On the classification of Clifford algebras and their relation to spinors in n dimensions

Nikos Salingaros
Mathematics Department, University of Iowa, Iowa City, Iowa 52242

(Received 19 June 1981; accepted for publication 21 August 1981)

A classification of all the Clifford algebras is given in terms of Kronecker products of the quaternion and dihedral groups. The relationship to spinors in n dimensions is explicitly determined. We show that the real Clifford algebra in Minkowski spacetime is distinct from both the algebra of Dirac matrices and the algebra of Majorana matrices, and cannot be realized by the spinor framework. The matrix representations of Clifford algebras are discussed, and are utilized to give a classification of the real forms of Lie algebras. We are thus able to relate Clifford, Lie, and spinor algebras in an intrinsic geometrical setting.

PACS numbers: 02.10.Tq

I. INTRODUCTION

The apparatus of theoretical physics consists in large part of certain algebraic structures which arise naturally in the description of physical phenomena. The Clifford algebras[1–4] appear to offer a framework for a unified setting of many of these algebras. (For various applications of Clifford algebras to physics, see references in Ref. 5.)

In the past, we have recast the Clifford algebras in a form suitable for calculations and manipulations in physics, by using the basis of differential forms in a Riemannian space.[5] The advantage is that we were able to utilize the entire apparatus from the theory of differential forms (exterior product, duality, and the geometrical interpretation) in addition to the useful manipulatory properties of the Clifford algebras. Our classification of all Clifford algebras included many of the algebras that have appeared in physics, and placed them in a useful geometrical setting.[5]

The purpose of the present paper is to further the task begun in Ref. 5 towards providing a simple and unified algebraic framework for use in physics. First, we show how the Clifford algebras can be related in a very simple manner to Kronecker (tensor) products of the quaternion and dihedral groups. This provides a means of directly constructing the larger Clifford algebras, and obtaining relations between them. The Classification theorem (Theorem 5) gives the structure of all the Clifford algebras. The content of this theorem is given in an easily accessible manner in Tables I and II, which include, in particular, the relationship of Clifford algebras to Riemannian spaces.

Second, we relate some of the Clifford algebras to spinors in n dimensions.[6–8] What is important is that many Clifford algebras *cannot* be related to the spinor framework (Theorem 6). An example is the quaternion algebra H. Another physically important example is the real Clifford algebra in Minkowski spacetime $A^{1,3} \approx N_4$ [elsewhere called $H(2)$ or $M_2(H)$]. This algebra is distinct from both the algebra of Dirac matrices and the algebra of Majorana matrices (which we demonstrate here) and cannot be realized by the spinor framework. This point has not been discussed previously.

Next, we determine the Lie algebra which corresponds to each Clifford algebra, in two distinct ways. First, we give the orthogonal Lie algebras $SO(r,s)$ realized by the Clifford algebras via the Lie bracket (Table III). These algebras are in fact larger than the spin (n) algebras which are obtained from the usual spinor construction, and include spin (n) as a subalgebra.

Second, we give the matrix representation space of each Clifford algebra, which in turn enables us to obtain the full matrix Lie algebra corresponding to each Clifford algebra (Theorem 9). The previously determined relationship to the orthogonal Lie algebras, along with this result, gives us a useful list of Lie algebra isomorphisms (Table IV). In addition, Theorem 9 combined with the Classification theorem (Theorem 5) determines the real forms $SL(k;H)$ and $SL(2k;\mathbb{R})$ of the complex Lie algebra $SL(2k;\mathbb{C})$. These last results, while not new, demonstrate the usefulness of the construction by the simplicity and ease with which they are obtained.

A novel feature of our discussion is the utilization of the abelian algebra $\mathbf{\Omega}$ (elsewhere denoted $\mathbb{R} \oplus \mathbb{R}$) in a key manner in the classification of Clifford algebras. The algebra $\mathbf{\Omega}$ has properties akin to the complex field $\mathbb{C}$, yet $\mathbf{\Omega}$ is not a conventional field. The properties of $\mathbf{\Omega}$ were discussed in detail in Ref. 9.

II. THE DIFFERENTIAL FORM REALIZATION OF CLIFFORD ALGEBRAS

A Clifford algebra[1–4] is defined via the anticommuting bases,

$$e^i e^j = -e^j e^i, \quad i \neq j, \quad i,j = 1,\ldots,n,$$
$$(e^i)^2 = +1 \text{ or } -1. \tag{1}$$

We have previously joined the theory of Clifford algebras to the theory of differential forms by identifying the basis e^i with the basis one-forms $dx^i = \sigma^i$ of a flat n-dimensional Riemannian space.[5] The metric of this space is $g^{ij} = (\sigma^i, \sigma^j)$ and has diagonal entries either $+1$ or -1 (all off-diagonal entries are zero). We can realize the product (1) in terms of the basis one-forms, the metric, and the Grassmann–Cartan exterior product $\wedge$. (For the apparatus of differential forms, see Ref. 10.) This realization is discussed at length in Ref. 5. The product was called the "vee product," and was denoted by v. For the basis one-forms the vee prod-

 0022-2488/82/010001-07$02.50

uct obeys the following rules:

$$\sigma^i \vee \sigma^j = \sigma^i \wedge \sigma^j, \quad i \neq j,$$
$$\sigma^i \vee \sigma^i = g^{ii} \text{ (no sum)}. \tag{2}$$

By repeated application of the vee product, we can generate all the basis p forms. In particular, we single out the volume element $\omega^n = \sigma^1 \wedge \cdots \wedge \sigma^n = \sigma^1 \vee \cdots \vee \sigma^n$, which is important in the following section. In each n-dimensional space there are 2^n basis forms, all of which can be manipulated with the vee product. The vee product between a basis r form and a basis s form is given in two steps, as follows: (i) Identify the k indices that the two basis forms have in common, and permute them into canonical form (below), (ii) contract between identical indices.

This prescription gives the general definition of the vee product between two specific basis forms.

Definition 1:

$$(\sigma^{i_1} \wedge \cdots \wedge \sigma^{i_r}) \vee (\sigma^{j_1} \wedge \cdots \wedge \sigma^{j_s})$$
$$= (-1)^{\pi_r}(-1)^{\pi_s}(\sigma^{\lambda_1} \wedge \cdots \wedge \sigma^{\lambda_{r-k}} \wedge \sigma^{\nu_1} \wedge \cdots \wedge \sigma^{\nu_k}),$$
$$\vee(\sigma^{\nu_k} \wedge \cdots \wedge \sigma^{\nu_1} \wedge \sigma^{\mu_1} \wedge \cdots \wedge \sigma^{\mu_{s-k}})$$
$$= (-1)^{\pi_r}(-1)^{\pi_s} g^{\nu_1\nu_1} \cdots g^{\nu_k\nu_k} \sigma^{\lambda_1} \wedge \cdots \wedge \sigma^{\lambda_{r-k}} \wedge \sigma^{\mu_1} \wedge \cdots \wedge \sigma^{\mu_{s-k}}. \tag{3}$$

Here the factors of $(-1)^{\pi_r}$ and $(-1)^{\pi_s}$ arise, respectively, from the permutations

$$\begin{pmatrix} i_1 \cdots i_k & i_{k+1} \cdots i_r \\ \lambda_1 \cdots \lambda_{r-k} & \nu_1 \cdots \nu_k \end{pmatrix}, \tag{4a}$$

$$\begin{pmatrix} j_1 \cdots j_k & j_{k+1} \cdots j_s \\ \nu_k \cdots \nu_1 & \mu_1 \cdots \mu_{s-k} \end{pmatrix}. \tag{4b}$$

The crucial difference between the vee product and the well-known exterior product is that the vee product of two forms is always another form, whereas the exterior product of two forms is zero if the sum of their ranks exceeds the dimension of the space. Another important property is the existence of a unique inverse in vee: every basis form has an inverse equal to the form itself up to a sign.

These properties demonstrate that the set of all basis forms define a finite group under the vee-product, called the "vee group" of that particular Riemannian space.[5] We have shown that the group algebra over $\mathbb{R}$ of each "vee" group is isomorphic (modulo Z_2) to the real Clifford algebra corresponding to that space.[5]

The Z_2 grading is due to the following observation. In the finite vee group, one must consider positive and negative basis forms as distinct elements, while in the corresponding group algebra over $\mathbb{R}$, the distinction between positive and negative bases is not made. The group algebra must therefore be divided by the group of two elements $\{1, -1\} \approx Z_2$.

The elements of the Clifford algebra in this realization are antisymmetric tensor fields expanded on the basis p forms, or equivalently, linear combinations of the p forms. These possess algebraic properties above and beyond those expected from the usual exterior algebra, in particular the possibility of division, since the basis is endowed with the vee product. The vee product between the basis (2,3) gives a set of rules for the product of the field components. This product has been previously discussed in Ref. 5 and is not needed in the present analysis.

III. THE VEE-GROUP STRUCTURE

In this section, we give a construction of the vee group of differential forms G^n in each n-dimensional Riemannian space by using group-theoretical methods.

We first recall some results from Refs. 1, 4, and 5.

In the vee multiplication, the volume element ω^n anticommutes with all the forms in G^n when n = even, and commutes when n = odd. Hence, the center of each group will contain the elements $\{1, -1\}$ for n = even, and $\{1, -1, \omega^n, -\omega^n\}$ when n = odd. The actual groups defined by the corresponding centers depend upon whether $(\omega^n)^2 = +1$ or -1, which is in turn determined by the signature and dimension of the metric in each case. The center of each group of forms is given by the following theorem.[5] Here, Z_2 is the cyclic group of order 2; Z_4 is the cyclic group of order 4, isomorphic to the complex group; and D_2 is the dihedral group of order 4, isomorphic to the Gauss–Klein Veergruppe $Z_2 \otimes Z_2$.

Theorem 1: The center of the vee group G^n is isomorphic to the finite group Z_2, when n = even; Z_4, when n = odd and $(\omega^n)^2 = -1$; and $Z_2 \otimes Z_2 = D_2$ when n = odd and $(\omega^n)^2 = +1$.

Using Theorem 1, we gave in Ref. 5 a key result of the group structure as the following theorem:

Theorem 2: The factor group G^n modulo the center of G^n, is the abelian group $(Z_2)^n = Z_2 \otimes \cdots \otimes Z_2$ (n times); and is given by the three distinct cases:

$$n = \text{even}, \quad G^n/Z_2 = (Z_2)^n = (D_2)^{n/2}, \tag{5a}$$

$$n = \text{odd}, \quad (\omega^n)^2 = -1: \; G^n/Z_4 = (Z_2)^{n-1} = (D_2)^{(n-1)/2}, \tag{5b}$$

$$n = \text{odd}, \quad (\omega^n)^2 = 1: \; G^n/D_2 = (Z_2)^{n-1} = (D_2)^{(n-1)/2}. \tag{5c}$$

We proceed to apply some general group-theoretical results to the construction of the vee groups. First, since the vee group G^n is of order 2^{n+1}, it is referred to as a "2-group" in the mathematical literature.[11] Second, Theorem 2 demonstrates that the vee groups G^n are "extra-special 2-groups," defined as follows[11]:

Definition 2: G is an "extra-special 2-group" iff G/center (G) = abelian, and center (G) is a 2-group.

We now recall a general theorem from the theory of group representations[11]:

Theorem 3: Every extra-special p-group G is the Kronecker product of nonabelian p groups of order p^3, and so has order p^{2m+1} for some m.

In the special case of interest, $p = 2$ for the vee groups G^n. There are only two nonisomorphic nonabelian vee groups of order $2^3 = 8$, and they are the quaternion group Q_4 and the dihedral group of order 8, D_4.[5]

In Ref. 5, Table IV, we gave the following result (as $N_1 \otimes N_1 = N_2 \otimes N_2$).

$$Q_4 \otimes Q_4 = D_4 \otimes D_4. \tag{6}$$

Hence, all mixed Kronecker products of m copies of the groups Q_4 and D_4 will reduce to either $(D_4)^m$ or $Q_4 \otimes (D_4)^{m-1}$.

Putting these results together, we obtain the corollary to the Frobenius–Schur theorem.[11,12]

Theorem 4: For any extra-special 2-group G of order 2^{2k+1},

$$G \approx (D_4)^k \text{ or } G \approx Q_4 \otimes (D_4)^{k-1}.$$

Theorem 4 can be rewritten as a corollary after identifying the extra-special 2-groups with the vee groups. From Theorems 1 and 2, Definition 1, and recalling the order of G^n as 2^{n+1}, we have the

Corollary to Theorem 4: For any vee group G^{2k},

$$G^{2k} \approx (D_4)^k \text{ or } G^{2k} \approx Q_4 \otimes (D_4)^{k-1}. \tag{7}$$

This corollary will be utilized in the next section to construct all the Clifford algebra $A^{p,q}$ as the groups algebras of the vee groups. Previously we derived some relations between the vee groups which were listed in Table IV, Ref. 5. They in fact provided explicit proofs of special cases of Theorem 4 and its corollary.

IV. CLASSIFICATION OF CLIFFORD ALGEBRAS

The Ref. 5, we identified the Clifford algebras $A^{p,q}$, corresponding to each Riemannian space $M^{p,q}$, by direct construction. [The notation used in the following: the metric of the space $M^{p,q}$ contains p plus ones and q minus ones on the diagonal, zeros elsewhere, and it is used in the definition of the vee product (2,3). The dimension of the space is $n = p + q$.]

The most remarkable result of the construction is the demonstration that certain algebras of the same dimension but distinct signature are in fact isomorphic. These isomorphisms prompted the notation for the Clifford algebras introduced in Ref. 5, and which is discussed below. In this section, we show how these isomorphisms are a direct consequence of the Frobenius–Schur theorem.[12] This appears to be the first simple explanation of this fundamental property of the Clifford algebras.

A novel feature of our discussion is the utilization of the abelian algebra $\mathbf{\Omega}$ in a key manner. $\mathbf{\Omega}$ is isomorphic to the Clifford algebra $A^{1,0}$ which is generated by a basis one-form ω with the property that $\omega \vee \omega = +1$. Every element $a \in \mathbf{\Omega}$ is of the form $a = x + \omega y$; $x, y \in \mathbb{R}$. It is easy to see the similarity to complex algebra $\mathbb{C}$, which is itself isomorphic to the Clifford algebra $A^{0,1}$. A detailed discussion of the properties of $\mathbf{\Omega}$, as well as the reason why $\mathbf{\Omega}$ is not a conventional field, was given in Ref. 9.

We present below as Theorem 5, the construction and classification of all the Clifford algebras. The notation is as follows: $N_2 = H$ is the familiar quaternion algebra and N_1 is an algebra with three anticommuting elements whose squares are $+1$, $+1$, -1, respectively.[13] N_1 [elsewhere denoted $\mathbb{R}(2)$ or $M_2(\mathbb{R})$] is isomorphic to the group algebra (Z_2-graded) of the dihedral group D_4.[5] Finally note that the definition of the algebras N_5 and N_6 given in Ref. 5 has been interchanged.

Theorem 5:

(a) The Clifford algebras are constructed from the real algebras with three anticommuting elements N_1 and $N_2 = H$ (which are of dimension 4), and from the abelian algebras $\mathbf{\Omega}$ and $\mathbb{C}$ (which are of dimension 2).

(b) There are only two nonisomorphic algebras of dimension 2^{2k}. They are labelled as N_m, m = odd or even, and they are obtained as Kronecker products of the algebras N_1 and $N_2 = H$ as follows:

$$N_{2k-1} = (N_1)^k, \tag{8a}$$

and

$$N_{2k} = N_2 \otimes (N_1)^{k-1}. \tag{8b}$$

(c) There are only three nonisomorphic algebras of dimension 2^{2k+1}. They are labelled as S_k, k = integer, and Ω_m, m = odd or even, and are given by the Kroenecker product of the N algebras with $\mathbf{\Omega}$ and $\mathbb{C}$.

$$S_k = N_{2k} \otimes \mathbb{C} \approx N_{2k-1} \otimes \mathbb{C}, \tag{9a}$$

$$\Omega_{2k-1} = N_{2k-1} \otimes \mathbf{\Omega}, \tag{9b}$$

$$\Omega_{2k} = N_{2k} \otimes \mathbf{\Omega}. \tag{9c}$$

(d) The identification of the Clifford algebras with each Riemannian space is given by their position in Table I. The table can be indefinitely extended downwards for the S_k algebras with k = integer, and for the N_m and Ω_m algebras, with m = odd or even.

(e) The periodicity of the Clifford algebras is generated by N_4, the Clifford algebra in Minkowski spacetime.

The proof of Theorem 5 will be indicated here: part (b) follows from translating the corollary to Theorem 4 [Eq. (7)] from the vee groups to the Clifford algebras; part (c) combines results first obtained in Refs. 9 and 14, which essentially follow from Theorem 1 of this paper; part (d) is an extension of the results of Ref. 5; part (e) is intrinsically related to the Bott periodicity,[3] and can be graphically deduced from Table I.

There is a large amount of information contained in this theorem, which we proceed to discuss in stages. First, we indicate which of the Clifford algebras are otherwise known in physics.

Of the N algebras, N_1 is related to the elementary (real) spinors,[13] while N_2 is isomorphic to the quaternion algebra H.[5,13] N_3 was first realized by the 4×4 real matrices of Majorana.[5,14] N_4 is isomorphic to the Clifford algebra in Minkowski spacetime, which has been discussed in Ref. 5 and 14–16.

Of the S algebras, S_1 is isomorphic to the algebra of Pauli matrices,[5,13] while S_2 is isomorphic to the algebra of Dirac matrices.[5,14] These two algebras are related to the complex spinors of dimension 2 and 4, respectively, as is well known (see Sec. 5).

Of the Ω algebras, only one has been given an explicit realization: Ω_2 (elsewhere denoted $H \oplus H$) is isomorphic to the "biquaternions" of Clifford.[9]

We note that the N, S, and Ω algebras are frequently referred to in terms of their matrix representation space. This is discussed in Sec. 7 of this paper. Theorem 8 [Eq. (19)] can be used to relate the matrix notation of the Clifford algebras to the one discussed here.

The Clifford algebra corresponding to a Riemannian space can be found from its position on the triangular grid of

TABLE I. Classification of Clifford algebras in each Riemannian space.

Dimension of Algebra		Dimension of Riemann Space
1	$\mathbb{R}$	n = 0
2	Ω $\mathbb{C}$	1
4	N_1 N_1 $N_2=H$	2
8	S_1 Ω_1 S_1 Ω_2	3
16	N_4 N_3 N_3 N_4 N_4	4
32	Ω_4 $S_2=D$ Ω_3 $S_2=D$ Ω_4 $S_2=D$	5
64	N_6 N_6 N_5 N_5 N_6 N_6 N_5	6
128	S_3 Ω_6 S_3 Ω_5 S_3 Ω_6 S_3 Ω_5	7
256	N_7 N_8 N_8 N_7 N_7 N_8 N_8 N_7 N_7	8
512	Ω_7 S_4 Ω_8 S_4 Ω_7 S_4 Ω_8 S_4 Ω_7 S_4	9
1024	N_9 N_9 N_{10} N_{10} N_9 N_9 N_{10} N_{10} N_9 N_9 N_{10}	10
⋮	⋮ ⋮ ⋮	⋮

Table I. The coordinates of the grid (p,q) correspond to the metric of the Riemannian space $M^{p,q}$; each row has the same dimension $n = p + q$, and the entries are ordered from left to right as $(p = n, q = 0)$, $(p = n - 1, q = 1)$,..., $(p = 0, q = n)$.

To give a specific example, the algebra $A^{2,2}$ is defined by the anticommutation relation $\{\sigma^\mu,\sigma^\nu\} = 2g^{\mu\nu}, \mu, \nu = 1,...,4$, with the metric $g^{\mu\mu} = \text{diagonal}\,(+1, +1, -1, -1)$. $A^{2,2}$ is related to a Riemannian space of dimension $n = 4$, and is an algebra of dimension $2^4 = 16$. The entries on the $n = 4$ line of Table I correspond to the algebras $A^{4,0}, A^{3,1}, A^{2,2}, A^{1,3}$, and $A^{0,4}$, respectively. Hence, the Clifford algebra $A^{2,2}$ corresponds to N_3, which happens to be isomorphic to the algebra of the Majorana matrices.[5]

As an example of the isomorphism between algebras of the same dimension but distinct metric, the Clifford algebras of dimension 32 are given in Table I, row $n = 5$, as

$$A^{5,0} \approx \Omega_4,\quad A^{4,1} \approx S_2,\quad A^{3,2} \approx \Omega_3,$$
$$A^{2,3} \approx S_2,\quad A^{1,4} \approx \Omega_4,\quad A^{0,5} \approx S_2. \tag{10}$$

From (10) we have the result that the Clifford algebras $A^{4,1}$, $A^{2,3}$, and $A^{0,5}$ are all isomorphic to the Dirac algebra $S_2 = D$.

A remark is necessary on the nature of the Clifford algebra isomorphisms. In the differential form realization, the elements of each Clifford algebra are antisymmetric tensor fields, or linear combinations thereof. A sharp distinction is made between tensor fields of distinct rank, because of the geometry. Hence each Clifford algebra $A^{p,q}$ is geometrically distinct, for each distinct space $M^{p,q}$. However, when one ignores the geometrical construction (i.e., when one has an abstract or matrix representation), one can distinguish algebras only if their underlying vee-groups are distinct. The Clifford algebras that are algebraically distinct are in fact those given by Theorem 5. One should therefore always keep in mind that the concept of an isomorphism is an algebraic one and not a geometrical one.

Theorem 5 gives relations which have direct physical relevance. For example, it clears up the old question concerning the relationship of the Dirac algebra to the spaces with metric $(+1, +1, +1, -1)$ and $(-1, -1, -1, +1)$. From Table I, row $n = 4$, we identify the algebras corresponding to these two metrics as $A^{3,1} \approx N_3$ and $A^{1,3} \approx N_4$, respectively. Hence the different metrics give rise to two *distinct* algebras. N_3 is isomorphic to algebra of the Majorana matrices, while N_4 is isomorphic to the Clifford algebra in Minkowski spacetime. The Dirac algebra $D = S_2$ appears only on the row $n = 5$ of Table I, and therefore corresponds to a Clifford algebra in *five* dimensions. This surprising result can be explained by Theorem 5. Equation (9a) for $k = 2$ gives the relation $S_2 = N_4 \otimes \mathbb{C} \approx N_3 \otimes \mathbb{C}$. Hence, the Dirac algebra is isomorphic to the complexification of the real Clifford algebra in Minkowski spacetime N_4, or equivalently, to the complexification of the algebra of the real Majorana matrices N_3.[14] Since field theory usually involves complex quantities, the above distinction is not noticed in actual practice. When one deals with real spinor fields, however, the difference is crucial (see also the following section).

The structural properties of the Clifford algebras are summarized in Table II, where we have listed the distinct

TABLE II. Structure of the Clifford algebras in terms of finite groups.

Dimension of algebra	Algebra	Corresponding vee group	Dimension of corresponding Riemann space
1	$\mathbb{R}$	Z_2	0
2	Ω	$D_2 = Z_2 \otimes Z_2$	1
	$\mathbb{C}$	Z_4	
2^{2k}	N_{2k-1}	$(D_4)^k$	$2k$
$(k = 1,2,3...)$	N_{2k}	$Q_4 \otimes (D_4)^{k-1}$	
2^{2k+1}	S_k	$Z_4 \otimes (D_2)^k$	$2k + 1$
	Ω_{2k-1}	$Z_2 \otimes (D_4)^k$	
	Ω_{2k}	$Z_2 \otimes Q_4 \otimes (D_4)^{k-1}$	

Clifford algebras in each dimension, along with the corresponding vee groups. Once an algebra has been identified from Table I, one can obtain its underlying group structure from Table II.

We note that in previous treatments, the abelian algebra $\mathbf{\Omega}$ was not utilized in the same capacity as the complex field $\mathbb{C}$. Instead, the vector space isomorphism $\mathbf{\Omega} = \mathbb{R} \oplus \mathbb{R}$ was employed to give the $\mathbf{\Omega}$ algebras as $N_k \oplus N_k$ instead of $N_k \otimes \mathbf{\Omega} = N_k \oplus \omega N_k$.[9] While the identification $N_k \oplus N_k$ is correct in the matrix representation space,[3,4] it should not be used in the present discussion, as it conceals the vee-group structure of the algebra $\mathbf{\Omega}$ (see discussion in Ref. 9).

The connection to other work[3,4] is made by noting that Theorem 5 implies the following recursion relations obtained by iterating Eq. (8), and using Eq. (6).

$$N_{2k-1} \otimes N_2 \approx N_{2k} \otimes N_1 \approx N_{2k+2}, \quad (k \geqslant 1) \qquad (11a)$$

$$N_{2k-1} \otimes N_1 \approx N_{2k} \otimes N_2 \approx N_{2k+1}. \qquad (11b)$$

Specific cases of these relations were derived in Ref. 5 (Table IV), by direct calculation.

With the identification $A^{2,0} \approx N_1$ and $A^{0,2} \approx N_2 = H$ (Table I), and the application of (11a) and (11b) and (9a), (9b), and (9c) to the edges of Table I, we can verify the following relations for compact spaces.

$$A^{0,n} \otimes A^{2,0} \approx A^{n+2,0}, \qquad (12a)$$

$$A^{n,0} \otimes A^{0,2} \approx A^{0,n+2}. \qquad (12b)$$

These relations were used as a starting point for the construction of Clifford algebras in Ref. 3.

V. SPINORS AND SPIN ALGEBRAS

In this section, we show how spinors in n dimensions are related to some of the Clifford algebras. The main result is the demonstration that there exist Clifford algebras which cannot be directly related to the spinor formalism; among them is the real Clifford algebra in Minkowski spacetime N_4.

A second result is that even in those cases where Clifford algebras are related to spinors, we can realize a related Lie algebra that is in fact larger than that obtained by the usual spinor methods. This has the practical consequence that we can derive isomorphisms between Lie algebras useful in physics in a very simple manner (Table IV).

We have explicitly shown the relationship between the elementary (real two-component) spinors ψ and the Clifford algebra N_1 in Ref. 13. This relationship can be written as

$$\psi \leftrightarrow N_1. \qquad (13)$$

Spinors are objects in matrix representation space. The demonstration of relation (13) in Ref. 13. was given in the representation space of N_1, which is $\mathbb{R}(2)$ (2×2 real matrices). We can, however, utilize the group theoretical results of this paper in order to discuss spinors in n dimensions without having to enter the representation space of the Clifford algebras. (This will be done in turn in the following section after we have given the representation space of the Clifford algebras.)

Spinors in n dimensions are constructed from Kronecker products of the elementary spinors ψ as follows[2,6–8]:

$$\psi^1 = \psi; \quad \psi^n = (\psi)^n = \psi \otimes \cdots \otimes \psi, \ n \text{ times}. \qquad (14)$$

The complex spinors are obtained by complexifying the spinor field. In order to distinguish between real and complex spinors in n dimensions, we will occasionally specify the field, as $\psi^n(\mathbb{R})$ or $\psi^n(\mathbb{C})$.

Definition 3: $\psi^n(\mathbb{C}) = \psi^n(\mathbb{R}) \otimes \mathbb{C}$. (15)

The relationship of the Clifford algebras to spinors in n dimensions is given by the following theorem.

Theorem 6:

(a) The real spinors in $2k$ dimensions are related to the odd N algebras as

$$\psi^k(\mathbb{R}) \leftrightarrow N_{2k-1}. \qquad (16a)$$

(b) The complex spinors in $2k$ dimensions are related to the S algebras as

$$\psi^k(\mathbb{C}) \leftrightarrow S_k. \qquad (16b)$$

The proof of Theorem 6 is obtained from Theorem 5 and relationships (13)–(15). A few comments are in order.

Most important is the fact that the algebras N_{even} and Ω *cannot* be related to the spinor formalism. This includes the Clifford algebra in Minkowski spacetime, $A^{1,3} \approx N_4$.

The dimension of the spinor space corresponds to the dimension of the corresponding Riemann space in the real case (Table II); in the complex case the spinors are considered over the field $\mathbb{C}$, hence the dimension is one less than the corresponding Riemann space.

