutions given by Wainwright *et al.*,[9] but all the other metrics found in this paper are new physical solutions as far as we know.

Finally, we would like to mention the "intrinsic symmetries" technique, which allows us to obtain inhomogeneous solutions by imposing certain structures on the spatial part of the metric.

[1]C. B. Collins, Gen. Relativ. Gravit. **10**, 925 (1979).
[2]C. B. Collins and D. A. Szafron, J. Math. Phys. **20**, 2347 (1979).
[3]A. Krasinski, Gen. Relativ. Gravit. **13**, 1021 (1981).
[4]E. Martinez and J. L. Sanz, J. Math. Phys. **26**, 785 (1985).
[5]G. F. R. Ellis, "Relatisvic cosmology," in *General Relativity and Cosmology, XLVII Enrico Fermi Summer School Proceedings,* edited by R. K. Sachs (Academic, New York, 1971).
[6]C. W. Misner, K. S. Thorne, and J. A. Wheeler, in *Gravitation* (Freeman, San Francisco, 1973).
[7]J. W. York, "The initial value problem and dynamics," in *Gravitational Radiation, Les Houches 1982,* edited by N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983).
[8]J. Wainwright, J. Phys. A: Math. Gen. **14**, 1131 (1981).
[9]J. Wainwright, W. C. Ince, and B. J. Marshman, Gen. Relativ. Gravit. **10**, 259 (1979).

Anisotropic cosmological model in Nordtvedt's scalar–tensor theory of gravitation

A. Banerjee and N. Banerjee
Department of Physics, Jadavpur University, Calcutta-700032, India
N. O. Santos
Laboratoire de Physique Théorique, Institut Henri Poincaré, 11, rue Pierre et Marie Curie, Paris, France

(Received 26 March 1985; accepted for publication 31 May 1985)

Anisotropic cosmological models are considered in the light of the scalar–tensor theory of gravitation proposed by Nordtvedt. Special attention is paid to Bianchi type I models. The models consist of perfect fluid with the equation of state $p = \epsilon\rho$. The solutions are obtained in Dicke's conformally transformed units for empty space, as well as for $\epsilon = 1$ and $\frac{1}{3}$, assuming two separate functional relationships between ω and ϕ. Their properties are also compared with those of the models given in Brans–Dicke theory.

I. INTRODUCTION

In view of the recently verified large magnitude of the constant parameter ω in Brans–Dicke[1] (BD) theory it is argued that the experimental results in this theory should not differ much from those in Einstein's theory. In this background, the generalization of BD theory with the parameter ω as a variable quantity has of late drawn much attention. It has been claimed by Nordtvedt[2] that an accurate light deflection experiment and also the data on the rate of advance of perihelion of mercury could require $\omega' \neq 0$. Static solutions in this theory have been considered by Banerjee and Duttachoudhury,[3] Vanden Bergh,[4] and Rao and Reddy.[5] Barker[6] argued that there was no *a priori* reason to exclude the introduction of a long-range scalar field in the evolution of the universe with the possibility of ω being small at some stages of the evolution, making the results differ appreciably from those in Einstein's theory. Cosmological solutions in this theory for isotropic and anisotropic models were studied later by Vanden Bergh,[7] Bishop,[8] and Banerjee and Santos.[9,10]

In the present paper a homogeneous anisotropic model such as a Bianchi type I model is being reviewed in the background of the generalized scalar–tensor theory of Nordtvedt. The calculations are comparatively simple in Dicke's revised units,[11] which is the conformally transformed version of the original scalar–tensor theory of gravitation,[1] where the so-called gravitational constant is variable. In these units the standard Einstein's equations are satisfied. The scalar field plays the role of an additional material source, which, however, may formally be said to constitute a stiff fluid with density equal to the pressure. In this theory the equation of motion of a test particle is nongeodesic with varying rest mass. The scalar field is to be found from a separate equation and the knowledge of the exact solution for the scalar field ϕ is necessary to obtain solution in the original version with varying gravitational constant. It is also shown that for a homogeneous universe in the absence of rotation and for $(2\omega + 3) > 0$, the Raychaudhury equation leads to the appearence of the singularity in the model. This is because Hawking's energy condition is satisfied for $(2\omega + 3) > 0$ irrespective of whether ω is a constant as in Brans–Dicke theory (Raychaudhuri[12]) or ω is a function of the scalar field as in Nordtvedt's theory. The contribution of the scalar field to the energy density in some cases becomes even more dominant than the matter part near the singularity.

In the following sections we have made some general observation on the properties of homogeneous universes in Nordtvedt's scalar–tensor theory and then obtained exact solutons for Bianchi type I models in matter-free space, stiff fluid, and radiation. All these cases are discussed for two different choices of ω—one being Schwinger's relation (see Vanden Bergh[4]) and the other that of Barker. The results are analyzed in the background of those existing in Brans–Dicke theory. Particularly in the stiff fluid case ($p = \rho$) it is observed that in Nordtvedt's theory, with the special choice of ω either due to Schwinger or to Barker, the matter density under no circumstances remains finite as the singularity of zero volume approaches. This result is unlike the solution in Brans–Dicke theory with $\omega = \text{const}$.

In Sec. II we write the field equations in Dicke's revised units and discuss the behavior of the homogeneous anisotropic models in general. In Sec. III we integrate the field equations for a Bianchi type I model and analyze the solutions.

II. FIELD EQUATIONS

The field equations for Nordtvedt's scalar–tensor theory in the revised units of Dicke are

$$G_{\alpha\beta} \equiv R_{\alpha\beta} - \tfrac{1}{2} R g_{\alpha\beta} = -kT_{\alpha\beta} - \frac{(2\omega + 3)}{2} \frac{1}{\phi^2} (\phi_{,\alpha}\phi_{,\beta} - \tfrac{1}{2} g_{\alpha\beta}\phi^{,\mu}\phi_{,\mu}) . \tag{2.1}$$

The corresponding wave equation for the scalar field ϕ can be written as

 0022-2488/85/123125-06$02.50

$$\Box(\ln\phi)\equiv(\ln\phi)^{;\mu}{}_{;\mu}$$

$$=\frac{1}{(2\omega+3)}\left[kT-\frac{1}{\phi}\phi^{,\mu}\phi_{,\mu}\frac{d\omega}{d\phi}\right]. \tag{2.2}$$

In the above, the constant k stands for the usual $8\pi G_0$ with, however, the choice $c=1$. It is interesting to note that the field equations look like those in Einstein's theory with G_0 being a universal constant. Here $T_{\alpha\beta}$, representing the energy momentum tensor of a perfect fluid, is given by [with signature $(-+++)$]

$$T_{\alpha\beta}=(\rho+p)v_\alpha v_\beta+pg_{\alpha\beta}\,, \tag{2.3}$$

with the four-velocity satisfying, in comoving coordinates, the following conditions:

$$v^\mu v_\mu=-1 \quad\text{and}\quad v^\mu=\delta^\mu_0\,. \tag{2.4}$$

The ρ and p in (2.3) represent the fluid density and the pressure, respectively.

In view of the Bianchi identity, the divergence of the field equation (2.1) yields

$$kT^{\beta}_{\alpha;\beta}+\frac{\phi_{,\alpha}}{\phi}\left[\frac{1}{2}\phi_{,\beta}\phi^{,\beta}\frac{d\omega}{d\phi}+\frac{(2\omega+3)}{2}\left(\Box\phi-\frac{1}{\phi}\phi_{,\beta}\phi^{,\beta}\right)\right]=0\,, \tag{2.5}$$

which, with the use of the wave equation, yields

$$k\left[(\rho+p)_{,\beta}v^\beta v_\alpha+(\rho+p)\dot v_\alpha+(\rho+p)\theta v_\alpha+p_{,\alpha}+\tfrac{1}{2}(\phi_{,\alpha}/\phi)T\right]=0\,, \tag{2.6}$$

where θ stands for the usual expansion scalar $v^\beta{}_{;\beta}$ and $\dot v_\alpha=v_{\alpha;\beta}v^\beta$ is the acceleration vector. For a rotation-free spatially homogeneous space-time, the vorticity scalar $\omega=0$ and also

$$p_{,\mu}=av_\mu,\quad \rho_{,\mu}=bv_\mu,\quad \phi_{,\mu}=cv_\mu\,,$$

where a, b, and c are scalar functions and v_μ is the velocity vector, which is hypersurface orthogonal (see Raychaudhuri[12]). From (2.6) it immediately follows that $\dot v_\mu\propto v_\mu$ and since $\dot v^\mu v_\mu=0$ and $v^\mu v_\mu=-1$ along with $v^\mu=\delta^\mu_0$ for a comoving system, we have the acceleration vector $\dot v_\mu$ vanishing. The Raychaudhuri equation,[13] therefore, reduces for a homogeneous nonrotating universe to

$$\theta_{,\mu}v^\mu=-2\sigma^2-\tfrac{1}{3}\theta^2+R_{\mu\nu}v^\mu v^\nu\,, \tag{2.7}$$

where σ is the shear scalar. Now since we have written $\phi_{,\mu}=cv_\mu$ it follows that $\phi^{,\mu}\phi_{,\mu}=-c^2$ and in view of (2.1) one can write

$$\begin{aligned}R_{\mu\nu}-\tfrac{1}{2}Rg_{\mu\nu}&=-k\left[(\rho+p)v_\mu v_\nu+pg_{\mu\nu}+[(2\omega+3)/2k\phi^2](c^2v_\mu v_\nu+\tfrac{1}{2}g_{\mu\nu}c^2)\right]\\&=-k\left[(\rho+\rho_\phi+p+p_\phi)v_\mu v_\nu+g_{\mu\nu}(p+p_\phi)\right],\end{aligned} \tag{2.8}$$

where

$$\rho_\phi=p_\phi=\frac{(2\omega+3)}{4k\phi^2}c^2=-\frac{(2\omega+3)}{4k}\frac{\phi_{,\mu}\phi^{,\mu}}{\phi^2}\,. \tag{2.9}$$

From (2.8) it is not difficult to get the following relation:

$$R_{\mu\nu}v^\mu v^\nu=-(k/2)(\rho+\rho_\phi+3p+3p_\phi)\,. \tag{2.10}$$

For $(2\omega+3)>0$, ρ_ϕ and p_ϕ are both greater than zero and thus, along with the conditions that the fluid density and pressure are both positive, we have

$$R_{\mu\nu}v^\mu v^\nu<0\,,$$

that is, Hawking's energy condition[14] is satisfied. It follows from (2.7) that

$$\theta_{,\mu}v^\mu<0\,,$$

which, in comoving coordinates, reduces to the condition $\dot\theta<0$. It means that there is no minimum volume and one cannot avoid the singularity. These results will be apparent in specific models discussed in the following sections.

Before going onto the exact solutions in some special cases we make a few general observations about the nature of the solutions. The splitting of the energy momentum tensor into two different parts—one due to the perfect fluid and the other due to the scalar field in the field equations (2.8)—enables us to write the effective density $\bar\rho$ and the effective pressure $\bar p$ as (see Ruban and Finkelstein[15])

$$\bar\rho=\rho+\rho_\phi\quad\text{and}\quad\bar p=p+p_\phi\,.$$

Now, ρ_ϕ and p_ϕ being equal in magnitude, the solutions for the metric in otherwise empty space will be formally identical with the stiff fluid solutions up to a solution for the scalar field itself from the wave equation for the scalar field. The $p=\rho$ solutions in the scalar–tensor theory, being considered here in this paper, may also formally be identical with the stiff fluid solutions in general relativity. The scalar field constituting ρ_ϕ and p_ϕ does, however, satisfy a separate field equation. It is shown later in special cases that near the singularity the fluid density ρ and the energy density ρ_ϕ due to the scalar field exhibit different time behavior. Moreover, the simple solutions, which are obtained in Dicke's revised units in Einstein's framework, pass over to the original Nordtvedt or Brans–Dicke solutions in atomic units by conformal transformations involving the scalar field, that is, $g_{\mu\nu}=(1/\phi)\bar g_{\mu\nu}$. These are more natural with the principle of geodesic motion being fulfilled ($m_0=\text{const}$). When $\bar\rho$ and $\bar p$—the effective density and pressure—are treated as unknown along with the metric components and if there is a sufficient number of field equations to determine them, solutions can be obtained independently of any particular choice of ω as a function of the scalar field ϕ. But this specific relationship is required to be known explicitly before finding the solution for ϕ and thus in consequence to obtain solutions in original atomic units where G, the so-called gravitational constant, is no longer a constant. Solutions are more complicated in the original theory but can be generated from the solution obtained in the following sections when the explicit solutions of ϕ for different functional forms of ω are known.

We will now show that in general for all spatially homogeneous nonrotating universes where the trace of the energy momentum tensor for matter vanishes, the expressions for ρ_ϕ and p_ϕ are independent of any particular choice of the parameter ω as a function of ϕ in Nordtvedt's theory. When rotation vanishes, the homogeneous varieties are orthogonal to "t" lines in a suitable system and the metric can be written as

$$ds^2=-dt^2+g_{ij}\,dx^i\,dx^j\,, \tag{2.11}$$

where $i, j = 1, 2, 3$. Thus from (2.2) one gets, in view of the spatial homogeneity, the relation

$$-\left[\frac{\dot{\phi}}{\phi} l^3\right]_{,0} = \frac{l^3}{(2\omega+3)}\left[kT + \frac{\dot{\phi}^2}{\phi}\frac{d\omega}{d\phi}\right], \tag{2.12}$$

where l^3 stands for $\sqrt{-g}$. Now, in general, for homogeneous universes with ρ, p both constants on the homogeneous varieties, one can express l as $l = R(t)\, W(x^i)$, where the x^i's stand for space coordinates only (Banerji[16]) and so (2.12) reduces to

$$\left(\frac{\dot{\phi}}{\phi}\right)R^3 = -\int \frac{kT}{(2\omega+3)} R^3\, dt - \int \frac{(\dot{\phi}/\phi)}{(2\omega+3)}\frac{d\omega}{dt} R^3\, dt. \tag{2.13}$$

When the energy momentum tensor is traceless we get the relation

$$\frac{\dot{\phi}}{\phi} R^3 = -\int \frac{R^3(\dot{\phi}/\phi)}{(2\omega+3)} d\omega,$$

which also can be written in the form

$$\frac{d}{d\omega}\left(\frac{\dot{\phi}}{\phi} R^3\right) = -\frac{1}{(2\omega+3)}\left(\frac{\dot{\phi}}{\phi} R^3\right). \tag{2.14}$$

Integration of (2.14) yields

$$((\dot{\phi}/\phi)\, R^3)(2\omega+3)^{1/2} = \text{const},$$

and, using this, one finally arrives at the result

$$\rho_\phi = p_\phi = \frac{(2\omega+3)}{4k}\left(\frac{\dot{\phi}}{\phi}\right)^2 = \frac{A}{R^6}, \tag{2.15}$$

where A is an arbitrary constant. The expression (2.15) is therefore valid for matter-free space ($\rho = p = 0$) and also for radiation ($\rho = 3p$). Equation (2.13) cannot, in general ($T \neq 0$), be integrated without any knowledge about the integrand as a function of time. We will see that the same relation (2.15) appears in the special case of the Bianchi I model discussed in what follows in appropriate situations.

III. INTEGRATION OF THE FIELD EQUATIONS AND BEHAVIOR OF THE MODELS IN A BIANCHI TYPE I UNIVERSE

The metric for the Bianchi type I homogeneous cosmological model is

$$ds^2 = -dt^2 + e^{2\gamma}dx^2 + e^{2\theta}dy^2 + e^{2\psi}dz^2, \tag{3.1}$$

where γ, θ, and ψ are functions of "t" alone. The nontrivial field equations, according to (2.1) and (2.2), will be

$$G_0^0 = \tfrac{3}{2}(\dot{R}/R)^2 - \tfrac{1}{2}(\dot{\gamma}^2 + \dot{\theta}^2 + \dot{\psi}^2)$$
$$= k\rho + [(2\omega+3)/4](\dot{\phi}/\phi)^2, \tag{3.2a}$$

$$G_1^1 = \ddot{\theta} + \ddot{\psi} + \tfrac{3}{2}(\dot{R}/R)(\dot{\psi} + \dot{\theta} - \dot{\gamma}) + \tfrac{1}{2}(\dot{\gamma}^2 + \dot{\theta}^2 + \dot{\psi}^2)$$
$$= -kp - [(2\omega+3)/4](\dot{\phi}/\phi)^2, \tag{3.2b}$$

$$G_2^2 = \ddot{\gamma} + \ddot{\psi} + \tfrac{3}{2}(\dot{R}/R)(\dot{\gamma} + \dot{\psi} - \dot{\theta}) + \tfrac{1}{2}(\dot{\gamma}^2 + \dot{\theta}^2 + \dot{\psi}^2)$$
$$= -kp - [(2\omega+3)/4](\dot{\phi}/\phi)^2, \tag{3.2c}$$

$$G_3^3 = \ddot{\gamma} + \ddot{\theta} + \tfrac{3}{2}(\dot{R}/R)(\dot{\gamma} + \dot{\theta} - \dot{\psi}) + \tfrac{1}{2}(\dot{\gamma}^2 + \dot{\theta}^2 + \dot{\psi}^2)$$
$$= -kp - [(2\omega+3)/4](\dot{\phi}/\phi)^2, \tag{3.2d}$$

$$\Box(\ln\phi) = -(ln\,\phi)^{\cdot\cdot} - (3\dot{R}/R)(\ln\phi)^{\cdot}$$
$$= \frac{1}{(2\omega+3)}\left[k(3p-\rho) + \frac{\dot{\phi}^2}{\phi}\frac{d\omega}{d\phi}\right], \tag{3.3}$$

with $R^3 = \exp(\gamma + \theta + \psi)$.

The divergence relation (2.5) will yield for $\alpha = 0$ the relation

$$(\rho R^3)^{\cdot} + 3pR^2\dot{R} + \tfrac{1}{2}(\dot{\phi}/\phi)R^3(\rho - 3p) = 0. \tag{3.4}$$

It is interesting to note that relation (3.4) in Nordtvedt's theory is identical with the relation obtained in the Brans–Dicke theory where $\omega = \text{const}$ (see Raychaudhuri[12]).

We have five independent equations (3.2) and (3.3) and seven unknowns $\gamma, \theta, \psi, \rho, p, \phi$, and ω. In what follows, we assume two relations, an equation of state relating p with ρ and another equation connecting ω and ϕ.

From (3.2b) and (3.2d), subtracting one from the other and integrating, we obtain

$$\dot{\theta} - \dot{\gamma} = C_1/R^3, \quad \dot{\psi} - \dot{\theta} = C_2/R^3, \quad \dot{\gamma} - \dot{\psi} = C_3/R^3, \tag{3.5}$$

where C_1, C_2, C_3 are constants satisfying the condition $C_1 + C_2 + C_3 = 0$.

Now we proceed to study the following different cases.

A. Case I: $p = \rho = 0$

This is a case of empty space, for which Eq. (3.3) reduces to

$$\frac{\ddot{\phi}}{\dot{\phi}} - \frac{\dot{\phi}}{\phi} + \frac{3\dot{R}}{R} + \tfrac{1}{2}\left(\frac{2\dot{\omega}}{2\omega+3}\right) = 0,$$

integration of which yields immediately

$$(\dot{\phi}/\phi)(2\omega+3)^{1/2} = D_1/R^3, \tag{3.6}$$

where D_1 is a constant of integration. The same relation was obtained in Sec. II. Equation (3.6) is valid independent of the nature of ω, including the Brans–Dicke case where $\omega = \text{const}$.

Adding (3.2a), (3.2c), and (3.2d) and subtracting (3.2b) from the result we get

$$2\ddot{\gamma} + \tfrac{3}{2}(\dot{R}/R)(3\dot{R}/R + 3\dot{\gamma} - \dot{\theta} - \dot{\psi}) = 0,$$

i.e.,

$$\ddot{\gamma} + 3(\dot{R}/R)\dot{\gamma} = 0,$$

which integrates to yield

$$\dot{\gamma} = a_1/R^3. \tag{3.7a}$$

From different combination of the field equations, one obtains

$$\dot{\theta} = a_2/R^3, \tag{3.7b}$$

$$\dot{\psi} = a_3/R^3. \tag{3.7c}$$

In the above, a_1, a_2, a_3 are constants of integration. Adding (3.2a) and (3.2b) and integrating, it is possible to obtain

$$\dot{\gamma} = 3\dot{R}/R + D_2/R^3, \tag{3.8}$$

where D_2 is another constant of integration. Substituting a_1/R^3 for $\dot{\gamma}$ in Eq. (3.8) we obtain

$$3\dot{R}/R = (a_1 - D_2)/R^3,$$

which in turn on integration leads us, after a suitable choice

for the time origin, to the solution

$$R^3 = D_3 t\,, \tag{3.9}$$

where $D_3 = a_1 - D_2$.

The solutions for the scalar field ϕ can be obtained from (3.6) once $\omega(\phi)$ is known as an explicit function of ϕ. The solution in the BD case, that is, for constant magnitude of ω, is simple and straightforward. In the following we examine the situation for two different functional forms for ω, mentioned previously.

(a) Schwinger's relation: (See Vanden Bergh.[4]) Here $(2\omega + 3) = 1/\alpha\phi$, α being a constant quantity. The general solution is, in view of (3.6),

$$\phi^{-1/2} = -(D_1/D_3)(\alpha^{1/2}/2)[3 \ln R + a]\,, \tag{3.10}$$

where a is a constant of integration.

(b) Barker's relation: In this case $(2\omega + 3) = 1/(\phi - 1)$, so that integration of (3.6) yields

$$(\phi - 1)^{1/2} = \tan[(D_1/2D_3)(3 \ln R + b)]\,, \tag{3.11}$$

b being another constant of integration. One should note that in Dicke's revised version the solutions for the metric tensor and the scalar field as given in (3.8)–(3.11) are quite simple. It is not difficult to go back to the original version of Brans–Dicke atomic units, where masses remain constant and G varies, and find the metric using the transformation relations

$$g_{\alpha\beta} = (1/\phi)\bar{g}_{\alpha\beta}\,,$$

where $\bar{g}_{\alpha\beta}$ and $g_{\alpha\beta}$ are metrics in revised units and the original atomic units, respectively. The solutions given above for ϕ in variable G theory can be shown to be identical with those given earlier by Banerjee and Santos.[9,10] When the spatial volume vanishes (that is, R is vanishingly small), the shear scalar σ is infinitely large. The geometric shear is defined in the usual way (see Banerjee and Santos[9]) and gives the measure of anisotropy. The effective energy density of the scalar field for the $\rho = p = 0$ case is found to be

$$\rho_\phi = ((2\omega + 3)/4k)(\dot{\phi}/\phi)^2\,, \tag{3.12}$$

which in view of (3.6) shows that $\rho_\phi \propto R^{-6}$. The energy density due to the scalar field is infinitely large when $R \to 0$ or at the initial epoch $t = 0$. Further from (3.9) we see that $\dot{R} \neq 0$ for any finite value of R and also $\ddot{R} < 0$ indicating that the singularity of zero proper volume cannot be eliminated in such models.

B. Case II: $p = \rho$

This particular case in BD theory was studied by Nariai[17] in atomic units and later by Raychaudhuri[18] in Dicke's revised units. It is interesting to investigate if in Nordtvedt's theory there is any distinct change in the behavior of the model.

Proceeding exactly in the same way as had been done for the case of empty space, we obtain

$$\dot{\gamma} = b_1/R^3\,, \tag{3.13a}$$

$$\dot{\theta} = b_2/R^3\,, \tag{3.13b}$$

$$\dot{\psi} = b_3/R^3\,, \tag{3.13c}$$

where b_1, b_2, and b_3 are constants of integration. Adding (3.2a) and (3.2b) we obtain after integration

$$\dot{\gamma} = 3\dot{R}/R + F_1/R^3\,, \tag{3.14}$$

F_1 being a constant of integration. Using (3.13a) and (3.14) we obtain after integration, with a suitable choice of the origin of the time coordinate,

$$R^3 = F_2 t\,, \tag{3.15}$$

where $F_2 = b_1 - F_1$. Using the relations (3.5) and (3.15) in (3.2a),

$$k\rho + [(2\omega + 3)/4](\dot{\phi}/\phi)^2 = F/3R^6\,, \tag{3.16}$$

where $F = F_2^2 - C_1^2 - C_3^2 - C_1 C_3$. Another relation connecting ρ, R, and ϕ is obtained by putting $p = \rho$ in (3.4) and integrating in the form

$$\rho R^6/\phi = \text{const.} \tag{3.17}$$

Now using (3.16) in (3.3) and putting $p = \rho$ one gets a relation

$$-(2\omega + 3)[(\ln\phi)^{\cdot\cdot} + (F_2/R^3)(\ln\phi)^{\cdot}]$$
$$= \left\{-\frac{(2\omega + 3)}{2} + \phi\frac{d\omega}{d\phi}\right\}\{(\ln\phi)^{\cdot}\}^2 + \frac{2F}{3R^6}\,. \tag{3.18}$$

When $\omega = \text{const}$, that is, in the Brans–Dicke theory, Eq. (3.18) yields on integration, using (3.15),

$$(\dot{\phi}/\phi)t = m(1 - t^m/\beta)/(1 + t^m/\beta)\,, \tag{3.19}$$

where β is an arbitrary constant appearing on integration, and from (3.16) one gets for the matter density

$$k\rho = \frac{F}{3R^6} - \frac{(2\omega + 3)}{4}\frac{m^2}{t^2}\left(\frac{1 - t^m/\beta}{1 + t^m/\beta}\right)^2,$$

where $m^2 = (4/3F_2^2)(F/(2\omega + 3))$. Taking $m > 0$ and allowing $t \to 0$ one gets a relation

$$k\rho = \frac{F}{3F_2^2 t^2} - \frac{(2\omega + 3)}{4}\frac{m^2}{t^2}\left[1 - \frac{4}{\beta}t^m + \cdots\right],$$

and it is clear from above that ρ remains finite even when the spatial volume at this state ($t \to 0$) vanishes if $m = 2$. The shear and the expansion scalar, however, attain infinitely large magnitudes. This feature of the solution was previously noted by Nariai and Raychaudhuri.

For positive values of m other than 2 we have both ρ and ρ_ϕ increasing infinitely (as $t \to 0$) as $1/t^2$ or $1/R^6$.

Now considering our problem in Nordtvedt's scalar–tensor theory, we make two choices for ω as functions of the scalar field—Schwinger's relation and Barker's relation—exactly in the previous manner and proceed to find solutions for the scalar field.

(a) For Schwinger's choice $(2\omega + 3) = 1/(\alpha\phi)$, the relation (3.18) reduces to

$$\frac{\ddot{\phi}}{\phi} - 2\left(\frac{\dot{\phi}}{\phi}\right)^2 + \frac{F_2}{R^3}\left(\frac{\dot{\phi}}{\phi}\right) = -\frac{2}{3}F_\alpha\frac{\phi}{R^6}\,. \tag{3.20}$$

Now defining $\tau = R^3 = F_2 t$ and $\mu = 1/\phi$ we can write (3.20) in the form

$$\frac{d^2\mu}{d\tau^2} + \frac{1}{\tau}\frac{d\mu}{d\tau} = \frac{a}{\tau^2}\,, \tag{3.21}$$

with $a = 2F_\alpha/3F_2^2$. Now writing y for $(d\mu/d\tau)$ Eq. (3.21) reduces to a linear first-order differential equation

$$\frac{dy}{d\tau} + \frac{1}{\tau}y = \frac{a}{\tau^2}\,,$$

the solution of which is given by

$$y = \frac{d\mu}{d\tau} = \frac{F_3}{\tau} + \frac{a}{\tau}\ln\tau,$$

which in turn on further integration yields

$$\phi^{-1} = F_3 \ln\tau + (a/2)(\ln\tau)^2 + F_4, \tag{3.22}$$

where F_4 is a constant of integration. Using (3.22) in (3.16) we have

$$k\rho = kp = F/3\tau^2 - (F_2^2/4\alpha\tau^2)(F_3 + a\ln\tau)^2 \times [(F_3 + (a/2)\ln\tau)\ln\tau + F_4]^{-1}. \tag{3.23}$$

In the limit as $t\to 0$ we have $\tau\to 0$ (that is, $R\to 0$) and it follows from (3.22) that the scalar field ϕ vanishes. The singularity exists at this limit because (3.23) shows that $\rho \sim 1/(\tau\ln\tau)^2$ which increases to an infinitely large magnitude as $\tau\to 0$. The situation is similar to that in the Brans–Dicke case ($\omega = \text{const}$) except for the situation that by a suitable choice of some constant parameters we can keep the mass density finite in the latter case even when the spatial volume vanishes (see Raychaudhuri[18]). In (3.23), for the special case $F_3^2 = 2aF_4$, the mass density is always zero and we get empty space. Again as for the energy density due to the scalar field we note that it is, in view of (3.16), equal to $(F/3k\tau^2 - \rho)$, which, following the previous analysis, goes to infinity (as $\tau\to 0$) as $1/\tau^2$. Thus near the singularity the scalar field energy dominates in comparison with the matter density.

(b) Barker's relation for ω and ϕ is $(2\omega + 3) = 1/(\phi - 1)$, so that the relation (3.13) is now

$$\frac{\ddot\phi}{\phi} - \left(\frac{\dot\phi}{\phi}\right)^2 + \frac{F_2}{R^3}\left(\frac{\dot\phi}{\phi}\right) = \left(\frac{\dot\phi}{\phi}\right)^2\left[\frac{1}{2} + \frac{\phi}{2(\phi-1)}\right] + \frac{2F}{3R^6}(\phi - 1). \tag{3.24}$$

The differential equation (3.24) can be solved proceeding in the following manner. We write $\tau = R^3 = F_2 t$ and $b = 2F/3F_2^2$ and define a new variable $\mu = (1/\phi - 1)$ so that Eq. (3.24) can be written, omitting a few intermediate steps, as

$$\mu\frac{d^2\mu}{d\tau^2} - \frac{1}{2}\left(\frac{d\mu}{d\tau}\right)^2 + \frac{\mu}{\tau}\left(\frac{d\mu}{d\tau}\right) = -\frac{b}{\tau^2}\mu^2. \tag{3.25}$$

Further transformations like $y^2 = \pm\mu\tau$ reduces (3.25) to a very simple form

$$\tau^2\frac{d^2y}{d\tau^2} + \left(\frac{1}{4} + \frac{b}{2}\right)y = 0, \tag{3.26}$$

which is a simplified version of the Euler differential equation. The solutions are the following:

$$y/\sqrt{\tau} = \begin{cases} A\cos\{(b/2)^{1/2}\ln\tau\} + B\sin\{(b/2)^{1/2}\ln\tau\}, \\ \quad b > 0, \\ A\tau^{(|b|/2)^{1/2}} + B\tau^{-(|b|/2)^{1/2}}, \quad b < 0, \\ A + B\ln\tau, \quad b = 0. \end{cases} \tag{3.27}$$

The constants A, B in three different solutions are completely independent and arbitrary. One can obtain the value for ϕ from (3.27) using the transformation relation $y^2/\tau = \pm\mu$. The positive or negative sign is to be taken according as ϕ is less than or greater than unity. In order that the total energy density $(\rho + \rho_\phi) \geqslant 0$, we must have, from (3.16), $F \geqslant 0$ and correspondingly $b \geqslant 0$. For $b \geqslant 0$ [that is, for $(\rho + \rho_\phi) > 0$], ϕ remains always finite because the sine and cosine functions in (3.27) are bounded in magnitude and for the same reason it is not difficult to show that both ρ and ρ_ϕ explode to infinity as $1/R^6$ as $R\to 0$. For $b = 0$, we have $(\rho + \rho_\phi) = 0$ and both have infinities of the same order as the spatial volume approaches zero except for their signs. For $(2\omega + 3) > 0$ in Barkar's theory, $\phi > 1$ and $\rho_\phi > 0$, so that the fluid density $\rho < 0$, whereas for $(2\omega + 3) < 0$ we have the reverse situation.

Lastly, in the stiff fluid case also, since from (3.15) we have $\ddot R < 0$ and $\dot R \neq 0$ for any finite magnitude of R, there is no lower bound and there exists a point of singularity.

From the above results it is clear that one can find the solution for the metric in the case of an empty space and also in the case of a stiff fluid without knowing the functional dependence of ω on ϕ. But the functional form of ω is necessary for obtaining the solutions for ϕ. Moreover, in the stiff fluid case, we observe that the behavior of the energy densities due to the scalar field and matter are different for different choices of ω. From the results obtained in this section we find that for Schwinger's choice of ω, $\rho \sim 1/(\tau\ln\tau)^2$, and $\rho_\phi \sim 1/\tau^2$ for $R\to 0$, that is, the scalar field dominates over matter near the singularity, whereas for Barker's choice, both ρ and ρ_ϕ explode to infinity as $1/\tau^2$ as $R\to 0$. These results are in accordance with the general discussions in the previous section.

C. Case III: $p = \frac{1}{3}\rho$

This is a radiation case. It is possible to integrate Eq. (3.3) directly and obtain a relation like

$$(\dot\phi/\phi)(2\omega + 3)^{1/2} = G_1/R^3, \tag{3.28}$$

G_1 being a constant of integration. This relation had already been obtained in Sec. II for a rotation-free homogeneous universe with $T = 0$. The trace of the field equations (3.2a)–(3.2d) give

$$2(\ddot\gamma + \ddot\theta + \ddot\psi) + (\dot\gamma^2 + \dot\theta^2 + \dot\psi^2) + \tfrac{9}{2}(\dot R/R)^2 + \tfrac{3}{2}(\dot R/R)(\dot\gamma + \dot\theta + \dot\psi) = -[(2\omega + 3)/2](\dot\phi/\phi)^2. \tag{3.29}$$

Using (3.5) and the fact that $R^3 = \exp(\gamma + \theta + \psi)$, Eq. (3.29) takes the form

$$\frac{6\ddot R}{R} + 3\left(\frac{\dot R}{R}\right)^2 + \frac{\frac{1}{2}G_1 + C_1^2 + C_3^2}{R^6} + 3\dot\gamma^2 + 2\dot\gamma\frac{(C_1 - C_3)}{R^3} = 0,$$

and consequently $\dot\gamma$ can be written as

$$\dot\gamma = \frac{1}{2}\left[-\frac{2}{3}\frac{(C_1 - C_3)}{R^3} \pm \left\{\frac{4}{9}\frac{(C_1 - C_3)^2}{R^6} - 8\frac{\ddot R}{R} - 4\left(\frac{\dot R}{R}\right)^2 - \frac{4G}{3R^6}\right\}^{1/2}\right]. \tag{3.30}$$

Again from (3.5) we have

$$\dot\gamma = \dot R/R - (C_1 - C_3)/3R^3. \tag{3.31}$$

Equating (3.30) and (3.31) one gets

$$\frac{\ddot{R}}{R}+\left(\frac{\dot{R}}{R}\right)^2+\frac{G-\frac{1}{3}(C_1-C_3)^2}{6R^6}=0\,, \tag{3.32}$$

where $G=\frac{1}{2}G_1^2+C_1^2+C_3^2$ and

$$G-\tfrac{1}{3}(C_1-C_3)^2=\tfrac{1}{2}G_1^2+\tfrac{2}{3}(C_1^2+C_3^2+C_1C_3)>0\,,$$

since it can be shown that $C_1^2+C_3^2+C_1C_3>0$. Now writing u for $\dot{R}$ and using derivatives with respect to R in place of time, the derivative (3.32) can be written as

$$\frac{d(u^2)}{dR}+\frac{2}{R}(u^2)+\tfrac{1}{3}\left[G-\tfrac{1}{3}(C_1-C_3)^2\right]\frac{1}{R^5}=0\,. \tag{3.33}$$

The general solution of this equation is

$$u^2=(\dot{R})^2=G_2/R^2+\tfrac{1}{6}\left[G-\tfrac{1}{3}(C_1-C_3)^2\right]R^4\,, \tag{3.34}$$

G_2 being an integration constant. Writing the symbol G_3 for $\frac{1}{6}\left[G-\frac{1}{3}(C_1-C_3)^2\right]$ and integrating (3.34) in the next step,

$$t+G_4=\begin{cases}\dfrac{1}{2G_2}\{R(G_2R^2+G_3)^{1/2}\}-\dfrac{G_3}{2G_2}\ln\left[RG_2^{1/2}\right. \\ \qquad\left.+(R^2G_2+G_3)^{1/2}\right]\,, \quad G_2>0\,, \\ \dfrac{1}{2G_2}\{R(G_2R^2+G_3)^{1/2}\}-\dfrac{G_3}{2G_2}\dfrac{1}{(-G_2)^{1/2}} \\ \qquad\times\arcsin\left[R(-G_2/G_3)^{1/2}\right]\,, \quad G_2<0\,,\end{cases} \tag{3.35}$$

G_4 being a constant of integration.

The solutions (3.35) give us $R(t)$ as an explicit function of time. Integration of Eq. (3.23) yields

$$\int(2\omega+3)^{1/2}\frac{d\phi}{\phi}=G_1\int\frac{dt}{R^3}\,.$$

Replacing the time variable by the variable R and utilizing (3.34) in the above relation we can write

$$\int(2\omega+3)^{1/2}\frac{d\phi}{\phi}=G_1\int\frac{dR}{R(G_2R^2+G_3)^{1/2}}\,. \tag{3.36}$$

The integral on the right-hand side of (3.36) is different in two different cases $G_2>0$ and $G_2<0$. After integration (3.36) yields

$$\int(2\omega+3)^{1/2}\frac{d\phi}{\phi}=\begin{cases}\dfrac{G_1}{2(G_3)^{1/2}}\ln\dfrac{(G_2R^2+G_3)^{1/2}-G_3^{1/2}}{(G_2R^2+G_3)^{1/2}+G_3^{1/2}}\,, & \text{for } G_2>0\,, \\ \dfrac{G_1}{2(G_3)^{1/2}}\ln\dfrac{G_3^{1/2}-(G_2R^2+G_3)^{1/2}}{G_3^{1/2}+(G_2R^2+G_3)^{1/2}}\,, & \text{for } G_2<0\,.\end{cases} \tag{3.37}$$

The integral on the left-hand side of (3.37) can, however, be found provided one knows ω as an explicit function of the scalar field ϕ. The Brans–Dicke case is simple and one integrates the left-hand side for $\omega=\text{const}$ to get $(2\omega+3)^{1/2}\ln\phi$. Two different cases in Nordtvedt's theory are (a) Schwinger's relation is $(2\omega+3)=1/(\alpha\phi)$ and the left-hand side of (3.37) is

$$\int(2\omega+3)^{1/2}\frac{d\phi}{\phi}=-\frac{2}{(\alpha\phi)^{1/2}}\,;$$

and (b) Barker's relation is $(2\omega+3)=1/(\phi-1)$ and we have

$$\int(2\omega+3)^{1/2}\frac{d\phi}{\phi}=2\arctan(\phi-1)^{1/2}\,.$$

The above results along with (3.37) express ϕ as functions of the variable R and in turn as functions of the time t in view of (3.35). In the present case, where the matter content is in the form of radiation with the equation of state $p=\frac{1}{3}\rho$, one can easily conclude from the general relation (3.4) that $\rho\propto 1/R^4$, whereas the energy density due to the scalar field is, from (3.28),

$$\rho_\phi=\frac{(2\omega+3)}{4k}\left(\frac{\dot{\phi}}{\phi}\right)^2=\frac{G_1^2}{4kR^6}\,,$$

so that $\rho_\phi\propto 1/R^6$.

Above analysis indicates that at $R\to 0$ both the matter density and the energy density due to the scalar field increase to indefinitely large magnitude—the latter increasing at a much faster rate than the former, and as the singularity is approached, ρ_ϕ dominates over ρ. One can remark that the situation here is exactly analogous to that in BD theory and this is fundamentally due to the fact that Eq. (3.4) and (3.28) are valid independently of the choice of ω. From (3.32) we observe that, as $G-\frac{1}{3}(C_1-C_3)^2$ is positive, $\ddot{R}/R<0$ and hence $(\ln R^3)^{\cdot\cdot}<0$ indicating that there is no lower bound of R^3 and one cannot avoid the singularity. This observation is in keeping with the discussions in Sec. II, where we observed that the singularity is unavoidable from the consideration of Hawking's energy condition.

The case $G_2<0$ provides an interesting situation. From (3.34), one can have $\dot{R}=0$ at some finite value of $R=R_1$, where

$$G_3/R_1^4=-G_2/R_1^2\,.$$

As $(\ln R^3)^{\cdot\cdot}<0$, we have a miximum at this point and the model recollapses into the singularity after this maximum. This happens only if $G_2<0$. On the other hand, if $G_2>0$, there is no such maximum at any stage of evolution and $R\to 0$ is the point of singularity.

ACKNOWLEDGMENTS

N. B. and N. O. S. acknowledge the financial support from the Council of Scientific and Industrial Research (C. S. I. R.) and C. A. P. E. S. (Brazil), respectively.

[1] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[2] K. Nordtvedt, Jr., Astrophys. J. **161**, 1059 (1970).
[3] A. Banerjee, and S. B. Duttachoudhury, J. Math. Phys. **21**, 2799 (1980).
[4] N. Vanden Bergh, Gen. Relativ. Gravit. **14**, 17, (1982).
[5] V. U. M. Rao, and D. R. K. Reddy, Gen. Relativ. Gravit. **14**, 1017 (1982).
[6] B. M. Barker, Astrophys. J. **219**, 5 (1978).
[7] N. Vanden Bergh, Gen. Relativ. Gravit. **15**, 441 (1983).
[8] N. T. Bishop, Mon. Not. R. Astron. Soc. **176**, 241 (1976).
[9] A. Banerjee, and N. O. Santos, Phys. Rev. D. **23**, 2111 (1981).
[10] A. Banerjee and N. O. Santos, J. Phys. A: Math. Gen. **14**, 2829 (1981).
[11] R. H. Dicke, Phys. Rev. **125**, 2163 (1962).
[12] A. K. Raychaudhuri, *Theoretical Cosmology* (Clarendon, Oxford, 1979).
[13] A. K. Raychaudhuri, Phys. Rev. **86**, 90 (1955).
[14] S. Hawking, G. F. R. Ellis, *Large Scale Structure of Spacetime* (Cambridge U. P., Cambridge, 1973).
[15] V. A. Ruban, and A. M. Finkelstein, Gen. Relativ. Gravit. **6**, 601 (1975).
[16] S. Banerji, Prog. Theor. Phys. **39**, 365 (1968).
[17] H. Nariai, Prog. Theor. Phys. **47**, 1824 (1972).
[18] A. K. Raychaudhuri, Prog. Theor. Phys. **53**, 1360 (1975).

An extension of quaternionic metrics to octonions

S. Marques[a)]
Instituto de Física Teórica, Rua Pamplona, 145, São Paulo, Brazil
C. G. Oliveira[b)]
Departamento de Física, Universidade de Brasília, 70.910 Brasília, DF, Brazil

(Received 22 May 1984; accepted for publication 11 June 1985)

A treatment of a non-Riemannian geometry including internal complex, quaternionic, and octonionic space is made. Then, an interpretation of this geometry for the nonsymmetric theory of Einstein–Schrödinger, and for the unified theory of Borchsenius is showed. Finally, field equations in the extended octonionic geometry of space-time are obtained through a minimal action principle.

I. INTRODUCTION

The manifold where general relativity is defined is constituted by the space-time with a locally flat symmetric metric on which, through parallel transport of vectors, a symmetrical connection, $\{{}^{\rho}_{\mu\nu}\} = \{{}^{\rho}_{\nu\mu}\}$ is defined, and, as a function of that, the space-time curvature. In nonsymmetric theories[1] we define, on the space-time manifold, parallel transport of vectors through nonsymmetric connections. The metric contains a real symmetric part as in general relativity and a skew-symmetric part, taken by Einstein[1,2] as proportional to the electromagnetic field tensor. Actually, the interpretation for the skew-symmetric part for the metric as an electromagnetic field tensor have been proved to be physically incorrect,[3] the objections being overcome in the new interpretation developed by Moffat,[4] where the nonsymmetric metric is taken as a nonsymmetric gravitational field. In spite of this, we use here the interpretation of Einstein, as well as that the skew-symmetric part of the metric follows the interpretation given in the Borchsenius theory.[5]

The symmetry group acting on the space-time manifold is the "manifold mapping group" (MMG). In addition to the MMG we can associate to each point objects that, besides transforming through space-time mappings, can transform by the effect of internal mappings. The symmetry groups formed with these internal mappings, called "internal groups," are, in many important applications, local Lie groups. In the extended geometrical treatment, which permits the inclusion of Yang–Mills fields in unified theory (the Borchsenius theory[5]), the SU (2) group will be of special importance. The SU (2) algebra can be reinterpreted through a quaternionic algebra.

The main goal of this work is to establish geometrical properties of the full space to which we refer above. Next, we reinterpret it through a quaternionic algebra. As an obvious generalization of the quaternionic geometry we then suggest a (split) octonionic geometry.[6,7] Each one of these four internal spaces is allowed through a mathematical theorem, by Hurwitz.[8] This will be done in Secs. II–V. As a matter of completion, we will obtain in Sec. VI, field equations through a minimal action principle.

The departure point of the geometrical theory presented in this work is, besides general relativity, the nonsymmetrical theory of Einstein–Schödinger in its complex formulation. However, there is another possible formulation for the latter, realized over the algebra of real numbers (see Hlavaty[9]). Indeed, Einstein used the complex formulation only in his initial works.[10] In the real formulation of nonsymmetric theory, the algebraic structure imposed on the space-time manifold is realized through the *real* tensor $g_{\mu\nu}$, which can be written as

$$g_{\mu\nu} = h_{\mu\nu} + k_{\mu\nu}. \tag{1.1}$$

We define

$$g_{\mu\nu} \overset{\text{def}}{=} h_{\mu\nu} - k_{\mu\nu} = h_{\nu\mu} + k_{\nu\mu} = g_{\nu\mu}. \tag{1.2}$$

For the real connection $\Gamma^{\nu}{}_{\lambda\mu}$, a function of $g_{\mu\nu}$, it is assumed that the following equation is true:

$$\partial_{\omega} g_{\lambda\mu} - \Gamma^{\alpha}{}_{\lambda\omega} g_{\alpha\mu} - \Gamma^{\alpha}{}_{\omega\mu} g_{\lambda\alpha} = 0. \tag{1.3}$$

Then, it is easily proved that

$$\bar{\Gamma}^{\nu}{}_{\lambda\mu} = \Gamma^{\nu}{}_{\mu\lambda} \tag{1.4}$$

(Einstein assumes this relation), where $\bar{\Gamma}^{\lambda}{}_{\mu\nu}$ is a function of $\bar{g}_{\mu\nu}$. This is the same as saying that Eq. (1.3) above is true even though we substitute $g_{\alpha\beta}$, $\Gamma^{\lambda}{}_{\mu\nu}$ for $\bar{g}_{\alpha\beta}$, $\bar{\Gamma}^{\lambda}{}_{\mu\nu}$.

The only remarkable difference between the real and complex formulation of nonsymmetric theory is that the choice of complex quantities will allow us to obtain the nonsymmetric theories in terms of (complex) vierbeins, which is not possible in closed form if the metric tensor is real.[11]

II. A METRICAL GEOMETRY

Consider the non-Riemannian space-time of the Einstein nonsymmetric theory associated with an n-dimensional internal space. Let

$$ds^2 = (1/n)\,\mathrm{Tr}\,(G_{\mu\nu}\,dx^{\mu}\,dx^{\nu}), \tag{2.1}$$

where

$$G_{\mu\nu} = (G_{\mu\nu}{}^{a}{}_{b}(x)), \quad a,b = 1,\ldots,n, \tag{2.2}$$

is a matrix of internal space such that

$$(1/n)\,\mathrm{Tr}\,G_{\mu\nu} = g_{\mu\nu}. \tag{2.3}$$

Here $g_{\mu\nu}$ is the metric of the Einstein–Schrödinger asymmetric theory (or Moffat–Boal theory[12]). Therefore, (2.1) is the

[a)] Present address: Physical Department, Universidade de Brasília, CEP 70.910, Brasília, Brazil.
[b)] Deceased (1931–1985).

 0022-2488/85/123131-09$02.50

line element on the curved space-time. We further impose that

$$G^{\dagger}_{\mu\nu} = G_{\nu\mu}, \tag{2.4}$$

where the "†" operation is the Hermitian conjugation in internal space, and the inverse $G_{\mu\nu}$ is defined by

$$G_{\mu\alpha}G^{\mu\nu} = \delta^{\nu}_{\alpha}\mathbb{1} \tag{2.5}$$

in this order. Using (2.4) and (2.5) we also obtain

$$G^{\nu\mu}G_{\alpha\mu} = \delta^{\nu}_{\alpha}\mathbb{1}. \tag{2.6}$$

Since $G_{\mu\nu}$ is an object with two matrix indexes in internal space, we will restrict ourselves from now on to the internal space of 2×2 matrices, whose symmetry group is SU (2). Every object in this space can be written as a linear combination of four linearly independent matrices (τ_i, $i = 0, 1, 2, 3$), where $\tau_0 = \mathbb{1}_2$ and $\tau_i^{+} = \tau_i$, $i = 1, 2, 3$. With this, the "metric" (2.2) can be written as

$$G_{\mu\nu} = G_{\mu\nu i}(x)\tau_i, \quad i = 0,1,2,3. \tag{2.7}$$

The symmetry conditions (2.4) and the restrictions (2.3) imply that

$$G_{\mu\nu} = g_{\mu\nu}\tau_0 + q_{\mu\nu i}\tau_i, \quad i = 1,2,3, \tag{2.8}$$

where

$$g_{\mu\nu} = g_{\mu\nu} + i\kappa F_{\mu\nu},$$
$$q_{\mu\nu i} = \tfrac{1}{2}(i\epsilon\kappa^2/h)\underset{\vee}{f}_{\mu\nu i}. \tag{2.9}$$

In these formulas $\kappa = -2\hbar/e$ ($c = G = 1$) is a universal constant such that, in the limit $\kappa \to 0$, the Einstein–Maxwell–Yang–Mills theory is obtained (see Brochsenius[5]), e is the elementary electric charge, ϵ is the elementary isotopic charge, and $F_{\mu\nu}$ is the Maxwell tensor and $f_{\mu\nu i}$ represents the Yang–Mills field strength in the Moffat–Boal theory and Borchsenius theory.

III. AN AFFINE GEOMETRY

Let

$$A^{\mu}(x) = a^{\mu}_i(x)\tau_i, \quad i = 0,1,2,3, \tag{3.1}$$

where the $a^{\mu}_i(x)$ are the components of $A^{\mu}(x)$ in the internal matrix space. The parallel transport in space-time of an object with space-time and internal indices will affect both aspects. Since we are working with a non-Riemannian manifold, the resultant space-time connection can be taken as that of the Einstein nonsymmetric theory, and the space-time covariant derivative is given by

$$A^{\mu}_{\ ;\nu} = A^{\mu}_{\ ,\nu} + \Omega^{\mu}_{\ \alpha\nu}A^{\alpha}, \tag{3.2}$$

where

$$\Omega^{\mu}_{\ \alpha\nu} = \Omega^{*\mu}_{\ \ \nu\alpha},$$
$$\Omega^{\mu}_{\ \alpha\nu} = \Omega^{\mu}_{\ \alpha\nu} + iK^{\mu}_{\ \underset{\vee}{\alpha\nu}}, \tag{3.3}$$

such that

$$\Omega'^{\lambda}_{\ \mu\nu}(x') = \frac{\partial x'^{\lambda}}{\partial x^{\alpha}}\frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial x^{\gamma}}{\partial x'^{\nu}}\Omega^{\alpha}_{\ \underset{-}{\beta\gamma}} + \frac{\partial x'^{\lambda}}{\partial x^{\delta}}\frac{\partial^2}{\partial x'^{\mu}}\frac{x^{\delta}}{\partial x'^{\nu}},$$
$$\underset{\vee}{K'}^{\lambda}_{\ \mu\nu}(x') = \frac{\partial x'^{\lambda}}{\partial x^{\alpha}}\frac{\partial x^{\beta}}{\partial x'^{\mu}}\frac{\partial x^{\gamma}}{\partial x'^{\nu}}\underset{\vee}{K}^{\alpha}_{\ \beta\gamma}. \tag{3.4}$$

In the Einstein–Schrödinger nonsymmetric theory,[1] the connection used in the calculation of the field equations is the "Schrödinger connection"

$$\theta^{\rho}_{\ \mu\nu} = \Omega^{\rho}_{\ \mu\nu} - (2/i\kappa)\delta^{\rho}_{\mu}A_{\nu}, \tag{3.5}$$

where the A_{ν} are the electromagneticlike potentials, which are given in terms of $\Omega^{\rho}_{\ \mu\nu}$ as

$$A_{\nu} = -\tfrac{1}{3}(i\kappa)\underset{\vee}{\Omega}^{\rho}_{\ \rho\nu}. \tag{3.6}$$

The $\Omega^{\rho}_{\ \mu\nu}$ is a nonsymmetric connection, and therefore, there exist two types of covariant derivatives: a " + " derivative

$$A^{\mu}_{\ \underset{+}{;\nu}} = A^{\mu}_{\ ,\nu} + \Omega^{\mu}_{\ \alpha\nu}A^{\alpha}, \tag{3.7}$$

and a " − " derivative

$$A^{\mu}_{\ \underset{-}{;\nu}} = A^{\mu}_{\ ,\nu} + \Omega^{\mu}_{\ \nu\alpha}A^{\alpha}. \tag{3.8}$$

The same definition can be used for any type of tensor.

The space-time curvature can be obtained through the difference

$$A^{\mu}_{\ \underset{+\,+}{;\alpha\beta}} - A^{\mu}_{\ \underset{+\,+}{;\beta\alpha}},$$

where A^{μ} is given by (3.1). Performing this difference we obtain

$$R^{\sigma}_{\ \mu\nu\rho} = (\partial_{\rho}\Omega^{\sigma}_{\ \mu\nu} - \Omega^{\sigma}_{\ \alpha\nu}\Omega^{\alpha}_{\ \mu\rho}) - (\partial_{\nu}\Omega^{\sigma}_{\ \mu\rho} - \Omega^{\sigma}_{\ \alpha\rho}\Omega^{\alpha}_{\ \mu\nu}). \tag{3.9}$$

Again we must perform the transition $R^{\sigma}_{\ \mu\nu\rho} \to R^{\sigma}_{\ \mu\nu\rho}\,\tau_0$ for calculations in the matrix notation. We observe that this curvature was obtained using " + + " derivatives, which were used by Einstein in his nonsymmetric theory.[2] It is possible to find other curvatures with vector covariant derivatives using " + " and " − " types of derivatives; however, these curvatures are interrelated by algebraic relations to the above expression (3.9).

Defining the internal vector $\psi^{a} = \psi^{a}(x)$, $a = 1, 2$, the "internal" covariant derivatives are given by

$$\psi^{a}_{\ \|\mu} = \psi^{a}_{\ ,\mu} + \Gamma_{\mu}^{\ a}{}_{b}\psi^{b}. \tag{3.10}$$

The affinity $\Gamma_{\mu} = (\Gamma_{\mu}^{\ a}{}_{b}(x))$ is the object which makes $\psi^{a}{}_{\|\mu}$ transform like a vector under transformations in the internal space. In the case of an isotopic gauge,[13] the physically relevant part of the field Γ_{μ} is of the form

$$\Gamma_{\mu} = i\mathbf{C}_{\mu}\cdot\boldsymbol{\tau} = -i(\epsilon/\hbar)\mathbf{b}_{\mu}\cdot\boldsymbol{\tau}, \tag{3.11}$$

where ϵ is the elementary isotopic charge. From now on, we use this form for Γ_{μ}. The (internal) transformation law of Γ_{μ} is

$$\Gamma'_{\mu} = U(x)\Gamma_{\mu}U^{-1}(x) - \frac{\partial U(x)}{\partial x^{\mu}}U^{-1}(x), \tag{3.12}$$

where the $U(x)$ are the internal transformation matrices belonging to the local SU (2) group. In the curved space-time, Γ_{μ} transforms like a vector.

The internal curvature is defined in the usual form, i.e., from the difference

$$\psi^{a}_{\ \|\mu\nu} - \psi^{a}_{\ \|\nu\mu} = P_{\mu\nu}^{\ \ a}{}_{b}\psi^{b}. \tag{3.13}$$

Here $P_{\mu\nu}^{\ \ a}{}_{b}$ is the curvature in the internal space,

$$P_{\mu\nu} = \Gamma_{\mu,\nu} - \Gamma_{\nu,\mu} - [\Gamma_\mu, \Gamma_\nu]. \tag{3.14}$$

We have

$$P_{\mu\nu} = -P_{\nu\mu}. \tag{3.15}$$

After (3.7) and (3.10), the *total* covariant derivative of the space-time vector $A^\mu(x)$, defined in (3.1) is

$$A^\mu{}_{|\alpha}(x) = A^\mu{}_{,\alpha} + \Omega^\mu{}_{\rho\alpha} A^\rho + [\Gamma_\alpha, A^\mu]. \tag{3.16}$$

With the definition (3.16), we can also obtain a "total curvature" through the difference

$$A^\mu{}_{|\alpha\beta} - A^\mu{}_{|\beta\alpha} = \mathbb{R}^\mu{}_{\lambda\alpha\beta} A^\lambda - A^\mu P_{\alpha\beta} - 2A^\mu{}_{|\rho} \Omega^\rho{}_{\alpha\beta}. \tag{3.17}$$

The $\mathbb{R}^\mu{}_{\lambda\alpha\beta}$ is a "mixing" of space-time and internal curvatures:

$$\mathbb{R}^\mu{}_{\lambda\alpha\beta} = \left(\mathfrak{T}^\mu{}_{\lambda\alpha,\beta} + \mathfrak{T}^\mu{}_{\rho\beta}\mathfrak{T}^\rho{}_{\lambda\alpha}\right) - \left(\mathfrak{T}^\mu{}_{\lambda\beta,\alpha} + \mathfrak{T}^\mu{}_{\rho\alpha}\mathfrak{T}^\rho{}_{\lambda\beta}\right)$$

$$= R^\mu{}_{\lambda\alpha\beta}\tau_0 + \delta^\mu{}_\lambda P_{\alpha\beta}, \tag{3.18}$$

with

$$\mathfrak{T}^\rho{}_{\nu\alpha} = \Omega^\rho{}_{\nu\alpha} + i\delta^\rho{}_\nu \mathbf{C}_\alpha \cdot \boldsymbol{\tau}. \tag{3.19}$$

The dot under the index indicates the covariant derivative index. Therefore, $\mathbb{R}^\mu{}_{\lambda\alpha\beta}$ is called the "total curvature" and it is this "curvature" that Borchsenius makes use of in his calculations for obtaining field equations in a unified theory of gravitation, electromagnetism, and Yang–Mills fields.[5]

IV. QUATERNION INTERPRETATION OF THE GEOMETRY

Let X be a set of numbers. Consider a set X^2 of pairs of numbers X, in which the addition is defined in the usual form and the multiplication is defined by the equation

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 \cdot x_2 - \bar{y}_2 \cdot y_1, y_1 \cdot \bar{x}_2 + y_2 \cdot x_1),$$

with

$$(\overline{x, y}) = (\bar{x}, -\bar{y}).$$

Then the set X^2 is called the "double" of the set X. The "double" of real numbers (dim $n = 1$) gives the complex numbers (dim $n = 2$). The "double" of complex numbers gives the quaternions (dim $n = 4$), and the "double" of quaternions gives the octonions (dim $n = 8$). These four sets of numbers are distinguished from all other possible ones by the Hurwitz theorem.[8]

We can observe the following facts: general relativity theory is carried out over the set of real numbers and the Einstein–Schrödinger nonsymmetric theory, and so the Moffat–Boal theory, can be carried out on the set of complex numbers. We will see now that the Borchsenius theory can be reinterpreted via quaternions, when these are realized by means of the Pauli matrices. This is carried out easily by performing the transition

$$\sigma_i \to i\omega_i, \quad i = 1,2,3,$$

where $\sigma_i \equiv \tau_i$ are the Pauli matrices and

$$\omega_i = (1/i)\sigma_i, \quad i = 1,2,3,$$

$$\omega_0 = \sigma_0 = \tau_0 = \mathbf{1}_2. \tag{4.1}$$

The "numbers" ω_i, $i = 1, 2, 3$, $\omega_0 = \mathbf{1}_2$, with a product defined by

$$\omega_i \omega_j = \epsilon_{ijk}\omega_k - \delta_{ij}\omega_0, \tag{4.2}$$

are the generators of a quaternionic algebra $\mathbb{Q}$.

Therefore, the collection of tensors that are matrices in the internal space with the local symmetry group SU (2) can be written as quaternions when we consider the space-time derivatives continuing unchanged. Thus the internal covariant derivatives will now be what we call quaternionic covariant derivatives or $\mathbb{Q}$ derivatives. Then, if K is a quaternion, the $\mathbb{Q}$ derivative acting on K is defined by

$$K_{\|\mu} = K_{,\mu} + [\Gamma_\mu, K], \tag{4.3}$$

where

$$\Gamma_\mu = i\mathbf{C}_\mu \cdot \boldsymbol{\tau} \equiv i\mathbf{C}_\mu \cdot \boldsymbol{\sigma} = -\mathbf{C}_\mu \cdot \boldsymbol{\omega}. \tag{4.4}$$

As a consequence, the $\mathbb{Q}$-covariant derivative is always represented by the commutator $[\Gamma_\mu, K]$.

In this way, the metric and the curvature will be rewritten as

$$G_{\mu\nu} = g_{\mu\nu}\omega_0 + (1/i)q_{\mu\nu_i}\omega_i, \quad i = 1,2,3, \tag{4.5}$$

$$P_{\mu\nu} = [(C_{\nu k,\mu} - C_{\mu k,\nu} - 2\epsilon_{ijk}C_{\mu i}C_{\nu j})\omega_k]. \tag{4.6}$$

The space-time curvature obviously retains the same form.

Therefore, we can see that an interpretation of the geometrical objects of the Borchsenius theory via quaternions is a direct one. With the above interpretation we can suppose that there exists a more general theory that will be carried out on an algebra of dim $n = 8$, an octonionic algebra (O). It is this generalization that we intend to study in the next section. (The definition and properties of octonionic algebra are introduced in Appendix A.) The present analysis suggests that the use of complex quantities in the Einstein–Schrödinger theory is not merely a mathematical artifice.

V. THE GEOMETRY IN THE OCTONIONIC SPACE

We consider now a mathematical manifold in which we have the non-Riemannian space-time of the Einstein nonsymmetric theory, but with an associated internal space described by octonions. We will call this space an "octonionic space." We will consider here only the "split octonions."