Our result differs in a significant way from other work in that we were able to maintain a clear distinction between the real and complex algebras and spinors. This is not always possible using the traditional spinor methods. Consider for example the construction of the Dirac spinors $\psi^2(\mathbb{C})$ (related to the Dirac algebra $D = S_2$) as the Kronecker product of two copies of the Pauli spinors $\psi^1(\mathbb{C})$ (related to the Pauli algebra S_1 by Theorem 6). From Eq. (14) the relation is

$$\psi^1(\mathbb{C}) \otimes \psi^1(\mathbb{C}) = \psi^2(\mathbb{C}). \qquad (17)$$

It is clear that the analogous construction for *real* spinors gives the Majorana spinors $\psi^2(\mathbb{R})$ (which are related to the Majorana algebra N_3 by Theorem 6). The relation for the real case is

$$\psi^1(\mathbb{R}) \otimes \psi^1(\mathbb{R}) = \psi^2(\mathbb{R}). \qquad (18)$$

It is therefore impossible to obtain spinors corresponding to the real Clifford algebra in Minkowski spacetime N_4; we can only obtain their complexification, which by Eq. (9a) and (15) are the Dirac spinors. Furthermore, the algebras N_3 and N_4 do *not* have the same properties; for instance, we show below that the corresponding Lie algebras are distinct (Table III) (see also the discussion in Ref. 14).

We can now discuss the relation between spinors in n dimensions and Clifford algebras by constructing the associated Lie algebras. From the spinors, we construct the "spin" Lie algebras spin (n), which are locally isomorphic to SO(n). Globally, spin (n) provides the double covering of SO(n). The relationship to the Clifford algebras is given by the following theorem.

Theorem 7:

(a) The two-forms of the Clifford algebra $A^{p,q}$ corresponding to the Riemann space $M^{p,q}$ provide a representation of the Lie algebra spin $(p,q) \approx \mathrm{SO}(p,q)$, via the Lie bracket.

(b) The (inner) automorphism group of the tensor fields in the Riemannian space $M^{p,q}$ is the Lie group corresponding to the Lie algebra $\mathrm{SO}(p,q)$. This is known as the "Clifford group of automorphisms."

Parts (a) and (b) are separately well known.[17,18]

We have previously drawn attention to the fact that one can include all the differential form basis of a Clifford algebra $A^{p,q}$ in realizing a Lie algebra via the Lie bracket, and not just the two-forms. This was done in Ref. 5, where we determined the Lie algebras corresponding to each Clifford algebra. It is obvious that the Lie algebra obtained by including all the basis forms must necessarily be larger than the Lie algebra given by Theorem 7, and must include the Lie algebra of Theorem 7 as a subalgebra.

These have been directly constructed in Ref. 5 by computing the Killing–Cartan form of the enveloped rotation group. The results for the first few algebras, which are of interest in physics, are listed in Table III. For comparison, we have also listed the Lie algebras given by Theorem 7.

Because the Lie algebras corresponding to the Clifford algebras are topological, they are better described in terms of the matrix representation space. This has been done in the following section [Theorem 9, Eq. (23)].

We remark moreover on why the full rotation algebras of Table III have not been singled out previously. The reason for this is that in the usual spinor construction, one is naturally led to the group of automorphisms, hence to Theorem 7. Second, the matrix representations of the automorphism algebras in fact exhaust the representation space and one cannot represent the larger full algebras in the same space. We have been able to construct the full algebras in Table III only by utilizing the differential forms basis (see Ref. 5).

VI. MATRIX REPRESENTATIONS AND LIE ALGEBRAS

In this section we give the matrix representation space of each Clifford algebra. This makes possible an alternative discussion on the relationship between spinors in n dimensions and Clifford algebras. We next determine the Lie algebra associated with each Clifford algebra, and derive some useful isomorphisms. In particular, we show how the results of this paper can be utilized to give the real forms of certain Lie algebras.

TABLE III. Rotation algebras represented by Clifford algebras.

Riemann space	Lie algebra from full Clifford algebra	Lie algebra from two-forms only
$M^{2,0}$	SO(2,1)	SO(2)
$M^{1,1}$	SO(1,2)	SO(1,1)
$M^{0,2}$	SO(0,3)	SO(0,2)
$M^{3,0}$	SO(3,1)	SO(3)
$M^{2,1}$	SO(2,2)	SO(2,1)
$M^{1,2}$	SO(1,3)	SO(1,2)
$M^{0,3}$	SO(0,4)	SO(0,3)
$M^{4,0}$	SO(5,1)	SO(4)
$M^{3,1}$	SO(3,3)	SO(3,1)
$M^{2,2}$	SO(3,3)	SO(2,2)
$M^{1,3}$	SO(1,5)	SO(1,3)
$M^{0,4}$	SO(1,5)	SO(0,4)

The matrix representation space of the Clifford algebras has been determined in the classic paper of Atiyah, Bott, and Shapiro,[3] which we recall here. Let $\mathbb{F}(k)$ [also denoted as $M_k(\mathbb{F})$], be the $k \times k$ matrix with entries from the field $\mathbb{F}$. $\mathbb{F}$ is either $\mathbb{R}$, $\mathbb{C}$, $\mathbf{\Omega}$, or H. Since $\mathbf{\Omega}$ is not a conventional field, we employ the vector space isomorphism $\mathbf{\Omega} \approx \mathbb{R} \oplus \mathbb{R}$ (see Ref. 9). The matrix representation space of the Clifford algebras is obtained by combining the results of Sec. IV with those of Ref. 3, and is given as follows.

Theorem 8: The matrix representation space of each Clifford algebra is

$$N_{2k-1} \approx \mathbb{R}(2k), \tag{19a}$$

$$N_{2k} \approx H(k), \tag{19b}$$

$$S_k \approx \mathbb{C}(2k), \tag{19c}$$

$$\Omega_{2k-1} \approx \mathbf{\Omega}(2k) \approx \mathbb{R}(2k) \oplus \mathbb{R}(2k), \tag{19d}$$

$$\Omega_{2k} \approx (\mathbf{\Omega} \otimes H)(k) \approx H(k) \oplus H(k). \tag{19e}$$

These are the irreducible representations. For calculational purposes, we can obtain reducible representations in terms of the familiar real and complex matrix algebras $\mathbb{R}(n)$ and $\mathbb{C}(n)$, for the cases $H(k)$, $\mathbf{\Omega}(k)$, and $\mathbf{\Omega} \otimes H(k)$, by recalling the well-known inclusion relations

$$H(k) \subset \mathbb{C}(2k) \tag{20a}$$

and

$$\mathbf{\Omega}(k) \approx \mathbb{R}(k) \oplus \mathbb{R}(k) \subset \mathbb{R}(2k). \tag{20b}$$

Hence, for cases (19b), (19d), and (19e) we have from identities (20a), and (20b) reducible representations for the following:

$$N_{2k} \subset \mathbb{C}(2k), \tag{21a}$$

$$\Omega_{2k-1} \subset \mathbb{R}(4k), \tag{21b}$$

$$\Omega_{2k} \subset \mathbb{C}(4k). \tag{21c}$$

This analysis shows two things:

(i) The only Clifford algebras with irreducible representations over $\mathbb{R}$ and $\mathbb{C}$ are in fact those corresponding to spinors in n dimensions, and

(ii) if one constructs (as is traditionally done) matrix representations over $\mathbb{R}$ and $\mathbb{C}$, then one cannot distinguish between the following sets of algebras in representation space. (Recall that $\Omega_0 = \mathbf{\Omega}$.)

$$\{N_{2k-1}; \Omega_{k-1}\}, \quad k = 1,2,3,... \tag{22a}$$

$$\{S_k; N_{2k}; \Omega_k\}. \tag{22b}$$

TABLE IV. Isomorphisms among Lie algebras.

$\mathrm{SO}(2,1) \approx \mathrm{SL}(2;\mathbb{R})$
$\mathrm{SO}(3) \approx \mathrm{SL}(1;H)$
$\mathrm{SO}(3,1) \approx \mathrm{SL}(2;\mathbb{C})$
$\mathrm{SO}(2,2) \approx \mathrm{SL}(2;\mathbb{R}) \oplus \mathrm{SL}(2;\mathbb{R})$
$\mathrm{SO}(4) \approx \mathrm{SL}(1;H) \oplus \mathrm{SL}(1;H)$
$\mathrm{SO}(5,1) \approx \mathrm{SL}(2;H)$
$\mathrm{SO}(3,3) \approx \mathrm{SL}(4;\mathbb{R})$

Since spinors are vectors in representation space, it is impossible to construct spinors corresponding to the cases (19b), (19d), and (19e). This provides a demonstration of Theorem 6, in representation space.

From Theorem 8, we can obtain the Lie algebra corresponding to each Clifford algebra. We use the construction of the Lie algebra SL(k,$\mathbb{F}$) defined via the Lie bracket on the full matrix algebra $\mathbb{F}(k)$ [or $M_k(\mathbb{F})$]. The restriction to matrices of unit determinant results in no loss of generality. The Lie algebras corresponding to the Clifford algebras are given by the following theorem.

Theorem 9: The Lie algebra corresponding to every Clifford algebra of Table I is given as

$$\begin{aligned} &N_{2k-1} \sim \mathrm{SL}(2k;\mathbb{R}),\\ &N_{2k} \sim \mathrm{SL}(k;H),\\ &S_k \sim \mathrm{SL}(2k;\mathbb{C}),\\ &\Omega_{2k-1} \sim \mathrm{SL}(2k;\mathbf{\Omega}) \approx \mathrm{SL}(2k;\mathbb{R}) \oplus \mathrm{SL}(2k;\mathbb{R}),\\ &\Omega_{2k} \sim \mathrm{SL}(k;\mathbf{\Omega} \otimes H) \approx \mathrm{SL}(k;H) \oplus \mathrm{SL}(k;H). \end{aligned} \tag{23}$$

It is possible to relate the correspondence between Clifford algebras and Lie algebras in representation space to the group-theoretical discussion of the previous section. By comparing the rotation algebras given separately in Table III with the Lie algebras given by Theorem 9, Eq. (23), we can give isomorphisms between the Lie algebras of small order. For each rotation algebra in Table III, identify the corresponding Clifford algebra in Table I, then find the Lie algebra from (23) to obtain the following well-known relations (Table IV). [Recall that $\mathrm{SO}(p,q) \approx \mathrm{SO}(q,p)$.]

This result is given as an example of the utility and application of the methods of this paper. We note that these relations (Table IV) cannot be obtained via the usual spinor methods.

The above connection between Clifford and Lie algebras makes possible yet another useful observation. All the structural relations for the Clifford algebras, Eq. (9) and (11), can be translated to give relations among the Lie algebras. For instance, Eq. (9a) with identification (23) gives

$$\mathrm{SL}(2k;\mathbb{C}) \approx \mathrm{SL}(k;H) \otimes \mathbb{C} \approx \mathrm{SL}(2k;\mathbb{R}) \otimes \mathbb{C}. \tag{24}$$

This is precisely the determination of the two real forms $\mathrm{SL}(2k;\mathbb{R})$ and $\mathrm{SL}(k;H)$ of the complex Lie algebra $\mathrm{SL}(2k;\mathbb{C})$, as is well known.[19,20]

The recursion relations (11) translate into

$$\mathrm{SL}(2k;\mathbb{R}) \otimes \mathrm{SL}(1;H) \approx \mathrm{SL}(k;H) \otimes \mathrm{SL}(2;\mathbb{R}) \approx \mathrm{SL}(k+1;H), \tag{25a}$$

$$\mathrm{SL}(2k;\mathbb{R}) \otimes \mathrm{SL}(2;\mathbb{R}) \approx \mathrm{SL}(k;H) \otimes \mathrm{SL}(1;H) \approx \mathrm{SL}(2k+2;\mathbb{R}). \tag{25b}$$

These recursion relations (25) can be used to obtain many useful structural identities for the Lie algebras.

This demonstrates the utility of the construction given in this paper, since it enables us to obtain nontrivial results in a very simple manner. Conversely, our discussion indicates that the classic work of Cartan on the classification of complex Lie algebras and their real forms in many ways anticipated later results on Clifford algebras.

VII. CONCLUSION

In this paper, we gave a simple classification of all the Clifford algebras in terms of their underlying group structure (Tables I and II). An interesting result is that by constructing all the Clifford algebras over the real field $\mathbb{R}$, we obtained the complex Clifford algebras in odd-dimensional Riemannian spaces.

We then compared the Clifford algebras to spinors in n dimensions, and showed that two classes of Clifford algebras cannot be related to spinors. Among these is the real Clifford algebra in Minkowski spacetime $A^{1,3} \approx N_4$, which is distinct from both the algebra of Dirac matrices, and the algebra of Majorana matrices. This discussion showed how the Clifford algebras provide a broader framework than the traditional spinor methods in the description of physical tensor fields.

A discussion of the Lie algebras related to the Clifford algebras resulted in a very simple derivation of isomorphisms between some of the Lie algebras useful in physics (Tables III and IV).

In conclusion, we believe that this paper has clarified the relationship between the algebraic frameworks on which much of physics is done. We have also provided a direct method for the utilization of the unfamiliar Clifford algebras in the construction of physical models.

ACKNOWLEDGMENTS

Thanks are due to Dr. K. C. Hannabuss for his remarks, which were the origin of Theorem 2 in this paper. I would also like to thank Dr. N. Marmaridis for many informative discussions.

[1]W. K. Clifford, "On the Classification of Geometric Algebras," paper XLIII, in Mathematical Papers of W. K. Clifford, edited by R. Tucker (MacMillan, London, 1882).
[2]P. K. Rashevskii, Am. Math. Soc. Translations (Ser. 2) **6**, 1 (1957).
[3]M. F. Atiyah, R. Bott, and A. Shapiro, Topology **3** (Suppl. 1), 3 (1964).
[4]M. Karoubi, *K-Theory* (Springer, Berlin, 1979); see also Ref. 18.
[5]N. Salingaros, J. Math. Phys. **22**, 226 (1981).
[6]E. Cartan, *The Theory of Spinors* (Hermann, Paris, 1966).
[7]R. Brauer and H. Weyl, Am. J. Math. **57**, 425 (1935).
[8]A. Pais, J. Math. Phys. **3**, 1135 (1962).
[9]N. Salingaros, "Algebras With Three Anticommuting Elements. II," J. Math. Phys. **22**, 2096 (1981).
[10]H. Flanders, *Differential Forms* (Academic, New York, 1963).
[11]L. Dornhoff, *Group Representation Theory, Part A* (Dekker, New York, 1971).
[12]F. G. Frobenius and I. Schur, Sitzungber. Preuss. Akad. Wiss. (Berlin), 186 (1906).
[13]Y. Ilamed and N. Salingaros, "Algebras With Three Anticommuting Elements. I ," J. Math. Phys. **22**, 2091 (1981).
[14]N. Salingaros, in "Proceedings of the third Workshop on Lie-Admissible Formations," University of Massachusetts at Boston, 1980, Hadronic J. **4**, 949 (1981).
[15]N. Salingaros and M. Dresden, Phys. Rev. Lett. **43**, 1 (1979).
[16]N. Salingaros, "Electromagnetism and the Holomorphic Properties of Spacetime," J. Math. Phys. **22**, 1919 (1981).
[17]H. Boerner, *Representations of Groups* (North-Holland, Amsterdam, 1967).
[18]I. R. Porteous, *Topological Geometry* (Van Nostrand Reinhold, London, 1969).
[19]J. Tits, *Tabellen zu den Einfachen Lie Gruppen und Ihren Darstellungen*, Springer Lecture Notes in Mathematics No. 40 (Springer, Berlin, 1967).
[20]I. G. MacDonald, in *Representation Theory of Lie Groups*, edited by M. F. Atiyah, London Math. Soc. Lecture Notes No. 34 (Cambridge U.P., Cambridge, 1979).

Modified fourth-order Casimir invariants and indices for simple Lie algebras

Susumu Okubo
Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627

(Received 16 June 1981; accepted for publication 21 August 1981)

The fourth-order indices for Lie algebras have been defined and studied by Patera, Sharp, and Winternitz. We show that it may be more convenient to modify the original definition and that the modified fourth-order indices are intimately related to eigenvalues of symmetrized fourth-order Casimir invariants. Explicit expressions for these quantities are given and we also find a quartic trace identity involving the generic element of these Lie algebras. We discuss the triality principle for the Lie algebra D_4 in connection with identical vanishing of the modified fourth-order index for this algebra.

PACS numbers: 02.20.Sv

1. INTRODUCTION AND SUMMARY OF MAIN RESULTS

Let L be a simple Lie algebra over the complex field or more generally over any algebraically closed field of characteristic zero. Let $\{\rho\}$ be a representation of L. The second-order index of Dynkin[1] is then defined by

$$l_2(\rho) = \sum_M (M,M), \tag{1.1}$$

where the summation is over all weights M of the representation $\{\rho\}$ and (M,M) is the standard symmetric bilinear form[2] in the root space of L. The Dynkin index has many nice properties. Let $\{\rho_A\}$ and $\{\rho_B\}$ be two irreducible representations of L and decompose the product $\{\rho_A\}\otimes\{\rho_B\}$ into a direct sum of N irreducible components as

$$\{\rho_A\}\otimes\{\rho_B\} = \sum_{j=1}^{N}\oplus\{\rho_j\}. \tag{1.2}$$

Then, $l_2(\rho)$ satisfies

$$d(\rho_A)l_2(\rho_B) + d(\rho_B)l_2(\rho_A) = \sum_{j=1}^{N} l_2(\rho_j). \tag{1.3}$$

Here, $d(\rho)$ designates the dimension of the irreducible representation $\{\rho\}$. Also, it has been noted in an earlier paper[3] which will hereafter be referred to as (I), that we have a quadratic-sum rule

$$4\frac{l_2(\rho_A)l_2(\rho_B)}{d(\rho_0)} + d(\rho_A)d(\rho_B)\left[\frac{l_2(\rho_A)}{d(\rho_A)} + \frac{l_2(\rho_B)}{d(\rho_B)}\right]^2 = \sum_{j=1}^{N}\frac{1}{d(\rho_j)}[l_2(\rho_j)]^2, \tag{1.4}$$

where $\{\rho_0\}$ hereafter designates the adjoint representation of L.

Patera, Sharp, and Winternitz[4] introduced notion of higher-order indices. In particular, the fourth-order index $l_4(\rho)$ is defined by

$$l_4(\rho) = \sum_M (M,M)^2. \tag{1.5}$$

Then for the decomposition Eq. (1.2), they showed the validity of

$$d(\rho_A)l_4(\rho_B) + d(\rho_B)l_4(\rho_A) + \frac{2(n+2)}{n}l_2(\rho_A)l_2(\rho_B) = \sum_{n=1}^{N} l_4(\rho_j), \tag{1.6}$$

where n is the rank of the simple Lie algebra L. Numerical values of $l_2(\rho)$ and $l_4(\rho)$ for many low-dimensional irreducible representations $\{\rho\}$ of any simple Lie algebras with rank less than eight have been tabulated by McKay and Patera.[5] Moreover, many interesting properties of these indices have recently been found by Patera and his coworkers.[6]

We note the following. Defining the modified fourth-order index $\bar{l}_4(\rho)$ by

$$\bar{l}_4(\rho) = l_4(\rho) - \frac{(n+2)d(\rho_0)}{n[d(\rho_0)+2]}\left[\frac{l_2(\rho)}{d(\rho)} - \frac{1}{6}\frac{l_2(\rho_0)}{d(\rho_0)}\right]l_2(\rho), \tag{1.7}$$

Eqs. (1.3), (1.4), and (1.6) imply the validity of

$$d(\rho_A)\bar{l}_4(\rho_B) + d(\rho_B)\bar{l}_4(\rho_A) = \sum_{j=1}^{N}\bar{l}_4(\rho_j), \tag{1.8}$$

which shares the same simpler structure as Eq. (1.3) for $l_2(\rho)$. Moreover, we can show that $\bar{l}_4(\rho)$ is identically zero, i.e.,

$$\bar{l}_4(\rho) = 0 \tag{1.9}$$

for any irreducible representation $\{\rho\}$ of all exceptional Lie algebras G_2, F_4, E_6, E_7, and E_8, as well as for A_1, A_2, B_2, and D_4. The validity of Eq. (1.9) for these algebras, except for B_2 and D_4, has been proved in I. As we shall see in Sec. 4, the validity for D_4 is connected with the so-called triality principle[7] in D_4. These facts suggest that $\bar{l}_4(\rho)$ rather than $l_4(\rho)$ may have the more basic properties. One purpose of this note is to show first that $\bar{l}_4(\rho)$ is indeed intimately connected with fourth-order Casimir invariants of L and that higher order indices may also be defined in terms of suitable higher order Casimir invariants. Moreover, we will generalize the quartic trace identity found in I to all classical Lie algebras.

For our purpose, it is convenient to classify all simple Lie algebras into the following three categories:

$$\begin{aligned}&\text{(i) } A_1, A_2, G_2, F_4, E_6, E_7, E_8;\\ &\text{(ii) } A_n (n\geqslant 3), B_n (n\geqslant 2), C_n (n\geqslant 2), D_n (n\geqslant 5);\\ &\text{(iii) } D_4.\end{aligned} \tag{1.10}$$

Next, let $t_1, t_2, \dots, t_{d_0}$ with $d_0 \equiv d(\rho_0)$ be an ordered basis of the simple (abstract) Lie algebra L with commutation relation

$$[t_\mu, t_\nu] \equiv C_{\mu\nu}^{\lambda} t_\lambda, \quad (\mu,\nu = 1,2,\dots,d_0), \tag{1.11}$$

where $C_{\mu\nu}^{\lambda}$ $(\mu,\nu = 1,2,\dots,d_0)$ are the structure constants of L with respect to this basis. Also, we adopt hereafter the usual

 0022-2488/82/010008-13$02.50

summation convention about repeated Greek indices, unless it is otherwise stated. Let x_μ ($\mu = 1,2,...,d_0$) be representation matrices of t_μ in an arbitrary but fixed irreducible representation $\{\lambda\}$ which we call reference representation and set

$$h_{\mu_1\mu_2\cdots\mu_p} = \frac{1}{p!}\sum_P \operatorname{Tr}(x_{\mu_1}x_{\mu_2}\cdots x_{\mu_p}), \tag{1.12}$$

where the summation is over $p!$ permutations P of p indices $\mu_1,\mu_2,\cdots\mu_p$. Clearly, $h_{\mu_1\mu_2\cdots\mu_p}$ is completely symmetric in p indices $\mu_1,\mu_2,\ldots\mu_p$. We have $h_\mu = 0$ for $p = 1$. For $p = 2$, it is well known[2] that $h_{\mu\nu} = \operatorname{tr}(x_\mu x_\nu)$ is independent of the choice of a particular representation $\{\lambda\}$, apart from a multiplicative constant. Hence, we set

$$h_{\mu\nu} = \operatorname{tr}(x_\mu x_\nu) = C_0(\lambda)g_{\mu\nu} \tag{1.13}$$

for a nonzero constant $C_0(\lambda)$, where $g_{\mu\nu}$ is the Killing form, with its inverse $g^{\mu\nu}$. In Sec. 3 we will choose a special normalization $C_0(\lambda) = 1$ with $\{\lambda\}$ being the basic (i.e., lowest-dimensional) representation of L. We can now raise and/or lower Greek indices as usual by means of $g^{\mu\nu}$ and $g_{\mu\nu}$ and set

$$I_p = h^{\mu_1\mu_2\cdots\mu_p}t_{\mu_1}t_{\mu_2}\cdots t_{\mu_p}, \tag{1.14}$$

where the product on the right side refers to that in the universal enveloping algebra[8] $U(L)$ of L. We can readily verify[3,9] that I_p are Casimir invariants of L. Let $\{\rho\}$ be hereafter a generic irreducible representation of L and let X_μ be representation matrices of t_μ in $\{\rho\}$. We write the common eigenvalue of I_p in $\{\rho\}$ as $I_p(\rho)$ so that

$$h^{\mu_1\mu_2\cdots\mu_p}X_{\mu_1}X_{\mu_2}\cdots X_{\mu_p} = I_p(\rho)E, \tag{1.15}$$

where E is the identity matrix in $\{\rho\}$. Since $h_{\mu_1\cdots\mu_p}$ depends upon the choice of the reference representation $\{\lambda\}$, we often write $I_p(\rho)$ as $I_p(\rho;\lambda)$ whenever we want to emphasize its dependence upon $\{\lambda\}$. Taking the trace of both sides of Eq. (1.15), we then find a reciprocity relation

$$I_p(\rho;\lambda)d(\rho) = I_p(\lambda;\rho)d(\lambda). \tag{1.16}$$

In view of Eq. (1.13), the second-order Casimir invariant I_2 is essentially unique, apart from a multiplicative normalization constant which depends upon $\{\lambda\}$. Similarly, the symmetrized third-order Casimir invariant I_3 is again unique[10,11] in this sense. Actually, it is known[10,11] that $I_3 = 0$ identically for all simple Lie algebras except for the algebra A_n ($n\geqslant 2$). However, the situation is different for the fourth-order Casimir invariant I_4. In fact, $I_4(\rho;\lambda)$ changes in general its form when we change $\{\lambda\}$. This is related to the fact that the square of I_2 is also a fourth-order Casimir invariant. Hence, if I_4 is a fourth-order Casimir invariant, then so is

$$I'_4 = I_4 + C(I_2)^2$$

for arbitrary constant C. In order to remove this ambiguity, we proceed as follows. We call the reference representation $\{\lambda\}$ exceptional, if we can find a constant C such that we have an identity

$$\begin{aligned} h_{\mu\nu\alpha\beta} &= C\{h_{\mu\nu}h_{\alpha\beta} + h_{\mu\alpha}h_{\nu\beta} + h_{\mu\beta}h_{\nu\alpha}\} \\ &= CC_0^2\{g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}\}. \end{aligned} \tag{1.17}$$

Otherwise, we call $\{\lambda\}$ nonexceptional. We know[3] that any irreducible representation $\{\lambda\}$ of the type (i) algebras of Eq. (1.10), e.g., A_1, A_2, G_2, F_4, E_6, E_7, and E_8 is always exceptional. This is related to the fact that any fourth-order Casimir invariant I_4 of these algebras is simply proportional to $(I_2)^2$. In other words, for type (i) algebras, we have no genuine fourth-order Casimir invariant[11] so that the problem does not arise. For the type (ii) as well as type (iii) Lie algebras, we define

$$g_{\mu\nu\alpha\beta} = [2 + d(\rho_0)]h_{\mu\nu\alpha\beta} - \frac{H(\lambda)}{I_2(\lambda)}\{g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}\}, \tag{1.18}$$

where we have set

$$H(\rho) = [I_2(\rho)]^2 - \tfrac{1}{6}I_2(\rho_0)I_2(\rho) \tag{1.19}$$

for a later purpose. It is simple to see the orthogonality relation

$$g^{\mu\nu}g^{\alpha\beta}g_{\mu\nu\alpha\beta} = h^{\mu\nu}h^{\alpha\beta}g_{\mu\nu\alpha\beta} = 0. \tag{1.20}$$

Moreover, $g_{\mu\nu\alpha\beta}$ is identically zero for all type (i) Lie algebras. At any rate, we now define the modified fourth-order Casimir invariant J_4 by

$$J_4 = g^{\mu\nu\alpha\beta}t_\mu t_\nu t_\alpha t_\beta \tag{1.21a}$$

so that its eigenvalue $J_4(\rho)$ is given by

$$J_4(\rho) = [2 + d(\rho_0)]I_4(\rho) - 3\frac{d(\rho_0)}{d(\lambda)[I_2(\lambda)]^2}H(\lambda)H(\rho). \tag{1.21b}$$

When we want to emphasize the dependence of $J_4(\rho)$ upon the reference representation $\{\lambda\}$, we write it as $J_4(\rho,\lambda)$. Then, the reciprocity relation (1.16) leads to

$$d(\rho)J_4(\rho;\lambda) = d(\lambda)J_4(\lambda;\rho) \tag{1.22}$$

when we note Eq. (1.13). For all type (i) algebras, we have identically

$$J_4(\rho) = 0. \tag{1.23}$$

For all types (ii) and (iii) Lie algebras, we choose $\{\lambda\}$ to be nonexceptional. Then J_4 is not identically zero. Moreover, for the type (ii) cases, J_4 is independent of the particular choice of nonexceptional representation $\{\lambda\}$, except for an overall normalization constant, just as I_2 and I_3 are. For the type (iii) algebra, i.e., D_4, the situation is more involved. Actually, D_4 possesses one more fourth-order Casimir invariant $\hat{I}_4$, in addition to J_4 and $(I_2)^2$. Discussion of this case will be given in Sec. 4.