Any octonion $A^\mu(x)$ with a space-time index, can be written in terms of the (split) generators as

$$A^\mu(x) = a^\mu(x)u_0^* + b^\mu(x)u_0 + k_i^\mu(x)u_i^* + l_i^\mu(x)u_i, \quad i = 1,2,3, \tag{5.1}$$

where the coefficients of $A^\mu(x)$ transform like vectors under space-time transformations.

Through arguments similar to those used in the treatment of objects that are matrices in the internal space, we can conclude that a space-time covariant derivative of $A^\mu(x)$ must be of the form

$$A^\mu{}_{+;\alpha} = A^\mu{}_{,\alpha} + \Omega^\mu{}_{\rho\alpha}A^\rho,$$

$$A^\mu{}_{-;\alpha} = A^\mu{}_{,\alpha} + \Omega^\mu{}_{\alpha\rho}A^\rho, \tag{5.2}$$

where now

$$\Omega^\mu{}_{\rho\alpha} \equiv \Omega^{\dot{\mu}}{}_{\rho\alpha}\mathbf{1}. \tag{5.3}$$

The $\Omega^{\mu}{}_{\rho\alpha}$ is the affinity of the nonsymmetric theory and $\mathbb{1} = u_0^* + u_0$ is the unity element of the split octonion algebra. The space-time curvature is again given by $R^{\sigma}{}_{\mu\nu\rho} \cdot \mathbb{1}$, where $R^{\sigma}{}_{\mu\nu\rho}$ is the curvature of nonsymmetric theory.

Let $\mathbb{K}$ be an octonion. The octonionic covariant derivative, or $\mathbb{O}$-derivative, of $\mathbb{K}$ is defined by

$$\mathbb{K}_{\|\mu} = \mathbb{K}_{,\mu} + |\mathfrak{T}_{\mu},\mathbb{K}|, \tag{5.4}$$

where $\mathfrak{T}_{\mu}$ is the "octonionic affinity." This means that $\mathfrak{T}_{\mu}$ is the object which makes $\mathbb{K}_{\|\mu}$ transform like an *octonion* under $\mathbb{O}$ transformations [see Appendix B, where the reason for the nonappearance of parenthesis in (5.5) is explained]

$$\mathbb{K}' = \mathbb{U}\mathbb{K}\mathbb{U}^{-1},$$
$$\mathbb{K}'_{\|\mu} = \mathbb{U}\mathbb{K}_{\|\mu}\mathbb{U}^{-1}, \tag{5.5}$$

and

$$\mathfrak{T}'_{\mu} = \mathbb{U}\mathfrak{T}_{\mu}\mathbb{U}^{-1} - \frac{\partial\mathbb{U}}{\partial x^{\mu}}\mathbb{U}^{-1}, \tag{5.6}$$

where the $\mathbb{U}(x)$ are octonions which define local (octonionic) transformations, isomorphic to the rotation group O_3, which means that they are SU (2)-like octonionic transformations. The octonion $\mathbb{U}^{-1}$ is defined as being identical to $\bar{\mathbb{U}}$, the conjugate of $\mathbb{U}$. See Appendix B for a more detailed treatment on the properties of $\mathbb{O}$ transformations.

We are taking the "doubling" of the quaternionic algebra, which forms the split octonions with realization via Pauli matrices. Therefore, it is logical to *impose* that $\mathfrak{T}_{\mu}$ be a trace-free Zorn matrix, which we suppose to be of the Yang–Mills type:

$$\mathfrak{T}_{\mu} = -L_{\mu_i}u^*{}_i - K_{\mu_i}u_i$$
$$\equiv \begin{pmatrix} O_2 & \mathbf{L}_{\mu}\cdot\boldsymbol{\omega} \\ -\mathbf{K}_{\mu}\cdot\boldsymbol{\omega} & O_2 \end{pmatrix}, \quad i = 1,2,3. \tag{5.7}$$

In the limit $\mathbf{K}_{\mu}\rightarrow\mathbf{L}_{\mu}$ we reobtain the Yang–Mills affinity, as should be expected.

The total derivative of an octonion $A_{\mu}{}^{\rho}$, with two space-time indices, is defined then as

$$A_{\underset{+}{\mu}}{}^{\overset{\rho}{+}}{}_{|\alpha} = A_{\mu}{}^{\rho}{}_{,\alpha} + \Omega^{\rho}{}_{\sigma\alpha}A_{\mu}{}^{\sigma} - \Omega^{\sigma}{}_{\mu\alpha}A_{\sigma}{}^{\rho} + [\mathfrak{T}_{\alpha},A_{\mu}{}^{\rho}]. \tag{5.8}$$

The octonionic curvature is obtained by a similar process, by performing the difference

$$\mathbb{K}_{\|\mu\nu} - \mathbb{K}_{\|\nu\mu} = \mathbb{P}_{\mu\nu}\mathbb{K} - \mathbb{K}\mathbb{P}_{\nu\mu} + \{\mathfrak{T}_{\mu},\mathfrak{T}_{\nu},\mathbb{K}\}, \tag{5.9}$$

where $\{\mathfrak{T}_{\mu}, \mathfrak{T}_{\nu}, \mathbb{K}\}$ is the associator of fields $\mathfrak{T}_{\mu}$, $\mathfrak{T}_{\nu}$, $\mathbb{K}$, and

$$\mathbb{P}_{\mu\nu} = \mathfrak{T}_{\mu,\nu} - \mathfrak{T}_{\nu,\mu} - [\mathfrak{T}_{\mu},\mathfrak{T}_{\nu}] \tag{5.10}$$

is called an octonionic curvature, or "$\mathbb{O}$ curvature." In terms of components it is given by

$$\begin{aligned}\mathbb{P}_{\mu\nu} = {} & (L_{\mu_i}K_{\nu_i} - L_{\nu_i}K_{\mu_i})u^*{}_0 + (K_{\mu_i}L_{\nu_i} - K_{\nu_i}L_{\mu_i})u_0 \\ & + (L_{\nu k,\mu} - L_{\mu k,\nu} - 2\epsilon_{ijk}K_{\mu i}K_{\nu j})u_k^* \\ & + (K_{\nu k,\mu} - K_{\mu k,\nu} - 2\epsilon_{ijk}L_{\mu i}L_{\nu j})u_k, \\ & i,j,k = 1,2,3.\end{aligned} \tag{5.11}$$

In the limit $\mathbf{K}_{\mu}\rightarrow\mathbf{L}_{\mu}$ we reobtain the quaternionic expression.

If A^{μ} is an arbitrary octonionion with a space-time index, the total octonionic curvature is obtained through the difference

$$\begin{aligned}A^{\mu}{}_{\underset{++}{|\alpha\beta}} - A^{\mu}{}_{\underset{++}{|\beta\alpha}} = {} & R^{\mu}{}_{\lambda\alpha\beta}A^{\lambda} + [\mathbb{P}_{\alpha\beta},A^{\mu}] \\ & + 6\{\mathfrak{T}_{\alpha},\mathfrak{T}_{\beta},A^{\mu}\} - 2A^{\mu}{}_{\underset{+}{|\rho}}\Omega^{\rho}{}_{\underset{\vee}{\alpha\beta}},\end{aligned} \tag{5.12}$$

where $\Omega^{\rho}{}_{\underset{\vee}{\alpha\beta}} = \frac{1}{2}(\Omega^{\rho}{}_{\alpha\beta} - \Omega^{\rho}{}_{\beta\alpha})$. The curvatures $R^{\mu}{}_{\lambda\alpha\beta}$ and $\mathbb{P}_{\alpha\beta}$ can be put together in a same expression as we will see below.

Defining

$$\mathfrak{T}^{\rho}{}_{\mu\nu} = \Omega^{\rho}{}_{\mu\nu} + \delta^{\rho}{}_{\mu}\Gamma_{\nu} \tag{5.13}$$

(the dot under the index ν recalls the covariant derivative index that was used), we can rewrite (5.12) in the form

$$\begin{aligned}& A^{\mu}{}_{\underset{++}{|\alpha\beta}} - A^{\mu}{}_{\underset{++}{|\beta\alpha}} \\ & = (\mathbb{R}^{\mu}{}_{\lambda\alpha\beta})A^{\lambda} + A^{\mu}\mathbb{P}_{\alpha\beta} + \{\mathfrak{T}^{\mu}{}_{\rho\alpha},\mathfrak{T}^{\rho}{}_{\lambda\beta},A^{\lambda}\} \\ & \quad - \{\mathfrak{T}^{\mu}{}_{\rho\beta},\mathfrak{T}^{\rho}{}_{\lambda\alpha},A^{\lambda}\} + 4\{\mathfrak{T}_{\alpha},\mathfrak{T}_{\beta},A^{\mu}\} \\ & \quad - 2A^{\mu}{}_{\underset{+}{;\rho}}\Omega^{\rho}{}_{\underset{\vee}{\alpha\beta}}.\end{aligned} \tag{5.14}$$

Here, $\mathbb{R}^{\mu}{}_{\lambda\alpha\beta}$ is the "octonionic total curvature tensor"

$$\mathbb{R}^{\mu}{}_{\lambda\alpha\beta} = (\mathfrak{T}^{\mu}{}_{\lambda\alpha,\beta} + \mathfrak{T}^{\mu}{}_{\rho\beta}\mathfrak{T}^{\rho}{}_{\lambda\alpha}) - (\mathfrak{T}^{\mu}{}_{\lambda\beta,\alpha} + \mathfrak{T}^{\mu}{}_{\rho\alpha}\mathfrak{T}^{\rho}{}_{\lambda\beta}). \tag{5.15}$$

Expanding (5.15), we obtain

$$\mathbb{R}^{\mu}{}_{\lambda\alpha\beta} = R^{\mu}{}_{\lambda\alpha\beta}\cdot\mathbb{1} + \delta^{\mu}_{\lambda}\mathbb{P}_{\alpha\beta}. \tag{5.16}$$

We define the "line element" on the octonionic space as

$$ds^2 = \tfrac{1}{4}\mathrm{Tr}(\mathbb{G}_{\mu\nu}\,dx^{\mu}\,dx^{\nu}), \tag{5.17}$$

where $\mathbb{G}_{\mu\nu}$ is the "octonionic metric tensor" on the octonionic space. This can be written in a more general form as

$$\mathbb{G}_{\mu\nu} = s_{\mu\nu_0}(x)u^*{}_0 + s_{\mu\nu_i}(x)u^*{}_i + r_{\mu\nu_0}(x)u_0 + r_{\mu\nu_i}(x)u_i \tag{5.18}$$

or

$$\mathbb{G}_{\mu\nu} = \begin{pmatrix} s_{\mu\nu_0} & -\mathbf{s}_{\mu\nu}\cdot\boldsymbol{\omega} \\ \mathbf{r}_{\mu\nu}\cdot\boldsymbol{\omega} & r_{\mu\nu_0}\omega_0 \end{pmatrix} = \mathbb{G}_{\mu\nu}(s,r) \tag{5.19}$$

in the Zorn matrix notation. With the above definition for $\mathbb{G}_{\mu\nu}$, the line element may be written, from (5.17), as

$$ds^2 = \tfrac{1}{2}(s_{\mu\nu_0} + r_{\mu\nu_0})dx^{\mu}\,dx^{\nu}, \tag{5.20}$$

which gives the line element of the nonsymmetric theory when $r_{\mu\nu_0}\rightarrow s_{\mu\nu_0}$.

The octonionic conjugation is defined in Appendix A. We will define now the "Hermitian conjugate" of any octonion A as

$$\begin{aligned}A^{\dagger} & = a^*u_0 + b^*u^*{}_0 - x^*{}_iu^*{}_i - j^*{}_iu_i \\ & = \begin{pmatrix} b^*\omega_0 & x^*{}_i\omega_i \\ -y^*{}_i\omega_i & a^*\omega_0 \end{pmatrix},\end{aligned} \tag{5.21}$$

where we take the complex conjugate of the coefficients of $\bar{A}$.

The octonionic Hermitian conjugate of the "metric" $G_{\mu\nu}$ is

$$\mathbb{G}^{\dagger}_{\mu\nu}(s,r)=\begin{pmatrix} r^{*}_{\mu\nu_0}\omega_0 & s^{*}_{\mu\nu_i}\omega_i \\ -r^{*}_{\mu\nu_i}\omega_i & s^{*}_{\mu\nu_0\omega_0}\end{pmatrix}$$
$$=\bar{\mathbb{G}}_{\mu\nu}(s^{*},r^{*}). \tag{5.22}$$

Considering, as in the case of nonsymmetric and quaternionic theory, that the fields s and r are such that

$$s^{*}_{\mu\nu_0}=s_{\nu\mu_0}, \qquad r^{*}_{\mu\nu}=r_{\nu\mu_0},$$
$$\underset{\vee}{\mathbf{s}}{}^{*}_{\mu\nu}=\underset{\vee}{\mathbf{s}}_{\mu\nu}=-\underset{\vee}{\mathbf{s}}_{\nu\mu}, \qquad \underset{\vee}{\mathbf{r}}{}^{*}_{\mu\nu}=\underset{\vee}{\mathbf{r}}_{\mu\nu}=-\underset{\vee}{\mathbf{r}}_{\nu\mu}, \tag{5.23}$$

we find a symmetry relation for $\mathbb{G}_{\mu\nu}$:

$$\mathbb{G}^{\dagger}_{\mu\nu}(s,r)=\begin{pmatrix} r_{\nu\mu_0}\omega_0 & -s_{\nu\mu_i}\omega_i \\ r_{\nu\mu_i}\omega_i & s_{\nu\mu_0}\omega_0\end{pmatrix}=G_{\nu\mu}\begin{pmatrix} r \\ s\end{pmatrix}. \tag{5.24}$$

[Here the notation $\binom{r}{r}$ remembers simply the column places of fields r and s in the Zorn matrix $\mathbb{G}_{\mu\nu}$.] In our case, we want to have $r_{\mu\nu_0}=s_{\mu\nu_0}$, because we are studying the octonionic space, which consists of a non-Riemannian space-time of nonsymmetric theory associated to internal octonionic space. We have then

$$r_{\mu\nu_0}=s_{\mu\nu_0}=\underset{-}{g}_{\mu\nu}+i\kappa\underset{\vee}{F}_{\mu\nu}, \tag{5.25}$$

and we can say that the fields $\mathbf{s}_{\mu\nu}$ and $\mathbf{r}_{\mu\nu}$ are of the Yang–Mills form, but are different from each other. In the limit $\mathbf{r}\to\mathbf{s}$ we reobtain the quaternionic case. If we take (5.25), the symmetry condition (5.24) can be written as

$$\mathbb{G}^{\dagger}_{\mu\nu}(s,r)=\mathbb{G}_{\nu\mu}(s,r). \tag{5.26}$$

There exists an inverse $\mathbb{G}^{\mu\nu}$ of $\mathbb{G}_{\mu\nu}$ such that

$$\mathbb{G}_{\mu\alpha}\mathbb{G}^{\mu\nu}=\mathbb{G}_{\mu\alpha}(s,r)\mathbb{G}^{\mu\nu}(r,s)=\delta^{\nu}_{\alpha}\mathbb{1}, \tag{5.27}$$

in this order, as in quaternionic case. The expression (5.27) also can be written, because of the property (5.24) [or (5.26)], in the form

$$\mathbb{G}^{\nu\mu}\begin{pmatrix} r \\ s\end{pmatrix}\mathbb{G}_{\alpha\mu}\begin{pmatrix} r \\ s\end{pmatrix}=\delta^{\nu}_{\alpha}, \tag{5.28}$$

or, if we take (5.25),

$$\mathbb{G}^{\nu\mu}(s,r)\mathbb{G}_{\alpha\mu}(s,r)=\delta^{\nu}_{\alpha}. \tag{5.29}$$

Herewith we complete the principal geometrical considerations for an octonionic space. Table I summarizes the principal geometrical objects of Riemannian space-time, non-Riemannian space-time of the nonsymmetric theory, quaternionic space-time of the Borchsenius theory, and octonionic space-time defined above.

VI. FREE FIELD EQUATIONS IN AN OCTONIONIC UNIFIED THEORY

Bonnor and Moffat–Boal[12] suggested a modification in the Einstein Lagrangian of the nonsymmetric theory by introducing a new term

$$\mathcal{L}_{\mathrm{MB}}=\mathcal{L}_{\mathrm{ES}}+(4\pi G/k^2c^4)\underset{\vee}{\mathcal{G}}{}^{\mu\nu}g_{\nu\mu}, \tag{6.1}$$

where $\mathcal{L}_{\mathrm{ES}}$ is the Einstein-Schrödinger Lagrangian.[2] The obtained field equations reduce themselves to the Einstein–Maxwell equations in the limit $k\to 0$. Borchsenius has worked an extension to a theory that includes the Yang–Mills field.[13] The Lagrangian in this case is

$$\mathcal{L}_{\mathrm{KB}}=\mathrm{Tr}(\mathcal{G}^{\mu\nu}R_{\mu\nu}+[1/(i\kappa)^2]\underset{\vee}{\mathcal{G}}{}^{\mu\nu}g_{\mu\nu}), \tag{6.2}$$

where the objects in (6.2) are now 2×2 matrices. The density $\mathcal{G}^{\mu\nu}$ is

$$\mathcal{G}^{\mu\nu}=wg^{\mu\nu},$$
$$w=[-\tfrac{1}{4}\mathrm{Tr}(\det g_{\mu\nu})]^{1/2}. \tag{6.3}$$

Again he uses here a universal constant $i\kappa=-2i\hbar/e$ as a result of which, in the limit $\kappa\to 0$, he obtains, besides the Einstein–Maxwell equations, the Yang–Mills field equations.

We saw in previous sections the interpretation of the geometry of the Borchsenius theory via a quaternion algebra, as well as its extension to an octonionic-split geometry. We will obtain now the free-field equations for this octonionic-split geometry.

The "metric tensor" $\mathbb{G}_{\mu\nu}$ was defined as

$$\mathbb{G}_{\mu\nu}=(\underset{-}{g}_{\mu\nu}+i\kappa\underset{\vee}{F}_{\mu\nu})(u_0+u^{*}_0)+\underset{\vee}{\mathbf{s}}_{\mu\nu}\cdot\mathbf{u}^{*}+\underset{\vee}{\mathbf{r}}_{\mu\nu}\cdot\mathbf{u}, \tag{6.4}$$

where $\underset{\vee}{\mathbf{r}}_{\mu\nu}$ and $\underset{\vee}{\mathbf{s}}_{\mu\nu}$ are fields of the Yang–Mills type. Therefore,

$$\underset{\vee}{\mathbf{s}}_{\mu\nu}=-\kappa\mu\underset{\vee}{\mathbf{k}}_{\mu\nu}, \qquad \underset{\vee}{\mathbf{r}}_{\mu\nu}=-\kappa\mu\underset{\vee}{\mathbf{l}}_{\mu\nu}, \tag{6.5}$$

where $i\kappa$ is the universal constant used by Borchsenius, $\mu=-\epsilon/e$, e and ϵ being the elementary quantum of electrical and isotopic charge, respectively. The "density" $\mathfrak{G}^{\mu\nu}$ is

$$\mathfrak{G}^{\mu\nu}=w\mathbb{G}^{\mu\nu}, \tag{6.6}$$
$$w=[-\tfrac{1}{4}\mathrm{Tr}(\det\mathbb{G}_{\mu\nu})]^{1/2}, \tag{6.7}$$

where the trace operation (Tr) is carried out on the Zorn matrix of $\mathbb{G}_{\mu\nu}$, and the determinant of the matrix in the world index is taken, μ, ν.

The free-field octonionic Lagrangian is taken as

$$\mathcal{L}=\mathrm{Tr}(\mathfrak{G}^{\mu\nu}\mathbb{R}_{\mu\nu}+[1/(i\kappa)^2]\underset{\vee}{\mathfrak{G}}{}^{\mu\nu}\mathbb{G}_{\mu\nu}), \tag{6.8}$$

where $\mathbb{R}_{\mu\nu}$ is given by (5.15). This Lagrangian is invariant under an Einstein λ transformation (see Borchsenius[5]). Besides, there is also invariance under octonionic local transformations. By the use of the Palatini variational method over an action constructed with the above Lagrangian,

$$\mathcal{A}=\int\mathcal{L}\,d^4x, \tag{6.9}$$

we obtain the field equations

$$\mathbf{R}^{*}_{\mu\nu}=0, \tag{6.10}$$
$$-\mathfrak{G}^{\mu\nu}{}_{\alpha}-\Omega^{\nu}{}_{\lambda\alpha}\mathfrak{G}^{\mu\lambda}-\Omega^{\mu}{}_{\alpha\lambda}\mathfrak{G}^{\lambda\nu}+\Omega^{\lambda}{}_{\alpha\lambda}\mathfrak{G}^{\mu\nu}$$
$$-\tfrac{2}{3}\delta^{\nu}{}_{\alpha}\underset{\vee}{\Omega}{}^{\rho}{}_{\lambda\rho}\mathfrak{G}^{\mu\lambda}-[\mathfrak{T}_{\alpha},\mathfrak{G}^{\mu\nu}]=0, \tag{6.11}$$

where

$$\mathbf{R}^{*}_{\mu\nu}=\mathbf{R}_{\mu\nu}+\mathbf{I}_{\mu\nu}, \tag{6.12}$$
$$\mathbf{I}_{\mu\nu}=(-1/2\kappa^2)\Big[\underset{\alpha\beta}{\mathbb{G}}{}^{\beta\alpha}_{\mu\alpha}(\underset{\vee}{\mathbb{G}}{}^{\beta\alpha}_{\beta\nu})+(\underset{\alpha\beta}{\mathbb{G}}{}^{\beta\alpha}_{\mu\alpha}\underset{\vee}{\mathbb{G}}{}^{\beta\alpha}_{\beta\nu})\mathbb{G}_{\beta\nu}$$
$$+\tfrac{1}{2}\{\underset{\alpha\beta}{\mathbb{G}}{}_{\mu\nu}(\underset{\vee}{\mathbb{G}}{}^{\beta\alpha}_{\alpha\beta})+(\underset{\alpha\beta}{\mathbb{G}}{}_{\mu\nu}\underset{\vee}{\mathbb{G}}{}^{\beta\alpha}_{\alpha\beta})\}+\mathbb{G}_{\mu\nu}\Big]. \tag{6.13}$$

TABLE I. General table for symbols used in the text.

	General relativity theory (R algebra)	Nonsymmetric theory (C algebra)	Borchsenius or quaternionic theory (Q algebra)	Octonionic theory (O algebra)
Line element	$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu$	$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu$ $= \underset{-}{g}_{\mu\nu}\,dx^\mu\,dx^\nu$	$ds^2 = \frac{1}{2}\,\mathrm{Tr}\,(G_{\mu\nu}\,dx^\mu\,dx^\nu)$ $= \underset{-}{g}_{\mu\nu}\,dx^\mu\,dx^\nu$ "Tr" acting upon Q matrices	$ds^2 = \frac{1}{4}\,\mathrm{Tr}\,(\mathbf{G}_{\mu\nu}\,dx^\mu\,dx^\nu)$ $= \frac{1}{2}(s_{\mu\nu_0} + r_{\mu\nu_0})dx^\mu\,dx^\nu$ "Tr" acting upon Zorn matrices
Metric	$g_{\mu\nu} = g_{\nu\mu}$	$g_{\mu\nu} = \underset{-}{g}_{\mu\nu} + i\kappa \underset{v}{F}_{\mu\nu}$ $\underset{v}{F}_{\mu\nu}$ = Maxwell tensor $g^*_{\mu\nu} = g_{\nu\mu}$	$G_{\mu\nu} = (\underset{-}{g}_{\mu\nu} + i\kappa \underset{\vee}{F}_{\mu\nu})\omega_0$ $+ \underset{\vee}{\mathbf{f}}_{\mu\nu}\cdot\omega$ $\underset{\vee}{\mathbf{f}}_{\mu\nu}$ = Yang-Mills field $G^\dagger{}_{\mu\nu} = G_{\nu\mu}$ "†" acting upon Q matrices	$\mathbf{G}_{\mu\nu} = s_{\mu\nu_0}u_0^* + r_{\mu\nu_0}u_0$ $+ \underset{\vee}{\mathbf{s}}_{\mu\nu}\cdot\mathbf{u}^* + \underset{\vee}{\mathbf{r}}_{\mu\nu}\cdot\mathbf{u}$ $\mathbf{G}_{\mu\nu}(s, r)$ $s_{\mu\nu_0}$, $r_{\mu\nu_0}$: nonsymmetric-theory-like $\underset{\vee}{\mathbf{s}}_{\mu\nu}$, $\underset{\vee}{\mathbf{r}}_{\mu\nu}$: Yang–Mills-like $\mathbf{G}^\dagger{}_{\mu\nu}(s, r) = \bar{\mathbf{G}}_{\mu\nu}(s^*, r^*)$ $= \mathbf{G}_{\nu\mu}(s, r)$, when $s_{\mu\nu_0} = r_{\mu\nu_0} = \underset{-}{g}_{\mu\nu} + i\kappa \underset{\vee}{F}_{\mu\nu}$ "†" acting upon Zorn matrices
Affinity	$\Gamma^\rho{}_{\mu\nu} = \{{}^{\rho}_{\mu\nu}\} = \Gamma^\rho{}_{\nu\mu}$	$\Omega^\rho{}_{\mu\nu} = \theta^\rho{}_{\mu\nu} - \frac{2}{i\kappa}\delta^\rho_\mu A_\nu$ $A_\nu = -\frac{1}{3}(i\kappa)\Omega^\rho{}_{\underset{\vee}{\nu\rho}}$ $\theta^\rho{}_{\mu\nu}$ = Schrödinger connection $\underset{v}{\theta}{}^\rho{}_{\nu\mu} = 0$	$\mathfrak{T}^\rho{}_{\mu\nu} = \Omega^\rho{}_{\mu\nu}\omega_0 + \delta^\rho_\mu \mathbf{C}_\nu\cdot\omega$ $\mathbf{C}_\nu\cdot\omega$ = Q affinity	$\mathfrak{T}^\rho{}_{\mu\nu} = \Omega^\rho{}_{\mu\nu}(u_0 + u_0^*)$ $+ \delta^\rho_\mu(\mathbf{L}_\nu\cdot\mathbf{u}^* + \mathbf{K}_\nu\cdot\mathbf{u})$ $\mathbf{L}_\nu\cdot\mathbf{u}^* + \mathbf{K}_\nu\cdot\mathbf{u}$ = O affinity
Curvature	$R^\sigma{}_{\rho\mu\nu} = R^\sigma{}_{\rho\mu\nu}(\{{}^{\rho}_{\mu\nu}\})$	$R^\sigma{}_{\rho\mu\nu} = R^\sigma{}_{\rho\mu\nu}(\Omega^\rho{}_{\mu\nu})$	$\mathbf{R}^\sigma{}_{\rho\mu\nu} = \mathbf{R}^\sigma{}_{\rho\mu\nu}(\mathfrak{T}^\rho{}_{\mu\nu})$ $= R^\sigma{}_{\rho\mu\nu}(\Omega^\rho{}_{\mu\nu})\omega_0 + \delta^\sigma_\rho P_{\mu\nu}$ $P_{\mu\nu}$ = Q curvature	$\mathbf{R}^\sigma{}_{\rho\mu\nu} = \mathbf{R}^\sigma{}_{\rho\mu\nu}(\mathfrak{T}^\rho{}_{\mu\nu})$ $= R^\sigma{}_{\rho\mu\nu}(\Omega^\sigma{}_{\mu\nu})(u_0^* + u_0)$ $+ \delta^\sigma_\rho \mathbf{P}_{\mu\nu}$ $\mathbf{P}_{\mu\nu}$ = O curvature

If, in (6.11), we use the Schrödinger connection, $\theta^\rho{}_{\mu\nu}$,

$$\Omega^\rho{}_{\mu\nu} = \theta^\rho{}_{\mu\nu} - \tfrac{2}{3}\,\Omega_\nu\delta^\rho{}_\mu\,, \tag{6.14}$$

$$\Omega_\nu = (3/i\kappa)A_\nu\,, \tag{6.15}$$

A_ν the electromagnetic vector potential, and $\theta^\rho{}_{\mu\nu}$ a space-time connection such that

$$\theta_\mu = \theta^\rho{}_{\underset{\vee}{\mu\rho}} = 0\,, \tag{6.16}$$

we obtain,

$$-\mathfrak{G}^{\mu\nu}{}_{,\alpha} - \theta^\nu{}_{\lambda\alpha}\mathfrak{G}^{\mu\lambda} - \theta^\mu{}_{\alpha\lambda}\mathfrak{G}^{\lambda\nu} + \theta^\lambda{}_{\underset{-}{\alpha\lambda}}\mathfrak{G}^{\mu\nu} - [\mathfrak{T}_\alpha, \mathfrak{G}^{\mu\nu}] = 0\,, \tag{6.17}$$

or, in Einstein notation,

$$\overset{\mu\nu}{\mathfrak{G}}{}^{+\,-}{}_{;\alpha} + [\mathfrak{T}_\alpha, \mathfrak{G}^{\mu\nu}] = 0\,, \tag{6.18}$$

equivalently,

$$\overset{\mu\nu}{\mathfrak{G}}{}^{+\,-}{}_{|\alpha} = 0\,, \tag{6.19}$$

the O-total-derivative of $\mathfrak{G}^{\mu\nu}$. For $\mathfrak{T}_\mu$ we have, according to Borchsenius, and from (5.7),

$$\mathbf{L}_\mu = (-\epsilon/\hbar)\mathbf{c}_\mu\,, \qquad \mathbf{K}_\mu = (-\epsilon/\hbar)\mathbf{d}_\mu\,. \tag{6.20}$$

Also placing (6.13) in (6.12) we obtain

$$R^*{}_{\mu\nu}(\theta) = 0\,, \tag{6.21}$$

$$R^*{}_{\underset{\vee}{\mu\nu}}(\theta) + \tfrac{2}{3}(\Omega_{\mu,\nu} - \Omega_{\nu,\mu})(u_0 + u^*{}_0) - \underset{\vee}{\mathbf{P}}_{\mu\nu} = 0\,, \tag{6.22}$$

with $\underset{\vee}{\mathbf{P}}_{\mu\nu}$ given in (5.10). Antisymmetrizing (6.17) in μ and ν and contracting in μ and α, we obtain

$$\overset{\mu\nu}{\underset{\vee}{\mathfrak{G}}}{}_{,\nu} + \left[\mathfrak{T}_\nu, \overset{\mu\nu}{\underset{\vee}{\mathfrak{G}}}\right] = 0\,. \tag{6.23}$$

Again we have

$$\overset{\mu\nu}{\mathbf{G}}{}^{\vee} = i\kappa \overset{\mu\nu}{F}{}^{\vee}(u_0 + u^*_0) + \overset{\mu\nu}{\mathbf{s}}{}^{\vee} \cdot \mathbf{u}^* + \overset{\mu\nu}{\mathbf{r}}{}^{\vee} \cdot \mathbf{u}\,, \tag{6.24}$$

which gives, for (6.23),

$$(w\overset{\mu\nu}{F}{}^{\vee})_{,\nu} = 0\,, \tag{6.25}$$

$$(w\overset{\mu\nu}{\mathbf{s}}{}^{\vee})_{,\nu} - 2\mathbf{K}_\nu \wedge \overset{\mu\nu}{\mathbf{r}}{}^{\vee} w = 0\,, \tag{6.26}$$

$$(w\overset{\mu\nu}{\mathbf{r}}{}^{\vee})_{,\nu} - 2L_\nu \wedge \overset{\mu\nu}{\mathbf{s}}{}^{\vee} w = 0\,,$$

which are the "current components" of the Maxwell equations, in the case of (6.25), and Yang–Mills-like equations for (6.26). In the limit $\overset{\mu\nu}{\mathbf{s}}{}^{\vee} \to \overset{\mu\nu}{\mathbf{r}}{}^{\vee} = -\kappa\mu \overset{\mu\nu}{\mathbf{f}}{}^{\vee}$, and $\mathbf{K}_\mu \to \mathbf{L}_\mu = (-\epsilon/\hbar)\mathbf{b}_\mu$, we reobtain the quaternionic Borchsenius equations.

VII. CONCLUSION

Table I shows a collection of objects, namely, the metric, the affinity, and the curvature. We can observe there that these objects maintain their "forms" when we go from general relativity to the Einstein nonsymmetric theory, and then go to quaternions (Borchsenius theory) and octonions. The automorphism group of quaternions is SU(2), which is homomorphic to the rotation group O_3. In the case of split octonions, the automorphism group is the split G_2 (an exceptional Lie group). We showed by means of what we called "octonionic transformations" that these are homomorphic to the rotation group O_3. These "O transformations" are necessary to obtain the octonionic field equations, (6.19) and (6.21)–(6.23). The component equations (6.25) and (6.26), of Eq. (6.23), decompose in Maxwell equations and two equations of the type of Yang–Mills. These last ones are a sort of "doubling" of the quaternionic Yang–Mills equations, but with a symmetric mixing of component terms. Equation (6.18), or (6.19), is the octonionic generalization of the corresponding (quaternionic) Borchsenius field equations [see Eq. (3.20), Ref. (5)]. The Moffat–Boal theory has proved to have ghost poles and tachyons[14] in the weak-field approximation.

The same must happen with non-Abelian theories, like that of Borchsenius (quaternionic theory) and again, with the extension to the split-octonion field theory, developed here, but the analysis of this part is out of the scope of this paper. An important aspect of Q and O field theory that may be remembered here is the relation to local quaternionic and octonionic spinors.[15] For example, under the SU(3) subgroup of split-G_2, leaving u_0 and u^*_0 invariant, the three split octonions (u_1,u_2,u_3) transform like a unitary triplet (quarks) and the complex conjugate octonions (u^*_1, u^*_2, u^*_3) transform like a unitary antitriplet (antiquarks).[16]

APPENDIX A: THE CAYLEY ALGEBRA. REALIZATION VIA ZORN MATRICES

The octonions algebra has eight dimensions and its base vectors, e_0, e_i, $i = 1,\ldots,7$, satisfy the product law

$$e_0 e_i = e_i e_0 = e_i\,,$$
$$e_i e_j = -\delta_{ij} e_0 + \epsilon_{ijk} e_k\,. \tag{A1}$$

Now ϵ_{ijk} is an object completely skew-symmetric with seven nonzero elements: ϵ_{123}, ϵ_{516}, ϵ_{624}, ϵ_{435}, ϵ_{471}, ϵ_{572}, and ϵ_{673}. This algebra is also called the Cayley algebra.[8] It is neither commutative nor associative, but belongs to the class of alternative algebras, with the property that for any three octonions the associator of x,y,z, is given by

$$\{x,y,z\} = (xy)z - x(yz)\,. \tag{A2}$$

This changes sign when any two of its arguments change position. The quantities x,y,z, are called Cayley numbers. The Cayley algebra with the case given in (A1) belongs to the class of division algebras (real base), but it can also be presented as a "split algebra" if we use a new basis defined on the complex field. This is given as

$$u_0 = \tfrac{1}{2}(e_0 + ie_7), \qquad u^*_0 = \tfrac{1}{2}(e_0 - ie_7)\,,$$
$$u_i = \tfrac{1}{2}(e_i + ie_{i+3}), \quad u^*_i = \tfrac{1}{2}(e_i - ie_{i+3})\,, \qquad i = 1,2,3\,. \tag{A3}$$

From this definition follows the multiplication table

$$i,j,k = 1,2,3$$

$$\begin{aligned}
&u^*_i u^*_j = \epsilon_{ijk} u_k\,, && u_i u^*_j = -\delta_{ij} u_0\,,\\
&u_i u_j = \epsilon_{ijk} u^*_k\,, && u^*_i u_j = -\delta_{ij} u^*_0\,,\\
&u_i u_0 = 0\,, && u^*_i u^*_0 = 0\,,\\
&u_0 u_i = u_i\,, && u^*_0 u^*_i = u^*_i\,,\\
&u_0{}^2 = u_0\,, && u^*_0{}^2 = u^*_0\,,\\
&u_i u^*_0 = u_i\,, && u^*_i u_0 = u^*_i\,,\\
&u^*_0 u_i = 0\,, && u_0 u^*_i = 0\,,\\
&\multicolumn{2}{c}{u_0 u^*_0 = u^*_0 u_0 = 0\,.}
\end{aligned} \tag{A4}$$

It is to our interest to find a convenient realization for the basis elements $\{u_0,u_i,u^*_0,u^*_i\}$ through the use of Pauli matrices. This is possible by the identifications

$$u_0 = \begin{pmatrix} 0 & 0 \\ 0 & \omega_0 \end{pmatrix}, \qquad u^*_0 = \begin{pmatrix} \omega_0 & 0 \\ 0 & 0 \end{pmatrix},$$
$$u_i = \begin{pmatrix} 0 & 0 \\ \omega_i & 0 \end{pmatrix}, \quad u^*_i = \begin{pmatrix} 0 & -\omega_i \\ 0 & 0 \end{pmatrix}, \qquad i = 1,2,3\,. \tag{A5}$$

ω_0 and ω_i are given in (4.1) and (4.2). For an arbitrary split octonion A, we have

$$A = au^*_0 + bu_0 + x_i u^*_i + y_i u_i = \begin{pmatrix} a & -\mathbf{x} \\ \mathbf{y} & b \end{pmatrix}. \tag{A6}$$

The operation of octonionic conjugation is defined as

$$\overline{A} = bu^*_0 + au_0 - x_i u^*_i - y_i u_i = \begin{pmatrix} a & \mathbf{x} \\ -\mathbf{y} & b \end{pmatrix}. \tag{A7}$$

The norm of A in then given by

$$A\bar{A} = \bar{A}A = (ab + \mathbf{x}\cdot\mathbf{y})\cdot\mathbf{1}\,, \tag{A8}$$

where $\mathbf{1}$ is the identity element of algebra, $\mathbf{1} = 1\cdot u^*{}_0 + 1\cdot u_0$. The "matrices" written above are called "Zorn matrices" or "vectorial matrices."[17] If we define the product of two Zorn matrices as

$$AB = \begin{pmatrix} a & -\mathbf{x} \\ \mathbf{y} & b \end{pmatrix}\begin{pmatrix} c & -\mathbf{u} \\ \mathbf{v} & d \end{pmatrix} = \begin{pmatrix} ac - \mathbf{x}\cdot\mathbf{v} & -(a\mathbf{u} + d\mathbf{x} + \mathbf{y}\wedge\mathbf{v}) \\ c\mathbf{y} + b\mathbf{v} + \mathbf{x}\wedge\mathbf{u} & bd - \mathbf{y}\cdot\mathbf{u} \end{pmatrix}, \tag{A9}$$

the multiplication table for the u's is reproduced in this Zorn matrix notation. Besides, we have the following properties for octonions:

$$A\cdot\mathbf{1} = \mathbf{1}\cdot A = A\,, \quad \overline{AB} = \bar{B}\,\bar{A}\,, \tag{A10}$$

$$A + \bar{A} = \mathrm{Tr}(A)\cdot\,, \quad \mathrm{Tr}(AB) = \mathrm{Tr}(BA)\,,$$

where the trace operation (Tr) is performed on Zorn matrices. From (A10), and using the definition of the associator given in (A2) we have that

$$\mathrm{Tr}[(AB)C] = \mathrm{Tr}[A(BC)] = \mathrm{Tr}[ABC]\,, \tag{A11}$$

and thus the trace operation on a product of Zorn matrices follows the cyclic order of factors.

In general, the Cayley complex (split) algebra contain seven Euclidean vectorial subalgebras, as well as seven subalgebras of quaternions. This property follows from multiplication rules of the complex (split) base given in (A4) and from the definition (A3).

Finally we must observe that we are always using $A = Z(A)$, the Zorn matrix of the octonion A, because the Zorn algebra is isomorphic to the split ocotonionic algebra.

APPENDIX B: INTERNAL TRANSFORMATIONS FOR QUATERNIONS AND OCTONIONS

1. Quaternionic transformations

If we have a complex unimodular matrix A,

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

then

$$A^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

The conjugation operation of quaternionions and octonions is equivalent to the inversion of a unimodular matrix.

Let U be a quaternionic transformation matrix, or Q transformation,

$$U = m_0\omega_0 + m_i\omega_i\,, \tag{B1}$$

which is at the same time a transformation matrix of the SU(2) group, i.e., it satisfies the condition

$$U^{-1} = U^\dagger\,, \tag{B2}$$

where $U^\dagger = U^{T*}$ is the Hermitian conjugate of U. It is easy to show that

$$U^\dagger \equiv \bar{U}^* = m^*{}_0\omega_0 - m^*{}_i\omega_i\,, \tag{B3}$$

where $\bar{U}$ is the quaternionic conjugate of U, defined as

$$\bar{U} = m_0\omega_0 - m_i\omega_i\,. \tag{B4}$$

Thus U, being a unimodular matrix, must satisfy the property of the inverse unimodular matrix expressed above, i.e., $U^{-1} \equiv \bar{U}$, which, with the condition (B2) for matrices of the SU(2) group, results in the coefficients m_0, m_i, $i = 1,2,3$, from expansion (B1), being real. Therefore

$$\bar{U}U = U\bar{U} = U^{-1}U = UU^{-1} = (m_0^2 + m_i^2)\omega_0 = \mathbf{1}\,, \tag{B5}$$

because

$$m_0^2 + m_i^2 = \det U = \mathbf{1}\,. \tag{B6}$$

We must have then

$$U = e^{i\lambda\cdot\sigma} = e^{-\lambda\cdot\omega}\,, \tag{B7}$$

where $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ are real parameters.

The $\mathbb{Q}$ transformation of a quaternion ψ is defined by

$$\psi' = U\psi U^{-1}\,. \tag{B8}$$

If

$$\psi = \alpha_0\omega_0 + \alpha_i\omega_i\,, \tag{B9}$$

ψ' is given in terms of components, after (B8),

$$\begin{aligned}\psi' = {} & (m_0\alpha_0 m_0 + m_0\alpha_k m_k + m_k\alpha_0 m_k - m_k\alpha_k m_0)\omega_0 \\ & + [m_0\alpha_p m_0 - m_0\alpha_0 m_p + m_p\alpha_0 m_0 + m_k\alpha_k m_p \\ & + \epsilon_{klp}(m_k\alpha_l m_0 - m_0\alpha_k m_l) \\ & - \epsilon_{ijk}\epsilon_{klp} m_i\alpha_j m_l]\omega_p\,.\end{aligned} \tag{B10}$$

The symmetry group of this transformation is SU(2), which is homomorphic to the rotation group O_3. This is given through the relation[18]

$$U\omega_i U^{-1} = R_{ij}\omega_j\,, \quad i,j = 1,2,3\,. \tag{B11}$$

In terms of components it is

$$m_0^2\delta_{ip} + m_i m_p - 2\epsilon_{ijp} m_0 m_j + \epsilon_{ijk}\epsilon_{klp} m_l m_j = R_{ip}\,. \tag{B12}$$

For the local $\mathbb{Q}$ transformations, which are used in the space-time connected to internal quaternionic space, we must have

$$U = U(x) = m_0(x)\omega_0 + m_i(x)\omega_i \tag{B13}$$

and

$$U(x) = e^{-\lambda(x)\cdot\omega}\,, \tag{B14}$$

where now, $\lambda(x) = (\lambda_1(x)\lambda_2(x), \lambda_3(x))$ are real functions. Also, in this case, the coefficient of ψ in (B9) are functions of space-time coordinates.

2. Octonionic transformations

When we consider octonions, we may define an octonionic transformation, or $\mathbb{O}$ transformation, by means of the octonion $\mathbb{U}$, in our case, split

$$\mathbb{U} = p_0 u_0{}^* + p_i u^*{}_i + q_0 u_0 + q_i u_i\,, \quad i = 1,2,3\,, \tag{B15}$$

where p_0, p_i, q_0, q_i, $i = 1,2,3$, are real coefficients. Then,

$$\bar{\mathbb{U}} = q_0 u^*{}_0 - p_i u^*{}_i + p_0 u_0 - q_i u_i \tag{B16}$$

is the conjugate of $\mathbb{U}$. Comparing with quaternions, we can define

$$\mathbb{U}^{-1} \equiv \overline{\mathbb{U}} . \tag{B17}$$

Actually,

$$\overline{\mathbb{U}}\mathbb{U} = \mathbb{U}\widetilde{\mathbb{U}} = \mathbb{U}^{-1}\mathbb{U} = \mathbb{U}^{-1}\mathbb{U} = (p_0 q_0 + p_i q_i)\mathbb{1} , \tag{B18}$$

which will be equal to $\mathbb{1}$ if

$$p_0 q_0 + p_i q_i = 1 . \tag{B19}$$

In this case we have for $\mathbb{U}$,

$$\mathbb{U} = e^{-\delta \cdot \mathbf{u}^*} - \gamma \cdot \mathbf{u} , \tag{B20}$$

where δ and γ are real parameters.

For an extension of the case valid for quaternions, the $\mathbb{O}$ transformation of an octonion Ψ is defined by

$$\Psi' = \tfrac{1}{2}\left[(\mathbb{U}\Psi)\mathbb{U}^{-1} + \mathbb{U}(\Psi\mathbb{U}^{-1})\right] . \tag{B21}$$

We can show easily that

$$(\mathbb{U}\Psi)\mathbb{U}^{-1} = \mathbb{U}(\Psi\mathbb{U}^{-1}) , \tag{B22}$$

which simplifies the relation (B21) above to

$$\Psi' = \mathbb{U}\Psi\mathbb{U}^{-1} . \tag{B23}$$

If ψ is given in terms of components by

$$\psi = \rho_0 u^*{}_0 + \rho_i u^*{}_i + \kappa_0 u_0 + \kappa_i u_i , \tag{B24}$$

we have, after (B23),

$$\begin{aligned}\Psi' = {} & (p_0\rho_0 q_0 + p_0\rho_k q_k + p_k\kappa_0 q_k - p_k\kappa_k q_0)u^*{}_0 \\ & + [\, p_0\rho_p p_0 - p_0\rho_0 p_p + p_p\kappa_0 p_0 + p_p\kappa_k p_k \\ & + \epsilon_{klp}(p_0\kappa_l q_k - q_0\kappa_k q_l) - \epsilon_{ijk}\epsilon_{klp}\, p_i\, \rho_j q_l\,]u^*{}_p \\ & + (q_0\kappa_0 p_0 + q_0\kappa_k p_k + q_k\rho_0 p_k - q_k\rho_k p_0)u_0 \\ & + [\, q_0\kappa_p q_0 - q_0\kappa_0 q_p + q_p\rho_0 q_0 + q_k\rho_k q_p \\ & + \epsilon_{klp}(p_k\rho_l q_0 - p_0\rho_k p_l) - \epsilon_{ijk}\epsilon_{klp} q_i\kappa_j\, p_l\,]u_p . \end{aligned} \tag{B25}$$

We can obtain a relation similar to (B8), for O transformations:

$$\mathbb{U}(u^*{}_i + u_i)\mathbb{U}^{-1} = k_{ij}u^*{}_j + l_{ij}u_j, \quad i,j = 1,2,3 . \tag{B26}$$

In terms of components it is

$$p_0^2\delta_{ip} + p_i p_p - \epsilon_{ilp}q_j(p_0 + q_0) + \epsilon_{ijk}\epsilon_{klp}\, p_j q_l = k_{ip} , \tag{B27}$$

$$q_0^2\delta_{ip} + q_i q_p - \epsilon_{ilp}\, p_j(p_0 + q_0) + \epsilon_{ijk}\epsilon_{klp} q_j\, p_l = l_{ip} ,$$

with the additional condition

$$p_0 q_i - p_i q_0 = 0 . \tag{B28}$$

From (B27) we see clearly that in the limit $q_{0,i} \to p_{0,i}$, the $\mathbb{O}$ transformation, through $\mathbb{U}$, is equivalent to a $\mathbb{Q}$ transformation, through $\mathbb{U}$. Therefore, the $\mathbb{O}$ transformation is homomorphic to the rotation group O_3 and so it must be SU(2)-like (in our case). This is due to a mixture of terms observed in (B27). Besides we observe a certain symmetry, with regard to the positioning of components terms, in expressions (B25) and (B27).

For the local $\mathbb{O}$ transformations, which are used in the octonionic space, we must have

$$\Psi = \Psi(x) = \rho_0(x)u^*{}_0 + \rho_i(x)u^*{}_i + \kappa_0(x)u_0 + \kappa_i(x)u_i ,$$

$$\mathbb{U} = \mathbb{U}(x) = p_0(x)u^*{}_0 + p_i(x)u^*{}_i + q_0(x)u_0 + q_i(x)u_i ,$$

and

$$\mathbb{U}(x) = e^{-\delta(x)\cdot\mathbf{u}^* - \gamma(x)\cdot\mathbf{u}} ,$$

where $\delta(x)$ and $\gamma(x)$ are now real functions of the space-time coordinates.

[1] A. Einstein, Rev. Mod. Phys. **20**, 35 (1948); M. A. Tonnelat, *Les Theories Unitaires de l'Électromagnétisme e de la Gravitation* (Gauthier–Villars, Paris, 1965).

[2] A. Einstein, *The Meaning of Relativity* (Princeton U.P., Princeton, NJ, 1955), Apendix 2.

[3] Mann and Moffat [R. B. Mann and J. W. Moffat, Phys. Rev. D **26**, 1858 (1982)] and Kunstatter, Leivo, and Savaria [G. Kunstatter, P. Leivo, and P. Savaria, Class. Quantum Grav. **1**, 7 (1984)] showed that the physical sector of the nonsymmetric theory, based on the Einstein–Strauss field equations [A. Einstein and E. G. Strauss, Ann. Math. **47**, 731 (1946)], corresponds to spin-2 and spin-0, i.e., the skew part $g_{\mu\nu}$ must have spin-0 in the weak field approximation, and therefore cannot be interpreted as the electromagnetic spin-1 field $F_{\mu\nu}$.

[4] J. W. Moffat, Phys. Rev. D **19**, 3554 (1979); J. W. Moffat, J. Math. Phys. **21**, 1798, (1980).

[5] K. Borchsenius, Phys. Rev. D **13**, 2707 (1976).

[6] We want to thank Professor L. A. Maia, Department of Mathematics, Universidade de Brasília, for important elucidations on the aspects of octonionic algebra.

[7] In a recent paper, Moffat [J. W. Moffat, J. Math. Phys. **25**, 347 (1984)] treats an octonionic Riemannian geometry. However, he takes into account only the geometrical formulation of the nonsymmetric theory of gravitation, which is not the same as the one we are exposing in this work, that treats an extension of the unified (quaternionic) Borchsenius theory of gravitation, electromagnetism, and Yang–Mills fields.

[8] A. Cayle, "*Note on system of imaginaires*," Phys. Mag. (London) **30**, 157 (1847); *The Collected Mathematical Papers I* (Cambridge, U. P. Cambridge, 1889), p. 301; M. Gunaydin and F. Gursey, J. Math. Phys. **14**, 1651 (1973); L. A. Maia, "Sobre a Analiticidade de Funções em uma Álgebra de Divisão," thesis, Universidade de Brasília, Departamento de Matemática (1982).

[9] V. Hlavaty, *Geometry of Einstein's Unified Field Theory* (Nordhoff, Groningen, 1957). See also the work of Kunstatter and Yates [G. Kunstatter and R. Yates, J. Phys. A **14**, 847 (1981)]. More recently, there is the work of Kunstatter, Moffat, and Malzan [G. Kunstatter, J. W. Moffat, and J. Malzan, J. Math. Phys. **24**, 886 (1983)], which extends general relativity to the algebra of hyperbolic complex numbers.

[10] A. Einstein, Ann. Math. **46**, 578 (1945); A. Einstein and E. G. Strauss, Ann. Math. **47**, 731 (1946).

[11] C. G. Oliveira, in *II Escola de Cosmologia e Gravitação*, edited by M. Novello (*João Pessoa*,Brazil, 1980); Rev. Brasil. Fís. **10**, 121 (1980).

[12] J. W. Moffat and D. H. Boal, Phys. Rev. D **11**, 1375 (1975); W. B. Bonnor, Proc. R. Soc. London Ser. A **226**, 336 (1954).

[13] R. L. Mills and C. N. Yang, Phys. Rev. **96**, 191 (1954).

[14] R. B. Mann and J. W. Moffat, Phys. Rev. D **26**, 1858 (1982).

[15] D. Finkelstein, J. M. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **3**, 207 (1962); D. Finkelstein, J. M. Jauch, and D. Speiser, J. Math. Phys. **4**, 136 (1963); D. Finkelstein, J. M. Jauch, S. Schiminovich, and D. Speiser, J. Math. Phys. **4**, 788 (1963). The last of these papers about quaternionic quantum mechanics studies the properties of a quaternion Hilbert space local to the space-time manifold.

[16] M. Günaydin and F. Gürsey, J. Math. Phys. **14**, 1651 (1973); M. Günaydin and F. Gürsey, Phys. Rev. D **9**, 3387 (1974); M. Günaydin, J. Math. Phys. **17**, 1875 (1976).

[17] M. Zorn, Abh. Math. Sem. Univ. Hamburg **8**, 123 (1931).

[18] J. Aharoni, *The Special Theory of Relativity* (Oxford U. P., London, 1965), 2nd Ed., Eq. (11.6), p. 297.

Two-mode para-Bose coherent states

G. M. Saxena
Time and Frequency Section, National Physical Laboratory, New Delhi-110 012, India
C. L. Mehta[a)]
Department of Physics, Kansas State University, Manhattan, Kansas 66506

(Received 6 April 1984; accepted for publication 24 May 1985)

The construction of eigenstates of the square of annihilation operators for a two-mode para-Bose system is reported. Bose coherent states can be deduced from these eigenstates as a special case. These states are termed para-Bose coherent states. These states are degenerate. The expansion of the coherent states in terms of two-mode para-Bose energy eigenstates has been obtained and their salient properties are discussed. Also discussed is the uncertainty product of the square of position and momentum operators in the para-Bose coherent states for a two-mode system.

I. INTRODUCTION

We discuss in this paper the construction and properties of two-mode para-Bose coherent states. These states are defined as eigenstates of the square of annihilation operators.[1] It may be recalled that for normal Bose and one-mode para-Bose systems,[2] the coherent state is defined as the eigenstate of the annihilation operator a_i, i.e.,

$$a_i|z_1, z_2 \cdots z_i \cdots\rangle = z_i|z_1, z_2 \cdots z_i \cdots\rangle \tag{1.1a}$$

(normal Bose case) and

$$a|z, L\rangle = z|z, L\rangle \tag{1.1b}$$

(one-mode para-Bose case). Here L is a positive integer known as the order of parastatistics.[3] For bosons, $L = 1$, while for parabosons, $L > 1$. The minimum energy of a para-Bose system depends on L. For a Bose system ($L = 1$) the operators belonging to different modes commute among themselves, i.e.,

$$[a_i, a_j] = 0. \tag{1.2}$$

As a result it is possible to find simultaneous eigenstates of $a_i (i = 1, 2, ...)$. For a one-mode para-Bose system it is again possible to define a coherent state as an eigenstate of a [cf. Eq. (1.1b)]. However, for a para-Bose system with two or higher modes, we are confronted with the problem of the non-commuting nature of the annihilation operators belonging to different modes. Therefore simultaneous eigenstates of annihilation operators cannot be constructed. On the other hand, the square of annihilation operators and $J_{jk} = \frac{1}{2}[a_k, a_j]$ commute among themselves for a multimode para-Bose system, i.e.,

$$[a_i^2, a_j^2] = 0, \quad [a_j^2, J_{jk}] = 0, \tag{1.3}$$

and we define para-Bose coherent states as the simultaneous eigenstates of a_i^2 and J_{ik} $(i,k = 1,2,...)$. We shall discuss only the two-mode para-Bose coherent state, a representative of the multimode system. The contents of the paper are arranged as follows. In Sec. II we shall define the two-mode para-Bose coherent state and obtain its expansion in terms of energy eigenstates. The matrix elements of Hamiltonian and the squares of the position and the momentum operators will be obtained in the coherent state, and the familiar results of the normal Bose and one-mode para-Bose cases will be deduced as special cases. In Sec. III, the uncertainty relation for the squares of position and momentum operators with their expectation values in the coherent state will be discussed.

II. DEFINITION OF PARA-BOSE COHERENT STATES

We define the two-mode para-Bose coherent state $|z_1, z_2, z, L\rangle$ as an eigenstate of a commuting set of operators a_1^2, a_2^2, and $J = \frac{1}{2}[a_2, a_1]$ ($J = 0$ for normal bosons) with eigenvalues z_1^2, z_2^2, and z, respectively, i.e.,

$$a_1^2|z_1, z_2, z, L\rangle = z_1^2|z_1, z_2, z, L\rangle, \tag{2.1a}$$

$$a_2^2|z_1, z_2, z, L\rangle = z_2^2|z_1, z_2, z, L\rangle, \tag{2.1b}$$

and

$$J|z_1, z_2, z, L\rangle = z|z_1, z_2, z, L\rangle. \tag{2.1c}$$

We shall see that coherent states $|z_1, z_2, z, L\rangle$ are not orthogonal. The expansion of $|z_1, z_2, z, L\rangle$ will be obtained in terms of energy eigenstates,[4,5] $|nlm\rangle$ for the two-mode para-Bose system. The energy eigenstate $|n,l,m\rangle$ is defined as

$$|nlm\rangle = \frac{1}{\sqrt{N_{nlm}}}(\{a_1^+, a_2\})^m (a_2^+)^l J^{+n}|0\rangle \quad (l > m). \tag{2.2}$$

Here $|0\rangle$ is the vacuum state such that

$$\{a_i, a_j^+\}|0\rangle = L\delta_{ij}|0\rangle, \tag{2.3}$$

with L being the order of the parastatistics. Here J^+ is the Hermitian conjugate of J, and J^+ creates one particle each in both the modes. Its nonvanishing nature for parabosons gives rise to antisymmetrical states. For the state $|nlm\rangle$ the number of excitations of first and second modes are $n + m$ and $l + n - m$, respectively. Hence there is degeneracy if only $(m + n)$ and $(l + n - m)$ are fixed. It is because $J^+ \neq 0$. The Bose energy states are nondegenerate as $J^+ = 0$. The normalization constant[5] N_{nlm} is given by

[a)] Permanent address: Department of Physics, Indian Institute of Technology, Hauz Khas, New Delhi-110 016, India.

 0022-2488/85/123140-06$02.50

$$N_{nlm} = \frac{l!m!(l-\Lambda_l)!!(n+l+1-\Lambda_{n+l})!!(n+l+L-2+\Lambda_{n+l})!!(n-\Lambda_n)!!(n+L-3+\Lambda_n)!!}{(l+1-\Lambda_l)!!(l-m)!(L-2)!}, \tag{2.4}$$

where Λ_k is the projection to odd k values, i.e.,

$$\Lambda_k = \tfrac{1}{2}[1-(-1)^k] = 0, \quad \text{if } k \text{ is even}, \tag{2.5}$$
$$\Lambda_k = \tfrac{1}{2}[1-(-1)^k] = 1, \quad \text{if } k \text{ is odd}.$$

The states $|nlm\rangle$ are orthonormal, i.e.,

$$\langle n'l'm'|nlm\rangle = \delta_{nn'}\delta_{ll'}\delta_{mm'}. \tag{2.6}$$

Let us define the expansion of $|z_1, z_2, z, L\rangle$, in terms of energy eigenstates $|nlm\rangle$, as

$$|z_1, z_2, z, L\rangle = \sum_{nlm} C_{nlm}|nlm\rangle \quad (l>m). \tag{2.7}$$

Here C_{nlm} is the expansion coefficient and is a function of z_1, z_2, z and L. The Hilbert–Fock space of our two-mode system spanned by states $|nlm\rangle$ is composed of the following subspaces, which are invariant with respect to all the generators of symplectic $Sp_{4,R}$ algebra: (i) the subspace $S_1(L)$ spanned by all the vectors $|nlm\rangle$ having both n and l even; (ii) the subspace $S_2(L)$ spanned by all the state vectors having n odd and l even; and (iii) the subspace $S_3(L)$ spanned by all the vectors having l odd.

Thus the expansion of $|z_1, z_2, z, L\rangle$ in $|nlm\rangle$ will yield degenerate eigenstates of a_i^2 $(i=1,2)$ and J because of invariance of the subspaces with respect to a_i^2. It will be a trivial exercise to verify that even for bosons and the one-mode para-Bose system the eigenstates of a_i^2 are degenerate. But if the eigenstates of the square of the annihilation operator are allowed to be eigenstates of the annihilation operator as well for bosons and the one-mode para-Bose system then the degeneracy vanishes. By definition, the two-mode para-Bose coherent state is an eigenstate of a_1^2, a_2^2, and J. It can be verified that, on putting $n=0$ in the expansion (2.7), the eigenvalue, z, of J for the eigenstate $|z_1, z_2, z, L\rangle$ becomes zero and the states $|z_1, z_2, z=0, L\rangle$ span the subspace of the Hilbert space. Therefore for the two-mode para-Bose coherent state $|z_1, z_2, z, L\rangle$ to span the complete Hilbert space, it is necessary that it should be an eigenstate of three commuting operators a_1^2, a_2^2, and J. These three operators ensure the completeness of the two-mode para-Bose coherent states.

Now let us consider the expansion of states $|z_1, z_2, z, L\rangle$ in terms of basis $|nlm\rangle$ using Eq. (2.2), (2.4), (2.6), and (2.7). In the process of expansion we obtain the following recurrence relations.

(I) *For even n and l values,*

$$z_1^2 C_{nlm} = \left[\frac{(m+2)(m+1)(n+l+L)(n+l+3)}{(l+3)(l+1)}\right]^{1/2} C_{n,l+2,m+2} - \left[\frac{(m-l+1)(m-l)(n+2)(n+L-1)}{(l+1)(l-1)}\right]^{1/2} C_{n+2,l-2,m}, \tag{2.8}$$

$$z_2^2 C_{nlm} = \left[\frac{(n+L+1)(n+l+3)(l-m+2)(l-m+1)}{(l+3)(l+1)}\right]^{1/2} C_{n,l+2,m} - \left[\frac{m(m-1)(n+2)(n+L-1)}{(l+1)(l-1)}\right]^{1/2} C_{n+2,l-2,m-2}, \tag{2.9}$$

$$zC_{nlm} = [(n+L+1)(L+n-1)]^{1/2} C_{n+1,l,m}. \tag{2.10}$$

(II) *For even n and odd l,*

$$z_1^2 C_{nlm} = \frac{2}{l(l+2)}[(m+1)(l-m)(n+l+2)(L+n-1)]^{1/2} C_{n+1,l,m+1} + (l+2)^{-1}[(m+2)(m+1)(n+l+2)(n+L+l+1)]^{1/2} C_{n,l+2,m+2} - l^{-1}[(l-m)(l-m-1)(n+2)(n+L-1)]^{1/2} C_{n+2,l-2,m}, \tag{2.11}$$

$$z_2^2 C_{nlm} = \frac{-2}{l(l+2)}[m(l-m+1)(l+n+2)(L+n-1)]^{1/2} C_{n+1,l,m-1} + (l+2)^{-1}[(l-m+2)(l-m+1)(n+l+2)(n+l+L+1)]^{1/2} C_{n,l+2,m} - l^{-1}[(n+2)m(m-1)(L+n-1)]^{1/2} C_{n+2,l-2,m-2}, \tag{2.12}$$

$$zC_{nlm} = -[(n+L-1)(n+l+2)]^{1/2} C_{n+1,l,m}. \tag{2.13}$$

(III) *For odd n and even l,*

$$z_1^2 C_{nlm} = \left[\frac{(m+2)(m+1)(n+l+2)(n+l+L+1)}{(l+1)(l+3)}\right]^{1/2} C_{n,l+2,m+2} - \left[\frac{(l-m)(l-m-1)(n+1)(n+L)}{(l-1)(l+3)}\right]^{1/2} C_{n+2,l-2,m}, \tag{2.14}$$

$$z_2^2 C_{nlm} = \left[\frac{(n+L+l+1)(n+l+2)(l-m+2)(l-m+1)}{(l+1)(l+3)}\right]^{1/2} C_{n,l+2,m}$$
$$- \left[\frac{m(m-1)(n+L)n+1)}{(l-1)(l+1)}\right]^{1/2} C_{n+2,l-2,m-2}, \tag{2.15}$$
$$zC_{nlm} = [(n+1)(n+l+2)]^{1/2} C_{n+1,l,m}. \tag{2.16}$$

(IV) *For odd n and odd l case,*

$$z_1^2 C_{nlm} = \frac{2}{l(l+2)} [(m+1)(l-m)(n+l+L)(n+1)]^{1/2} C_{n+1,l,m+1}$$
$$+ (l+2)^{-1}[(m+2)(m+1)(n+l+L)(n+l+3)]^{1/2} C_{n,l+2,m+2}$$
$$- l^{-1}[(l-m-1)(l-m)(n+1)(L+n)]^{1/2} C_{n+2,l-2,m}, \tag{2.17}$$
$$z_2^2 C_{nlm} = \frac{-2}{1(1+2)} [(n+1)(n+l+L)m(l-m+1)]^{1/2} C_{n+1,l,m-1}$$
$$- l^{-1}[(n+1)(n+L)m(m-1)]^{1/2} C_{n+2,l-2,m-2}$$
$$+ (l+2)^{-1}[(l-m+1)(l-m+2)(n+L+l)(n+l+3)]^{1/2} C_{n,l+2,m}, \tag{2.18}$$
$$zC_{nlm} = -[(n+1)(L+l+n)]^{1/2} C_{n+1,l,m}. \tag{2.19}$$

On solving the recurrence relations (2.8)–(2.19) we obtain, for all even and odd values of n and l,

$$C_{n,l,m} = (1-2\mu) z^n \left[\frac{(l+1-\Lambda_l)!!(l+L-2+\Lambda_l)!!(L-3)!!}{(n-\Lambda_n)!!(l+n+1-\Lambda_{n+l})!!(n+L-3+\Lambda_n)!!(l+n+L-2+\Lambda_{n+l})!!}\right]^{1/2} C_{0,l,m}, \tag{2.20}$$

where μ is defined as

$\mu = 0$, for any combination of n and l when both are not odd;

$\mu = 1$, when n and l both are odd.

We can also write $C_{0,l,m}$ for all values of l and m as

$$C_{0,l,m} = (z_1)^m (z_2)^{l-m}$$
$$\times \left[\frac{\Gamma([(l+1)/2] + \frac{1}{2})\Gamma(L/2)\Gamma(\frac{1}{2})}{2^l \Gamma([m/2]+1)\Gamma([(l-m)/2]+1)\Gamma([(l-m+1)/2]+\frac{1}{2})\Gamma([(m+1)/2]+\frac{1}{2})\Gamma([(l+1)/2]+L/2)}\right]^{1/2}$$
$$\times C_{000} F_{l,m}(z, z_1, z_2). \tag{2.21}$$

Here,

$$C_{0,2,1} \cong (z_1 z_2/\sqrt{L}) C_{000}$$

is taken and $[Y]$ stands for the largest integer smaller than or equal to Y.
The $F_{lm}(z, z_1, z_2)$ are defined as

$$F_{2l,2m} = 1 + \sum_{p=1}^{m} \frac{2^p m!(2l-2p-1)!!(l-m)!}{(m-p)!p!(2l-1)!!(l-m-p)!} \left(\frac{z}{z_1 z_2}\right)^{2p}, \tag{2.22a}$$
$$F_{2l,2m+1} = 1 + \sum_{p=1}^{m} \frac{2^p m!(2l-2p-1)!!(l-m-1)!}{(m-p)!p!(2l-1)!!(l-m-p-1)!} \left(\frac{z}{z_1 z_2}\right)^{2p}, \tag{2.22b}$$
$$F_{2l+1,2m} = 1 + \sum_{p=0}^{m-1} \frac{2^{p+1} m!(2l-2p-1)!!(l-m)!}{(m-p-1)!p!(2l+1)!!(l-m-p)!} \left(\frac{z}{z_1 z_2}\right)^{2p+1} + \sum_{q=1}^{m} \frac{2^q m!(2l-2q+1)!!(l-m)!}{(m-q)!q!(2l+1)!!(l-m-q)!} \left(\frac{z}{z_1 z_2}\right)^{2q}, \tag{2.22c}$$
$$F_{2l+1,2m+1} = 1 + \sum_{p=1}^{m} \frac{2^p m!(2l-2p+1)!!(l-m)!}{(m-p)!p!(2l+1)!!(l-m-p)!} \left(\frac{z}{z_1 z_2}\right)^{2p} - \sum_{q=0}^{m} \frac{2^{q+1} m!(2l-2q-1)!!(l-m)!}{q!(m-q)!(2l+1)!!(l-q-m-1)!} \left(\frac{z}{z_1 z_2}\right)^{2q+1}. \tag{2.22d}$$

From Eqs. (2.20) and (2.21) and the above derivations we may obtain C_{nlm} for all values of n, l, and m.

The C_{000} can be evaluated using the normalization property of the state $|z_1, z_2, z\rangle$, i.e.,

$$\langle z_1, z_2, z | z_1, z_2, z\rangle = 1. \tag{2.23}$$

Up to an arbitrary phase it comes out to be

$$C_{000}=\left[\sum_{nlm}|z_1|^{2m}|z_2|^{2l-2m}|z|^{2n}\left\{\frac{(1-2\mu)(l+1-\Lambda_l)!!(L+l-2+\Lambda_l)!!}{(n-\Lambda_n)!!(l+n+1-\Lambda_{n+l})!!(L+n-3+\Lambda_n)!!(l+n+L-2+\Lambda_{n+l})!!}\right.\right.$$
$$\left.\left.\times\frac{(L-3)!!\Gamma([(l+1)/2]+\frac{1}{2})\Gamma(L/2)\Gamma(\frac{1}{2})}{2^l\Gamma([m/2]+1)\Gamma([(l-m)/2]+1)\Gamma([l-m+1)/2]+\frac{1}{2})\Gamma([(m+1)/2]+\frac{1}{2})}\right\}\Gamma([(l+1)/2]+L/2)\right]^{-1/2}$$
$$\times|F_{l,m}(z,z_1,z_2)|^{-1}. \tag{2.24}$$

If we put $L=1$ in the above equation with $n=0$, and $F_{l,m}=1$, then

$$C_{000}=\left[\sum_{l,m}|z_1|^{2m}|z_2|^{2l-2m}\left\{\frac{(\Gamma(\frac{1}{2}))^2}{2^l\Gamma([m/2]+1)\Gamma([(l-m)/2]+1)\Gamma([(l-m+1)/2]+\frac{1}{2})\Gamma([m+1)/2]+\frac{1}{2})}\right\}\right]^{-1/2}$$
$$=\left[\sum_{l,m}|z_1|^{2m}|z_2|^{2l-2m}\frac{1}{m!(l-m)!}\right]^{-1/2}=e^{-(|z_1|/2)^2}e^{-(|z_2|/2)^2}, \tag{2.25}$$

which is the familiar expression of the normal boson case.

We shall now obtain the matrix elements of energy (free Hamiltonian), square of position, and momentum operators in these para-Bose coherent states.

First we shall express the Hamiltonian, square of position, and momentum operators in terms of annihilation and creation operators. We have for two-mode para-Bose system,

$$H=\tfrac{1}{2}\{a_1,a_1^+\}+\tfrac{1}{2}\{a_2,a_2^+\}, \tag{2.26}$$
$$q_i^2=\tfrac{1}{2}(\{a_i,a_i^+\}+a_i^2+a_i^{+2}), \tag{2.27a}$$
$$p_i^2=\tfrac{1}{2}(\{a_i,a_i^+\}-a_i^2-a_i^{+2})\quad(i=1,2). \tag{2.27b}$$

Using Eqs. (2.7), (2.26), and (2.27) we obtain the matrix elements of H, q_i^2, and p_i^2, which are given by the following equations:

$$\langle z_1,z_2,z|H|z_1,z_2,z\rangle=\sum_{nlm}|C_{nlm}|^2(2n+l+L), \tag{2.28}$$
$$\langle z_1,z_2,z|q_1^2|z_1,z_2,z\rangle=\frac{1}{2}(z_1^2+z_1^{*2})+\sum_{nlm}|C_{nlm}|^2\left(n+m+\frac{L}{2}\right), \tag{2.29a}$$
$$\langle z_1,z_2,z|q_2^2|z_1,z_2,z\rangle=\frac{1}{2}(z_2^2+z_2^{*2})+\sum_{nlm}|C_{nlm}|^2\left(n+l-m+\frac{L}{2}\right), \tag{2.29b}$$
$$\langle z_1,z_2,z|p_1^2|z_1,z_2,z\rangle=\sum_{nlm}|C_{nlm}|^2\left(n+m+\frac{L}{2}\right)-\frac{1}{2}(z_1^2+z_1^{*2}), \tag{2.30a}$$

and

$$\langle z_1,z_2,z|p_2^2|z_1,z_2,z\rangle=\sum_{nlm}|C_{nlm}|^2\left(n+1-m+\frac{L}{2}\right)-\frac{1}{2}(z_2^2+z_2^{*2}). \tag{2.30b}$$

The $C_{n,l,m}$ can be substituted from Eqs. (2.20) and (2.21). Here we have also made use of the fact that $|nlm\rangle$ is an eigenstate of $\frac{1}{2}\{a_1,a_1^+\}$ and $\frac{1}{2}\{a_2,a_2^+\}$ with eigenvalues $(n+m+L/2)$ and $(n+l-m+L/2)$, respectively.

We shall now deduce the familiar results of the normal Bose system and of the one-mode para-Bose system from the two-mode para-Bose system as special cases.

(i) Normal Bose system: In the normal Bose case, $L=1$, $n=0$, and $J=\frac{1}{2}[a_2,a_1]=0$. On putting these values in Eqs. (2.20)–(2.22), we obtain

$$C_{0,l,m}=z_1^m z_2^{l-m}[(\Gamma(\tfrac{1}{2}))^2/2^l\Gamma([m]/2+1)\Gamma([l-m/2]+1)\Gamma([l-m+1]/2+\tfrac{1}{2})\Gamma([m+1]/2+\tfrac{1}{2})]^{1/2}. \tag{2.31}$$

We know that

$$K!=2^K\Gamma([K]/2+1)\Gamma([K+1]/2+\tfrac{1}{2})/\Gamma(\tfrac{1}{2}).$$

So Eq. (2.31) can then be rewritten as

$$C_{0,l,m}=z_1^m z_2^{l-m}[1/m!(l-m)!]^{-1/2}C_{000}. \tag{2.32}$$

This is the familiar result obtained from the expansion

$$|z_1,z_2\rangle=\sum_{l,m}C_{l,m}|l,m\rangle$$

for the normal Bose system. The constant C_{000} is given by (2.25).

(ii) One-mode para-Bose system: If we put either $m=0$, or $l=m$ (and $L\neq 1$), we can realize the one-mode para-Bose system. Here again $n=0$ and $F_{lm}=1$. We obtain from Eqs. (2.20) and (2.21) the relations

(i) for $m = 0$,

$$C_{0,l,0} = z_2^l \left[\Gamma(L/2)/2^l \Gamma([l/2]+1)\Gamma([l+1/2]+L/2)\right]^{1/2} C_{000}\,; \tag{2.33a}$$

(ii) for $l = m$,

$$C_{0,l,l} = z_1^l \left[\Gamma(L/2)/2^l \Gamma([l/2]+1)\Gamma([l+1/2]+L/2)\right]^{1/2} C_{000}\,. \tag{2.33b}$$

These results are in agreement with the results of the one-mode para-Bose system.[2]

We just observed that the eigenstates of a_i^2 $(i = 1,2)$ for the two-mode para-Bose system readily give the familiar results for normal and one-mode para-Bose coherent states. We can, therefore, say that the above deductions also fulfill the requirement for the states $|z_1, z_2, z, L\rangle$ to be the coherent state.

III. POSITION-MOMENTUM UNCERTAINTY RELATION

For a multimode para-Bose system the position q_i (momentum p_i) operators belonging to different modes do not commute. Consequently their simultaneous eigenstates cannot be constructed. However, the square of the position q_i^2 (momentum p_i^2) operators commute and we consider them observables. Their simultaneous eigenstates can be defined. The commutator of the square of the position (q_i^2) and the momentum (p_i^2) operators is not a C number, i.e.,

$$[q_i^2, p_i^2] = \eta\{2\eta(a_i^{+2} - a_i^2)\} \quad (\eta = \sqrt{-1})\,. \tag{3.1}$$

The operator within the curly bracket on the right-hand side (rhs) is a Hermitian operator. We know that for the Hermitian operators A, B, and C, which satisfy the commutator relation

$$[A,B] = \eta C\,, \tag{3.2}$$

the uncertainties $\langle(\Delta A)^2\rangle$ and $\langle(\Delta B)^2\rangle$ satisfy the inequality[6]

$$\langle(\Delta A)^2\rangle\langle(\Delta B)^2\rangle \geqslant \tfrac{1}{4}\langle C\rangle^2\,. \tag{3.3}$$

As a consequence for the operators q_i^2 qand p_i^2 we have the inequality

$$\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle \geqslant \tfrac{1}{4}\langle 2\eta(a_i^{+2} - a_i^2)\rangle^2\,. \tag{3.4}$$

Taking the averaged value of the uncertainties in the square of the position and momentum operators in the para-Bose coherent states $|z_1, z_2, z, L\rangle$ for a two-mode case, we should have

$$\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle \geqslant -(z_i^{*2} - z_i^2)^2 \quad (i = 1,2)\,. \tag{3.5}$$

The rhs of the inequality depends on the state and it is a positive quantity. The para-Bose coherent states will be minimum uncertainty product states only if relation (3.5) is an equality and $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle/|2\eta(z_i^2 - z_i^2)|$ is a minimum. In the para-Bose coherent states

$$\left(|z_1, z_2, z, L\rangle = \sum C_{nlm}|nlm\rangle\right)$$

the uncertainties $\langle(\Delta q_1^2)^2\rangle$ and $\langle(\Delta p_1^2)^2\rangle$ have the values

$$\langle(\Delta q_1^2)^2\rangle = \langle q_1^4\rangle - \langle q_1^2\rangle^2$$
$$= \left[\sum_{nlm}|C_{nlm}|^2(n+m)^2 + \sum_{nlm}|C_{nlm}|^2(n+m) \times \left\{1 - \sum_{nlm}|C_{nlm}|^2(n+m)\right\} + z_1^2 + z_1^{*2} + \frac{L}{2}\right], \tag{3.6}$$

and

$$\langle(\Delta p_1^2)^2\rangle = \langle p_1^4\rangle - \langle p_1^2\rangle^2$$
$$= \left[\sum_{nlm}|C_{nlm}|^2(n+m)^2 + \sum_{nlm}|C_{nlm}|^2(n+m) \times \left\{1 - \sum_{nlm}|C_{nlm}|^2(n+m)\right\} - z_1^2 - z_1^{*2} + \frac{L}{2}\right]. \tag{3.7}$$

Similarly $\langle(\Delta q_2^2)^2\rangle$ and $\langle(\Delta p_2^2)^2\rangle$ can be evaluated. The uncertainty products $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle$ $(i = 1,2)$ in the para-Bose coherent state are then given by

$$\langle(\Delta q_1^2)^2\rangle\langle(\Delta p_1^2)^2\rangle$$
$$= \left[\left\{\sum_{nlm}|C_{nlm}|^2(n+m) \times\left(1+n+m - \sum_{nlm}|C_{nlm}|^2(n+m)\right) + \frac{L}{2}\right\}^2 - z_1^4 - z_1^{*4} - 2|z_1|^4\right], \tag{3.8}$$

and

$$\langle(\Delta q_2^2)^2\rangle\langle(\Delta p_2^2)^2\rangle$$
$$= \left[\left\{\sum_{nlm}|C_{nlm}|^2(n+l-m) \times\left(1+n+l-m - \sum_{nlm}|C_{nlm}|^2(n+l-m)\right) + \frac{L}{2}\right\}^2 - z_2^4 - z_2^{*4} - 2|z_2|^4\right]. \tag{3.9}$$

The values of $|C_{nlm}|^2$ can be substituted from Eqs. (2.20)–(2.22).

The value of $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle$ $(i = 1,2)$, obtained from Eqs. (3.8) and (3.9), when substituted in the relation (3.5), will not make it an equality. The rhs of the relation (3.5) depends on the state. Therefore the uncertainty product $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle$ or the quantity $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle \times [\,|\langle 2\eta(a_i^{+2} - a_i^2)\rangle|\,]^{-1}$ will not be the minimum for the two-mode para-Bose coherent states. The above product depends on the state even for the normal coherent state since the commutator $[q_i^2, p_i^2]$ is an operator. However, for the normal coherent state (defined as an eigenstate of a_i^2 and a_i) the uncertainty product $\langle(\Delta q_i)^2\rangle\langle(\Delta p_i)^2\rangle$ is the minimum.

But for two- or higher-mode para-Bose coherent states the uncertainty product $\langle(\Delta q_i)^2\rangle\langle(\Delta q_i)^2\rangle$ does not have any significance, as for a para-Bose system q_i^2 and p_i^2 are considered observables.

IV. CONCLUSION

We have defined in this paper eigenstates of a_1^2, a_2^2, and J and called them two-mode para-Bose coherent states. The nomenclature "para-Bose coherent states" has been used for these states as normal coherent states, and the one-mode para-Bose coherent states can be deduced from them as special cases. The two-mode para-Bose coherent state is an eigenstate, not of two but three operators, as otherwise these states are not complete and will span only a subspace of the Hilbert space of the para-Bose system. We also observe that because of the operator J, different modes are not independent of each other, unlike in the normal coherent states. For the para-Bose coherent state the uncertainty product or the quantity $\langle(\Delta q_i^2)^2\rangle\langle(\Delta p_i^2)^2\rangle/|\langle 2\eta(a_i^{+2}-a_i^2)\rangle|$ is not the minimum. It depends on the state and L, the order of parastatistics.

[1]G. M. Saxena, Ph.D. thesis, I.I.T., Delhi, 1983.
[2]J. K. Sharma, C. L. Mehta, and E. C. G. Sudarshan, J. Math. Phys. **19**, 2089 (1978).
[3]H. S. Green, Phys. Rev. **90**, 270 (1953).
[4]C. Alabiso, F. Duimio, and J. L. Redondo, Nuovo Cimento A **61**, 766 (1969).
[5]G. M. Saxena and C. L. Mehta, Pramāna **24**, 583 (1985).
[6]S. Kumar and C. L. Mehta, Pramana **10**, 75 (1978).

On the energy–time conjugation in quantum physics

F. Bailly
Laboratoire de Physique des Solides, Centre National de la Recherche Scientifique, 1, Place Aristide Briand, 92195 Meudon Principal Cedex, France

(Received 18 February 1985; accepted for publication 31 May 1985)

The energy–time conjugation is discussed in terms of mutually Laplace conjugated positive variables. The quantum statistical distribution functions in the energy are thus put into correspondence with "draw" distributions in the time, represented by sums of Dirac δ distributions. Time averages on correlation functions correspond to ensemble averages in energy. Conversely energy coupling of systems can be represented by a special operation on the δ distributions in time. For this aim, the *connection* of distributions is introduced, which enables one to take into account in some "multiplicative" way their simultaneous and cooperative effects. The Appendix is entirely devoted to the definition and properties of these connections and to some aspects of their algebra that make them suitable for treating some fundamental problems bound to the necessity of accounting interactive effects of singularities.

I. INTRODUCTION

In quantum mechanics, the use of the Fourier transform is very frequent, leading to the consideration of mutually conjugated variables.

For instance, the momentum components of a quantum object (p_x, p_y, p_z) appear to be conjugated to the position components of this object (x, y, z); or in other terms, the reciprocal lattice of wave-vector components k_x, k_y, k_z appears to be the conjugated one of a space lattice with fundamental vector components x, y, z. Expressing correlations between Fourier conjugated variables leads also to the well-known uncertainty relations[1] $(\Delta p_x \Delta x \geqslant \hbar, \ldots)$.

The situation is quite similar (as suggested by the relativistic point of view) to that between the energy and the time variables, but the physical meaning of this conjugation is not exactly the same: energy is a dynamical variable of the considered system and the time is a parameter.

Therefore, if position and momentum play a symmetric role in the description of the system, being measurable at the same given time t, energy and time in contrast do not play the same symmetrical role: the uncertainty on the energy (ΔE) is related to a duration (Δt), which is characteristic of the evolution of the system (for instance, a quantic transition or a lifetime of a resonance).

Moreover, we have to consider that the time variable, due to its possible irreversible character, does not present the same symmetry features as the position ones. Such a remark can also be related to the fact that the energy, for its part, has to be considered as essentially positive (excluding of course the case of the antiparticles for which, if one may speak of negative energies, we also have to speak of past-oriented time, as in Feynman time-oriented graphs).

Given these considerations, it could be interesting to raise the following question: if to a particular distribution of positions (like a lattice) there corresponds a conjugated distribution in momenta—or wave vectors—(like a reciprocal lattice) by the means of a Fourier transform, which involves all the relevant space (from $-\infty$ to $+\infty$), what type of distribution in a temporal space would correspond to a distribution in energy of an ensemble of particles, both spaces allowing only positive values for their variables, using for this aim, for instance, a Laplace transform[2] (which can be considered as a particular, restricted, Fourier transform)? In order to answer such a question two approaches are possible, which lead, of course, to the same results. Either we consider a special distribution in time variables and we "Laplace-transform" it (in the same manner as special positions of a lattice permit by the means of a Fourier transform to construct, for instance, Brillouin's zone in a reciprocal lattice), or we choose a particularly significant energy distribution for our quantum objects and we "inverse-Laplace-transform" it. The physical meaning of this second way is more evident and for this reason we will use it in the following.

II. ENERGY DISTRIBUTIONS AND THEIR INVERSE LAPLACE TRANSFORMS

Indeed we know quite natural relationships between the distribution in energy of quantum objects and their available energy levels, which are represented by their statistical distribution functions.

Considering the particles as indistinguishable, the function $\varphi(E_r)$ representing the occupation number of the rth energy level E_r is

$$\varphi(E_r) = (e^{(E_r-\mu)\beta} + \epsilon)^{-1}. \tag{1}$$

Here $\varphi(E_r)$ obeys the normalization condition $\Sigma_r \varphi(E_r) = N$, where N is the total number of particles; $\beta = (k_B T)^{-1}$ is related to the temperature T (k_B is the Boltzmann constant); μ is either the Fermi level (for fermions) or the chemical potential (for bosons), both related to the condition $\Sigma_r \varphi(E_r) E_r = E$ (the total energy); ϵ is a constant, $+1$ or -1, depending on the type of the statistics: Fermi–Dirac for fermions (submitted to the exclusion principle) and Bose–Einstein for bosons, respectively.

In the case of a distinguishable particle statistics, $\varphi(E_r)$ takes the form $\simeq e^{-\beta E_r}$ (Maxwell–Boltzmann distribution). It is clear that such an expression can be derived from (1) putting $\epsilon = 0$ directly in it or considering a degeneracy when $(E_r - \mu)\beta \gg 1$; but it is worthwhile to remark that the important fact lies in the physical (and conceptual) aspect of the

 0022-2488/85/123146-09$02.50

difference between discernability or undiscernability of the elements themselves much more than in a limiting procedure for the distribution function.

Developing (1) for a Bose–Einstein distribution ($\epsilon = -1$ and $\mu \leqslant 0$) we obtain

$$\varphi_B(E_r) = \sum_{k=0}^{\infty} e^{\beta(k+1)(\mu - E_r)}. \tag{2a}$$

In the case of a Fermi–Dirac distribution ($\epsilon = +1$), the situation is a little more complicated, because we have to distinguish between two possibilities: $E_r > \mu$ and $E_r < \mu$.

In the first one ($\varphi \equiv \varphi_{F_+}$) we have

$$\varphi_{F_+} = \sum_{k=0}^{\infty} (-1)^k e^{\beta(k+1)(\mu - E_r)} \tag{2b}$$

and in the second ($\varphi \equiv \varphi_{F_-}$):

$$\varphi_{F_-} = \sum_{k=0}^{\infty} (-1)^k e^{\beta k(E_r - \mu)}. \tag{2c}$$

For a Maxwell–Boltzmann distribution we simply have $\varphi_M \simeq e^{-\beta E_r}$.

Now we have to choose a relevant so-called Laplace variable in such a way that, as far as possible, these expressions are mutually comparable. Choosing the energy E_r directly would work without difficulty for (2a) and (2b) because of the minus sign in the exponential in front of the argument E_r, but it would cause trouble in (2c), which shows another formal behavior. (In the following, however, we will sometimes make this choice in order to illustrate some examples in the two restricted cases.) To be more general, we will choose $s_r = |\Delta E_r| = |E_r - \mu|$ for this Laplace variable to which correspond, respectively, the functions

$$\Psi_B(s_r) = \sum_{k=0}^{\infty} e^{-\beta(k+1)s_r}, \tag{3a}$$

$$\Psi_{F_+}(s_{r_+}) = \sum_{k=0}^{\infty} (-1)^k e^{-\beta(k+1)s_{r_+}}, \tag{3b}$$

$$\Psi_{F_-}(s_{r_-}) = \sum_{k=0}^{\infty} (-1)^k e^{-\beta k s_{r_-}} \tag{3c}$$

(and of course $\Psi_M \simeq e^{-\beta s_r}$).

At this point it may be asked how to distinguish basically the two eventualities (3b) and (3c) for which the variable s_r does not mean exactly the same physical quantity ($E_r - \mu$ in the first case and $\mu - E_r$ in the second). We propose here a physical answer coming from the fact that for a zero temperature ($T \to 0, \beta \to \infty$) Ψ_{F_+} has to be equal to zero and Ψ_{F_-} to 1 (Fermi degeneracy). This feature has a clear correspondence in the "reciprocal space" we introduce now.

Let us consider indeed the functions Ψ as Laplace transforms (with variable s_r) of some functions f (with variable t). It is easy to obtain by inverse Laplace transformation the following results:

$$f_B(t) = \sum_{k=0}^{\infty} \delta(t - \alpha\beta(k+1)), \tag{4a}$$

$$f_{F_+}(t) = \sum_{k=0}^{\infty} (-1)^k \delta(t - \alpha\beta(k+1)), \tag{4b}$$

$$f_{F_-}(t) = \sum_{k=0}^{\infty} (-1)^k \delta(t - \alpha\beta k), \tag{4c}$$

where δ is the Dirac distribution.

We write the operation of the Laplace transformation as

$$\Psi(s_r) = \int_0^{\infty} e^{-s_r t/\alpha} f(t)\,dt.$$

Therefore α is a constant with the dimension of an action (it may be bound to $\hbar$, the Planck constant). From this point of view of the dimensionalities, it is to be noted that the "functions" f have the dimension of the inverse of a time, i.e., a frequency (through the δ Dirac distribution), and that their transforms Ψ (occupation numbers) have no dimension as requested by our conventions. [We have also $f_M \simeq \delta(t - \alpha\beta)$.] The "functions" f are represented in Fig. 1.

The two following remarks may be made at this stage.

(1) The distinction between the situations generated by Ψ_{F_+} and Ψ_{F_-} is very easy to specify in the t-variable space: it corresponds indeed to the impossibility or possibility of the "occupation" (0 or infinite value for f), respectively, of the $t = 0$ state. This point will be discussed in the next paragraph.

(2) The values of the integrals of f, depending on t, are interesting to calculate. We have, with $G = \int_0^t f(t)\,dt$,

$$G_B(t) = k, \quad \text{for } k \leqslant t/\alpha\beta < k+1, \tag{5a}$$

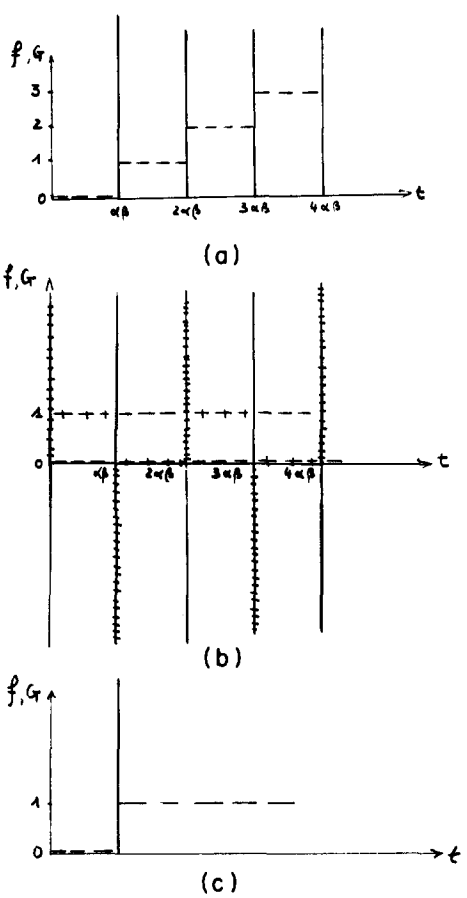


FIG. 1. (a) Inverse Laplace transform for Bose–Einstein statistics (—) and its integral (– – –). (b) Inverse Laplace transform for the Fermi–Dirac statistics in the two cases mentioned in the text: (+ + + +), under the Fermi level; (—), above the Fermi level and their integrals (+ + +) and (– – –), respectively. (c) Inverse Laplace transform for the Maxwell–Boltzmann statistics (—) and its integral (– – –).

$$G_{F_+}(t)=0, \quad \text{for } 2p<t/\alpha\beta<2p+1,$$
$$G_{F_+}(t)=1, \quad \text{for } 2p+1<t/\alpha\beta<2p+2, \tag{5b}$$

and conversely

$$G_{F_-}(t)=1, \quad \text{for } 2p<t/\alpha\beta<2p+1,$$
$$G_{F_-}(t)=0, \quad \text{for } 2p+1<t/\alpha\beta<2p+2, \tag{5c}$$
$$G_M(t)=Y(t-\alpha\beta),$$

where Y is the Heaviside distribution.

III. COMMENTARY

From these results it is conceivable to consider the t space as the space of the drawings of the different particle *species* according to their possible assignations at definite states to which correspond the energy levels in the "reciprocal" energy space. Denoting the different combinations of δ distributions as "draw functions" (Fig. 1) their integrals G can be interpreted as the possibility for a given species to be characterized by the corresponding occupation number for the available states. Thus the Fermi–Dirac statistics permit only the repeating of the numbers 0 and 1, according to the exclusion principle, while the Bose–Einstein statistics authorize all the possible numbers, draw after draw, for one state. The Maxwell–Boltzmann statistics, for their part, permit numbers 0 and 1, but once only, due to the fact that its elements are all distinguishable; so when a draw has taken place for one of them, then no more is possible for the considered species.

Coming back now to the question of the distinction between two situations for the Fermi–Dirac statistics (f_{F_+} and f_{F_-}), it is easy to interpret it in the drawings space we have introduced. Indeed it may be considered that the difference lies essentially in the initial state at $t=0$: has a draw taken place or not initially? If yes, then the state is full whatever the temperature [$T=0, \beta\to\infty$, cf. (4c)]; if not, then the state is empty whatever the time (as measured by k), because $\alpha\beta$ is infinite at $T=0$ [cf. (4b)]. In the energy space it corresponds to the fact that at the absolute zero temperature all the energy states under the Fermi level have a probability equal to 1 to be occupied, and the energy states above the Fermi level have the same probability to be empty.

IV. ENERGY AVERAGING AND TIME CORRELATIONS

From a statistical distribution point of view, what we are interested in is also mainly the possibility of taking averages for physical quantities. In this way some important remarks can follow from the results we have presented here.

In the energy space the individual operations for taking averages are (1) to multiply the statistical function by some other one corresponding to the quantity to be averaged, and (2) to integrate the product over the relevant energy values.

What types of operations correspond to these ones in the reciprocal t space of draws?

To the first corresponds no more a multiplication but a convolution, which can be interpreted as the establishing of a time (or draw) correlation. If $g(t)$ is the inverse Laplace transform of a function $\gamma(s)$, then the correspondence

$$\gamma(s)\varphi(s)\to\int_0^t g(\tau)f(t-\tau)d\tau=g_c(t) \tag{6}$$

holds.

The second operation (integration in the s space) is related to another one in the t space depending on the limits of the integral, for instance,

$$\int_s^\infty \gamma(s)\varphi(s)ds\to\frac{1}{t}g_c(t). \tag{7}$$

It is interesting to note that to an ensemble average (on the energy) in (7) (left-hand side) corresponds something like a time average (or a drawing average) in the right-hand side. This fact could be possibly related to the ergodic properties for such systems.

Taking into account the special form of f (as sums of a weighted δ-distribution), g_c can be expressed as

$$g_c(t)=\sum_{k=0 \text{ or } 1}^{K} a_k\, g(t-\alpha\beta k), \quad K<t/\alpha\beta<K+1.$$

The coefficients a_k depend on the nature of the statistics; so does also the lower bound for k.

Being careful in the regularization of the involved integrals, it is then possible to show that to a density of states $D(E_r)$ proportional to $\sqrt{E_r}$—in order to define the number of particles as $N=\int_0^\infty D(E_r)\varphi(E_r)dE_r$ and in this way the Fermi level or the chemical potential—corresponds a distribution in the t space of the form $\Sigma_k b_k(t-\alpha\beta k)^{-3/2}$. It is then possible to represent in this manner different particular phenomena such as the Bose–Einstein condensation in this specific case.

V. ENERGY COUPLING AND DISTRIBUTION CONNECTIONS IN TIME

Conversely it may be asked if some correlation in the energy space may arise from a specific process in the draw space. Of course usually there is no particular problem to create such correlations in the Laplace conjugated variable. It is sufficient indeed to consider the direct product of two functions, say $f_1(t)\times f_2(t)$ in the time coordinate and its Laplace transform can be written in the form of a convolution product like $(1/2i\pi)\int_{c-i\infty}^{c+i\infty}\varphi_1(z)\varphi_2(s-z)dz$, $\varphi_1(s)$ and $\varphi_2(s)$ being the Laplace transforms of $f_1(t)$ and $f_2(t)$, respectively, all the poles of the functions being at the "left" of c.

But in the specific case of our distributions we have two reasons to consider such a procedure as irrelevant.

(1) Direct product of functions in the time coordinate are not possible because we have no more functions but Dirac distributions, the products of which are not defined.

(2) It would be more interesting and useful to be able to assign some operation in the time coordinate system leading to a coupling like $\int_0^E\varphi(E')\varphi(E-E')dE'$ between the energy probability functions, for instance, or for any other functions developable as a series of exponentials of the energy, i.e., involving a Dirac distribution in time. It is possible to specify such a combination defining an operation between Dirac distributions, which we call *connection* (see the Appendix where the definitions and properties are given and results are obtained that we will use here).

For our purpose it is sufficient to know that given a connection (indicated by a dot ●) between $\delta(x-x_0)$ and $\delta(x-x_1)$, then, if $x_0, x_1 \in [a,b]$,

$$\int_a^b \delta(x-x_0) \bullet \delta(x-x_1) f(x)dx = \frac{f(x_0)-f(x_1)}{x_0-x_1} \tag{8a}$$

[cf. Eqs. (A7a) and (A7b) of the Appendix] and that at the limit $x_1 \to x_0$, for instance, we have indeed, f being a "good" function,

$$\int_a^b \delta(x-x_0) \bullet \delta(x-x_0) f(x)dx = \int_a^b \delta^{[2]}(x-x_0) f(x)dx = -\int_a^b \delta'(x-x_0) f(x)dx = f'(x_0)$$

[cf. Eq. (A5) of the Appendix], which means that the expression (8a) is true for all $x_0, x_1 \in [a,b]$ (also when $x_1 = x_0$).

Taking the Laplace transform of connections is then equivalent to calculating

$$\int_0^\infty \delta(x-x_0) \bullet \delta(x-x_1) e^{-px} dx = \frac{e^{-px_0} - e^{-px_1}}{x_0 - x_1}, \tag{8b}$$

for all $x_1, x_0 \geqslant 0$.

Let us apply this procedure in our case of the draw distribution. To be specific (and also because it is less complicated) we choose the Bose–Einstein function as the basic statistical distribution with E_r as the Laplace variable [cf. Eq. (2a)].

In the time coordinate we have then the corresponding distribution:

$$f_{1B}(t) = \sum_{k=0}^{\infty} e^{\beta(k+1)\mu} \delta(t - \alpha\beta(k+1)). \tag{9}$$

Suppose that some function $l(t)$ is developed under the form

$$l(t) = \sum_{m=0}^{\infty} a_m \delta(t - \alpha\beta m),$$

to which corresponds by Laplace transform, in the energy space, the exponential development

$$\lambda(E) = \sum_{m=0}^{\infty} a_m e^{-m\beta E}.$$

Then coupling $l(t)$ to $f_{1B}(t)$ by means of the connections leads to the expression

$$C(t) = l(t) \bullet f_{1B}(t) = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} a_m e^{\beta\mu(k+1)} \delta(t - \alpha\beta(k+1)) \bullet \delta(t - \alpha\beta m). \tag{10a}$$

Laplace-transforming (10a) and taking into account Eq. (8b), we obtain $\Gamma(E)$ as[3]

$$\Gamma(E) = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} a_m e^{\beta\mu(k+1)} \frac{e^{-\beta E(k+1)} - e^{-\beta E m}}{\alpha\beta(k+1-m)} = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} a_m e^{\beta\mu(k+1)} \frac{e^{-\beta E(k+1)}}{\alpha} \times \int_0^E e^{-\beta E'(m-k-1)} dE'. \tag{10b}$$

And finally performing the summations over k and m, we get

$$\Gamma(E) = \frac{1}{\alpha} \int_0^E \varphi_B(E')\lambda(E-E')dE'. \tag{11a}$$

Equation (11a) is the desired expression for a correlation in energy. Of course, if $\lambda(E)$ itself corresponds to a Bose–Einstein distribution $\varphi_B(E)$, i.e., if $a_m = e^{\beta\mu m}$ for $m \neq 0$ and $a_0 = 0$, then (11a) takes the form of a self-correlated statistical distribution:

$$\frac{1}{\alpha} \int_0^E \varphi_B(E')\varphi_B(E-E')dE'. \tag{11b}$$

VI. CONCLUSION

These examples are easily extended to other cases. The important feature lies in the fact that we have been able to construct correlations in both mutually conjugated spaces: energy and time.

We have to remark that relations like (11a) and (11b) are typical of results for coupled systems, each of them having independently its definition and individuality before the coupling process, if we neglect the proper interaction parameters.[4]

Thus it appears that the use of the connections in the draw space (i.e., in the time coordinate) between the Dirac distribution plays the role of establishing such a coupling between the systems represented each one by their own distributions developments. This method enables us to represent the coupling directly in the space where the draw takes place (or where the relevant time variable is flowing).

In particular it is to be noted that (11a) and (11b) play an important role in passing, by coupling between systems, from a microcanonical ensemble to a canonical ensemble in statistical mechanics (and thus also in defining a system in interaction with a thermostat[4]).

Saying it in another way, to the question, "What would correspond, in the time coordinate space, to the coupling in energy of physical systems?," we may answer in terms of distribution connections.

Only a few special energy distribution functions in energy have been considered in this paper and their inverse Laplace transforms physically interpreted. However, we want to point out the fact that in principle every distribution function in energy, developable as a series of decreasing exponentials of the energy (or of some simple function of the energy as Laplace variable), may be transformed and interpreted in the same way by the means of δ distributions and their connections in the time, as it appears considering the general method we have used here.

Indeed, writing, for instance, (whatever p_n)

$$\varphi(E) = \sum_n p_n e^{-\beta n E},$$

leads to an inverse transform in time of the shape

$$f(t) = \sum_n p_n \delta(t - \alpha\beta n),$$

and then all the considerations we have developed above can be applied.

More generally these types of results hold for any correspondence between energy and time through the Laplace transform, but the interesting fact of having to introduce connections for δ distributions may disappear.

Before terminating this conclusion we want to discuss a last important point: What happens with the uncertainty relation between time and energy in this representation?

Of course this relation is a physical one directly associated with the existence of the action constant and is not bound to a special representation of the correspondence between energy and time, but it may be interesting to discuss briefly how it appears in our particular representation. To do it we will follow two different, but corresponding, types of considerations.

Usually in the frame of the Fourier transform representation a qualitative argument is to be found in accounting the rate of periodic oscillations contributing to plane monochromatic waves in a wave packet and it is shown that there is a relationship (whose notations are evident) such as $\Delta E \simeq (\partial E / \partial p)\Delta p$.

As $\partial E / \partial p = v$ and $\Delta t \simeq \Delta x / v$, one obtains then $\Delta E \Delta t \simeq \Delta x \Delta p_x \gtrsim h$ (the last inequality having been demonstrated earlier).

In our case such a qualitative argument can be used too, but we have now to consider the relevant contributions in evolutive, nonstationary processes. Such processes are well represented, for instance, by an exponential decreasing evolution in the time involving a relaxation time (or a lifetime) τ ($\simeq e^{-t/\tau}$).

Similar considerations as above for stationary oscillating situations show that these relevant contributions to such processes are indeed fulfilling the same types of conditions and so $\Delta E \Delta t$ must be greater than h. (A more precise and deeper discussion applicable in our case may be found also in Ref. 5.)

However, a more rigorous result can be established, starting from another point of view.

Considering, indeed, the well-known quantum mechanical commutators or their classical counterparts, the Poisson brackets, these operators obey, for time t and energy E, the obvious relations

$$\left[\frac{\partial}{\partial t}, t\right] = \left[\frac{\partial}{\partial E}, E\right] = 1.$$

Now we recall that in our Laplace transformations the operator $\partial / \partial t$ is transformed into a multiplicator (E) (dropping out an inessential additive constant for the transformed function at $t = 0$) and that, conversely, the operator $\partial / \partial E$ is inverse-transformed into the multiplicator $(-t)$ (cf. also the rules of the symbolic operational calculus and similar correspondences for Fourier transformations).

We obtain thus, as in the usual representations where these commutators take place (neglecting here the dimensional factor h),

$$[E,t] = -[t,E] = 1.$$

Now, as discussed by most of the authors on quantum mechanics (see also Ref. 7), the "fourth" uncertainty relation, between time and energy does not possess the same status as the three others (between position and impulsion), for if the energy represents, indeed, a physical observable of the system to which corresponds a quantum Hermitian operator, the time for its part is by no way such an observable but much more an external parameter with no corresponding operator.

Thus in order to derive mathematically such a relation, it is necessary to come back to the earliest bases of the operator algebra of quantum mechanics, i.e., the classical Poisson's brackets, which describe precisely the evolution in time of some physical quantities and to take into account directly, as done for instance by Morse and Feshbach,[8] the effect of such commutation relations on the mean square deviation of the considered quantities.

Let A and B be such quantities; then, to the value k of the Poisson's bracket between them ($[A,B] = k$) corresponds the value $k/2$ of the product of their mean square deviation ($\Delta A \Delta B \geqslant k/2$) with $\Delta A = [\langle A^2 \rangle - \langle A \rangle^2]^{1/2}$ and similarly for ΔB (the $\langle \cdots \rangle$'s indicate an average).

Thus we can consider that our representation is conservative with respect to the uncertainty relation for energy and time, due to the structure of the Laplace transformation.

ACKNOWLEDGMENTS

I want to acknowledge my colleagues G. Cohen-Solal and R. Mosseri for many interesting and encouraging discussions, and F. Raymond for having pointed out some important features of the results I present here.

I have to mention especially the aid of P. Degond (Centre de Mathématiques Appliquées de l'Ecole Polytechnique) and of F. Golse (Centre de Mathématiques Appliquées de l'Ecole Normale Supérieure) in controlling the mathematical part of this work and in pursuing enlightening discussions on this subject.

APPENDIX: CONNECTIONS OF DISTRIBUTIONS

In the text we have encountered the need for, but also the difficulty in having to define something like a "multiplication" of distributions (in the drawing space), which could lead to energy correlations for the probability functions in the energy space.

In order to construct such correlation functions in energy we have been lead to introduce the operation of *connection* between distributions.

In all the following this connection will be noted by a dot $\bullet$, z represents a complex variable, x a real variable (with values $x_0, x_1, \ldots$), and f is a function of the complex variable as good as desired.

1. Connections at one point

Given a point on the real axis x_0, characterized as a pole *singularity*, we define the connections at this point.

Let Γ be a closed path in the complex plane. If x_0 lies in its interior it will be noted as Γ_0 and as Γ_e otherwise. Let us

consider x_0 as a pole of order $k+1$ in the plane. Denoting by $f^{(k)}(x_0)$ the k th derivative of f in x_0 (the superscript (k) will always denote a k th derivative), we have

$$f^{(k)}(x_0)=\frac{(-1)^k}{2i\pi}\oint_{\Gamma_0} f(z)\left(\frac{1}{z-x_0}\right)^{(k)}dz \tag{A1}$$

(for $\Gamma=\Gamma_e$, the integral is, of course, zero).

Similarly if $[a,b]$ is a closed interval on the real axis containing x_0, we have

$$f^{(k)}(x_0)=(-1)^k\int_a^b \delta^{(k)}(x-x_0)f(x)dx \tag{A2}$$

(if the interval does not contain x_0, the integral is again zero).

Comparing (A1) and (A2) suggests associating a Dirac distribution at x_0 on the real axis to a pole at x_0 in the complex plane and at the same time the interval $[a,b]$ to the path Γ.

Consider now the expression $(1/(z-x_0))^{k+p+2}$. We can interpret it either as a product, $(1/(z-x_0))^{k+1}\times(1/(z-x_0))^{p+1}$, or as a $(k+p+1)$th derivative of $1/(z-x_0)$,

$$\frac{(-1)^{k+p+1}}{(k+p+1)!}\left(\frac{1}{z-x_0}\right)^{(k+p+1)}.$$

Under the first aspect we have then

$$\left(\frac{1}{z-x_0}\right)^{k+p+2}=\frac{(-1)^k}{k!}\left(\frac{1}{z-x_0}\right)^{(k)}\times\frac{(-1)^p}{p!}\left(\frac{1}{z-x_0}\right)^{(p)} \tag{A3a}$$

and under the second one

$$\left(\frac{1}{z-x_0}\right)^{k+p+2}=\frac{(-1)^{k+p+1}}{(k+p+1)!}\left(\frac{1}{z-x_0}\right)^{(k+p+1)}. \tag{A3b}$$

Associate now, as suggested by the similarity of (A1) and (A2), the distribution $\delta^{(k)}(x-x_0)$ to $(1/(z-x_0))^{(k)}$ as operators, and similarly for the others; we define then the *operation of connection* ● such as (A3a) ≡ (A3b), and consequently

$$\frac{(-1)^k}{k!}\delta^{(k)}(x-x_0)\bullet\frac{(-1)^p}{p!}\delta^{(p)}(x-x_0)\sim\frac{(-1)^{k+p+1}}{(k+p+1)!}\delta^{(k+p+1)}(x-x_0). \tag{A4}$$

The sign $\sim$ denotes an operational equivalence.

In particular we have, with $k=p=0$,

$$\delta(x-x_0)\bullet\delta(x-x_0)\equiv\delta^{[2]}(x-x_0)\sim-\delta'(x-x_0).$$

And, more generally reiterating the operations, we obtain from (A4),

$$\delta^{[n]}(x-x_0)\sim\frac{(-1)^{(n-1)}}{(n-1)!}\delta^{(n-1)}(x-x_0), \tag{A5}$$

the superscript $[n]$ $(n\geqslant 1)$ represents the n times reiteration of the connection at x_0 and of course $\delta^{[1]}(x-x_0)\equiv\delta(x-x_0)$. [Considering such an operatorial equivalence makes the connections $\delta^{[n]}(x-x_0)$ comparable to the functional $(x-x_0)_+^{-n}$ usually introduced in the distribution algebras, whose residue is exactly the same.[6]]

The relations (A4) and (A5) contain all the needed properties of the connection at one point for the purposes of the text. But some of them are interesting in themselves and deserve to be made explicit or generalized. Such a generalization will be postponed until subsection 3 of this Appendix.

Before this we have to discuss the definition and the effect of the operation of connections between several distinct points.

2. Connections between several points

These points are always considered as *pole singularities* on the real axis. Two of them, x_0 and x_1, are sufficient to express the principle of the procedure used, which is similar in many aspects to that we have used in Sec. A 1.

Principally, it is sufficient to split the expression $1/(z-x_0)(z-x_1)$ in the two additive terms $[1/(x_0-x_1)](1/(z-x_0)-1/(z-x_1))$ and then to establish the former correspondences.

Let

$$\varphi(\Gamma)=\frac{1}{2i\pi}\oint_\Gamma\frac{f(z)}{(z-x_0)(z-x_1)}dz. \tag{A6}$$

If $x_0,x_1\notin\Gamma$,

$$\Gamma\equiv\Gamma_e \quad\text{and}\quad \varphi(\Gamma_e)=0.$$

If x_0 or $x_1\in\Gamma$,

$$\Gamma\equiv\Gamma_0 \quad\text{or}\quad \Gamma\equiv\Gamma_1 \quad\text{and}\quad \begin{cases}\varphi(\Gamma_0)=\dfrac{f(x_0)}{x_0-x_1},\\ \varphi(\Gamma_1)=\dfrac{f(x_1)}{x_1-x_0}.\end{cases}$$

If x_0 and $x_1\in\Gamma$,

$$\Gamma\equiv\Gamma_{0,1} \quad\text{and}\quad \varphi(\Gamma_{0,1})=[f(x_0)-f(x_1)]/(x_0-x_1).$$

Let us consider now the interval $[a,b]$ and let us define the connection $\delta(x-x_0)\bullet\delta(x-x_1)$ such as

$$\Psi(x_0,x_1)=\int_a^b\delta(x-x_0)\bullet\delta(x-x_1)f(x)dx \tag{A7a}$$

and

$$\text{if } x_0,x_1\notin[a,b],\ \text{then}\ \Psi\equiv\Psi_e=\varphi(\Gamma_e)=0,$$

$$\text{if } x_0 \text{ or } x_1\in[a,b],\ \text{then}\ \begin{cases}\Psi\equiv\Psi_0=\varphi(\Gamma_0)=f(x_0)/(x_0-x_1),\\ \text{or}\\ \Psi\equiv\Psi_1=\varphi(\Gamma_1)=f(x_1)/(x_1-x_0),\end{cases} \tag{A7b}$$

$$\text{if } x_0 \text{ and } x_1\in[a,b],\ \text{then}\ \Psi\equiv\Psi_{0,1}=\varphi(\Gamma_{0,1})=\frac{f(x_0)-f(x_1)}{x_0-x_1}.$$

Taking the limit $x_1\to x_0$ in $\Psi_{0,1}$ shows that the definition (A7a) and (A7b) is compatible with the definition of the connections at the point x_0 of the Appendix subsection 1

$$\delta^{[2]}(x-x_0)\sim-\delta'(x-x_0).$$

It is to be noted that (choosing for f a constant) the following result holds:

$$\int_{-\infty}^{+\infty} \delta(x-x_0) \bullet \delta(x-x_1)dx = 0.$$

But with x_0 (or x_1) $\in [a,b]$ and x_1 (or x_0) $\notin [a,b]$,

$$\int_a^b \delta(x-x_0) \bullet \delta(x-x_1)dx = \frac{1}{x_0-x_1}$$

$$\left(\text{or } \frac{1}{x_1-x_0}, \text{respectively}\right).$$

It is easy to generalize all these definitions and results for any number of connections between any number of pole-singular points in such a way that we assume the compatibility of these definitions with the definitions of subsection 1 of this Appendix for connections at one point, when a limiting procedure is applied for collapsing two or more points together.

Thus $\Pi_{i=1}^n \bullet$, representing the composition of the connections between n different points $x_i \in [a,b]$, gives

$$\int_a^b \prod_{i=1}^n \bullet\, \delta(x-x_i) f(x)dx = \sum_{j=1}^n \frac{f(x_j)}{\Pi_{i\neq j}(x_j-x_i)}.$$

Of course if one point, say x_k, does not appear in the considered interval $[a,b]$, the corresponding term in the sum $(j=k)$ vanishes in this relation [as if $f(x_k)$ were zero], but remains still among the values of i in the products of the denominator.

We want also to point out the important fact that if the δ distributions are "weighted" by some functions of x, then the connections compose these functions in a multiplicative form.

For instance, if we have to deal with two singularities at different points, each one weighted by a good function in the interval of interest, say $f(x)$ at x_0 and $g(x)$ at x_1, then $(x_0, x_1 \in [a,b])$

$$\int_a^b \delta(x-x_0) f(x) \bullet \delta(x-x_1)\, g(x)dx$$

$$= [f(x_0)g(x_0) - f(x_1)g(x_1)]/(x_0-x_1),$$

which makes evident the multiplication of the functions.

This result is closely related to the correspondence in the complex plane for which

$$\frac{1}{2i\pi}\oint_{\Gamma_{01}} \frac{f(z)g(z)}{(z-x_0)(z-x_1)}dz$$

$$= [f(x_0)g(x_0) - f(x_1)g(x_1)]/(x_0-x_1)$$

may be considered as the relevant Cauchy integral of the product of the one pole function in two different points, $f(z)/(z-x_0)$ and $g(z)/(z-x_1)$. [We remark that such an approach makes it necessary to distinguish essentially the two functionals $f(x)\delta(x-x_0)$ and $f(x_0)\delta(x-x_0)$, when referred at the operation of connection. For in the second case we would have obtained a different result, namely 0.] Thus in the scope of the connection algebra, weighting a distribution becomes equivalent to connecting it with a regular function. We have to indicate here, however, that connecting two regular functions together leads to a vanishing result as shown in the following, due to the fact that no more singularities appear and thus there are no more intervals in which a singularity exists (cf. subsection 3).

Such an approach can be useful for all the problems where the properties of several points, considered as singular ones, have to be correlated simultaneously, i.e., from a "multiplicative" point of view of their mutual effects (and of their concurrent effects on a common fact). Physical examples are, for instance, perturbations, transitions between states, and, of course, correlations such as these we consider in the text, particularly for couplings of systems.

3. Some generalizations for connections and their algebra

Connections at one point behave like usual "multiplication" in their "power" form used in (A5). In fact from Eqs. (A4) and (A5) it follows that

$$\delta^{[k+1]}(x-x_0) \bullet \delta^{[p+1]}(x-x_0) \sim \delta^{[k+p+2]}(x-x_0). \tag{A8}$$

So do the connections with respect to the derivation as it is seen, writing successively

$$(\delta^{[n]}(x-x_0))'$$

$$= \frac{(-1)^{n-1}}{(n-1)!}\delta^{(n)}(x-x_0) = \frac{-n!}{(n-1)!}\delta^{[n+1]}(x-x_0)$$

$$= -n\delta^{[n-1]}(x-x_0) \bullet \delta^{[2]}(x-x_0)$$

$$= n\delta'(x-x_0) \bullet \delta^{[n-1]}(x-x_0). \tag{A9}$$

Other interesting properties follow from (A5). Multiplying by $(-1)^{n-1}$ and summing over n, we obtain, for the right-hand side $(x_0 \in [a,b])$,

$$\int_a^b \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(n-1)!}\delta^{(n-1)}(x-x_0)f(x)dx$$

$$= \sum_{n=1}^{\infty} \frac{f^{(n-1)}(x_0)}{(n-1)!} = e^{d/dx_0}f(x_0). \tag{A10}$$

The last equality is formal, introducing the operator e^{d/dx_0} for the development $\Sigma_{m=0}^{\infty}(1/m!)(d^m/dx_0^m)$. This operator is also a *translator* T_{+1} for f at x_0. We have indeed

$$T_{+1}f(x_0) = e^{d/dx_0}f(x_0) = f(x_0+1), \tag{A11}$$

as can be shown directly by the Taylor development of $f(x_0+1)$ around x_0.

[Dropping the factor $(-1)^{n-1}$ in (A10) would have lead us to the definition of e^{-d/dx_0} and to the translator T_{-1} such as $T_{-1}f(x_0) = f(x_0-1)$.]

Now taking into account the left-hand side of (A5), we may write

$$\int_a^b \sum_{n=1}^{\infty} \delta^{[n]}(x-x_0)f(x)dx$$

$$\equiv \int_a^b \delta(x-x_0) \bullet \sum_{n=1}^{\infty} \delta^{[n-1]}(x-x_0)f(x)dx$$

$$= f(x_0+1). \tag{A12}$$

Following the same analogy as before with the usual powers, multiplications, and developments, for connections, we define the new operator as follows:

$$\sum_{n=1}^{\infty} \delta^{[n-1]}(x-x_0) \sim [1-\delta(x-x_0)]^{[-1]}. \tag{A13}$$

So (A12) takes the form

$$\int_a^b \delta(x-x_0) \bullet [1-\delta(x-x_0)]^{[-1]} f(x)dx = f(x_0+1). \tag{A14}$$

(Dropping—introducing—again the factor $(-1)^{n-1}$ in (A12) leads us then to define in place of (A13) the operator $[1+\delta(x-x_0)]^{[-1]}$ and to the corresponding $f(x_0-1)$ in place of (A14).)

Such results are not very surprising if we recall the initial correspondence between the Dirac distribution and the poles in the complex plane that permitted us to construct the connections.

In fact, the integral

$$\rho(\Gamma) = \frac{1}{2i\pi} \oint_\Gamma \frac{f(z)dz}{(z-x_0)(1-1/(z-x_0))} \tag{A15a}$$

can be rewritten as

$$\rho(\Gamma) = \frac{1}{2i\pi} \oint \frac{f(z)dz}{z-(x_0+1)}, \tag{A15b}$$

and, if the point x_0+1 lies in the interior of Γ ($\Gamma \equiv \Gamma_0^+$),

$$\rho(\Gamma_0^+) = f(x_0+1) \tag{A15c}$$

(and $=0$ otherwise).

But formally if we associate the distribution $\delta(x-x_0)$ on the real axis at each occurrence of the pole $(z-x_0)^{-1}$, the expression (A15) may be expressed in terms of connections exactly as in (A14):

$$\int_a^b \delta(x-x_0) \bullet (1-\delta(x-x_0))^{[-1]} f(x)dx = f(x_0+1)$$

[and conversely for $(1+1/(z-x_0))^{-1}$ in (A15a) leading to $f(x_0-1)$].

What should be noted here is the fact that, with respect to the connections operation, it is possible to write

$$\delta(x-x_0) \bullet [1-\delta(x-x_0)]^{[-1]} \sim \delta(x-x_0-1) \tag{A16}$$

(and also $\delta(x-x_0) \bullet [1+\delta(x-x_0)]^{[-1]} \sim \delta(x-x_0+1)$).

But the first condition on the interval ($x_0 \in [a,b]$), which corresponded to the fact that the pole x_0 had to lie in the interior of Γ, is to be changed to a second one, x_0+1 (or x_0-1, depending of our choice) $\in [a,b]$ corresponding to the new position of the pole relatively to the path Γ_0^+ (or Γ_0^-, respectively). Saying it in another form, we would get the same result by moving the boundaries of the first definition interval $[a,b]$ to a second one $[a+1,b+1]$ (or $[a-1,b-1]$, respectively). Thus the term translator appears as well adapted for describing all the conditions of the system, including the definition domains.

From the point of view of possible physical correspondences with usual operators, it may be worthwhile to indicate that such exponential-differential operators as we have encountered here are frequently found in the time-dependent evolution of quantum situations, for instance, as a complex evolution operator like $e^{i\hbar(\partial/\partial t)}$.

It is also interesting to indicate that expressions like (A16), for instance, may be manipulated as if they were true multiplications. Actually, following the same procedure as in (A15a) and (A15b), which lead us to (A16), it is very easy to show that we can treat this type of connection algebra in a so-called "multiplicative" way such as, for instance,

$$\delta(x-x_1) \bullet [1-x_2\delta(x-x_0)]^{[-1]} \bullet [1+x_2\delta(x-x_0-x_2)]^{[-1]} \sim \delta(x-x_1), \tag{A17}$$

which could be symbolically written (keeping in mind that we have always to work with the δ distribution, i.e., pole singularities in the complex plane)

$$[1-x_2\delta(x-x_0)]^{[-1]} \bullet [1+x_2\delta(x-x_0-x_2)]^{[-1]} \sim I, \tag{A18a}$$

I being the identity for this algebra ($=1$).

Indeed, to maintain the correspondence between the effects of the poles in the complex plane and that of the Dirac singularities on the real axis, we have to recall that the interval $[a,b]$ on the axis has to correspond to the contour around the considered pole. If the pole disappears the integral vanishes, the result of the integration being reduced to zero. In the same way, the interval $[a,b]$ has to collapse and to be reduced to zero ($[a',a']$ for instance), because no more simple singularity appears in its interior; so the integral taken between a' and a' vanishes too. Maybe, it would even be more appropriate for a single pole at x_0 to write the integral under the form line $\lim_{\epsilon \to 0} \int_{x_0-\epsilon}^{x_0+\epsilon} \cdots$ instead of $\int_a^b \cdots$; but it is sufficient to keep in mind that it is necessary, for performing a nonzero integration in this type of algebra, to preserve the existence of a δ singularity in the integration range (conserving the equivalence with the Cauchy integral).

In these examples, the importance of specifying the *two* prescriptions—positions of poles, definition of the related considered intervals—appears very clearly, if we want to work with the connections.