We now define modified Dynkin indices $D^{(p)}(\rho)$ ($p = 2,3,4$) by

$$\begin{aligned} D^{(2)}(\rho) &= d(\rho)I_2(\rho), \\ D^{(3)}(\rho) &= d(\rho)I_3(\rho), \\ D^{(4)}(\rho) &= d(\rho)J_4(\rho). \end{aligned} \tag{1.24}$$

Then, for the product decomposition Eq. (1.2), we will prove in the next section the validity of

$$d(\rho_A)D^{(p)}(\rho_B) + d(\rho_B)D^{(p)}(\rho_A) = \sum_{j=1}^{N} D^{(p)}(\rho_j) \tag{1.25}$$

for $p = 2,3,4$. We may note that $D^{(2)}(\rho)$ is essentially equivalent to $l_2(\rho)$ since

$$D^{(2)}(\rho) = (D^{(2)}(\lambda)/l_2(\lambda))l_2(\rho) = \text{const}\times l_2(\rho), \tag{1.26}$$

while $D^{(3)}(\rho)$ is nothing but the anomaly coefficient[10,12] of the triangular anomaly in grand unified gauge theory. Since Eq. (1.25) for $p=4$ has the same form as Eq. (1.8) for $\bar{I}_4(\rho)$, we may guess that they must be related to each other. We shall prove in the next section that indeed we have

$$\bar{I}_4(\rho) = C D^{(4)}(\rho) \tag{1.27}$$

for a constant C, which may depend upon the reference representation $\{\lambda\}$ but *not* on $\{\rho\}$. However, we can not necessarily express $D^{(4)}(\rho)$ in terms of $\bar{I}_4(\rho)$ since the constant C could be zero for some cases. Indeed, this happens for the case of the Lie algebras D_4 and B_2, so that we have $\bar{I}_4(\rho)=0$ also for this case, though $D^{(4)}(\rho)\neq 0$ in general. For this reason, we made the distinction between $\bar{I}_4(\rho)$ and $D^{(4)}(\rho)$. Also, for the Lie algebra D_4, we can define additional fourth-order index $\hat{D}^{(4)}(\rho)$ by

$$\hat{D}^{(4)}(\rho) = d(\rho)\hat{I}_4(\rho), \tag{1.28}$$

which also satisfies Eq. (1.25). C in Eq. (1.27) will be computed in the Appendix.

Let t and X be generic elements of L and its representation in the generic irreducible representation $\{\rho\}$, respectively. Expressing t as

$$t = \xi^\mu t_\mu \tag{1.29}$$

for some complex numbers ξ^μ $(\mu = 1,2,\ldots,d_0)$, we have, of course,

$$X = \xi^\mu X_\mu. \tag{1.30}$$

It is well known that we have

$$\mathrm{Tr}(X_\mu X_\nu) = \frac{d(\rho)I_2(\rho)}{d(\lambda)I_2(\lambda)} h_{\mu\nu} = \frac{D^{(2)}(\rho)}{D^{(2)}(\lambda)} h_{\mu\nu}. \tag{1.31}$$

Therefore, we find

$$\mathrm{Tr}X^2 = C_2(t)D^{(2)}(\rho) \tag{1.32}$$

when we set

$$C_2(t) = \frac{1}{D^{(2)}(\lambda)} h_{\mu\nu}\xi^\mu\xi^\nu.$$

Note that $C_2(t)$ depends upon $\{\lambda\}$ but *not* on $\{\rho\}$. Similarly, for the Lie algebra $A_n\,(n\geqslant 2)$, we find

$$\mathrm{Tr}X^3 = C_3(t)D^{(3)}(\rho). \tag{1.33}$$

If $\{\rho_1\}$ and $\{\rho_2\}$ are any two irreducible representation of L, then these relations imply

$$\mathrm{Tr}^{(2)}X^p/\mathrm{Tr}^{(1)}X^p = D^{(p)}(\rho_2)/D^{(p)}(\rho_1) \tag{1.34}$$

for $p=2$ and 3, where $\mathrm{Tr}^{(j)}$ $(j=1,2)$ refers to the trace with respect to the space $\{\rho_j\}$. This relation has been utilized recently[13] to prove uniqueness of grand-unified groups SU(5) and SO(10). Similarly, we can prove

$$\mathrm{Tr}X^4 - K(\rho)(\mathrm{Tr}X^2)^2 = C_4(t)D^{(4)}(\rho) \tag{1.35}$$

when we set

$$K(\rho) = \frac{d(\rho_0)}{2[2+d(\rho_0)]d(\rho)}\left\{6 - \frac{I_2(\rho_0)}{I_2(\rho)}\right\}. \tag{1.36}$$

The relation (1.35) is valid for all simple Lie algebras, except for D_4, as well as some special class of irreducible representations $\{\rho\}$ for D_4, which will be specified in Sec. 4. For all types (i) algebras, we have

$$\mathrm{Tr}X^4 = K(\rho)(\mathrm{Tr}X^2)^2 \tag{1.37}$$

for these algebras, reproducing the result of (I). Also, some class of irreducible representation $\{\rho\}$ for D_4 satisfy Eq. (1.37). For other cases, Eq. (1.35) implies the validity of

$$\frac{\mathrm{Tr}^{(2)}X^4 - K(\rho_2)[\mathrm{Tr}^{(2)}X^2]^2}{\mathrm{Tr}^{(1)}X^4 - K(\rho_1)[\mathrm{Tr}^{(1)}X^2]^2} = \frac{D^{(4)}(\rho_2)}{D^{(4)}(\rho_1)} \tag{1.38}$$

for two irreducible representations $\{\rho_1\}$ and $\{\rho_2\}$ of type (ii) Lie algebras. Some applications of Eq. (1.38) will be given elsewhere.

In this work, we adopt the lexicographical ordering of simple root systems as in Ref. 4 as well as in Ref. 14. Writing the corresponding fundamental weight system as $\{\Lambda_1,\Lambda_2,\ldots,\Lambda_n\}$, the highest weight Λ of any irreducible representation is given by

$$\Lambda = m_1\Lambda_1 + m_2\Lambda_2 + \cdots + m_n\Lambda_n, \tag{1.39}$$

in terms of nonnegative integers m_j $(j=1,2,\ldots,n)$. We often use the notation $\{\Lambda\}$ for $\{\rho\}$ whenever its highest weight Λ is known.

2. DERIVATION OF MAIN IDENTITY

First we will prove the validity of Eq. (1.25). Let $X_\mu^{(A)}$ and $X_\mu^{(B)}$ be representation matrices of t_μ in irreducible representations $\{\rho_A\}$ and $\{\rho_B\}$, respectively, and set

$$X_\mu = X_\mu^{(A)}\otimes E_B + E_A\otimes X_\mu^{(B)}, \tag{2.1}$$

which defines the representation matrix of t_μ in the product space $\{\rho_A\}\otimes\{\rho_B\}$. Here, E_A and E_B are unit matrices in the respective spaces. Computing both sides of

$$h^{\mu\nu}\mathrm{Tr}(X_\mu X_\nu) = I_2(\rho_A\otimes\rho_B),$$

of course this immediately gives the well-known result

$$d(\rho_A)d(\rho_B)\{I_2(\rho_A)+I_2(\rho_B)\} = \sum_{j=1}^{N} d(\rho_j)I_2(\rho_j), \tag{2.2}$$

since $\mathrm{Tr}(X_\mu)=0$. Similarly, calculating both sides of

$$h^{\mu\nu\alpha\beta}\mathrm{Tr}(X_\mu X_\nu X_\alpha X_\beta) = I_4(\rho_A\otimes\rho_B),$$

we find

$$d(\rho_A)d(\rho_B)\left\{I_4(\rho_A)+I_4(\rho_B)+\frac{6}{d(\lambda)[I_2(\lambda)]^2}H(\lambda)I_2(\rho_A)I_2(\rho_B)\right\} = \sum_{j=1}^{N} d(\rho_j)I_4(\rho_j), \tag{2.3}$$

where $H(\lambda)$ is defined by Eq. (1.19). In deriving Eq. (2.3) we used

$$h_{\mu\nu} = C_0(\lambda)g_{\mu\nu},\ h^{\mu\nu} = C_0(\lambda)g^{\mu\nu}, \tag{2.4a}$$

$$C_0^2(\lambda) = \frac{d(\lambda)}{d(\rho_0)} I_2(\lambda), \tag{2.4b}$$

as well as a relation

$$h^{\mu\nu\alpha\beta}h_{\mu\nu}h_{\alpha\beta} = d(\lambda)H(\lambda), \tag{2.5}$$

which will be proved shortly. Also, we have[3] a quadratic relation

$$d(\rho_A)d(\rho_B)\left\{4\frac{I_2(\rho_A)I_2(\rho_B)}{d(\rho_0)} + [I_2(\rho_A)+I_2(\rho_B)]^2\right\} = \sum_{j=1}^{N} d(\rho_j)[I_2(\rho_j)]^2, \tag{2.6}$$

which is equivalent to Eq. (1.4). Therefore, when we set

$$J_4(\rho) = [2+d(\rho_0)]I_4(\rho) - 3\frac{d(\rho_0)}{d(\lambda)[I_2(\lambda)]^2}H(\lambda)H(\rho), \tag{2.7}$$

Eqs. (2.2), (2.3), and (2.6) lead to

$$d(\rho_A)d(\rho_B)\{J_4(\rho_A)+J_4(\rho_B)\} = \sum_{j=1}^{N} d(\rho_j)J_4(\rho_j), \tag{2.8}$$

which is equivalent to Eq. (1.25) for $p=4$. The case for $p=3$ can be proved similarly. The reciprocity relation Eq. (1.22) can be proved from Eq. (1.16) and Eq. (2.4).

Hereafter in this note, $\{\rho\}$ always refers to the generic irreducible representation of L with X_μ being the matrix representations of t_μ in $\{\rho\}$. For simplicity, let us set

$$H_{\mu\nu\alpha\beta}(\rho) = \frac{1}{4!}\sum_P \operatorname{Tr}(X_\mu X_\nu X_\alpha X_\beta), \tag{2.9}$$

which is completely symmetric in indices μ,ν,α, and β. If we set $\{\rho\} = \{\lambda\}$, this gives

$$h_{\mu\nu\alpha\beta} = H_{\mu\nu\alpha\beta}(\lambda). \tag{2.10}$$

When we note a trivial identity

$$\operatorname{Tr}([X_\lambda, X_\mu X_\nu X_\alpha X_\beta]) = 0,$$

we find

$$C^{\tau}_{\lambda\mu}H_{\tau\nu\alpha\beta}(\rho) + C^{\tau}_{\lambda\nu}H_{\mu\tau\alpha\beta}(\rho) + C^{\tau}_{\lambda\alpha}H_{\mu\nu\tau\beta}(\rho) + C^{\tau}_{\lambda\beta}H_{\mu\nu\alpha\tau}(\rho) = 0. \tag{2.11}$$

Let V denote a vector space spaned by all completely symmetric quartic forms $K_{\mu\nu\alpha\beta}$ satisfying the condition

$$C^{\tau}_{\lambda\mu}K_{\tau\nu\alpha\beta} + C^{\tau}_{\lambda\nu}K_{\mu\tau\alpha\beta} + C^{\tau}_{\lambda\alpha}K_{\mu\nu\tau\beta} + C^{\tau}_{\lambda\beta}K_{\mu\nu\alpha\tau} = 0. \tag{2.12}$$

Then the validity of Eq. (2.11) implies

$$H_{\mu\nu\alpha\beta}(\rho)\in V.$$

Similarly, when we set

$$K^{(0)}_{\mu\nu\alpha\beta} = \tfrac{1}{3}[g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}], \tag{2.13}$$

we can easily prove

$$K^{(0)}_{\mu\nu\alpha\beta}\in V \tag{2.14}$$

when we note that

$$C^{\alpha}_{\lambda\mu}g_{\alpha\nu} \equiv f_{\lambda\mu\nu} \tag{2.15}$$

is completely antisymmetric in the three indices λ,μ,ν. In (I), we proved that

$$\operatorname{Dim}V = \begin{cases} 1, \text{ for type (i) algebras} \\ 2, \text{ for type (ii) algebras} \quad , \\ 3, \text{ for type (iii) algebra } D_4 \end{cases} \tag{2.16}$$

where Dim means the dimension of V. As we noted in (I), the validity of Eq. (2.16) is intimately related to the fact that types (i)–(iii) Lie algebras have[11] precisely 0, 1, and 2 fundamental fourth-order Casimir invariants, respectively, in the universal enveloping algebra $U(L)$.

Next, let us set

$$G_{\mu\nu\alpha\beta}(\rho) = \frac{1}{4!}\sum_P \operatorname{Tr}(X_\mu X_\nu X_\alpha X_\beta) - \frac{1}{3}K(\rho)\{(\operatorname{Tr}X_\mu X_\nu)(\operatorname{Tr}X_\alpha X_\beta) + (\operatorname{Tr}X_\mu X_\alpha)(\operatorname{Tr}X_\nu X_\beta) + (\operatorname{Tr}X_\mu X_\beta)(\operatorname{Tr}X_\nu X_\alpha)\}, \tag{2.17}$$

where $K(\rho)$ is given by Eq. (1.36). We may rewrite Eq. (2.17) also as

$$G_{\mu\nu\alpha\beta}(\rho) = H_{\mu\nu\alpha\beta}(\rho) - \frac{1}{2+d(\rho_0)}\frac{d(\rho)H(\rho)}{d(\lambda)I_2(\lambda)} \times\{g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha}\}. \tag{2.18}$$

It may be stressed that $G_{\mu\nu\alpha\beta}(\rho)$ is really independent of the choice of the reference representation $\{\lambda\}$ in view of Eq. (2.4). We can readily verify

$$h^{\mu\nu}h^{\alpha\beta}H_{\mu\nu\alpha\beta}(\rho) = d(\rho)H(\rho) \tag{2.19}$$

on the basis of Eqs. (2.4), (2.15), and (1.31). This also gives the orthogonality condition

$$g^{\mu\nu}g^{\alpha\beta}G_{\mu\nu\alpha\beta}(\rho) = K^{(0)\mu\nu\alpha\beta}G_{\mu\nu\alpha\beta}(\rho) = 0. \tag{2.20}$$

Then Eqs. (2.5) and (1.20) are special cases of Eqs. (2.19) and (2.20), respectively, for $\{\rho\} = \{\lambda\}$. The modified fourth-order Casimir invariant $J_4(\rho)$ is evidently given by

$$D^{(4)}(\rho) = d(\rho)J_4(\rho) = [2+d(\rho_0)]h^{\mu\nu\alpha\beta}G_{\mu\nu\alpha\beta}(\rho) = g^{\mu\nu\alpha\beta}H_{\mu\nu\alpha\beta}(\rho). \tag{2.21}$$

From this, we can derive Eq. (1.22).

Let us now first consider the type (i) Lie algebras of Eq. (1.10). For this case, we have Dim $V=1$. Since we have

$$K^{(0)\mu\nu\alpha\beta}K^{(0)}_{\mu\nu\alpha\beta} = \tfrac{1}{3}d(\rho_0)[d(\rho_0)+2]\neq 0,$$

it follows that $K^{(0)}_{\mu\nu\alpha\beta}\neq 0$. Therefore, any $K_{\mu\nu\alpha\beta}$ and especially $G_{\mu\nu\alpha\beta}(\rho)$ must be proportional to $K^{(0)}_{\mu\nu\alpha\beta}$. However, the orthogonality condition Eq. (2.20) then requires

$$G_{\mu\nu\alpha\beta}(\rho) = 0 \tag{2.22}$$

for all type (i) algebras, so that $J_4(\rho)=0$. For type (ii) cases, we have Dim $V=2$ and hence $G_{\mu\nu\alpha\beta}(\rho)$ must be a linear combination of $g_{\mu\nu\alpha\beta}$ and $K^{(0)}_{\mu\nu\alpha\beta}$, when we note that both are linearly independent of each other for a choice of $\{\lambda\}$ being nonexceptional. Again, the orthogonality condition

Eq. (2.20) requires then that we have

$$G_{\mu\nu\alpha\beta}(\rho) = \frac{B(\rho)}{2+d(\rho_0)} g_{\mu\nu\alpha\beta} \tag{2.23}$$

for all type (ii) Lie algebras, where $B(\rho)$ is a constant. We may regard all type (i) Lie algebras also as special cases of Eq. (2.23), with $B(\rho)=0$. Multiplying both sides of Eq. (2.23) by $h^{\mu\nu\alpha\beta}$ and noting Eq. (2.21), we get

$$d(\rho)J_4(\rho) = B(\rho)d(\lambda)J_4(\lambda).$$

If we have $J_4(\lambda)=0$, then this requires $J_4(\rho)=0$ for all generic irreducible representations $\{\rho\}$. This is clearly not possible for type (ii) Lie algebras, as we will see from explicit computation of $J_4(\rho)$ in Sec. 3. Therefore, we conclude $J_4(\lambda)\neq 0$ for any nonexceptional reference representation $\{\lambda\}$ of type (ii) Lie algebras, and hence

$$B(\rho) = \frac{d(\rho)J_4(\rho)}{d(\lambda)J_4(\lambda)} = \frac{D^{(4)}(\rho)}{D^{(4)}(\lambda)} \tag{2.24}$$

for any type (ii) Lie algebras. Then we may rewrite Eq. (2.23) as

$$\frac{[2+d(\rho_0)]}{d(\rho)J_4(\rho)} G_{\mu\nu\alpha\beta}(\rho) = \frac{1}{d(\lambda)J_4(\lambda)} g_{\mu\nu\alpha\beta}. \tag{2.25}$$

In particular, we have $g_{\mu\nu\alpha\beta} = [2+d(\rho_0)]G_{\mu\nu\alpha\beta}(\lambda)$. As we already remarked just after Eq. (2.18), $G_{\mu\nu\alpha\beta}(\rho)$ does not depend upon $\{\lambda\}$. Therefore, Eq. (2.25) implies that $g_{\mu\nu\alpha\beta}$ is really independent of a particular choice of $\{\lambda\}$, apart from a multiplicative constant. This, together with Eq. (2.21), proves the uniqueness of the modified fourth-order Casimir invariant $J_4(\rho)$ for type (ii) Lie algebras, apart from an overall normalization constant, as long as we choose $\{\lambda\}$ to be nonexceptional.

Now multiplying both sides of Eq. (2.23) or Eq. (2.25) by $\xi^\mu\xi^\nu\xi^\alpha\xi^\beta$, we immediately obtain Eq. (1.35). In order to derive Eq. (1.27), let h_j $(j=1,2,\dots,n)$ and e_α be the standard Cartan–Weyl basis. Here, n is the rank of the Lie algebra. Let H_j $(j=1,2,\dots,n)$ be representation matrices of h_j in $\{\rho\}$. Then, $l_2(\rho)$ and $l_4(\rho)$ are evidently expressed as

$$l_2(\rho) = \sum_{j,k=1}^{n} g^{jk}\,\mathrm{Tr}(H_jH_k),$$
$$l_4(\rho) = \mathrm{Tr}\Big(\sum_{j,k=1}^{n} g^{jk}H_jH_k\Big)^2. \tag{2.26}$$

Then the modified fourth-order index $\bar{l}_4(\rho)$ is given by

$$\bar{l}_4(\rho) = \sum_{i,j,k,l=1}^{n} g^{ij}g^{kl}G_{ijkl}(\rho), \tag{2.27}$$

when we note

$$[H_i,H_j]=0,\quad \sum_{j=1}^{n} g^{ij}g_{jk} = \delta^i_k,$$
$$\mathrm{Tr}(H_iH_j) = \frac{l_2(\rho)}{n} g_{ij},\quad (i,j,k=1,2,\dots,n), \tag{2.28}$$

and when we set $X_\mu = H_i$, $X_\nu = H_j$, $X_\alpha = H_k$, and $X_\beta = H_l$ in Eq. (2.17), and note Eq. (1.36). Similarly, when we multiply both sides of Eq. (2.25) by $g^{ij}g^{kl}$, we find the desired formula

$$\bar{l}_4(\rho) = \frac{D^{(4)}(\rho)}{D^{(4)}(\lambda)} \bar{l}_4(\lambda) = \text{const}\times D^{(4)}(\rho). \tag{2.29}$$

This proves the validity of Eq. (1.27) for the type (ii) Lie algebras. Similarly, starting with Eq. (2.22), we find $\bar{l}_4(\rho)=0$ for type (i) Lie algebras. The multiplicative constant in Eq. (2.29) is computed in the Appendix.

The validity of Eq. (2.25) also implies that a necessary and sufficient condition for $\{\lambda\}$ to be exceptional for type (ii) Lie algebras is to have $J_4(\lambda)=0$. For example, $\{\lambda\}=\{\Lambda_2\}$ for C_4 is exceptional since we have $J_4(\Lambda_2)=0$ for C_4, as will be shown in the next section. If $\{\lambda\}$ is exceptional, we have also

$$\mathrm{Tr}x^4 = K(\lambda)(\mathrm{Tr}x^2)^2$$

for $x=\xi^\mu x_\mu$ of type (ii) Lie algebras.

Finally for the type (ii) Lie algebras, the uniqueness of $J_4(\rho;\lambda)$ except for its normalization, together with the reciprocity relation Eq. (1.22), implies the validity of

$$J_4(\rho;\lambda') = \frac{d(\lambda')J_4(\lambda';\lambda)}{d(\lambda)J_4(\lambda;\lambda)} J_4(\rho;\lambda) \tag{2.30}$$

for any two reference representations $\{\lambda\}$ and $\{\lambda'\}$. Therefore, once we know $J_4(\rho;\lambda)$ for a given nonexceptional $\{\lambda\}$, we can calculate $J_4(\rho;\lambda')$ for any $\{\lambda'\}$.

For the remaining case of the type (iii) Lie algebra D_4, we have $\mathrm{Dim}\,V=3$, so that the argument presented in this section must be modified accordingly. This case will be discussed in detail in Sec. 4.

3. EIGENVALUES OF FOURTH-ORDER CASIMIR INVARIANTS

The purpose of this section is to give explicit expressions for $J_4(\rho)$ in terms of n nonnegative integers m_j $(1\leqslant j\leqslant n)$ specifying the highest weight Λ of $\{\rho\}$ as in Eq. (1.39). Since we have $J_4(\rho)=0$ identically for all type (i) Lie algebras, it suffices to consider only classical Lie algebras A_n, B_n, C_n, and D_n.

Using the lexicographical ordering of simple roots as in Ref. 5 and/or Ref. 14, the basic, or defining, representation of all type (ii) Lie algebras is given by $\{\Lambda_1\}$, which is nonexceptional. Also, the adjoint representation $\{\rho_0\}$ is found to be nonexceptional for all type (ii) algebras, although the latter is exceptional for D_4. Therefore, for our purpose, we could use either $\{\Lambda_1\}$ or $\{\rho_0\}$ for the choice of the reference representation $\{\lambda\}$, since $J_4(\rho;\lambda)$ must be independent of $\{\lambda\}$ except for overall normalization as in Eq. (2.30). In this section, we use $\{\lambda\}=\{\Lambda_1\}$ to be definite, and normalize

$$g_{\mu\nu} = h_{\mu\nu} = \mathrm{tr}(x_\mu x_\nu); \tag{3.1}$$

hence

$$d(\Lambda_1)I_2(\Lambda_1) = d(\rho_0), \tag{3.2}$$

$$l_2(\rho) = n\,\frac{d(\rho)I_2(\rho)}{d(\rho_0)}. \tag{3.2'}$$

Moreover, all formulas given in this section are also applicable to Lie algebras A_1, A_2, and D_4, which are not of type (ii), as well as D_3 where $D_3 = A_3$.

Let us now set

$$b_{\mu_1\mu_2\cdots\mu_p} = \mathrm{Tr}(x_{\mu_1} x_{\mu_2} \cdots x_{\mu_p}),$$

$$I_p^{\mathrm{NS}}(\rho) E = b^{\mu_1\mu_2\cdots\mu_p} X_{\mu_1} X_{\mu_2} \cdots X_{\mu_p}.$$

Then $I_p^{\mathrm{NS}}(\rho)$ is the eigenvalue of nonsymmetrized pth order Casimir invariant I_p^{NS} in the irreducible representation $\{\rho\}$. The explicit expressions for $I_p^{\mathrm{NS}}(\rho)$ have been given by many authors,[15–22] from which we can compute $I_4(\rho)$ and $J_4(\rho)$. However, the calculation is very tedious and complicated, although it is straightforward. We report results of the calculation following the notation and method given in Ref. 20.

A. Algebra A_n $(n \geqslant 1)$

We embed the Lie algebra A_n into the Lie algebra of the unitary group U$(n+1)$ whose irreducible representation is labeled[23] by $n+1$ integers f_j $(1 \leqslant j \leqslant n+1)$, satisfying

$$f_1 \geqslant f_2 \geqslant f_3 \geqslant \cdots \geqslant f_{n+1}. \tag{3.3}$$

Then the nonnegative integers m_j $(1 \leqslant j \leqslant n)$ in Eq. (1.39) specifying the generic irreducible representation $\{\rho\}$ of the algebra A_n are related to the f_j's by

$$m_j = f_j - f_{j+1} \quad (1 \leqslant j \leqslant n). \tag{3.4}$$

For simplicity, we set hereafter

$$N = n+1, \quad N \geqslant 2, \tag{3.5}$$

so that we are dealing with the Lie algebra A_{N-1} of the SU(N) group. Following the notation of Ref. 10, we set

$$\sigma_j = f_j + \frac{1}{2}(N+1) - j - \frac{1}{N}\sum_{k=1}^{N} f_k,$$
$$(1 \leqslant j \leqslant N), \tag{3.6}$$

which satisfy conditions

$$\sigma_1 > \sigma_2 > \cdots > \sigma_N, \tag{3.7a}$$

$$\sum_{j=1}^{N} \sigma_j = 0. \tag{3.7b}$$

The σ_j are related to the m_j $(1 \leqslant j \leqslant N-1)$ by

$$m_j = \sigma_j - \sigma_{j+1} - 1 \quad (1 \leqslant j \leqslant N-1), \tag{3.8a}$$

$$\sigma_j = \frac{1}{N}\left\{ \sum_{k=j}^{N} (N-k)(m_k+1) - \sum_{k=1}^{j-1} k(m_k+1) \right\}. \tag{3.8b}$$

Now, eigenvalues of symmetrized Casimir invariants $I_2(\rho)$, $I_3(\rho)$, and $I_4(\rho)$ are calculated[24] as

$$I_2(\rho) = \sum_{j=1}^{N} \{(\sigma_j)^2 - [\sigma_j^{(0)}]^2\},$$

$$I_3(\rho) = \sum_{j=1}^{N} (\sigma_j)^3,$$

$$I_4(\rho) = \sum_{j=1}^{N} \{(\sigma_j)^4 - [\sigma_j^{(0)}]^4\} - \frac{2N^2-3}{6} I_2(\rho)$$
$$= \sum_{j=1}^{N} (\sigma_j)^4 - \frac{2N^2-3}{6} \sum_{j=1}^{N} (\sigma_j)^2$$
$$+ \frac{1}{720} N(N^2-1)(11N^2-9)), \tag{3.9}$$

where we have set

$$\sigma_j^{(0)} = \tfrac{1}{2}(N+1) - j, \tag{3.10}$$

so that

$$\sum_{j=1}^{N} [\sigma_j^{(0)}]^2 = \frac{1}{12} N(N^2-1),$$

$$\sum_{j=1}^{N} [\sigma_j^{(0)}]^4 = \frac{1}{240} N(N^2-1)(3N^2-7),$$

$$\sum_{j=1}^{N} \sigma_j^{(0)} = \sum_{j=1}^{N} [\sigma_j^{(0)}]^3 = 0. \tag{3.11}$$

Then $J_4(\rho)$ is calculated to be

$$J_4(\rho) = (N^2+1) \sum_{j=1}^{N} \{(\sigma_j)^4 - [\sigma_j^{(0)}]^4\}$$
$$- \frac{2N^2-3}{N} \left\{ \left[\sum_{j=1}^{N} (\sigma_j)^2 \right]^2 - \left[\sum_{j=1}^{N} (\sigma_j^{(0)})^2 \right]^2 \right\}$$
$$= (N^2+1) \sum_{j=1}^{N} (\sigma_j)^4 - \frac{2N^2-3}{N} \left[\sum_{j=1}^{N} (\sigma_j)^2 \right]^2$$
$$+ \frac{1}{720} N(N^2-1)(N^2-4)(N^2-9). \tag{3.12}$$

We note that $J_4(\rho)$ contains only quartic polynomials of σ_j's in addition to a constant. We can verify the fact that $J_4(\rho)$ is identically zero for $N=2$ and $N=3$, corresponding to A_1 and A_2, since we have an identity $\frac{1}{2}(a^2 + b^2 + c^2)^2 = a^4 + b^4 + c^4$ for any three numbers a, b, and c satisfying $a+b+c=0$.