Under these conditions it is then possible to write formally the operatorial equality [derived in a "multiplicating" way from (A18a)]

$$[1-x_2\delta(x-x_0)]^{[-1]} \sim 1+x_2\delta(x-x_0-x_2) \tag{A18b}$$

or any other equality of this type and submitted to similar restrictions.

Under the same conditions it is possible to introduce (in the spirit of the connections we used up to now) another type of distribution: the form that derives Dirac's δ. It is well known that the Heaviside distribution $Y(x-x_0)$ ($=1$ for $x>x_0$, $=0$ for $x<x_0$) has a $\delta(x-x_0)$ for its first derivative, but it corresponds to no more poles at x_0, only to a discontinuity. This may be made evident by the following considerations.

Write $\delta(x-x_0)$ as a formal derivative $X'(x-x_0)$ of a given unknown but without pole $X(x-x_0)$. Then,

$$\lim_{\epsilon \to 0} \int_{x_0-\epsilon}^{x_0+\epsilon} \delta(x-x_0) f(x)dx = f(x_0) = \lim_{\epsilon \to 0} \int_{x_0-\epsilon}^{x_0+\epsilon} X'(x-x_0) f(x)dx. \tag{A19a}$$

Integrating by parts,

$$f(x_0) = \lim_{\epsilon \to 0} \left\{ [X(x-x_0)f(x)]_{x_0-\epsilon}^{x_0+\epsilon} - \int_{x_0-\epsilon}^{x_0+\epsilon} X(x-x_0) f(x)dx \right\}. \tag{A19b}$$

As there is no pole in $X(x-x_0)$ between $x_0-\epsilon$ and $x_0+\epsilon$ [and $f(x)$ is a function as good as desired], then the second term of (A19b), i.e., the integral, becomes zero when $\epsilon \to 0$.

The integrated part remains and we have

$$f(x_0)=\lim_{\epsilon\to 0}\left[X(\epsilon)f(x_0+\epsilon)-X(-\epsilon)f(x_0-\epsilon)\right]$$

$$=f(x_0)\lim_{\epsilon\to 0}\left[X(\epsilon)-X(-\epsilon)\right]$$

and

$$\lim_{\epsilon\to 0}\left[X(\epsilon)-X(-\epsilon)\right]=1.$$

Thus X can be made equivalent to the Heaviside Y.

Now going back to the correspondence for the δ's with poles in the complex plane, it can be argued that the corresponding $1/(z-x_0)$ is the derivative of the function $\log(z-x_0)$ (even if this function is a multivalued one, which necessitates a limitation of the range of the variable or a cut in the plane). The singularity shown at x_0 by this function can then be brought into correspondence with the discontinuity induced by $Y(x-x_0)$ on the real axis.

All these considerations can be extended easily to the case of several singularities. The only important remark we want to point out in this latter case is the essential difference between the usual convolution algebra involving the distributions and the connection algebra we present here. Actually in the convolution algebra the composed effect of two different distributions always gives a result equal to zero, whatever the functions on which the operation is applied; this fact is basically different in the connection algebra as shown by the definitions (A7a) and (A7b). Thus the connections really take into account the simultaneous effect of two or more distinct pole singularities.

However, it must be emphasized that some precautions are to be taken in working, for instance, with derivatives or "powers" among the connections.

As an example, we have $(x_1,x_2\in[a,b])$

$$\int_a^b \delta'(x-x_1)\bullet\delta(x-x_2)f(x)dx$$

$$=-\frac{f'(x_1)}{x_1-x_2}-\frac{f(x_2)}{(x_1-x_2)^2}+\frac{f(x_1)}{(x_1-x_2)^2}.$$

The first two terms on the right-hand side seem to be completely natural [taking into account the fact that $\delta'(x-x_1)\sim-\delta^{[2]}(x-x_1)$] but it may be asked about the third one. For explaining its presence we have to remember that the $\delta'(x-x_1)$ [or conversely $\delta^{[2]}(x-x_1)$] operates on *all* functions *and* distributions under the integral (in the same way as for a product) and thus also on the $\delta(x-x_2)$, the effect of which is precisely to introduce this term.

This point appears very clearly in writing down the corresponding underlying Cauchy integral (which permitted us to construct the connections) $(x_1,x_2\in\Gamma)$:

$$\frac{1}{2i\pi}\oint\frac{f(z)}{(z-x_1)^2(z-x_2)}dz$$

$$=-\frac{f'(x_1)}{x_1-x_2}+\frac{f(x_1)}{(x_1-x_2)^2}-\frac{f(x_2)}{(x_1-x_2)^2}.$$

It follows from this remark that we have to be careful in calculations, particularly when there appears, for instance, a "power" of distributions among the connections, this "power" being operatorially equivalent to a derivative [Eq. (A5)].

From a similar point of view, in introducing the Heaviside distribution it may be asked if it is possible to define an operator like $\delta^{[-p]}(x-x_0)$, i.e., something like an inverse in our connection algebra.

Actually it is not difficult to do it if we remain careful in the definitions of the integration ranges (or in the contours of the Cauchy integrals).

Following the correspondence in the complex plane let us associate the function $(z-x_0)^p$ to $\delta^{[-p]}(x-x_0)$ on the real axis and any closed contour around x_0 then

$$\oint_\Gamma f(z)(z-x_0)^p\,dz=0.$$

Consequently we also have (for all $p\geqslant 0$)

$$\lim_{\epsilon\to 0}\int_{x_0-\epsilon}^{x_0+\epsilon}\delta^{[-p]}(x-x_0)f(x)dx=0. \tag{A20}$$

More generally and taking into account the earlier results for connections at one point, we write

$$\lim_{\epsilon\to 0}\int_{x_0-\epsilon}^{x_0+\epsilon}\delta^{[n-p]}(x-x_0)f(x)dx$$

$$=\begin{cases}0, & \text{for } n\leqslant p,\\ f^{(n-p-1)}(x_0)/(n-p-1)!, & \text{for } n>p.\end{cases} \tag{A21a}$$

In a similar way we obtain for two points $(x_1\neq x_0)$,

$$\lim_{\epsilon\to 0}\int_{x_1-\epsilon}^{x_1+\epsilon}\delta^{[-p]}(x-x_0)\bullet\delta^{[n]}(x-x_1)f(x)dx$$

$$=f^{(n-1)}(x_1)/(n-1)!(x_1-x_0)^p,$$

and, more generally $(x_1,x_2\in[a,b]\neq x_0)$,

$$\int_a^b\delta^{[-p]}(x-x_0)\bullet\delta(x-x_1)\bullet\delta(x-x_2)f(x)dx$$

$$=\left[(x_1-x_0)^pf(x_1)-(x_2-x_0)^pf(x_2)\right]/(x_1-x_2). \tag{A21b}$$

Results like (A21a) and (A21b) show that such an inverse for our distribution (in the spirit of the connections) does not lead to a result always equivalent to zero as (A20) suggests.

[1] A. Messiah, *Mécanique quantique* (Dunod, Paris, 1962); L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1965).

[2] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953); R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1962).

[3] The presence of the factor α comes from the fact that the argument of the connected δ distributions is of the dimension of a time, introducing thus a time^{-1} dimension coefficient, and so does the correlation function too, if λ (as φ) has the dimension of a number. For avoiding it we would have to deal with dimensionless variables like βE and $\tau/\alpha\beta$, for instance, but the physical meaning would become less evident.

[4] A. Y. Khinchin, *Mathematical Foundations of Quantum Statistics* (Graylock, Albany, NY, 1960).

[5] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1958).

[6] I. M. Guelfand and G. E. Chilov, *Les distributions* (Dunod, Paris, 1962).

[7] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Mécanique quantique* (Hermann, Paris, 1973).

[8] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

Dirac quantization of a three-dimensional gauge theory

A. Burnel and M. Van Der Rest-Jaspers
Service de Physique Théorique et Mathématique, Institut de Physique, Université de Liège, Sart Tilman, B-4000 Liege 1, Belgium[a]

(Received 25 April 1985; accepted for publication 5 June 1985)

A model recently proposed by Hagen is examined from the point of view of Dirac quantization of constrained systems. This model exhibits interesting particular features for the Dirac method itself. Among them are the odd number of second-class constraints and the fact that, when a gauge is fixed, constraints result from compatibility conditions between Lagrange multipliers. From the point of view of the model itself, the invalidity of the axial gauge in the non-Abelian case is obtained by comparing the effective Hamiltonians for two different values of the arbitrary spacelike vector.

I. INTRODUCTION

Recently, Hagen[1] proposed a three-dimensional gauge model in order to have a better understanding of the concepts involved in gauge theories. He studied this model from the conventional Lagrangian point of view in the Coulomb and axial gauges and concluded that the axial gauge is, in the non-Abelian case, in conflict with relativistic invariance. We want to study this model in the framework of Dirac[2] quantization of constrained systems in order to have a better understanding firstly of the method itself and secondly of the Hagen model.

First of all, the number of second-class constraints is three. This means that one of these constraints must be transformed into a first-class quantity with respect to the other two. *A priori*, the selection of this constraint from the set of the second-class constraints can be arbitrary. However, if we want the theory to look like a usual gauge theory, it appears that the primary second-class constraints must be kept as such in contrast to the secondary second-class constraint, which must be transformed into a first-class quantity. Other choices lead to difficulties in the gauge-fixing procedure.

Secondly, when a class-I gauge[3] like the Coulomb or the axial gauge is fixed through a Lagrange multiplier inside the Lagrangian, the Dirac algorithm leads to an overdetermination of some multipliers accompanying the primary second-class constraints in the Hamiltonian. Compatibility of these equations implies additional constraints which are also obtained if the starting point of the gauge-fixing procedure is the effective Hamiltonian, where primary second-class constraints are strongly realized. In other words, the procedure of constraint determination does not stop with equations fixing the Lagrange multipliers. Compatibility of these equations must also be required.

The third point we want to point out in our analysis concerns the invalidity of the axial gauge in the non-Abelian case. Here, it is easily proven by comparing the effective Hamiltonian for two different values of the arbitrary spacelike vector. In the case of convenient class-I gauges, the effective Hamiltonian describes the physical system and must be independent of the gauge choice. Such is the case for the Coulomb and axial gauges in the Abelian case although the proof is not trivial. Our proof rests heavily on the current conservation $\partial^\mu j_\mu = 0$, which does not hold in the non-Abelian case, where the Coulomb gauge is also not convenient for the same reasons it is not in Yang–Mills theory.

We organize our work as follows. In Sec. II, we develop the Dirac formulation of the Hagen model. We discuss the gauge fixing in Sec. III, where we successively study Coulomb, axial, temporal, and relativistic gauges and compare the effective Hamiltonian in Coulomb and axial gauges. In Sec. IV, we study the non-Abelian case and, in particular, the invalidity of the axial gauge.

II. THE HAGEN MODEL IN THE DIRAC FORMALISM

Let us start with the Lagrangian[1]

$$\mathcal{L} = \tfrac{1}{2}\phi^\mu \epsilon_{\mu\nu\rho}\,\partial^\rho \phi^\nu + \phi^\mu j_\mu, \tag{1}$$

where the Greek indices take the values 0, 1, 2. The metric tensor is $g_{\mu\nu} = (1, -1, -1)$, $\epsilon_{\mu\nu\rho}$ is a completely antisymmetric tensor ($\epsilon_{012} = 1$), and j_ν is a conserved current. Under a gauge transformation

$$\phi_\mu \to \phi_\mu + \partial_\mu \lambda, \tag{2}$$

we have

$$\begin{aligned}\delta\mathcal{L} &= \tfrac{1}{2}\partial^\mu\lambda\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\nu + \partial^\mu\lambda j_\mu \\ &= \partial^\mu(\tfrac{1}{2}\lambda\epsilon_{\mu\nu\rho}\,\partial^\mu\phi^\nu + \lambda j_\mu),\end{aligned} \tag{3}$$

i.e., the Lagrangian is invariant up to a four-divergence, which is sufficient for applying the Noether theorem. There is thus a gauge invariance.

Let us now go to the Hamiltonian. We have

$$\pi^1 = \frac{\partial\mathcal{L}}{\partial\phi_1} = \frac{1}{2}\phi^2, \tag{4}$$

$$\pi^2 = \frac{\partial\mathcal{L}}{\partial\phi_2} = -\frac{1}{2}\phi^1, \tag{5}$$

$$\pi^0 = \frac{\partial\mathcal{L}}{\partial\phi_0} = 0. \tag{6}$$

None of these relations is an evolution equation. They are constraints and the total Hamiltonian density reads

$$\begin{aligned}\mathcal{H}_T &= u^0\pi^0 + u^1(\pi^1 - \tfrac{1}{2}\phi^2) + u^2(\pi^2 + \tfrac{1}{2}\phi^1) \\ &\quad - \tfrac{1}{2}\phi^\mu\epsilon_{\mu\nu i}\,\partial^i\phi^\nu - \phi^\nu j_\nu,\end{aligned} \tag{7}$$

where u^0, u^1, u^2 are arbitrary functions of ϕ_μ, π^μ, and j_μ. Since the constraints (4)–(6) must hold for any time, we have

$$0 = \partial_0\pi^0 = \{\pi^0, H_T\} = \epsilon_{0ji}\,\partial^j\phi^i + j_0 = 0, \tag{8}$$

[a] Postal address: Institut de Physique au Sart Tilman, Bâtiment B.5, B-4000 Liege 1, Belgium.

 0022-2488/85/123155-05$02.50

$$0 = \partial_0(\pi^1 - \tfrac{1}{2}\phi^2) = u^2 + \partial^2\phi^0 + j_1 = 0, \tag{9}$$

$$0 = \partial_0(\pi^2 + \tfrac{1}{2}\phi^1) = -u^1 - \partial^1\phi^0 + j_2 = 0. \tag{10}$$

Equation (8) is a new constraint while (9) and (10) can be used as equations determining the Lagrange multipliers u^1, u^2.

Again the constraint (8) must be true for any time. Its Poisson brackets with the Hamiltonian lead to

$$\partial^1 u^2 - \partial^2 u^1 + \partial^0 j^0 = 0, \tag{11}$$

i.e., a condition on the Lagrange multipliers u^1, u^2 which is a trivial consequence of (9) and (10) and the current conservation.

A problem occurs with the constraints in this example. While (6) is a first-class constraint, the Poisson brackets between $K_1 \equiv \pi^1 - \frac{1}{2}\phi^2$, $K_2 \equiv \pi^2 + \frac{1}{2}\phi^1$, $K_3 \equiv \epsilon_{0ji}\,\partial^i\phi^j + j_0$ do not vanish, i.e., we have a system with three second-class constraints. In order to use the Dirac algorithm, one of these constraints must be transformed into a first-class quantity with respect to the other two, according to the rule

$$K'_a = K_a - \{K_a, K_\beta\} C^{-1}_{\beta\gamma} K_\gamma, \tag{12}$$

where

$$C_{\beta\gamma} = (\{K_\beta, K_\gamma\}), \tag{13}$$

β, γ being any pair among the indices 1,2,3. *A priori* the choice of this first-class constraint is arbitrary.

If we transform K_3 into a first-class constraint and realize strongly K_1 and K_2, we are led to the Hamiltonian

$$\mathscr{H}_3 = u^0\pi^0 - \phi^0(\partial^2\phi^1 - 2\,\partial^1\pi^1 + j^0) - \phi^1 j_1 - 2\pi^1 j_2, \tag{14}$$

while the nonvanishing Dirac brackets involving ϕ^1, π^1, ϕ^0, and π^0 are

$$\{\phi_1(x), \pi^1(y)\}_{D, x_0 = y_0} = \tfrac{1}{2}\delta^{(2)}(\mathbf{x} - \mathbf{y}), \tag{15}$$

$$\{\phi_0(x), \pi^0(y)\}_{D, x_0 = y_0} = \delta^{(2)}(\mathbf{x} - \mathbf{y}). \tag{16}$$

As usual, the first-class constraints $\pi^0 = 0$ and

$$\partial^2\phi^1 - 2\,\partial^1\pi^1 + j^0 = 0 \tag{17}$$

are not compatible with these brackets. There are two undetermined functions u^0 and ϕ^0 in the Hamiltonian, one of them (ϕ_0) being the variable canonically conjugate to the primary first-class constraint π^0 as is usual in traditional gauge theories.

If, instead of K_3, we now transform K_1 into a first-class constraint, the Hamiltonian becomes

$$\mathscr{H}_1 = u^0\pi^0 + u^1(\pi^1 - \tfrac{1}{2}\phi^2) - \phi^1 j_1 - \phi^2 j_2, \tag{18}$$

where ϕ^2 must be expressed in terms of ϕ^1 and j_0 through Eq. (8), i.e.,

$$\phi^2 = (\partial^1)^{-1}(\partial^2\phi^1 + j^0). \tag{19}$$

Here and in the following, the operator $(\partial^1)^{-1}$ will be defined by

$$[(\partial^1)^{-1} f](x) = -\tfrac{1}{2}\int d^2y\, \epsilon(x^1 - y^1)\delta(x^2 - y^2) f(y). \tag{20}$$

It is clear that other definitions could also work provided $\partial^1(\partial^1)^{-1} f = f$. The Hamiltonian $\mathscr{H}_1$ is nonlocal and there are again two arbitrary functions u^0 and u^1. However, u^1 cannot be considered as the variable canonically conjugate to the primary first-class constraint, so that the theory is quite different from a usual gauge theory. The same situation clearly occurs with K_2. As it will be seen in the following section, troubles occur with these choices, so that we will be unambiguously led to the choice of K_3 as the first-class constraint.

III. GAUGE FIXING

In order to quantize the theory defined by the Hamiltonian $\mathscr{H}_3$ or $\mathscr{H}_1$, it is necessary to fix a gauge. We make different choices suggested by the usual Maxwell theory.

A. Coulomb gauge

Originally defined by $\partial^i\phi^i = 0$, it is transformed into

$$\partial^1\phi^1 + 2\,\partial^2\pi^1 = 0, \tag{21}$$

in order to apply to the Hamiltonian $\mathscr{H}_3$. We have

$$\{\partial^1\phi^1(x) + 2\,\partial^2\pi^1(x),\quad \partial^2\phi^1(y) - 2\,\partial^1\pi^1(y) + j^0(y)\}_{x_0 = y_0} = \Delta\delta^{(2)}(\mathbf{x} - \mathbf{y}) \tag{22}$$

and

$$\partial_0(\partial^1\phi^1 + 2\,\partial^2\pi^1) = \{\partial^1\phi^1 + 2\,\partial^2\pi^1,\ H_3\}_{x_0 = y_0} = -\Delta\phi^0 + \partial^1 j^2 - \partial^2 j^1 = 0. \tag{23}$$

Equation (23) fixes the undetermined parameter ϕ^0; u^0 is also fixed by taking

$$\partial_0(\Delta\phi^0 - \partial^1 j^2 + \partial^2 j^1) = 0, \quad \text{i.e.,}\quad \Delta u^0 = \partial_0(\partial^1 j^2 - \partial^2 j^1). \tag{24}$$

The two constraints (6) and (17) can now be realized strongly, leading us to an effective Hamiltonian

$$\mathscr{H}_{3\,\text{eff}} = -j_0\Delta^{-1}(\partial^1 j^2 - \partial^2 j^1), \tag{25}$$

which is a nonlocal interaction between material fields. This is the physical content of the Hagen gauge theory.

If we want to use the Coulomb gauge with the Hamiltonian $\mathscr{H}_1$, it must be transformed into

$$\partial^1\phi^1 + (\partial^1)^{-1}[(\partial^2)^2\phi^1 + \partial^2 j_0] = 0, \tag{26}$$

or, by applying ∂^1, into

$$\Delta\phi^1 + \partial^2 j_0 = 0. \tag{27}$$

It is not necessary to use such a complicated expression and it is preferable to work with $\phi^1 = 0$, i.e., the axial gauge. Then

$$\partial_0\phi^1 = \{\phi^1, H_1\} = -u^1 = 0 \tag{28}$$

fixes the undetermined parameter u^1. However, u^0 is not fixed here and π^0 cannot be realized strongly. The gauge choice can be completed by taking $u^0 = 0$ and the effective Hamiltonian is

$$\mathscr{H}_{1\,\text{eff}} = -j_0(\partial^1)^{-1} j_2, \tag{29}$$

which is apparently different from $\mathscr{H}_{3\,\text{eff}}$. We will, however, show later that $\mathscr{H}_{1\,\text{eff}} = \mathscr{H}_{3\,\text{eff}}$.

For a complete understanding of the Dirac method and for making a clear choice between $\mathscr{H}_3$ and $\mathscr{H}_1$, it is interesting to introduce the gauge condition inside the Lagrangian with the help of a Lagrange multiplier

$$\mathscr{L} = \tfrac{1}{2}\phi^\mu\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\nu + \phi^\nu j_\nu - S\,\partial^i\phi^i. \tag{30}$$

This is a theory with four primary constraints

$$\pi^0 = \pi^1 - \tfrac{1}{2}\phi^2 = \pi^2 + \tfrac{1}{2}\phi^1 = \pi_S = 0. \tag{31}$$

The total Hamiltonian is

$$\mathscr{H}_T = u^0\pi^0 + u^1(\pi^1 - \tfrac{1}{2}\phi^2) + u^2(\pi^2 + \tfrac{1}{2}\phi^1) + u_S\pi_S - \phi_0(\partial^2\phi^1 - \partial^1\phi^2 + j_0) - \phi^i j_i + S\,\partial^i\phi^i, \tag{32}$$

and the Dirac algorithm leads to two secondary constraints

$$\partial^2\phi^1 - \partial^1\phi^2 + j_0 = \partial^i\phi^i = 0 \tag{33}$$

and to four equations on u^1 and u^2

$$u^2 + \partial^2\phi^0 + j^1 + \partial^1 S = 0, \tag{33a}$$

$$-u^1 - \partial^1\phi^0 + j^2 + \partial^2 S = 0, \tag{33b}$$

$$\partial^2 u^1 - \partial^1 u^2 - \partial^k j^k = 0, \tag{33c}$$

$$\partial^i u^i = 0, \tag{33d}$$

while u_0 and u_S are not fixed. Additional constraints, however, result from the compatibility between Eqs. (33). They read

$$\Delta S = \Delta\phi^0 + \partial^2 j^1 - \partial^1 j^2 = 0. \tag{34}$$

They imply

$$\Delta u_S = \Delta u^0 - \partial_0(\partial^1 j^2 - \partial^2 j^1) = 0. \tag{35}$$

We recover in this way the results obtained from the Hamiltonian $\mathscr{H}_3$. This shows that the use of $\mathscr{H}_1$ and $\mathscr{H}_2$ is not natural and should be rejected.

Let us note that the constraints (34) are necessary for the consistency of the method. The Hagen Lagrangian in the Coulomb gauge thus provides us with an example where the constraint search is not stopped with equations fixing the multipliers. Compatibility of these equations must be required.

B. Axial gauge

We follow the same procedure by taking now $n_i\phi_i = 0$ or

$$n^1\phi^1 + 2n^2\pi^1 = 0 \tag{36}$$

as gauge condition. We have

$$\{n^1\phi^1(x) + 2n^2\pi^1(x),\ \partial^2\phi^1(y) - 2\partial^1\pi^1(y) + j^0(y)\}_{D,x_0=y_0} = \mathbf{n}\cdot\boldsymbol{\partial}\delta^{(2)}(\mathbf{x}-\mathbf{y}) \tag{37}$$

and

$$\mathbf{n}\cdot\boldsymbol{\partial}\phi_0 = n^2 j^1 - n^1 j^2 \tag{38}$$

as the resulting class-II gauge condition.

The effective Hamiltonian is therefore

$$\mathscr{H}_{A\,\mathrm{eff}} = j_0(\mathbf{n}\cdot\boldsymbol{\partial})^{-1}(n^2 j^1 - n^1 j^2). \tag{39}$$

C. Equivalence between different axial and Coulomb gauges

Such an equivalence is not evident by comparing the effective Hamiltonians. It must, however, hold since both axial and Coulomb gauges are convenient. The effective Hamiltonian involves only physical objects and we cannot change the theory by changing the gauge. We make the calculations in a formal way and we start with

$$H_{3\,\mathrm{eff}} = \int d^2x\,\mathscr{H}_{3\,\mathrm{eff}} = -\int d^2x\, j_0\Delta^{-1}(\partial^1 j^2 - \partial^2 j^1) = -\langle j_0,\ \Delta^{-1}(\partial^1 j^2 - \partial^2 j^1)\rangle, \tag{40}$$

where $\langle f,g\rangle$ represents the scalar product of the two functions f, g. We use successively

$$\partial^1(\partial^1)^{-1} = \mathbf{1} = \Delta\Delta^{-1} \tag{41}$$

and

$$[\Delta^{-1},(\partial^1)^{-1}] = 0 \tag{42}$$

to get

$$H_{3\,\mathrm{eff}} = -\langle j_0,\ \Delta^{-1}(\partial^1)^{-1}((\partial^1)^2 j^2 - \partial^1\partial^2 j^1)\rangle = -\langle j_0,\ (\partial^1)^{-1} j^2\rangle + \langle j_0,\ \Delta^{-1}(\partial^1)^{-1}\partial^2(\partial^1 j^1 + \partial^2 j^2)\rangle. \tag{43}$$

Using current conservation,

$$H_{3\,\mathrm{eff}} = -\langle j_0,\ (\partial^1)^{-1} j^2\rangle + \langle j_0,\ \Delta^{-1}(\partial^1)^{-1}\partial^2\partial_0 j_0\rangle. \tag{44}$$

If we consider now

$$S = -\int dx_0\, H_{3\,\mathrm{eff}}, \tag{45}$$

due to the antisymmetry property of the operator $\Delta^{-1}(\partial^1)^{-1}\partial^2\partial_0$, we easily get

$$S = \int d^3x\, j_0(x)(\partial^1)^{-1}(x-y) j^2(y), \tag{46}$$

which is the action in the axial gauge $\mathbf{n} = (1,0)$. By the same reasoning it is possible, using current conservation, to go from the axial gauge with $\mathbf{n} = (1,0)$ to the axial gauge with $\mathbf{n} = (0,1)$. As a consequence, the Hamiltonians in any of the axial gauges and in the Coulomb gauge are equivalent. This is a result which is necessary for the consistency of the gauge choice.

D. Temporal gauge

The temporal gauge is characterized by $\phi_0 = 0$. It is a consistent gauge if we start from $\mathscr{H}_3$. It leads to the Hamiltonian

$$\mathscr{H}_{3t} = -\phi^1 j_1 - 2\pi^1 j_2, \tag{47}$$

and, as usual, it is necessary to impose

$$\partial^2\phi^1 - 2\,\partial^1\pi^1 + j_0 = 0 \tag{48}$$

as a constraint on physically acceptable states in order to recover the starting theory. Again, $\phi_0 = 0$ is not sufficient to fix the gauge if we start with $\mathscr{H}_1$ or $\mathscr{H}_2$ and, as for the Coulomb gauge, if we incorporate the temporal gauge inside the Lagrangian through a Lagrange multiplier, we are unavoidably led to $\mathscr{H}_{3t}$. This again confirms the inadequacy of the choices $\mathscr{H}_1$ or $\mathscr{H}_2$.

E. Relativistic gauges

We develop briefly only the Lagrangian formulation and introduce the gauge condition through a Lagrange multiplier. The starting Lagrangian is

$$\mathscr{L} = \tfrac{1}{2}\phi^\mu\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\nu - S\,\partial^\mu\phi_\mu + \tfrac{1}{2}aS^2 + \phi^\mu j_\mu, \tag{49}$$

where a is the usual gauge parameter but is here dimensioned. Field equations are

$$\epsilon^{\mu\nu\rho}\,\partial_\rho\phi_\mu - \partial^\nu S - j^\nu = 0 \tag{50}$$

and

$$\partial^\mu\phi_\mu = aS. \tag{51}$$

Equation (50) implies

$$\Box S = 0, \tag{52}$$

i.e., S is a free field.

As in Maxwell theory, the unwanted states are eliminated by a condition on physical states which reads

$$S^{(+)}|\psi_{\text{phys}}\rangle = 0. \tag{53}$$

The relativistic gauges give us a local manifestly relativistic formulation of a particular nonlocal interaction between material fields. In our context, we do not see any particular interest in pursuing this discussion, so that we go to the following point.

IV. THE NON-ABELIAN CASE

It is possible to make a non-Abelian extension of the Hagen model. The Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}\phi^\mu_a\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\nu_a + (g/6)\phi^\mu_a\epsilon_{\mu\nu\rho}\,f_{abc}\phi^\rho_b\phi^\nu_c + \phi^a_\mu j^\mu_a, \tag{54}$$

where f_{abc} are the antisymmetric structure constants of the involved non-Abelian compact *Lie* algebra. Here $\mathcal{L}$ is invariant, up to a four-divergence, under the gauge transformations

$$\phi^{\,a}_\mu \to \phi^{\,a}_\mu + \partial_\mu\omega^a + gf^{abc}\phi^{\,b}_\mu\omega^c, \tag{55}$$

provided the current satisfies

$$D^\mu_{ab}\,j^b_\mu = 0, \tag{56}$$

where the covariant derivative D^μ_{ab} is given by

$$D^\mu_{ab} = \partial^\mu\delta_{ab} + gf_{abc}\phi^{\,\mu}_c. \tag{57}$$

We can easily reproduce in the present case the Dirac analysis of theory described by (1). The primary constraints are

$$\pi^1_a - \tfrac{1}{2}\phi^{\,2}_a = \pi^2_a + \tfrac{1}{2}\phi^{\,1}_a = \pi^0_a = 0. \tag{58}$$

There are also secondary constraints

$$\epsilon_{0ij}(\partial^j\phi^{\,i}_a + (g/2)f_{abc}\phi^j_b\phi^{\,i}_c) + j^a_0 = 0, \tag{59}$$

which will be transformed into first-class quantities with respect to the two multiplets of second-class primary constraints. After the strong realization of the constraints, the Hamiltonian reads

$$\mathcal{H}_T = u^0_a\pi^a_0 - \phi^{\,0}_a(\partial^2\phi^{\,1}_a - 2\,\partial^1\pi^1_a + 2gf_{abc}\pi^1_b\phi^{\,1}_c + j^a_0) + \phi^{\,1}_a j^1_a + 2\pi^1_a j^2_a. \tag{60}$$

The theory manifests the same difficulties as the Yang–Mills theory, i.e., the Coulomb gauge is not acceptable and the relativistic gauges cannot be realized in a Lagrangian way without Faddeev–Popov ghosts. This last fact can be seen in a very simple way. The Lagrangian in the relativistic gauges is

$$\mathcal{L} = \tfrac{1}{2}\phi^{\,\mu}_a\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^{\,\nu}_a + (g/6)\epsilon_{\mu\nu\rho}\phi^{\,\mu}_a f_{abc}\phi^{\,\rho}_b\phi^{\,\nu}_c + \phi^{\,a}_\mu j^\mu_a - S_a\,\partial^\mu\phi^{\,a}_\mu + \tfrac{1}{2}aS^2. \tag{61}$$

It gives rise to the equations of motion

$$\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\mu_a - \partial_\nu S_a - (g/2)\epsilon_{\mu\nu\rho}\phi^\mu_c f_{cba}\phi^\rho_b = j^a_\nu \tag{62}$$

and

$$\partial^\mu\phi^a_\mu = aS^a, \tag{63}$$

from which we deduce, using (57),

$$D^{\,\nu}_{ba}\,\partial_\nu S_a = 0. \tag{64}$$

Since the operator $D^\nu\,\partial_\nu$ is not Hermitian, such an equation can never give rise to a unitary S matrix. The fulfillment of this last requirement leads either to Faddeev–Popov ghosts or to non-Lagrangian equations of motion

$$\epsilon_{\mu\nu\rho}\,\partial^\rho\phi^\mu_a - D^{\,ab}_\nu S_b - (g/2)\epsilon_{\mu\nu\rho}\phi^\mu_c f_{cba}\phi^\rho_b = j^a_\nu. \tag{65}$$

Only the axial gauge seems to give a simple formulation of the theory. It develops exactly as in the Abelian theory and we are led to the effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = j^a_0(\mathbf{n}\cdot\boldsymbol{\partial})^{-1}(n^2j^1_a - n^1j^2_a). \tag{66}$$

There is a big difference with the Abelian case, which makes the axial gauge inconvenient here. It depends on the particular choice of $\mathbf{n}$ as it can be seen if we take successively $\mathbf{n} = (0,1)$ and $\mathbf{n} = (1,0)$. In the Abelian case, the current conservation $\partial^\mu j_\mu = 0$ allowed us to show the equivalence between the two Hamiltonians. In the non-Abelian case, the current satisfies instead $D^\mu_{ab}\,j^b_\mu = 0$ and it is impossible to reproduce the same calculation. Axial gauges are therefore not acceptable. This result was already stated by Hagen in a different way. He has shown the nonrelativistic covariance of this gauge. Here, we deduce it very simply from the nonequivalence between the Hamiltonians resulting from two different values for $\mathbf{n}$, which is also a manifestation of the noncovariance.

V. CONCLUSIONS

We have made a Hamiltonian analysis of the Hagen model. Our analysis illustrates how the Dirac algorithm should be used when an odd number of second-class constraints occur. It also shows that the search for constraints is not stopped when a Lagrange multiplier is determined. It may happen that consistency conditions for the Lagrange multipliers determination produce additional constraints. From the point of view of the model itself, it is clear that it furnishes a way to study a particular nonlocal interaction as resulting from a gauge theory. In the non-Abelian case, the model shows essentially the same features as Yang–Mills theory and cannot serve as an exemplary tool to solve the Yang–Mills problems. An interesting particular result is the invalidity of the axial gauge in the non-Abelian case. This invalidity seems, however, to be restricted to the Hagen model. The physical meaning of this *non-Abelian* model can therefore be questioned since there is no gauge in which the meaning is apparent.

It is also of interest to compare the Hagen model with a similar model including a kinetic term $-1/4mF_{\mu\nu}F^{\mu\nu}$. Such a model has been extensively studied by Deser, Jackiw, and Templeton[4] in both Abelian and non-Abelian cases. They also give an interesting topological discussion of $\frac{1}{2}\varphi^\mu\epsilon_{\mu\nu\rho}\,\partial^\rho\varphi^\nu$ and its non-Abelian partner, which appear to generate a mass term in the field equations. This dynamical

model does not present any particularity with respect to the usual constrained theories. There is also no trouble with the non-Abelian version since the physical meaning can be obtained in the gauge characterized by the gauge-fixing term[5] $-(1/2m)\,\partial_k A_0{}^\alpha\,\partial_k A_0{}^\alpha$, where, in the asymptotically free limit, the gauge field part of the Hamiltonian describes a massive scalar field. Formally, the Hagen model is obtained by taking an infinite value for the mass m. This limit is, however, singular in the sense that the constraint system must be handled with the care mentioned in this paper. Moreover, the natural gauge-fixing term also disappears in this limit and this impeaches us to find, in the non-Abelian case, a gauge in which the physical meaning is apparent and described by the sole Hamiltonian.

[1]C. R. Hagen, Ann. Phys. (NY) **157**, 342 (1984); Phys. Rev. D **31**, 331 (1985).
[2]P. A. M. Dirac, *Lectures on Quantum Mechanics, Belfer Graduate School* (Yeshiva U., New York, 1964). See also A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Dynamics, Belfer Graduate School* (Yeshiva U., New York, 1964); and N. Mukunda and E. C. G. Sudarshan, *Classical Dynamics. A Modern Perspective* (Wiley, New York, 1973); K. Sundermeyer, *Constrained Dynamics, Lecture Notes in Physics*, Vol. 169 (Springer, Berlin, 1982).
[3]A. Burnel, Phys. Rev. D **26**, 442 (1982).
[4]S. Deser, R. Jackiw, and S. Templeton, Ann. Phys. (NY) **140**, 372 (1982).
[5]A. Burnel, Phys. Rev. D (to be published).

Color analysis, theory of Γ-graded integrable evolution equations, and super Nijenhuis operators

Robert Trostel
Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstr. 4-5, Sekr. HAD 1, 1000-Berlin 12, Federal Republic of Germany and The Institute for Basic Research, 96 Prescott Street, Cambridge, Massachusetts 02138

(Received 31 December 1984; accepted for publication 23 May 1985)

Using the generalized Grassmann variables or color variables, a theory of Γ-graded integrable evolution equations is presented by elevating the treatments of Magri, Gel'fand–Dorfman, and Fuchssteiner of nonlinear integrable bosonic evolution equations to the Γ-graded case. As an example, it is shown that Kupershmidt's super-KdV is characterized by a Z_2-graded Nijenhuis operator compatible with the underlying Z_2-graded Hamiltonian structure.

I. INTRODUCTION AND MAIN RESULTS

Solitons and integrable nonlinear evolution equations play an important role in several branches of theoretical physics.[1,2] After the famous discovery of the particlelike behavior of solutions of the Korteweg–de Vries (KdV) equation in 1967 by Gardner *et al.*,[3] a great interest has arisen in studying the mathematical foundations of nonlinear evolution equations characterized by an infinite set of independent conservation quantities. Several mathematical concepts such as Bäcklund transformations, Lax pairs, the inverse scattering transform, and Hamiltonian and bi-Hamiltonian structures[4,5] have been shown to be useful and effective tools in understanding and describing integrable nonlinear evolution equations. The recent contributions of Magri,[6] Gel'fand and Dorfman,[7] Fuchssteiner,[8–10] Fokas,[11] Fuchssteiner and Fokas,[12,13] and Fuchssteiner and Oevel,[14] represent an additional culminating point in the mathematical, historical process of understanding and revealing the algebraic and differential geometric structures behind the integrable nonlinear evolution equations. They found that in most of the cases, the soliton equations are moreover characterized by recursion operators,[15] by hereditary symmetries, i.e., by Nijenhuis operators,[6–14] which in addition turn out to be compatible with the underlying symplectic structure. However, there are important nonlinear evolution equations showing hereditary symmetries but lacking a Hamiltonian structure (Burger's equation[10]) and there are cases where a bi-Hamiltonian formulation need not give rise to Nijenhuis operators (Hirota–Satsuma equation[5]).

Recently, there has also been a great interest in studying the variational formulations of dynamical systems, in particular from the viewpoint of the possibility to mutate corresponding prescribed sets of evolution equations. By definition, the mutation consists in premultiplying a prescribed set of evolution equations by invertible integrating matrices (called genotopic or isotopic operators), which lead to variationally self-adjoint evolution equations (see Santilli[16–18] and others).

In Ref. 19, we have given an explicit example of Santilli's method in the area of soliton equations by finding out that the exponential of a scalar multiple of the Nijenhuis operator constitutes an invertible integrating operator, thus leading to a class of different variational formulations of the corresponding integrable dynamical systems.

The question may arise if the technique described above can be elevated to the supersymmetric case, i.e., to a situation where a physical system is described by commuting and anticommuting variables and which is governed by a coupled set of bosonic and fermionic field equations. In a very recent and remarkable paper, Kupershmidt[20] has presented a special coupled set of Z_2-graded evolution equations, called super-KdV equations, characterized by superintegrability under super Poisson brackets. His paper constitutes a decisive step with respect to the extension of theories showing particlelike behavior to the Z_2-graded case.

The purpose of our paper mainly consists in presenting some initiating steps concerning the extension of several familiar concepts such as the "Hamiltonian formalism," the "Birkhoffian formalism," the "Nijenhuis operator," the "symplectic two-form," "Poisson involutivity," etc., to the case of dealing with an associative Γ-graded algebra A of generalized Grassmann numbers (or color numbers)[21] characterized by a finite Abelian grading group Γ and equipped with a Γ-compatible commutation factor σ: $\Gamma\times\Gamma \to \mathbb{C}\setminus\{0\}$. This means that we here present the theoretical framework for Γ-graded Nijenhuis tensors compatible with Γ-graded Hamiltonian structures and that we therefore lay down and prepare the foundations for colored or Γ-graded soliton equations.

At this stage, it might be appropriate to remind the reader of the fact[21] that a color analysis constitutes a natural generalization of the usual Z_2-graded Grassmann analysis. It is nearly superfluous to recall that the latter is undoubtedly important in conventional Z_2-graded, i.e., supersymmetric theoretical physics according to Martin,[22] Volkov and Akulov,[23] Wess and Zumino,[24] Salam and Strathdee,[25] Berezin,[26] Ramond,[27] Casalbuoni,[28] Neveu and Schwarz,[29] Kostant,[30] Corwin, Ne'eman and Sternberg,[31] Rogers,[32] Dell and Smolin,[33] Jadczyk and Pilch,[34] Kac,[35] de Witt,[36] and others.

It has moreover been found and emphasized in Ref. 21 that the variational formulations of dynamical systems, the Hamiltonian and Poisson bracket formalism, can naturally be elevated to the Γ-graded case, so that it is evident to study properties of generalized soliton equations from the outset in

 0022-2488/85/123160-12$02.50

the abstract Γ-graded theoretical framework. Its use is also suggested and recommended if one wishes to generalize the statistics in quantum field theory (see Ohnuki and Kamefuchi,[37–39] Omote and Kamefuchi,[40] and Levine and Tomozawa[41,42]) or if one deals with the recently introduced generalized Lie (super)algebras, also called color (super)algebras (e.g., Kac,[35] Rittenberg and Wyler,[43,44] Lukierski and Rittenberg,[45] Agrawala,[46] Green and Jarvis,[47] Scheunert,[48] and others). Even in the conventional $\mathbf{Z}_2$-graded case, we arrive at a couple of apparently new propositions that generalize the corresponding bosonic treatments of integrable evolution equations[6–14,19] to the mixed bosonic fermionic theoretical framework. As one example of our theory, we show that Kupershmidt's super KdV possesses a $\mathbf{Z}_2$-graded Nijenhuis tensor compatible with the $\mathbf{Z}_2$-graded Hamiltonian structure thus permitting us to derive a $\mathbf{Z}_2$-graded Santilli's isotopic degree of freedom in the $\mathbf{Z}_2$-graded variational formulation of the super KdV by using the exponential of the $\mathbf{Z}_2$-graded Nijenhuis operator as an integrating operator for the super KdV equations.

II. COLOR ANALYSIS

Let $\mathbf{A}$ be a Γ-graded associative Banach algebra over the complex numbers $\mathbb{C}$, with unit[49]

$$\mathbf{A}=\bigoplus_{\gamma\in\Gamma}\mathbf{A}_\gamma,\quad \mathbf{A}_\gamma\mathbf{A}_\delta\subset\mathbf{A}_{\gamma+\delta},\quad \forall\gamma,\delta\in\Gamma,$$

where the $\mathbf{A}_\gamma$ are the Banach subspaces of $\mathbf{A}$ and where $o\in\Gamma$ denotes the neutral element in the finite Abelian grading group Γ. Moreover, $\mathbf{A}$ is characterized by a commutation function

$$\sigma:\ \Gamma\times\Gamma\to\mathbb{C}\setminus\{0\},\quad \sigma(\alpha,\beta)\,\sigma(\beta,\alpha)=1,$$
$$\sigma(\alpha,\beta)\,\sigma(\alpha,\gamma)=\sigma(\alpha,\beta+\gamma),\quad \forall\alpha,\beta,\gamma\in\Gamma,$$
$$q^1_\alpha\,q^2_\beta=\sigma(\alpha,\beta)\,q^2_\beta\,q^1_\alpha,\quad \forall q^1_\alpha\in\mathbf{A}_\alpha,\quad q^2_\beta\in\mathbf{A}_\beta,$$

where the last equation can be found, for example, in Ref. 43. The conditions on the commutation function σ entail the relations

$$\sigma(\alpha,0)=\sigma(0,\alpha)=1,\quad \sigma(\alpha,\alpha)=\pm 1,$$
$$\sigma(\alpha,\beta)=\sigma(-\alpha,-\beta)=\sigma(\beta,-\alpha)$$
$$=\sigma(-\beta,\alpha),\quad \forall\alpha,\beta\in\Gamma,$$

which indicate that one deals with ordinary bosonic or fermionic numbers in each subspace $\mathbf{A}_\gamma$ of $\mathbf{A}$. Let now $[n_\Gamma]:=(n_\gamma)_{\gamma\in\Gamma}$ denote a set of positive integers $n_\gamma\in\mathbf{N}_0$ satisfying $n_\gamma=n_{-\gamma}$. We introduce the Banach space (equipped with the product topology)

$$\mathbb{G}^{[n_\Gamma]}:=\bigoplus_{\gamma\in\Gamma}(\mathbf{A}_\gamma)^{n_\gamma}$$

being an $\mathbf{A}_0$ module. Elements of $\mathbb{G}^{[n_\Gamma]}$ are denoted by

$$x,y,z\in\mathbb{G}^{[n_\Gamma]},\quad x=(q^i_\alpha)^{i\in\{1,\dots,n_\alpha\}}_{\alpha\in\Gamma},$$
$$y=(y^j_\gamma)^{j\in\{1,\dots,n_\gamma\}}_{\gamma\in\Gamma},\quad z=(z^k_\delta)^{k\in\{1,\dots,n_\delta\}}_{\delta\in\Gamma},$$
$$q^i_\alpha,\ y^i_\alpha,\ z^i_\alpha\in\mathbf{A}_\alpha.$$

In what follows, we shall be concerned with the set $\mathbb{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathbf{A})$ of $\mathbf{A}$-valued functions f being infinitely many times continuously $\mathbf{A}_0$ differentiable.[49] If $(\lambda\in\mathbb{R})$,

$$\left(\frac{\partial}{\partial q^i_\gamma}f\right)_x=\left(\frac{\partial f}{\partial q^i_\gamma}\right)_x,$$
$$\frac{d}{d\lambda}f_{x+\lambda y}|_{\lambda=0}=\sum_{\gamma\in\Gamma}\sum_{i=1}^{n_\gamma}y^i_\gamma\left(\frac{\partial}{\partial q^i_\gamma}f\right)_x$$

denotes the left derivative of the function f with respect to $q^i_\gamma\in\mathbf{A}_\gamma$ at $x\in\mathbb{G}^{[n_\Gamma]}$, we can derive the generalized or color Heisenberg commutation rules[49]

$$\frac{\partial}{\partial q^i_\alpha}\frac{\partial}{\partial q^j_\beta}=\sigma(\alpha,\beta)\frac{\partial}{\partial q^j_\beta}\frac{\partial}{\partial q^i_\alpha},$$
$$\frac{\partial}{\partial q^i_\alpha}q^j_\beta=\delta_i{}^j\,\delta_\alpha{}^\beta+\sigma(\beta,\alpha)\,q^j_\beta\frac{\partial}{\partial q^i_\alpha},$$
$$\forall\alpha,\beta\in\Gamma,\quad \forall i\in\{1,\dots,n_\alpha\},\quad \forall j\in\{1,\dots,n_\beta\}.$$

Note for the following that the generalized Leibnitz rule

$$\frac{\partial}{\partial q^i_\alpha}(f_\beta g_\gamma)=\frac{\partial f_\beta}{\partial q^i_\alpha}g_\gamma+\sigma(\beta,\alpha)f_\beta\frac{\partial g_\gamma}{\partial q^i_\alpha}$$

is valid for arbitrary

$$f_\beta\in\mathbb{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathbf{A}_\beta),\quad g_\gamma\in\mathbb{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathbf{A}_\gamma).$$

We point out that

$$\frac{\partial f_\beta}{\partial q^i_\alpha}\in\mathbb{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathbf{A}_{\beta-\alpha})$$

holds. The Γ-graded associative Banach algebra $\mathbf{A}$ can, for example, be constructed as follows. Take $|\Gamma|=\nu+\mu$ infinite-dimensional vector spaces $\mathscr{V}_\alpha$ with countable basis $\{e_\alpha i\}i\in\mathbb{N}$ for each $\alpha\in\Gamma$, where $|\Gamma|$ denotes the finite number of elements of Γ, where ν denotes the finite number of bosonic elements γ_i of Γ [i.e., satisfying $\sigma(\gamma_i,\gamma_i)=+1$], and where μ denotes the corresponding number of fermionic elements β_j of $\Gamma[\sigma(\beta_j,\beta_j)=-1]$. The tensor space $T(\mathscr{V})$ over $\mathscr{V}=\oplus_{\alpha\in\Gamma}\mathscr{V}_\alpha$ is introduced and the ideal I generated by the elements of the form $v_\alpha\otimes w_\delta-\sigma(\alpha,\delta)w_\delta\otimes v_\alpha$, $v_\alpha\in\mathscr{V}_\alpha$, $w_\delta\in\mathscr{V}_\delta$, $\alpha,\delta\in\Gamma$ is constructed. One arrives at the well-defined associative algebra $\mathscr{A}:=T(\mathscr{V})/I$[48,50] and $\mathbf{A}$ is taken to be the complex linear hull of the linear independent elements $E^{[0]}:=1\in\mathscr{A}$ and

$$E^{[N]}:=\underset{\gamma_1}{(e_1)^{n^1_1}}\underset{\gamma_1}{(e_2)^{n^2_1}}\cdots\underset{\gamma_1}{(e_{L_1})^{n^{L_1}_1}}\cdots\underset{\gamma_\nu}{(e_1)^{n^1_\nu}}\cdots\underset{\gamma_\nu}{(e_{L_\nu})^{n^{L_\nu}_\nu}}$$
$$\cdot\underset{\beta_1}{(e_{j^1_1})}\cdots\underset{\beta_1}{(e_{j^{k_1}_1})}\cdots\underset{\beta_\mu}{(e_{j^1_\mu})}\cdots\underset{\beta_\mu}{(e_{j^{k_\mu}_\mu})}\in\mathscr{A},$$

where

$$n^1_1,n^2_1,\dots,n^{L_1}_1;\ n^1_2,\dots,n^{L_2}_2;\dots;n^1_\nu,\dots,n^{L_\nu}_\nu\in\mathbf{N}_0,$$
$$j^1_1<j^2_1<\cdots<j^{K_1}_1,\dots,j^1_\mu<\cdots<j^{K_\mu}_\mu\in\mathbb{N},$$

where $L_1,\dots,L_\nu,K_1,\dots,K_\mu\in\mathbf{N}_0$ and where a certain counting procedure has been adopted, $N\in\mathbb{N}$. If M denotes

$$M:=\max\{|\sigma(\alpha,\delta)|;\ \alpha,\delta\in\Gamma\}\geqslant 1,$$

$\mathbf{A}$ is made into a Γ-graded associative Banach algebra over the complex numbers $\mathbb{C}$, with unit and equipped with a Γ-

compatible commutation factor σ by taking the norm

$$\|x\| := |x_0| + M^{\nu+\mu-1}\sum_{N=1}^{\infty}|x_N| < \infty,$$

$$x = \sum_{N=0}^{\infty} x_N E^{[N]} \in \mathbf{A}, \quad x_N \in \mathbf{C}.$$

A tedious calculation shows that A in addition has the important property that each $\mathbf{A}_0$-linear bounded linear map f from $\mathbf{A}_\alpha$ to $\mathbf{A}_\delta$, i.e., satisfying

$$f \in \mathscr{L}_{\mathbf{A}_0}(\mathbf{A}_\alpha, \mathbf{A}_\delta), \ \alpha,\delta \in \Gamma,$$

$$f(x_0 \cdot y_\alpha) = x_0 \cdot f(y_\alpha), \quad \forall x_0 \in \mathbf{A}_0, \quad y_\alpha \in \mathbf{A}_\alpha$$

is already generated by a unique element $z_{\delta-\alpha} \in \mathbf{A}_{\delta-\alpha}$ so that $f(y_\alpha) = y_\alpha z_{\delta-\alpha}$ holds. We however find it advisable to weaken the additional topological condition expressed in Ref. 49, p. 1545, third line from the bottom to

$$\|z_{\delta-\alpha}\| < M^{\nu+\mu}\|f\|,$$

with corresponding modifications in Theorems VI.12–VI.14 in Ref. 49.

III. Γ-GRADED BIRKHOFFIAN FORMALISM

For the following, let us consider a prescribed set of infinitely many times continuously $\mathbf{A}_0$-differentiable $\mathbf{A}_{-\gamma}$-valued functions

$$\mathscr{E}_{j,-\gamma}: \quad [t_a,t_b]\times\mathbf{G}^{[n_\Gamma]}\times\mathbf{G}^{[n_\Gamma]} \to \mathbf{A}_{-\gamma},$$
$$j = 1,\ldots,n_\gamma, \quad \gamma\in\Gamma, \quad t_a < 0 < t_b, \quad t_a,t_b \in \mathbf{R},$$

called the set of evolution terms associated with a certain physical system of differential type 1 concerning the time derivatives. The special elements $\underline{x}$ of $\mathbf{Q} = C^{\mathscr{N}}([t_a,t_b], \mathbf{G}^{[n_\Gamma]})$, $\mathscr{N} > 4$ [i.e., of the Banach space Q of all at least $\mathscr{N}$ times continuously differentiable maps from $[t_a,t_b]$ to $\mathbf{G}^{[n_\Gamma]}$ with standard norm $\|\underline{x}\| := \max\left\{\|\overset{(l)}{x}(t)\|;\ l = 0,1,\ldots,\mathscr{N};\ t\in[t_a,t_b]\right\}$, where $\overset{(l)}{x}(t) = (\overset{(l)}{q}{}^i_\alpha(t))^{i\in\{1,\ldots,n_\alpha\}}_{\alpha\in\Gamma} \in \mathbf{G}^{[n_\Gamma]}$], which satisfy

$$\mathscr{E}_{j,-\gamma}(t,x(t),\dot{x}(t)) = 0,$$
$$\forall j\in\{1,\ldots,n_\gamma\}, \quad \forall\gamma\in\Gamma, \quad \forall t\in[t_a,t_b], \tag{1}$$

constitute the solution set $\mathscr{S} \subset \mathbf{Q}$ consisting of all physical curves within a fixed time interval $[t_a,t_b]$ of the corresponding physical system. As in the pure bosonic case,[16,18] we now want to study the conditions on the functional structure of a prescribed set of evolution terms $\mathscr{E}_{j,-\gamma}$ so that they result from an action functional, i.e., so that they turn out to be generalized variationally self-adjoint. According to Ref. 21, the generalized variational self-adjointness conditions read

$$\mathscr{O} = \frac{\partial\mathscr{E}_{j,-\gamma}}{\partial q^i_\beta} + \sigma(\beta,\gamma) \times\left[-\frac{\partial\mathscr{E}_{i,-\beta}}{\partial q^j_\gamma} + \frac{d}{dt}\frac{\partial\mathscr{E}_{i,-\beta}}{\partial\dot{q}^j_\gamma}\right], \tag{2a}$$

$$\mathscr{O} = \frac{\partial\mathscr{E}_{j,-\gamma}}{\partial\dot{q}^i_\beta} + \sigma(\beta,\gamma)\frac{\partial\mathscr{E}_{i,-\beta}}{\partial\dot{q}^j_\gamma}, \tag{2b}$$
$$\forall i\in\{1,\ldots,n_\beta\}, \quad \forall\beta\in\Gamma,$$
$$\forall j\in\{1,\ldots,n_\gamma\}, \quad \forall\gamma\in\Gamma, \quad \forall\underline{x}\in\mathbf{Q},$$

or more abstract[49]

$$d_{\mathbf{Q}_\square}[\omega_1]_{\mathbf{Q}_\square} = 0, \tag{3}$$

where $\omega_1 \in \mathbf{C}^\infty_{\mathbf{A}_0}(\mathbf{Q}, \mathscr{L}_{\mathbf{A}_0}(\mathbf{Q};\mathbf{A}_0))$, defined by

$$\omega_{1_{\underline{y}}}(\underline{y}) := \int_{t_a}^{t_b} dt \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} y^j_\gamma(t)\cdot\mathscr{E}_{j,-\gamma}(t,x(t),\dot{x}(t)), \tag{4}$$

denotes the "physical one-form" on the curve space Q (see Ref. 49), where $\mathbf{Q}_\square$ denotes the Banach subspace (of Q) consisting of all at least $\mathscr{N}$-times continuously differentiable maps from $[t_a,t_b]$ to $\mathbf{G}^{[n_\Gamma]}$, which vanish with all their derivatives up to $\mathscr{N}-2$ at the boundary t_a and t_b where $C^\infty_{\mathbf{A}_0}(\mathbf{Q}, \mathscr{L}^{\text{anti}}_{\mathbf{A}_0}(\mathbf{Q}^p;\mathbf{A}_0))$ denotes the vector space consisting of all infinitely many times continuously $\mathbf{A}_0$-differentiable maps from Q to the completely antisymmetric bounded $p-\mathbf{A}_0$-linear maps from Q to $\mathbf{A}_0$, $p\in\mathbf{N}_0$ (generalized p forms).[49]

Now using (2), one can easily show the following theorem.

Theorem 1: $\mathscr{E}_{j,-\gamma}$ is generalized variationally self-adjoint exactly iff there exist infinitely many times continuously $\mathbf{A}_0$-differentiable $\mathbf{A}_{-\gamma-\delta}$-valued functions $\psi_{j,-\gamma;i,-\delta}$: $[t_a,t_b]\times\mathbf{G}^{[n_\Gamma]} \to \mathbf{A}_{-\gamma-\delta}$ and also corresponding $\mathbf{A}_{-\delta}$-valued functions $\phi_{i,-\delta}$: $[t_a,t_b]\times\mathbf{G}^{[n_\Gamma]} \to \mathbf{A}_{-\delta}$ so that

$$\mathscr{E}_{j,-\gamma}(t,x(t),\dot{x}(t)) = \sum_{\delta\in\Gamma}\sum_{k=1}^{n_\delta}\dot{q}^k_\delta(t)\psi_{k,-\delta;j,-\gamma}(t,x(t)) + \phi_{j,-\gamma}(t,x(t)), \tag{5}$$
$$\forall j\in\{1,\ldots,n_\gamma\} \quad \forall\gamma\in\Gamma, \quad \forall\underline{x}\in\mathbf{Q},$$

holds, where in addition $\psi_{k,-\delta;j,-\gamma}$ and $\phi_{j,-\gamma}$ satisfy the σ-closure condition

$$0 = \sigma(\gamma,\delta)\frac{\partial}{\partial q^i_\delta}\psi_{k,-\alpha;j,-\gamma} + \sigma(\alpha,\gamma)\frac{\partial}{\partial q^j_\gamma}\psi_{i,-\delta;k,-\alpha} + \sigma(\delta,\alpha)\frac{\partial}{\partial q^k_\alpha}\psi_{j,-\gamma;i,-\delta}, \tag{6}$$

the σ antisymmetry

$$0 = \psi_{i,-\delta;j,-\gamma} + \sigma(\delta,\gamma)\psi_{j,-\gamma;i,-\delta}, \tag{7}$$

and

$$0 = \frac{\partial}{\partial q^i_\delta}\phi_{j,-\gamma} - \sigma(\delta,\gamma)\frac{\partial}{\partial q^j_\gamma}\phi_{i,-\delta} + \sigma(\delta,\gamma)\frac{\partial\psi_{j,-\gamma;i,-\delta}}{\partial t}. \tag{8}$$

Proof: According to (2a), $\partial^2\mathscr{E}_{i,-\beta}/\partial\dot{q}^k_\alpha\,\partial\dot{q}^j_\gamma$ must identically vanish because no other $\ddot{q}^k_\alpha$ term appears in Eq. (2a). Equation (2b) immediately leads to the σ-antisymmetry relation (7). Exploiting once more (2a), using the commutation rules like $(\partial/\partial q^j_\gamma)\dot{q}^i_\alpha = \sigma(\alpha,\gamma)\dot{q}^i_\alpha\,\partial/\partial q^j_\gamma$ and using the conditions for the Γ-compatible commutation factor σ (see Ref. 21), one finally arrives at (5), (6), and (8). ■

Theorem 2: The physical one-form ω_1 (4) is generalized variationally self-adjoint with respect to $(\mathbf{Q},\mathbf{Q}_\square)$, i.e., ω_1 sat-

isfies (3) exactly iff there exist two, in general, time-dependent forms

$$\omega_{2}\in C^{\infty}_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}^{\text{anti}}_{A_0}((\mathbb{G}^{[n_\Gamma]})^2; A_0)),$$

$$\omega_{1}\in C^{\infty}_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}_{A_0}(\mathbb{G}^{[n_\Gamma]}; A_0)),$$

being C^∞ in all of their arguments so that

$$\omega_{1\underline{x}}(\underline{y}) = \int_{t_a}^{t_b} dt \left[(\omega_2)_{x(t)}(y(t),\dot{x}(t)) - (\omega_1)_{x(t)}(y(t))\right] \tag{9}$$

holds for all $\underline{x},\underline{y}\in Q$, where ω_1 and ω_2 satisfy

$$d\omega_2 = 0, \tag{10}$$

$$d\omega_1 = \frac{\partial}{\partial t}\omega_2. \tag{11}$$

Proof: Using Theorem 1, we define

$$(\omega_2)_x(y,z) = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{\delta\in\Gamma}\sum_{k=1}^{n_\delta} y^j_\gamma z^k_\delta \psi_{k,-\delta;j,-\gamma}(t,x),$$

$$(\omega_1)_x(y) = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} y^j_\gamma (-1)\phi_{j,-\gamma}(t,x),$$

$$x = (q^i_\alpha)^{i\in\{1,\ldots,n_\alpha\}}_{\alpha\in\Gamma},\quad y = (y^j_\gamma)^{j\in\{1,\ldots,n_\gamma\}}_{\gamma\in\Gamma}, \tag{12}$$

$$z = (z^k_\delta)^{k\in\{1,\ldots,n_\delta\}}_{\delta\in\Gamma}\in\mathbb{G}^{[n_\Gamma]},$$

and observe that (7) corresponds to $(\omega_2)_x(y,z) = -(\omega_2)_x(z,y)$, $\forall x,y,z\in\mathbb{G}^{[n_\Gamma]}$ that (6) corresponds to $d\omega_2 \overset{(7)}{=} 0$ and that finally (8) corresponds to $d\omega_1 = (\partial/\partial t)\omega_2$. ■

Comment: It might appear surprising that in contrast to Theorem 1, Theorem 2 seems to have nothing to do with any grading Γ or with any Γ-compatible commutation factor. But this is absolutely not the case. First of all, the forms ω_1, ω_2 considered by us are not usual forms but they are A_0 valued and A_0 linear. Second, they are infinitely many times continuously generalized superdifferentiable, i.e., A_0 differentiable (see Ref. 49), in contrast to the conventional forms, which are only C^∞ sections in antisymmetrized tensor products of the cotangent bundle. In addition, we refer to Ref. (49), p. 1550, where we have shown that the exterior differential d and the Poincaré operator K are compatible with the category of generalized superdifferentiable A_0-valued and A_0-linear forms. We now refer to Ref. 16 for the bosonic version of the next theorem.