We note that the fundamental representations $\{\Lambda_j\}$ $(1 \leqslant j \leqslant N-1)$ correspond to completely antisymmetric tensor representations, while $\{k\Lambda_1\}$ for $(k \geqslant 1)$ are completely symmetric representations in the sense of Young's tableau.[23] For possible applications to particle physics, which will be reported elsewhere, we will give below an explicit formula for eigenvalues of fourth-order Casimir invariants in these representations.

$$I_4(\Lambda_j) = \frac{N+1}{6N^3} j(N-j)\{N^2(N^2+6N+6)$$
$$- 6(N^2+3N+3) j(N-j)\} \quad (1 \leqslant j \leqslant N-1). \tag{3.13a}$$

$$J_4(\Lambda_j) = \frac{(N+1)(N+2)(N+3)}{6N} j(N-j)$$
$$\times \{N(N+1) - 6j(N-j)\} \quad (1 \leqslant j \leqslant N-1). \tag{3.13b}$$

$$J_4(k\Lambda_1) = \frac{(N-1)(N-2)(N-3)}{6N} k(N+k)$$
$$\times \{N(N-1) + 6k(N+k)\} \quad (k \geqslant 1). \tag{3.13c}$$

Also, we note that

$$I_2(\Lambda_j) = \frac{N+1}{N} j(N-j), \quad (1 \leqslant j \leqslant N-1),$$

$$I_3(\Lambda_j) = \frac{(N+1)(N+2)}{2N^2} j(N-j)(N-2j) \quad (1 \leqslant j \leqslant N-1),$$

$$I_2(\rho_0) = 2N,$$

$$J_4(\Lambda_1) = \frac{1}{6N}(N^2-1)(N^2-4)(N^2-9),$$

$$J_4(\rho_0) = \frac{1}{3} N(N^2-4)(N^2-9). \tag{3.14}$$

Since $J_4(\rho_0)\neq 0$, for A_n $(n\geqslant 3)$, the adjoint representation $\{\rho_0\}$ for A_n $(n\geqslant 3)$ is not exceptional. However, for $N=8$, we find $J_4(\Lambda_2)=0$ so that $\{\Lambda_2\}$ for A_7 is exceptional. Therefore, we must have a special identity

$$\mathrm{Tr}X^4=\tfrac{1}{12}(\mathrm{Tr}X^2)^2 \tag{1.37'}$$

for any generic element X of the Lie algebra A_7 [SU(8) group] in the irreducible representation $\{\Lambda_2\}$. But this fact is not accidental for the following reason. The 56-dimensional representation $\{\Lambda_6\}$ of the exceptional algebra E_7 decomposes[5] into a direct sum $\{\Lambda_2\}\oplus\{\Lambda_6\}$ of its subalgebra A_7. Noting that $\{\Lambda_6\}$ is contragradient to $\{\Lambda_2\}$ for A_7, the validity of the quartic identity (1.37') for $\{\Lambda_2\}$ of A_7 follows from the corresponding relation Eq. (1.37) for E_7. Similarly, restricting ourselves to SO(8) and/or Sp(8) subgroups of SU(8), we expect to have the validity of the same relation (1.37') for the irreducible representation $\{\Lambda_2\}$ of these subalgebras by similar reasoning. This fact will be verified shortly for both C_4 and D_4.

Corresponding to the decomposition

$$\{\Lambda_1\}\otimes\{\Lambda_1\}=\{\Lambda_2\}\oplus\{2\Lambda_1\},$$

we must have a sum rule

$$2[d(\Lambda_1)]^2J_4(\Lambda_1)=d(\Lambda_2)J_4(\Lambda_2)+d(2\Lambda_1)J_4(2\Lambda_1)$$

by Eq. (2.8). The validity of this identity can be verified easily by the numerical results of Eq. (3.13) together with

$$d(\Lambda_j)=\frac{N(N-1)\cdots(N-j+1)}{j!},\quad (1\leqslant j\leqslant N-1).$$

$$d(k\Lambda_1)=\frac{N(N+1)\cdots(N+k-1)}{k!},\quad (k\geqslant 1).$$

Note that the general dimensional formula for $d(\rho)$ is given by

$$d(\rho)=\frac{\Pi_{j<k}^{N}(\sigma_j-\sigma_k)}{1!2!\cdots(N-1)!}=\prod_{j<k}^{N}\frac{\sigma_j-\sigma_k}{\sigma_j^{(0)}-\sigma_k^{(0)}}.$$

It may be interesting to note that $I_2(j\Lambda_1)$, $I_3(j\Lambda_1)$, $J_4(j\Lambda_1)$, and $d(j\Lambda_1)$ for the completely symmetric representation $\{j\Lambda_1\}$ are identical in form to $I_2(\Lambda_j)$, $I_3(\Lambda_j)$, $J_4(\Lambda_j)$, and $d(\Lambda_j)$ for the completely antisymmetric one $\{\Lambda_j\}$, except possibly for signs, when we formally change the dimension N to $-N$.

B. Algebra B_n $(n\geqslant 2)$

We use again the Weyl's symbol f_j $(1\leqslant j\leqslant n)$ so that

$$f_j=m_j+m_{j+1}+\cdots+m_{n-1}+\tfrac{1}{2}m_n. \tag{3.15}$$

All f_j's are simultaneously integers or half-integers, corresponding to tensor or spinor representations of B_n. We set

$$l_j=f_j+n-j+\tfrac{1}{2}\quad(1\leqslant j\leqslant n),$$
$$l_j^{(0)}=n-j+\tfrac{1}{2}\quad(1\leqslant j\leqslant n). \tag{3.16}$$

Then we find

$$2I_2(\rho)=\sum_{j=1}^{n}[(l_j)^2-(l_j^{(0)})^2]$$
$$=\sum_{j=1}^{n}(l_j)^2-\frac{1}{12}n(4n^2-1), \tag{3.17a}$$

$$8I_4(\rho)=\sum_{j=1}^{n}[(l_j)^4-(l_j^{(0)})^4]$$
$$-\frac{1}{6}(2n-1)(4n+1)$$
$$\times\sum_{j=1}^{n}[(l_j)^2-(l_j^{(0)})^2]$$
$$=\sum_{j=1}^{n}(l_j)^4-\frac{1}{6}(2n-1)(4n+1)$$
$$\times\sum_{j=1}^{n}(l_j)^2+\frac{n(4n^2-1)}{720}(44n^2-20n+11), \tag{3.17b}$$

$$8J_4(\rho)=(2n^2+n+2)\sum_{j=1}^{n}[(l_j)^4-(l_j^{(0)})^4]$$
$$-(4n+1)\left\{\left[\sum_{j=1}^{n}(l_j)^2\right]^2-\left[\sum_{j=1}^{n}(l_j^{(0)})^2\right]^2\right\}$$
$$=(2n^2+n+2)\sum_{j=1}^{n}(l_j)^4-(4n+1)\left[\sum_{j=1}^{n}(l_j)^2\right]^2$$
$$+\frac{1}{360}n(n^2-1)(4n^2-1)(2n+3)(2n-7). \tag{3.17c}$$

We observe that $J_4(\rho)$, in contrast to $I_4(\rho)$, contains only quartic polynomials of l_j. The same property is also shared by other Lie algebras C_n and D_n, as we will see.

We also give expressions for $J_4(\Lambda_j)$ and $J_4(k\Lambda_1)$ below.

$$24J_4(\Lambda_j)=(n+1)(2n+3)j(2n+1-j)$$
$$\times\{(n+1)(2n+1)-3j(2n+1-j)\}\quad(1\leqslant j\leqslant n-1), \tag{3.18a}$$
$$24J_4(\Lambda_n)=-\tfrac{1}{8}n(n^2-1)(4n^2-1)(2n+3)\quad(j=n), \tag{3.18b}$$
$$24J_4(k\Lambda_1)=(n-1)(2n-1)k(2n-1+k)$$
$$\times\{(2n^2-n+3)+3k(2n-1+k)\}\quad(k\geqslant 1), \tag{3.18c}$$

for $\Lambda=\Lambda_j$ $(1\leqslant j\leqslant n)$ and $\Lambda=k\Lambda_1$ $(k\geqslant 1)$. Again we can verify the validity of

$$2[d(\Lambda_1)]^2J_4(\Lambda_1)=d(\Lambda_2)J_4(\Lambda_2)+d(2\Lambda_1)J_4(2\Lambda_1), \tag{3.19}$$

corresponding to

$$\{\Lambda_1\}\otimes\{\Lambda_1\}=\{\Lambda_2\}\oplus\{2\Lambda_1\}\oplus\{0\}, \tag{3.20}$$

for B_n $(n\geqslant 3)$, where $\{0\}$ refers to the trivial representation. Also, we note

$$I_2(\rho_0)=I_2(\Lambda_2)=2n-1,\quad(n\geqslant 3),$$
$$I_2(\Lambda_1)=n,$$
$$12J_4(\rho_0)=(n^2-1)(2n+3)(2n-1)(2n-7)\neq 0.$$

We remark that, for $n=13$ and $j=6$, we find $J_4(\Lambda_6)=0$ and hence that $\{\Lambda_6\}$ for B_{13} is exceptional. Finally, the dimension formula is

$$d(\rho)=\frac{2^n}{1!3!\cdots(2n-1)!}\left(\prod_{i=1}^{n}l_i\right)\prod_{j<k}^{n}[(l_j)^2-(l_k)^2],$$

so that

$$d(\Lambda_j)=\frac{(2n+1)!}{j!(2n+1-j)!}\quad(1\leqslant j\leqslant n-1),$$
$$d(k\Lambda_1)=\frac{2n+2k-1}{k!}\frac{(2n+k-2)!}{(2n-1)!}\quad(k\geqslant 1).$$
$$d(\Lambda_n)=2^n\quad(j=n).$$

C. Lie algebra $C_n (n \geqslant 2)$

We set

$$f_j = m_j + m_{j+1} + \cdots + m_n, \tag{3.21a}$$

$$l_j = f_j + n - j + 1, \tag{3.21b}$$

$$l_j^{(0)} = n - j + 1, \tag{3.21c}$$

for $j = 1,2,\ldots,n$. Then we calculate

$$2I_2(\rho) = \sum_{j=1}^{n} [(l_j)^2 - (l_j^{(0)})^2] = \sum_{j=1}^{n} (l_j)^2 - \frac{1}{6} n(n+1)(2n+1), \tag{3.22a}$$

$$\begin{aligned} 8I_4(\rho) &= \sum_{j=1}^{n} [(l_j)^4 - (l_j^{(0)})^4] - \frac{1}{3}(n+1)(4n+1) \times \sum_{j=1}^{n} [(l_j)^2 - (l_j^{(0)})^2] \\ &= \sum_{j=1}^{n} (l_j)^4 - \frac{1}{3}(n+1)(4n+1) \sum_{j=1}^{n} (l_j)^2 + \frac{1}{90} n(n+1)(2n+1)[29n^2 + 64n + 32], \end{aligned} \tag{3.22b}$$

as well as

$$\begin{aligned} 8J_4(\rho) &= (2n^2 + n + 2) \sum_{j=1}^{n} [(l_j)^4 - (l_j^{(0)})^4] - (4n+1)\left\{\left[\sum_{j=1}^{n} (l_j)^2\right]^2 - \left[\sum_{j=1}^{n} (l_j^{(0)})^2\right]^2\right\} \\ &= (2n^2 + n + 2) \sum_{j=1}^{n} (l_j)^4 - (4n+1)\left[\sum_{j=1}^{n} (l_j)^2\right]^2 + \frac{1}{180} n(n^2 - 1)(4n^2 - 1)(2n+3)(n+4). \end{aligned} \tag{3.23}$$

We calculate also

$$24J_4(\Lambda_j) = (n+1)(2n+3)j(2n+2-j) \times \{(2n^2 + 3n + 4) - 3j(2n+2-j)\}, \tag{3.24a}$$

$$24J_4(k\Lambda_1) = (n-1)(2n-1)k(2n+k) \times \{n(2n-1) + 3k(2n+k)\} \tag{3.24b}$$

for the completely antisymmetric representation $\{\Lambda_j\}$ $(1 \leqslant j \leqslant n)$ and for the completely symmetric representation $\{k\Lambda_1\}$ $(k \geqslant 1)$. We can check the validity of Eq. (3.19) corresponding to Eq. (3.20) for this case, using Eq. (3.24). We note also

$$I_2(\rho_0) = I_2(2\Lambda_1) = 2(n+1),$$

$$I_2(\Lambda_1) = \tfrac{1}{2}(2n+1),$$

$$J_4(\rho_0) = J_4(2\Lambda_1) = \tfrac{1}{6}(n^2 - 1)(2n-1)(2n+3) \times (n+4) \neq 0.$$

In particular, the adjoint representation $\{\rho_0\}$ is nonexceptional. We have, however, $J_4(\Lambda_2) = 0$ for $n = 4$ so that $\{\lambda\} = \{\Lambda_2\}$ for C_4 is exceptional. Therefore, for $\{\rho\} = \{\Lambda_2\}$ of C_4, we have the validity of Eq. (1.37).

The dimensional formula for C_n is

$$d(\rho) = \frac{1}{1!3!\cdots(2n-1)!}\left(\prod_{i=1}^{n} l_i\right) \times \prod_{j<k}^{n} [(l_j)^2 - (l_k)^2].$$

In particular, we have

$$d(\Lambda_j) = \frac{2(n+1-j)}{j!} \frac{(2n+1)!}{(2n+2-j)!} \quad (1 \leqslant j \leqslant n),$$

$$d(k\Lambda_1) = \frac{(2n+k-1)!}{k!(2n-1)!} \quad (k \geqslant 1).$$

D. Lie algebra $D_n (n \geqslant 3)$

We define

$$f_j = m_j + m_{j+1} + \cdots + m_{n-2} + \tfrac{1}{2}(m_{n-1} + m_n), \quad 1 \leqslant j \leqslant n-2, \tag{3.25}$$

$$f_{n-1} = \tfrac{1}{2}(m_{n-1} + m_n), \quad f_n = \tfrac{1}{2}(-m_{n-1} + m_n),$$

and set

$$l_j = f_j + n - j, \qquad l_j^{(0)} = n - j, \tag{3.26}$$

for $1 \leqslant j \leqslant n$. We then calculate

$$2I_2(\rho) = \sum_{j=1}^{n} [(l_j)^2 - (l_j^{(0)})^2] = \sum_{j=1}^{n} (l_j)^2 - \frac{1}{6} n(n-1)(2n-1), \tag{3.27a}$$

$$\begin{aligned} 8I_4(\rho) &= \sum_{j=1}^{n} [(l_j)^4 - (l_j^{(0)})^4] - \frac{1}{3}(n-1)(4n-1) \sum_{j=1}^{n} [(l_j)^2 - (l_j^{(0)})^2] \\ &= \sum_{j=1}^{n} (l_j)^4 - \frac{1}{3}(n-1)(4n-1) \times \sum_{j=1}^{n} (l_j)^2 + \frac{1}{90} n(n-1)(2n-1) \times [11n^2 - 16n + 8], \end{aligned} \tag{3.27b}$$

as well as

$$\begin{aligned} 8J_4(\rho) &= (2n^2 - n + 2) \sum_{j=1}^{n} [(l_j)^4 - (l_j^{(0)})^4] - (4n-1)\left\{\left[\sum_{j=1}^{n} (l_j)^2\right]^2 - \left[\sum_{j=1}^{n} (l_j^{(0)})^2\right]^2\right\} \\ &= (2n^2 - n + 2) \sum_{j=1}^{n} (l_j)^4 - (4n-1)\left[\sum_{j=1}^{n} (l_j)^2\right]^2 + \frac{1}{180} n(n^2 - 1)(4n^2 - 1)(n-4)(2n-3). \end{aligned} \tag{3.28}$$

For the special case $\Lambda = \Lambda_j$ $(1 \leqslant j \leqslant n)$ and $\Lambda = k\Lambda_1$ $(k \geqslant 1)$, we find

$$24J_4(\Lambda_j) = (n+1)(2n+1)j(2n-j) \times \{n(2n+1) - 3j(2n-j)\} \quad (1 \leqslant j \leqslant n-2), \tag{3.29a}$$

$$24J_4(\Lambda_n) = 24J_4(\Lambda_{n-1}) = -\frac{1}{8}n(n^2-1)(4n^2-1)$$
$$\times(2n-3), \; (j = n-1 \text{ and } n) \tag{3.29b}$$
$$24J_4(k\Lambda_1) = (n-1)(2n-3)k\,[2n-2+k\,]$$
$$\times\{(2n^2-3n+4)+3k\,(2n-2+k\,)\}, \tag{3.29c}$$

from which we again verify the validity of Eq. (3.19). Setting $j=2$ in Eq. (3.29a), we find

$$I_2(\rho_0) = 2(n-1), \quad I_2(\Lambda_1) = \tfrac{1}{2}(2n-1),$$
$$J_4(\rho_0) = \tfrac{1}{6}(n^2-1)(2n+1)(2n-3)(n-4), \tag{3.30}$$
$$J_4(\Lambda_1) = \tfrac{1}{24}(n^2-1)(4n^2-1)(2n-3),$$

so that $J_4(\rho_0)\neq 0$ for $n\geqslant 5$, but $J_4(\rho_0)=0$ for D_4. The dimensional formula for D_n is given by

$$d(\rho) = \frac{2^{n-1}}{2!4!\cdots(2n-4)!}\sum_{j<k}^{n}[(l_j)^2-(l_k)^2]$$
$$= \prod_{j<k}^{n}\frac{(l_j)^2-(l_k)^2}{(l_j^{(0)})^2-(l_k^{(0)})^2}.$$

In particular, we have

$$d(\Lambda_j) = \frac{(2n)!}{j!(2n-j)!} \quad (1\leqslant j\leqslant n-2),$$
$$d(k\Lambda_1) = \frac{2(n+k-1)}{k!}\,\frac{(2n+k-3)!}{(2n-2)!}, \quad k\geqslant 1,$$
$$d(\Lambda_n) = d(\Lambda_{n-1}) = 2^{n-1}.$$

Again, we find that $I_2(\Lambda_j)$, $J_4(\Lambda_j)$, and $d(\Lambda_j)$ of $C_n(D_n)$ will transform into $I_2(j\Lambda_1)$, $J_4(j\Lambda_1)$, and $d(j\Lambda_1)$ of D_n (C_n), apart from signs when we make a formal change $n\to -n$, provided that we restrict ourselves to $1\leqslant j\leqslant n-2$.

E. Equivalences $B_2 = C_2$ and $D_3 = A_3$

It is well known that we have equivalences $B_2 = C_2$ and $A_3 = D_3$. Here we will discuss the effects of these equivalences on Casimir invariants.

First consider B_2 and C_2. We label all relevant quantities, such as l_j in B_2, as $\bar{l}_j$ by adding bars, while unbarred quantities refer to those of the algebra C_2. The correspondence $B_2 = C_2$ implies

$$\bar{m}_2 = m_1, \quad \bar{m}_1 = m_2, \tag{3.31}$$

which is effected by interchange of two simple roots α_1 and α_2 in Dynkin diagrams of these algebras. Then we find

$$\bar{l}_1 = \tfrac{1}{2}(l_1+l_2), \quad \bar{l}_2 = \tfrac{1}{2}(l_1-l_2), \tag{3.32}$$

and can verify that

$$\bar{I}_2(\rho) = \tfrac{1}{2}I_2(\rho),$$
$$\bar{J}_4(\rho) = -\tfrac{1}{4}J_4(\rho) \tag{3.33}$$

from expressions for these quantities given in this section. The differences between normalization factors for I_2 and I_4 in Eq. (3.33) are due to the fact that $\{\lambda\} = \{\Lambda_1\}$ of B_2 transforms to $\{\Lambda_2\}$ but *not* $\{\Lambda_1\}$ of C_2 by Eq. (3.31).

Similarly, we use the barred symbols such as $\bar{m}_j$ for the algebra A_3, while unbarred ones refer to the algebra D_3. Then, an inspection of Dynkin diagrams of A_3 and D_3 requires identification of

$$\bar{m}_1 = m_2, \quad \bar{m}_2 = m_1, \quad \bar{m}_3 = m_3, \tag{3.34}$$

so that we find

$$\sigma_1 = \tfrac{1}{2}(l_1+l_2-l_3),$$
$$\sigma_2 = \tfrac{1}{2}(l_1+l_3-l_2),$$
$$\sigma_3 = \tfrac{1}{2}(l_2+l_3-l_1), \tag{3.35}$$
$$\sigma_4 = -\tfrac{1}{2}(l_1+l_2+l_3),$$

for σ_j $(1\leqslant j\leqslant 4)$ of the Lie algebra A_3. We can verify the validity of the identities

$$\sigma_1^2+\sigma_2^2+\sigma_3^2+\sigma_4^2 = l_1^2+l_2^2+l_3^2, \tag{3.36a}$$
$$\sigma_1^4+\sigma_2^4+\sigma_3^4+\sigma_4^4 = \frac{3}{4}(l_1^2+l_2^2+l_3^2)^2 - \frac{1}{2}(l_1^4+l_2^4+l_3^4), \tag{3.36b}$$

so that we have

$$\bar{I}_2(\rho) = 2I_2(\rho),$$
$$\bar{J}_4(\rho) = -4J_4(\rho). \tag{3.37}$$

4. LIE ALGEBRA D_4 AND TRIALITY

We have to discuss the case of D_4 separately in view of $\operatorname{Dim} V = 3$. It is convenient for our purpose to write the Lie algebra D_4 in a non-Cartan form

$$[X_{ab}, X_{cd}\,] = \delta_{ad}X_{bc} + \delta_{bc}X_{ad} - \delta_{ac}X_{bd} - \delta_{bd}X_{ac} \tag{4.1}$$
$$X_{ab} = -X_{ba} \tag{4.2}$$

where latin indices a,b,c,d assume eight values 1,2,...,8. Correspondingly, we often write

$$X_\mu = X_{ab}, \quad \mu = (a,b\,). \tag{4.3}$$

Choosing $\{\lambda\} = \{\Lambda_1\}$ again, which is the eight-dimensional defining (or basic) representation, the matrix element of x_μ is given by

$$(x_{ab}\,)_{cd} = \delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}, \tag{4.4}$$

while the Killing metric tensor $g_{\mu\nu}$ also has the same form

$$g_{(ab),(cd)} = 2[\delta_{ad}\delta_{bc} - \delta_{ac}\delta_{bd}\,], \tag{4.5}$$

for our normalization Eq. (3.1). We define $g_{\mu\nu\alpha\beta}$ and J_4 by Eqs. (1.18) and (1.21), respectively. Now, D_4 possesses one additional fourth-order Casimir invariant $\hat{I}_4$, which can be constructed as follows. Let

$$\mu_j = (a_j, b_j), \quad j = 1,2,3,4 \tag{4.6a}$$

and set

$$e_{\mu_1\mu_2\mu_3\mu_4} = \epsilon_{a_1b_1a_2b_2a_3b_3a_4b_4}, \tag{4.6b}$$

where $\epsilon_{a_1\cdots b_4}$ is the completely antisymmetric Levi–Civita symbol in eight-dimensional space. Then, we set

$$\hat{I}_4(\rho)E = \frac{1}{4!}\,e^{\mu_1\mu_2\mu_3\mu_4}X_{\mu_1}X_{\mu_2}X_{\mu_3}X_{\mu_4}$$
$$= \frac{1}{4!2^4}\,\epsilon^{a_1b_1a_2b_2a_3b_3a_4b_4}$$
$$\times X_{a_1b_1}X_{a_2b_2}X_{a_3b_3}X_{a_4b_4}. \qquad (4.7)$$

Using the same notation as is given in the previous section, $\hat{I}_4(\rho)$ is then given by[15,16,20]

$$\hat{I}_4(\rho) = l_1l_2l_3l_4, \quad l_1 > l_2 > l_3 > |l_4|. \qquad (4.8)$$

Now, we note that $g_{\mu\nu\alpha\beta}$, $e_{\mu\nu\alpha\beta}$, and $K^{(0)}_{\mu\nu\alpha\beta} = \frac{1}{3}(g_{\mu\nu}g_{\alpha\beta} + g_{\mu\alpha}g_{\nu\beta} + g_{\mu\beta}g_{\nu\alpha})$ are mutually orthogonal in the sense that

$$g^{\mu\nu\alpha\beta}K^{(0)}_{\mu\nu\alpha\beta} = e^{\mu\nu\alpha\beta}K^{(0)}_{\mu\nu\alpha\beta} = g^{\mu\nu\alpha\beta}e_{\mu\nu\alpha\beta} = 0, \qquad (4.9)$$

and these three quantities now span the vector space V. Setting

$$D^{(4)}(\rho) = d(\rho)J_4(\rho),$$
$$\hat{D}^{(4)}(\rho) = d(\rho)\hat{I}_4(\rho), \qquad (4.10)$$

both satisfy Eq. (1.25), so that we have two modified fourth-order indices.

We now define $G_{\mu\nu\alpha\beta}(\rho)$ by Eq. (2.18), as before. However, since $\operatorname{Dim} V = 3$, we have to modify Eq. (2.23) as

$$[2 + d(\rho)]G_{\mu\nu\alpha\beta}(\rho) = B(\rho)g_{\mu\nu\alpha\beta} + C(\rho)e_{\mu\nu\alpha\beta} \quad (4.11)$$

for some constants $B(\rho)$ and $C(\rho)$. Multiplying both sides of Eq. (4.11) by $g^{\mu\nu\alpha\beta}$ and $e^{\mu\nu\alpha\beta}$ and noting the orthogonality relation Eq. (4.9), we obtain

$$d(\rho)J_4(\rho) = B(\rho)d(\lambda)J_4(\lambda), \qquad (4.12a)$$

$$[2 + d(\rho_0)]d(\rho)\hat{I}_4(\rho) = \frac{1}{4!}\,e^{\mu\nu\alpha\beta}e_{\mu\nu\alpha\beta}C(\rho). \quad (4.12b)$$

Since for $\{\lambda\} = \{\Lambda_1\}$, we have $J_4(\lambda)\neq 0$, as we will see from Eq. (3.29a), the expression for $B(\rho)$ is the same as before and is given by Eq. (2.24).

Next, choosing $X_\mu = H_i$, $X_\nu = H_j$, $X_\alpha = H_k$ and $X_\beta = H_l$, multiplying both sides of Eq. (4.11) by $g^{ij}g^{kl}$ and noting that

$$\sum_{i,j,k,l=1}^{4} g^{ij}g^{kl}e_{ijkl} = 0,$$

we find that the term proportional to $C(\rho)$ vanishes. Therefore, from Eq. (2.27), we obtain again Eq. (2.29), and hence Eq. (1.27), i.e.,

$$\bar{I}_4(\rho) = CD^{(4)}(\rho) \qquad (4.13)$$

for the present case also, where C is a constant.

Now, from the table of McKay and Patera,[5] we calculate $\bar{I}_4(\Lambda_1) = 0$ for $\{\rho\} = \{\Lambda_1\}$. On the other hand, Eq. (3.30) requires $D^{(4)}(\Lambda_1)\neq 0$. Therefore, Eq. (4.13) requires that $C = 0$, and we conclude that

$$\bar{I}_4(\rho) = 0 \qquad (4.14)$$

identically for all irreducible representations $\{\rho\}$ of the Lie algebra D_4. Actually, the validity of Eq. (4.14) is not accidental but is intimately connected with the existence of the triality principle[7] for the Lie algebra D_4. We explain this fact below. It is well known[2,8] that the Lie algebra D_4 is very exceptional in the sense that it alone among all simple Lie algebras has the maximum number of outer automorphisms. Let

$$G = \frac{\text{outer-automorphism group}}{\text{inner-automorphism group}} \qquad (4.15)$$

be the quotient group of outer automorphisms over inner automorphisms. Then G is the identity for A_1, G_2, F_4, E_7, E_8, B_n $(n\geqslant 2)$, and C_n $(n\geqslant 2)$, while G is the cyclic group Z_2 for A_n $(n\geqslant 2)$, D_n $(n\geqslant 5)$ and E_6. However, G is Z_3 for D_4, where Z_p is the cyclic group of p objects. Moreover, it is known that G is isomorphic to the invariance group of Dynkin diagrams for these Lie algebras.

As we see from the Dynkin diagram of D_4 (see Fig. 1), then $G = Z_3$ is identified as the permutation group of three simple roots α_1, α_3, and α_4, while the center root α_2 remains invariant. To see this more clearly, let us introduce $K^\mu_{\ \nu}$, $R^{\mu\nu} = -R^{\nu\mu}$, and $R_{\mu\nu} = -R_{\nu\mu}$ $(\mu,\nu = 1,2,3,4)$ as follows[20,25]

$$K^\mu_{\ \nu} = \tfrac{1}{2}\{-X_{\mu\nu} - X_{\mu+4,\nu+4} - i[X_{\mu,\nu+4} - X_{\mu+4,\nu}]\},$$
$$R^{\mu\nu} = -R^{\nu\mu} = \tfrac{1}{2}\{X_{\mu\nu} - X_{\mu+4,\nu+4}$$
$$- i[X_{\mu,\nu+4} + X_{\mu+4,\nu}]\}, \qquad (4.16)$$
$$R_{\mu\nu} = -R_{\nu\mu} = \tfrac{1}{2}\{-X_{\mu\nu} + X_{\mu+4,\nu+4}$$
$$- i[X_{\mu,\nu+4} + X_{\mu+4,\nu}]\}.$$

Then they satisfy commutation relations

$$[K^\mu_{\ \nu}, K^\alpha_{\ \beta}] = \delta^\mu_\beta K^\alpha_{\ \nu} - \delta^\alpha_\nu K^\mu_{\ \beta},$$
$$[K^\mu_{\ \nu}, R^{\alpha\beta}] = -\delta^\alpha_\nu R^{\mu\beta} - \delta^\beta_\nu R^{\alpha\mu},$$
$$[K^\mu_{\ \nu}, R_{\alpha\beta}] = \delta^\mu_\alpha R_{\nu\beta} + \delta^\mu_\beta R_{\alpha\nu}, \qquad (4.17)$$
$$[R_{\mu\nu}, R^{\alpha\beta}] = \delta^\alpha_\mu K^\beta_{\ \nu} + \delta^\beta_\nu K^\alpha_{\ \mu} - \delta^\alpha_\nu K^\beta_{\ \mu} - \delta^\beta_\mu K^\alpha_{\ \nu},$$
$$[R_{\mu\nu}, R_{\alpha\beta}] = 0 = [R^{\mu\nu}, R^{\alpha\beta}],$$

for $\mu,\nu,\alpha,\beta = 1,2,3,4$. The Cartan subalgebra elements H_j $(j = 1,2,3,4)$ may then be identified as

$$H_j = K^j_{\ j}(\text{no summation on } j), \quad j = 1,2,3,4, \qquad (4.18)$$

while all other $K^\mu_{\ \nu}$ $(\mu\neq\nu)$, $R_{\mu\nu}$, and $R^{\mu\nu}$ correspond to E_α's for some nonzero root α. If v^+ is the maximal (or highest) vector[2] in the irreducible representation $\{\rho\}$, then we have

$$H_jv^+ = f_jv^+, \qquad (4.19)$$

where f_j $(j = 1,2,3,4)$ are given by Eq. (3.25). We now define two outer automorphisms, π and σ, in D_4 as follows. First, we define π by

(i) $\mu\neq 4$, $\nu\neq 4$,

$$\pi(K^\mu_{\ \nu}) = K^\mu_{\ \nu}, \quad \pi(R^{\mu\nu}) = R^{\mu\nu},$$
$$\pi(R_{\mu\nu}) = R_{\mu\nu},$$

(ii) $\mu = 4$, $\nu\neq 4$,

$$\pi(K^4_{\ \nu}) = -R_{4\nu}, \quad \pi(R_{4\nu}) = -K^4_{\ \nu},$$

(iii) $\mu\neq 4$, $\nu = 4$,

$$\pi(K^\mu_{\ 4}) = R^{\mu 4}, \quad \pi(R^{\mu 4}) = K^\mu_{\ 4},$$

(iv) $\mu = \nu = 4$,

$$\pi(K^4_{\ 4}) = -K^4_{\ 4}. \qquad (4.20)$$

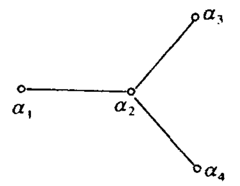


FIG. 1. Numbering of simple roots for the Dynkin diagram of D_4.

We can verify that π is an outer automorphism of D_4. Now, in view of Eqs. (4.19) and (4.20), π induces a mapping of weight systems of D_4 into itself in the same representation space $\{\rho\}$, keeping the maximal vector v^+ intact. In particular,

$$\pi: f_j \to f_j, \quad l_j \to l_j \quad (j \neq 4),$$
$$f_4 \to -f_4, \quad l_4 \to -l_4 \quad (j=4). \tag{4.21}$$

Then, by Eq. (3.25), this is equivalent to interchanging of m_3 and m_4 while m_1 and m_2 remain unchanged. Therefore, we can interpret π to be the element of the permutation group Z_3 which interchanges two simple roots α_3 and α_4 in the Dynkin diagram.

Next, consider another mapping σ defined by

$$\begin{aligned} K^{\mu}_{\ \nu} &\to -K^{\nu}_{\ \mu} + \frac{1}{2}\delta^{\nu}_{\mu} \sum_{\lambda=1}^{4} K^{\lambda}_{\ \lambda}, \\ R^{\mu\nu} &\to -\frac{1}{2}\sum_{\alpha,\beta=1}^{4} \epsilon_{\mu\nu\alpha\beta} R^{\alpha\beta}, \\ R_{\mu\nu} &\to -\frac{1}{2}\sum_{\alpha,\beta=1}^{4} \epsilon^{\mu\nu\alpha\beta} R_{\alpha\beta}, \end{aligned} \tag{4.22}$$

for $\mu,\nu = 1,2,3,4$, where $\epsilon_{\mu\nu\alpha\beta} = \epsilon^{\mu\nu\alpha\beta}$ is the completely antisymmetric Levi–Civita symbol in four-dimensional space. Although σ is also an outer automorphism of D_4, this is not so convenient for our purpose, since the maximal vector v^+ can be seen to be *not* invariant under σ. For this reason, we consider another mapping τ defined as follows. Let W be an operation which interchanges labels 1 and 4 and 2 and 3 ($1 \leftrightarrow 4$ and $2 \leftrightarrow 3$) and set

$$\tau = W\sigma. \tag{4.23a}$$

We simply remark that W corresponds to an inner automorphism. The explicit operation of τ in D_4 is now given by

$$\begin{aligned} K^{\mu}_{\ \nu} &\to -K^{\bar\nu}_{\ \bar\mu} + \frac{1}{2}\delta^{\nu}_{\mu} \sum_{\lambda=1}^{4} K^{\lambda}_{\ \lambda}, \\ R^{\mu\nu} &\to -\frac{1}{2}\sum_{\alpha,\beta=1}^{4} \epsilon_{\bar\mu\bar\nu\bar\alpha\bar\beta} R^{\bar\alpha\bar\beta}, \\ R_{\mu\nu} &\to -\frac{1}{2}\sum_{\alpha,\beta=1}^{4} \epsilon^{\bar\mu\bar\nu\bar\alpha\bar\beta} R_{\bar\alpha\bar\beta}, \end{aligned} \tag{4.23b}$$

where $\bar\mu = 5-\mu$ etc. so that

$$\bar 1 = 4,\ \bar 2 = 3,\ \bar 3 = 2, \text{ and } \bar 4 = 1.$$

We can verify that τ is an outer automorphism of D_4. Studying the effect of τ on the maximal vector v^+, we see that it induces the mapping

$$\begin{aligned} f_1 &\to \bar f_1 = \tfrac{1}{2}(f_1 + f_2 + f_3 - f_4), \\ f_2 &\to \bar f_2 = \tfrac{1}{2}(f_1 + f_2 - f_3 + f_4), \\ f_3 &\to \bar f_3 = \tfrac{1}{2}(f_1 - f_2 + f_3 + f_4), \\ f_4 &\to \bar f_4 = \tfrac{1}{2}(-f_1 + f_2 + f_3 + f_4), \end{aligned} \tag{4.24a}$$

or equivalently,

$$\begin{aligned} l_1 &\to \bar l_1 = \tfrac{1}{2}(l_1 + l_2 + l_3 - l_4), \\ l_2 &\to \bar l_2 = \tfrac{1}{2}(l_1 + l_2 - l_3 + l_4), \\ l_3 &\to \bar l_3 = \tfrac{1}{2}(l_1 - l_2 + l_3 + l_4), \\ l_4 &\to \bar l_4 = \tfrac{1}{2}(-l_1 + l_2 + l_3 + l_4), \end{aligned} \tag{4.24b}$$

or

$$l_\mu \to \bar l_\mu = \tfrac{1}{2}(l_1 + l_2 + l_3 + l_4) - l_{\bar\mu}. \tag{4.24c}$$

Note that the ordering relation

$$\bar l_1 > \bar l_2 > \bar l_3 > |\bar l_4| \tag{4.25}$$

is still preserved. In terms of m_j's, this is equivalent to the interchange of $m_1 \leftrightarrow m_3$, so that we interpret τ to imply the permutation of two simple roots α_1 and α_3. In terms of π and τ, the action of the six elements of Z_3 are given by

$$\begin{aligned} &1. && \text{I:identity}, \\ &2. && \pi: \alpha_3 \leftrightarrow \alpha_4, \\ &3. && \tau: \alpha_1 \leftrightarrow \alpha_3, \\ &4. && \pi\tau\pi = \tau\pi\tau: \alpha_1 \leftrightarrow \alpha_4, \\ &5. && \tau\pi: \alpha_1 \to \alpha_3 \to \alpha_4 \to \alpha_1, \\ &6. && \pi\tau: \alpha_1 \to \alpha_4 \to \alpha_3 \to \alpha_1. \end{aligned} \tag{4.26}$$

Let us label the irreducible representation $\{\rho\}$ as (f_1, f_2, f_3, f_4). Then, for example, we find

$$\pi: \begin{cases} (1,0,0,0) \leftrightarrow (1,0,0,0) \\ (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \leftrightarrow (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2}) \end{cases},$$
$$\tau: \begin{cases} (1,0,0,0) \leftrightarrow (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2}) \\ (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \leftrightarrow (\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) \end{cases}.$$

In other words, three eight-dimensional irreducible representations, corresponding to the vector $(1,0,0,0)$, the spinor $(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})$, and the mirror spinor $(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2})$, interchange among themselves by actions of Z_3. This is one of the well-known manifestations[26] of the triality principle[7] of D_4. The dimension $d(\rho)$ is, of course, invariant under Z_3.

We first observe that $I_2(\rho)$ is invariant under Z_3 if we note

$$\bar l_1^2 + \bar l_2^2 + \bar l_3^2 + \bar l_4^2 = l_1^2 + l_2^2 + l_3^2 + l_4^2, \tag{4.27}$$

where $\bar l_j$ are defined by Eq. (4.24b). However, the fourth-order Casimir invariants are *not* invariant. We can readily see

$$\pi: \begin{cases} J_4(\rho) \to J_4(\rho) \\ \hat I_4(\rho) \to -\hat I_4(\rho) \end{cases}. \tag{4.28}$$

However, by τ, we find

$$\sum_{j=1}^{4} (\bar l_j)^4 = -\frac{1}{2}\sum_{j=1}^{4}(l_j)^4 + \frac{3}{4}\Big[\sum_{j=1}^{4} l_j^2\Big]^2 - 6 l_1 l_2 l_3 l_4, \tag{4.29a}$$

$$\bar{l}_1\bar{l}_2\bar{l}_3\bar{l}_4 = -\frac{1}{8}\sum_{j=1}^{4}(l_j)^4 + \frac{1}{16}\Big[\sum_{j=1}^{4} l_j^2\Big]^2 + \frac{1}{2} l_1 l_2 l_3 l_4, \tag{4.29b}$$

after some calculations. Since Eq. (3.28) for $n = 4$ gives

$$J_4(\rho) = \frac{15}{8}\Big\{2\sum_{j=1}^{4}(l_j)^4 - \Big[\sum_{j=1}^{4}(l_j)^2\Big]^2\Big\} \tag{4.30}$$

for D_4, Eq. (4.29) implies

$$\pi:\begin{cases} J_4(\rho) \to -\frac{1}{2}J_4(\rho) - \frac{45}{2}\hat{I}_4(\rho) \\ \hat{I}_4(\rho) \to -\frac{1}{30}J_4(\rho) + \frac{1}{2}\hat{I}_4(\rho) \end{cases}. \tag{4.31}$$

Then Eqs. (4.28) and (4.31) show that $J_4(\rho)$ and $\hat{I}_4(\rho)$ [and hence $D^{(4)}(\rho)$ and $\hat{D}^{(4)}(\rho)$] form a basis of two-dimensional irreducible representations of Z_3. Note that if we had used $I_4(\rho)$ instead of $J_4(\rho)$, this conclusion would not apply. This is another indication of the naturalness of the modified fourth-order Casimir invariant $J_4(\rho)$ in contrast to $I_4(\rho)$. Also, this implies the impossibility of finding the unique fourth-order Casimir invariant for D_4, which is independent of the reference representation $\{\lambda\}$.

Now, let us return to the discussion of $\bar{I}_4(\rho)$. When we note $g_{jk} = c'\delta_{jk}$ for a constant c' which does not concern us here, we find

$$c'\sum_{j,k=1}^{4} g^{jk}H_jH_k = \sum_{j=1}^{4}(K^j_j)^2 = \sum_{j=1}^{4}(\bar{K}^j_j)^2.$$

Therefore, $\bar{I}_4(\rho)$ is invariant under Z_3 since its invariance under π is also evident. Then, the validity of Eq. (4.13) requires $C = 0$ and hence $\bar{I}_4(\rho) = 0$ identically, since $D^{(4)}(\rho)$ belongs to a doublet representation of Z_3. This is the reason why $\bar{I}_4(\rho) = 0$ is not accidental but is related to the triality.

Also, we remark the following. Any irreducible representation $\{\rho\} = (f_1, f_2, f_3, f_4)$ satisfying

$$f_1 = f_2 + f_3, \quad f_4 = 0 \tag{4.32}$$

is invariant under Z_3 since $\Lambda = f_3(\Lambda_1 + \Lambda_3 + \Lambda_4) + (f_2 - f_3)\Lambda_2$, so that in view of Eq. (4.31), we must have

$$J_4(\rho) = \hat{I}_4(\rho) = 0.$$

This can be directly verified[27] also from Eqs. (4.8) and (4.30). Then Eq. (4.11) requires $G_{\mu\nu\alpha\beta}(\rho) = 0$ for any such representation. In particular, we have the validity of Eq. (1.37), i.e.,

$$\mathrm{Tr}X^4 = K(\rho)(\mathrm{Tr}X^2)^2 \tag{4.33}$$

for any irreducible representation $\{\rho\}$ satisfying the condition Eq. (4.32). Especially the adjoint representation $\{\rho_0\} = (1,1,0,0)$ must satisfy Eq. (4.33), as was already noted by Cvitanovic[28] some time ago.

As another application of triality, we note that if $f_4 = 0$, then a similar consideration of Eq. (4.11) leads to the validity of Eq. (1.35). However, we can generalize this into the following type of representations:

$$\begin{aligned} \{\rho_1\} &= (f_1, f_2, f_3, 0), \\ \{\rho_2\} &= (\bar{f}_1, \bar{f}_2, \bar{f}_3, \bar{f}_4), \\ \{\rho_3\} &= (\bar{f}_1, \bar{f}_2, \bar{f}_3, -\bar{f}_4), \end{aligned} \tag{4.34a}$$

where $\bar{f}_4$ satisfies the condition

$$\bar{f}_4 = \bar{f}_1 - \bar{f}_2 - \bar{f}_3. \tag{4.34b}$$

Then Eq. (1.35) is valid for three types of representations: $\{\rho_1\}$, $\{\rho_2\}$, and $\{\rho_3\}$. The reason is as follows. By the action of Z_3, the type $\{\rho_1\}$ representation transforms into $\{\rho_2\}$ and/or $\{\rho_3\}$. Then we identify

$$\begin{aligned} \bar{f}_1 &= \tfrac{1}{2}(f_1 + f_2 + f_3), \\ \bar{f}_2 &= \tfrac{1}{2}(f_1 + f_2 - f_3), \\ \bar{f}_3 &= \tfrac{1}{2}(f_1 - f_2 + f_3), \\ \bar{f}_4 &= \tfrac{1}{2}(-f_1 + f_2 + f_3). \end{aligned} \tag{4.35}$$

But for $\{\rho_1\}$, we have $\hat{I}_4(\rho_1) = 0$, so that Eq. (4.31) requires $\hat{I}_4(\rho_j) = \frac{1}{15}J_4(\rho_j)$ for $j = 2$ and 3. Then, from Eqs. (4.11) and (4.12), we find the validity of Eq. (1.35) also for $\{\rho_2\}$ and $\{\rho_3\}$, although the value of $C_4(t)$ changes.

We note that if Eq. (4.32) is satisfied, then $\{\rho_1\} = \{\rho_2\} = \{\rho_3\}$. Also if we set $f_2 = f_3 = 0$ and $f_1 = p$ in Eq. (4.35), then three representations

$$(p,0,0,0),$$

$$\Big(\frac{p}{2}, \frac{p}{2}, \frac{p}{2}, \frac{p}{2}\Big),$$

and

$$\Big(\frac{p}{2}, \frac{p}{2}, \frac{p}{2}, -\frac{p}{2}\Big)$$

transform among themselves under Z_3. The case $p = 1$ corresponds to the eight-dimensional representation.

In concluding this section, we simply note that

$$4\sum_{j=1}^{4}(\bar{l}_j)^6 - 5\Big[\sum_{j=1}^{4}(\bar{l}_j)^4\Big]\Big[\sum_{k=1}^{4}(\bar{l}_k)^2\Big] = 4\sum_{j=1}^{4}(l_j)^6 - 5\Big[\sum_{j=1}^{4}(l_j)^4\Big]\Big[\sum_{k=1}^{4}(l_k)^2\Big] \tag{4.36}$$

is invariant under Z_3. This fact is relevant to the study of sixth-order Casimir invariants of D_4, which can again be classified by actions of Z_3.

ACKNOWLEDGMENTS

This paper is in part supported by the U. S. Department of Energy under Contract No. DE-AC02-76ER13065.

APPENDIX

As we noted in Eq. (1.27), we have

$$\bar{I}_4(\rho) = Cd(\rho)J_4(\rho). \tag{A1}$$

Since the constant C is independent of the generic irreducible representation $\{\rho\}$, we can compute it by

$$C = \bar{I}_4(\Lambda_1)/d(\Lambda_1)J_4(\Lambda_1), \tag{A2}$$

where we can calculate $J_4(\Lambda_1)$ by the formulas of Sec. 3. We may evaluate $l_2(\Lambda_1)$ and $l_4(\Lambda_1)$ directly from the defining equations (1.1) and (1.5). However, our normalization condition Eqs. (3.2) and (3.2′) imply

$$l_2(\Lambda_1) = n, \tag{A3}$$

which differs by a factor of 2 for algebras B_n ($n \geqslant 2$) and D_n ($n \geqslant 3$) from those adopted by McKay and Patera.[5] In order to make a definite comparison to results of Ref. 5, we renormalize our inner product by

$$(\alpha,\alpha)_{\max} = 2 \tag{A4}$$

for simple roots, as in their paper, while formally retaining explicit expressions for $J_4(\rho)$ given in Sec. 3. With this understanding, we recalculate and find

1. A_n **($n \geqslant 1$)**

$$l_2(\Lambda_1) = n = N - 1, \quad l_4(\Lambda_1) = \frac{n^2}{n+1} = \frac{(N-1)^2}{N}, \tag{A5}$$

$$\bar{l}_4(\Lambda_1) = \frac{(N-1)(N-2)(N-3)}{3(N^2+1)},$$

$$C = \frac{2}{(N+1)(N+2)(N+3)(N^2+1)}.$$

2. B_n **($n \geqslant 2$)**

$$l_2(\Lambda_1) = l_4(\Lambda_1) = 2n,$$

$$\bar{l}_4(\Lambda_1) = \frac{4n(n-1)(n-2)}{3(2n^2+n+2)},$$

$$C = \frac{16(n-2)}{(n+1)(2n+3)(4n^2-1)(2n^2+n+2)}.$$

3. C_n **($n \geqslant 2$)**

$$l_2(\Lambda_1) = n, \quad l_4(\Lambda_1) = \frac{n}{2},$$

$$\bar{l}_4(\Lambda_1) = \frac{n(n-1)(n-2)}{3(2n^2+n+2)},$$

$$C = \frac{4(n-2)}{(n+1)(2n+3)(4n^2-1)(2n^2+n+2)}.$$

4. D_n **($n \geqslant 3$)**

$$l_2(\Lambda_1) = l_4(\Lambda_1) = 2n,$$

$$\bar{l}_4(\Lambda_1) = \frac{4n(n-1)(n-4)}{3(2n^2-n+2)},$$

$$C = \frac{16(n-4)}{(n+1)(2n-3)(4n^2-1)(2n^2-n+2)}.$$

From these, we see $\bar{l}_4(\rho) = 0$ identically again for D_4. However, we find also $\bar{l}_4(\rho) = 0$ for B_2 and C_2. The reason behind the validity for the latter algebras is rather obscure in contrast to that for D_4, which has been discussed in Sec. 4.

[1]E. B. Dynkin, Mat. Sbornik, N. S. **30**, 349 (1952) [Amer. Math. Soc. Transl. Ser. **2**, Vol. 6, 111 (1957)].

[2]J. E. Humphrey, *Introduction to Lie Algebras and Representation Theory*, Graduate Text in Math., No. 9 (Springer, Berlin, 1972).

[3]S. Okubo, J. Math. Phys. **20**, 586 (1979). This paper will be referred to as (I).

[4]J. Patera, R. T. Sharp, and P. Winternitz, J. Math. Phys. **17**, 1972 (1976), Erratum **18**, 1519 (1977). We have modified their notation $I^{(2p)}$ to $l_{2p}(\rho)$ since we reserve $I_p(\rho)$ for eigenvalues of pth-order Casimir invariants.

[5]W. G. McKay and J. Patera, *Tables of Dimensions, Indices, and Branching Rules for Representations of Simple Lie Algebras* (Dekker, New York, 1981).

[6]J. Patera and R. T. Sharp, J. Math. Phys. (to appear); J. McKay, J. Patera, and R. T. Sharp, U. Montreal Report CRMA-1011 (1981).

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

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

[9]G. Racah, Princeton Lecture Note (1951); B. Gruber and L. O'Raifeartaigh, J. Math. Phys. **5**, 1796 (1964).

[10]S. Okubo, Phys. Rev. D **16**, 3528 (1977).

[11]A. Borel and C. Chevalley, Amer. Math. Soc. Mem. **14**, 1 (1955); C. Warner, *Harmonic Analysis on Semi-Simple Lie Groups* (I) (Springer, New York, 1972) p. 144.

[12]J. Banks and H. Georgi, Phys. Rev. D **14**, 1159 (1976).

[13]Y. Tosa and S. Okubo, Phys. Rev. D **23**, 3058 (1981).

[14]S. Okubo, Phys. Rev. D **18**, 3792 (1978).

[15]J. D. Louck, Los Alamos Scientific Lab. Rep. (1960) (unpublished).

[16]A. M. Perelomov and V. S. Popov, Yad. Fiz. **3**, 924, 1127 (1966) [Sov. J. Nucl. Phys. **3**, 679, 819 (1966)]; V. S. Popov and A. M. Perelomov, *ibid.*, **5**, 693 (1967) [*ibid.*, **5**, 489 (1967)].

[17]J. D. Louck and L. C. Biedenharn, J. Math. Phys. **11**, 2368 (1970); H. S. Green, *ibid.*, **12**, 2107 (1971); S. Okubo, *ibid.*, **16** 528 (1975); M. C. K. Aguilera–Navarro and V. C. Aguilera–Navarro, *ibid.*, **17**, 1173 (1976).

[18]M. X. F. Wong and H. Y. Yeh, J. Math. Phys. **16**, 1239 (1975).

[19]C. O. Nwachuku and M. A. Rashid, J. Math. Phys. **17**, 1611 (1976).

[20]S. Okubo, J. Math. Phys. **18**, 2382 (1977).

[21]S. A. Edwards, J. Math. Phys. **19**, 164 (1978).

[22]M. J. Englefield and R. C. King, J. Phys. A **13**, 2297 (1980).

[23]H. Weyl, *Classical Groups* (Princeton U. P., Princeton, N. J. 1939).

[24]These expressions for $I_2(\rho)$ and $I_3(\rho)$ differ by a factor of 2 from those given in Ref. 10. This is of course due to a difference of normalization for $g_{\mu\nu}$.

[25]S. Okubo, Phys. Rev. C **10**, 2048 (1974).

[26]For another way of formulating the triality and for its connection to the octonion algebra, see M. Günaydin and F. Gürsey, J. Math. Phys. **14**, 1651 (1973).

[27]We note an identity $[a^2 + b^2 + c^2]^2 - 2(a^4 + b^4 + c^4) = (a + b + c) \times (a + b - c)(b + c - a)(c + a - b)$ for any three numbers a, b, and c. This identity also follows immediately from Eq. (4.29b) when we set $l_4 = 0$, while Eq. (4.29a) will essentially reproduce Eq. (3.36b) for $l_4 = 0$. The validity of this relation is intimately related by Eq. (1.37) to the Lie algebra A_2 as well as the pseudo-octonion algebra, as we noted in Ref. 3.

[28]P. Cvitanovic, private communication (1979) where he used a method developed by P. Cvitanovic, Phys. Rev. D **14**, 1536 (1976) and U. Oxford Report 40/77 (1977) (unpublished).

Gribov ambiguities from the bifurcation theory viewpoint

Edward Malec
Institute of Physics, Jagellonian University, Reymonta 4, 30-059 Kraków 16, Poland

(Received 3 June 1980; accepted for publication 14 November 1980)

The connection between singularities of the Faddeev–Popov determinant and the local gauge degeneracy is discussed. A criterion relating the two phenomena is given. It is proved that a whole neighborhood of the local gauge copy is filled with copies of transverse potentials.

PACS numbers: 02.30.Jr

1. INTRODUCTION

Let us define a gauge as a section of the connection bundle[1] A/Z over A/G, where A denotes the space of connections and Z the center of the compact semisimple gauge group G. Usually one defines the gauge by imposing the Coulomb condition upon potentials A_i^a,

$$\partial_i A_i^a = 0. \tag{1}$$

This work is devoted to analysis of the local uniqueness of gauge from the standpoint of bifurcation theory. We deal with the Euclidean manifolds with a boundary as well as with R^3 and R^4 noncompact spaces.

Gribov[2] argued that:

(i) if the Faddeev–Popov determinant is singular at some transverse potentials [i.e., ones satisfying (1)], then the condition (1) does not assure uniqueness—there are local degeneracies;

(ii) the gauge is locally unique for sufficiently weak potentials.