Theorem 3 (generalized Pfaffian action principle): Let $\mathscr{E}_{j,-\gamma}$ be generalized variationally self-adjoint, or equivalently stated, let the physical one-form $\omega_1(4)$ satisfy $d_{Q_\square}[\omega_1]_{Q_\square} = 0$, then the action functional $\omega_0\in C^\infty_{A_0}(Q,A_0)$ up to boundary terms (see Ref. 49) reads

$$\omega_{0\underline{x}} = K(\omega_1)_{\underline{x}} = \int_0^1 d\tau\, \omega_{1\tau\underline{x}}(\underline{x})$$

$$= \int_0^1 d\tau \int_{t_a}^{t_b} dt \left[(\omega_2)_{\tau x(t)}(x(t),\tau\dot{x}(t)) - (\omega_1)_{\tau x(t)}(x(t))\right]$$

$$= \int_{t_a}^{t_b} dt \left[(\nu_1)_{x(t)}(\dot{x}(t)) - (\nu_0)_{x(t)}\right], \tag{13}$$

where

$$\nu_1 := K(\omega_2)\in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}; \mathscr{L}_{A_0}(\mathbb{G}^{[n_\Gamma]}; A_0)),$$

$$\nu_0 := K(\omega_1)\in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}; A_0), \tag{14}$$

satisfy

$$d\nu_1 = \omega_2,\quad d\nu_0 = \omega_1 - \frac{\partial}{\partial t}\nu_1. \tag{15}$$

Proof: Using the lemma of Poincaré $d_{Q_\square}K_{Q_\square} + K_{Q_\square}d_{Q_\square} = \text{id}$ for the $Q_\square$-equivalence classes $[\omega_p]_{Q_\square}$ of generalized p forms $\omega_p\in C^\infty_{A_0}(Q,\mathscr{L}^{\text{anti}}_{A_0}(Q^p;A_0))$ (see Ref. 49), one obtains $[\omega_1]_{Q_\square} = d_{Q_\square}[\omega_0]_{Q_\square}$ with $\omega_0 := K(\omega_1)$ because K and d are compatible with the $Q_\square$-equivalence structure.[49] By virtue of Theorem 1 or 2, we dispose of a certain functional structure of the evolution terms $\mathscr{E}_{j,-\gamma}$, which finally permits us to understand (13) and to derive (15), once more using the lemma of Poincaré $dK + Kd = \text{id}$ but now applied to (time-dependent) generalized p forms

$$\omega_p\in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}^{\text{anti}}_{A_0}((\mathbb{G}^{[n_\Gamma]})^p; A_0)).$$ ■

Using (12), one also has the following expression for the generalized symplectic potential

$$(\nu_1)_x(y) = \int_0^1 d\tau\, \tau \sum_{\gamma\in\Gamma}\sum_{\delta\in\Gamma}\sum_{k=1}^{n_\delta}\sum_{j=1}^{n_\gamma} q^j_\gamma y^k_\delta \times \psi_{k,-\delta;j,-\gamma}(t,\tau x), \tag{16}$$

and for the generalized Birkhoffian

$$(\nu_0)_x = -\int_0^1 d\tau \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} q^j_\gamma \phi_{j,-\gamma}(t,\tau x), \tag{17}$$

where we have chosen the same designation "Birkhoffian" for ν_0 in our Γ-graded case (with arbitrary finite Abelian grading group Γ and with a corresponding arbitrary but Γ-compatible commutation factor) as it has been introduced and chosen by Santilli[16,18] in the pure bosonic case, i.e., if $\Gamma = Z_1 = \{0\}$. Notice, if we specialize in the case $\Gamma = Z_2 = \{0,1\}$ and $\sigma(0,0) = \sigma(0,1) = \sigma(1,0) = -\sigma(1,1) = 1$, we would consequently arrive at a "super-Birkhoffian-formalism."[51]

IV. Γ-GRADED HAMILTONIAN FORMALISM

In what follows, we shall be concerned with generalized Hamiltonian systems, which are by definition characterized first by generalized variationally self-adjoint evolution terms $\mathscr{E}_{j,-\gamma}$, being explicitly independent of time,

$$\mathscr{E}_{j,-\gamma}(x,\dot{x}) = \sum_{\delta\in\Gamma}\sum_{k=1}^{n_\delta} \dot{q}^k_\delta \psi_{k,-\delta;j,-\gamma}(x) + \phi_{j,-\gamma}(x), \quad (18)$$

and which are characterized second by the invertibility of the $\mathbf{A}_{-\delta-\gamma}$-valued functions $\psi_{k,-\delta;j,-\gamma}\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_{-\delta-\gamma})$, i.e., by the existence of $(\psi^{-1})^{k,-\delta;j,-\gamma}\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_{\delta+\gamma})$ satisfying

$$\sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} \psi_{k,-\delta;j,-\gamma}(x)(\psi^{-1})^{j,-\gamma;l,-\beta}(x) = \delta_k{}^l\delta_\delta{}^\beta,$$

$$\sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} (\psi^{-1})^{l,-\beta;j,-\gamma}(x)\psi_{j,-\gamma;k,-\delta}(x) = \delta_l{}^k\delta_\beta{}^\delta. \quad (19)$$

Proposition 4: Presupposing (19), Eqs. (6) and (7) are equivalent to

$$0 = (\psi^{-1})^{j,-\gamma;k,-\delta} + \sigma(\gamma,\delta)\sigma(\gamma,\gamma)\sigma(\delta,\delta) \times(\psi^{-1})^{k,-\delta;j,-\gamma}, \quad (20)$$

$$\begin{aligned} 0 = \sum_{\delta\in\Gamma}\sum_{r=1}^{n_\delta}\Big[&\sigma(\beta,\alpha)\sigma(\gamma,\beta)\sigma(\gamma,\gamma)(\psi^{-1})^{i,-\alpha;r,-\delta} \\ &\times\frac{\partial}{\partial q^r_\delta}(\psi^{-1})^{j,-\beta;k,-\gamma} \\ &+ \sigma(\gamma,\beta)\sigma(\alpha,\gamma)\sigma(\alpha,\alpha)(\psi^{-1})^{j,-\beta;r,-\delta} \\ &\times\frac{\partial}{\partial q^r_\delta}(\psi^{-1})^{k,-\gamma;i,-\alpha} \\ &+ \sigma(\alpha,\gamma)\sigma(\beta,\alpha)\sigma(\beta,\beta)(\psi^{-1})^{k,-\gamma;r,-\delta} \\ &\times\frac{\partial}{\partial q^r_\delta}(\psi^{-1})^{i,-\alpha;j,-\beta}\Big]. \end{aligned} \quad (21)$$

The proof of Proposition 4 is tedious but straightforward and is left to the interested reader.

We now define

$$\begin{aligned} &\omega_2^{-1}(\rho_1,\rho_1') \\ &:= \sum_{\gamma\in\Gamma}\sum_{\alpha\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{k=1}^{n_\alpha}(\rho_1)_{j,-\gamma}(\psi^{-1})^{j,-\gamma;k,-\alpha}(\rho_1')_{k,-\alpha}, \\ &\forall\rho_1,\rho_1'\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathscr{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0)), \end{aligned} \quad (22)$$

$$\rho_{1_x}(y) = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} y^j{}_\gamma(\rho_1)_{j,-\gamma}(x),$$

$$(\rho_1)_{j,-\gamma}\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{A}_{-\gamma}),$$

and the "generalized time-evolution bracket"

$$\{G_0,F_0\} := \omega_2^{-1}(dG_0,dF_0), \quad \forall G_0,F_0\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0). \quad (23)$$

We also need the concept "generalized Hamiltonian vector field"

$$(dG_0)^\# := \sum_{\gamma\in\Gamma}\sum_{\alpha\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{k=1}^{n_\alpha}\frac{\partial G_0}{\partial q^j_\gamma}(\psi^{-1})^{j,-\gamma;k,-\alpha}\cdot\frac{\partial}{\partial q^k_\alpha}, \quad (24)$$

for any $\mathbf{A}_0$-valued function $G_0\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0)$.

Theorem 5: Presupposing (19), Eqs. (6) and (7) are equivalent to

$$\omega_2^{-1}(\rho_1,\rho_1') = -\omega_2^{-1}(\rho_1',\rho_1), \quad \forall\rho_1,\rho_1'\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathscr{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0)) \quad (25)$$

and

$$\{F_0,\{G_0,H_0\}\} + \{G_0,\{H_0,F_0\}\} + \{H_0,\{F_0,G_0\}\} = 0, \quad \forall F_0,G_0,H_0\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0). \quad (26)$$

Proof: The following formulas are valid:

$$\begin{aligned} &\omega_2^{-1}(\rho_1,\rho_1') + \omega_2^{-1}(\rho_1',\rho_1) \\ &\overset{(22)}{=} \sum_{\gamma\in\Gamma}\sum_{\alpha\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{i=1}^{n_\alpha}(\rho_1)_{j,-\gamma}[(\psi^{-1})^{j,-\gamma;i,-\alpha} \\ &\quad + \sigma(\alpha,\alpha)\sigma(\gamma,\alpha)\sigma(\gamma,\gamma)(\psi^{-1})^{i,-\alpha;j,-\gamma}](\rho_1')_{i,-\alpha}, \end{aligned} \quad (27)$$

$$\begin{aligned} &\{F_0,\{G_0,H_0\}\} + \text{cycl} \\ &\overset{(7)\text{or}(25)}{=} \sum\frac{\partial F_0}{\partial q^j_\gamma}\frac{\partial G_0}{\partial q^k_\beta}\frac{\partial H_0}{\partial q^r_\delta}\sigma(\delta,\gamma) \\ &\quad\times\sum_{\alpha\in\Gamma}\sum_{i=1}^{n_\alpha}\Big[\sigma(\beta,\gamma)\sigma(\delta,\beta)\sigma(\delta,\delta)(\psi^{-1})^{j,-\gamma;i,-\alpha} \\ &\quad\times\frac{\partial}{\partial q^i_\alpha}(\psi^{-1})^{k,-\beta;r,-\delta} \\ &\quad + \sigma(\delta,\beta)\sigma(\gamma,\delta)\sigma(\gamma,\gamma)(\psi^{-1})^{k,-\beta;i,-\alpha} \\ &\quad\times\frac{\partial}{\partial q^i_\alpha}(\psi^{-1})^{r,-\delta;j,-\gamma} \\ &\quad + \sigma(\gamma,\delta)\sigma(\beta,\gamma)\sigma(\beta,\beta)(\psi^{-1})^{r,-\delta;i,-\alpha} \\ &\quad\times\frac{\partial}{\partial q^i_\alpha}(\psi^{-1})^{j,-\gamma;k,-\beta}\Big], \end{aligned} \quad (28)$$

so that the proof becomes obvious by using Proposition 4. ■

Proposition 6: $d\omega_2 = 0$ and (19) imply

$$\omega_2(\cdot,(dG_0)^\#) = dG_0, \quad (29)$$

$$\omega_2^{-1}(dG_0,\cdot) = (dG_0)^\#, \quad (30)$$

$$[(dG_0)^\#,(dF_0)^\#]_- = (d\{G_0,F_0\})^\#. \quad (31)$$

The proofs of (29) and (30) are easy. According to Theorem 5, we have $([(dG_0)^\#,(dF_0)^\#]_- - (d\{G_0,F_0\})^\#)\times(H_0) = 0$, $\forall G_0,F_0,H_0\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0)$. We then choose $H_0 = \Sigma q^i_\alpha y_{i,-\alpha}$ with arbitrary but constant $y_{i,-\alpha}\in\mathbf{A}_{-\alpha}$. This proves (31). ■

The generalized Hamiltonian evolution equations now read

$$\dot{q}^i_\beta = (d\nu_0)^{\#i}_{x(t)\beta} = [\omega_2^{-1}(d\nu_0,\cdot)_{x(t)}]^i_\beta, \quad (32)$$

and are obtained by putting the evolution terms (18) zero and by using (19). Equation (32) is nothing more than the Γ-graded flow equation of the generalized Hamiltonian vector field $(d\nu_0)^\#$ with generalized Hamiltonian $\nu_0\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbf{A}_0)$. The Γ-graded (local) flow $\Phi_t\in\mathbf{C}^\infty_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{G}^{[n_\Gamma]})$ (being $\mathbf{C}^\infty$ in all of its arguments; for brevity, we do not use the, of course,

more precise flow-box notation[52]) satisfies

$$\Phi_{t+s} = \Phi_t \circ \Phi_s, \quad \forall t,s,t+s \in [t_a,t_b], \quad \Phi_0 = \mathrm{id}.$$

As in the pure bosonic case,[53] we have the following theorem.

Theorem 7: Let the generalized two-form

$$\omega_2 \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}^{\mathrm{anti}}_{A_0}((\mathbb{G}^{[n_\Gamma]})^2; A_0)),$$

$$\omega_{2_x}(y,z) := \sum_{\gamma\in\Gamma}\sum_{\delta\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{k=1}^{n_\delta} y^j_\gamma z^k_\delta \psi_{k,-\delta;j,-\gamma}(x)$$

be closed and let (19) be valid. Let now Φ be C^∞ and let $\Phi_u \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{G}^{[n_\Gamma]})$ be the (local) Γ-graded flow of a generalized vector field $X = \Sigma X^j_\beta(\cdot)\cdot\partial/\partial q^j_\beta \in \mathfrak{X}^{A_0}_0(\mathbb{G}^{[n_\Gamma]})$, $X^j_\beta(\cdot) \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]};A_\beta)$. Then, Φ_u are generalized canonical transformations, i.e., $\Phi^*_u(\omega_2) = \omega_2$, $\forall u\in\mathbb{R}$, if and only if X is generalized Hamiltonian, i.e., $X = (dF_0)^\#$, where $F_0 \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]};A_0)$.

Proof: The formulas

$$L_X = di_X + i_X d \tag{33a}$$

and

$$\frac{d}{du}\Phi^*_u = \Phi^*_u L_X, \tag{33b}$$

where L denotes the Lie derivative, also apply in the case of dealing with generalized p forms $\omega_p \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}^{\mathrm{anti}}_{A_0}((\mathbb{G}^{[n_\Gamma]})^p; A_0))$. ■

Theorem 8: The Lie algebra of A_0-valued conservation quantities with respect to the generalized Hamiltonian ν_0 is isomorphic (modulo constant) to the Lie algebra of infinitesimal generalized canonical transformations which leave the generalized Hamiltonian ν_0 invariant.

Proof: The Lie algebra of A_0-valued conservation quantities is equipped with the generalized time evolution bracket (23) as product. The isomorphism (modulo constants) is then given by $F_0 \to (dF_0)^\#$. ■

V. GENERALIZATION OF SANTILLI'S ISOTOPIC ACTION FUNCTIONALS AND OF NIJENHUIS TENSORS

The question may arise if there are invertible generalized tensors $\mathscr{T} \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}, \mathscr{L}_{A_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{G}^{[n_\Gamma]}))$, which map generalized vector fields $X = \Sigma X^j_\beta(\cdot)(\partial/\partial q^j_\beta)$, where $X^j_\beta(\cdot) \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]};A_\beta)$, via

$$\mathscr{T}_x(X_x) = \sum_{\gamma\in\Gamma}\sum_{\beta\in\Gamma}\sum_{k=1}^{n_\gamma}\sum_{j=1}^{n_\beta} X^k_\gamma(x)\, {}^t k^j_{\gamma\beta}(x)\frac{\partial}{\partial q^j_\beta} \tag{34}$$

to generalized vector fields, where ${}^t k^j_{\gamma\beta}(\cdot) \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}; A_{-\gamma+\beta})$, so that the mutated physical one-form $\omega^{[\mathscr{T}]}_1 \in C^\infty_{A_0}(\mathbb{Q}, \mathscr{L}_{A_0}(\mathbb{Q};A_0))$ defined by

$$\omega^{[\mathscr{T}]}_{1_x}(\underline{y}) := \int_{t_a}^{t_b} dt \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} (\mathscr{T}_{x(t)}(y(t)))^j_\gamma \times \mathscr{E}_{j,-\gamma}(x(t),\dot{x}(t)), \tag{35}$$

which of course describes the same physical evolutions as ω_1 does (because of the presupposed invertibility of $\mathscr{T}$), remains generalized variationally self-adjoint, i.e.,

$$d_{Q_\square}[\omega_1]_{Q_\square} = 0 \tag{36}$$

and

$$d_{Q_\square}[\omega^{[\mathscr{T}]}_1]_{Q_\square} = 0. \tag{37}$$

According to Theorem 2, we observe that besides

$$\omega_{1_x}(\underline{y}) = \int_{t_a}^{t_b} dt\, [(\omega_2)_{x(t)}(y(t),\dot{x}(t)) - (\omega_1)_{x(t)}(y(t))], \tag{38}$$

$d\omega_1 = 0$, $d\omega_2 = 0$,

the following conditions for

$$\omega^{[\mathscr{T}]}_{1_x}(\underline{y}) = \int_{t_a}^{t_b} dt\, [(\omega_2)_{x(t)}(\mathscr{T}_{x(t)}(y(t)),\dot{x}(t)) - (\omega_1)_{x(t)}(\mathscr{T}_{x(t)}(y(t)))] \tag{39}$$

must hold:

$$-\omega_2(\mathscr{T}(X),Y) = \omega_2(\mathscr{T}(Y),X),$$
$$d(\omega_2\circ\mathscr{T}) = 0, \quad d(\omega_1\circ\mathscr{T}) = 0. \tag{40}$$

This means that we ask if there are generalized Hamiltonian systems (i.e., $d\omega_2 = 0$, $d\omega_1 = 0$, ω_2^{-1} exists), which remain generalized Hamiltonian after applying a generalized Santilli's integrating tensor $\mathscr{T}$ to the set of the original generalized Hamiltonian evolution terms. As in the pure bosonic case,[19] we shall try to construct for a restricted class of generalized Hamiltonian systems characterized by Γ-graded integrability, not only one integrating tensor $\mathscr{T}$ but a λ-parametric class of tensors $\mathscr{T} = \exp(\lambda\mathscr{N})$, $\lambda\in A_0$, where $\mathscr{N}$ is a Γ-graded Nijenhuis tensor.

Definition 9: A generalized tensor field

$$\mathscr{N} \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}; \mathscr{L}_{A_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{G}^{[n_\Gamma]})),$$

which maps generalized vector fields

$$X = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} X^j_\gamma(\cdot)\frac{\partial}{\partial q^j_\gamma} \in \mathfrak{X}^{A_0}_0(\mathbb{G}^{[n_\Gamma]}),$$

via

$$\mathscr{N}_x(X_x) = \sum_{\gamma\in\Gamma}\sum_{\beta\in\Gamma}\sum_{j=1}^{n_\gamma}\sum_{k=1}^{n_\beta} X^j_\gamma(x)\mathscr{N}^{jk}_{\gamma\beta}(x)\frac{\partial}{\partial q^k_\beta},$$

$$\mathscr{N}^{jk}_{\gamma\beta}(\cdot) \in C^\infty_{A_0}(\mathbb{G}^{[n_\Gamma]}; A_{-\gamma+\beta}),$$

to generalized vector fields $\mathscr{N}(X) \in \mathfrak{X}^{A_0}_0(\mathbb{G}^{[n_\Gamma]})$ and which obeys

$$\mathbf{L}_{\mathscr{N}(X)}\mathscr{N}(Y) = \mathscr{N}(\mathbf{L}_X\mathscr{N}(Y) - \mathbf{L}_Y\mathscr{N}(X)) - \mathscr{N}^2\mathbf{L}_X(Y), \tag{41}$$

$$\forall X = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} X^j_\gamma(\cdot)\frac{\partial}{\partial q^j_\gamma},$$

$$Y = \sum_{\gamma\in\Gamma}\sum_{j=1}^{n_\gamma} Y^j_\gamma(\cdot)\frac{\partial}{\partial q^j_\gamma} \in \mathfrak{X}^{A_0}_0(\mathbb{G}^{[n_\Gamma]})$$

is called a Γ-graded Nijenhuis tensor.

We emphasize that the bosonic Nijenhuis tensors play an important role in the theory of nonlinear integrable evolu-

tion equations.[6–14] As in the pure bosonic case,[7,8] one shows that each formal power series $F(\mathcal{N}) = \Sigma_{n=0}^{\infty} a_n (\mathcal{N})^n$ with $a_n \in \mathbf{A}_0$ remains a formal Γ-graded Nijenhuis tensor. We shall later take the well-defined function

$$\exp(\lambda\mathcal{N}) := \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \mathcal{N}^n, \quad \lambda \in \mathbf{A}_0,$$

also because of the obvious invertibility of $\exp(\lambda\mathcal{N})$ via $\exp(-\lambda\mathcal{N}) = (\exp(\lambda\mathcal{N}))^{-1}$. If $\mathcal{N}$ is a Γ-graded Nijenhuis tensor, the following formula can be proved by induction:

$$\begin{aligned} &\mathbf{L}_{\mathcal{N}(X)}(\mathcal{N}^n(Y)) - \mathcal{N}(\mathbf{L}_X(\mathcal{N}^n(Y))) \\ &\quad = \mathcal{N}[\mathbf{L}_{\mathcal{N}(X)}(\mathcal{N}^{n-1}(Y)) - \mathcal{N}(\mathbf{L}_X(\mathcal{N}^{n-1}(Y)))], \quad (42) \\ &\qquad \forall X, Y \in \mathfrak{X}_0^{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}), \quad n \in \mathbf{N}, \end{aligned}$$

from which

$$\begin{aligned} &\mathbf{L}_{\mathcal{N}(X)}(\mathcal{N}^n(Y)) - \mathcal{N}(\mathbf{L}_X(\mathcal{N}^n(Y))) \\ &\quad = \mathcal{N}^n(\mathbf{L}_{\mathcal{N}(X)}(Y) - \mathcal{N}(\mathbf{L}_X Y)) \quad (43) \end{aligned}$$

follows.

Proposition 10: (Compare this proposition to Magri[6] in the pure bosonic case.) Let X be a generalized vector field $X = \Sigma X_\gamma^j(\cdot)(\partial/\partial q_\gamma^j) \in \mathfrak{X}_0^{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]})$; $X_\gamma^j(\cdot) \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}; \mathbf{A}_\gamma)$ with the property

$$\mathbf{L}_X \mathcal{N} = 0, \quad (44)$$

where $\mathcal{N}$ is a Γ-graded Nijenhuis tensor, then

$$\mathbf{L}_{\mathcal{N}^n(X)} \mathcal{N} = 0, \quad \forall n \in \mathbf{N}, \quad (45)$$

is valid.

Proof:

$$\begin{aligned} (\mathbf{L}_{\mathcal{N}^n(X)}\mathcal{N})(Y) &= \mathbf{L}_{\mathcal{N}^n(X)}(\mathcal{N}(Y)) - \mathcal{N}(\mathbf{L}_{\mathcal{N}^n(X)} Y) \\ &\overset{(43)}{=} \mathcal{N}^n(\mathbf{L}_X(\mathcal{N}(Y)) - \mathcal{N}(\mathbf{L}_X Y)) \\ &= \mathcal{N}^n((\mathbf{L}_X\mathcal{N})(Y)) = 0. \quad \blacksquare \end{aligned}$$

Proposition 11: (Compare this to Magri[6] in the pure bosonic case.) Let $\mathcal{N}$ be a Γ-graded Nijenhuis tensor and let $X \in \mathfrak{X}_0^{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]})$ be a generalized vector field which leaves $\mathcal{N}$ invariant, i.e.,

$$\mathbf{L}_X \mathcal{N} = 0,$$

then, the following formula is valid:

$$\forall n,p \in \mathbf{N}_0: \quad [\mathcal{N}^n(X), \quad \mathcal{N}^p(X)]_- = 0, \quad (46)$$

i.e., all generalized vector fields $\mathcal{N}^n(X)$ are involution and therefore lead to commutable Γ-graded flows.

The proof is evident by using (45):

$$\begin{aligned} &[\mathcal{N}^n(X), \quad \mathcal{N}^p(X)]_- \\ &\quad = \mathbf{L}_{\mathcal{N}^n(X)} \mathcal{N}^p(X) \\ &\quad = (\mathbf{L}_{\mathcal{N}^n(X)}\mathcal{N})\mathcal{N}^{p-1}(X) + \cdots + \mathcal{N}^{p-1}(\mathbf{L}_{\mathcal{N}^n(X)}\mathcal{N})(X) \\ &\qquad + \mathcal{N}^p(\mathbf{L}_{\mathcal{N}^n(X)} X) \\ &\overset{(45)}{=} \mathcal{N}^p(\mathbf{L}_{\mathcal{N}^n(X)} X) = -\mathcal{N}^p(\mathbf{L}_X \mathcal{N}^n(X)) = 0. \quad \blacksquare \end{aligned}$$

Definition 12: The triple $(\omega_2, \nu_0, \mathcal{N})$ is called a generalized Hamiltonian system with compatible Γ-graded Nijenhuis tensor $\mathcal{N}$ if the following formulas:

$$d(\omega_1 \circ \mathcal{N}) = 0, \quad (47)$$

$$\omega_2(\mathcal{N}(X), Y) = \omega_2(X, \mathcal{N}(Y)), \quad (48)$$

i.e.,

$$\begin{aligned} &\omega_2 \circ \mathcal{N} \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}^{\text{anti}}((\mathbb{G}^{[n_\Gamma]})^2; \mathbf{A}_0)), \\ &d(\omega_2 \circ \mathcal{N}) = 0, \quad (49) \\ &\omega_1 = d\nu_0, \quad \omega_2 = d\nu_1 \quad (50) \end{aligned}$$

are valid, and if ω_2^{-1} exists.

Theorem 13: (Compare this to Refs. 6, 7, and 54 in the pure bosonic case.) Let $\omega_1 \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}; \mathbf{A}_0))$ and let $\mathcal{N}$ be a Γ-graded Nijenhuis tensor. Let

$$d\omega_1 = 0, \quad d(\omega_1 \circ \mathcal{N}) = 0.$$

Then, we obtain

$$d(\omega_1 \circ \mathcal{N}^n) = 0, \quad \forall n \in \mathbf{N}. \quad (51)$$

Proof: The following identity holds for each

$$\begin{aligned} &\tilde{\omega}_1 \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}; \mathbf{A}_0)), \\ &\tilde{\mathcal{N}} \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}; \mathbb{G}^{[n_\Gamma]})): \\ &d(\tilde{\omega}_1 \circ \tilde{\mathcal{N}}^2)(Y,Z) \\ &\quad = d(\tilde{\omega}_1 \circ \tilde{\mathcal{N}})(Y, \tilde{\mathcal{N}}(Z)) + d(\tilde{\omega}_1 \circ \tilde{\mathcal{N}})(\tilde{\mathcal{N}}(Y), Z) \\ &\qquad - d\tilde{\omega}_1(\tilde{\mathcal{N}}(Y), \tilde{\mathcal{N}}(Z)) \\ &\qquad - \tilde{\omega}_1(\mathbf{L}_{\tilde{\mathcal{N}}(Y)}\tilde{\mathcal{N}}(Z) - \tilde{\mathcal{N}}[L_Y\tilde{\mathcal{N}}(Z) - \mathbf{L}_Z\tilde{\mathcal{N}}(Y)] \\ &\qquad + \tilde{\mathcal{N}}^2 \mathbf{L}_Y Z), \quad (52) \end{aligned}$$

$$\forall Y = \sum Y_\gamma^j(\cdot)\frac{\partial}{\partial q_\gamma^j},$$

$$Z = \sum Z_\gamma^j(\cdot)\frac{\partial}{\partial q_\gamma^j} \in \mathfrak{X}_0^{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}),$$

so that (51) can easily be proved by putting $\tilde{\mathcal{N}} = \mathcal{N}$, $\tilde{\omega}_1 := \omega_1 \circ \mathcal{N}^m$, $m = 0, 1, \ldots$, successively. $\blacksquare$

Corollary 14: The same presuppositions as in the previous theorem lead to

$$d(\omega_1 \circ f(\mathcal{N})) = 0$$

for all formal power series

$$f(\mathcal{N}) = \sum_{n=0}^{\infty} a_n \mathcal{N}^n, \quad a_n \in \mathbf{A}_0.$$

In particular, we obtain

$$d(\omega_1 \circ \exp(\lambda\mathcal{N})) = 0, \quad \forall \lambda \in \mathbf{A}_0. \quad (53)$$

Proposition 15: Let

$$\begin{aligned} &\tilde{\omega}_2 \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}^{\text{anti}}((\mathbb{G}^{[n_\Gamma]})^2; \mathbf{A}_0)), \\ &\tilde{\mathcal{N}} \in \mathbf{C}_{\mathbf{A}_0}^\infty(\mathbb{G}^{[n_\Gamma]}, \mathcal{L}_{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}; \mathbb{G}^{[n_\Gamma]})), \end{aligned}$$

with the property

$$\tilde{\omega}_2(\tilde{\mathcal{N}}(X), Y) = \tilde{\omega}_2(X, \tilde{\mathcal{N}}(Y)), \quad (54)$$

$$\forall X = \sum X_\gamma^j(\cdot)\frac{\partial}{\partial q_\gamma^j},$$

$$Y = \sum Y_\gamma^j(\cdot)\frac{\partial}{\partial q_\gamma^j} \in \mathfrak{X}_0^{\mathbf{A}_0}(\mathbb{G}^{[n_\Gamma]}),$$

then

$$\tilde{\omega}_2(f(\tilde{\mathcal{N}})(X),Y)=\tilde{\omega}_2(X,f(\tilde{\mathcal{N}})(Y))$$

is valid for all formal power series

$$f(\tilde{\mathcal{N}})=\sum a_n\tilde{\mathcal{N}}^n,\quad a_n\in\mathbb{A}_0,$$

especially, we have

$$\tilde{\omega}_2(\exp(\lambda\tilde{\mathcal{N}})(X),Y)=\tilde{\omega}_2(X,\exp(\lambda\tilde{\mathcal{N}})(Y)). \tag{55}$$

In this case the notation $\tilde{\omega}_2\circ f(\tilde{\mathcal{N}})$ does make sense. Let now $\tilde{\omega}_2$ and $\tilde{\mathcal{N}}$ obey the same presuppositions as in the previous proposition. The identity

$$\begin{aligned}
&d(\tilde{\omega}_2\circ\tilde{\mathcal{N}}^2)(X,Y,Z)\\
&\quad=\tfrac{2}{3}[d(\tilde{\omega}_2\circ\tilde{\mathcal{N}})(\tilde{\mathcal{N}}(X),Y,Z)+d(\tilde{\omega}_2\circ\tilde{\mathcal{N}})(X,\tilde{\mathcal{N}}(Y),Z)\\
&\qquad+d(\tilde{\omega}_2\circ\tilde{\mathcal{N}})(X,Y,\tilde{\mathcal{N}}(Z))]\\
&\qquad-\tfrac{1}{3}[d\tilde{\omega}_2(\tilde{\mathcal{N}}(X),\tilde{\mathcal{N}}(Y),Z)+d\tilde{\omega}_2(\tilde{\mathcal{N}}(X),Y,\tilde{\mathcal{N}}(Z))\\
&\qquad+d\tilde{\omega}_2(X,\tilde{\mathcal{N}}(Y),\tilde{\mathcal{N}}(Z))]\\
&\qquad-\tfrac{1}{3}\tilde{\omega}_2(\mathbf{L}_{\tilde{\mathcal{N}}(X)}\tilde{\mathcal{N}}(Y)-\tilde{\mathcal{N}}[\mathbf{L}_X\tilde{\mathcal{N}}(Y)-\mathbf{L}_Y\tilde{\mathcal{N}}(X)]\\
&\qquad+\tilde{\mathcal{N}}^2\mathbf{L}_XY,Z)\\
&\qquad+\tfrac{1}{3}\tilde{\omega}_2(\mathbf{L}_{\tilde{\mathcal{N}}(X)}\tilde{\mathcal{N}}(Z)-\tilde{\mathcal{N}}[\mathbf{L}_X\tilde{\mathcal{N}}(Z)-\mathbf{L}_Z\tilde{\mathcal{N}}(X)]\\
&\qquad+\tilde{\mathcal{N}}^2\mathbf{L}_XZ,Y)\\
&\qquad-\tfrac{1}{3}\tilde{\omega}_2(\mathbf{L}_{\tilde{\mathcal{N}}(Y)}\tilde{\mathcal{N}}(Z)-\tilde{\mathcal{N}}[\mathbf{L}_Y\tilde{\mathcal{N}}(Z)-\mathbf{L}_Z\tilde{\mathcal{N}}(Y)]\\
&\qquad+\tilde{\mathcal{N}}^2\mathbf{L}_YZ,X)
\end{aligned}\tag{56}$$

permits us to prove the following theorem.

Theorem 16: Let

$$\omega_2\in\mathsf{C}^\infty_{\mathbb{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathscr{L}^{\mathrm{anti}}_{\mathbb{A}_0}((\mathbb{G}^{[n_\Gamma]})^2;\mathbb{A}_0)),$$

$$\mathcal{N}\in\mathsf{C}^\infty_{\mathbb{A}_0}(\mathbb{G}^{[n_\Gamma]},\mathscr{L}_{\mathbb{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{G}^{[n_\Gamma]}))$$

obey

$$d\omega_2=0,\quad \omega_2(\mathcal{N}(X),Y)=\omega_2(X,\mathcal{N}(Y)),$$

$$d(\omega_2\circ\mathcal{N})=0,$$

and let $\mathcal{N}$ be a Γ-graded Nijenhuis tensor, then

$$d(\omega_2\circ f(\mathcal{N}))=0, \tag{57}$$

for all formal power series

$$f(\mathcal{N})=\sum a_n\mathcal{N}^n,\quad a_n\in\mathbb{A}_0.$$

In particular, we arrive at the formula

$$d(\omega_2\circ\exp(\lambda\mathcal{N}))=0,\quad \forall\lambda\in\mathbb{A}_0. \tag{58}$$

We point out that formula (56) is the Γ-graded counterpart of the corresponding formula (in the pure bosonic case) 5.10, p. 352, in the article by Frölicher and Nijenhuis,[54] putting $\mathbf{L}=\mathbf{M}=\tilde{\mathcal{N}}$, $l=m=1$, and taking into account

$$\tilde{\mathcal{N}}\wedge\tilde{\mathcal{N}}=\tilde{\mathcal{N}}^2,\quad \tilde{\omega}_2\wedge\tilde{\mathcal{N}}=2\tilde{\omega}_2\circ\tilde{\mathcal{N}},$$

$$(\tilde{\omega}_2\wedge\tilde{\mathcal{N}})\wedge\tilde{\mathcal{N}}=4\tilde{\omega}_2\circ\tilde{\mathcal{N}}^2,$$

$$\begin{aligned}
&[\tilde{\mathcal{N}},\tilde{\mathcal{N}}](X,Y)\\
&\quad=2[\mathbf{L}_{\tilde{\mathcal{N}}(X)}\tilde{\mathcal{N}}(Y)-\tilde{\mathcal{N}}(L_X\tilde{\mathcal{N}}(Y)-\mathbf{L}_Y\tilde{\mathcal{N}}(X))\\
&\qquad+\tilde{\mathcal{N}}^2\mathbf{L}_XY].
\end{aligned}$$

If we now collect all our results, we arrive at the following theorem.

Theorem 17: Let now $(\omega_2,\nu_0,\mathcal{N})$ constitute a generalized Hamiltonian system with compatible Γ-graded Nijenhuis tensor $\mathcal{N}$ (Definition 12), then

$$d(\omega_1\circ\exp(\lambda\mathcal{N}))=0,\quad d(\omega_2\circ\exp(\lambda\mathcal{N}))=0, \tag{59}$$

are valid, $\forall\lambda\in\mathbb{A}_0$. Consequently, the generalized tensor fields

$$\mathcal{T}:=\exp(\lambda\mathcal{N}) \tag{60}$$

constitute a λ-parametric class of generalized Santilli's isotopic operators, this means, that the mutated one-forms $\omega_1^{[\mathcal{T}]}\in\mathsf{C}^\infty_{\mathbb{A}_0}(\mathbb{Q},\mathscr{L}_{\mathbb{A}_0}(\mathbb{Q};\mathbb{A}_0))$ defined by (35) remain generalized variationally self-adjoint and of Hamiltonian type. The corresponding action functionals $\omega_0^{[\mathcal{T}]}$ read

$$\begin{aligned}
\omega_{0_x}^{[\mathcal{T}]}&=K(\omega_1^{[\mathcal{T}]})_x\\
&=\int_{t_a}^{t_b}dt\,[K(\omega_2\circ\exp(\lambda\mathcal{N}))_{x(t)}(\dot{x}(t))\\
&\qquad-K(\omega_1\circ\exp(\lambda\mathcal{N}))_{x(t)}],
\end{aligned}\tag{61}$$

with a λ-parametric class of generalized symplectic potentials

$$\nu_1^{[\lambda]}:=K(\omega_2\circ\exp(\lambda\mathcal{N})),\quad d\nu_1^{[\lambda]}=\omega_2\circ\exp(\lambda\mathcal{N}), \tag{62}$$

and with a λ-parametric class of generalized Hamiltonians

$$\nu_0^{[\lambda]}:=K(\omega_1\circ\exp(\lambda\mathcal{N})),\quad d\nu_0^{[\lambda]}=\omega_1\circ\exp(\lambda\mathcal{N}). \tag{63}$$

We emphasize that the limit $\lambda\to 0$ in (62) and (63) continuously leads to the initial generalized Hamiltonian ν_0 and to the initial generalized symplectic potential ν_1. We also point out that ω_1 and $\omega_1^{[\mathcal{T}]}$ describe the same physical evolutions, because $\omega_{1_x}=0$ is valid if and only if $\omega_{1_x}^{[\mathcal{T}]}=0$ is valid.

Theorem 18: Let again $(\omega_2,\nu_0,\mathcal{N})$ constitute a generalized Hamiltonian system with compatible Γ-graded Nijenhuis tensor $\mathcal{N}$ (Definition 12), then

$$\mathsf{L}_{(d\nu_0)^\#}(\exp(\lambda\mathcal{N}))=0, \tag{64}$$

$\forall\lambda\in\mathbf{A}_0$, is valid, where $(d\nu_0)^\#=\omega_2^{-1}(d\nu_0,\cdot)$. This means that $(d\nu_0)^\#$ constitutes an infinitesimal generalized canonical transformation, which leaves the Γ-graded Nijenhuis tensor $\mathcal{N}$ invariant. (Compare this result with Ref. 6 in the bosonic case.)

Proof: Because of

$$\begin{aligned}
&(\omega_2\circ\exp(\lambda\mathcal{N}))\,(X,(d\nu_0)^\#)\\
&\quad=\omega_2(\exp(\lambda\mathcal{N})(X),(d\nu_0)^\#)\\
&\quad\overset{(29)}{=}d\nu_0(\exp(\lambda\mathcal{N})(X))\\
&\quad\overset{(50)}{=}(\omega_1\circ\exp(\lambda\mathcal{N}))(X)\\
&\quad\overset{(63)}{=}d\nu_0^{[\lambda]}(X),
\end{aligned}$$

we find

$$di_{(d\nu_0)^\#}(\omega_2\circ\exp(\lambda\mathcal{N}))=0,$$

and because of (59) and (33a) we arrive at

$$\mathsf{L}_{(d\nu_0)^\#}(\omega_2\circ\exp(\lambda\mathcal{N}))=0.$$

Finally, because of Theorem (7) and as ω_2 is supposed to be invertible, we arrive at (64). ■

By virtue of Proposition 11, we obtain

$$[\exp(\lambda\mathcal{N})(d\nu_0)^{\#},\ \exp(\lambda'\mathcal{N})(d\nu_0)^{\#}]_{-}=0, \quad \forall\lambda,\lambda'\in\mathbb{A}_0. \tag{65}$$

Theorem 19: Let again $(\omega_2, \nu_0, \mathcal{N})$ denote a generalized Hamiltonian system with compatible Γ-graded Nijenhuis tensor $\mathcal{N}$, then

$$\exp(\lambda\mathcal{N})(d\nu_0)^{\#}=(d\nu_0^{[\lambda]})^{\#}. \tag{66}$$

This means $\exp(\lambda\mathcal{N})(d\nu_0)^{\#}$ constitutes a λ-parametric class of commutable generalized Hamiltonian vector fields, whose Hamiltonians $\nu_0^{[\lambda]}$ commute under the generalized time evolution bracket (23)

$$0=\{\nu_0^{[\lambda]},\ \nu_0^{[\lambda']}\}=\omega_2^{-1}\,(d\nu_0^{[\lambda]},\ d\nu_0^{[\lambda']}), \quad \forall\lambda,\lambda'\in\mathbb{A}_0. \tag{67}$$

Proof: We observe

$$\begin{aligned}&\omega_2(X,\exp(\lambda\mathcal{N})(d\nu_0)^{\#})\\ &\quad=\omega_2(\exp(\lambda\mathcal{N})(X),(d\nu_0)^{\#})\\ &\quad\overset{(29)}{=}d\nu_0(\exp(\lambda\mathcal{N})(X))\overset{(50)}{=}(\omega_1\circ\exp(\lambda\mathcal{N}))(X)\\ &\quad\overset{(63)}{=}d\nu_0^{[\lambda]}(X)\overset{(29)}{=}\omega_2(X,(d\nu_0^{[\lambda]})^{\#}).\end{aligned}$$

Finally, we consider

$$\begin{aligned}&\{\nu_0^{[\lambda]},\nu_0^{[\lambda']}\}\\ &\quad=(d\nu_0^{[\lambda]})^{\#}(\nu_0^{[\lambda']})\\ &\quad\overset{(29)}{=}\omega_2((d\nu_0^{[\lambda]})^{\#},(d\nu_0^{[\lambda']})^{\#})\\ &\quad=(\omega_2\circ\exp((\lambda+\lambda')\mathcal{N}))((d\nu_0)^{\#},(d\nu_0)^{\#})=0.\end{aligned}$$ ■

Corollary 20: Let again $(\omega_2, \nu_0, \mathcal{N})$ constitute a generalized Hamiltonian system with compatible Γ-graded Nijenhuis tensor $\mathcal{N}$ and let $\nu_{0;k}$ denote

$$\nu_{0;k}:=\frac{d^k}{d\lambda^k}(\nu_0^{[\lambda]})|_{\lambda=0}\in\mathbb{C}^{\infty}_{\mathbb{A}_0}(\mathbb{G}^{[n_\Gamma]};\mathbb{A}_0), \tag{68}$$

then, the system $\{\nu_{0;k}\}_{k=0,1,2,\ldots}$ constitutes a set of generalized conservation quantities with respect to the Hamiltonian ν_0 which are involutive under the generalized time evolution bracket:

$$\{\nu_0,\nu_{0;k}\}=0=\{\nu_{0;k},\nu_{0;p}\},\quad \forall k,p\in\mathbb{N}. \tag{69}$$

We finally stress and underline that the propositions and theorems exposed above constitute a generalization and extension of well-known facts in the pure bosonic calculus (if one, for instance, deals with integrable nonlinear evolution equations[6–14]) to the case if using a color analytic calculus, i.e., if one starts with an associative Γ-graded Banach algebra $\mathbb{A}$ with arbitrary but finite Abelian grading group Γ and equipped with a corresponding Γ-compatible commutation factor σ. Incidentally, we automatically comprise the conventional $\mathbb{Z}_2$-graded supersymmetric case also by choosing $\Gamma=\mathbb{Z}_2=\{0,1\}$, $\sigma(0,0)=\sigma(0,1)=\sigma(1,0)=-\sigma(1,1)=1$ and by taking $\mathbb{A}$ to be the Banach–Grassmann algebra of Rogers[32] and Jadzcyk and Pilch.[34] We point out that even in the conventional $\mathbb{Z}_2$-graded case, the concepts, "$\mathbb{Z}_2$-graded Nijenhuis tensor," "$\mathbb{Z}_2$-graded Santilli's integrating operators," and the corresponding theorems are apparently new. The following example shows that the very remarkable super-KdV (see Ref. 20) recently studied and introduced by Kupershmidt possesses indeed a $\mathbb{Z}_2$-graded Nijenhuis tensor being compatible with the $\mathbb{Z}_2$-graded Hamiltonian structure.

VI. $\mathbb{Z}_2$-GRADED NIJENHUIS TENSOR AND KUPERSHMIDT'S SUPER-KdV

As in the following, the example considered is a field theory, we now formally extend the results of Secs. II–V to the infinite-dimensional case. This means that we do not take the space $\mathbb{G}^{[n_\Gamma]}$ as configuration space but instead of $\mathbb{G}^{[n_\Gamma]}$ we necessitate the Schwarz space $\mathbb{H}^{[n_\Gamma]}$ consisting of all C^∞ maps from $\mathbb{R}^M$ to $\mathbb{G}^{[n_\Gamma]}$, which vanish with all their derivatives at infinity more rapidly than any power of $\|\vec{r}\|^{-1}$, $\vec{r}\in\mathbb{R}^M$.

We confess that the formal substitution of $\mathbb{G}^{[n_\Gamma]}$ by $\mathbb{H}^{[n_\Gamma]}$, the substitution of $\Sigma_{\gamma,i}$ by $\Sigma_{\gamma,i}\int d^M\vec{r}$, and the substitution of partial left derivatives $\partial/\partial q^i_\gamma$ by partial variational left derivatives $\delta/\delta q^i_\gamma(\vec{r})$ in all previous formulas in order to transcribe the results from discrete systems to field theories is a daring mathematical procedure, in particular, from the topological viewpoint, because of some subtle problems which arise if one deals with an analysis on non-normable vector spaces.[55–57] Because of the validity of the symmetry rule (see Keller[55]), however, we are not *algebraically* hindered to perform the transcription indicated above.

Section VI is therefore exclusively devoted to the applications of the algebraic part of the corresponding transcription of results and formulas contained in Secs. II–V to the field theoretical case. Let now $\mathbb{A}$ be the Banach–Grassmann algebra $\mathbb{B}$ (see Rogers[32] and Jadzcyk and Pilch[34])

$$\mathbb{B}=\mathbb{B}_0\oplus\mathbb{B}_1 \tag{70}$$

which is $\mathbb{Z}_2$ graded. The corresponding $\mathbb{Z}_2$-compatible commutation factor σ obeys $\sigma(0,0)=\sigma(0,1)=\sigma(1,0)=-\sigma(1,1)=1$. Let us consider the $\mathbb{B}_0$ module $\mathbb{G}^{[1,1]}:=\mathbb{B}_0^1\oplus\mathbb{B}_1^1$ and the corresponding Schwarz space $\mathbb{H}^{[1,1]}$ over $\mathbb{R}$. Elements of $\mathbb{H}^{[1,1]}$ will be denoted by

$$\underline{x}=(u(r),\varphi(r))_{r\in\mathbb{R}},\quad \underline{y}=(y(r),\eta(r))_{r\in\mathbb{R}},\quad \underline{z}=(z(r),\zeta(r))_{r\in\mathbb{R}}\in\mathbb{H}^{[1,1]}. \tag{71}$$

In what follows, we also need the vector space $\mathbb{Q}$ consisting of all maps $\underline{x}$, $\underline{y}$ from $\mathbb{R}$ to $\mathbb{H}^{[1,1]}$

$$\mathbb{Q}\ni\underline{x}:t\to\underline{x}(t)=(u(t,r),\varphi(t,r))_{r\in\mathbb{R}}\in\mathbb{H}^{[1,1]},\quad \mathbb{Q}\ni\underline{y}:t\to\underline{y}(t)=(y(t,r),\eta(t,r))_{r\in\mathbb{R}}\in\mathbb{H}^{[1,1]}, \tag{72}$$

having in addition the following properties.

(a) The maps $u(\cdot,\cdot)$, $\varphi(\cdot,\cdot)$, $y(\cdot,\cdot)$, $\eta(\cdot,\cdot)$ are C^∞ in all of their arguments.

(b) The maps

$$\frac{\partial^n u(t,\cdot)}{\partial t^n},\quad \frac{\partial^n\varphi(t,\cdot)}{\partial t^n},\quad \frac{\partial^n y(t,\cdot)}{\partial t^n},\quad \frac{\partial^n\eta(t,\cdot)}{\partial t^n},$$

with all their spatial derivatives vanish uniformly with respect to all elements t of a compact neighborhood of $\tilde{t}$ (de-

pending on $\tilde{t}$) at spacelike infinity more rapidly than any power of $|r|^{-1}$, $r\in\mathbf{R}$, for each $\tilde{t}$ and $n\in\mathbf{N}_0$.

According to Kupershmidt,[20] the super-KdV reads

$$u_t(t,r)=\partial_r\left[3u(t,r)^2-u_{rr}(t,r)+3\varphi(t,r)\varphi_r(t,r)\right],$$
$$\varphi_t(t,r)=3u_r(t,r)\varphi(t,r)+6u(t,r)\varphi_r(t,r)-4\varphi_{rrr}(t,r), \tag{73}$$

where

$$u_t(t,r)=\frac{\partial u(t,r)}{\partial t},\quad \varphi_{rrr}(t,r)=\frac{\partial^3\varphi(t,r)}{\partial r^3},\ \text{etc.}$$

Equation (73) can equivalently be rewritten as

$$\left(\int_{-\infty}^{r}-\frac{1}{2}\int_{-\infty}^{\infty}\right)ds\,u_t(t,s)$$
$$=3u(t,r)^2-u_{rr}(t,r)+3\varphi(t,r)\varphi_r(t,r),$$
$$\varphi_t(t,r)=3u_r(t,r)\varphi(t,r)+6u(t,r)\varphi_r(t,r)-4\varphi_{rrr}(t,r). \tag{74}$$

One easily confirms that

$$\omega_{2_x}(\underline{y}_x,\underline{z}_x):=\int_{-\infty}^{+\infty}dr\,y_x(r)\left(\int_{-\infty}^{r}-\frac{1}{2}\int_{-\infty}^{+\infty}\right)ds\,z_x(s)$$
$$+\int_{-\infty}^{+\infty}dr\,\eta_x(r)\zeta_x(r) \tag{75}$$

leads to a $\mathbf{Z}_2$-graded symplectic two-form ω_2,

$$\omega_{2_x}(\underline{y}_x,\underline{z}_x)=-\omega_{2_x}(\underline{z}_x,\underline{y}_x),\quad d\omega_2=0, \tag{76}$$

where $\underline{y}_x,\underline{z}_x\in\mathbf{H}^{[1,1]}$ depend infinitely many times functional $\mathbb{B}_0$-differentiable on $\underline{x}\in\mathbf{H}^{[1,1]}$. Instead of $\underline{y}_x,\underline{z}_x$, we write the $\mathbf{Z}_2$-graded vector fields Y_x,Z_x in functional differential geometric notation

$$Y_x=\int_{-\infty}^{+\infty}dr\,y_x(r)\frac{\delta}{\delta u(r)}+\int_{-\infty}^{+\infty}dr\,\eta_x(r)\frac{\delta}{\delta\varphi(r)},$$
$$Z_x=\int_{-\infty}^{+\infty}dr\,z_x(r)\frac{\delta}{\delta u(r)}+\int_{-\infty}^{+\infty}dr\,\zeta_x(r)\frac{\delta}{\delta\varphi(r)}. \tag{77}$$

The following $\mathbf{Z}_2$-graded tensor field $\mathcal{N}$ is introduced

$$\mathcal{N}_x(Y_x):=\int_{-\infty}^{+\infty}dr\Big[4u(r)y_x(r)$$
$$+2u_r(r)\left(\int_{-\infty}^{r}-\frac{1}{2}\int_{-\infty}^{+\infty}\right)ds\,y_x(s)-\partial_{rr}y_x(r)$$
$$+3\varphi(r)\partial_r\eta_x(r)+\varphi_r(r)\eta_x(r)\Big]\frac{\delta}{\delta u(r)}$$
$$+\int_{-\infty}^{+\infty}dr\Big[3\varphi(r)y_x(r)$$
$$+2\varphi_r(r)\left(\int_{-\infty}^{r}-\frac{1}{2}\int_{-\infty}^{+\infty}\right)ds\,y_x(s)$$
$$+4u(r)\eta_x(r)-4\partial_{rr}\eta_x(r)\Big]\frac{\delta}{\delta\varphi(r)}, \tag{78}$$

which maps $\mathbf{Z}_2$-graded vectors fields to $\mathbf{Z}_2$-graded vector fields. A tedious calculation shows that $\mathcal{N}$ is a $\mathbf{Z}_2$-graded Nijenhuis tensor, that means that $\mathcal{N}$ satisfies

$$\mathsf{L}_{\mathcal{N}(Y)}\mathcal{N}(Z)=\mathcal{N}\left[\mathsf{L}_Y\mathcal{N}(Z)-\mathsf{L}_Z\mathcal{N}(Y)\right]-\mathcal{N}^2\mathsf{L}_YZ, \tag{79}$$

for all $\mathbf{Z}_2$-graded vector fields Y,Z [see (77)]. Let us now put (74) in $\mathbf{Z}_2$-graded Hamiltonian form. We find that (74) is equivalent to

$$\omega_{2_{\underline{x}(t)}}(\underline{y},\dot{\underline{x}}(t))=dv_{0_{\underline{x}(t)}}(\underline{y}),\quad \forall\underline{y}\in\mathbf{H}^{[1,1]}, \tag{80}$$

with $\mathbf{Z}_2$-graded Hamiltonian

$$v_{0_x}=\int_{-\infty}^{+\infty}dr\Big[u(r)^3+\frac{1}{2}u_r(r)^2+3u(r)\varphi(r)\varphi_r(r)$$
$$-2\varphi(r)\varphi_{rrr}(r)\Big], \tag{81}$$

according to Kupershmidt.[20] We now want to show that

$$d(dv_0\circ\mathcal{N})=0,\quad d(\omega_2\circ\mathcal{N})=0 \tag{82}$$

are valid.

Proof: We calculate

$$K(dv_0\circ\mathcal{N})_x$$
$$=\int_0^1 d\tau(dv_0\circ\mathcal{N})_{\tau x}(\underline{x})$$
$$=\int_{-\infty}^{+\infty}dr\Big[\frac{5}{2}u(r)^4+5u(r)u_r(r)^2+\frac{1}{2}u_{rr}(r)^2$$
$$+15u(r)^2\varphi(r)\varphi_r(r)+20u(r)\varphi_r(r)\varphi_{rr}(r)$$
$$+8\varphi_{rr}(r)\varphi_{rrr}(r)+15\varphi(r)\varphi_{rr}(r)u_r(r)\Big]$$

(where K is the Poincaré operator) and find that

$$dK(dv_0\circ\mathcal{N})=dv_0\circ\mathcal{N}$$

is valid. A short calculation confirms the validity of

$$\omega_2(\mathcal{N}(Y),Z)=\omega_2(Y,\mathcal{N}(Z)),$$

for all $\mathbf{Z}_2$-graded vector fields Y,Z [see (77)]. We then calculate

$$K(\omega_2\circ\mathcal{N})_x(\underline{z})$$
$$=\int_0^1 d\tau\,\tau(\omega_2\circ\mathcal{N})_{\tau x}(\underline{x},\underline{z})$$
$$=\int_{-\infty}^{+\infty}dr\Big[\Big(\frac{4}{3}u(r)^2+\frac{2}{3}u_r(r)\partial_r^{-1}u-\frac{2}{3}\varphi_r(r)\varphi(r)\Big)$$
$$\times\partial_r^{-1}z-\frac{1}{2}u(r)z_r(r)+\frac{2}{3}\varphi_r(r)\partial_r^{-1}u\zeta(r)$$
$$+\frac{7}{3}u(r)\varphi(r)\zeta(r)-2\varphi(r)\zeta_{rr}(r)\Big],$$

where we have abbreviated

$$\partial_r^{-1}u:=\left(\int_{-\infty}^{r}-\frac{1}{2}\int_{-\infty}^{+\infty}\right)ds\,u(s),$$

etc. We again find that

$$dK(\omega_2 \circ \mathcal{N}) = \omega_2 \circ \mathcal{N}$$

is valid, so that (82) is proved. ■

According to Theorem 17, we know that $\exp(\lambda \mathcal{N})$ is a $\mathbf{Z}_2$-graded integrating operator, thus leading to a $\mathbf{Z}_2$-graded isotopic degree of freedom within the $\mathbf{Z}_2$-graded Hamiltonian description of Kupershmidt's super-KdV. The λ-parametric class of the $\mathbf{Z}_2$-graded symplectic potentials reads

$$(\nu_1^{[\lambda]})_{\underline{x}} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int_0^1 d\tau\, \tau((\omega_2)_{\tau\underline{x}} \circ \mathcal{N}_{\tau\underline{x}}^n)(\underline{x}) \tag{83}$$

and the corresponding λ-parametric class of the $\mathbf{Z}_2$-graded Hamiltonians reads

$$(\nu_0^{[\lambda]})_{\underline{x}} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int_0^1 d\tau((d\nu_0)_{\tau\underline{x}} \circ \mathcal{N}_{\tau\underline{x}}^n)(\underline{x}) \tag{84}$$

being continuously deformable to ν_1 and ν_0 via $\lambda \to 0$. According to Theorem 20, we know that the

$$(\nu_{0;n})_{\underline{x}} = \int_0^1 d\tau((d\nu_0)_{\tau\underline{x}} \circ \mathcal{N}_{\tau\underline{x}}^n)(\underline{x}) \tag{85}$$

commute under the generalized time evolution bracket, which in our case reads

$$\begin{aligned}\{F_0, G_0\} &= \omega_2^{-1}(dF_0, dG_0) \\ &= \int_{-\infty}^{+\infty} dr\, \partial_r \left(\frac{\delta F_0}{\delta u(r)}\right)\left(\frac{\delta G_0}{\delta u(r)}\right) \\ &\quad + \int_{-\infty}^{+\infty} dr \left(\frac{\delta F_0}{\delta \varphi(r)}\right)\left(\frac{\delta G_0}{\delta \varphi(r)}\right),\end{aligned} \tag{86}$$

for suitable infinitely many times formal functional $\mathbb{B}_0$-differentiable and $\mathbb{B}_0$-valued functionals F_0, G_0 on $\mathbf{H}^{[1,1]}$. Besides the original Hamiltonian $\nu_0 = \nu_{0;0}$, the next conservation quantity $\nu_{0;1}$ reads

$$\begin{aligned}(\nu_{0;1})_{\underline{x}} = \int_{-\infty}^{+\infty} dr \Big[&\frac{5}{2} u(r)^4 + 5u(r)u_r(r)^2 + \frac{1}{2} u_{rr}(r)^2 \\ &+ 15u(r)^2 \varphi(r)\varphi_r(r) + 20u(r)\varphi_r(r)\varphi_{rr}(r) \\ &+ 8\varphi_{rr}(r)\varphi_{rrr}(r) + 15\varphi(r)\varphi_{rr}(r)u_r(r)\Big].\end{aligned} \tag{87}$$

VII. CONCLUDING REMARKS

Using the calculus with color numbers, i.e., with variables having unusual commutation properties, we have presented the initiating steps towards the construction of a theory of Γ-graded integrable evolution equations under the viewpoint of generalized Hamiltonian systems compatible with Γ-graded Nijenhuis operators. As an example of our theory, we have found that Kupershmidt's super-KdV fits in our scheme by presenting a $\mathbf{Z}_2$-graded Nijenhuis tensor compatible with the $\mathbf{Z}_2$-graded Hamiltonian structure of the super KdV. The explicit construction of Γ-graded integrable evolution equations with grading groups Γ different from $\mathbf{Z}_1$, $\mathbf{Z}_2$ is under current investigation and will be published elsewhere.

ACKNOWLEDGMENTS

It is a sincere pleasure to thank Professor Dr. R. M. Santilli for the kind invitation to the Second Workshop on Hadronic Mechanics, held at Como, Italy in 1984 and for fruitful suggestions via correspondence. I am delighted to thank Professor Dr. B. Fuchssteiner and Dr. W. Oevel for stimulating discussions and for a kind invitation. I also thank Professor Dr. H. Bucka, Professor Dr. K-E. Hellwig, Professor Dr. S. Hess, Professor Dr. W. Muschik, Professor Dr. U. Scherz, and my father Professor Dr. Rudolf Trostel for kind help and warm encouragement. Kind comments of the referee are as well gratefully acknowledged.

[1] R. K. Bullough, P. J. Caudrey, G. L. Lamb, Jr., D. W. McLaughlin, H. M. Gibbs, M. Toda, R. Hirota, A. C. Newell, V. E. Zakharov, M. Wadati, F. Calogero, A. Degasperis, S. P. Novikov, L. D. Faddeev, and A. H. Luther, *Solitons, Topics in Current Physics* (Springer, Berlin, 1980).

[2] R. M. Miura, H. E. Moses, R. Hermann, A. R. Bishop, H. C. Yuen, B. M. Lake, K. E. Lonngren, H. Ikezi, R. D. Parmentier, D. W. McLaughlin, A. C. Scott, J. H. Batteh, J. D. Powell, G. S. Deem, and N. J. Zabusky, *Solitons in Action* (Academic, New York, 1978).

[3] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. **19**, 1095 (1967).

[4] B. A. Kupershmidt, J. Math. Phys. **22**, 449 (1981).

[5] B. Fuchssteiner, Prog. Theor. Phys. **68**, 1082 (1982).

[6] F. Magri, *Lecture Notes in Physics*, Vol. 120 (Springer, Berlin, 1980), pp. 233–263.

[7] I. M. Gel'fand and I. Ya. Dorfman, Functional Anal. Appl. **13**, 248 (1979); **14**, 223 (1980).

[8] B. Fuchssteiner, Prog. Theor. Phys. **65**, 861 (1981).

[9] B. Fuchssteiner, Lett. Math. Phys. **4**, 177 (1980).

[10] B. Fuchssteiner, Nonlinear Anal. Theor. Meth. Appl. **3**, 849 (1979).

[11] A. S. Fokas, J. Math. Phys. **21**, 1318 (1980).

[12] B. Fuchssteiner and A. S. Fokas, Physica D **4**, 47 (1981).

[13] A. S. Fokas and B. Fuchssteiner, Nonlinear Anal. Theor. Meth. Appl. **5**, 423 (1981).

[14] B. Fuchssteiner and W. Oevel, J. Math. Phys. **23**, 358 (1982).

[15] P. J. Olver, J. Math. Phys. **18**, 1212 (1977).

[16] R. M. Santilli, *Foundations of Theoretical Mechanics*, Vols. I and II (Springer, Berlin, 1978 and 1983), p. 20 and pp. 32, 35, respectively.

[17] R. M. Santilli, Ann. Phys. **103**, 354, 409 (1977); **105**, 227 (1977).

[18] R. M. Santilli, Hadronic J. **1**, 223, 574 (1978).

[19] R. Trostel, Hadronic J. **6**, 650 (1983).

[20] B. A. Kupershmidt, Phys. Lett. A **102**, 213 (1984).

[21] R. Trostel, J. Math. Phys. **25**, 3183 (1984).

[22] J. L. Martin, Proc. R. Soc. London Ser. A **251**, 536 (1959).