It should be noted that the singularity of the Faddeev–Popov determinant corresponds to a nonzero solution of our equation (4), which is the result of the linearization of the transversality condition (3). It should be stressed, that Gribov's statement (i) is not at all obvious, since solutions to linearized equations may not be tangent to any curve of exact solutions to a full nonlinear equation (see a counter example in Berger[3]). The validity of (i) will be corroborated only partially in this paper, under some assumptions about the multiplicity of solutions to the linearized Eq. (4). This will be done in Theorem 0: its possible generalization including the noncompact R^n case is discussed in Sec. 5.

We will show that the locally degenerate potentials can have nonvanishing measure in path integral quantization (that fact was pointed out by Moncrief,[4] but he dealt with global copies and used different techniques). Two examples of copied potentials are contained in Sec. 3.

The local uniqueness for weak potentials, so important from the perturbation theory viewpoint, is well known and probably proved previously (its generalization for noncompact R case was done essentially by Moncrief[4]). I will discuss it in Sec. 4 and 5 only for completeness.

2. MAIN RESULTS

Let us define the element of the gauge group G by $g = \exp(-i\boldsymbol{\alpha}(\mathbf{x})\cdot\boldsymbol{\sigma})$, where $\boldsymbol{\sigma}$ is the generator of the Lie algebra $\bar{G}$. We assume a basis for the Lie algebra $\bar{G}$ in which the structure constants f^a_{bc} are completely antisymmetric. The potentials transform under the gauge transformation according to a pseudotensor rule:

$$(hA_i)^g = hg^{-1}A_i g + ig^{-1}\partial_i g. \tag{2}$$

The parameter h may be interpreted as the intensity of the potential A_i, after suitable renormalization. Such redefinition of A_i is always possible and it has the following advantage: that we may use bifurcation theory methods treating h as a bifurcation parameter. The bifurcating solutions g obtained below are analytical functions of h.

The Coulomb condition is, inserting (2) into (1):

$$\partial_i(hg^{-1}A_i^b\sigma_b g + ig^{-1}\partial_i g)^a = 0. \tag{3}$$

Suppose that A_i^a and α^a are of class $C^\infty(\Omega)$ and on the boundary

$$\alpha^a(\partial\Omega) = 0.$$

(As a matter of fact I should prove our theorems previously in suitable Hölder or Sobolev spaces and then use the Sobolev embedding theorem[6] to show that obtained solutions α^a are of class C^∞. I omit such technical details now and elsewhere).

Expanding in (3) the functions g near $\alpha = 0$ give (including terms to first order)

$$\Delta\alpha^a + 2hf^a_{cd}A_i^c\partial_i\alpha^d = 0. \tag{4}$$

If these equations have a nontrivial solution at $h = h_0$, vanishing on the boundary, and the critical values of h are isolated, then the full nonlinear equation (3) may have bifurcating solutions $g \neq 1$, $g(\partial\Omega) = 1$.

The linearized operator D defined above [the Frechet derivative of (3) at $\alpha^a = 0$] is skew-symmetric, with elliptic symbol[5,6] so its kernel is finite, $\dim \ker D < \infty$, (ii) $\ker D^* = \operatorname{coker} D = \ker D$.[5–7] Under our assumptions about the differentiability of potentials and gauge transformations, taking into account (i) and (ii) we conclude that $\ker D \cap \operatorname{Range} D = 0$, and the operator D acts bijectively between $C^\infty(\Omega)/\ker D \to C^\infty(\Omega)/\ker D$.

Hence, as a direct consequence of the theorem 4.2.3.[5] we get:

Theorem 1: If Eq. (4) have an odd number of solutions at $h = h_0$, then Eq. (3) have nontrivial solutions $\alpha^a(\mathbf{x}, h - h_0)$ of class C^∞, such that $\alpha^a \to 0$ as $h \to h_0$. These solutions are approximated to first-order by some combinations of the elements of $\ker D$.

 0022-2488/82/010021-03$02.50

The case of simple multiplicity may be examined in more detail by using the Lyapunov–Schmidt procedure.[7] The computations yield

$$\alpha^a = f^a(h - h_0)\times\text{const} + O(h - h_0), \tag{5}$$

where $f^a \in \ker D$, $O(h - h_0)/(h - h_0) \to 0 \quad (h \to h_0)$.

In order to avoid misunderstandings we stress that for transverse A_i^a and h fixed there is only one element g near the unity of gauge group, such that Eq. (3) hold (assuming simple multiplicity). The fixed transverse potential hA_i with h sufficiently near h_0 has a transverse gauge copy $(hA_i)^g$ where $g = \exp(-i\alpha^a\sigma_a)$, and consequently a curve of transverse copied potentials $(hA_i)^g$ parametrized by $(h - h_0)$ corresponds to a curve hA_i [if $(h - h_0)$ is small enough]. Both curves originate from the point $h_0 A_i^\alpha$ which has no small gauge copy itself.

Now we prove

Theorem 2: The bifurcating solution $g = \exp(-i\alpha^a\sigma_a)$, where α^a is taken from (5), is locally unique.

Proof: Our thesis follows from the implicit function theorem which asserts that bifurcating solutions are unique in a sufficiently small neighborhood of critical points (h_0, α^a).[5,6,7] These results are valid in two, three, and four Euclidean dimensions.

It follows from the Theorem 2 that the Frechet derivative D does not vanish at (hA_i^a, α^a) if $h \neq h_0$, and $h \approx h_0$(α^a is the bifurcating solution). Therefore the implicit function theorem assures the existence of a locally unique C^∞ continuation $g(A_i^a)$ on some open transverse neighborhood V of $h_0 A_i^a$. {Note that the operator defined by

$$\phi^b(hA_i^a, \overline{\alpha}^a) = \partial_i\,[\exp(+i\overline{\alpha}^a\sigma_a)hA_i\exp(-i\overline{\alpha}^a\sigma_a) + i\exp(+i\overline{\alpha}^a\sigma_a)\partial_i\exp(-i\overline{\alpha}^a\sigma_a)^b\,]$$

acting between (C^∞ transverse potentials) $\times$ (C^∞ Lie algebra $\bar{G}$ valued functions)$\to$(C^∞ Lie algebra $\bar{G}$ valued functions), annihilates (hA_i^a, α^a), but $\partial\phi^a/\partial\overline{\alpha}^b\,|_{\overline{\alpha}^b = \alpha^b}$ is nonzero if $h = h_0$}.

Hence we conclude that a whole neighborhood of each local gauge copy $(hA_i^a)^g$ is filled with copies of transverse potentials near $h_0 A_i^a$, $\bar{A}_i^a \in V$. This fact means that the degenerate potentials can have nonvanishing measure in path integral quantization. Section 3 presents two families of such potentials. Moncrief[4] obtained the result, but contrary to the way it was presented above, he requires that the global copies of small potentials are known; since they are not obtainable by linearization, it is rather difficult to find some explicit examples.

3. EXAMPLES

[Throughout this section we put $G = \mathrm{SU}(2)$].

(a) Let Ω belong to R^3, with a sphere as a boundary. Suppose

$$A_i^a = \delta_{a1} kr\epsilon_{ikl}x_k\partial_l\cos\upsilon, \tag{6}$$

where δ_{a1} is the Kronecker's symbol.

Then Eq. (4) are

$$\Delta f^2(\vec{x}) - 2hk(r)\epsilon_{ikl}\,x_k\,\partial_l\,\cos\upsilon\,\partial_i f^3(\mathbf{x}) = 0,$$
$$\Delta f^3(\vec{x}) + 2hk(r)\epsilon_{ikl}\,x_k\,\partial_l\,\cos\upsilon\,\partial_i f^2(\mathbf{x}) = 0,$$
$$\Delta f^1(\vec{x}) = 0. \tag{7}$$

The boundary conditions are $f^a(\partial\Omega) = 0$. f^1 is zero, while f^2 and f^3 may be found in the form

$$f^2(\vec{x}) = \sin\varphi P_l^1(\upsilon)f^2(r),$$
$$f^3(\vec{x} = \cos\varphi P_l^1(\upsilon)f^3(r). \tag{8}$$

The equations for $f^2(r)$ and $f^3(r)$ are

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{l+1}{r^2}\,l\right)f^2 + \frac{2hk(r)f^3}{r} = 0,$$
$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - l\,\frac{l+1}{r^2}\right)f^3. + \frac{2hk(r)f^2}{r} = 0. \tag{9}$$

Put $k(r) = -r$ and $l = 1$. Then $f^2(r) = f^3(r) = r^{-\frac{1}{2}}\mathscr{J}_{3/2}(\sqrt{2h}\;r)$, where $\mathscr{J}_{3/2}$ is the Bessel function. The boundary condition $\mathscr{J}_{3/2}(\sqrt{2h}\;R_0) = 0$ effects a quantization of h. Nonzero solutions exist for sufficiently large h and a smallest possible value of h, say h_0, corresponds to a simple multiplicity. The solution bifurcating from $g = 1$ is (to first order)

$$g(\mathbf{x}) = \exp[-i(h - h_0)\text{const}P_1^1(\upsilon)\mathscr{J}_{3/2}(\sqrt{2h}\;r)] \times(\sigma_2\sin\varphi + \sigma_3\cos\varphi.)] \tag{10}$$

(b) The same analysis in a two-dimensional case, under the assumption

$$A_i^a = \delta_{a1}\epsilon_{ik}x_k k(r),\ \epsilon_{ik} = -\epsilon_{ki},\ \epsilon_{12} = 1, \tag{11}$$

yields the following form of the solutions to the linearized equations:

$$f^1(\mathbf{x}) = 0,$$
$$f^2(\mathbf{x}) = \cos\varphi f(r),$$
$$f^3(\mathbf{x}) = \sin\varphi f(r), \tag{12}$$

where $f(r)$ is a solution of

$$f'' + \frac{1}{r}f' - \frac{1}{r^2}f - \frac{2hk(r)}{r}f = 0,\quad f(R_0) = 0. \tag{13}$$

(R_0 is a radius of a boundary).

Suppose $k(r) = \text{const} = c$ (such choice corresponds to a constant magnetic field along the third axis in the three-dimensional context). Then $f = \mathscr{J}_2(2\sqrt{-2hcr})$, and the critical values of h are given by $\mathscr{J}_2(2\sqrt{-2hcR_0}) = 0$.

The bifurcating solution is

$$g(\mathbf{x}) = \exp\left[-i(h - h_0)\text{const}\mathscr{J}_2(2\sqrt{-2hc\surd r})\right. \times(\partial_2\cos\varphi + \partial_3\sin\varphi.)] \tag{14}$$

4. THE COMPACT CASE

The Frechet derivative of the operator defined in (3) at $h = 0$, $\alpha^a = 0$ is simply the Laplacian. The Laplacian acts isomorphically between $C_\infty^0(\Omega$ (spaces of functions vanishing on the boundary $\partial\Omega) \to C^\infty(\Omega)$, and more generally, between $C_0^{k+2+d} \to C^{k+d}$ (Hölder) or $W_0^{k+2} \to W^k$ (Sobolev) spaces.[5,6,7] So from the implicit function theorem we obtain:

Theorem 3: The gauge exists locally for sufficiently weak potentials [i.e., there exists such neighborhood of $g = 1$ where Eq. (3) has no solutions $g \neq 1$].

5. THE NONCOMPACT CASE

Consider the case when Ω is replaced by R^N. The Hölder or Sobolev spaces are not correct now.[8] We will use those of Cantor–Nirenberg–Walker, $M^p_{s,\delta}(R^n, R^m)$.[8] $M^p_{s\delta}$ is the completion of C_0^∞ functions wiht compact support in the following norm $\| \ \|_{s,p,\delta} = \Sigma_{|\alpha|<s} \left| (\sqrt{1+r^2})^{|\alpha|+\delta} D^\alpha f \right|_p$ where $\delta \in R$, $s \geqslant 0$ an integer, $\|_p$ denotes the usual L_p norm on R^n and D^α is $\partial_{|\alpha|}/\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}$ corresponding to a multi-index α. The Laplacian acts isomorphically between $M^p_{s\delta} \rightarrow M^p_{s-2,\delta+2}$, for $p > n/(n-2)$, $1/p + 1/p' = 1$, $-n/p < \delta < -2 + n/p'$ (see Cantor[8]).

Theorem 4: For sufficiently weak fields the gauge exists locally in three or four Euclidean dimensions.

Proof: This is a direct consequence of the above mentioned Laplacian property, for $p > 3$ (in three dimensions) or $p > 2$ (in four dimensions); our thesis follows now from the implicit function theorem, as in Theorem 3.

Note that this result was obtained essentially by Moncrief[4] who proved injectivity of our operator D for small potentials A_i^a (that is all that we used).

The results of Sec. 2 cannot be obtained in the noncompact case so simply. The methods used beforehand do not work since the zeroes of the Frechet derivative D are not isolated in the parameter space h. One must then put a subsidiary condition (correctly stated); if the linearized equations have nonzero solutions at discrete values h, it is possible to get result analogous to our Theorem 1 (it is necessary to prove beforehand the splitting property[4] of D in the spaces $M^p_{s,\delta}$).

6. REMARKS

The Gribov ambiguities include also phenomena of nonlocal nature (e.g., Gribov showed that even the potential $A = 0$ has a nontrivial gauge copy, far from the unity element); they are not explainable in the framework of bifurcation theory. One can study bifurcation from $g \neq 1$; especially interesting seems to be that case when a global g connects $A = 0$ and a nonzero transverse field. Unfortunately, such g is not known explicitly as yet, so the analytical method is inapplicable.

[1] J. M. Singer, Commun. Math. Phys. **60**, 7 (1978); M. Daniel and C. Viallet, PAR LPTHE preprint (March 1979).
[2] V. N. Gribov, Nucl. Phys. B **138**, 1 (1978).
[3] M. S. Berger, in *Bifurcation Theory and Nonlinear Eigenvalue Problems*, edited by S. Antman and J. B. Keller (Benjamin, New York, 1969).
[4] V. Moncrief, J. Math. Phys. **20**, 579 (1979).
[5] M. S. Berger, *Nonlinearity and Functional Analysis* (Academic, New York, 1977).
[6] L. Nirenberg, *Topics in Nonlinear Functional Analysis* (Courant Institute of Mathematical Sciences, Lecture Notes, New York, 1974.
[7] M. Vajnberg and V. Trenogin, *Theory of Branching of Solutions of Nonlinear Equations* (Noordhoff, Groningen, 1974).
[8] M. Cantor, Indiana Univ. Math. J. **24**, 897 (1975).

Lippmann–Schwinger equations for a scalar macroscopic field in an anisotropic stratified medium

Ya. A. Iosilevskii
The Israel National Oceanographic Institute, Tel–Shikmona, P.O.B. 8030, Haifa, Israel[a)]
Electrical Engineering Department, College of Engineering, Drexel University, Philadelphia, Pennsylvania 19104

(Received 10 June 1980; accepted for publication 23 October 1980)

We develop a general approach to solve the transmission problem for a scalar macroscopic field in an anisotropic stratified medium. The method is based on a chainlike set of functional equations of the Lippmann – Schwinger type. A typical example of the fields under consideration is the wave field of a particle in the effective mass tensor approximation.

PACS numbers: 03.40.Kf, 03.65.Nk

I. INTRODUCTION

In this paper, we develop a general approach based on equations of the Lippmann–Schwinger type to solve the problem of the propagation of scalar waves of any kind in an anisotropic plane-stratified medium. In the general case, the system under consideration consists of an arbitrary number of macroscopic crystalline layers of arbitrary thickness. The layers can be different from one another in their crystal symmetries, orientations of the crystallographic axes relative to the separation planes and in their physical characteristics. A typical example of the fields under consideration is the wave field of a particle in the effective mass tensor approximation.

For any dynamic field in a stratified medium (as well as in any other medium, homogeneous or inhomogeneous), two interrelated problems of practical and theoretical interest can be formulated, viz., the problem of the field of an arbitrary "extraneous" source (the "forced" field) and the problem of a free field. The former is reduced to the problem of calculating the Green's function (i.e., the field of a point source) for the given system. This problem was solved in Ref. 1. The problem of a free field is considered in the present paper.

In a stratified medium, two kinds of states of a scalar field are possible. A state of the first kind arises as a result of all conceivable scatterings (i.e., multiple reflections and refractions by all existing interfaces) of an initial plane wave impinging at an arbitrary angle on the nearest separation plane. The field in such a state is different from zero in the whole infinite or in some semi-infinite space and can therefore be called nonlocalized. On the contrary, the field in a state of the second kind is different from zero only in a region which is limited from both ends in the direction normal to the interfaces. Such a state can propagate in the form of a traveling wave only in the directions parallel to the interfaces and can therefore be called localized. In this paper, we consider both kinds of states of a free scalar field.

In Ref. 1, our approach was based on a chainlike set of functional equations of Dyson's type in the mixed coordinate-propagation vector representation. In the present paper, in order to solve the problem of the propagation of a free macroscopic field in a stratified medium as described above, we also formulate a chainlike set of functional equations in the mixed coordinate-propagation vector representation. However, these equations are now equations for the scattering amplitudes, and not for the Green's function, and should therefore be of the Lippmann–Schwinger type.[2] Thus, we generalize the conventional collision theory[2,3] to the problem which is, in fact, classical.

The structure of the set of Lippman–Schwinger equations proves to be similar to that of Dyson's equation. This similarity can be used in two ways. Firstly, the set of Lippmann–Schwinger equations can be solved straightforwardly by essentially the same method as suggested in Ref. 1. Secondly, the scattering amplitudes can be expressed in terms of the corresponding total Green's function. Thus the problem of a free field reduces to the problem of the field of a point source. Naturally, the final results for the scattering amplitudes obtained in these two ways are the same.

The method of the Lippmann–Schwinger equations developed in this paper for a scalar field in an anisotropic stratified media is of importance from both the practical and theoretical points of view. Practically, this is an extremely effective method of finding the eigenstates of the field in an anisotropic stratified system and also of solving the corresponding transmission problem. The most essential advantage of our approach over conventional methods is that our approach enables us to reduce the amount of calculations needed for solving each specific problem to a minimum. This is demonstrated in Sec. VII by a rather complicate example of an arbitrary anisotropic three-layer medium. Theoretically, our method is a generalization of the conventional collision theory[2,3] to the problems which are essentially classical. Also, the scalar field is the simplest field in mathematical physics. The problem of the propagation of the acoustic or electromagnetic field in an anisotropic stratified medium can be solved by the same method but is technically much more complicated (compare, e.g., Refs. 1 and 4). Therefore, aside from the interest which results obtained present in themselves, the considerations of this paper can serve as an introduction to the corresponding theory of the more com-

[a)]Permanent address.

plicated fields just mentioned. This theory will be discussed elsewhere.

II. FORMULATION OF THE PROBLEM

A. The basic equation

Let us consider a free scalar macroscopic field of frequency ω,

$$\psi(\mathbf{r},t) = \psi(\mathbf{r};\omega)e^{-i\omega t}, \quad \omega > 0, \tag{2.1}$$

in an anisotropic stratified medium. The class of the fields and the properties of the media under consideration are assumed to be the same as in Ref. 1.

Regarding a stratified system as a particular case of a spatially inhomogeneous medium, one can write the equation for $\psi(\mathbf{r};\omega)$ in the form

$$[w(\mathbf{r}) + \nabla^i\epsilon^{ik}(\mathbf{r})\nabla^k]\psi(\mathbf{r};\omega) = 0, \tag{2.2}$$

where

$$w(\mathbf{r}) = E - U(\mathbf{r}) + i\gamma(\mathbf{r}), \ \gamma(\mathbf{r}) \to \pm 0, \tag{2.3}$$

$\epsilon^{ik}(\mathbf{r})$ is a material characteristic of the medium ($\epsilon^{ik} = \epsilon^{ki}$), $U(\mathbf{r})$ is a potential, and $\gamma(\mathbf{r})$ is a small phenomenological parameter. The form of E depends on whether the original time-dependent equation for $\psi(\mathbf{r},t)$ is of Schrödinger's type or a conventional classical equation, viz.,

$$E = \begin{cases} \hbar\omega & \text{for Schrödinger's field,} \quad (2.4) \\ \omega^2 & \text{for a classical field.} \quad (2.5) \end{cases}$$

Similarly, γ should be associated with the inverse lifetime of the state of energy E in the quantum case, or with the absorption of the field in the classical case.

We deal with a medium consisting of anisotropic flat layers of arbitrary thicknesses in contact. In the general case, the layers are assumed to be different from one another in their energy-like (or frequencylike) parameters w_μ, material characteristics ϵ_μ^{ik}, and in the orientations of the crystallographic axes with respect to the separation planes. The quantities w_μ and ϵ_μ^{ik} are constants. The integer subscript μ numbers the layers successively.

Choosing the x-axis along a normal to the interfaces, we have[1]

$$w(\mathbf{r}) = \sum_{\mu=0}^{n} s_\mu(x)w_\mu, \quad \epsilon^{ik}(\mathbf{r}) = \sum_{\mu=0}^{n} s_\mu(x)\epsilon_\mu^{ik}, \tag{2.6}$$

where

$$\begin{aligned} &s_\mu(x) = \theta(x - d_{\mu-1}) - \theta(x - d_\mu), \\ &d_{\mu-1} < d_\mu, \ \mu = 1,\cdots,n-1; \\ &s_0(x) = \theta(d_0 - x), \quad s_n(x) = \theta(x - d_{n-1}), \\ &d_0 = -\infty, \ d_n = \infty. \end{aligned} \tag{2.7}$$

We can also write

$$\psi(\mathbf{r};\omega) = \psi(x)e^{i\mathbf{f}_\parallel\cdot\mathbf{r}_\parallel}, \quad \psi(x) \equiv \psi(x;\mathbf{f}_\parallel,\{w_\nu\}), \tag{2.8}$$

where the mark $\parallel$ stands for the orthogonal projection of a vector onto the yz-plane (e.g., onto any one of the interfaces). As a result, Eq. (2.2) becomes

$$\sum_{\nu=0}^{n} [w_\nu s_\nu(x) + \epsilon_\nu^{ik}\partial_x^i s_\nu(x)\partial_x^k]\psi(x) = 0, \tag{2.9}$$

where

$$\partial_x^k = \delta^{k1}\frac{\partial}{\partial x} + if_\parallel^k, \quad \mathbf{f}_\parallel = (0,f^2,f^3) = (0,f^y,f^z), \tag{2.10}$$

$$w_\mu = E_\mu + i\gamma_\mu, \quad E_\mu = E - U_\mu, \quad \gamma_\mu \to \pm 0. \tag{2.11}$$

In this paper, we retain the main notation of Ref. 1. In particular, $x = d_\mu$ is the separation plane between the $(\mu-1)$th and the μth layers, $s_\mu(x)$ is the shape function of the μth layer, $\theta(x) = \frac{1}{2}(1 + \operatorname{sgn} x)$ is the unit step function, n is the total number of interfaces. The total number of layers is equal to $n+1$. According to Eqs. (2.7), the zeroth ($\mu = 0$) and the last ($\mu = n$) layers are assumed to be semi-infinite spaces $x < d_0$ and $x > d_{n-1}$, respectively. In Eqs. (2.8) and (2.9), the subscript ν is used instead of μ.

From the complex conjugate of Eq. (2.9) and from the definitition of $\psi(x)$ [see Eqs. (2.8)], it follows that

$$\psi^*(x;\mathbf{f}_\parallel,\{w_\nu\}) = \psi(x;-\mathbf{f}_\parallel,\{w_\nu^*\}) = \psi(-x;\mathbf{f}_\parallel,\{w_\nu^*\}). \tag{2.12}$$

In the case of Schrödinger's field, the state ψ^* is time-reversed to ψ.[3] Hence, $\psi(x;\mathbf{f}_\parallel,\{w_\nu^*\})$ should, according to Eq. (2.12), be regarded as a state which is time-reversed to $\psi(x;-\mathbf{f}_\parallel,\{w_\nu\})$. By analogy, we retain this definition also for a classical field. It should, however, be remembered that the original time-dependent equation for a classical field is of the second order in $\partial/\partial t$ and therefore this definition is purely formal.

B. Waves in a homogeneous medium

If $\epsilon^{ik}(\mathbf{r}) = \epsilon_\mu^{ik}$ and $U(\mathbf{r}) = U_\mu$ in the whole infinite space, Eq. (2.1) becomes

$$(\mathscr{E}_\mu + \epsilon_\mu^{ik}\nabla^i\nabla^k)\phi_\mu(\mathbf{r}) = 0, \tag{2.13}$$

where we have substituted $\mathscr{E}_\mu$ for w_μ, and $\phi_\mu(\mathbf{r})$ for $\psi(\mathbf{r};\omega)$. Equation (2.13) gives the eigenfunctions and eigenvalues of the field in an infinite medium of the μth kind, viz.,

$$\phi_\mu(\mathbf{r}) = a_\mu e^{i\mathbf{f}\cdot\mathbf{r}}, \quad \mathscr{E}_\mu = \mathscr{E}_\mu(\mathbf{f}) = \epsilon_\mu^{ik}f^if^k > 0, \tag{2.14}$$

where $\mathbf{f}$ is a propagation vector, and a_μ is a normalization coefficient. The quadratic form $\mathscr{E}_\mu(\mathbf{f})$ is assumed to be positive definite.

The eigenfrequencies $\omega = \omega_\mu(\mathbf{f})$ can be found by combining the relation

$$E = \mathscr{E}_\mu(\mathbf{f}) + U_\mu \tag{2.15}$$

with Eq. (2.4) or (2.5). At the same time, the group velocity of a wave packet,

$$v_{\mu g}^i(\mathbf{f}) = \partial\omega_\mu(\mathbf{f})/\partial f^i, \tag{2.16}$$

is expressed in terms of the quantity

$$v_\mu^i(\mathbf{f}) = \partial\mathscr{E}_\mu(\mathbf{f})/\partial f^i = 2\epsilon_\mu^{ik}f^k \tag{2.17}$$

by a simple relation. This is

$$v_{\mu g}^i(\mathbf{f}) = \begin{cases} \hbar^{-1}v_\mu^i(\mathbf{f}) & \text{for Schrödinger's field,} \quad (2.18) \\ \frac{1}{2}E^{-1/2}v_\mu^i(\mathbf{f}) & \text{for a classical field,} \quad (2.19) \end{cases}$$

where the square root is positive ($E = \omega^2 > 0$).

C. The stationary scattering problem

Our purpose is to represent the stationary state of the field in a stratified medium, which is formed as a result of all possible multiple scatterings (i.e., reflections and refractions) by all existing interfaces of a plane wave coming from infinity. We assume that such a wave of frequency ω comes from the semi-infinite space $x < d_0$ and strikes the interface $x = d_0$ at an arbitrary angle. We can therefore write

$$\psi(x) = \phi_0(x - d_0) + \chi_0(x), \;\; x < d_0, \tag{2.20}$$

where $\phi_0(x)$ is the incident wave in the $(x, \mathbf{f}_\parallel)$ representation $\chi_0(x)$ is associated with the wave reflected into the interior of the region $x < d_0$.

In accordance with Eqs. (2.8), (2.14), and (2.15), the incident wave is given by the relations

$$\phi_0(x) = a_0 e^{if_*^x x}, \quad E = \mathscr{E}_0(\mathbf{f}_*) + U_0, \quad f_*^i = \delta^{i1} f_*^x + f_\parallel^i, \tag{2.21}$$

where $\mathbf{f}_*$ is its propagation vector, and E is connected with the frequency of the wave by Eq. (2.4) or (2.5). The magnitude of the normalization coefficient $a_0 = a_0(\mathbf{f}_*)$ is unimportant in the problem under consideration, because we are interested only in the ratio of the wave amplitude in each layer to a_0. The constant phase factor $\exp(-if_*^x d_0)$, which appears in the expression for $\phi_0(x - d_0)$ [see Eq. (2.20)], is introduced for convenience and can, in principle, be included in a_0.

In addition to the characteristics of the "bare" plane wave, which are given by Eqs. (2.21), we introduce the sign index of the x-component of its group velocity, viz.,

$$s_* \equiv \mathrm{sgn} v_*^x = \mathrm{sgn} v_{0g}^x(\mathbf{f}_*), \tag{2.22}$$

where

$$v_*^i \equiv v_0^i(\mathbf{f}_*) = \partial \mathscr{E}_0(\mathbf{f}_*)/\partial f_*^i = 2\epsilon_0^{ik} f_*^k, \tag{2.23}$$

in accordance with Eq. (2.17). The fact that the signs of $v_0^x(\mathbf{f}_*)$ and $v_{0g}^x(\mathbf{f}_*)$ are the same follows from Eqs. (2.18) and (2.19).

It is clear that

$$s_* = +1 \quad \text{for forward propagation in time,} \tag{2.24}$$

$$s_* = -1 \quad \text{for backward propagation in time.} \tag{2.25}$$

The case of $s_* = +1$ corresponds to the stationary scattering problem as was formulated at the beginning of this subsection. If $s_* = -1$, we arrive at the stationary state which should be regarded as a time-reversed state, in accordance with the definition of Sec. IIA (compare with conventional collision theory[3]). In this paper, we consider both kinds of stationary solutions of the problem.

It should be noted that, generally, there is no single-valued correlation between the signs of $v_0^x(\mathbf{f}_*)$ and f_*^x. At each given s_*, the component f_*^x can, in principle, have either sign. Only in the case of an isotropic medium, $\mathrm{sgn} f_*^x = s_*$. In fact, the sign of $v_0^x(\mathbf{f}_*)$, but not of f_*^x, is important in the problem under consideration.

D. Boundary conditions

At each separation plane, say $x = d_\mu$, the quantity $\psi(x)$ satisfies the boundary conditions

$$\psi(d_\mu - 0) = \psi(d_\mu + 0) \equiv \psi(d_\mu), \tag{2.26}$$

$$\Pi(d_\mu - 0) = \Pi(d_\mu + 0) \equiv \Pi(d_\mu), \tag{2.27}$$

where, by definition,

$$\Pi(x) = -\epsilon_\mu^{xk} \partial_x^k \psi(x), \quad d_{\mu-1} < x < d_\mu, \quad \mu = 0,1,\ldots,n. \tag{2.28}$$

Equation (2.26) means that, at any separation plane, the field under consideration is assumed to be a continuous function of the position vector. Integrating both sides of Eq. (2.9) with respect to x between $d_\mu - \delta$ and $d_\mu + \delta$, and taking into account Eq. (2.26) and the definition of $\Pi(x)$, Eq. (2.28), we arrive at Eq. (2.27) as $\delta \to +0$.

If the nth region is impenetrable to the given field, we have

$$\psi(x) = 0, \quad x > d_{n-1}. \tag{2.29}$$

As a consequence, Eq. (2.26) for $\mu = n$ takes the form

$$\psi(d_{n-1} - 0) \equiv \psi(d_n) = 0, \tag{2.30}$$

while Eq. (2.27) for $\mu = n$ becomes ineffective. Thus, we arrive at the scattering problem for a stratified semi-infinite space $x < d_{n-1}$ with the rigid boundary condition at $x = d_{n-1}$ as given by Eq. (2.30).

E. Equations of motion

Making use of Eqs. (2.7), and (2.27), and (2.28), and taking into account that $d\theta(x)/dx = \delta(x)$, one can rewrite Eq. (2.9) in the form of the set of differential equations

$$(w_\mu + \epsilon_\mu^{ik} \partial_x^i \partial_x^k) \psi(x) = 0, \quad d_{\mu-1} < x < d_\mu, \quad \mu = 0,1,\ldots,n. \tag{2.31}$$

Since

$$[\mathscr{E}_0(\mathbf{f}_*) + \epsilon_0^{ik} \partial_x^i \partial_x^k] \phi_0(x) = 0, \tag{2.32}$$

substitution of Eq. (2.20) into Eq. (2.31) with $\mu = 0$ gives

$$(w_0 + \epsilon_0^{ik} \partial_x^i \partial_x^k) \chi_0 = 0, \;\; x < d_0, \;\; \mu = 0, \tag{2.33}$$

where

$$w_0 = E_0 + i\gamma_0 = \mathscr{E}_0(\mathbf{f}_*) + i\gamma_0 \tag{2.34}$$

[for Eqs. (2.32) and (2.34), see Eqs. (2.10), (2.11), and (2.13)–(2.15)]. Equations (2.31) with $\mu = 1,2,\ldots,n$ and Eqs. (2.33) form the complete set of equations of motion for the problem under consideration.

F. The quantum-mechanical interpretation

Writing

$$\epsilon_\mu^{ik} = \tfrac{1}{2}\hbar^2 (m_\mu{}^{-1})^{ik}, \tag{2.35}$$

we can regard Eq. (2.9) as a time-independent equation of Schrödinger's type for quasiparticles in the effective mass tensor approximation. The original time-dependent equation for $\psi(\mathbf{r},t)$ can obviously be written

$$i\hbar \partial\psi/\partial t = \hat{H}\psi, \tag{2.36}$$

where

$$\hat{H} = \nabla^i \epsilon^{ik}(\mathbf{r})\nabla^k + U(\mathbf{r}), \quad U(\mathbf{r}) = \sum_{\mu=0}^{n} s_\mu(x) U_\mu \qquad (2.37)$$

[for $\epsilon^{ik}(\mathbf{r})$ and $s_\mu(x)$, see Eqs. (2.6) and (2.7)].

Making use of Eqs. (2.36) and (2.37), we arrive in the usual way (see e.g., Ref. 5, pp. 56, 57) at the continuity equation

$$\partial|\psi|^2/\partial t = \nabla^i j^i, \qquad (2.38)$$

where the current density is written as

$$j^i = (i/\hbar)\epsilon^{ik}(\psi \nabla^k \psi^* - \psi^* \nabla^k \psi) \quad [\epsilon^{ik} = \epsilon^{ik}(\mathbf{r}), \ \psi = \psi(\mathbf{r},t)]. \qquad (2.39)$$

An asterisk stands for complex conjugation. In developing Eq. (2.39), we have taken into account that $\hat{H}^* = \hat{H}$ and $\epsilon^{ik} = \epsilon^{ki}$.

Combination of Eq. (2.39) with Eqs. (2.1), (2.6), and (2.8) gives

$$\begin{aligned} j^i &= j^i(x;\mathbf{f}_\parallel,\{w_\nu\}) \\ &= \frac{i}{\hbar}\left[\sum_{\nu=0}^{n} s_\nu(x)\epsilon_\nu^{ik}\right][\psi(x)\partial_x^{k*}\psi^*(x) - \psi^*(x)\partial_x^k \psi(x)], \end{aligned} \qquad (2.40)$$

where ∂_x^k is defined by Eq. (2.10). Owing to Eqs. (2.26)–(2.28), $j^i = j^i(x)$ proves to be continuous at each interface, i.e.,

$$j^i(d_\mu - 0) = j^i(d_\mu + 0), \quad \mu = 0,1,\dots,n-1, \qquad (2.41)$$

as it should be.

Substitution of ϵ_μ^{ik} for ϵ^{ik} and $\phi_\mu(\mathbf{r})$ for ψ [see Eqs. (2.13) and (2.14)] into Eq. (2.39) gives the current density of quasiparticles in an infinite medium of the μth kind,

$$j_\mu^i = |a_\mu|^2 v_{\mu g}^i(\mathbf{f}), \qquad (2.42)$$

where $v_{\mu g}^i(\mathbf{f})$ is the group velocity as defined by Eqs. (2.17) and (2.18). Setting $\mu = 0$ and $\mathbf{f} = \mathbf{f}_*$ in Eq. (2.42), we obtain the current density of the incident particles or, more generally, the "bare" current density. Making use of the relation

$$\begin{aligned} |\mathbf{v}_{\mu g}(\mathbf{f})|^2 &\equiv v_{\mu g}^i(\mathbf{f}) v_{\mu g}^i(\mathbf{f}) \\ &= 4\hbar^{-2}(\epsilon_\mu^2)^{ik} f^i f^k = \hbar^2 (m_\mu^{-2})^{ik} f^i f^k, \end{aligned} \qquad (2.43)$$

we can normalize j_μ^i to any desired intensity.

III. THE RETARDED AND THE ADVANCED "STANDARD" GREEN'S FUNCTIONS

Like conventional collision theory,[3] our formalism is based on the use of the properly defined *retarded* and *advanced* "standard" Green's functions in the $(x,\mathbf{f}_\parallel)$-representation. Below, we introduce these functions and discuss their main properties.

The "standard" (unperturbed) Green's function for the μth medium in the $(x,\mathbf{f}_\parallel)$-representation, $D_\mu(x) \equiv D_\mu(x;\mathbf{f}_\parallel,w_\mu)$, is, by definition, the solution of the equation[1]

$$(w_\mu + \epsilon_\mu^{ik}\partial_x^i \partial_x^k) D_\mu(x) = \delta(x), \quad -\infty < x < \infty, \qquad (3.1)$$

where ∂_x^i is given by Eq. (2.10), and w_μ is an arbitrary complex parameter. It was shown in Ref. 1 that

$$D_\mu(x) = -(2\beta_\mu)^{-1} e_\mu(x), \qquad (3.2)$$

$$T_\mu(x) = -\epsilon_\mu^{xk}\partial_x^k D_\mu(x) = -\tfrac{1}{2} s_x e_\mu(x), \qquad (3.3)$$

where

$$e_\mu(x) = \exp(i f_\mu^{(s_x)x} x), \quad s_x = \operatorname{sgn} x, \qquad (3.4)$$

and $f^x = f_\mu^{(s_x)x} = f_\mu^{(s_x)}(\mathbf{f}_\parallel,w_\mu)$ is the solution of the equation

$$\mathscr{E}_\mu(\mathbf{f}) \equiv \epsilon_\mu^{ik} f^i f^k = w_\mu \qquad (3.5)$$

with respect to f^x, which satisfies the condition

$$\operatorname{sgn} \operatorname{Im} f_\mu^{(s_x)x} = s_x. \qquad (3.6)$$

In addition to $D_\mu(x)$ itself, the quantity $T_\mu(x)$ as defined by Eq. (3.3) is of fundamental importance in our formalism.

Equation (3.5) is quadratic in f^x and has therefore two solutions. These can be written

$$f^x = f_\mu^{(s)x} = f_\mu^{(s)x}(\mathbf{f}_\parallel,w_\mu) = p_\mu + i s q_\mu, \quad s = \pm 1, \qquad (3.7)$$

where

$$\begin{aligned} &p_\mu = -(\epsilon_\mu^{xx})^{-1}\epsilon_\mu^{xk} f_\parallel^k, \quad q_\mu = (\epsilon_\mu^{xx})^{-1}\beta_\mu, \\ &\beta_\mu = (\epsilon_\mu^{xx})^{1/2}(u_\mu - w_\mu)^{\frac{1}{2}}, \end{aligned} \qquad (3.8)$$

$$\begin{aligned} &u_\mu = \lambda_{\mu\parallel}^{ik} f_\parallel^i f_\parallel^k > 0, \quad \lambda_{\mu\parallel}^{ik} = \epsilon_\mu^{ik} - (\epsilon_\mu^{xx})^{-1}\epsilon_\mu^{xi}\epsilon_\mu^{xk}, \\ &\{i,k\} = \{y,z\}, \end{aligned} \qquad (3.9)$$

$$\operatorname{sgn} \operatorname{Re} q_\mu = \operatorname{sgn} \operatorname{Re} \beta_\mu = \operatorname{sgn}(\epsilon_\mu^{xx})^{1/2} = +1. \qquad (3.10)$$

Inserting Eq. (3.7) into Eq. (3.6) and taking into account Eq. (3.10), we find that $s = s_x$.

The quantity u_μ has a simple physical meaning. Let us find the minimum of $\mathscr{E}_\mu(\mathbf{f})$ as a function of $f^x(\mathbf{f}_\parallel = \text{const})$. From the equation $v_\mu^x(\mathbf{f}) = 0$ [see Eq. (2.17)], it follows that

$$\begin{aligned} &f^x = f_{\mu\min}^x = -(\epsilon_\mu^{xx})^{-1}\epsilon_\mu^{xk} f_\parallel^k = p_\mu, \\ &f_{\mu\min}^i = \delta^{i1} f_{\mu\min}^x + f_\parallel^i = \delta^{i1} p_\mu + f_\parallel^i. \end{aligned} \qquad (3.11)$$

After simple reduction, we find that the desired quantity can be written

$$\mathscr{E}_{\mu\min}(\mathbf{f}_\parallel) \equiv \mathscr{E}_\mu(\mathbf{f}_{\mu\min}) = u_\mu, \qquad (3.12)$$

where u_μ is exactly that as defined by Eqs. (3.9). Since, by assumption, $\mathscr{E}_\mu(\mathbf{f})$ is positive for any $\mathbf{f}$, the quadratic form $u_\mu = u_\mu(\mathbf{f}_\parallel)$ is also positive definite.

In all further considerations, we assume that $w_\mu = E_\mu + i\gamma_\mu$, in accordance with Eqs. (2.11). If $\gamma_\mu \to +0$, the quantity $D_\mu(x)$ can be called the retarded Green's function. If $\gamma_\mu \to -0$, we deal with the advanced Green's function.

In order to find the retarded and the advanced Green's functions, we write

$$\begin{aligned} &f^x = \bar{f}^x + i\eta, \\ &f^i = \delta^{i1} f^x + f_\parallel^i, \quad \bar{f}^i = \delta^{i1}\bar{f}^x + f_\parallel^i, \end{aligned} \qquad (3.13)$$

where

$$\begin{aligned} \bar{f}^x = \bar{f}_\mu^{(s)x} &= \lim_{\gamma_\mu \to \pm 0} f_\mu^{(s)x}(\mathbf{f}_\parallel, E_\mu + i\gamma_\mu) \\ &= p_\mu + i s(\epsilon_\mu^{xx})^{-1}\bar{\beta}_\mu. \end{aligned} \qquad (3.14)$$

We also write

$$\beta_\mu = \bar{\beta}_\mu(1 + i\delta_\mu), \qquad (3.15)$$

where

$$\bar{\beta}_\mu = \lim_{\gamma_\mu \to \pm 0} (\beta_\mu)_{w_\mu = E_\mu + i\gamma_\mu}$$

$$= \begin{cases} (\epsilon_\mu^{xx})^{1/2}\sqrt{u_\mu - E_\mu} \quad \text{if } E_\mu < u_\mu, & (3.16) \\ -is_{\gamma_\mu}(\epsilon_\mu^{xx})^{1/2}\sqrt{E_\mu - u_\mu} \quad \text{if } E_\mu > u_\mu, & (3.17) \end{cases}$$

$$\delta_\mu = [2(E_\mu - u_\mu)]^{-1}\gamma_\mu, \quad s_{\gamma_\mu} = \text{sgn}\gamma_\mu, \tag{3.18}$$

with the square roots being positive. The sign factors on the right-hand sides of Eqs. (3.16) and (3.7) are chosen so that β_μ, Eq. (3.15), satisfies the condition (3.10).

The quantity η can be found if we insert Eq. (3.15) into Eq. (3.7) $[q_\mu = (\epsilon_\mu^{xx})^{-1}\beta_\mu$, see Eqs. (3.8)] and compare the resulting expression for f^x with that given by the first of Eqs. (3.13). Thus we obtain

$$\eta = \eta_\mu^{(s)} = -is\gamma_\mu/2\bar{\beta}_\mu = \gamma_\mu/v_\mu^x(\bar{\mathbf{f}}_\mu^{(s)}). \tag{3.19}$$

In writing the final result in Eq. (3.19), we have made use of the relation

$$v_\mu^x(\bar{\mathbf{f}}_\mu^{(s)}) = 2(\epsilon_\mu^{xx}\bar{f}_\mu^{(s)x} + \epsilon_\mu^{xk}f_\parallel^{xk}) = 2is\bar{\beta}_\mu, \tag{3.20}$$

which follows from Eq. (2.17), in view of the definitions of $\bar{f}_\mu^{(s)i}$ [see Eqs. (3.14) and (3.15)] and p_μ [see Eqs. (3.8)].

Equations (3.14) and (3.19) can alternatively be derived if we insert f^i as defined by Eqs. (3.13) into Eq. (3.5) and then write Eq. (3.5) in the linear approximation in η as

$$\mathscr{E}_\mu(\bar{\mathbf{f}}) + i\eta v_\mu^x(\bar{\mathbf{f}}) = E_\mu + i\gamma_\mu$$

[for $v_\mu^x(\bar{\mathbf{f}})$, see Eq. (2.17)]. Thus we recover Eq. (3.19). Also, we get $\mathscr{E}_\mu(\bar{\mathbf{f}}) = E_\mu$, which is equivalent to Eq. (3.14).

In order to identify $f^x = f_\mu^{(s)x}$ with $f_\mu^{(s_x)x}$ one should set $s = s_x$ in Eqs. (3.13)–(3.17), (3.19), and (3.20).

From Eqs. (3.16) and (3.17), it follows that there is a difference between the retarded and the advanced Green's functions only if $E_\mu > u_\mu$. In this case, Eqs. (3.14) and (3.20) become

$$\bar{f}^x = \bar{f}_\mu^{(s)x} \equiv \bar{f}_\mu^{[\sigma_\mu]x} = p_\mu + \sigma_\mu(\epsilon_\mu^{xx})^{-1}|\beta_\mu|,$$
$$v_\mu^x(\bar{\mathbf{f}}) = 2\sigma_\mu|\beta_\mu|, \quad E_\mu > u_\mu, \tag{3.21}$$

where

$$\sigma_\mu = ss_{\gamma_\mu}, \quad s, s_{\gamma_\mu} = \pm 1. \tag{3.22}$$

Since $\bar{f}^x$ and $v_\mu^x(\bar{\mathbf{f}}_\mu)$ so obtained are real, combination of Eq. (3.6) with Eqs. (3.13) and (3.19) gives

$$\text{sgn}v_\mu^x(\bar{\mathbf{f}}) = s_x s_{\gamma_\mu} \quad \text{if } E_\mu > u_\mu \tag{3.23}$$

[for s_{γ_μ}, see Eqs. (3.18)]. This relation justifies the above definition of the retarded and the advanced Green's functions: If $s_{\gamma_\mu} = +1$, the quantity $D_\mu(x)$ is associated with waves coming out from the plane $x = 0$ ($\text{sgn}v_\mu^x = \text{sgn}x$); if $s_{\gamma_\mu} = -1$, the waves are incoming ($\text{sgn}v_\mu^x = -\text{sgn}x$).

From the first of Eqs. (3.21), it follows that

$$\bar{f}_\mu^{[-\sigma]x} = -\bar{f}_\mu^{[\sigma]x} + 2p_\mu = \bar{f}_\mu^{[\sigma]x} - 2(\epsilon_\mu^{xx})^{-1}\epsilon_\mu^{xk}\bar{f}_\mu^{[\sigma]k},$$
$$\sigma = \pm 1. \tag{3.24}$$

In writing this result, we have taken into account the definition of p_μ as given in Eqs. (3.8).

In tensor notation independent of the system of coordinates chosen, Eq. (3.24) can be written

$$\bar{f}_\mu^{[\sigma]i} = \tilde{P}_\mu^{ik}\bar{f}_\mu^{[\sigma]k} = \bar{f}_\mu^{[\sigma]i} - 2n^i(\mathbf{v}_\mu\cdot\bar{\mathbf{f}}_\mu^{[\sigma]}), \tag{3.25}$$

where

$$P_\mu^{ik} = \delta^{ik} - 2v_\mu^i n^k,$$
$$v_\mu^i = (\epsilon_\mu^{ab}n^a n^b)^{-1}\epsilon_\mu^{ik}n^k, \tag{3.26}$$

and $\mathbf{n}$ is a unit vector in the x-direction. A tilde stands for transposition.

According to Refs. 1 and 6, P_μ^{ik} as defined by Eqs. (3.26) is the operator of the generalized image in the plane $\mathbf{r}\cdot\mathbf{n} = 0$ which is the boundary of an anisotropic half-space ($\mathbf{r}\cdot\mathbf{n} < 0$ or $\mathbf{r}\cdot\mathbf{n} > 0$) of the μth kind. In other words, the generalized image $\mathbf{r}_{\text{img}}$ of a space point $\mathbf{r}$ belonging to the half-space is written as

$$r_{\text{img}}^i = P_\mu^{ij}r^j = r^i - 2v_\mu^i(\mathbf{r}\cdot\mathbf{n}). \tag{3.27}$$

If now one defines the generalized image $\mathbf{f}_{\text{img}}$ of a propagation vector $\mathbf{f}$ by requiring that $\mathbf{r}\cdot\mathbf{f} = \mathbf{r}_{\text{img}}\cdot\mathbf{f}_{\text{img}}$, the result is

$$f_{\text{img}}^i = \tilde{P}_\mu^{ij}f^j = f^i - 2n^i(\mathbf{v}_\mu\cdot\mathbf{f}). \tag{3.28}$$

This can readily be verified by means of the relation

$$P_\mu^{ij}P_\mu^{jk} = P_\mu^{ij}\tilde{P}^{kj} = \delta^{ik}. \tag{3.29}$$

Calculations of $v_\mu^i(\mathbf{f}_{\text{img}})$ by Eqs. (2.17) and (2.38) give

$$v_\mu^i(\mathbf{f}_{\text{img}}) \equiv 2\epsilon_\mu^{ij}f_{\text{img}}^j = 2\epsilon_\mu^{ij}\tilde{P}_\mu^{jk}f^k = 2P_\mu^{ij}\epsilon_\mu^{jk}f^k$$
$$= P_\mu^{ij}v_\mu^j(\mathbf{f}) = v_\mu^i(\mathbf{f}) - 2v_\mu^i(\mathbf{v}_\mu(\mathbf{f})\cdot\mathbf{n}). \tag{3.30}$$

In developing this result, we have made use of the relation

$$\epsilon_\mu^{jk}\tilde{P}_\mu^{jk} = P_\mu^{ij}\epsilon_\mu^{jk} = \epsilon_\mu^{ik} - 2\epsilon_\mu^{ab}n^a n^b v_\mu^i v_\mu^k, \tag{3.31}$$

which follows from Eqs. (3.26). Thus, under the generalized image transformation, the quantity $\mathbf{v}_\mu(\mathbf{f})$ and hence, the group velocity [see Eqs. (2.18) and (2.19)] are transformed like a position vector [compare Eqs. (3.27) and (3.30)].

Let us multiply Eq. (3.28) by v_μ^i and Eq. (3.30) by n^i. Since $\mathbf{v}_\mu\cdot\mathbf{n} = 1$, this gives

$$\mathbf{v}_\mu\cdot\mathbf{f}_{\text{img}} = -\mathbf{v}_\mu\cdot\mathbf{f}, \quad \mathbf{n}\cdot\mathbf{v}_\mu(\mathbf{f}_{\text{img}}) = -\mathbf{n}\cdot\mathbf{v}_\mu(\mathbf{f}). \tag{3.32}$$

As a consequence, Eqs. (3.28) and (3.30) are invariant under the permutations $\mathbf{f}_{\text{img}} \rightleftarrows \mathbf{f}$ and $\mathbf{v}_\mu(\mathbf{f}_{\text{img}}) \rightleftarrows \mathbf{v}_\mu(\mathbf{f})$, respectively, in agreement with Eq. (3.29). For an isotropic medium, we have

$$\epsilon_\mu^{ik} = \epsilon_\mu\delta^{ik}, \quad v_\mu^i = n^i \tag{3.33}$$

and therefore $f_{\text{img}}^x = -f^x$, as it should be.

Comparing Eq. (3.25) with Eq. (3.28), we conclude that the vector $\bar{f}_\mu^{[\sigma]i}$ can be regarded as the generalized image of the vector $\bar{f}_\mu^{[-\sigma]i}$ in the plane $\mathbf{r}\cdot\mathbf{n} = 0$. The physical meaning of the generalized image is as follows (see also Sec. V A): If $\mathbf{f} = \bar{\mathbf{f}}_\mu^{[\sigma]}$ is the propagation vector of a plane wave which propagates in an anisotropic half-space, $\mathbf{r}\cdot\mathbf{n} < 0$ or $\mathbf{r}\cdot\mathbf{n} > 0$, of the μth kind and strikes the plane $\mathbf{r}\cdot\mathbf{n} = x = 0$ separating this half-space from a medium of a different kind, then $\mathbf{f}_{\text{img}} = \bar{\mathbf{f}}_\mu^{[-\sigma]}$ is the propagation vector of the reflected wave. The x-components of group velocities of the incident and reflected waves are given by the second of Eqs. (3.21), which is in agreement with the second of Eqs. (3.32) ($n^i = \delta^{ix}$).

We recall that the "bare" plane wave is given in the region $x < d_0$, which corresponds to $\mu = 0$. Without loss of generality, we can assume that $d_0 = 0$. In view of Eq. (2.34), Eq. (3.17) with $\mu = 0$ becomes

$$\bar{\beta}_0 = -is_{\gamma_0}|\epsilon_0^{xk} f_*^k| = -is_{\gamma_0} s_* \epsilon_0^{xk} f_*^k = -\tfrac{1}{2} i s_{\gamma_0} s_* v_*^x. \tag{3.34}$$

In developing the final result in Eq. (3.34), we have taken into account Eqs. (2.22) and (2.23).

Substitution of Eq. (3.34) together with the expression for p_0 [see Eqs. (3.8)] into Eq. (3.21) with $\mu = 0$ and $s = s_\mu$ gives

$$\bar{f}_0^{(s_x)x} \equiv \bar{f}_0^{[s_x s_{\gamma_0}]x} \equiv \bar{f}_0^{\{\bar{s}\}x} = \bar{f}_*^x - (1-\bar{s})(\epsilon_0^{xx})^{-1}\epsilon_0^{xk} f_*^k, \tag{3.35}$$

where

$$\bar{s} = s_x s_{\gamma_0} s_* \tag{3.36}$$

and f_*^k is defined in Eqs. (2.21). Equation (3.35) can be written in covariant form as

$$\bar{f}_0^{(s_x)i} \equiv \bar{f}^{[s_x s_{\gamma_0}]i} \equiv \bar{f}_0^{\{\bar{s}\}i} = f_*^i - (1-\bar{s})(\mathbf{v}_0 \cdot \mathbf{f}_*) n^i \tag{3.37}$$

or, equivalently,

$$\bar{f}_0^{\{-1\}i} = \bar{P}_0^{ik} f_*^k = f_*^i - 2n^i(\mathbf{v}_0 \cdot \mathbf{f}_*), \bar{f}_0^{\{+1\}i} = f_*^i, \tag{3.38}$$

where P_0^{ik} and v_0^i are defined by Eqs. (3.26). Thus, $\bar{f}_0^{\{-1\}i}$ is the generalized image of f_*^i in the plane $\mathbf{r} \cdot \mathbf{n} = x = 0$.

Setting $f_{\rm img}^i = \bar{f}_\mu^{\{-1\}i}, f^i = f_*^i$, and $\mu = 0$ in the second of Eqs. (3.32), and taking into account that $n^i = \delta^{ix}$, we have

$$v_0^x(\bar{\mathbf{f}}_0^{\{-1\}}) = -v_0^x(\mathbf{f}_*) = -v_*^x. \tag{3.39}$$

As a consequence, Eq. (3.23) with $\bar{\mathbf{f}} = \bar{\mathbf{f}}_0^{\{-1\}}$ and $\mu = 0$ becomes $s_x s_{\gamma_0} = -s_*$, in agreement with the fact that $\bar{s} = -1$ [see Eqs. (2.22), (2.23), (3.35), and (3.36)]. Since $s_x = -1 (x < d_0 = 0)$, the relation obtained reduces to $s_{\gamma_0} = s_*$.

This result is a particular case of the general condition

$$s_{\gamma_\mu} = s_*, \quad \mu = 0,1,\ldots,n, \tag{3.40}$$

which, evidently, must be imposed on s_{γ_μ} [see the criteria (2.24), (2.25), and the discussion following Eq. (3.23)].