[23] D. V. Volkov and V. P. Akulov, Phys. Lett. B **46**, 109 (1973).

[24] J. Wess and B. Zumino, Phys. Lett. B **66**, 361 (1977).

[25] A. Salam and J. Strathdee, Nucl. Phys. B **76**, 477 (1974).

[26] F. A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966).

[27] P. Ramond, Phys. Rev. D **3**, 2415 (1971).

[28] R. Casalbuoni, Nuovo Cimento A **33**, 115, 389 (1976).

[29] A. Neveu and J. H. Schwarz, Nucl. Phys. B **31**, 86 (1971).

[30] B. Kostant, *Lecture Notes in Mathematics*, Vol. 570 (Springer, Berlin, 1977), p. 177.

[31] L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).

[32] A. Rogers, J. Math. Phys. **21**, 1352 (1980).

[33] J. Dell and L. Smolin, Commun. Math. Phys. **66**, 197 (1979).

[34] A. Jadczyk and K. Pilch, Commun. Math. Phys. **78**, 373 (1981).

[35] V. G. Kac, Adv. Math. **26**, 8 (1977).

[36] B. de Witt, *Supermanifolds, Cambridge Monographs on Mathematical Physics* (Cambridge U. P., London, 1984).

[37] Y. Ohnuki and S. Kamefuchi, *Quantum Field Theory and Parastatistics* (Springer, Berlin, 1982).

[38] Y. Ohnuki and S. Kamefuchi, Physica A **114**, 416 (1982).

[39]Y. Ohnuki and S. Kamefuchi, Nuovo Cimento A **70**, 435 (1982).
[40]M. Omote and S. Kamefuchi, *Lecture Notes in Physics*, Vol. 94 (Springer, Berlin, 1979), p. 452.
[41]R. Y. Levine and Y. Tomozawa, Ann. Phys. **149**, 457 (1983).
[42]R. Y. Levine and Y. Tomozawa, Phys. Rev. D **28**, 1358 (1983); J. Math. Phys. **24**, 1603 (1983).
[43]V. Rittenberg and D. Wyler, Nucl. Phys. B **139**, 189 (1978).
[44]V. Rittenberg and D. Wyler, J. Math. Phys. **19**, 2193 (1978).
[45]J. Lukierski and V. Rittenberg, Phys. Rev. D **18**, 385 (1978).
[46]V. K. Agrawala, Hadronic J. **4**, 444 (1981).
[47]H. S. Green and P. D. Jarvis, J. Math. Phys. **24**, 1681 (1983).
[48]M. Scheunert, J. Math. Phys. **24**, 2658, 2671, 2681 (1983).
[49]R. Trostel, Hadronic J. **6**, 1518 (1983).
[50]Y. Kobayashi and S. Nagamachi, J. Math. Phys. **25**, 3367 (1984).
[51]R. M. Santilli (private communication).
[52]R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, Tensor Analysis, and Applications* (Addison–Wesley, Reading, MA, 1983), pp. 188 and 193.
[53]R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, New York, 1978), 2nd ed., p. 189.
[54]A. Frölicher and A. Nijenhuis, Indag. Math. **18**, 338 (1956).
[55]H. H. Keller, *Lecture Notes in Mathematics*, Vol. 417 (Springer, Berlin, 1974), p. 90.
[56]S. Yamamuro, *Lecture Notes in Mathematics*, Vol. 374 (Springer, Berlin, 1974).
[57]A. Frölicher and W. Bucher, *Lecture Notes in Mathematics*, Vol. 30 (Springer, Berlin, 1966).

Spectrum doubling and double-valuedness

Richard Stacey
Department of Mathematics, Birmingham University, P. O. Box 363, Birmingham B15 2TT, England

(Received 27 March 1985; accepted for publication 7 June 1985)

The effect of making the lattice Dirac operator the square root of the Laplacian is investigated. The doubling of the fermion spectrum is then matched by that of the boson spectrum, unless bosons are restricted to a double-spaced sublattice. Fermion spectrum doubling is found to be a necessary consequence of the double-valuedness of the square root.

I. INTRODUCTION

In the continuum the Dirac operator $\eth$ and the Laplacian $\Box$ are related by

$$\eth\cdot\eth = \Box. \tag{1}$$

Equation (1) can be formally inverted:

$$\eth = \sqrt{\Box}. \tag{2}$$

The Dirac operator can therefore be described or defined as the "square root" of the Laplacian. In this paper we will examine the effect of enforcing Eqs. (1) and (2) on a lattice. We hope in this way to cast light on the lattice fermion doubling problem.[1]

First (Sec. II) we start from the naive lattice fermion prescription with its doubled (2^d-fold) spectrum. Enforcing Eq. (1) we get a lattice boson prescription with an identical 2^d doubling (Sec. III). To eliminate boson doubling we have to restrict the boson fields to a double-spaced sublattice.

In Sec. IV we start from a standard undoubled lattice Laplacian and use Eq. (2) to define a lattice Dirac operator. The resulting theory has an undoubled fermion spectrum. Unfortunately, it is ill defined, with multiple sign ambiguities. Now these are inevitable when we take square roots, and occur in the continuum also. The crucial difference is that on a lattice we cannot specify a single set of signs consistently. To get a well-defined theory we are forced to introduce fermion spectrum doubling. We end up with the theory and conclusions of Sec. III all over again.

In Sec. VI we look for a deeper understanding. We observe that difficulties identical to those found in Sec. IV beset the definition of the square-root function on $\mathbb{C}$. Now this analog has a known mathematical solution, in the theory of Riemann surfaces.[2] Fermion spectrum doubling emerges from our analysis as a natural and inevitable consequence of the double-valuedness of the square root. This result holds in the continuum also (Sec. VII). Perhaps the true resolution of the lattice's problems lies in a better understanding of the continuum theory. In Sec. VIII we comment on our assumptions and state our conclusions.

II. LATTICE FERMION DOUBLING

The lattice fermion doubling problem is seen at its simplest on a cubic lattice with the derivative prescription

$$\frac{\partial\psi}{\partial x_\mu} \to \frac{1}{2a}\{\psi(x + a\hat{\mu}) - \psi(x - a\hat{\mu})\}. \tag{3}$$

The Fourier transform of Eq. (3) is

$$p_\mu\tilde{\psi}(p) \to [\sin(p_\mu a)/a]\,\tilde{\psi}(p). \tag{4}$$

An arbitrary translation-invariant derivative prescription is characterized by the "lattice momentum" $D_\mu(p)$ that replaces p_μ in this transform. For instance, the lattice Feynman rules involve $\not{D}(p)$ rather than $\not{p}$. In our example,

$$D_\mu(p) \equiv a^{-1}\sin(p_\mu a). \tag{5}$$

Spectrum doubling arises because $D_\mu(p)$ has two zeros in the Brillouin zone (defined as any $2\pi/a$ range for p_μ). There are in all 2^d energy zeros giving 2^d fermions, with equal numbers of each handedness.

A good way of seeing the 2^d degeneracy is to solve

$$\frac{\partial\psi}{\partial x_\mu} = 0, \quad \forall\mu. \tag{6}$$

Because of the $2a$ spacing over which $\partial\psi/\partial x_\mu$ is measured [Eq. (3)] there are 2^d double-spaced sublattices, on each of which ψ is constant, but the constants are independent. In the continuum theory there is a single constant, and one fermion field.

We cannot eliminate doubling by restricting attention to a single sublattice. Observe that, since handedness is given by

$$\operatorname{sgn}\left(\prod_{i=1}^{d}\frac{\partial D_i(p)}{\partial p_i}\right),$$

a transformation

$$p_j \to p_j + \pi/a, \quad p_k \to p_k \quad (k \neq j) \tag{7}$$

exchanges left- and right-handed components. The subfields defined on each of the double-spaced sublattices [e.g. $\Sigma\exp\{i\Sigma_j p_j(2n_j a)\}\,\psi(2n_i, \ldots, 2n_d)$] are either even (our example) or odd under the transformation of Eq. (7). They are therefore "mixed fermion" fields of the form $L \pm R$, with L and R equal up to handedness. Such fields cannot be associated with one handedness or one fermion species.

III. LATTICE BOSON DOUBLING

Theories of bosons involve the second derivative of the field Φ. Using the prescription of Eq. (3) twice, we get

$$\frac{\partial^2\Phi}{\partial x_\mu{}^2} = \frac{1}{2a}\left\{\frac{\partial\Phi}{\partial x_\mu}(x + a\hat{\mu}) - \frac{\partial\Phi}{\partial x_\mu}(x - a\hat{\mu})\right\}$$

$$= \frac{1}{4a^2}\{\Phi(x + 2a\hat{\mu}) + \Phi(x - 2a\hat{\mu}) - 2\Phi(x)\}, \tag{8}$$

 0022-2488/85/123172-04$02.50

which Fourier transforms to

$$p_\mu{}^2\tilde{\Phi}(p) \to [\sin(p_\mu a)/a]^2\tilde{\Phi}(p). \tag{9}$$

The theory is characterized by the "lattice momentum-squared" $D^2(p)$. In this case

$$D^2(p) = \sum_\mu \left(\frac{\sin(p_\mu a)}{a}\right)^2. \tag{10}$$

Clearly, Eqs. (4) and (10), with the usual γ algebra, give the desired relation [Eq. (1)] between δ and $\Box$. Equally clearly, $D^2(p)$ has 2^d zeros in the Brillouin zone, giving exactly the same doubling problem for bosons as for fermions.

Boson spectrum doubling is easily removed. Because of the $\pm 2a\hat{\mu}$ terms in Eq. (8) the 2^d boson fields lie on the 2^d distinct double-spaced sublattices discussed in Sec. II. They do not mix like fermions. If we restrict the boson field and its gauge couplings to one of the sublattices (e.g., $x_i = 2an_i, \forall i$), the doubling disappears.

IV. DOUBLE-VALUEDNESS

The standard undoubled lattice boson prescription replaces Eq. (8) with

$$\frac{\partial^2\Phi}{\partial x_\mu{}^2} \to \frac{1}{a^2}\{\Phi(x + a\hat{\mu}) + \Phi(x - a\hat{\mu}) - 2\Phi(x)\}, \tag{11}$$

leading to a lattice momentum-squared

$$D'^2(p) = \sum_\mu \left[\frac{2}{a}\sin\left(\frac{p_\mu a}{2}\right)\right]^2. \tag{12}$$

Preserving the square-root relation of Eq. (2), we are led to a new lattice momentum

$$D'_\mu(p) = (2/a)\sin(p_\mu a/2). \tag{13}$$

This prescription gives a single energy zero in the Brillouin zone, however the latter is specified.

Unfortunately, $D_\mu'(p)$ is not well defined. It changes sign under

$$p_\mu \to p_\mu + 2\pi/a.$$

This double-valuedness results in a 2^d ambiguity in the specification of the Dirac operator $\not{D}'(p)$.

Such ambiguities are not peculiar to the lattice. In the continuum,

$$\sqrt{p^2} = \pm p_1\gamma_1 \pm \cdots \pm p_d\gamma_d \tag{14}$$

shows the same problem. To overcome it we simply specify a single sign in front of each p_i over the whole of momentum space. A similar approach on the lattice would be to choose a particular Brillouin zone, e.g.,

$$p_\mu \in (-\pi/a, +\pi/a), \quad \forall\mu, \tag{15}$$

and stick to it. However, such a restriction cannot be consistently enforced in the presence of gauge interactions. Suppose we have a fermion loop, and restrict the momentum in one fermion line according to Eq. (15). The other fermion momenta are completely specified by that in the given line and by the external (photon) momenta. In general, they will range over $(-\pi/a + k, +\pi/a + k)$ for some $k \neq 0$, proving our assertion of inconsistency.

V. RETURN OF DOUBLING

A general solution of the Dirac equation using $D'_\mu(p)$ will have a $4\pi/a$ period in each p_μ. To get a well-defined theory with solutions of period $2\pi/a$ we must take symmetric combinations of the form

$$\tilde{\psi}_s \equiv \tilde{\psi}(p) + \tilde{\psi}(p + 2\pi/a). \tag{16}$$

If p is chosen to make $\tilde{\psi}(p)$ left-handed then $\tilde{\psi}(p + 2\pi/a)$ is right-handed. We have spectrum doubling.

The theory obtained this way is actually a subset of the standard doubled theory (Secs. II and III) on a lattice of spacing $\frac{1}{2}a$. Consider the following.

(i) $D_\mu'(p)$ [Eq. (13)] is just $D_\mu(p)$ [Eq. (5)] with $a \to \frac{1}{2}a$.

(ii) In order to define $D_\mu'(p)$ in terms of fermion fields $\psi(x)$ we have to use the derivative prescription

$$\frac{\partial\psi}{\partial x_\mu} \to \frac{1}{a}\{\psi(x + a\hat{\mu}) - \psi(x)\} \tag{17}$$

to define $\dfrac{\partial\psi}{\partial x_\mu}$ at $x + \frac{1}{2}a\hat{\mu}$. If we do this then

$$p_\mu\tilde{\psi}(p) \to i\sum_x e^{ip(x + (1/2)a\hat{\mu})}\frac{\partial\psi}{\partial x_\mu} = (2/a)\sin(p_\mu a/2)\,\tilde{\psi}(p), \tag{18}$$

as required. The $\frac{1}{2}a$-spaced lattice is unavoidable.

(iii) The Fourier transform of Eq. (16) is ($d = 1$)

$$\psi(na) = \int_0^{2\pi/a} e^{ip\cdot na}\left(\tilde{\psi}(p) + \tilde{\psi}\left(p + \frac{2\pi}{a}\right)\right)dp = \int_0^{4\pi/a} e^{ip\cdot na}\tilde{\psi}(p)\,dp. \tag{19}$$

The final expression is exactly what we expect in the doubled $\frac{1}{2}a$-spaced theory.

It should be clear that the particular subset of the $\frac{1}{2}a$-spaced theory selected by Eq. (16) is that where all fields except those on the $x_i = n_i a$ sublattice are constrained to vanish. The remaining subfield gives a symmetrized mixed fermion theory (cf. end of Sec. II). This is useless if we want fermions of a specific handedness. If the theory makes no distinction (e.g., quantum electrodynamics) it might be useful. Otherwise, since we already have all the effects of doubling, we might as well have the full 2^d-doubled $\frac{1}{2}a$-spaced theory. Note that bosons [using Eq. (11)] live on a double-spaced sublattice exactly as in Sec. III.

VI. COVERING SPACES AND AN ANALOG

In the search for deeper understanding we will now consider a closely analogous problem. The attempt to define a square-root function on the complex plane is beset by difficulties identical to ours in Sec. IV. Again there are sign ambiguities, which cannot be resolved consistently. If we follow one choice of sign through a 2π rotation about the origin we arrive at the opposite sign. We will consider possible approaches on $\mathbb{C}$, and see how they can be applied to the lattice. Note that here and in Sec. VII we consider only $d = 1$ for simplicity. The generalization to arbitrary d is straightforward.

A. Complex analysis

One option on $\mathbb{C}$ is to cut the plane from 0 to ∞, and define a consistent square-root function on the cut plane. This approach is crude and generally unsatisfactory. The position of the cut is arbitrary, and varies when z is rotated; there are unwelcome discontinuities across the cut; and the topology of the plane is altered.

A much more satisfactory understanding of all "multivalued" functions is provided by the theory of Riemann surfaces.[2] According to this the square-root function has for its range not $\mathbb{C}$ but a surface over $\mathbb{C}$. Two points on the surface lie above each point z of $\mathbb{C}$. These correspond to the two choices of $\sqrt{z}$. If we could define the two branches consistently over all $\mathbb{C}$ then the surface would be like two distinct copies of $\mathbb{C}$. We cannot, and the surface is connected. A 2π rotation about the origin on $\mathbb{C}$ is covered by a line on the surface whose end points are distinct, but lie above the same point on $\mathbb{C}$. A further 2π rotation on $\mathbb{C}$ is required to give a closed loop on the covering surface. Observe that in this approach the square-root function is essentially $1 \to 2$, taking $z \in \mathbb{C}$ to two points $(\pm\sqrt{z})$ on the covering space. It is natural that the square root should have this property, since it is the inverse of a function (the square) that is $2 \to 1$.

B. Lattice analogs

The cut plane approach is analogous to the SLAC derivative prescription[3] and others like it.[4] The discontinuities in lattice momenta cause serious problems.[5] This approach is not a success, and we will not consider it further.

The ideas of the Riemann surface are easily applied to the lattice map

$$(4/a^2)\sin^2(pa/2) \to (2/a)\sin(pa/2). \tag{20}$$

The Brillouin zone on which p and $(4/a^2)\sin^2(pa/2)$ are defined is topologically a circle S^1 of radius $1/a$ (period $2\pi/a$). The square root $(2/a)\sin(pa/2)$ is defined on a space above $S^1(1/a)$ that is like the edge of a Möbius band of radius $1/a$. If it had been possible to define the two branches consistently everywhere, this space would have been the border of a twist-free band [i.e., two copies of $S^1(1/a)$].

Once again we are dealing with a $1 \to 2$ function. The consistency problems of Sec. IV arose solely because we tried to treat it as a $1 \to 1$ function.

The edge of the Möbius band is topologically a circle S^1 of radius $2/a$. The lattice momentum $(2/a)\sin(pa/2)$ is properly defined on $S^1(2/a)$ rather than on the Brillouin zone $S^1(1/a)$. The Dirac equation defines ψ in terms of the lattice momentum, so ψ too lives on the space $S^1(2/a)$. If we are to interpret $\psi((2/a)\sin(pa/2))$ as a function on the Brillouin zone $S^1(1/a)$ we must associate each p with *two* fields, $\psi((2/a)\sin(pa/2))$ and $\psi((-2/a)\sin(pa/2))$. That is, we have a doubled spectrum. *Doubling occurs because the Dirac operator is defined as a square root, and square roots are properly* $1 \to 2$ *functions.*

VII. CONTINUUM SPECTRUM DOUBLING

We can duplicate the analysis of Sec. VI in the continuum. Equation (2) implies that it is not the momentum p that appears in the Dirac equation but

$$\sqrt{p^2} = \pm p.$$

The square root is defined on a covering space over momentum space $(\mathbb{R}^1)$ that is like the edge of an infinite band. Following Sec. VI we define ψ also on the edge of the band. Once more there are two fields associated with each momentum p, $\psi(p)$ and $\Psi(-p)$, each satisfying a conventional Dirac equation. We are led to a theory that is just the $a \to 0$ limit of the lattice theory of Sec. VI, with the doubled spectrum of that theory.

Of course it is possible to avoid doubling in the continuum by restricting attention to one component of the covering space. Whether it is legitimate to do so depends on how fundamental Eq. (2) is.

VIII. COMMENTS AND CONCLUSIONS

Assuming that the square-root definition of δ or $\not{p}$ is fundamental, we have found the following.

(a) When bosons and fermions appear together on a lattice the fermions are 2^d doubled, while the bosons must be restricted to a double-spaced sublattice if they are not to be doubled also (Secs. III and V). This result could be important for models like that of Ref. 6, which have bosonic and fermionic excitations of the same field.

(b) When the space on which the lattice momentum lives is defined properly, and the fields in the Dirac equation are defined on that space, then fermion doubling is inevitable (Sec. VI).

Finally, we must comment on our fundamental assumption [Eq. (2)].

(i) If δ were well understood it would be simple to use Eq. (1) to define $\Box$, and there would be no ambiguities. However, it is the Laplacian which is well understood, and the theory of bosons, so that is where we must start. We need Eq. (2). It is notable that Kähler fermions[7] start from Eq. (2) to get a different sort of square root of $\Box$. Since lattice bosons are well understood and unproblematic it is natural to copy the continuum analysis from the beginning and use Eq. (2) on the lattice.

(ii) It is possible that such properties of ψ as its double-valuedness under rotations are connected with the square-root ambiguities discussed above.

(iii) The square-root definition has the virtue that it *makes sense* of lattice fermion doubling [conclusion (b)]. Against this may be set the fact that the same analysis leads to fermion doubling in the continuum (Sec. VII). This last fact may not be the demerit it seems. Perhaps it indicates that the lattice's problems are rooted in the continuum, which is where we should look for their solution. We look forward to further comments.

[1] K. G. Wilson, Phys Rev. D **10**, 2445 (1974); H. B. Nielsen and M. Ninomiya, Nucl. Phys. B **185**, 20 (1980); L. H. Karsten, Phys. Lett. B **104**, 315 (1981).

[2]See, e.g., O. Forster, *Lectures on Riemann Surfaces* (Springer, New York, 1981) or H. M. Farkus and I. Kra, *Riemann Surfaces* (Springer, New York, 1980). (N. B. The ideas used in Sec. VI are fully described there.)
[3]S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D **14**, 1627 (1976).
[4]R. Stacey, Phys. Rev. D **26**, 468 (1982).
[5]J. M. Rabin, Phys. Rev. D **24**, 3218 (1981); R. Stacey, Z. Phys. C **10**, 75 (1983).
[6]The sine–Gordon model. See T. H. R. Skyrme, Proc. R. Soc. London Ser. A **247**, 260 (1958); S. Coleman, Phys. Rev. D **11**, 2088 (1975).
[7]E. Kähler, Rend. Mat. (Roma) **21**, 425 (1962).

Differential characters: The Dirac monopole as an example

Robert Coquereaux[a)]
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, New York 11794

(Received 27 March 1985; accepted for publication 21 June 1985)

Choosing the Dirac monopole as an example, the theory of differential characters is sketched and the quantization condition is recovered in a new way without considering singularities and without using a global formulation of gauge fields (i.e., without using fiber bundle techniques).

I. INTRODUCTION

Assume that the space (time) is a manifold of dimension n and suppose that, while studying some physical system, there is a contribution to the Lagrangian of a quantity A which cannot be associated with a globally defined n form on M. One possible solution is to subdivide M into patches $\{U_i\}$ and to choose a collection of locally defined quantities $\{A_i\}$ but then one also has to add a contribution to the action associated with the change of patches (a recent paper[1] actually generalizes these ideas and advocates the use of Čech cohomology in relation with topological quantization). Another possibility that we want to describe here is based on the fact that only phase factors (and not phases) are physically meaningful: thinking of a Bohm–Aharanov experiment, we are therefore led to *define* a magnetic monopole as a map $\hat{f}$ from loops on the two-sphere S^2 to the group $U(1) = R/Z$ (remember that, classically, this is a static problem with spherical symmetry and that a test particle moves on a two-dimensional sphere centered on the monopole). Actually, we will impose a quite severe restriction: although $\hat{f}$ is *not* a differential form, we want it to admit a kind of differential (a smooth two-form) in order to be able to write Stokes' theorem. It happens that the structure of the set of all such maps $\hat{f}$ [call it $\hat{H}_1(S^2, R/Z)$] can be computed easily, without using forms with singularities and even without using a global formulation of gauge fields. However, its structure could also be computed by considering one-forms with smooth derivatives and point singularities in S^2 (interpreted as gauge potentials) or by probing the structure of S^2 by putting U(1) bundles over it and classifying those bundles (via the first Chern class); the link between these last two approaches being provided by the pullback on S^2 of the secondary (Chern–Simons) characteristic classes, which are only locally defined on the base S^2. More generally (and in plain terms), a differential character of order p is an object which has nice derivatives and which assigns a U(1) phase factor to any p-cycle of M (for example, a S^p sphere included in M). To our knowledge, differential characters have never been used in physics and they are not even very popular among mathematicians since the basic reference[2] is not published. However, it is clear that the whole study of anomalies in gauge theories (in particular, the geometrical meaning of the Wess–Zumino effective Lagrangian, the Witten quantization condition, etc.) could fit into the framework of the theory of differential characters; this actually motivates the present article, which will be mainly expository. We will start with a short study of differential forms with singularities which have been considered historically in the first place, both in physics (Dirac string) and in mathematics.[3] There we give a first definition of differential characters (following Ref. 4), which is not really canonical but is quite intuitive. Then we give (following Ref. 2) an axiomatic definition of differential characters and derive a few of their properties; in order to illustrate these ideas, we recover the quantization condition for the Dirac monopole in a way which relies neither on the study of singularities (Dirac string) nor on fiber bundle techniques (as in Wu and Yang[5,6]). Finally, we show how differential characters can also be obtained from Chern–Simon classes and we recover the usual discussion based on the study of the U(1) bundles over S^2.

The abstract actually summarizes the content of section III which is the main section of the paper and can be read independently of the others.

II. DIFFERENTIAL FORMS WITH SINGULARITIES

Let M be a manifold of dimension n and let $f = f_k$ be a k-form defined on the complement of some $(p = n - k - 1)$-dimensional polyhedron d_p. We can think of f as a differential form with singularities but we impose that the singular set of f [call it sing(f)] is included in d_p; notice that sing(f) needs not to be a manifold. In the example of the Dirac monopole, we would take $M = S^2$, f_1, a one-form, and $d_0 = \text{sing}(f)$, a point on S^2. We suppose, moreover, that there exists a differential form $\omega = \omega_{k+1}$ or order $k+1$, defined on all M such that $\int_{c_{k+1}} \omega = \int_{\partial c_{k+1}} f$ whenever the boundary of the (otherwise arbitrary) $k+1$ chain c_{k+1} does not meet d_p. In other words, we set (by Stokes' theorem) $\omega = df$ whenever df is defined and assume that ω is smooth on M: we are therefore not interested in all possible k-forms with singularities but only in those which have smooth derivatives. In the example of the Dirac monopole, ω would be the Maxwell field. We would like to define an object $\hat{f}$, associated with f, which could be integrated on all the cycles of M, even on those which intersect the singular set of f; the idea is the following: first notice that for a given arbitrary cycle z_k (even intersecting the singular set), it is possible to find a cycle z'_k not intersecting the singular set such that $z_k = \partial c'_{k+1} + z'_k$ for some $k+1$ chain c'_{k+1}; for example, with $n = 2$, $k = 1$, we are in the situation depicted by Fig. 1 (the cross × indicates a point singularity). If f had no singularity, we would be allowed to write Stokes' theorem as

a) Permanent address: Centre de Physique Theorique, Section 2, Centre National de la Recherche Scientifique Luminy, Case 907, Marseille 13288, France.

 0022-2488/85/123176-04$02.50

FIG. 1. A singularity on S^2.

$$\int_{c_2'} df = \int_{\partial c_2'} f = \int_{z_1 - z_1'} f = \int_{z_1} f - \int_{z_1'} f, \tag{1}$$

and we would get, in particular,

$$\int_{z_1} f = \int_{z_1'} f + \int_{c_2'} df. \tag{2}$$

We are therefore tempted to define the object $\hat{f}$ acting on arbitrary k-cycles as the k-dimensional generalization of the right-hand side (rhs) of Eq. (2). However, if f possesses singularities, the above prescription is ambiguous; to see it, let us return to the (Dirac) example $n = 2$, $k = 1$ and choose the cycle z_1'' depicted by Fig. 2; a careless use of Stokes' theorem would lead to

$$\int_{z_1} f = \int_{z_1''} f + \int_{c_2''} df. \tag{3}$$

Let us then evaluate the difference between the rhs of Eq. (2) and Eq. (3); we get

$$\left(\int_{z_1''} f + \int_{c_2''} \omega\right) - \left(\int_{z_1'} f + \int_{c_2'} \omega\right) = \int_c \omega - \int_{\partial c} f, \tag{4}$$

where $c = c_2'' - c_2'$ is the two-chain depicted by Fig. 3. The right-hand side of (4) does not vanish because of the singularity and is a deviation from Stokes' formula, which we can call the residue of f relative to c. In this particular case, if $\int \omega$ is an integer n and if we call e the oriented chain which satisfies $c + e = [S^2]$ (intuitively e is the oriented complement of c in S^2), then $\partial c = -\partial e$ and we can write

$$n = \int_{S^2} \omega = \int_c \omega + \int_e \omega = \int_c \omega + \int_e df$$
$$= \int_c \omega + \int_{\partial e} f = \int_c - \int_{\partial c} f$$

The right hand side of (4) is therefore an integer; *the right-hand-side of (2) is then defined up to an integer and we have to reduce* mod Z *to get a single valued map.* Let bar denote reduction mod Z, then, we define

$$\langle z_1, \hat{f} \rangle = \overline{\int_{z_1'} f + \int_{c_2'} \omega} \in \mathrm{U}(1) = \frac{R}{Z}. \tag{5}$$

Recall that $\omega = df$ whenever df is defined. More generally, if c is a $k+1$ form and if f has no singularity on the boundary

FIG. 2. A singularity on S^2.

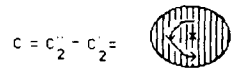


FIG. 3. A singularity on S^2.

∂c, the rhs of (4) is linked to the number $\#(\mathrm{sing}(f) \cap c)$ of singularities of f contained in c (taking into account the possible multiplicities and the orientation); notice that in order for this number to be a finite integer in a generic case, we have chosen p (the dimension of the polyhedron d_p) satisfying $p + (k+1) = n$ and also ω as a form with integer periods. The above prescription (5) can than be generalized[4] and we define the differential character $\hat{f}$ associated with f as the map from (arbitrary) k-cycles z_k to U(1) by the equation

$$\langle z_k, \hat{f} \rangle = \overline{\int_{z_k} f + \int_{c_{k+1}'} df} \in \mathrm{U}(1) = \frac{R}{Z}, \tag{6}$$

where z_k' is any cycle not intersecting the singular set of f and where c_{k+1}' is such that $\partial c_{k+1}' = z_k - z_k'$. In practice, it is convenient to write $\int_{z_k} \hat{f} = \langle z_k, \hat{f} \rangle$, but one should keep in mind that this is not a real number and that $\hat{f}$ is not a differential form. Finally, one should also notice that the correspondence $f \to \hat{f}$ is not one to one; indeed, it is clear that if η is a smooth, closed ($d\eta = 0$) k-form with integer periods, then f and $f + \eta$ define the same differential character $\hat{f}$; the representation of differential characters by forms with singularities is therefore not canonical. Also, the amount of information contained in $\hat{f}$ is not clear at that point: we saw that f determines $\hat{f}$ but we will see that $\hat{f}$ actually determines $\omega (= df)$ and also another quantity associated with the ambiguity at the singularity (it will turn out to be an integer cohomology class). These two remarks justify the following (axiomatic) definition.

III. DIFFERENTIAL CHARACTERS

As already announced in the introduction, a map of $\hat{f}$ from k-cycles to $R/Z = \mathrm{U}(1)$ is called a differential character[2] if it is a homeomorphism $[\hat{f}(z_1 + z_2) = \hat{f}(z_1) + \hat{f}(z_2)]$ and if there is a differential form ω (of degree $k+1$) such that $\overline{\omega} = \hat{f}\partial$ (where bar still denotes reduction mod Z). With more familiar notations, we would write $\int_{z_1 + z_2} f = \int_{z_1} \hat{f} + \int_{z_2} \hat{f}$ and $\overline{\omega} = d\hat{f}$, although $\hat{f}$ is not a differential form. It should be clear that such restrictions are quite severe and that not any U(1) valued map on k-cycles (even a homeomorphism) can be extended (via ω) on $k+1$ chains (their boundary being k-cycles). We now want to study a few properties of $\hat{f}$ but at this point it becomes convenient to use the terminology of simplicial (co)homology with coefficients in Z (or R or R/Z) because we do not want to miss a possible torsion phenomena which would be overlooked by the use of De Rham cohomology. Let us choose Q, a real cochain (a homomorphism from chains to R) such that $\overline{Q}$ coincides with $\hat{f}$ on k-cycles; then $d\overline{Q} = \overline{\omega}$ (by definition of $\overline{\omega}$); dQ will therefore differ from ω by some integral cochain c (a homomorphism from chains to Z) and one gets $dQ = \omega - c$. Let us now show that ω and c are necessarily closed: since $d^2 = 0$, we get $d\omega = dc$, which would imply that $d\omega$ has integral values, but it is clear that a differential

form, even if it has integral periods (i.e., integral values when integrated over cycles) cannot have integral values (when integrated on any chain); therefore, $d\omega = dc = 0$. Via integration, we can consider ω as a real cochain, the relation $\omega - c = dQ$ implies then that ω and c (considered as real) belong to the same cohomology class (call it $[\omega]$). Let us now show that ω has integral periods: we know that $\hat{f}\partial$ vanishes on k-cycles z(since $\partial z = 0$) but $d\bar{Q} = f\partial$, therefore $d\bar{Q}$ also vanishes on cycles which implies that dQ has integer values when evaluated on cycles. The same will be true for $\omega = dQ + c$ (since c has integer values on anything). The choice of Q was, by no means, unique; it remains to show how the above is modified by choosing Q', another real cochain coinciding with $\hat{f}$ on k-cycles: then $\bar{Q} - \bar{Q}' = 0$ on cycles and therefore $Q - Q'$ takes integer values on cycles; this implies first that $d(Q - Q')$ is an integral $(k+1)$ cochain and that $Q - Q' = t + dw$ where t is an integral cochain and w is some $k-1$ real cochain (the dw contribution will vanish we evaluate $Q - Q'$ on cycles); we can now write $Q' = \omega' - c'$ as we did for Q and we reach the conclusion that $dQ - dQ' = dt = (\omega - \omega') - (c - c')$ is an integral cochain. Here again, we use the property that a differential form cannot take integer values when evaluated on any chain; this shows first that $\omega' = \omega$, next that $c' - c = dt$. In other words, ω is uniquely determined and c,c' belong to the same integer cohomology class (call it $u = [c]$); actually, we have already seen that ω and c (considered as real) belong to the same real cohomology class. The main conclusion of this analysis is that a differential character $\hat{f}$ of degree k determines (i) a closed differential form ω of degree $k+1$, with integral periods; and (ii) an integer cohomology class u of degree $k+1$. The following exact sequences provide a nice bookkeeping device which helps one to remember most of the above:

$$0 \to Z \to R \to R/Z \to 0, \tag{7}$$

$$0 \to H^k(R/Z) \to \hat{H}^k(R/Z) \xrightarrow{\delta_1} \Lambda_Z^{k+1} \to 0, \tag{8}$$

$$0 \to \Lambda^k/\Lambda_Z^k \to \hat{H}(R/Z) \xrightarrow{\delta_2} H^{k+1}(z) \to 0. \tag{9}$$

Here, $H^k(R)$, $H^k(Z)$, $H^k(R/Z)$ denote cohomology groups of M with coefficient in $R,Z,R/Z$; $\hat{H}^k(R/Z)$ is the group of differential characters, Λ^k is the set of differential forms of degree k, and Λ_Z^k is the subset of closed k-forms with periods lying in Z; also, we set $\omega \doteq \delta_1(\hat{f})$ and $u \doteq \delta_2(\hat{f})$ for $f \in \hat{H}^k(R/Z)$. These exact sequences also provide a very efficient tool that we can use to compute $\hat{H}^k(R/Z)$ as we shall see below.

We now return to physics and to the magnetic monopoles by defining a Dirac monopole as a differential character of degree 1 on the two-sphere S^2. In other words, we want to classify the possible phase factors on S^2. Mathematically, S^2 comes in because S^2 is a deformation retract of $R^3 - \{0\}$ and has therefore the same cohomology; we would work directly with $R^3 - \{0\}$ as well with a monopole sitting at the origin. To compute $\hat{H}^1$, we use the exact sequence (8); it is well known that, for a sphere, only the first and the last cohomology groups are nontrivial; also, a sphere has no torsion, hence, for S^2, we get $0 = H^1(Z) = H^1(R) = H^1(R/Z)$, $H^2(Z) = Z$ and we get the short sequence $0 \to \hat{H}^1 \xrightarrow{\delta_1} \Lambda_Z^2 \to 0$; therefore, δ_1 is one to one in this case. In other words, a Dirac monopole is fully characterized by a differential form of degree 2 on S^2 (physically, the magnetic field of the monopole), which, when integrated over S^2, gives an integer (all other two-cycles are just multiples of S^2): we recover the usual results. Notice that we can write $\hat{H}^1 \simeq \Lambda_Z^2 \simeq Z + d\Lambda^1$, indeed a two-form with integral periods on S^2 is, in particular, characterized (up to the d of a one-form) by the value it takes on the fundamental cycle (S^2 itself); of course, $d\Lambda^1$ gives a zero contribution by Stokes' theorem ($\partial[S^2] = 0$). This corresponds to the fact that for a fixed integer n, we can write any monopole magnetic field (solution or not of Maxwell's equations) as the sum of an integer multiple of the volume form ($F = (n/2)\sin\theta\, d\theta_\Lambda\, d\phi$, this solves Maxwell's equations) and of the differential of any one-form. Before ending this paragraph, and in order to complete the study of differential characters on S^2, we would just like to mention that, using Eq. (8) and (9), one shows that $\hat{H}^0$ is isomorphic with the set of all smooth maps from S^2 to $U^1 = R/Z$ and that $\hat{H}^2 = R/Z = \mathrm{U}(1)$. Notice that we have obtained these results without dealing with a possible point singularity on S^2 (or with string singularity on $R^3 - \{0\}$) and that we did not make any use of the classification of U(1) bundles yet: we now want to study this last link.

IV. RELATION WITH PRINCIPAL BUNDLES

Another way of obtaining the same results relies on the classification of U(1) bundles over S^2; this method became standard (see, for instance, Ref. 7) after the work of Ref. 5: the first Chern class for a U(1) principal bundle P is $c_1(P) = -[F/2\pi]$, F being the curvature of an arbitrary connection; the integral of c_1 for a U(1) bundle over S^2 is the integer giving the monopole charge $\int_{S^2} c_1 = -n$; if $n = 0$, we have the trivial bundle $P = S^2 \times \mathrm{U}(1)$, if $n = 1$, we have the U(1) Hopf fibration of $S^3 = \mathrm{SU}(2)$ over S^2, etc. Notice that F is a closed form on S^2, the base of the bundle, and also closed on P; however, although exact on P [F is the d of the connection form for a U(1) bundle], it is not exact on S^2: one has to choose an open covering of S^2 with two patches H_+, H_- and write $F = dA_+$ on H_+, $F = dA_-$ on H_-, the one-form A_+, associated with the northern hemisphere H_+ has a singularity at the south pole; if we now want to define globally on S^2, an object associated with $A_\pm$, we have to consider differential characters (either defined from forms with singularities as in Sec. II or abstractly, as in Sec. III). More generally, it is possible to construct differential characters out of principle bundle P *with* connection θ: let Ω be the curvature form of θ and I, an invariant polynomial on G (the structure group of P), then the cohomology class of $I(\Omega)$ is a characteristic class; $I(\Omega)$ is closed on M (the base of the bundle P) and on P, it is usually not exact on M (it represents a nontrivial cohomology class); however, is is well known that it is exact on P. One can then find a form $Q(\theta)$, globally defined on P such that $I(\Omega) = dQ(\theta)$. To construct $Q(\theta)$ and other Chern–Simons forms, one may use the so-called transgression formula.[8,9] It is clear that, when projected on the basis, the form $Q(\theta)$ is not a globally defined object: one has to use a local section to project it down and the above equality holds only locally. If we want to define globally on

the base M, an object corresponding to Q, one has to use the concept of differential characters: the reduction mod Z of Chern–Simons forms $Q(\theta)$ agrees, on the cycles of P, with the lifts (using the projection π^*) of differential characters on the base; for more details about these properties (in particular, if one wants to study what happens when the bundle P above M can be extended to a manifold N such that $\partial N = M$), one should consult Ref. 2. In our particular example (the Dirac monopole), the situation is quite simple: since we are interested in the first Chern class (whose representatives are proportional to the curvature of a chosen connection), the Chern–Simons form coincides in this simple case with the connection itself; if we call $\hat{A}$ the differential character under study (i.e., the Dirac monopole under study), we see, with the notation of Sec. III that $\delta_1(\hat{A}) = F$, the curvature two-form and that $\delta_2(\hat{A}) = c_1(P) = [F]$, the first Chern class which indeed is an integer class. In this example, $\delta_2(\hat{A})$ can be deduced from $\delta_1(\hat{A})$, this is not necessarily so in general; to build a very general example, one should use a bundle (P,M,G) admitting a non-torsion-free classifying space BG, choose an integer characteristic class u and an invariant polynomial I such that u and $[\omega] = [I(\Omega)]$, for some connection θ, coincide at the real level, then finally proceed to the construction of the differential characters via the mod Z reduction of the Chern–Simons class $Q(\theta)$ associated with $I(\Omega)$.

V. CONCLUSION

Besides the intrinsic interest of casting a new light on an old subject (the Dirac magnetic monopole), we hope that this sketchy discussion of the theory of differential characters will have convinced the reader of its potential impact in physics; in particular, the remarks made in Sec. IV show why the analysis of anomalies in gauge theories could be done by studying the set of differential characters associated to the space of connections modulo gauge transformations. Notice that an analogous theory can be held by merely replacing Z by Q, the rational numbers, and therefore $U(1) = R/Z$ by R/Q, this remark may indicate that differential characters could also provide a natural mathematical framework for the study of other physical phenomena (like the quantum Hall effect).

ACKNOWLEDGMENTS

I would like to thank my friends K. Pilch for many discussions and especially T. Januszkiewicz for his patient explanations. I would also like to thank Professor C. N. Yang and the members of ITP for their hospitality.

This work was supported in part by National Science Foundation Grant No. 81-09110 A03.

[1] O. Alvarez, "Topological quantization and cohomology," preprint LBL 18658, 1984.

[2] J. Cheeger and J. Simons, "Differential characters and geometric invariants," Mathematical preprint, Stony Brook, 1973.

[3] C. B. Allendoerfer and Y. Eells, Jr., "On the cohomology of smooth manifolds," Comment. Math. Helv. **12**, 165 (1957).

[4] J. Cheeger, *Multiplication of Differential Characters, in Symposia Mathematica* (INDAM, Bologna, 1973) Vol. XI.

[5] T. T. Wu and C. N. Yang, "Concept of nonintegrable phase factors and global formulation of gauge fields," Phys. Rev. D **12**, 3845 (1975).

[6] T. T. Wu and C. N. Yang, "Dirac monopole without string: Monopole harmonics," Nucl. Phys. B **107**, 365 (1976).

[7] T. Eguchi, P. B. Gilkey, and A. J. Hanson, "Gravitation, gauge theories and differential geometry," Phys. Rep. **66**, 213 (1980).

[8] R. Stora, "Algebraic structure and topological origin of anomalies," Cargese Lectures, 1983 (Plenum, New York, 1983).

[9] B. Zumino, "Chiral anomaly and differential geometry," Les Houches Lectures, 1983 (North-Holland, Amsterdam, 1983).

Spinorial infinite equations fitting metric-affine gravity

A. Cant[a] and Y. Ne'eman[b]
Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel[c]

(Received 12 December 1984; accepted for publication 12 June 1985)

Two different approaches are used to construct infinite-component spinor equations based on the multiplicity-free irreducible representations of $\overline{\mathrm{SL}}(4,\mathbb{R})$. These "manifield" equations are SL(2,C) invariant; they exist in special relativity, and can directly be coupled to gravitation in the metric-affine theory, i.e., in Einstein's general relativity with nonpropagating torsion and nonmetricity. In the first approach the maximal compact subgroup $\overline{\mathrm{SO}}(4)$ of $\overline{\mathrm{SL}}(4,\mathbb{R})$ is "physical." A vector operator X^{μ} is constructed directly in the infinite-dimensional reducible representation $\mathscr{D}^{\mathrm{disc}}(\frac{1}{2},0) \oplus \mathscr{D}^{\mathrm{disc}}(0,\frac{1}{2})$. In the second approach, SL(2,C) and a vector operator γ^{μ} are embedded directly in $\overline{\mathrm{SL}}(4,\mathbb{R})$ via the Dirac representation. A manifield equation is then constructed (in a manner analogous to the Majorana equation) by taking an infinite-dimensional irreducible multiplicity-free representation of SL(4,R), spinorial in j_1, in the (j_1,j_2) reduction over $\overline{\mathrm{SO}}(4)$. Both manifields can fit the observed mass spectrum.

I. INFINITE COMPONENT FIELDS

Relativistic quantum field theory exploits the concept of a local field as the fundamental dynamical object, with the particle aspect emerging as the offspring. The particles span unitary irreducible representations of the Poincaré group ISO(3,1) and its double covering $\overline{\mathrm{ISO}}(3,1)$. Fields, on the other hand, transform as finite—and thus nonunitary—representations: of GL(4,R) when tensorial, or of SL(2,C) for spinor fields. The latter group appears here as the double covering of the Lorentz group, i.e., $\mathrm{SL}(2,\mathbb{C}) = \overline{\mathrm{SO}}(3,1)$(quantum probabilities do not involve phases and thus allow the double covering). The nonunitarity of the representations [or non-Hermiticity of the relevant matrices of the Lorentz or GL(4,R) algebras] does not matter physically: the Lagrangian's Hermiticity requires the addition of the complex conjugate expression, and the non-Hermitian parts of the Noether-theorem-generated densities cancel.[1] As a result, the special Lorentz transformations in SL(2,C), for instance, have only orbital components with the pieces $\int d^3x(\bar{\psi}\sigma_{0i}\psi + \mathrm{H.c.})$ canceling. Boosting an electron state thus contributes only to the kinetic energy. The same type of cancellation occurs for the (noncompact) deformation generators in GL(4,R).

The compact subalgebras of GL(4,R) or SL(2,C) being the only ones to contribute to the physical currents and generator observables, why do we need the full groups altogether? The action and its Lagrange density have to be globally invariant under the (active) Poincaré group. When we include gravity we require invariance under the (passive) general covariance group (the diffeomorphisms Δ with local dependence of the transformations). The latter is realized nonlinearly over the linear subgroup GL(4,R); we thus have to use "world tensors" and the equivalence principle can be fulfilled in the easiest manner by keeping them in special relativity, too. This involves regarding GL(4,R) as "GL(3,1;R)," i.e., introducing the Minkowski metric linearly and identifying accordingly the SO(3,1) subgroup, with the special Lorentz transformations given by symmetrical matrices that do not belong to the SO(4) maximal compact subgroup of GL(4,R). Alternatively, one may define $x^4 = ix^0$ (the "Pauli metric") and identify the orthogonal matrices of the compact SO(4) with the physical Lorentz group, as we demonstrate in (5.3). One can then ask, in either case, for global Lorentz invariance and ensure that this be manifest invariance.

For spinor fields with a finite number of components, the transition to Δ does not exist [there is no finite spinorial representation of $\overline{\mathrm{SL}}(4,R)$] and the spinor components are invariant under GL(4,R). Global (active) covariance under the double-covered Poincaré group is formally ensured by SL(2,C). For all fields $\overline{\mathrm{SO}}(3,1) = \mathrm{SL}(2,\mathbb{C}) \subset \overline{\mathrm{ISO}}(3,1)$, the double covering of the Poincaré group is the global covariance group in the final result. This is thus the covariance group of special relativity both for particles and for fields. Of course there is the additional advantage of a smooth manifestly invariant classical fields' limit, where the particle aspect does not enter, and neither does unitarity.

Infinite-component fields, however, as they correspond to unitary representations of SL(2,C) and to Hermitian infinite matrices of the sl(2,C) algebra will, in contradistinction, yield "internal" contributions to the special Lorentz transformations. In this case, the boosts will excite the spin variable, too, and may thus contribute to the potential energy (i.e., connect to a different mass). Such infinite-component fields were first introduced by Majorana,[2] who used the only two irreducible representations of SL(2,C) for which an invariant (linear) first-order wave equation of the form[3]

$$(X^{\mu}\partial_{\mu} + ik)\,\psi(x) = 0 \tag{1.1}$$

can be written. The operators X^{μ} ($\mu = 0,1,2,3$) close on the Lie algebra $\mathrm{sp}(4,\mathbb{R}) \cong \mathrm{so}(3,2)$. The so-called "ladder representation" of Sp(4,R) is unitary and splits into the direct sum of the two Majorana representations. The algebra sp(4,R) was suggested as a spectrum-generating algebra[4] for hadronic,

[a] Present address: Department of Theoretical Physics, Research School of Physical Sciences, Australian National University, Canberra ACT 2601, Australia.
[b] Also on leave from University of Texas, Austin, Texas 78712.
[c] Wolfson Chair Extraordinary in Theoretical Physics.

 0022-2488/85/123180-10$02.50

nuclear, or other excitations, following the reintroduction of Majorana's work by Fradkin.[5] However, difficulties arose due to the presence of a continuous set of solutions with spacelike momenta, in addition to the discrete spectrum—which is itself not realistic since states of higher spin have a smaller rest mass. Dirac[6] recently rediscovered these equations and further developed the formalism.

II. GRAVITY: THE EINSTEIN, EINSTEIN–CARTAN AND METRIC-AFFINE THEORIES

(i) Einstein's theory is Riemannian, i.e., it precludes the propagation of either torsion or nonmetricity. Only the metric field $g_{\mu\nu}(x)$ propagates. Alternatively, we may use the tetrad fields $e^a_\mu(x)$, with

$$g_{\mu\nu}(x) = e^a_\mu(x)\, e^b_\nu(x)\eta_{ab}, \tag{2.1}$$

where η_{ab} is the Minkowski metric $(+1,-1,-1,-1)$.

In the above, the Latin indices a,b represent components of the four-vector representation of the anholonomic group. In Einstein's theory with spinor matter fields,[7] or in Einstein–Cartan theory,[8] this is SL(2,C) acting on the local frames.

The tetrad fields had to be introduced[7] in gravity after the discovery of the electron's spin, in order to cope with half-integer spin fields. In differential geometry they describe a general moving frame, i.e., a set of one-forms θ^a defined over some region U of space-time:

$$\theta^a = e^a_\mu\, dx^\mu. \tag{2.2}$$

At each point $x = \chi \in U$, the θ^a serve as local "coordinates," inertial at $\mathscr{P}$. From the principle of equivalence, i.e., a smooth transition to special relativity, when the gravitational field is extinguished, we now get a requirement of *local* $\mathrm{SL}(2,\mathbb{C}) = \overline{\mathrm{SO}}(3,1)$ invariance of the locally inertial coordinate system at each point: the frame is orthonormal. The spinor field carries a $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation of this (anholonomic) local Lorentz group, but is invariant under the diffeomorphisms (general coordinate transformations). Ordinary tensor fields vary under the (passive) action of the (holonomic) diffeomorphism group Δ and its affine [GA(4,R)] and linear [GL(4,R)] subgroups but are scalar under the anholonomic Lorentz group. To recapture their variation under the active anholonomic transformations of the local Lorentz group (and thus to satisfy the principle of equivalence) they have to be contracted with the tetrads:

$$\phi^{ab\cdots}_{cd\cdots}(x) = e^a{}_\mu(x)\, e^b{}_\nu(x)\cdots(e^{-1})^\rho_c(x) \times (e^{-1})^\sigma_d(x)\cdots\phi^{\mu\nu\cdots}_{\rho\sigma\cdots}. \tag{2.3}$$

They would then become world scalars (i.e., invariant under the holonomic Δ). General relativity with spinors is thus rewritten in a manner which makes Δ act trivially on all fields.

This treatment was presented in most textbooks as if it was required by the (erroneous) assumption (to which we return later) that there can be no world spinors, i.e., that the diffeomorphism group has no double covering $\bar{\Delta}$. In any case this is irrelevant for the Dirac field, as there are indeed no finite-dimensional unitary bivalued representations of SL(4,R), GL(4,R), or GA(4,R), or of the diffeomorphism group Δ. Finite spin fields are thus treated anholonomically only, as objects belonging to the tangent manifold, Minkowskian for a theory obeying the equivalence principle.

So the discovery of half-integer spin did not modify Einstein's theory, but it required reexpressing the gravitational field in terms of tetrads rather than the metric, the latter now appearing as a higher construct.

(ii) In Einstein–Cartan gravity, as developed by Sciama, Kibble, Trautman, and Hehl,[8] space-time is allowed to carry torsion, as well as curvature. Applying the Poincaré group double covering as a local gauge on the anholonomic indices, curvature is seen as the field strength of the SL(2,C) Lorentz connection $\omega^{ab}_\mu(x)$ and torsion as that of the translation gauge field, represented by the tetrad $e^a_\mu(x)$ (the "fundamental form"):

$$R_{\mu\nu}{}^{ab} := \partial_\nu\omega_\mu{}^{ab} - \partial_\mu\omega_\nu{}^{ab} + \omega_\nu{}^a{}_c\omega^{cb}_\mu - \omega_\mu{}^a{}_c\omega^{cb}_\nu, \tag{2.4}$$

$$S_{\mu\nu}{}^a := \partial_\nu e^a_\mu - \partial_\mu e^a_\nu + \omega_\nu{}^a{}_c e^c_\mu - \omega_\mu{}^a{}_c e^c_\nu. \tag{2.5}$$

Holonomically, torsion introduces an antisymmetric piece in the Einstein connection $\Gamma_\mu{}^\rho{}_\nu$, in addition to the symmetric Christoffel symbol,

$$S_\mu{}^\rho{}_\nu = (e^{-1})^\rho_a\, S^a_{\mu\nu} = \tfrac{1}{2}(\Gamma_\mu{}^\rho{}_\nu - \Gamma_\nu{}^\rho{}_\mu). \tag{2.6}$$

Considering gravity heuristically as a gauge theory of the Poincaré group, one would thus have expected to deal with two gauge fields (both with spin $J=2$), i.e., ω^{ab}_μ for SL(2,C) and e^a_μ for the translations. In the Einstein–Cartan version of gravity, varying the Lagrangian with respect to both yields the two equations

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R_\rho{}^\rho = kE_{\mu\nu}, \quad k := 8\pi c^{-4}G \tag{2.7}$$

(Einstein's equation) and, with $\hat{S}_{\mu\nu\rho} := S_{\mu\nu\rho} + g_{\rho\mu}\, S_\nu{}^\sigma{}_\sigma - g_{\rho\nu}S_\mu{}^\sigma{}_\sigma$,

$$\hat{S}_{\mu\nu\rho} = k\Sigma_{\mu\nu\rho} \tag{2.8}$$

(Cartan's equation), where $E_{\mu\nu}$ is the energy-momentum density tensor and $\Sigma_{\mu\nu\rho}$ the angular-momentum density tensor.

Einstein's (and the Einstein–Cartan) Lagrangian for the gravitational field is linear and contains only one derivative (from 2.4). Ths is why (2.8) is just an algebraic equation and only implies a substitution of torsion by spin. The Einstein equation (2.7) contains curvature (2.4) and through it the connection ω and through (2.6) the torsion S, as can be seen by writing the holonomic expression for the connection for a four-dimensional Riemannian differential manifold with torsion

$$\Gamma_{\mu\nu}{}^\rho \equiv g^{\rho\sigma}\Delta^{\alpha\beta\gamma}_{\nu\tau\sigma}\,(\tfrac{1}{2}\partial_\alpha\, g_{\beta\gamma} - g_{\gamma\delta}S_{\alpha\beta}{}^\delta), \tag{2.9}$$

$$\Delta^{\alpha\beta\gamma}_{\nu\tau\sigma} := \delta^\alpha_\nu\delta^\beta_\tau\delta^\gamma_\sigma + \delta^\alpha_\tau\delta^\beta_\sigma\delta^\gamma_\nu - \delta^\alpha_\sigma\delta^\beta_\nu\delta^\gamma_\tau,$$

with

$$D_\mu\, g_{\nu\rho} := -Q_{\mu\nu\rho} = 0. \tag{2.10}$$

The first term in Γ is the Riemannian connection, and the second is the torsion contribution. In fact, Eq. (2.9) results from substituting (2.6) and (2.8) in (2.5) and solving for Γ. Substituting spin for torsion in (2.7) simply adds a term quadratic in spin on the right-hand side.[8] Here, D_μ is the covar-

iant derivative, with connection Γ and $Q_{\mu\nu\rho}$ is the nonmetricity tensor.

Thus, even though we have allowed torsion, it does not propagate. It is confined to the regions where the spin density exists. The effective theory is thus still Einstein's, except for the spin–spin term to be added to $E_{\mu\nu}$.

(iii) The metric-affine theory[9] allows the most general differentiable manifold L_4, with a connection (allowing parallel transfer) and a metric (allowing local measurements of angles and distances). Expression (2.1) does not vanish, and $\Gamma_{\mu\nu}{}^{\rho}$ in (2.9) acquires an additional term in the parentheses, $\frac{1}{2} Q_{\alpha\beta\gamma}$. The local gauge group on the anholonomic indices is GL(4,R), deforming the tetrad frames. We use a gravitational Lagrangian in which the connection is now this complete affine connection, and with a new term added,

$$\mathcal{L}(g,\partial g,\Gamma,\partial\Gamma) = (-\det g)^{1/2}(g^{\mu\nu}R_{\sigma\mu\nu}{}^{\sigma} + \beta Q_{\alpha}Q^{\alpha}),$$
$$\beta \neq 0, \quad Q_{\alpha} := \tfrac{1}{4} Q_{\alpha\gamma}{}^{\gamma} \tag{2.11}$$

(Q_{α} is known as the Weyl vector), we get as a third field equation,

$$Q_{\alpha} = k\Upsilon_{\alpha}, \tag{2.12}$$

where Υ_{α} is the scale current, a reducible component of $\Upsilon_{\alpha\beta\gamma}$, the hypermomentum tensor density

$$(\det -g)^{1/2}\Upsilon_{\rho}^{\mu\nu} \equiv -\frac{\partial\mathcal{L}}{\partial(\partial_{\nu}\psi)}\lambda^{\mu}{}_{\rho}\psi,$$

where the $\lambda^{\mu}{}_{\rho}$ are the matrices of the GL(4,R) algebra. Thus nonmetricity Q does not propagate, and is confined to the regions of nonvanishing deformation-current or scale-current density. The energy momentum tensor density acquires a new term quadratic in the scale current.

III. THE DOUBLE COVERING OF GL(4,R)

The anholonomic group acting on the local frames has thus been enlarged to form $SL(2,C) = \overline{SO}(3,1)$ in Einstein's theory with spinors and in Einstein–Cartan theory, or to GL(4,R) in the metric-affine theory.[9] The (erroneous) universal impression among physicists that GL(4,R) possesses no double covering[10] seemed to restrict the application of metric-affine gravity to bosonic matter.

The existence of a double covering $\overline{GL}(n,\mathbb{R})$ was realized in physics in 1977.[11] This implied the existence of spinor-type fields transforming (whether fermonic or bosonic) as "bandor"[12,13] unitary infinite-dimensional representations of the (meta-) linear, affine, and diffeomorphism groups; under reduction of these covering groups to the covering group of the orthogonal subgroup SO(3) the fields decompose into representations of $SU(2) = \overline{SO}(3)$. It had been conjectured[12] that hadrons with their Regge excitation bands could be described by such bandor irreducible unitary representations of $GL(3,\mathbb{R}) \subset GL(4,\mathbb{R})$. It was now proposed[13] that such a description should also fit their interaction with gravity. The physical interpretation of the $\overline{GL}(n,R)$ currents was clarified and it was suggested that in metric-affine gravity, spinor-matter fields indeed appear as infinite-dimensional unitary representations of the anholonomic $\overline{GL}(4,\mathbb{R})$ acting on the tetrad indices.[14]

The term *polyfield* or *manifield* was suggested. It was also pointed out that since the diffeomorphism group is realized through (nonlinear) group coordinates over the linear $GL(n,\mathbb{R})$ subgroup, manifields could also be considered as providing for world spinors,[11] i.e., holonomic spinors, whether in Einstein or in affine gravity.[15] In this role the representations correspond physically to the double covering of the Greek-indexed coordinate (holonomic) linear group $\overline{GL}(4,\mathbb{R}) \subset \overline{\Delta} = \overline{\mathrm{Diff}}(4,\mathbb{R})$, in contradistinction to the above anholonomic $\overline{GL}(4,\mathbb{R})$ acting on the tetrad (Latin) indices, in the metric-affine theory.

We thus have three gravitational roles for such manifields: (a) anholonomic spinor matter fields in the metric-affine theory,[14] (b) holonomic world spinors in "classical" Einstein gravity,[11] and (c) holonomic world spinors in affine gravity.[15]

Mickelsson[16] has constructed a wave equation fitting case (c). His equation is $\overline{GL}(4,\mathbb{R})$ invariant; when the gravitational field is extinguished, it preserves global $\overline{GL}(4,\mathbb{R})$invariance, i.e., it does not obey the principle of equivalence. On the other hand, it could fit in an affine theory with a basic non-Minkowski microscopic structure of the space-time manifold, perhaps with macroscopic spontaneous breakdown to Minkowski space-time. Such models have only been discussed qualitatively[15] to date. Another (technical) reason why we do not favor a $\overline{GL}(4,\mathbb{R})$ invariant equation is that—as we shall see—the "bandor" representations do not allow the construction of such an equation.

In this article, we propose two distinct ways of meeting case (a). The manifield equations we construct are of the form (1.1). Although they involve unitary representations of $\overline{GL}(4,\mathbb{R})$, they are only SL(2,C) invariant and thus have a good equivalence-principle limit. *They can be used as more infinite-component field equations in special relativity and conventional tetrad gravity, or [role (a)] as spinor matter manifields in metric-affine gravity.* Our equations are in close analogy with the Dirac equation, and, as for the Dirac case, the gravitational field enters through the inverse-tetrad fields

$$\bar{\psi}X^{a}\partial_{a}\psi \to \bar{\psi}X^{a}D_{a}\psi = \bar{\psi}X^{a}(e^{-1})_{a}^{\mu}(\partial_{\mu} + \omega_{\mu})\psi, \tag{3.1}$$

where D_a is the anholonomic covariant derivation and ω_{μ} is the connection. In Einstein gravity with Dirac spinor fields we have

$$\omega_{\mu} = \omega_{\mu}{}^{bc}\lambda_{bc}, \tag{3.2}$$

with λ_{bc} a finite-dimensional nonunitary matrix representation of the sl(2,C) algebra. In metric-affine gravity λ_{ab} is a unitary infinite-dimensional matrix representation of the gl(4,R) algebra. When gravitation is introduced, the λ_{ab} take the six SL(2,C) (nonlinear) values for Riemannian space-time, or the full 16 (matrix) values for metric-affine gravity.

In the next section we shall summarize the properties of multiplicity-free representations of SL(4,R); Secs. V and VI discuss the formation of wave equations according to two quite distinct approaches. In each case we propose infinite-component fields. These manifields may thus provide the correct mode through which the sequences of hadron excitations interact with gravity.[13,14] Both fit role (a) but only the manifield (5.8) may fulfill role (b).

IV. $\overline{\mathrm{SL}}(4,\mathbb{R})$ AND ITS REPRESENTATION

The unitary irreducible representations of the group $\overline{\mathrm{SL}}(3,\mathbb{R})$ have been constructed and listed.[17,18] Those of $G=\overline{\mathrm{SL}}(4,\mathbb{R})$ have been studied[15,17,19,20] though a complete description is still lacking. The representations of G, which are multiplicity-free on reduction to the maximal compact subgroup $K=\overline{\mathrm{SO}}(4)=\mathrm{SU}(2)\times\mathrm{SU}(2)$, have a particularly simple form and were constructed explicitly.[15,19]

We use the basis $\{j_i^{(\chi)}|\chi=1,2;\ i=1,2,3\}$ for k, with

$$[j_i^{(\chi)},j_j^{(\eta)}]=i\delta_{\chi\eta}\epsilon_{ijk}j_k^{(\chi)}. \tag{4.1}$$

The remaining (noncompact) matrices in sl(4,$\mathbb{R}$) transform as the irreducible tensor operator Z of type (1,1) under K:

$$\begin{aligned}&[j_i^{(1)},Z_{jk}]=i\epsilon_{ijm}Z_{mk},\\&[j_i^{(2)},Z_{jk}]=i\epsilon_{ikm}Z_{jm},\\&[Z_{ij},Z_{km}]=-i(\delta_{jm}\epsilon_{ikn}j_n^{(1)}+\delta_{ik}\epsilon_{jmn}j_n^{(2)}).\end{aligned} \tag{4.2}$$

We also have the spherical basis, given in terms of the above Cartesian basis by

$$\begin{aligned}&j_0^{(\chi)}=j_3^{(\chi)},\quad j_\pm^{(\chi)}=j_1^{(\chi)}\pm ij_2^{(\chi)},\\&Z_{00}=Z_{33},\\&Z_{\pm1,0}=\mp(1/\sqrt{2})(Z_{13}\pm iZ_{23}),\\&Z_{0,\pm1}=\mp(1/\sqrt{2})(Z_{31}\pm iZ_{32}),\\&Z_{1,\pm1}=\pm\tfrac{1}{2}[(Z_{11}\mp Z_{22})+i(Z_{21}\pm Z_{12})].\end{aligned} \tag{4.3}$$

The commutation relations (4.2) become

$$\begin{aligned}&[j_0^{(1)},Z_{\alpha\beta}]=\alpha Z_{\alpha\beta},\quad [j_0^{(2)},Z_{\alpha\beta}]=\beta Z_{\alpha\beta},\\&[j_\pm^{(1)},Z_{\alpha\beta}]=(2-\alpha(\alpha\pm1))^{1/2}Z_{\alpha\pm1,\beta}\\&\qquad=\sqrt{2}\,Z_{\alpha\pm1,\beta}(1-\delta_{\pm1,\alpha}),\\&[j_\pm^{(2)},Z_{\alpha\beta}]=\sqrt{2}\,z_{\alpha,\beta\pm1,\beta}(1-\delta_{\pm1,\beta})\quad(\alpha,\beta=0,\pm1),\end{aligned} \tag{4.4}$$

with the remaining ones following from the so-called "sl(4,$\mathbb{R}$) condition"

$$[Z_{11},Z_{-1-1}]=-(j_0^{(1)}+j_0^{(2)}). \tag{4.5}$$

It is convenient to introduce here too, the basis used by Mickelsson[16] for gl(4,$\mathbb{R}$):

$$L_{rs}=e_{rs}-e_{sr},\quad A_{rs}=e_{rs}+e_{sr},\quad r,s=1,2,3,4, \tag{4.6}$$

where e_{rs} is the 4×4 matrix with 1 in the r,s position and all other elements zero. The L_{rs} span $k=\mathrm{so}(4)$, and we have

$$\begin{aligned}&[L_{rs},L_{tu}]=\delta_{st}L_{ru}-\delta_{rt}L_{su}-\delta_{su}L_{rt}+\delta_{ru}L_{st},\\&[L_{rs},A_{tu}]=\delta_{st}A_{ru}-\delta_{rt}A_{su}+\delta_{su}A_{rt}-\delta_{ru}A_{st},\\&[A_{rs},A_{tu}]=\delta_{st}L_{ru}+\delta_{rt}L_{su}+\delta_{su}L_{rt}+\delta_{ru}L_{st}.\end{aligned} \tag{4.7}$$

We also put

$$\begin{aligned}&L_k=L_{ij}+L_{k4}=2ij_k^{(1)},\quad M_k=L_{ij}-L_{k4}=2ij_k^{(2)},\\&\{ijk\}\text{ a cyclic permutation of }\{123\}.\end{aligned} \tag{4.8}$$

To construct multiplicity-free representations we take the subspace V of $L^2(K)$ with orthonormal basis

$$\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=[(2j_1+1)(2j_2+1)]^{1/2}D^{j_1}_{0m_1}D^{j_2}_{0m_2}. \tag{4.9}$$

Then, in the spherical basis,

$$\begin{aligned}&j_0^{(1)}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=m_1\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle,\\&j_0^{(2)}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=m_2\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle,\\&j_\pm^{(1)}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle\\&\quad=[j_1(j_1+1)-m_1(m_1\pm1)]^{1/2}\left|\begin{matrix}j_1 & j_2\\ m_1\pm1 & m_2\end{matrix}\right\rangle,\\&j_\pm^{(2)}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle\\&\quad=[j_2(j_2+1)-m_2(m_2\pm1)]^{1/2}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\pm1\end{matrix}\right\rangle,\end{aligned} \tag{4.10}$$

while the noncompact operators $Z_{\alpha\beta}$ are given by

$$\begin{aligned}&\left\langle\begin{matrix}j_1' & j_2'\\ m_1' & m_2'\end{matrix}\right|Z_{\alpha\beta}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle\\&\quad=(-1)^{j_1-m_1'+j_2-m_2'}\begin{pmatrix}j_1' & 1 & j_1\\ -m_1' & \alpha & m_1\end{pmatrix}\begin{pmatrix}j_2' & 1 & j_2\\ -m_2' & \beta & m_2\end{pmatrix}\\&\quad\times\langle j_1'j_2'||Z||j_1j_2\rangle.\end{aligned} \tag{4.11}$$

The reduced matrix elements are

$$\begin{aligned}&\langle j_1'j_2'||Z||j_1j_2\rangle\\&\quad=-i(-1)^{j_1'+j_2'}[(2j_1'+1)(2j_2'+1)(2j_1+1)\\&\qquad\times(2j_2+1)]^{1/2}(p_1+ip_2-\tfrac{1}{2}[j_1'(j_1'+1)\\&\qquad-j_1(j_1+1)+j_2'(j_2'+1)-j_2(j_2+1)])\\&\qquad\times\begin{pmatrix}j_1' & 1 & j_1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}j_2' & 1 & j_2\\ 0 & 0 & 0\end{pmatrix}\end{aligned} \tag{4.12}$$

and clearly they are nonzero only for the four possibilities

$$j_1'=j_1\pm1,\quad j_2'=j_2\pm1.$$

Strictly speaking, from (4.9), the values of j_1,j_2 should only be 0,1,2,.... But at this stage we can formally continue (4.10) and (4.11) to half-integer values of j_1,j_2 as well. The sl(4,$\mathbb{R}$) condition (4.5) must be rechecked. One can proceed to find the complete set of *all* the unitary irreducible multiplicity-free representations $\overline{\mathrm{SL}}(4,\mathbb{R})$ (see Ref. 15 and 19). We shall only need some of these representations. First, we have that class, belonging to the discrete series, which is *spinorial*: i.e., double valued for SL(4,$\mathbb{R}$), and quadruple valued for $\overline{\mathrm{SO}}(3,3)$ [note that $\mathrm{SL}(4,\mathbb{R})=\overline{\mathrm{SO}}(3,3)$, but single valued for $\overline{\mathrm{SL}}(4,\mathbb{R})$. Their K content has the structure of a triangular lattice (p_1,p_2 are Casimir invariants):

$$\begin{aligned}&D^{\text{discrete}}(j_1^0,j_2^0):\quad j_1^0=p_1+1,\quad j_1^0=0,\\&\qquad\qquad\text{or}\\&\qquad\qquad j_2^0=0,\qquad j_2^0=p_1+1,\\&p_1=-\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2},\tfrac{5}{2},\ldots,\quad p_2=0,\\&|j_1-j_2|>p_1+1.\end{aligned} \tag{4.13}$$

Second, we want to mention the ladder series[12] of tensorial bandors, i.e., single-valued for SL(4,$\mathbb{R}$), with K content as follows[19]:

$$\mathscr{D}^{\text{ladd}}(0,0;p_2):\{(0,0),(1,1),(2,2),\dots\},$$
$$\mathscr{D}^{\text{ladd}}(\tfrac{1}{2},\tfrac{1}{2};p_2)\{(\tfrac{1}{2},\tfrac{1}{2}),(\tfrac{3}{2},\tfrac{3}{2}),(\tfrac{5}{2},\tfrac{5}{2}),\dots\}, \tag{4.14}$$
$$p_1=-1,\quad p_2\in\mathbf{R}.$$

The second of these, rather surprisingly, turns out to be relevant for manifield equations (see Sec. VI). This representation was constructed in Ref. 20 in solving the strong-coupling model for the nucleon, for the value $p_2=0$.

V. SL(2,C)-INVARIANT WAVE EQUATIONS

We now turn to a consideration of some wave equations appropriate for the gravitational interactions of hadrons. The general type of equation we have in mind is (in momentum space, with no gravitational field yet present)

$$(X^{\mu}p_{\mu}-\kappa)\psi(p)=0, \tag{5.1}$$

where ψ takes its values in a Hilbert space V carrying a unitary multiplicity-free representation π of $\overline{\mathrm{SL}}(4,\mathbf{R})$, and κ is an SL(2,C)-invariant operator on V, possibly a function of $p^2=p^{\mu}p_{\mu}$ (this generality is sometimes needed when we look for realistic mass spectra—see Sec. VI). The X^{μ} ($\mu=0,1,2,3$) are linear operators on V. We demand only SL(2,C) invariance as discussed in Sec. III, so the X^{μ} transform as an SL(2,C) vector. Physically, we want an equation which provides a kind of "extended" Dirac field.

At this stage we are confronted by various choices, namely, (a) which is the "physical" Lorentz subgroup of $\overline{\mathrm{SL}}(4,\mathbf{R})$ and (b) how is it embedded? These points are by no means trivial, as we shall see.

Our embedding of $K=\overline{\mathrm{SO}}(4)$—as well as $\overline{\mathrm{SO}}(3,1)$—in $\overline{\mathrm{SL}}(4,\mathbf{R})$ has been the natural one, described by the Lie algebra branching rules sl(4,R)→so(4) or sl(2,C):

$$\begin{aligned}&\text{defining representation}\rightarrow(\tfrac{1}{2},\tfrac{1}{2}),\\ &\text{adjoint representation}\rightarrow(1,0)\oplus(0,1)\oplus(1,1).\end{aligned} \tag{5.2}$$

Since our representation π of $\overline{\mathrm{SL}}(4,\mathbf{R})$ is K finite, i.e., on reduction to K it contains the representation (j_1,j_2), but a finite number of times, it is most natural[15,16] to take the quantum numbers (j_1,j_2) to refer to the *physical* Lorentz group. This means using $x^4=ix^0$. The non-Hermiticity of these "physical" Lorentz generators does not affect the physics, as explained in Sec. I. For this solution the Lorentz boosts will again be purely orbital and contribute to the kinetic energy only. All of this is perfectly respectable, since only finite-dimensional representations of K are involved. But if we had taken directly the SL(2,C) subgroup then[16] π would not contain *any* finite-dimensional representations of SL(2,C); this case is usually ignored.

An important property of embedding (5.2) is that we must look *outside* sl(4,R) to find the required K vector X^{μ}. This is our first approach to wave equations, which is further discussed in this section. We refer to (5.2) as the *natural* embedding. It is based on an automorphism proved in Ref. 19.

However, there is a second approach, suggested to us by the case of the Majorana representations of SL(2,C). There is an embedding of SL(2,C) in SL(4,R) obtained via the Dirac representation:

$$\begin{aligned}&\mathrm{sl}(4,\mathbf{R})\rightarrow\mathrm{sl}(2,\mathbf{C}),\\ &\text{defining representation}\rightarrow(\tfrac{1}{2},0)\oplus(0,\tfrac{1}{2}),\\ &\text{adjoint representation}\rightarrow(1,0)\oplus(0,1)\oplus2(\tfrac{1}{2},\tfrac{1}{2})\oplus(0,0).\end{aligned} \tag{5.3}$$

[We shall show later that $\overline{\mathrm{SO}}(4)$ *cannot* be so embedded.] Now everything is quite different: we have two linearly independent SL(2,C) vectors $\gamma^{\mu},\gamma^{\mu'}$ in SL(4,R). So we can obtain automatically an SL(2,C)-invariant equation suitable for our purpose simply by taking an irreducible representation π of $\overline{\mathrm{SL}}(4,\mathbf{R})$.

It is important to realize that (5.3) does not provide a direct embedding of SO(3,1) in SL(4,R). Instead, SO(3,1) is embedded in SO(3,3) and SL(2,C) in $\mathrm{SL}(4,\mathbf{R})=\overline{\mathrm{SO}}(3,3)$.

We shall discuss this possibility (the *Dirac* embedding) in Sec. VI.

Let us come back to embedding (5.2). The condition that X^{μ} can be a K vector is

$$[L_{rs},X_t]=\delta_{st}X_r-\delta_{rt}X_s. \tag{5.4}$$

To express (5.4) in a form convenient for applying angular momentum algebra, we define the quantities X_{AB} ($A,B=\pm\frac{1}{2}$) by

$$(X_{AB})=\begin{pmatrix}X_1+iX_2 & -X_3-iX_4\\ -X_3+iX_4 & -X_1+iX_2\end{pmatrix}. \tag{5.5}$$

Then we see that the X_{AB} transform like the canonical basis for the K-vector representation $(\frac{1}{2},\frac{1}{2})$:

$$\begin{aligned}&[j_0^{(1)},X_{AB}]=A\,X_{AB},\quad [j_0^{(2)},X_{AB}]=B\,X_{AB},\\ &[j_{\pm}^{(1)},X_{AB}]=X_{A\pm1,B},\quad [j_{\pm}^{(2)},X_{AB}]=X_{A,B\pm1}.\end{aligned} \tag{5.6}$$

It is well known from the theory of Lorentz-invariant wave equations[21,22] that, in a candidate representation π of $\overline{\mathrm{SL}}(4,\mathbf{R})$, the matrix elements of X_{AB} are given by[23]

$$\begin{aligned}&\left\langle\begin{matrix}j_1' & j_2'\\ m_1' & m_2'\end{matrix}\middle|X_{AB}\middle|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle\\ &=(-1)^{j_1'-m_1'+j_2'-m_2'}\begin{pmatrix}j_1' & \frac{1}{2} & j_1\\ -m_1' & A & m_1\end{pmatrix}\\ &\times\begin{pmatrix}j_2' & \frac{1}{2} & j_2\\ -m_2' & B & m_2\end{pmatrix}\langle j_1'\,j_2'\rangle\,\langle j_1 j_2\rangle.\end{aligned} \tag{5.7}$$

They are nonzero only for the four possibilities $j_1'=j_1\pm\frac{1}{2}$, $j_2'=j_2\pm\frac{1}{2}$. It is immediately clear that, among the representations (4.13), the *only* possible unitary multiplicity-free spinorial representation π of $\overline{\mathrm{SL}}(4,\mathbf{R})$ that admits a K vector is the (reducible) combination suggested in Ref. 15. We have that

$$\pi=\mathscr{D}^{\text{disc}}(\tfrac{1}{2},0)\oplus\mathscr{D}^{\text{disc}}(0,\tfrac{1}{2}) \tag{5.8}$$

(so in each case $p_1=-\frac{1}{2}$, $p_2=0$), with the K content shown in Fig. 1. Here the dark (white) circles refer to $\mathscr{D}^{\text{disc}}(0,\frac{1}{2})(\mathscr{D}^{\text{disc}}(\frac{1}{2},0))$, and the only nonzero matrix elements of X_{AB} are between the K representations $(j,j+\frac{1}{2})$ and $(j+\frac{1}{2},j)$ ($j=0,1,2,\dots$), i.e., across the diagonal.

We want to remark here that our multiplicity-free representations do not allow the existence of a $\overline{\mathrm{SL}}(4,\mathbf{R})$ vector. The proof is given in Appendix A. Now the operators X^k in Mickelsson's wave equation[16] do transform as $\overline{\mathrm{SL}}(4,\mathbf{R})$ vec-

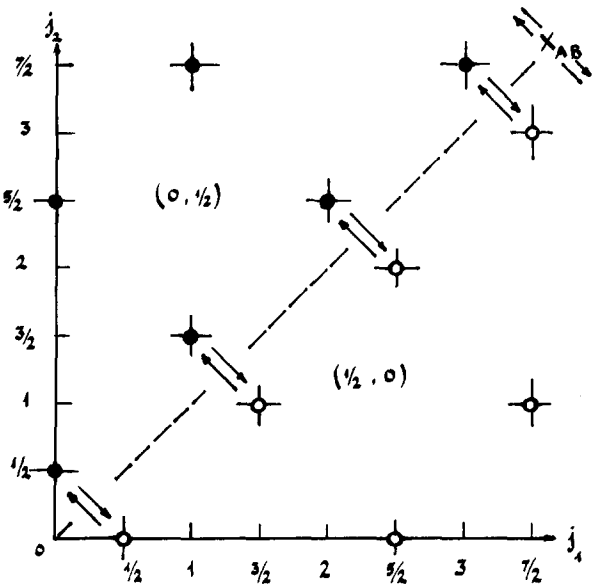


FIG. 1. Action of $X(\frac{1}{2},\frac{1}{2})$ on $\mathscr{D}^{\text{disc}}(\frac{1}{2},0) \oplus \mathscr{D}^{\text{disc}}(0,\frac{1}{2})$.

tor. However, his representations of $\overline{\text{SL}}(4,\mathbf{R})$ are not multiplicity-free, and the argument of Appendix A is no longer valid.

For our equation we write the reduced matrix elements in (5.7) as[24]

$$a_j = \langle j, j+\tfrac{1}{2} | \, |X| \, | j+\tfrac{1}{2}, j\rangle,$$
$$b_j = \langle j+\tfrac{1}{2}, j | \, |X| \, | j, j+\tfrac{1}{2}\rangle, \quad j = 0,1,2,\ldots . \tag{5.9}$$

The a_j and b_j can be arbitrary complex numbers. Thus, strictly speaking, we have a *family* of wave equations, each one described by a particular choice of these coupling constants (assumed nonzero). As far as SL(2,C) properties are concerned, each such system is an infinite set of decoupled equations for successively higher half-integral spins. Each constituent $(j, j+\frac{1}{2}) \rightleftharpoons (j+\frac{1}{2}, j)$ in general has the $2j+1$ spins: $2j+\frac{1}{2}, 2j-\frac{1}{2}, \ldots, \frac{1}{2}$. The gravitational field, in the form of the noncompact shear operators $Z_{\alpha\beta}$, will couple between these constituents, and also throw up new K representations so that altogether we recover the representation (5.8) of $\overline{\text{SL}}(4,\mathbf{R})$.

Although we are not concerned here with the Lie algebraic properties of the vector operator X^μ, we note[21] that the Lie algebra generated by the X^μ and k will be, for almost all choices of a_j, b_j,

$$\text{sp}(4,\mathbf{C}) \oplus \text{sp}(24,\mathbf{C}) \oplus \cdots \oplus \text{sp}(2(2j+1)(2j+2),\mathbf{C}) \oplus \cdots .$$

Including the $Z_{\alpha\beta}$ will no doubt generate an infinite-dimensional Lie algebra.

The mass spectrum, too, depends on the choice of a_j, b_j. Two equations for which the quantities a_j, b_j, $j = 0,1,2,\ldots$, coincide clearly have the same spectrum. The spectrum is given by

$$m = \kappa/\lambda \quad (\lambda \text{ a nonzero eigenvalue of } X^0), \tag{5.10}$$

if κ in (5.1) is a constant. More realistic mass spectra appear possible, as in (6.39) or (6.42). Spins coupled to zero eigenvalues of X^0 are excluded,[21] as they would have (infinite) unphysical masses. This may imply a need for subsidiary constraints.

Since (5.8) belongs to the double-covering $\overline{\text{SL}}(4,\mathbf{R})$, this manifield, though constructed so as to couple anholonomically to gravity, may also have a holonomic version.

VI. A WAVE EQUATION BASED ON THE DIRAC EMBEDDING

In this section we want to study the possible Lorentz-invariant wave equations obtained by considering suitable representations π of $\overline{\text{SL}}(4,\mathbf{R})$, where SL(2,C) is embedded according to (5.3). First of all we shall write down the (Lie algebra) embedding explicitly, directly using results of Ref. 25, where a general study was made of those real Lie algebras containing sl(2,C) and a vector operator.

The starting point is embedding of the *compact* algebras

$$\text{su}(4) \supseteq \text{su}(2) \oplus \text{su}(2),$$

provided by the Dirac representation $(\frac{1}{2},0) \oplus (0,\frac{1}{2})$. We use the fact that,[25,26] if g_0 is a real form of sl(4,C), obtained via the Weyl "unitary trick" from the involutive automorphism s (say) of su(4), then sl(2,C) is a subalgebra of g_0 if and only if

$$s(X,Y) = (Y,X), \quad \forall (X,Y) \in \text{su}(2) \oplus \text{su}(2). \tag{6.1}$$

From Ref. 25 we have the following result: sl(2,C) is embedded in sl(4,R), and s is the (outer) automorphism

$$s{:}\text{su}(4) \to \text{su}(4),$$
$$X \to N\bar{X}N^{-1} = -NX^{T}N^{-1}. \tag{6.2}$$

Note: su(4) consists of skew Hermitian matrices. The matrix $N \in \text{SU}(4)$ may be taken to be

$$N = \begin{pmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}. \tag{6.3}$$

One can check that s satisfies (6.1). But how do we know that the resulting real form is sl(4,R) [and not su(2,2) or su*(4), for example]? The reason is that[25]

$$U^{-1}N\bar{U} = I,$$

where

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} I & iI \\ -i\sigma^2 & -\sigma^2 \end{pmatrix} \in U(4), \tag{6.4}$$

and thus $s = \alpha^{-1}\theta\alpha$, where θ,α ae the automorphisms given by

$$\theta{:}\, X \to \bar{X}, \quad \alpha{:}\, X \to U^{-1}XU. \tag{6.5}$$

Since s is conjugate to θ, and θ clearly gives the real form sl(4,R) with the Cartan decomposition

$$\text{sl}(4,\mathbf{R}) \equiv k' \oplus p' = \text{so}(4) \oplus \{\text{real symmetric matrices}\}, \tag{6.6}$$

we see that we indeed have a realization of sl(4,R). Our Cartan decomposition is given by

$$g_0 = k \oplus p, \tag{6.7}$$

where the maximal compact subalgebra is

$$k=\left\{\left.\begin{pmatrix} ia & \beta & \gamma & 0 \\ -\bar\beta & ib & 0 & \gamma \\ -\bar\gamma & 0 & -ib & \beta \\ 0 & -\bar\gamma & -\bar\beta & ia \end{pmatrix}\right| a,b\in\mathbf{R};\beta,\gamma\in\mathbf{C}\right\}=\alpha^{-1}k'=Uk'U^{-1} \tag{6.8}$$

[isomorphism to so(4)]; while the noncompact generators are

$$p=\left\{\left.\begin{pmatrix} c & \delta & \tau & \theta \\ \bar\delta & -c & \eta & -\tau \\ \bar\tau & \bar\eta & -c & -\delta \\ \bar\theta & -\bar\tau & -\bar\delta & c \end{pmatrix}\right| c\in\mathbf{R};\delta,\tau,\theta,\eta\in\mathbf{C}\right\}=\alpha^{-1}p'=Up'U^{-1}, \tag{6.9}$$

which means that our realization is

$$g_0=\left\{\left.\begin{pmatrix} A & B \\ -\sigma^2\bar B\sigma^2 & \sigma^2\bar A\sigma^2 \end{pmatrix}\right| A,B\in\mathrm{gl}(2,\mathbf{C});\ \mathrm{Re\ Tr}(A)=0\right\}=\alpha^{-1}(\mathrm{sl}(4,\mathbf{R}))=U\,[\mathrm{sl}(4,\mathbf{R})]\,U^{-1}. \tag{6.10}$$

This realization is somewhat strange, but we can go over to the more familiar one by applying the isomorphism α. We shall always do this since we want to compare our embedding with the more familiar case of (5.2).

Our embedding proceeds via sp(4,R) as follows. For the compact algebra,

$$\mathrm{su}(2)\oplus\mathrm{su}(2)\subseteq\mathrm{usp}(4)=\mathrm{u}(4)\cap\mathrm{sp}(4,\mathbf{C}), \tag{6.11}$$

where

$$\mathrm{sp}(4,\mathbf{C})=\{X\in\mathrm{sl}(4,\mathbf{C})|BXB^{-1}=-X^T\} \tag{6.12}$$

and

$$B=\begin{pmatrix} -\sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}.$$

Then sl(2,C) is embedded in sp(4,R) (see Ref. 25). We take the automorphism $\tilde s$ of usp(4) given by

$$\tilde s\colon X\to MXM^{-1}, \tag{6.13}$$

where

$$M=i\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\in\mathrm{USp}(4)\quad(M^2=-I). \tag{6.14}$$

The resulting real form $\tilde g$ has Cartan decomposition

$$\tilde g=\tilde k\oplus\tilde p, \tag{6.15}$$

where

$$\tilde k=\left\{\left.\begin{pmatrix} -ia & \beta & id & 0 \\ -\bar\beta & -ia & 0 & id \\ id & 0 & ia & \beta \\ 0 & id & -\bar\beta & -ia \end{pmatrix}\right| a,d\in\mathbf{R};\beta\in\mathbf{C}\right\}\subseteq k \tag{6.16}$$

and

$$\tilde p=\left\{\left.\begin{pmatrix} b & \gamma & ic & \delta \\ -\bar\gamma & -b & \bar\delta & -ic \\ -ic & -\delta & -b & -\gamma \\ -\bar\delta & ic & -\bar\gamma & b \end{pmatrix}\right| b,c\in\mathbf{R}\quad\gamma,\delta\in\mathbf{R}\right\}\subseteq p. \tag{6.17}$$

Here $\tilde g$ is a realization of sp(4,R) contained in our realization g_0 of sl(4,R); the maximal compact subalgebra $\tilde k$ is isomorphic to u(2). Clearly s, as given by (6.2), is an extension of $\tilde s$, because if $X\in\mathrm{usp}(4)$

$$\begin{aligned} s(X)&=-NX^{T}N^{-1}\\ &=+NBXB^{-1}N^{-1}\\ &=MXM^{-1}=\tilde s(X)\quad(\text{since } NB=M). \end{aligned}$$

The isomorphism α given by (6.5) takes $\tilde g$ to

$$\mathrm{sp}'(4,\mathbf{R})=\{X\in\mathrm{sl}(4,\mathbf{R})|B'XB'^{-1}=-X^T\},$$

where $B'=U^TBU$, (6.18)

the more familiar realization.

We notice that so(4) *cannot* be embedded in sl(4,R) via the Dirac representation, since any two maximal compact subalgebras of sl(4,R) are conjugate under some automorphism, and so(4) is aleady embedded via the natural representation $(\frac12,\frac12)$. It is interesting to see how this result appears if we ask the general question: which real forms g_0 of sl(4,C) contain so(4) embedded via the Dirac representation? We discuss this in Appendix B: it turns out that su(2,2) and su*(4) are the only possibilities.

Coming back to our embedding $sl(2,\mathbb{C}) \subseteq sp(4,\mathbf{R}) \subseteq sl(4,\mathbf{R})$, we write the $sl(2,\mathbb{C})$ generators of rotations and Lorentz boosts as

$$H^k = \begin{pmatrix} i\sigma^k & 0 \\ 0 & i\sigma^k \end{pmatrix}, \quad F^k = \begin{pmatrix} -\sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \tag{6.19}$$

Using the Dirac matrices in the form

$$\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & -\sigma^k \\ \sigma^k & 0 \end{pmatrix}, \tag{6.20}$$

we introduce another vector operator, given by

$$\gamma^{0'} = -\gamma^5\gamma^0 = \begin{pmatrix} 0 & iI \\ -iI & 0 \end{pmatrix},$$
$$\gamma^{k'} = \gamma^5\gamma^k = \begin{pmatrix} 0 & i\sigma^k \\ i\sigma^k & 0 \end{pmatrix}, \tag{6.21}$$

where

$$\gamma^5 = \gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -iI & 0 \\ 0 & iI \end{pmatrix}.$$

Then, for $sp(4,\mathbf{R})$, we see that $\tilde{k}$ has been $\{H^k, i\gamma^0\}$ and $\tilde{p}$ has basis $\{F^{k'}, i\gamma^{k'}\}$. The remaining generators of $sl(4,\mathbf{R})$ are $i\gamma^{0'}$, $\gamma^5 \in k$, $i\gamma^{k'} \in p$. Note that $sl(4,R)$ contains two vector operators as expected from (5.3). But only *one* of these—in this case $i\gamma^\mu$—belongs to $sp(4,\mathbf{R})$, once the skew-symmetric form B is fixed.

Under the isomorphism $\alpha: g_0 \to sl(4,\mathbf{R})$ we have (in the notation of Sec. IV)

$$\begin{aligned} &H^1 \to -L_1, \quad H^2 \to L_3, \quad H^3 \to L_2, \\ &i\gamma^{0'} \to M_1, \quad i\gamma^{0'} \to M_3, \quad \gamma^5 \to -M_2, \end{aligned} \tag{6.22}$$

for the compact generators and

$$\begin{aligned} &F^1 \to Z_{12}, \quad F^2 \to -Z_{32}, \quad F_3 \to -Z_{22}, \\ &i\gamma^1 \to Z_{13}, \quad i\gamma^2 \to -Z_{33}, \quad i\gamma^3 \to -Z_{23}, \\ &i\gamma^{1'} \to Z_{11}, \quad i\gamma^{2'} \to -Z_{31}, \quad i\gamma^{3'} \to -Z_{21}, \end{aligned} \tag{6.23}$$

for the noncompact ones.

Notice that in our setup the physically relevant su(2) subalgebra is that spanned by **L**: in the approach taken in Refs. 15 and 16 it is that spanned by **L** + **M**. Also, the maximal compact subalgebra $k \cong so(4)$ has no physical role; though it is still mathematically relevant in the study of the representations of $sl(4,\mathbf{R})$. Again we stress that γ^μ, $\gamma^{\mu'}$ are *not* vectors under the so(4) subalgebra, but under the noncompact subalgebra $sl(2,\mathbb{C}) \cong so(3,1)$. It is also clear that $sl(3,\mathbf{R})$ does not fit into our scheme in such a way that its maximal compact subalgebra so(3) is spanned by **L**.

The Lie algebra $sl(4,\mathbf{R})$ is isomorphic to so(3,3). We can easily write the so(3,3) generators in terms of Dirac matrices as follows. Introducing the notation

$$\gamma^m = \gamma^1, \gamma^2, \gamma^3, -i\gamma^5, \gamma^0, -iI), \quad m = 1,2,3,4,5 \equiv 0,6, \tag{6.24}$$

we put

$$Q^{mn} = \tfrac{1}{2}\gamma^m\gamma^n. \tag{6.25}$$

Then we have the commutation relations of so(3,3),

$$[Q^{mn}, Q^{pq}] = g^{np}Q^{mq} - g^{mp}Q^{nq} - g^{nq}Q^{mp} + g^{mq}Q^{np}, \tag{6.26}$$

where the metric is

$$g^{mn} = \mathrm{diag}(-1,-1,-1,1,1,1). \tag{6.27}$$

The $sl(4,\mathbf{R})$ generators are identified as follows:

$$\begin{aligned} &Q^{ij} = \tfrac{1}{2}\gamma^i\gamma^j = -H^k \quad (ijk: \text{cyclic permutation of } \{1\,2\,3\}), \\ &Q^{i4} = \tfrac{1}{2} i\gamma^5\gamma^i = (i/2)\gamma^i, \\ &Q^{i5} = \tfrac{1}{2}\gamma^i\gamma^0 = -\tfrac{1}{2}F^i, \\ &Q^{\mu 6} = -\tfrac{1}{2} i\gamma^\mu, \\ &Q^{45} = -\tfrac{1}{2} i\gamma^5\gamma^0 = (i/2)\gamma^{0'}, \\ &Q^{46} = -\tfrac{1}{2}\gamma^5. \end{aligned} \tag{6.28}$$

These formulas are analogous to Barut's[27] four-dimensional realization of so(4,2); the only difference is that he takes γ^4 to be γ^5 and $g^{mn} = \mathrm{diag}(-1,-1,-1,-1,1,1)$.

It is interesting to compare our approach with Barut's theory[27,28] of the hadron spectrum using SO(4,2): in both cases $sl(2,\mathbb{C})$ is embedded via the Dirac representation. Barut was led to so(4,2) by the well-known properties of the hydrogen atom, which has a so(4) kinematical symmetry. We have the spectrum-generating algebra $sl(4,\mathbf{R}) \cong so(3,3)$. Kihlberg[29] has, in fact, suggested using so(3,3) for hadrons, with the maximal compact subalgebra $so(3) \oplus so(3)$ interpreted as the sum of spin and isospin algebras. In our approach, however, using $\overline{SL}(4,\mathbf{R})$, we have the gauge group of gravity naturally appearing. This is why we can speak of the gravitational interaction of hadrons.

Now we can produce Lorentz-invariant wave equations of the form (5.1), suitable for the description of the gravitational interactions of hadrons. One can say that our equations are extensions of Dirac's equation, since we used the Dirac representation of $sl(2,\mathbb{C})$. If we fix the vector operator to be γ^μ, then the equation is parity invariant. The parity operator P is essentially[25] the M of (6.14); it singles out the real form $sp(4,\mathbf{R})$. Parity invariance means that $i\gamma^0 \in \tilde{k}$. In the same way charge conjugation C is essentially the N of (6.3); it gives the real form $sl(4,\mathbf{R})$ and charge conjugation invariance means that $i\gamma^0 \in k$.

We have enlarged the $sp(4,\mathbf{R})$ algebra—whose ladder representations give the Majorana equations—to all of $sl(4,\mathbf{R})$, by taking the algebra generated by all the products of γ matrices (not just the commutators $[\gamma^\mu, \gamma^\nu]$, which close on $sp(4,R)$). Another way[9] of obtaining $sl(4,R)$ from Dirac's equation

$$(\gamma^\mu P_\mu - M)\,\psi(p) = 0 \tag{6.29}$$

is to let the mass term M be proportional to γ^5, and then take commutators of the γ^μ and M.

We can now take one of the unitary irreducible representations of $\overline{SL}(4,\mathbf{R})$ given by (4.13) to obtain the Lorentz-invariant wave equation

$$(\pi(\gamma^\mu p_\mu - \kappa)\,\psi(p) = 0, \tag{6.30}$$

where ψ takes its values in the Hilbert space V of the representation π. We could take κ to be $\pi(M)$ as the simplest approach, or even a general Lorentz-invariant operator-valued function of p^2.

Since the physical spin su(2) subalgebra is that spanned by **L**, the spin content for each representation π is easily

obtained: for each K representation (j_1,j_2) appearing, we have $(2j_2+1)$ copies of the SU(2) representation j_1. Clearly, in the present context of the Dirac embedding (5.3) the appropriate spinorial representations π are those that contain K representations (j_1,j_2) with half-integer j_1.

There are thus two candidate multiplicity-free representations (the method can, of course, be extended to non-multiplicity-free representations as well) with lowest spin $\frac{1}{2}$: (i) $\mathscr{D}^{\text{disc}}(\frac{1}{2},0)$ with spin content

$$(\tfrac{1}{2}) \oplus 3(\tfrac{3}{2}) \oplus 6(\tfrac{5}{2}) \oplus 10(\tfrac{7}{2}) \oplus \cdots, \tag{6.31}$$

(ii) $\mathscr{D}^{\text{ladd}}(\frac{1}{2},\frac{1}{2};p_2)$, $p_2 \in \mathbb{R}$, with spin content

$$2(\tfrac{1}{2}) \oplus 4(\tfrac{3}{2}) \oplus 6(\tfrac{5}{2}) \oplus \cdots. \tag{6.32}$$

Since $i\gamma^0 \in k$, integer values of j_2 as in $\mathscr{D}^{\text{disc}}(\frac{1}{2},0)$ may involve self-charge-conjugate states for zero eigenvalues λ of $i\gamma^0$. The representation $\mathscr{D}^{\text{ladd}}(\frac{1}{2},\frac{1}{2};p_2)$ on the other hand is symmetric in positive and negative energy states, like Dirac's spinor.

The Dirac embedding (5.3) is an embedding SO(3,1) $\subset$ SO(3,3) or SL(2,C) $\subset$ SL(4,$\mathbb{R}$). This is why the spinor nature of the equation and particles is not correlated with the single or double valuedness of the SL(4,$\mathbb{R}$) representation. For gravity, the Dirac embedding produces an anholonomic spinor and cannot be utilized for a holonomic ("world") spinor (see our discussion in Secs. I–III).

We would also like to know the SL(2,C) and Sp(4,$\mathbb{R}$) reduction, but this is not readily available from our infinitesimal approach. Certainly we have a direct sum of (infinite-dimensional) unitary irreducible representations: for example, we conjecture that the SL(2,C) decomposition of $\mathscr{D}^{\text{disc}}(\frac{1}{2},0)$ is

$$\{\tfrac{1}{2}, l^{(1)}\} \oplus 2\{\tfrac{3}{2}, l^{(2)}\} \oplus 3\{\tfrac{5}{2}, l^{(3)}\} \tag{6.33}$$

(in the notation of Ref. 21). We do not know what the labels $l^{(1)}$, $l^{(2)}$,... are. The first term in (6.33) may be the Majorana representation $\{\frac{1}{2},0\}$.

We are primarily interested in the mass spectrum of (6.3). Since $i\gamma^0$ belongs to the maximal compact subalgebra k of sl(4,R), there will be a discrete spectrum of rest masses (i.e., those corresponding to timelike momenta, $p^2>0$). It is easy to calculate the mass spectrum in a given case. First, we observe that, since

$$M_3 = T^{-1} M_1 T, \tag{6.34}$$

where

$$T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ -i & 0 & 0 & -1 \\ 0 & -1 & -i & 0 \end{bmatrix}, \tag{6.35}$$

we have, from (4.8),

$$\alpha(\gamma^0) = 2j_1^{(2)} = 2Tj_3^{(2)}T^{-1}. \tag{6.36}$$

Thus for a unitary multiplicity-free representation π of $\overline{\text{SL}}(4,\mathbb{R})$ we see from (4.1) that the spectrum of $\pi(\gamma^0)$ is given by

$$\lambda = 2m_2 = 2j_2, 2(j_2-1), \ldots, -2j_2, \tag{6.37}$$

where j_2 goes over all the su(2)$\times$su(2) representations (j_1,j_2) that occur in π. Note that for half-integer j_2 the equation is indeed symmetric in positive and negative energy states, like Dirac's equation. For integer j_2, $\pi(\gamma^0)$ will have one zero eigenvalue for each value of j_2.

The mass spectrum depends on the form of κ. If we take

$$\kappa = \beta I \quad (\beta \in \mathbb{R}), \tag{6.38}$$

then the spectrum of rest masses is given by

$$m = \beta/\lambda \quad [\lambda \text{ a nonzero eigenvalue of } \pi(\gamma^0)].$$

But this decreases as λ increases; states of higher m_2 and thus higher spins j_1 have a smaller mass as in the Majorana equation. It may be more realistic to take instead

$$\kappa = (\alpha p^2 + \beta) I, \quad \alpha, \beta \in \mathbb{R}. \tag{6.39}$$

So, choosing $\overset{0}{p} = (m,0,0,0)$, we have

$$(\pi(\gamma^0)m - \alpha m^2 - \beta)\, \psi(\overset{0}{p}) = 0,$$

i.e., the spectrum of rest masses is given by[27]

$$m[\lambda + (\lambda^2 - 4\alpha\beta)^{1/2}]/2\alpha, \tag{6.40}$$

which gives a better mass formula; in particular if $\beta = 0$ we get

$$m = \lambda/\alpha \tag{6.41}$$

and the mass is linear in λ. The observed Regge spectrum $m^2 \sim j_1$ with daughter trajectories is obtained by taking

$$\kappa = \{\alpha(p^2)^{3/2} + \beta\}\, I. \tag{6.42}$$

We observe that of the two "spinorial" equations (6.31) and (6.32), it is the ladder example that has nonsingular $\pi(\gamma^0)$, symmetric charge-conjugate (or negative-energy) states and can describe [with (6.42)] the physical mass spectrum. Its coupling to gravity is purely anholonomic and does not involve the double covering of SL(4,$\mathbb{R}$) and Δ.

APPENDIX A: LIMITATIONS ON X^ν AS $\overline{\text{SL}}(4,\mathbb{R})$ FOUR-VECTOR

In this appendix we shall show that, for the multiplicity-free representations other than (4.14), no $\overline{\text{SL}}(4,\mathbb{R})$ vector X^ν can be constructed (apart from the trivial case $X^\nu = 0$).

If X^ν is to be an $\overline{\text{SL}}(4,\mathbb{R})$ vector, then as well as (5.4), we must have

$$[A_{rs}, X_t] = \delta_{st} X_r + \delta_{rt} X_s. \tag{A1}$$

We can calculate the commutators $[Z_{\alpha\beta}, X_{AB}]$ in the spherical basis most simply by applying the Wigner–Eckart theorem for the tensor operator Z acting by commutation on the vector representation. Then the matrix elements are

$$\langle X_{CD} | Z_{\alpha\beta} | X_{AB} \rangle = 2i(-1)^{C+D} \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ -C & \alpha & A \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ -D & \beta & B \end{pmatrix} \tag{A2}$$

so that (A1) becomes

$$[Z_{\alpha\beta}, X_{AB}] = 2i(-1)^{A+B} \{(\tfrac{3}{4} - A(A+\alpha)) \times (\tfrac{3}{4} - B(B+\beta))\}^{1/2} X_{A+\alpha, B+\beta}, \tag{A3}$$

using the $3j$ symbols tabulated in Ref. 23. This result can also be obtained directly from (4.3) and (A1) if we use the relation

$$iZ_{ij} = \tfrac{1}{2}\delta_{ij}(A_{kk} - A_{44}) - (A_{ij} + \epsilon_{ijk}A_{k4}). \quad \text{(A4)}$$

Now if we take the commutator $[Z_{00}, X_{1/2\,1/2} = -iX_{1/2\,1/2}]$, for example, we see that the matrix elements

$$\left\langle \begin{matrix} j & j+\frac{1}{2} \\ m' & n' \end{matrix} \right| [Z_{00}, X_{1/2\,1/2}] \left| \begin{matrix} j+\frac{1}{2} & j \\ m & n \end{matrix} \right\rangle \quad \text{(A5)}$$

are zero because Z *never* couples (j_1, j_2) to itself in the representations (4.11) we have constructed except for (4.14) with $p_2 \neq 0$. Thus

$$\left\langle \begin{matrix} j & j+\frac{1}{2} \\ m' & n' \end{matrix} \right| X_{1/2\,1/2} \left| \begin{matrix} j+\frac{1}{2} & j \\ m & n \end{matrix} \right\rangle = 0, \quad \text{(A6)}$$

and since this is true for each direction of coupling in Fig. 1, we see that $X_{AB} \equiv 0$: no $\overline{\mathrm{SL}}(4,\mathbb{R})$ vector exists for our wave equation (5.8).

APPENDIX B: DIRAC EMBEDDING OF so(4) ⊂ sl(4,C)

Suppose that g_0 is a real form of sl(4,C) for which the maximal compact subalgebra k contains so(4) embedded via the Dirac representation. The g_0 arises from some involutive automorphism s of su(4) such that

$$s(X,Y) = (X,Y), \quad V(X,Y) \in \mathrm{su}(2) \oplus \mathrm{su}(2). \quad \text{(B1)}$$

There are two possibilities.

(a) If $s: X \to MXM^{-1}$ (inner) then (B1) gives

$$M = \begin{pmatrix} \alpha I & 0 \\ 0 & \beta I \end{pmatrix} \alpha^2 = \beta^2 = 1$$

and so

$$k = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \middle| \; A,B \in \mathrm{u}(2); \quad \mathrm{Tr}(A+B) = 0 \right\}$$

$$\cong \mathrm{su}(2) \oplus \mathrm{su}(2) \oplus \text{center of } k$$

and the real form is su(2,2). Since $M \in \mathrm{USp}(4)$ we have in fact an embedding $\mathrm{so}(4) \subseteq \mathrm{sp}(2,2)$ with $k \cong \mathrm{usp}(2) \oplus \mathrm{usp}(2)$.

(b) If s: $X \to N\bar{X}N^{-1}$, this gives

$$N = \begin{pmatrix} \alpha\sigma^2 & 0 \\ 0 & \beta\sigma^2 \end{pmatrix} \in \mathrm{SU}(4) \quad (\alpha^2 = \beta^2 = -1)$$

so $N\bar{N} = -|\alpha|^2 = -I$. Clearly $k = \mathrm{usp}(4)$ and this time the real form is su*(4).

So we can embed so(4) in either of these real forms; these possibilities do not concern us here. [Note that sl(4,ℝ) would have to come from $N\bar{N} = I$ in (b). This never happens.]

ACKNOWLEDGMENTS

We thank Professor S. Sternberg for his advice.

One of the authors (A.C.) is grateful to the Wolfson Chair Extraordinary of Theoretical Physics, Tel Aviv University, for its kind hospitality and financial support, and R. Kerner of the Département de Mécanique, Université de Paris, VI, for helpful discussions.

The other one of us (Y.N.) was supported in part by the U.S. Department of Energy Grant No. DE-AS05-76ER03992.

[1] See for example, Y. Ne'eman, *Algebraic Theory of Particle Physics* (Benjamin, New York, 1964), Chap. VIII.
[2] E. Majorana, Nuovo Cimento **9**, 335 (1932).
[3] D. T. Stoyanov and I. T. Todorov, J. Math. Phys. **9**, 2146 (1968).
[4] See Ref. 1, Chap. X.
[5] D. M. Fradkin, Am. J. Phys. **34**, 314 (1966).
[6] P. A. M. Dirac, in *Invited Papers, Tracts in Mathematics and Natural Sciences*, edited by M. Hamermesh (Gordon and Breach, New York, 1971), p. 1.
[7] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), Chap. XII, Sec. 5, p. 365.
[8] For a review, see F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester, Rev. Mod. Phys. **48**, 393 (1976).
[9] F. W. Hehl, G. D. Kerlick, and P. von der Heyde, Phys. Lett. B **63**, 446 (1976); Z. Naturforsch. A **31**, 111, 524, 823 (1976); E. A. Lord, Phys. Lett. A **65**, 1 (1978); F. W. Hehl, G. D. Kerlick, E. A. Lord, and L. L. Smalley, Phys. Lett. B **70**, 70 (1977); F. W. Hehl and G. D. Kerlick, GRG **9**, 691 (1978); F. W. Hehl, E. A. Lord, and L. L. Smalley, GRG **13**, 1037 (1981).
[10] See, for example, Ref. 7, p. 365; also, Ref. 8, p. 401, last paragraph of Sec. 1., etc.
[11] Y. Ne'eman, Ann. Inst. H. Poincaré A **28**, 369 (1978).
[12] H. Dothan, M. Gell-Mann, and Y. Ne'eman, Phys. Lett. **17**, 148 (1965).
[13] Y. Ne'eman, Proc. Nat. Acad. Sci. USA **74**, 4157 (1977).
[14] F. W. Hehl, E. A. Lord, and Y. Ne'eman, Phys. Lett. B **71**, 432 (1977); Phys. Rev. D **17**, 418 (1978).
[15] Y. Ne'eman and Dj. Šijački, Ann. Phys. (NY) **120**, 292 (1979); Proc. Nat. Acad. Sci. (USA) **76**, 561 (1979); **77**, 1761 (1980).
[16] J. Mickelsson, Comm. Math. Phys. **88**, 551 (1983).
[17] Y. Ne'eman, Found. Phys. **13**, 467 (1983). Note that the SL(n,R) have been studied intensively in mathematics (though their universal coverings have not). See, for example, D. A. Vogan, Jr., *Representations of Real Reductive Lie Groups* (Birkhauser, Boston, 1981); B. Speh, Math. Ann. **258**, 113 (1981).
[18] Dj. Šijački, J. Math. Phys. **16**, 298 (1975).
[19] Dj. Šijački and Y. Ne'eman, J. Math. Phys. **26**, 2457 (1985).
[20] Y. Dothan and Y. Ne'eman, in *Resonant Particles*, edited by B. A. Munir (Ohio Univ. Athens, Ohio, 1965), p. 17. Reprinted in *Symmetry Groups in Nuclear and Particle Physics*, edited by F. J. Dyson (Benjamin, New York, 1966), p. 287.
[21] I. M. Gelfand, R. A. Minlos, and Z. Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and their Applications* (Pergamon, Oxford, 1963).
[22] A. Cant and C. A. Hurst, J. Aust. Math. Soc. Ser. B **20**, 446 (1978).
[23] A. P. Yutsis, I. B. Levinson, and V. V. Vanegas, *Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962).
[24] H. J. Bhabha, Rev. Mod. Phys. **17**, 200 (1945).
[25] A. Cant. J. Math. Phys. **22**, 870, 878 (1981).
[26] J. F. Cornwell, Rep. Math. Phys. **2**, 239 (1971).
[27] A. O. Barut, in *Springer Tracts in Modern Physics, Ergebnisse der Exakten Naturwiss*, Vol. 50 (Springer, Berlin, 1969), pp. 1–28.
[28] A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (Polish Scientific Publishers, Warsaw, 1977).
[29] A. Kihlberg, Ark. Fys. **32**, 263 (1966).

Fluctuation–dissipation theorem for QCD plasma

J. Chakrabarti
Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403

(Received 11 October 1984; accepted for publication 23 May 1985)

We explore a quantum-chromodynamic (QCD) plasma in stationary nonequilibrium states assuming that the process of thermalization is governed by Fokker–Planck dynamics. The generalized thermodynamic potential appropriate to the state is obtained. A relationship is developed between the response function and the fluctuations in the stationary state.

I. INTRODUCTION

Recent calculations have suggested that there is a deconfinement transition that occurs in quantum chromodynamics (QCD) at temperatures of the order of a couple of hundred MeV. It also appears that such a phase transition is achievable in the laboratory in high-energy collisions of heavy nuclei.[1] It is therefore of interest to study this plasma state of matter.

Since there is a belief (and also perhaps a proof) that the colliding heavy nuclei would achieve thermal equilibrium within a time of about a fm/c, analyses of the plasma have been made under these equilibrium assumptions. A study of a QCD plasma away from thermal equilibrium broadens our appreciation of this state of matter.

It has recently been suggested that the thermalization of a QCD plasma is governed by the Fokker–Planck equation.[2,3] Since the coupling constant decreases as the momentum transfers increase, most of the parton collisons involve small exchanges of momentum. In this sense a parton in a plasma away from thermal equilibrium undergoes Brownian motion as it thermalizes. It should be remembered that a small fraction of collisions are hard and involve large momentum transfers and that the Fokker–Planck equation for thermalization of the plasma is, therefore, only a first approximation.

In this paper we study a QCD plasma away from thermal equilibrium. The phase-space distribution function is assumed to satisfy a Fokker–Planck equation. The stationary solution of the equation replaces the canonical distribution function $e^{-\beta H}$ ($H \equiv$Hamiltonian) of the equilibrium theory. We study the linear response theory and obtain a connection between the response function in terms of the stationary fluctuations. This is the fluctuation–dissipation theorem.[4]

II. PARTON DISTRIBUTION AND TRANSPORT EQUATIONS

We briefly review earlier work on the subject[2] here. Let us for definiteness assume that two heavy nuclei collide along the z axis at time $t = 0$. Instead of using z and t as coordinates we use τ and η defined as (see Fig. 1)

$$t = \tau \cosh \eta \quad \text{and} \quad z = \tau \sinh \eta. \tag{1}$$

Now τ is the proper time measured from the origin of the (t,z) coordinates. The transverse coordinates will not be used in our notation as all the effects of the transverse motions will be taken into account in the transverse mass of the partons.

Let Y be the rapidity of one of the nuclei and y the rapidity of a parton at the moment of collision. The initial parton distribution in a nucleon relevant for low P_t is known[2] and let us call it $Q(x)$. From x we can go to the rapidity variable by the use of

$$x = (m_T/M)e^{y-Y}, \tag{2}$$

where M is the mass of a nucleon and m_T the transverse mass of a parton. Thus the initial rapidity distribution of the partons at the moment of collision is known.

Let $F(\tau,\eta,y)$ be the phase space distribution function at (τ,η). The transport equation for the distribution is

$$V^\mu \partial_\mu F(\tau,\eta,y) = L(F), \tag{3}$$

where L is the collision operator.

The collision operator may be written in terms of the transport rate $T(y,y')$ that a parton of rapidity y gains y' to get to $y + y'$. Therefore, the collision operator in terms of T is given as

$$L(F) = \int dy' [T(y-y',y')F(\tau,\eta,y-y') - T(y,y')F(\tau,\eta,y)]. \tag{4}$$

Expanding the first term inside the integral around y and making the soft collision approximation we get the transport equation for $F(\tau,\eta,y)$ as[5]

$$V^\mu \partial_\mu F \equiv \frac{\partial}{\partial y}\left[\frac{\partial B}{\partial y} - A\right]F, \tag{5}$$

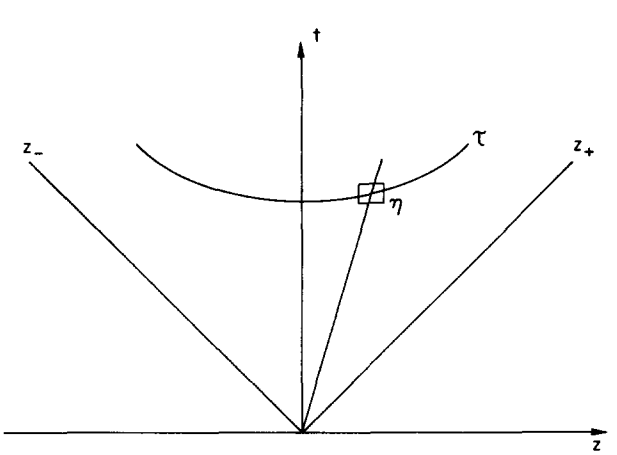


FIG. 1. The nuclei collide along the z axis beginning at time $t = 0$. The momentum distributions of the partons inside the nucleons prior to the collision are known from present experiments. The distributions evolve according to the Fokker–Planck equation subsequent to the collision.

 0022-2488/85/123190-03$02.50

where

$$B = \frac{1}{2}\int dy'\, y'^2 T(y,y') \tag{6}$$

and

$$A(y) = \int dy\, y' T(y,y'). \tag{7}$$

Instead of dealing with $F(\tau,\eta,y)$, it turns out to be more convenient to deal with the quantity f defined as

$$F(\tau,\eta,y) = \int_{-\infty}^{+\infty} P(\tau,\eta,y')f(\sigma,y-y')dy', \tag{8}$$

where $P(\tau,\eta,y')$ is the distribution in a collisionless plasma satisfying $V^\mu \partial_\mu P = 0$ and $\sigma = \tau\cosh(\eta - y')$. The function $P(\tau,\eta,y')$ is obtained from $Q(y)$.

Since the function P is arbitrary, Eq. (5) in F translates to

$$V^\mu \partial_\mu f(\tau,\eta,y-y') = L(f). \tag{9}$$

Using

$$V^\mu \partial_\mu = \frac{\partial}{\partial\sigma}, \tag{10}$$

we get

$$\frac{\partial}{\partial\sigma} f(\sigma,y) = \frac{\partial}{\partial y}\left[\frac{\partial B(y)}{\partial y} - A(y)\right] f(\sigma,y), \tag{11}$$

which is the Fokker–Planck equation for the function $f(\sigma,y)$.

The elimination of the explicit dependence on the initial collisionless distribution P makes the equation in f easy to handle. However, the dependence of P on η and y is small only in the central region and, therefore, the equation in f also pertains only to this domain.

In the central region[6] the function $B(y)$ may be approximated[2] by a constant B and may be eliminted from Eq. (11) by using the variables

$$\theta = B\sigma \text{ and } a(y) = A(y)/B. \tag{12}$$

Using these varaibles, we get

$$\frac{\partial}{\partial\theta} f(\theta,y) = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial y} - a(y)\right] f(\theta,y). \tag{13}$$

III. GENERALIZED THERMODYNAMIC POTENTIAL AND THE FLUCTUATION–DISSIPATION THEOREM[4]

The operator $A(y)$ may be obtained by assuming that the Liouville operator L acting on the equilibrium distribution gives zero. If the equilibrium distribution is assumed to be Maxwellian, then we get

$$a(y) = m_T\beta\sinh y, \tag{14}$$

where $\beta = 1/T$. The stationary solution $f_s(y)$ satisfies

$$\frac{\partial \ln f_s}{\partial y} = m_T\beta(\sinh y). \tag{15}$$

Thus,

$$f_s = f_0\, e^{m_T\beta\cosh y}. \tag{16}$$

The generalized thermodyanmic potential is obtained from[7]

$$f_s = e^{-\Phi_s}. \tag{17}$$

Thus,

$$\Phi_s \sim -m_T\beta\cosh y. \tag{18}$$

The average value of a dynamic variable X in a stationary state is

$$\langle X\rangle = \int dy\, X f_s(y). \tag{19}$$

The average value of X would change if we add a small perturbation as

$$\frac{\partial f(\theta,y)}{\partial\theta} = (L + \delta L)f(\theta,y). \tag{20}$$

It is easy to check that the system defined by Eq. (13) satisfies irreversibility and the condition of detailed balance.

If the system is disturbed at time θ_0 by adding a small perturbation term δL to the operator L, the distribution f changes to

$$f(\theta,y) = \left\{\exp\int_{\theta_0}^{\theta} d\theta'[L + \delta L(\theta')]\right\} f_s \tag{21}$$

$$\equiv\left[e^{L(\theta-\theta_0)} + \int_{\theta_0}^{\theta} d\theta'\, e^{L(\theta-\theta')}\delta L(\theta') \times e^{L(\theta'-\theta_0)} + \cdots\right] f_s$$

$$= f_s(y) + \int_{\theta_0}^{\theta} d\theta'\, e^{L(\theta-\theta')}\delta L(\theta') f_s + \cdots. \tag{22}$$

The change in the value of a dynamical variable may be approximated by keeping only the first term in δL. Thus,

$$\delta\langle X\rangle = \int [f(\theta,y) - f_s(y)]X(y)dy. \tag{23}$$

If the external perturbation is written as

$$\delta L = -K\frac{\partial}{\partial y}, \tag{24}$$

then, the response function R_x is defined as

$$\delta\langle X\rangle = \int_{-\infty}^{\theta} K(\theta')R_x(\theta-\theta')d\theta'. \tag{25}$$

Comparing Eqs. (22)–(25), we get

$$R_x(\theta) = -\int X(y)\exp(L\theta)\frac{\partial f_s}{\partial y}dy. \tag{26}$$

By using the properties of $f_s(y)$ it is now possible to express the response function in terms of the steady-state fluctuations. The function $f_s(y)$ satisfies

$$Lf_s = 0 \tag{27}$$

and

$$L[yf_s] = a(y)f_s. \tag{28}$$

Using these we may rewrite the response function as

$$R_x(\theta) = -\int X(y)\exp[L(\theta)]L(yf_s)dy \tag{29}$$

$$= -\frac{d}{d\theta}\int X(y)\exp[L\theta]yf_s(y)dy \tag{30}$$

$$= -\frac{d}{d\theta}\langle X(y,\theta)y\rangle. \tag{31}$$

This is the fluctuation dissipation theorem relating the response function R_x to the steady-state fluctuations.

IV. CONCLUSIONS

Under the assumption that the parton interactions are mostly soft during the process of thermalization in heavy ion collision, it has been suggested earlier[2] that the transport properties of a QCD plasma are of the Fokker–Planck variety. We have indicated how for such systems a generalized thermodynamic potential may be obtained. Also the responses of the plasma under small perturbations are related to the stationary fluctuations—the so-called fluctuation–dissipation theorem.

We believe that there is a rather vast area where these results may be applied. Propagation of disturbances in a plasma is a subject on which considerable work has already been done and more work seems possible. Further, it seems possible to arrive at the various moment sum rules *à la* Kubo.[8]

ACKNOWLEDGMENTS

We are grateful to Professor R. C. Hwa for providing us with many insights into the subject of a QCD plasma. We had interesting discussions with Professor Marvin Girardeau and Professor R. Mazo. Several conversations on the subject with Dr. E. Gozzi were useful. We acknowledge the hospitality of the theory group at Fermi Lab where most of this work was done.

This work was supported by the United States Department of Energy under Contract No. DE-AT06-76-ER 7004.

[1] For reviews see *Quark Matter Formation and Heavy Ion Collisions*, edited by M. Jacob and H. Satz (World Scientific, Singapore, 1982); and E. V. Shuryak, Phys. Rep. **61**, 71 (1980).

[2] R. C. Hwa, University of Oregon preprint OITS 263.

[3] For earlier work on transport properties of plasma see U. Heinz in Ref. 1, p. 439.

[4] R. Graham and F. Haake, *Springer Tracts in Modern Physics*, Vol. 66 (Springer, New York, 1973); S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

[5] To be precise the following approximation has been made:

$$T(y-y',y')F(y-y') = T(y,y')F(y) - y'\frac{\partial}{\partial y}[T(y,y')F(y)] + \frac{1}{2}y'^2\frac{\partial^2}{\partial y^2}[T(y,y')F(y)].$$

Thus,

$$L(F) = \int dy'\left[-y'\frac{\partial}{\partial y}(TF) + \frac{1}{2}y'^2\frac{\partial^2}{\partial y^2}(TF)\right].$$

Hence, Eq. (5) follows with A and B as defined in Eqs. (6) and (7).

[6] In going from the distribution $F(\tau,\eta,y)$ to $f(\tau,\eta,y-y')$ [see Eq. (8)], the dependence on P—the distribution in collisionless plasma—has been eliminated. P depends marginally on η and y only in the central rapidity region and, thus, the approximation pertains to this region only.

[7] The steady-state distribution is characterized by probability density f_s. When δL is switched on at θ_0 the probability density deviates from f_s. The thermodynamic potential plays the role of $(1/kT)H$, where H is the Hamiltonian.

[8] R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1954); R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).

Simple calculation of Löwdin's alpha function. II. Easier procedure for evaluating $b_{Kk}(L\,M|l)$, and vanishing of $h_{n,2n-l}(L\,M|l)$ for special values of l and n

Noboru Suzuki
Department of Materials Science, The University of Electro-Communications, Chofu-Shi, Tokyo 182, Japan

(Received 3 April 1985; accepted for publication 25 June 1985)

This paper subsequent to the one [J. Math. Phys. **25**, 1133 (1984)] (referred to as Part I) presents the following new results: It is found out that for $M = L$ and $L-1$ the coefficients $b_{Kk}(L\,M|l)$ in Löwdin's α-function have properties other than manifested in Part I. The expression for $b_{Kk}(L\,M|l)$ is shown to be equivalent to the one into which Sharma's expression, obtained in a different manner from that in Part I, is simplified by Rashid. The use of Rashid's expression leads to the recurrence formula for $b_{Kk}(L\,M|l)$ with respect to M only. This formula and the expression for the $b_{Kk}(L\,M|l)$ with $M = L$ provide an easier procedure for successively evaluating $b_{Kk}(L\,M|l)$ than in Part I. Furthermore, it is proved that the coefficients $h_{n,2n-l}(L\,M|l)$ in the asymptotic form of the α-function vanish for $l < l + M$ and for $n < l$.

I. INTRODUCTION

In the preceding paper,[1] which will be hereafter referred to as Part I, it is shown that Löwdin's α-function[2] derived from Silverstone and Moats' expansion formula[3] is expressed in a much easier form to calculate than those presented by several other investigators.[2,4,5] In the expression the coefficients $b_{Kk}(L\,M|l)$ defined in Sec. 2 of Part I, which appear in the α-function, are written in a simple form. Thanks to the simplicity, some of the properties of $b_{Kk}(L\,M|l)$ are manifested, and several recurrence formulas necessary for successively evaluating $b_{Kk}(L\,M|l)$ are derived. An asymptotic expression for the α-function is also obtained in a simple form. It is then proved from the property of $b_{Kk}(L\,M|l)$ with $k=0$ that the coefficients $h_{n,2n-l}(L\,M|l)$, appearing in the asymptotic expression, vanish for $l=0$ unless $l=0$.

The present paper, Part II, will amplify Part I by further investigating the following points: (i) whether any simplified expression for $b_{Kk}(L\,M|l)$ with a special value of M is available; (ii) whether any recurrence formula for $b_{Kk}(L\,M|l)$ with respect to only M, K, or k is obtainable; (iii) to prove that our expression for $b_{Kk}(L\,M|l)$ is equivalent to the one into which Sharma's expression,[5] obtained in a different manner from ours in Part I, is simplified by Rashid[6]; and (iv) to prove the vanishing of $h_{n,2n-l}(L\,M|l)$ for $n<l$, which is only referred to in Part I.

The following sections will present the useful results obtained through the investigation on the above points. In Sec. II it will be shown that $b_{Kk}(L\,M|l)$ for $M=L$ and $L-1$ are expressed, respectively, in only one-terms and from their expressions their properties, other than manifested in Part I, and the relation between them are found out. Section III will give the proof of the equivalence of our expression for $b_{Kk}(L\,M|l)$ to Rashid's expression. In Sec. IV it will be demonstrated that the recurrence formula for $b_{Kk}(L\,M|l)$, with respect to M only, is obtained using Rashid's expression, and this formula and the expression for $b_{Kk}(L\,M|l)$ with $M=L$ provide a much easier procedure for successively evaluating $b_{Kk}(L\,M|l)$ than in Part I. Section V will give the proof of the vanishing of $h_{n,2n-l}(L\,M|l)$ for $l<l+M$ and for $n<l$. Finally, in Sec. VI some remarks will be made on what has led to the results given in Part I and to be presented in this paper, and on the significance of the results.

II. PROPERTIES OF $b_{Kk}(L\,M|l)$ FOR $M=L$ AND $L-1$

Before starting the discussion in this section, we define the factorial for a half-integer as

$$(p-\tfrac{1}{2})!\equiv\begin{cases}(p-\frac{1}{2})(p-\frac{3}{2})\cdots\frac{1}{2}\,, \\ \quad\text{for a positive integer } p\,, \\ 1/[(-1)^{-p}(-p-\frac{1}{2})!]\,, \\ \quad\text{for a negative integer } p\,,\end{cases} \tag{1}$$

with $(-\frac{1}{2})!\equiv 1$. Thus the following relation holds:

$$(p-\tfrac{1}{2})!(-p-\tfrac{1}{2})! = (-1)^p\,. \tag{2}$$

Use is made of this factorial throughout Part II, since it facilitates all the discussions to be made hereafter.

The expression for $b_{Kk}(L\,M|l)$, Eq. (2.6) with Eqs. (2.4), (2.7), and (2.8) of Part I, is simplified using the above factorial as

$$b_{Kk}(L\,M|l) = (-1)^M[(L-M)!(L+M)!(l-M)!(l+M)!]^{1/2}\frac{(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!}{(L-\frac{1}{2})!K!(l-\frac{1}{2})!k!}$$
$$\times\sum_{\lambda=\lambda_{\min}}^{L+l}{}^{2}(-1)^{(L+l-\lambda)/2}\cdot C(L l\lambda;M\,-M)C(L l\lambda;0\,0)\frac{(-(L+l-\lambda)/2+K+k-\frac{1}{2})!}{((L+l+\lambda)/2-K-k)!}\,, \tag{3}$$

with $\lambda_{\min}=\max\{|L-l|, 2(K+k)-(L+l)\}$. Here $C(L l\lambda;\,M\,-M)$ and $C(L l\lambda;0\,0)$ are the Clebsch–Gordan coefficients.[7]

 0022-2488/85/123193-07$02.50

Throughout Part II, Eq. (3) is adopted as the expression for $b_{Kk}(L\,M\,|l)$, and only the $b_{Kk}(L\,M\,|l)$ with $l\geqslant L\geqslant M\geqslant 0$ are taken into consideration because there exist the symmetry relations between them, expressed in Eqs. (3.5) and (3.6) of Part I.

In this section it will be shown that $b_{Kk}(L\,M\,|l)$ for $M=L$ and $L-1$ have properties other than manifested in Sec. 3 of Part I.

First, the properties of $b_{Kk}(L\,M\,|l)$ with $M=L$ and their sums over K and over k will be investigated. Introducing into Eq. (3) the expressions[8] for $C(L\lambda;L\ -L)$ and $C(L\lambda;0\ 0)$, we obtain an expression for $b_{Kk}(L\,L\,|l)$ as

$$b_{Kk}(L\,L\,|l)=(-1)^{L}\frac{(L+l)!L!(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!}{K!(l-\frac{1}{2})!k!}$$
$$\times\sum_{\lambda=\lambda_{\min}\ 2}^{L+l}(\lambda+\tfrac{1}{2})\frac{((-L+l+\lambda)/2-\frac{1}{2})!(-(L+l-\lambda)/2+K+k-\frac{1}{2})!}{((L+l-\lambda)/2)!((L-l+\lambda)/2)!((L+l+\lambda)/2+\frac{1}{2})!((L+l+\lambda)/2-K-k)!}. \tag{4}$$

The summation over λ in Eq. (4) is carried out in Appendix A, and its result is given by Eq. (A4). The use of Eq. (A4) reduces Eq. (4) to

$$b_{Kk}(L\,L\,|l)=(-1)^{L}\frac{(L+l)!(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!(-L+K+k-\frac{1}{2})!}{K!(l-\frac{1}{2})!k!(L+l-K-k)!}. \tag{5}$$

Replacing k by $L+l-K-k$ in Eq. (5), we find that

$$b_{K\,L+l-K-k}(L\,L\,|l)=b_{Kk}(L\,L\,|l). \tag{6}$$

This reveals the symmetry property of $b_{Kk}(L\,L\,|l)$. Therefore, evaluation of all $b_{Kk}(L\,L\,|l)$ over the permissible values of K and k, seen in Eq. (2.12) of Part I, is not necessary. The sum of $b_{Kk}(L\,L\,|l)$ over K can be obtained by putting $\mu=L-k-\frac{1}{2}$, $\rho=L+l-k$, and $\nu-\rho=-L-\frac{1}{2}$ in Eq. (B1) of Appendix B. Then since $k\geqslant 0$, obviously $\mu+\nu-\rho=-k-1<0$. Hence, from the note in Appendix B, if and only if $\mu+\nu=L+l-2k-1\geqslant 0$ or $k\leqslant(L+l-1)/2$,

$$\sum_{K=0}^{L+l-k}b_{Kk}(L\,L\,|l)=0. \tag{7}$$

On the other hand, the sum of $b_{Kk}(L\,L\,|l)$ over k can be obtained by setting $\mu=L-K-\frac{1}{2}$, $\rho=L+l-K$, and $\nu-\rho=-l-\frac{1}{2}$ in Eq. (B1). Then, apparently $\mu+\nu-\rho=L-l-K-1<0$ because $l\geqslant L$ and $K\geqslant 0$. Thus, if and only if $\mu+\nu=2(L-K-\frac{1}{2})\geqslant 0$ or $K\leqslant L-\frac{1}{2}$,

$$\sum_{k=0}^{L+l-K}b_{Kk}(L\,L\,|l)=0. \tag{8}$$

The vanishing of these sums is helpful for checking whether the calculated values of $b_{Kk}(L\,L\,|l)$ are correct.

Second, the properties of $b_{Kk}(L\,M\,|l)$ with $M=L-1$ will be examined. Introduction of the expression for $C(L\lambda;L-1\ -L+1)$, obtained from that for $C(L\lambda;L\ -L)$ using the recurrence formula for $C(L\lambda;M\ -M)$ in M, Eq. (4.2) in Part I, into Eq. (3) yields an expression for $b_{Kk}(L\,L-1|l)$ as

$$b_{Kk}(L\,L-1|l)=(-1)^{L-1}\frac{(L-1)!(L+l-1)!(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!}{2\cdot K!(l-\frac{1}{2})!k!}$$
$$\times\sum_{\lambda=\lambda_{\min}\ 2}^{L+l}[\lambda(\lambda+1)+L(L-1)-l(l+1)](\lambda+\tfrac{1}{2})$$
$$\times\frac{((-L+l+\lambda)/2-\frac{1}{2})!(-(L+l-\lambda)/2+K+k-\frac{1}{2})!}{((L+l-\lambda)/2)!((L-l+\lambda)/2)!((L+l+\lambda)/2+\frac{1}{2})!((L+l+\lambda)/2-K-k)!}. \tag{9}$$

Separating the sum in Eq. (9) into two sums by utilizing the identity

$$\lambda(\lambda+1)=-(L+l+\lambda-2K-2k)(L+l-\lambda-2K-2k-1)+(L+l-2K-2k)(L+l-2K-2k-1), \tag{10}$$

reducing each of the two sums to a single term by using Eq. (A4), and then combining the two terms, we arrive at

$$b_{Kk}(L\,L-1|l)=(-1)^{L}(L+l-2K-2k)\frac{(L+l-1)!(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!(-L+K+k-\frac{1}{2})!}{K!(l-\frac{1}{2})!k!(L+l-K-k)!}. \tag{11}$$

From Eq. (11) it can be seen immediately that

$$b_{Kk}(L\,L-1|l)=0,\quad \text{for } K+k=(L+l)/2. \tag{12}$$

Also, compared with Eq. (5), the relation between $b_{Kk}(L\,L-1|l)$ and $b_{Kk}(L\,L\,|l)$ is easily obtained:

$$b_{Kk}(L\,L-1|l)=[(L+l-2K-2k)/(L+l)]\cdot b_{Kk}(L\,L\,|l). \tag{13}$$

Needless to say, this relation provides the means of evaluating $b_{Kk}(L\,L-1|l)$ from $b_{Kk}(L\,L\,|l)$. Further, introduction of the symmetry relation (6) into Eq. (13) leads to the following relation between $b_{Kk}(L\,L-1|l)$ and $b_{K\,L+l-K-k}(L\,L-1|l)$:

$$b_{K\,L+l-K-k}(L\,L-1|l)=-[(L+l-2k)/(L+l-2K-2k)]\cdot b_{Kk}(L\,L-1|l). \tag{14}$$

In particular, for $K=0$,

$$b_{0\,L+l-k}(L\,L-1|l)=-b_{0k}(L\,L-1|l)\,. \tag{15}$$

The property shown in Eq. (12) and the relation (14) are helpful for evaluating $b_{Kk}(L\,L-1|l)$.

III. EQUIVALENCE OF OUR EXPRESSION FOR $b_{Kk}(L\,M|l)$ TO RASHID'S

About ten years ago, Sharma derived an expression for $b_{Kk}(L\,M\,|l)$ [$b_k(KlLM)$ in his notation] in the conventional manner different from ours in Part I, which is written in the form of a quadruple sum.[5] Recently the expression has been simplified into the form of a single sum by Rashid[6] skillfully using Eq. (B1) and relation (2). In fitting Rashid's expression[9] to our definition for $b_{Kk}(L\,M\,|l)$ given in Sec. 2 of Part I and then replacing the summation index t in his expression by $L-M-s$, we may rewrite Rashid's expression in our notation as

$$b^{(R)}_{Kk}(L\,M\,|l)=\frac{(L-M)!(L+M)!(l-M)!(l+M)!(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!}{(L+l-K-k)!(L-\frac{1}{2})!K!(l-\frac{1}{2})!k!}$$
$$\times\sum_{s=0}^{L-M}\frac{(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M-s)!}\,. \tag{16}$$

In Eq. (16), to distinguish Rashid's expression from ours, (R) has been attached to the symbol b_{Kk} as a superscript.

Here a more general recurrence formula for $b_{Kk}(L\,M\,|l)$, expressed in Eq. (3), than (4.6) or (4.7) of Part I is derived in order to ascertain that $b^{(R)}_{Kk}(L\,M\,|l)$ satisfies it. It can be obtained by carrying out the same procedure as led to (4.6) of Part I under the consideration of the identity for $\lambda\,(\lambda+1)$, Eq. (10), which may be written as

$$(L-M)(l-M)b_{Kk}(L\,M+1|l)+(L+M)(l+M)b_{Kk}(L\,M-1|l)$$
$$=-4\{(K+1)(L-K-\tfrac{1}{2})b_{K+1\,k}(L\,M\,|l)\ \text{or}\ (k+1)(l-k-\tfrac{1}{2})b_{Kk+1}(L\,M\,|l)$$
$$-[(L+l-K-k)(K+k+\tfrac{1}{2})-\tfrac{1}{2}(Ll+M^2)]b_{Kk}(L\,M\,|l)\}\,. \tag{17}$$

Since it can be easily found that for $M=L$ and $L-1$, $b^{(R)}_{Kk}(L\,M\,|l)$ are, respectively, reduced to the same forms as Eqs. (5) and (11), the proof of the equivalence of $b^{(R)}_{Kk}(L\,M\,|l)$ to our $b_{Kk}(L\,M\,|l)$ is accomplished by showing that $b^{(R)}_{Kk}(L\,M\,|l)$ satisfy the same formula as Eq. (17).

First, the corresponding term $(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$ to the first one on the left-hand side of Eq. (17) is taken into account. In denoting the factor in front of the summation symbol in Eq. (16) by $\beta_{Kk}(L\,M\,|l)$, we may write it as

$$(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$$
$$=\beta_{Kk}(L\,M\,|l)\cdot(L+M+1)(l+M+1)\sum_{s=0}^{L-M-1}\frac{(-L+K+k+s-\frac{1}{2})!}{s!(L-M-1-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M+1-s)!}\,. \tag{18}$$

Multiplying each term in this sum by $1=[-(L+M+1-s)-2(-L+s-\frac{1}{2})]/(L-M-s)$ to divide this sum into two sums, and then doing each term in the second sum of the resulting two by $L+M+1=L+M+1-s+s$, we arrive at

$$(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$$
$$=-\beta_{Kk}(L\,M\,|l)\cdot(l+M+1)\Bigg\{(L+M+1)\sum_{s=0}^{L-M}\frac{(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M-s)!}$$
$$+2\Bigg[\sum_{s=0}^{L-M}\frac{(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{3}{2})!(L+M-s)!}$$
$$+\sum_{s=0}^{L-M}\frac{s\cdot(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{3}{2})!(L+M+1-s)!}\Bigg]\Bigg\}\,. \tag{19}$$

In Eq. (19), combination of the first sum and the first one in the brackets and multiplication of each term in the second sum in the brackets by $l+M+1=L+M+1-s+l-L+s$ lead to

$$(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$$
$$=-\beta_{Kk}(L\,M\,|l)\Bigg\{(l+M+1)\sum_{s=0}^{L-M}\frac{(-L+M+2s)(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M-s)!}$$
$$+2\Bigg[\sum_{s=0}^{L-M}\frac{s\cdot(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{3}{2})!(L+M-s)!}$$
$$+\sum_{s=1}^{L-M}\frac{(-L+K+k+s-\frac{1}{2})!}{(s-1)!(L-M-s)!(l-L-1+s)!(-L+s-\frac{3}{2})!(L+M+1-s)!}\Bigg]\Bigg\}\,. \tag{20}$$

Here, replacing the summation index s in the second sum in the brackets by $s+1$, then multiplying each term in the sum by $1=(L-M-s)/(L-M-s)$, and finally combining the three terms, we reach

$$(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$$
$$=-\beta_{Kk}(L\,M|l)\sum_{s=0}^{L-M}[2(L+l-K-k)s-(L-M)(2L+l+M-2K-2k)]$$
$$\times\frac{(-L+K+k+s-\frac{1}{2})!}{s!(L-M-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M-s)!}\,. \tag{21}$$

In utilizing the fact that the expression in the above brackets is equal to $2(L+l-K-k)(-L+K+k+s+\frac{1}{2})+(L+l-K-k)(2M-2K-2k-1)+(L-M)(l-M)$ and returning the expression in Eq. (21) to the one in terms of $b^{(R)}_{Kk}(L\,M|l)$, at last we obtain

$$(L-M)(l-M)b^{(R)}_{Kk}(L\,M+1|l)$$
$$=-2\{(K+1)(L-K-\tfrac{1}{2})b^{(R)}_{K+1\,k}(L\,M|l)$$
$$\text{or}\ (k+1)(l-k-\tfrac{1}{2})b^{(R)}_{K\,k+1}(L\,M|l)$$
$$+[(L+l-K-k)(M-K-k-\tfrac{1}{2})$$
$$+\tfrac{1}{2}(L-M)(l-M)]b^{(R)}_{Kk}(L\,M|l)\}\,. \tag{22}$$

Second, the corresponding term $(L+M)(l+M)\times b^{(R)}_{Kk}(L\,M-1|l)$ to the second one on the left-hand side of Eq. (17) is taken into consideration. The same manipulation as the above yields its expression in terms of $b^{(R)}_{Kk}(L\,M|l)$ as

$$(L+M)(l+M)b^{(R)}_{Kk}(L\,M-1|l)$$
$$=-2\{(K+1)(L-K-\tfrac{1}{2})b^{(R)}_{K+1\,k}(L\,M|l)$$
$$\text{or}\ (k+1)(l-k-\tfrac{1}{2})b^{(R)}_{K\,k+1}(L\,M|l)$$
$$+[(L+l-K-k)(-M-K-k-\tfrac{1}{2})$$
$$+\tfrac{1}{2}(L+M)(l+M)]b^{(R)}_{Kk}(L\,M|l)\}\,. \tag{23}$$

Consequently, combination of Eqs. (22) and (23) gives the same formula for $b^{(R)}_{Kk}(L\,M|l)$ as Eq. (17). Thus it just has been proven that Eq. (16) is an expression for $b_{Kk}(L\,M|l)$ equivalent to Eq. (3).

IV. EASIER PROCEDURE FOR EVALUATING $b_{Kk}(L\,M|l)$

In this section it will be shown, from the results in the previous sections, that an easier procedure for successively evaluating $b_{Kk}(L\,M|l)$ than in Part I can be found.

At the outset it should be noted that a new recurrence formula for $b_{Kk}(L\,M|l)$ with respect to only M can be derived by eliminating the term including $b^{(R)}_{K+1\,k}(L\,M|l)$ or $b^{(R)}_{K\,k+1}(L\,M|l)$ from both Eqs. (22) and (23). This formula is written as

$$(L+M)(l+M)b_{Kk}(L\,M-1|l)$$
$$=2M(L+l-2K-2k)b_{Kk}(L\,M|l)$$
$$+(L-M)(l-M)b_{Kk}(L\,M+1|l)\,. \tag{24}$$

Here the superscript (R) on the symbol b_{Kk} has been removed. Putting $M=L$ in Eq. (24) leads directly to the relation (13). The recurrence formula with respect to only K or k can be also obtained from the same equations (22) and (23). Its form is, however, rather complicated and thus less useful for computing $b_{Kk}(L\,M|l)$.

By the iterative use of Eq. (24) all the $b_{Kk}(L\,M|l)$ with $M\leqslant L-1$ can be calculated from $b_{Kk}(L\,L|l)$.

The evaluation of $b_{Kk}(L\,L|l)$ can be made using the recurrence formulas in K and in k derived immediately from Eq. (5). These formulas are written as

$$b_{K+1\,k}(L\,L|l)=-\frac{(L+l-K-k)(L-K-k-\frac{1}{2})}{(K+1)(L-K-\frac{1}{2})}\cdot b_{Kk}(L\,L|l)\,, \tag{25a}$$

and

$$b_{K\,k+1}(L\,L|l)=-\frac{(L+l-K-k)(L-K-k-\frac{1}{2})}{(k+1)(l-k-\frac{1}{2})}\cdot b_{Kk}(L\,L|l)\,. \tag{25b}$$

Thus, beginning with $b_{0\,0}(L\,L|l)$, we can calculate all $b_{Kk}(L\,L|l)$ over K and k, taking into account the symmetry relation (6).

Here it should be noted that, because of our definition for $b_{Kk}(L\,M|l)$, necessarily

$$b_{0\,0}(L\,M|l)=1\,, \tag{26}$$

independent of the value of M. This is easily ascertained from Eq. (16) and also by using Eqs. (5) and (24) for $K=k=0$.

In the above manner all $b_{Kk}(L\,M|l)$ are successively evaluated more easily than in Part I.

V. VANISHING OF $h_{n,2n-i}(L\,M|l)$ FOR $i<l+M$ AND FOR $n<l$

In the present section it will be proved that the coefficients $h_{n,s}(L\,M|l)$ appearing in the asymptotic form of the α-function, which are expressed in Eq. (5.6) of Part I,[10] vanish under the two independent conditions. Replacing the subscript s on $h_{n,s}(L\,M|l)$ with $2n-i$ and substituting Eq. (16) for $b_{Kk}(L\,M|l)$ in the expression for $h_{n,s}(L\,M|l)$, we may rewrite

$h_{n,s}(L\,M\,|\,l)$ as

$$
\begin{aligned}
&h_{n,2n-i}(L\,M\,|\,l)\\
&\quad=\frac{(L-M)!(L+M)!(l-M)!(l+M)!}{2\cdot(L-\frac{1}{2})!(l-\frac{1}{2})!}\\
&\quad\times\sum_{k=0}^{[i/2]}\sum_{K=[(i+1)/2]-k}^{L+l-k}\sum_{s=0}^{L-M}\frac{(n-k-\frac{1}{2})!(K-\frac{1}{2})!}{(n-k+\frac{1}{2})!([i/2]-k)!([(i+1)/2]-k-\frac{1}{2})!(K+k-[(i+1)/2])!(K+k-[i/2]-\frac{1}{2})!}\\
&\quad\times\frac{(L-K-\frac{1}{2})!(l-k-\frac{1}{2})!(-L+K+k+s-\frac{1}{2})!}{(L+l-K-k)!k!s!(L-M-s)!(l-L+s)!(-L+s-\frac{1}{2})!(L+M-s)!},
\end{aligned}
\tag{27}
$$

where n runs over all non-negative integers, and i from 0 to min $\{2n,2(L+l)\}$. Here $[i/2]$ denotes the quotient of $i/2$.

First, it will be shown that $h_{n,2n-i}(L\,M\,|\,l)$ for $i<l+M$ vanishes. Introducing into Eq. (27) the following relation, derived from Eq. (B1) by setting $\mu=L$, $\nu=-K-\frac{1}{2}$, and $\rho=k+s$,

$$
\begin{aligned}
&(K-\tfrac{1}{2})!(L-K-\tfrac{1}{2})!(-L+K+k+s-\tfrac{1}{2})!\\
&\quad=(-1)^{L+k+s}L!(k+s)!\sum_{t=0}^{\min\{L,k+s\}}[t!(L-t)!(k+s-t)!(-K-k-s+t-\tfrac{1}{2})!]^{-1},
\end{aligned}
\tag{28}
$$

and then extracting only the sum S_1 over K from the resulting equation, we may write S_1 as

$$
S_1=\sum_{K=[(i+1)/2]-k}^{L+l-k}\left[\left(K+k-\left[\frac{i+1}{2}\right]\right)!\left(K+k-\left[\frac{i}{2}\right]-\frac{1}{2}\right)!(L+l-K-k)!\left(-K-k-s+t-\frac{1}{2}\right)!\right]^{-1}.
\tag{29}
$$

Further, replacement of K with $\kappa-k+[(i+1)/2]$ leads to

$$
S_1=\sum_{\kappa=0}^{L+l-[(i+1)/2]}\left[\kappa!\left(L+l-\left[\frac{i+1}{2}\right]-\kappa\right)!\left(-\left[\frac{i+1}{2}\right]-s+t-\frac{1}{2}-\kappa\right)!\left(\left[\frac{i+1}{2}\right]-\left[\frac{i}{2}\right]-\frac{1}{2}+\kappa\right)!\right]^{-1}.
\tag{30}
$$

Putting $\mu=-[(i+1)/2]-s+t-\frac{1}{2}$, $\rho=L+l-[(i+1)/2]$, and $\nu-\rho=[(i+1)/2]-[i/2]-\frac{1}{2}$ in Eq. (B1) yields the result of the summation over κ in Eq. (30). Then it can be easily seen that $\mu+\nu-\rho=-[i/2]-1-s+t<0$ because $-s+t\leqslant\min\{L,[i/2]\}$. From the note in Appendix B, if and only if $\mu+\nu=L+l-i-s+t-1\geqslant 0$ or $-s+t\geqslant -L-l+i+1$, S_1 vanishes. Thus, since $-L-l+i\leqslant[i/2]$, only for $-s+t\leqslant\min\{L,-L-l+i\}$, S_1 remains nonzero. On the other hand, $\mu+\nu=L+l-i-s+t-1\geqslant l+M-i-1$ because $-s+t\geqslant -L+M$. Therefore, if $i<l+M$, then $\mu+\nu\geqslant 0$ for any permissible value of $-s+t$. Then S_1 vanishes, and accordingly $h_{n,2n-i}(L\,M\,|\,l)$ does. Here it should be noted that the condition $i<l+M$ involves the one for Eq. (5.8) of Part I to vanish. Consequently, only when $l+M\leqslant i\leqslant\min\{2n,2(L+l)\}$, S_1 takes a nonzero value, which, according to Eq. (B1), is written as

$$
S_1=(-1)^{L+l-[(i+1)/2]}\frac{([i/2]+s-t)!}{(-[(i+1)/2]-s+t-\frac{1}{2})!(L+l-[i/2]-\frac{1}{2})!(-L-l+i+s-t)!(L+l-[(i+1)/2])!}.
\tag{31}
$$

Second, it will be shown that $h_{n,2n-i}(L\,M\,|\,l)$ for $n<l$ vanishes. This vanishing is referred to in Part I, but has not yet been proven. Here the two relations

$$
\begin{aligned}
&\frac{(n-k-\frac{1}{2})!}{([(i+1)/2]-k-\frac{1}{2})!([i/2]-k)!}\\
&\quad=(n-[(i+1)/2])!(n-[i/2]-\tfrac{1}{2})!\\
&\quad\times\sum_{p=0}^{\min\{n-[(i+1)/2],[i/2]-k\}}[p!(n-[(i+1)/2]-p)!([i/2]-k-p)!([(i+1)/2]-[i/2]-\tfrac{1}{2}+p)!]^{-1},
\end{aligned}
\tag{32}
$$

and

$$
\frac{(k+s)!}{(k+s-t)!t!s!k!}=\sum_{q=\max\{0,-s+t\}}^{\min\{k,t\}}[q!(t-q)!(k-q)!(s-t+q)!]^{-1},
\tag{33}
$$

obtained using Eq. (B1), are introduced into the expression for $h_{n,2n-i}(L\,M\,|\,l)$ derived by inserting Eq. (31) into Eq. (27), and then, from the resulting expression, is extracted only the sum S_2 over k. It is written as

$$
\begin{aligned}
S_2&=\sum_{k=q}^{[i/2]-p}(-1)^k\frac{(l-k-\frac{1}{2})!}{(n-k+\frac{1}{2})!([i/2]-k-p)!(k-q)!}\\
&=(-1)^t\sum_k\left[(k-q)!\left(\left[\frac{i}{2}\right]-k-p\right)!\left(n-k+\frac{1}{2}\right)!\left(-l+k-\frac{1}{2}\right)!\right]^{-1}.
\end{aligned}
\tag{34}
$$

Further the replacement of $k-q$ with κ leads to

$$S_2=(-1)^l\sum_{\kappa=0}^{[i/2]-p-q}\left[\kappa!\left(\left[\frac{i}{2}\right]-p-q-\kappa\right)!\left(n-q+\frac{1}{2}-\kappa\right)!\left(-l+q-\frac{1}{2}+\kappa\right)!\right]^{-1}. \tag{35}$$

Putting $\mu=n-q+\frac{1}{2}$, $\rho=[i/2]-p-q$, and $\nu-\rho=-l+q-\frac{1}{2}$ in Eq. (B1), we can obtain the result of the summation over κ in Eq. (35). Then it can be seen immediately that $\mu+\nu=n-l+[i/2]-p-q$ and $\mu+\nu-\rho=n-l$. Here $n-l+[i/2]-p-q\geqslant\max\{-l+i-t, n-l\}$ because $p+q\leqslant[i/2]+\min\{n-i+t, 0\}$. Further, since $t\leqslant -l-M+i$,

$$\mu+\nu=n-l+[i/2]-p-q\geqslant\max\{M, n-l\}\geqslant 0. \tag{36}$$

The last equality or inequality in (36) arises from the restriction $M\geqslant 0$ taken in Sec. II. Because (36) holds for any set of the permissible values of p, q, and t, S_2 vanishes if $\mu+\nu-\rho=n-l<0$. Accordingly, then $h_{n,2n-i}(L\,M\,|\,l)$ disappears.

Thus the vanishing of $h_{n,2n-i}(L\,M\,|\,l)$ for $i<l+M$ and for $n<l$ has just been proven.

VI. CONCLUDING REMARKS

In Part I, the coefficients in the α-function were factorized into $\gamma(L\,M\,|\,l)$ expressed in Eq. (2.9) of Part I and $b_{Kk}(L\,M\,|\,l)$ so $b_{00}(L\,M\,|\,l)$ may equal a unity, and then, with respect to $b_{Kk}(L\,M\,|\,l)$, not the product $\gamma(L\,M\,|\,l)\times b_{Kk}(L\,M\,|\,l)$, their properties and the existence of some procedure for successively computing them were investigated. Consequently, several desirable results were obtained. Here in Part II, further investigation on the same subjects has been made, and some useful results have been obtained as given in the previous sections. It is noteworthy that the manipulation of $b_{Kk}(L\,M\,|\,l)$, not $\gamma(L\,M\,|\,l)\cdot b_{Kk}(L\,M\,|\,l)$, has yielded all of those results.

From a practical point of view, it is to be emphasized that, since the products $\gamma(L\,M\,|\,l)\cdot b_{Kk}(L\,M\,|\,l)$ and $\gamma(L\,M\,|\,l)\cdot h_{n,2n-i}(L\,M\,|\,l)$ are independent of the form of the radial part of a function to be expanded, they, once evaluated, can be used for any function as long as its angular part is expressed by a spherical harmonics. In this context, it is important that in Part I and especially in the present paper the procedure for successively evaluating $b_{Kk}(L\,M\,|\,l)$ has been equipped.

Any procedure for directly evaluating $h_{n,2n-i}(L\,M\,|\,l)$ by no use of $b_{Kk}(L\,M\,|\,l)$, however, has not yet been found. It is necessary to seek it further.

APPENDIX A: PERFORMANCE OF THE SUMMATION IN EQ. (4)

The sum in Eq. (4), denoted by S, may be rewritten as below by replacing $L+l-\lambda$ and $K+k$ by 2μ and j, respectively:

$$S=\sum_{\mu=0}^{\mu_{\max}}\frac{L+l-2\mu+\frac{1}{2}}{(L+l-\mu+\frac{1}{2})!\mu!}\cdot\frac{(l-\mu-\frac{1}{2})!(j-\mu-\frac{1}{2})!}{(L+l-j-\mu)!(L-\mu)!}, \tag{A1}$$

where $\mu_{\max}=\min\{L+l-j, L\}$. Now we perform the summation over μ in Eq. (A1). Separating the sum into two sums by using the identity $L+l-2\mu+\frac{1}{2}=(L+l-\mu+\frac{1}{2})-\mu$, replacing the summation index μ in the resulting second sum by $\mu+1$, and then combining the two sums, we obtain

$$\begin{aligned}S=(j-L-\tfrac{1}{2})\sum_{\mu=0}^{\mu_{\max}-1}&\frac{L+l-2\mu-\frac{1}{2}}{(L+l-\mu-\frac{1}{2})!\mu!}\cdot\frac{(l-\mu-\frac{3}{2})!(j-\mu-\frac{3}{2})!}{(L+l-j-\mu)!(L-\mu)!}\\&+\frac{(l-\mu_{\max}-\frac{1}{2})!(j-\mu_{\max}-\frac{1}{2})!}{(L+l-\mu_{\max}-\frac{1}{2})!\mu_{\max}!(L+l-j-\mu_{\max})!(L-\mu_{\max})!}.\end{aligned} \tag{A2}$$

Further, repeating the same procedure for the sum over μ as the above until all the terms in the sum disappear, we arrive at

$$\begin{aligned}S=(j-L-\tfrac{1}{2})!&\frac{(l-\mu_{\max}-\frac{1}{2})!(j-\mu_{\max}-\frac{1}{2})!}{(L+l-\mu_{\max}-\frac{1}{2})!}\\&\times\sum_{\kappa=0}^{\mu_{\max}}[(j-L-\mu_{\max}-\tfrac{1}{2}+\kappa)!\kappa!\\&\times(L+l-j-\kappa)!(L-\kappa)!]^{-1}.\end{aligned} \tag{A3}$$

Then putting $\mu=L$, $\rho=L+l-j$, and $\nu-\rho=j-L-\mu_{\max}-\frac{1}{2}$ in Eq. (B1) of Appendix B if $\mu_{\max}=L+l-j$, and otherwise interchanging μ and ρ in doing so, we find that the sum over κ is equal to

$$(L+l-\mu_{\max}-\tfrac{1}{2})!/[L!(l-\mu_{\max}-\tfrac{1}{2})!(L+l-j)!\times(j-\mu_{\max}-\tfrac{1}{2})!].$$

Consequently, Eq. (A3) is reduced to

$$S=(j-L-\tfrac{1}{2})!/[L!(L+l-j)!]. \tag{A4}$$

APPENDIX B: ADDITION THEOREM FOR BINOMIAL COEFFICIENTS

The addition theorem for the binomial coefficients ${}_\mu C_\kappa$ is expressed by

$$\sum_\kappa {}_\mu C_\kappa\cdot{}_\nu C_{\rho-\kappa}=\sum_\kappa {}_\mu C_{\rho-\kappa}\cdot{}_\nu C_\kappa={}_{\mu+\nu}C_\rho. \tag{B1}$$

In Eq. (B1), ρ and κ must be non-negative integers with the restriction that $\rho\geqslant\kappa$, while each of μ and ν, irrespective of being non-negative or not, may be an integer or a half-integer. Here it should be noted that, if and only if $\mu+\nu$ is a

non-negative integer and less than ρ, the sums in Eq. (B1) vanish because ${}_{\mu+\nu}C_{\rho} = 0$. For a non-negative integer or a half-integer μ, ${}_{\mu}C_{\kappa}$ may be written as

$$ {}_{\mu}C_{\kappa} = \mu!/[\kappa!(\mu-\kappa)!] , \tag{B2} $$

while for a negative integer μ,

$$ {}_{\mu}C_{\kappa} = (-1)^{\kappa}(-\mu+\kappa-1)!/[\kappa!(-\mu-1)!] . \tag{B3} $$

Here it is to be stressed that the expression (B2) is valid even for a half-integer, regardless of being non-negative or not, provided the definition (1) in the text is adopted. Substituting the expressions (B2) and/or (B3) for the binomial coefficients in Eq. (B1) according as the types of integers μ, ν, and $\mu+\nu$ take, we can obtain several formulas expressing the addition theorem for the respective cases.

[1]N. Suzuki, J. Math. Phys. **25**, 1133 (1984); **25**, 3135(E) (1984).
[2]P. O. Löwdin, Adv. Phys. **5**, 96 (1956).
[3]H. J. Silverstone and R. K. Moats, Phys. Rev. A **16**, 1731 (1977).
[4]M. P. Barnett and C. A. Coulson, Philos. Trans. R. Soc. London, Ser. A **243**, 221 (1951); Y. Tanabe, J. Phys. Soc. Jpn. **11**, 980 (1956); F. E. Harris and H. H. Michels, J. Chem. Phys. **43**, S165 (1965); R. R. Sharma, J. Math. Phys. **9**, 505 (1968); E. O. Steinborn and E. Filter, Theor. Chim. Acta **38**, 247 (1975); H. W. Jones and C. A. Weatherford, Int. J. Quantum Chem. **S12**, 483 (1978); I. I. Guseinov, J. Chem. Phys. **69**, 4990 (1978).
[5]R. R. Sharma, Phys. Rev. A **13**, 517 (1976).
[6]M. A. Rashid, J. Math. Phys. **22**, 271 (1981).
[7]For example, M. E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).
[8]See Eqs. (3.29) with (3.16) and Eq. (3.32) in Ref. 7.
[9]See Eq. (11) in Ref. 6.
[10]The 1/2 immediately after the closing bracket in Eq. (5.6) of Part I should be replaced by -1.

Hamiltonian picture of the free electron laser and unitary symmetries

G. Dattoli and R. Mignani[a)]
ENEA, Dip. TIB., Divisione Fisica Applicata, Centro Ricerche Energia Frascati, C. P. 65-00044 Frascati, Rome, Italy

(Received 17 July 1984; accepted for publication 21 June 1985)

It is shown that—in analogy to what has been recently suggested for a quantum n-level system—in a multimode analysis of the free electron laser (FEL) Hamiltonian picture, one can use the (n^2-1)-dimensional vector $\vec{F}$ of SU(n) to describe the main FEL dynamics. The FEL coherence properties are discussed by this group-theoretical approach, and a critical analysis of related Casimir invariants and conservation laws is given.

I. INTRODUCTION

It happens sometimes in physics that seemingly unrelated fields or techniques show unexpected connections. In most cases, this leads to a better understanding or to a clever mathematical formulation of the physical fields involved. An illuminating example is provided by the application of unitary symmetry schemes first to nuclear physics[1] and then to high-energy physics.[2]

More recently, the tools of unitary symmetries have been applied to the dynamics of an atomic n-level quantum system.[3–5] This kind of approach has allowed one to generalize the Bloch-like equations describing the interaction of a two-level system with radiation[6] to the case of multilevel systems—a problem long thought impossible to solve.[7] Moreover, a set of new, unforeseen constants of motion, able to give a deeper insight into the system's dynamics, have been found in this way.[3,4]

In recent times, the physics of the free electron laser (FEL) has gained more and more attention, both on the experimental and theoretical side.[8]

The FEL provides one more example of unsuspected connections among apparently remote branches of physics. Indeed, it has been shown by one of the present authors that, in the framework of the single-mode Hamiltonian picture,[10] the main features of the FEL dynamics are described by Bloch-like equations, in analogy with the case of a two-level system.

In this paper, we want to show that, in the hypothesis of a multimode FEL operation,[11] the FEL dynamical behavior can be accounted for by introducing a suitable (n^2-1)-dimensional vector, in full analogy with the results obtained for the n-level atomic dynamics.[4] A preliminary suggestion in this direction was put forward in Ref. 12, where the FEL coherence properties are studied in the multimode picture by exploiting the SU(n) invariance of the FEL Hamiltonian. Indeed, besides coherence, this group-theoretical approach will enable us to critically discuss the FEL constants of motion.

The paper is organized as follows: The explicit construction of the SU(3) vector, in the illustrative example of two laser beams, is presented in Sec. II; Sec. III is concerned with the vector picture of the FEL dynamics and its implications on coherence and conservation laws; the generalization to the SU(n) case and the classical limit are outlined in Sec. IV and a brief summary is given in Sec. V.

II. FEL QUANTUM HAMILTONIAN AND SU(3)

Let us start by considering the simplified case of two copropagating laser beams, with different wave vectors and intensities, undergoing FEL amplification. The quantum nonrelativistic Hamiltonian describing this process can be written as[11]

$$\begin{aligned} H = {} & \frac{p^2}{2m} + \hbar \sum_{j=1}^{3} \omega_j a_j^+ a_j + \hbar\Omega_{1,2} \\ & \times [a_1^+ a_2 e^{-i(k_1-k_2)z} + \text{h.c.}] \\ & + \hbar\Omega_{1,3} [a_1^+ a_3 e^{-i(k_1+k_3)z} + \text{h.c.}] \\ & + \hbar\Omega_{2,3} [a_2^+ a_3 e^{-i(k_2+k_3)z} + \text{h.c.}] \,. \end{aligned} \tag{2.1}$$

Here, a_l^+ and a_l ($l = 1,2$) are the creation and annihilation operators for the laser fields (assumed propagating in the positive z direction), while a_3^+ and a_3 refer to the undulator field—treated as a radiation field in the Weizsacker–Williams approximation[13]—moving in the negative z direction; the $\mathbf{k}_i$ are the wave vectors of the laser beams ($i = 1,2$) and of the undulator ($i = 3$), and $\omega_i = |\mathbf{k}_i|c$ the corresponding frequencies. Moreover, we have put

$$\Omega_{ij} = 2\pi c^2 r_0/(\omega_i\omega_j)^{1/2}V,$$

where $r_0 = e^2/mc^2$ is the classical electron radius and V the mode volume.

The first two terms in H describe the electron and the free field energy, respectively (the electron motion is assumed nonrelativistic in the chosen frame), while the other terms represent the interaction between electron and fields.

In comparison with the single-mode case,[9] the Hamiltonian in Eq. (2.1) contains two more interaction terms. However, we can see that a similar group-theoretical approach can be used, if one defines the following pseudospin vector operators $\mathbf{R}_{ij}$ ($l < j = 1,2,3$):

$$\begin{aligned} R_{lj}^{(1)} &= \frac{a_l^+ a_j S_{lj}}{2} + \text{h.c.}, \\ R_{lj}^{(2)} &= \frac{a_l^+ a_j S_{lj}}{2i} + \text{h.c.}, \\ R_{lj}^{(3)} &= \frac{(a_l^+ a_l - a_j^+ a_j)}{2}, \end{aligned} \tag{2.2}$$

[a)] Permanent address: Dipartimento di Fisica, I Università di Roma "La Sapienza", P. le A. Moro, 2-00185 Rome, Italy; Istituto Nazionale di Fisica Nucleare, Sezione di Roma, Rome, Italy; and Division of Physics, Institute for Basic Research, 96 Prescott Street, Cambridge, Massachusetts 02138.

 0022-2488/85/123200-04$02.50

where

$$S_{ij} = \exp(-iK_{ij}z) \tag{2.3}$$

and

$$K_{ij} = \begin{cases} k_i - k_j, & j = 2, \\ k_i + k_j, & j = 3. \end{cases} \tag{2.4}$$

Moreover, let us put

$$\mathbf{R}_{12} = \mathbf{T}, \quad \mathbf{R}_{13} = \mathbf{V}, \quad \mathbf{R}_{23} = \mathbf{U}, \tag{2.5}$$

on analogy of the usual definitions of isospin, V spin, and U spin in the standard SU(3) scheme (see Ref. 2).

Then, the Hamiltonian (2.1) becomes

$$H = \frac{p^2}{2m} + \hbar \sum_j \omega_j a_j^+ a_j + 2\hbar(\Omega_{12} T_1 + \Omega_{13} V_1 + \Omega_{23} U_1). \tag{2.6}$$

It is easy to realize that the dynamics of the system described by the Hamiltonian (2.6) can be specified by the time evolution of the pseudospin vectors **T**, **U**, and **V**. However, such a description is redundant, since—as we shall see—the FEL dynamics is fully determined by the equation of motion of a single, eight-component vector $\vec{\mathbf{F}}$ [as it trivially follows by embedding the three SU(2) algebras (2.5) in a single SU(3) algebra].

The vector $\vec{\mathbf{F}}$ we are concerned with can be written, e.g., in the form

$$\vec{\mathbf{F}} = (F_\alpha) = (T_1, T_2, T_3, V_1, V_2, U_1, U_2, M) \quad (\alpha = 1,\ldots,8), \tag{2.7}$$

where M is linked to the standard hypercharge by

$$M = (\sqrt{3}/2)Y = (a_1^+ a_1 + a_2^+ a_2 - 2a_3^+ a_3)/2\sqrt{3}. \tag{2.8}$$

Needless to say, the commutation rules of the components of $\vec{\mathbf{F}}$ specify an SU(3) algebra:

$$[F_\alpha, F_\beta] = i f_{\alpha\beta\gamma} F_\gamma \tag{2.9}$$

[$f_{\alpha\beta\gamma}$ being the usual structure constants of SU(3)]. However, let us notice that, according to the definitions (2.2), the SU(3) generators F_α also contain the electron variables (through the factors S_{ij}). Therefore, we also have to consider the commutator of the electron momentum p with F_α, which explicitly reads

$$[p, F_\alpha] = \begin{cases} -i\hbar K_{lm}^{(\alpha)} [F_{\alpha+1}(\delta_{\alpha 1} + \delta_{\alpha 4} + \delta_{\alpha 6}) \\ \qquad - F_{\alpha-1}(\delta_{\alpha 2} + \delta_{\alpha 5} + \delta_{\alpha 7})], \quad \alpha \neq 3,8, \\ 0, \quad \alpha = 3,8, \end{cases} \tag{2.10}$$

where

$$K_{lm}^{(\alpha)} = \begin{cases} K_{12}, & \alpha = 1,2, \\ K_{13}, & \alpha = 4,5, \\ K_{23}, & \alpha = 6,7, \end{cases} \tag{2.11}$$

and $\delta_{\alpha\beta}$ is the Kronecker delta. Thus, strictly speaking, the total invariance group of our system is not SU(3), but rather its semidirect product by U(1), SU(3) $\otimes$ U(1).

III. VECTOR PICTURE OF FEL DYNAMICS, COHERENCE AND CONSERVATION LAWS

It is now straightforward to get the Heisenberg equations of motion of the vector $\vec{\mathbf{F}}$, which specify the dynamics of the FEL process under study. By taking the commutators of $\vec{\mathbf{F}}$ with the Hamiltonian (2.6), we have

$$F'_\alpha = f_{\alpha\beta\gamma} \Omega_\beta F_\gamma - i\epsilon_{\alpha\beta} F_\beta, \tag{3.1}$$

where the prime denotes derivative with respect to the dimensionless time $\tau = t/\Delta t$ (Δt being the interaction time); the vector $\vec{\Omega}$ and the matrix $\epsilon_{\alpha\beta}$ are given, respectively, by

$$\vec{\Omega} = (\Omega_\alpha) = (2\Omega_{12}\Delta t, 0, -w_{12}, 2\Omega_{13}\Delta t, 0, 2\Omega_{23}\Delta t, 0, \tfrac{1}{2} w_{12} - w_{13})$$
$$[w_{ij} = (\omega_{ij} p \Delta t/mc) + \Delta_{ij}, \quad \omega_{ij} = K_{ij} c, \quad \Delta_{ij} = \omega_i - \omega_j], \tag{3.2}$$

and

$$\epsilon_{\alpha\beta} = \epsilon_{ij}^{(\alpha)} (\delta_{\alpha\beta} - \delta_{\alpha 3} - \delta_{\alpha 8}), \tag{3.3}$$

where $\epsilon_{ij}^{(\alpha)}$ is defined in a way analogous to $K_{ij}^{(\alpha)}$ of Eq. (2.11), with the matrix ϵ_{ij} given by

$$\epsilon_{ij} = (\hbar \Delta t/mc^2)\omega_{ij}^2 \tag{3.4}$$

and connected to the recoil of the electron.

Equation (3.1) consists of two parts. The first is a generalized rotation in SU(3) space, in full analogy with the results of Refs. 3 and 4. The second part plays the role of a dephasing term and is due to the noncommutativity of the electron variables. As it is easy to realize, such a term causes coherent states of the system to evolve into noncoherent ones. Therefore, we recover in a straightforward way the result[14] that Glauber coherence[15] is not preserved in the FEL quantum operation, and that this effect is strictly connected to the electron recoil.

Let us now discuss the conservation laws. First of all, the global symmetry SU(3) $\otimes$ U(1) of the system implies the (obvious) conservation of the total momentum of electron and fields:

$$\mathbf{p} + \sum_{j=1}^{3} \mathbf{k}_j a_j^+ a_j = \text{const} \tag{3.5}$$

(for further comments, see Ref. 11.) However, the most direct constants of motion deducible from our formalism are those related to the Casimir operators of the unitary group concerned. In this connection, let us notice that actually the invariance group of the Hamiltonian (2.6) is U(3), rather than SU(3), and, therefore, we have to consider three Casimir invariants, which can be comprised in the single formula[16]

$$C^{(p)} = \sum_{\underbrace{ji \cdots k}_{p} = 1}^{3} A_{ji} A_{ik} \cdots A_{kj} \quad (p = 1,2,3), \tag{3.6}$$

where

$$A_{ji} = a_j^+ a_i, \quad i \neq j, \tag{3.7}$$

$$A_{jj} = \begin{cases} a_j^+ a_j, & \text{for U(3)}, \\ a_j^+ a_j - \dfrac{1}{3} \displaystyle\sum_{k=1}^{3} a_k^+ a_k, & \text{for SU(3)} \end{cases} \tag{3.8}$$

(or suitable linear combinations of the above operators[17]). The first Casimir invariant [linked to the U(3) group] is thus given by

$$C^{(1)} = \sum_{k=1}^{3} a_k^+ a_k \tag{3.9}$$

and yields the FEL Manley–Rowe pseudoquantum rule

$$\sum_{i=1}^{2} n_L^i + n_U = \text{const}\,, \tag{3.10}$$

where n_L^i, n_U are the laser and undulator photon numbers, respectively (i.e., $C^{(1)}$ accounts for the conservation of the total number of photons). The second Casimir invariant is nothing but the length of the vector $\vec{\mathbf{F}}$. If a and b are the two non-negative integers labeling the irreducible representations of SU(3), the eigenvalues of $C^{(2)}$ explicitly read[18]

$$C^{(2)} = \tfrac{1}{12}(a-b)^2 + \tfrac{1}{4}(a+b)(a+b-4)\,. \tag{3.11}$$

Finally, the eigenvalues of the third Casimir invariant for any SU(3) irreducible representation are given by[18]

$$C^{(3)} = \tfrac{1}{18}(a-b)(2a+b+3)(2b+a+3)\,. \tag{3.12}$$

Although the two nonlinear Casimir operators $C^{(2)}$ and $C^{(3)}$ correspond, in principle, to new constants of motion, it is easy to see that, in the present case, they do not provide any new conservation law and simply restate Eq. (3.10). Indeed, all physically realizable states of our system must belong to a totally symmetric representation of SU(3), and it is well known[18] that, in this case [as immediately follows from Eqs. (3.11) and (3.12) for $b=0$], the two Casimir operators of SU(3) are related to each other and to the total number of photons. By the way, let us underline that, even in the case of the atomic n-level system considered in Refs. 3 and 4, the "new, nonlinear" constants of motion derived therein may no longer be considered as independent if one assumes that the initial state has a definite permutation symmetry. However, let us notice that there is (at least in principle) some difference between the conserved quantities defined in Refs. 3 and 4 and the Casimir invariants (3.11) and (3.12). Indeed, as pointed out recently,[19] the constants of motion expressed in terms of the elements of the density matrix do not coincide, in general, with the Casimir operators of U(n). To restore a complete analogy between the n-level atomic system and the multimode FEL operation would therefore require us to consider a multielectron theory of FEL and define a suitable density matrix by averaging on the electron and field variables.

IV. GENERALIZATION TO THE SU(n) CASE

The group-theoretical formalism developed in the previous sections for the three-mode FEL system is easily generalized to the n-mode case. The quantum nonrelativistic FEL Hamiltonian now reads[11]

$$H = \frac{p^2}{2m} + \sum_{j=1}^{n} \omega_j a_j^+ a_j \hbar \sum_{l<j=1}^{n} + \hbar \sum_{l<j=1}^{n} \Omega_{lj} a_l^+ a_j \exp(-iK_{lj}\, z) + \text{h.c.}, \tag{4.1}$$

where

$$K_{lj} = \begin{cases} k_l - k_j\,, & j \neq n\,, \\ k_l + k_j\,, & j = n\,, \end{cases} \tag{4.2}$$

and the index n refers to the undulator field.

In terms of the pseudospin vector operators $\mathbf{R}_{lj}$ ($l<j=1,\ldots,n$) [still defined by Eqs. (2.2)], (4.1) can be written as

$$H = \frac{p^2}{2m} + \sum_{j=1}^{n} \omega_j a_j^+ a_j + 2\hbar \sum_{l<j=1}^{n} \Omega_{lj} R_{lj}^{(1)}\,. \tag{4.3}$$

The (n^2-1)-dimensional vector $\vec{\mathbf{F}}$ describing the FEL dynamics may be chosen as follows:

$$\vec{\mathbf{F}} = (F_\alpha) = (R_{12}^{(1)}, R_{12}^{(2)}, \ldots, R_{n-1,n}^{(1)}, R_{n-1,n}^{(2)}, W_1, \ldots, W_{n-1}) \quad (\alpha = 1,\ldots,n^2-1)\,, \tag{4.4}$$

where we have assumed the standard form of the SU(n) diagonal generators W_j ($j=1,\ldots,n-1$):

$$W_j = [2j(j+1)]^{-1/2}(a_1^+ a_1 + \cdots + a_j^+ a_j - j a_j^+ a_j)\,. \tag{4.5}$$

The components of $\vec{\mathbf{F}}$ still satisfy the commutation rules (2.9), with $f_{\alpha\beta\gamma}$ being now the SU(n) structure constants. As to their commutators with the electron momentum p, they are given by a suitable generalization of Eq. (2.10), which is easily evaluated by taking into account that

$$[p, R_{lm}^{(1,2)}] = \mp i\hbar K_{lm} R_{lm}^{(2,1)}, \quad [p, W_j] = 0\,. \tag{4.6}$$

Then, it is easy to see that Eq. (3.1) still describes the dynamics of the system, with suitable generalizations of the definitions (3.2) and (3.3) of the [now (n^2-1)-dimensional] vector $\vec{\Omega}$ and of the matrix $\epsilon_{\alpha\beta}$. Obviously, the implications on Glauber coherence remain unchanged.

The constants of motion are given (apart from the conservation of total momentum and number of photons) by the Casimir invariants of SU(n):

$$C^{(p)} = \sum_{\underbrace{j_i \cdots k=1}_{p}}^{n} A_{ji} A_{i\cdot} \cdots A_{\cdot k} A_{kj} \quad (p=2,\ldots,n)\,, \tag{4.7}$$

where the operators A_{ij} are still defined by Eqs. (3.7) and (3.8) (with $3 \to n$).

As in the SU(3) case, the invariants $C^{(p)}$ give rise to a set of $(n-1)$ nonlinear conserved quantities, which are in principle independent, as long as one does not take explicitly into account the permutation symmetry of the system states. In fact, they all collapse into the single law of photon number conservation for totally symmetric SU(n) representations.

Eventually, we want to stress that the analysis we have carried out within a quantum framework can be also performed from a classical viewpoint. To this aim, one has merely to substitute everywhere the commutators by the Poisson brackets; e.g., by Eq. (2.9) one gets

$$[F_\alpha, F_\beta] = i f_{\alpha\beta\gamma} F_\gamma\,. \tag{4.8}$$

Moreover, one needs to take the classical image of the SU(n) generators, namely, for instance

$$F_1 = R_{12}^{(1)} \to (I_1 I_2)^{1/2} \cos[K_{12} z + (\varphi_1 - \varphi_2)]\,, \tag{4.9}$$

where I_j and φ_j are the action and phase, respectively, of the jth field.

The equation of motion (3.1) in the classical case becomes simply

$$F'_\alpha = f_{\alpha\beta\gamma} \Omega_\beta F_\gamma\,, \tag{4.10}$$

since no extra term due to the noncommutativity of the electron variables appears. Therefore, the effect of the FEL interaction in this case amounts to a mere rotation of the vector

$\vec{F}$ in the SU(n) space. In other words, the absence of the dephasing term causes coherent SU(n) states of the system to evolve into coherent states. Let us stress that this result holds true also in a "semiclassical" framework, i.e. when treating the electron classically (see Ref. 12).

We also want to notice that, according to Mukunda's theorem,[20] in the classical limit the FEL system under consideration possess invariance not only under SU(n), but also under $O(n+1)$.

As for the constants of motion, they are obtained by the Casimir operators by taking the limit (4.9). However, at the light of the considerations made in this section and in Sec. III, it is easy to see that in the classical limit, too, they do not provide any new physical conservation law.

V. SUMMARY

It has been shown that the invariance of the FEL quantum Hamiltonian in a multimode picture under U(n) $\otimes$ U(1) permits the description of the FEL dynamics in terms of the (n^2-1)-dimensional vector $\vec{F}$, built up by the SU(n) generators (in analogy to what has been recently done for an n-level atomic system). The equation of motion for $\vec{F}$ amounts to a rotation in SU(n) space plus a dephasing term (due to the noncommutativity of the electron variables). Then, one recovers in a straightforward way (without any recourse to explicit solutions) the result that coherence is not preserved under quantum FEL operation. Coherence is obviously regained in the classical and in the semiclassical limit (namely, when treating only the electron classically). The constants of motion [apart from the total momentum conservation), connected to the U(1) group] are given by the Casimir operators of U(n). However, it has been stressed that—if one takes into account the permutation symmetry of the system states—they do not provide any new physical conservation law and merely restate the FEL Manley–Rowe pseudoquantum rule (i.e., the conservation of the total number of photons). The way of performing the transition to the classical limit of the present group-theoretical approach has been explicitly outlined.

ACKNOWLEDGMENTS

We are greatly indebted to A. Dipace and A. Renieri for enlightening comments. Thanks are also due to F. Buccella, C. D. Cantrell, and W. B. Colson for very useful discussions. One of us (R. M.) acknowledges the warm hospitality of the E. N. E. A. Frascati Laboratories, where this work has been carried out.

[1]E. P. Wigner, Phys. Rev. **51**, 106 (1937); J. P. Elliot, Proc. R. Soc. London Ser. A **245**, 128 (1958); P. Franzini and L. Radicati, Phys. Lett. **6**, 322 (1963).
[2]See, e.g., M. Gell-Mann and Y. Ne'eman, *The Eightfold Way* (Benjamin, New York, 1964).
[3]J. N. Elgin, Phys. Lett. A **80**, 140 (1980).
[4]E. T. Hioe and J. H. Eberly, Phys. Rev. Lett. **47**, 838 (1981); Phys. Rev. A **25**, 2168 (1982).
[5]E. T. Hioe, Phys. Rev. A **28**, 879 (1983).
[6]R. P. Feynman, F. L. Vernon, Jr., and R. W. Hellwarth; J. Appl. Phys. **28**, 49 (1957).
[7]T. W. Mossberg and S. R. Hartmann, Phys. Rev. A **23**, 1271 (1981).
[8]G. Dattoli and A. Renieri, ENEA preprint, 1983 [to appear in *Laser Handbook*, edited by M. L. Stitch and M. S. Bass (North-Holland, Amsterdam), Vol. 4].
[9]G. Dattoli, Lett. Nuovo Cimento **27**, 247 (1980).
[10]A. Bambini and A. Renieri, Lett. Nuovo Cimento **21**, 399 (1978); A. Bambini, A. Renieri, and S. Stenholm, Phys. Rev. A **19**, 2013 (1979).
[11]G. Dattoli and A. Renieri, Nuovo Cimento B, **61**, 153 (1981).
[12]G. Dattoli, A. Renieri, and F. Romanelli, Opt. Commun. **35**, 245 (1980).
[13]J. M. J. Madey, J. Appl. Phys. **42**, 1906 (1971).
[14]See W. B. Becker and M. S. Zubairy, Phys. Rev. A **25**, 2200 (1982); G. Dattoli and M. Richetta, Opt. Commun. **50**, 165 (1984). See also Ref. 12 for a first application of SU(n) invariance to the study of coherence in FEL operation.
[15]Strictly speaking, the term "Glauber coherence" applies only to eigenstates of the annihilation operator, i.e., in the framework of the harmonic oscillator algebra. However, the concept of coherence can be extended to larger unitary groups. See F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6**, 2211 (1972).
[16]See, e.g., A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (Polish Scientific, Warsaw, 1980).
[17]For instance, in the SU(3) case the following choice of diagonal operators A_{ii} is currently done in the literature: $A_{11}=I_3+(Y/2)=Q$, $A_{22}=-I_3+(Y/2)=Q_V, A_{33}=-Y$, where Q is the charge and Q_V the scalar of the V-spin subalgebra. See, e.g., J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963).
[18]See, e.g., L. Fonda and G. C. Ghirardi, *Symmetry Principles in Quantum Physics* (Dekker, New York, 1970).
[19]F. T. Hioe, Phys. Rev. A **29**, 3434 (1984).
[20]N. Mukunda, Phys. Rev. **155**, 1383 (1967).

ERRATUM

Erratum: Lie transformation group solutions of the nonlinear one-dimensional Vlasov equation [J. Math Phys. 26, 1428 (1985)]

B. Abraham-Shrauner
Department of Electrical Engineering, Washington University, St. Louis, Missouri 63130

(Received 10 July 1985; accepted for publication 29 August 1985)

A parenthesis is missing to the right of the equal sign in front of $\rho\ddot{\rho}$ in Eq. (34). The functions $\sin \bar{t}$ ($\cos \bar{t}$) in the time derivative in Eq. (58) for I_1 (I_2) should be changed to $\cos \bar{t}$ ($\sin \bar{t}$). The bar over x in the partial derivative of $\bar{f}$ in Eq. (59) should be moved one space to the right. The term $(N^2/2)$ in the first and second lines below Eq. (59) should be N^2 only.

 0022-2488/85/123204-01$02.50