According to Eqs. (2.24), (2.25), and (3.40), one should use the retarded Green's functions when constructing the usual causal solutions of the stationary scattering problem, and the advanced Green's functions in constructing the time-reserved states. This is in agreement with conventional collision theory.[3]

IV. LIPPMAN–SCHWINGER EQUATIONS

In analogy with what was done in deriving the set of Dyson's equations in Ref. 1, we multiply the equation for the scattering amplitude in the μth layer [i.e., Eq. (2.33) if $\mu = 0$, and Eq. (2.31) if $\mu = 1,2,\ldots,n$] by the corresponding "standard" Green's function, $D_\mu(x_1 - x)$, where x_1 is an arbitrary fixed point on the x-axis. Then we integrate both sides of the resulting expression with respect to x between $d_{\mu-1} + \delta$ and $d_\mu - \delta$, respectively, thus obtaining

$$\int_{-\infty}^{d_0 - \delta} D_0(x_1 - x)(w_0 + \epsilon_0^{ik}\partial_x^i\partial_x^k)\chi_0(x)dx = 0, \ \mu = 0,$$

$$\int_{d_{\mu-1}+\delta}^{d_\mu - \delta} D_\mu(x_1 - x)(w_\mu + \epsilon_\mu^{ik}\partial_x^i\partial_x^k)\psi(x)dx = 0, \quad \mu = 1,2,\ldots,n,$$

where δ is a positive infinitesimal ($d_{-1} = -\infty,\ d_n = \infty$). In these equations, we integrate by parts the terms containing derivatives with respect to x [see Eqs. (2.10)]. Lastly, making use of Eq. (3.1) and the notation given by Eqs. (2.28) and (3.3), we arrive at the equations

$$s_0^{(\delta)}(x)\chi_0(x) - D_0(x - d_0 + \delta)\Gamma_0(d_0 - \delta) - T_0(x - d_0 + \delta)\chi_0(d_0 - \delta) = 0, \quad \mu = 0, \tag{4.1}$$

$$s_\mu^{(\delta)}(x)\psi(x) + D_\mu(x - d_{\mu-1} - \delta)\Pi(d_{\mu-1} + \delta) - D_\mu(x - d_\mu + \delta)\Pi(d_\mu - \delta) + T_\mu(x - d_{\mu-1} - \delta) \times \psi(d_{\mu-1} + \delta) - T_\mu(x - d_\mu + \delta)\psi(d_\mu - \delta) = 0, \quad \mu = 1,2,\ldots,n-1, \tag{4.2}$$

$$s_n^{(\delta)}(x)\psi(x) + D_n(x - d_{n-1} - \delta)\Pi(d_{n-1} + \delta) + T_n(x - d_{n-1} - \delta)\psi(d_{n-1} + \delta) = 0, \quad \mu = n, \tag{4.3}$$

where

$$\Gamma_0(x) = -\epsilon_0^{xk}\partial_x^k\chi_0(x) = \Pi(x) - \Lambda_0(x - d_0), \quad x < d_0, \tag{4.4}$$

$$\Lambda_0(x) = -\epsilon_0^{xk}\partial_x^k\phi_0(x) = -i\epsilon_0^{xk}f_*^k\phi_0(x) = -\tfrac{1}{2}iv_*^x\phi_0(x), \tag{4.5}$$

$$s_\mu^{(\delta)}(x) = \theta(x - d_{\mu-1} - \delta) - \theta(x - d_\mu + \delta), \quad \mu = 1,2,\ldots,n-1,$$
$$s_0^{(\delta)}(x) = \theta(d_0 - x - \delta), \quad s_n^{(\delta)}(x) = \theta(x - d_{n-1} - \delta). \tag{4.6}$$

In writing Eqs. (4.1)–(4.3), we have substituted x for x_1. In Eqs. (4.1) and (4.3), we have also taken into account that $D_\mu(\pm\infty) = 0$ and $T_\mu(\pm\infty) = 0$ ($d_{-1} = -\infty,\ d_n = \infty$). The final expression for $\Gamma_0(x)$ in Eq. (4.4) is apparent from Eqs. (2.20) and (2.28).

Inserting the expressions

$$\chi_0(d_0 - \delta) = \psi(d_0 - \delta) - \phi_0(-\delta),$$
$$\Gamma_0(d_0 - \delta) = \Pi(d_0 - \delta) - \Lambda_0(-\delta),$$

which follow from Eqs. (2.20) and (4.4), into Eq. (4.1), we obtain

$$s_0^{(\delta)}(x)\chi_0(x) - D_0(x - d_0 + \delta)\Pi(d_0 - \delta) - T_0(x - d_0 + \delta)\psi(d_0 - \delta) = -F_0(x - d_0 + \delta, -\delta), \quad \mu = 0, \tag{4.7}$$

where

$$F_0(x,x_0) = D_0(x)\Lambda_0(x_0) + T_0(x_0)\phi_0(x_0). \tag{4.8}$$

From Eqs. (3.2)–(3.4), it follows that

$$T_\mu(x) = D_\mu(x)R_\mu^{(s_x)}, \ R_\mu^{(s_x)} = D_\mu^{-1}(0)T_\mu(s_x \times 0) = s_x\beta_\mu. \tag{4.9}$$

Equation (4.8) can therefore be rewritten as

$$F_0(x,x_0) = D_0(x)\Phi_0^{(s_x)}(x_0), \tag{4.10}$$

where

$$\Phi_0^{(s_x)}(x_0) = D_0^{-1}(0)F_0(s_x \times 0, x_0). \tag{4.11}$$

Passing in Eq. (4.8) to the limit $x \to \pm 0$, and taking into account Eqs (3.2)–(3.4), (3.34), and (4.5), we obtain that

$$F_0(s_x \times 0, x_0) = -\tfrac{1}{2}(s_x + s_* s_{\gamma_0})\phi_0(x_0). \tag{4.12}$$

At the same time,

$$D_0(0) = -(2\beta_0)^{-1}, \tag{4.13}$$

where β_0 is understood as its limiting value $\bar{\beta}_0$ given by Eq. (3.34).

Combination of Eqs. (3.34) and (4.12) with Eq. (3.40) yields

$$\beta_0 = \bar{\beta}_0 = -i\epsilon_0^{xk}f_*^k = -\tfrac{1}{2}iv_*^x, \tag{4.14}$$

$$F_0(s_x \times 0, x_0) = -\theta(x)\phi_0(x_0). \tag{4.15}$$

Let us now pass in Eqs. (4.2), (4.3), and (4.7) to the $\lim \delta \to +0$ and take into account that

$$\lim_{\delta \to +0} s_\mu^{(\delta)}(x) = s_\mu(x), \quad x \neq d_{\mu-1}, d_\mu, \tag{4.16}$$

where $s_\mu(x)$ are defined by Eqs. (2.7). Then making use of Eqs. (2.26) and (2.27), we arrive at a chainlike set of functional equations of the Lippmann–Schwinger type, viz.,

$$\theta(d_0 - x)\chi_0(x) - A_0^{(1)}(x - d_0) = -F_0(x - d_0, 0), \; \mu = 0, \tag{4.17}$$

$$s_\mu(x)\psi(x) + A_\mu^{(\mu-1)}(x - d_{\mu-1}) - A_\mu^{(\mu+1)}(x - d_\mu) = 0, \quad \mu = 1,2,\ldots,n-1, \tag{4.18}$$

$$\theta(x - d_{n-1})\psi(x) + A_n^{(n-1)}(x - d_{n-1}) = 0, \; \mu = n, \tag{4.19}$$

where

$$A_\mu^{(\lambda)}(x - d_\nu) = D_\mu(x - d_\nu)\Pi(d_\nu) + T_\mu(x - d_\nu)\psi(d_\nu), \quad \lambda = \nu = \mu - 1 \text{ or } \mu = \lambda + 1, \nu = \mu, \tag{4.20}$$

and

$$F_0(x - d_0, 0) = \theta(x - d_0)D_0(x - d_0)\Phi_0^{(+1)}(0), \tag{4.21}$$

$$\Phi_0^{(+1)}(0) = -D_0^{-1}(0)\phi_0(0) = -a_0 D_0^{-1}(0) = 2a_0\beta_0, \tag{4.22}$$

in accordance with Eqs. (4.10), (4.11), and (4.15) [for $D_0(0)$, see Eqs. (4.13) and (4.14)].

The structure of the set of Lippmann–Schwinger equations is similar to the structure of Dyson's equations considered in Ref. 1. Each of Eqs. (4.17)–(4.19) relates to one definite layer. However, formally each of them holds in the whole infinite space ($-\infty < x < \infty$), except for the points on the separation planes ($x \neq d_\mu$, $\mu = 0,1,\ldots, n-1$). This is achieved by using the shape functions $s_\mu(x)$. The quantities $\psi(d_\mu)$ and $\Pi(d_\mu)$ in Eqs. (4.20) are unknowns to be found in the course of solving Eqs. (4.17)–(4.19).

It should be emphasized that Eqs. (4.17)–(4.19) determine both the usual "causal" solutions of the stationary scattering problem and the time-reversed states. The character of the solution depends on whether the retarded or the advanced "standard" Green's functions are used in Eqs. (4.20)–(4.22).

V. SOLUTION OF THE LIPPMANN–SCHWINGER EQUATIONS

Equations (4.17)–(4.19) can be solved by the method used in Ref. 1 for solving a similar set of Dyson's equations. Following this method, we set $x = x_0^+$ in Eq. (4.17), $x = x_\mu^-$ and $x = x_\mu^+$ in Eq. (4.18), and $x = x_n^-$ in Eq. (4.19), where x_μ^- and x_μ^+ are arbitrary fixed points satisfying the inequalities $x_\mu^- < d_{\mu-1}$ and $x_\mu^+ > d_\mu$, respectively. In particular, we can set

$$x = x_\mu^- = d_{\mu-1} - 0, \; x = x_\mu^+ = d_\mu + 0. \tag{5.1}$$

As a result, we obtain a chainlike set of $2n$ linear algebraic equations with respect to $2n$ unknowns $\psi(d_\mu)$ and $\Pi(d_\mu)$ ($\mu = 0,1,\ldots, n-1$). All these equations, except for the first, are homogeneous. Since

$$D_\mu(-0) = D_\mu(+0) = D_\mu(0), \quad T_\mu(-0) - T_\mu(+0) = 1 \tag{5.2}$$

[see Eqs. (3.2) and (3.3)], the same set of equations can also be obtained by the substitutions $x = d_{\mu-1} + 0$ and $x = d_\mu - 0$ instead of those given by Eqs. (5.1). Having found $\psi(d_\mu)$ and $\Pi(d_\mu)$, we can insert them back into Eqs. (4.17)–(4.19) and thus obtain the scattering amplitudes for all values of x.

In what follows, we consider a transformation which proves to be very useful in solving Eqs. (4.17)–(4.19). Also, we need this transformation in the further discussion (see the next section).

Setting $x = d_0 + 0$ in Eq. (4.17) and $x = d_{n-1} - 0$ in Eq. (4.19), we solve the resulting equations with respect to $\Pi(d_0)$ and $\Pi(d_{n-1})$, respectively. In view of Eqs. (4.20) and (4.21), we obtain

$$\Pi(d_0) = \Phi_0^{(+1)}(0) - R_0^{(+1)}\psi(d_0), \tag{5.3}$$

$$\Pi(d_{n-1}) = -R_n^{(-1)}\psi(d_{n-1}), \tag{5.4}$$

where $R_\mu^{(s)}$ is defined in Eqs. (4.9).

Substitution of Eq. (5.3) into Eqs. (4.17) and (4.18) with $\mu = 1$, and of Eq. (5.4) into Eqs. (4.18) with $\mu = n-1$ and (4.19) gives

$$\theta(d_0 - x)\chi_0(x) = \theta(d_0 - x)D_0(x - d_0)D_0^{-1}(0) \times [\psi(d_0) - \phi_0(0)], \quad \mu = 0, \tag{5.5}$$

$$s_1(x)\psi(x) - D_1(x - d_1)\Pi(d_1) - T_1(x - d_1)\psi(d_1) + \Theta_{01}^{(+1)}(x - d_0)\psi(d_0) = -D_1(x - d_1)\Phi_0^{(+1)}(0), \quad \mu = 1, \tag{5.6}$$

$$s_{n-1}(x)\psi(x) + D_{n-1}(x - d_{n-2})\Pi(d_{n-2}) + T_{n-1}(x - d_{n-2})\psi(d_{n-2}) - \Theta_{nn-1}^{(-1)}(x - d_{n-1})\psi(d_{n-1}) = 0, \quad \mu = n-1, \tag{5.7}$$

$$\theta(x - d_{n-1})\psi(x) = \theta(x - d_{n-1})D_n(x - d_{n-1}) \times D_n^{-1}(0)\psi(d_{n-1}), \quad \mu = n, \tag{5.8}$$

where

$$\Theta_{\mu\nu}^{(s)}(x) = T_\nu(x) - D_\nu(x)R_\mu^{(s)} = D_\nu(x)R_{\mu\nu}^{(s,s_\nu)}, \quad s = \pm 1, \tag{5.9}$$

$$R_{\mu\nu}^{(s_\mu,s_\nu)} = R_\nu^{(s_\nu)} - R_\mu^{(s_\mu)} = s_\nu\beta_\nu - s_\mu\beta_\mu, \quad s_\mu, s_\nu = \pm 1 \tag{5.10}$$

[see Eqs. (4.9)]. In Eqs. (5.5) and (5.8), the coefficient of $\psi(d_0)$ and $\psi(d_{n-1})$ appear originally in the form of $\Theta_{00}^{(+1)}(x - d_0)$ and $\Theta_{nn}^{(-1)}(x - d_{n-1})$, respectively. In view of Eqs. (4.9), (5.2), and (5.9), these can be written

$$\Theta_{\mu\mu}^{(s)}(x) = s\theta(-sx)D_\mu(x)D_\mu^{-1}(0), \quad (s = \pm 1, \; \mu = 0, n). \tag{5.11}$$

In deriving Eq. (5.5), we have also taken into account Eqs. (4.11) and (4.15).

Thus, we have eliminated the unknown constants $\Pi(d_0)$ and $\Pi(d_{n-1})$ and reduced the original set of $n+1$ Lippmann–Schwinger equations to a set of $n-1$ interdependent functional equations. These are Eq. (5.6), Eqs. (4.8) with $\mu = 2,3,\ldots,n-2$, and Eq. (5.7). Equations (5.5) and (5.8) are independent of those equations and of each other. On finding $\psi(d_0)$ and $\psi(d_{n-1})$, Eqs. (5.5) and (5.8) determine $\chi_0(x)$ ($x < d_0$) and $\psi(x)$ for $x > d_{n-1}$, respectively. Equation (5.5) for $x > d_0$ and Eq. (5.8) for $x > d_{n-1}$ become identities of the type $0 \equiv 0$.

Moving from the beginning and from the end of the reduced set of $n-1$ equations towards its middle, we can continue the process of successive elimination of the remaining unknown constants $\psi(d_\mu)$ and $\Pi(d_\mu)$ until the original set of Lippmann–Schwinger equations is completely solved.

Combination of Eqs. (4.20) with Eqs. (3.2)–(3.4) gives

$$A_\mu^{(\mu-1)}(x-d_{\mu-1}) = -a_0\alpha_\mu^{(+1)}\exp[if_\mu^{(+1)x}(x-d_{\mu-1})], \quad x > d_{\mu-1},$$
$$A_\mu^{(\mu+1)}(x-d_\mu) = a_0\alpha_\mu^{(-1)}\exp[if_\mu^{(-1)x}(x-d_\mu)], x<d_\mu, \tag{5.12}$$

where

$$\alpha_\mu^{(+1)} = (2\alpha_0\beta_\mu)^{-1}[\Pi(d_{\mu-1}) + \beta_\mu\psi(d_{\mu-1})],$$
$$\alpha_\mu^{(-1)} = -(2a_0\beta_\mu)^{-1}[\Pi(d_\mu) - \beta_\mu\psi(d_\mu)] \tag{5.13}$$

are constants independent of a_0. In Eqs. (5.12) and in all equations below, $f_\mu^{(s)x}$ is understood as its limiting value $\bar{f}_\mu^{(s)x}$ as $\gamma_\mu \to \pm 0$ (see Sec. III).

Inserting Eqs. (5.12) into Eqs. (4.17)–(4.19), we express the scattering amplitudes in the form of superpositions of plane waves, viz.,

$$\chi_0(x) = a_0\alpha_0^{(-1)}\exp[if_0^{(-1)x}(x-d_0)], \; x<d_0, \quad \mu=0, \tag{5.14}$$
$$\psi(x) = a_0\{\alpha_\mu^{(+1)}\exp[if_\mu^{(+1)x}(x-d_{\mu-1})] + \alpha_\mu^{(-1)}\exp[if_\mu^{(-1)x}(x-d_\mu)]\}, \quad d_{\mu-1} < x < d_\mu,$$
$$\mu = 1,2,\ldots,n-1, \tag{5.15}$$
$$\psi(x) = a_0\alpha_n^{(+1)}\exp[if_n^{(+1)x}(x-d_{n-1})], \quad x > d_{n-1}, \; \mu = n. \tag{5.16}$$

In writing Eq. (5.14), we have taken into account that $F_0(x-d_0,0) = 0$ if $x<d_0$ [see Eq. (4.21)].

Each given wave in Eqs. (5.15) and (5.16) is either travelling or spatially damped depending on whether the corresponding quantity $f_\mu^{(\pm 1)x} = \bar{f}_\mu^{(\pm 1)x}$ is real or complex (see Sec. III). However, in any case, we have

$$f_0^{(-1)x} = \bar{f}_0^{(-1)x} = f_*^x - 2(\mathbf{v}_0\cdot\mathbf{f}_*) = -f_*^x - 2(\epsilon_0^{xx})^{-1}\epsilon_0^{xk}f_\parallel^k \tag{5.17}$$

[see Eqs. (3.26), (3.35)–(3.38), and (3.40); $n^i = \delta^{ix}$]. Thus, $\chi_0(x)$ as given by Eq. (5.14) always corresponds to a traveling wave. According to Eq. (3.30), the group velocity of this wave is determined by the relation

$$v_0^i(\mathbf{f}_0^{(-1)}) = v_0^i(\mathbf{f}_*) - 2v_0^i(\mathbf{v}_0(\mathbf{f}_*)\cdot\mathbf{n}) = v_*^i - 2v_*^x(\epsilon_0^{xx})^{-1}\epsilon_0^{ix} \tag{5.18}$$

[see also Eqs. (2.18) and (2.19)]. If the medium is isotropic, we have $f_0^{(-1)x} = -f_*^x$, as it should be

From Eqs. (5.14) and (5.16), it follows that

$$\alpha_0^{(+1)} = a_0^{-1}\chi_0(d_0), \quad \alpha_n^{(-1)} = a_0^{-1}\psi(d_{n-1}). \tag{5.19}$$

Formally, these relations can be proved with the help of Eqs. (5.3), (5.4), and (5.13) [see also Eqs. (5.5) and (5.8)].

Equations (4.17)–(4.19) and the related discussion correspond to a stratified infinite space. In the case of a stratified semi-infinite space $x < d_{n-1}$, one should omit Eq. (4.19) and combine the last of Eqs. (4.18) (i.e., for $\mu = n-1$) with Eq. (2.30). The set of functional equations so obtained can be treated in the same way as above.

IV. THE REFLECTION AND TRANSMISSION COEFFICIENTS

We define the reflection (R) and transmission (T) coefficients of a stratified medium as

$$R = \lim_{(d_0-x)\to+\infty} \lim_{\gamma_0\to\pm 0} |j_{\text{refl}}^x(x)|/|j_{\text{in}}^x|,$$
$$T = \lim_{(x-d_{n-1})\to+\infty} \lim_{\gamma_n\to\pm 0} |j_{\text{tr}}^x(x)|/|j_{\text{in}}^x|, \tag{6.1}$$

where $\mathbf{j}_{\text{in}}$, $\mathbf{j}_{\text{refl}}(x)$, and $\mathbf{j}_{\text{tr}}(x)$ are the current densities of the incident, reflected, and transmitted waves, respectively. The limiting transitions in Eqs. (6.1) should be understood in the sense that

$$|\eta_0^{(-1)}| \ll (d_0-x)^{-1} \ll |f_0^{(-1)x}|, \quad x<d_0,$$
$$|\eta_n^{(+1)}| \ll (x-d_{n-1})^{-1} \ll |f_n^{(+1)x}|, \quad x>d_{n-1}, \tag{6.2}$$

where $f_\mu^{(s)x}$ and $\eta_\mu^{(s)}$ are defined by Eqs. (3.13)–(3.19) [see also Eq. (5.17)]. Thus the limiting transition $\gamma_0 \to \pm 0$ or $\gamma_n \to \pm 0$ must be performed before the corresponding limiting transition with respect to x.

It should be emphasized that all γ_μ must be of the same sign, in accordance with Eq. (3.40). Also, they should have been taken to be positive, in view of the criterion (2.24). However, in fact, the final results for R and T prove to be independent of the choice of the sign index s_*.

Substitution of $\phi_0(x)$ $(x<d_0)$, $\chi_0(x)$ $(x<d_0)$ or $\psi(x)$ $(x>d_{n-1})$ as given by Eq. (2.21), (5.14), or (5.16), for $\psi(x)$ into Eq. (2.40) yields

$$\mathbf{j}_{\text{in}} = |a_0|^2\mathbf{v}_{0g}(\mathbf{f}_*),$$
$$\mathbf{j}_{\text{refl}}(x) = |\chi_0(x)|^2\text{Re}\mathbf{v}_{0g}(\mathbf{f}_0^{(-1)}), \quad x<d_0, \tag{6.3}$$
$$\mathbf{j}_{\text{tr}}(x) = |\psi(x)|^2\text{Re}\mathbf{v}_{ng}(\mathbf{f}_n^{(+1)}), \quad x>d_{n-1},$$

where $\mathbf{v}_{ng}(\mathbf{f}_*)$, $\mathbf{v}_{0g}(\mathbf{f}_0^{(-1)})$, and $\mathbf{v}_{ng}(\mathbf{f}_n^{(+1)})$ are the group velocities of the incident, reflected and transmitted waves, respectively.

Inserting Eqs. (6.3) into Eqs. (6.1), and taking into account Eq. (2.18), we obtain

$$R = |\alpha_0^{(-1)}|^2|v_0^x(\bar{\mathbf{f}}_0^{(-1)})|/|v_0^x(\mathbf{f}_*)| = |\alpha_0^{(-1)}|^2, \tag{6.4}$$

$$T = \begin{cases} |\alpha_n^{(+1)}|^2|v_n^x(\bar{\mathbf{f}}_n^{(+1)})|/|v_0^x(\mathbf{f}_*)| = |\alpha_n^{(+1)}|^2|\beta_n|/|\beta_0| & \text{if } E_n > u_n, \quad (6.5) \\ 0 & \text{if } E_n < u_n \quad (6.6) \end{cases}$$

In developing Eqs. (6.4) and (6.5), we have made use of the second of Eqs. (3.21) and Eq. (4.14); β_0 and β_n are understood as $\bar{\beta}_0$ and β_n, respectivley. In order to prove Eq. (6.6), we note that, in the case of $E_n < u_n$, the quantity $f_n^{(+1)x}$ is complex and $v_n^i(\mathbf{f}_n^{(+1)})$ is pure imaginary as $\gamma_n \to \pm 0$ [see Eqs. (3.14), (3.16), and (3.20)]. Therefore both factors $|\psi(x)|^2$ and Re $\mathbf{v}_{0g}$ in the expression for $\mathbf{j}_{\text{tr}}(x)$ [see Eqs. (6.3)] vanish if x is large enough and $\gamma_n \to \pm 0$.

The quantities R and T as given by Eqs. (6.4)–(6.6) satisfy the relation

$$R = 1 - T. \tag{6.7}$$

This can be proved in the general form in analogy with what is done in quantum mechanics (see, e.g., Ref. 5, pp. 75–78). In each particular case, Eq. (6.7) can also be verified by straightforward calculations (see, e.g., the next section).

The above results were derived for Schrödinger's field. In the case of a classical field, the only difference is that one should use Eq. (2.19) instead of Eq. (2.18). Therefore Eqs. (6.1), (6.3), (6.4), and (6.6) remain in force. The same applies to Eq. (6.5) if we assume that $U_0 = U_n = 0$ and therefore

$$\mathscr{E}_0(\mathbf{f}_\bullet) = \mathscr{E}_n(\bar{\mathbf{f}}_n^{(+1)}) = E \equiv \omega^2 \tag{6.8}$$

[see Eqs. (2.11), (2.21), and (3.5)].

VII. EXAMPLES

A. Two half-spaces in contact ($n = 1$)

By way of example let us consider the simplest case of two semi-infinite spaces $x < d_0$ and $x > d_0$ is contact. This corresponds to $n = 1$. More complicated cases of a three-layer medium and a compound semi-infinite space will be considered in Secs. VIIB and VIIC.

Equating the right-hand sides of Eqs. (5.3) and (5.4) ($n = 1$) for $\Pi(d_0)$, we obtain

$$\psi(d_0) = R_{10}^{(-1,+1)-1}\Phi_0^{(+1)}(0) = 2a_0\beta_0/(\beta_0+\beta_1), \tag{7.1}$$

where use has been made of Eqs. (4.22), (4.13), and (5.10).

Combination of Eqs. (5.19) ($n = 1$) with Eqs. (2.20), (2.21), and (7.1) gives

$$\alpha_0^{(-1)} = (\beta_0-\beta_1)/(\beta_0+\beta_1), \quad \alpha_1^{(+1)} = 2\beta_0/(\beta_0+\beta_1). \tag{7.2}$$

As before, a bar over β_μ is omitted.

One can easily verify that R and T as given by Eqs. (6.4) and (6.5) ($n = 1$), subject to Eqs. (7.2), do satisfy Eq. (6.7). Also, in view of the definition of β_0 [see Eq. (4.14)] and β_1 [see Eqs. (3.16) and (3.17)], it follows from Eqs. (7.2) that $|\alpha_0^{(-1)}| = 1$ if $E_1 < u_1$, and $|\alpha_0^{(-1)}| < 1$ if $E_1 > u_1$.

If $|\beta_1| \to \infty$ (i.e., $U_1 \to \infty$ or $\epsilon_1^{xx} \to \infty$, $|q_1| < \infty$), Eqs. (7.2) become

$$\alpha_0^{(-1)} = -1, \quad \alpha_1^{(+1)} = 0, \tag{7.3}$$

which corresponds to the semi-infinite space $x < d_0$ with a rigid boundary $x = d_0$.

B. A three-layer medium ($n = 2$)

In the case of a medium consisting of three layers in contact, viz., a flat slab $d_0 < x < d_1$ and two semi-infinite spaces $x < d_0$ and $x > d_1$, one should set $n = 2$ in all general relations. Combining Eqs. (5.3) and (5.7) or (5.4) and (5.6), we obtain the same equation, namely

$$\begin{aligned} s_1(x)\psi(x) &+ \Theta_{01}^{(+1)}(x-d_0)\psi(d_0) - \Theta_{21}^{(-1)}(x-d_1)\psi(d_1) \\ &= -D_1(x-d_1)\Phi_0^{(+1)}(0), \quad \mu = 1. \end{aligned} \tag{7.4}$$

In order to solve this equation, we evaluate it at $x = d_0 - 0$ and $x = d_1 + 0$ in turn. Thus, we get the set of two linear algebraic equations

$$\begin{aligned} &\Theta_{01}^{(+1)}(-0)\psi(d_0) - \Theta_{21}^{(-1)}(-l_1)\psi(d_1) = -D_1(0)\Phi_0^{(+1)}(0), \\ &\Theta_{01}^{(+1)}(l_1)\psi(d_0) - \Theta_{21}^{(-1)}(+0)\psi(d_1) = -D_1(l_1)\Phi_0^{(+1)}(0), \end{aligned} \tag{7.5}$$

where, according to Eqs. (3.2), (5.9), and (5.10),

$$\begin{aligned} \Theta_{01}^{(+1)}(-0) &= (2\beta_1)^{-1}(\beta_1+\beta_0), \\ \Theta_{01}^{(+1)}(l_1) &= -(2\beta_1)^{-1}(\beta_1-\beta_0)e_1(l_1), \\ \Theta_{21}^{(-1)}(+0) &= -(2\beta_1)^{-1}(\beta_1+\beta_2), \\ \Theta_{21}^{(-1)}(-l_1) &= (2\beta_1)^{-1}(\beta_1-\beta_2)e_1(-l_1), \end{aligned} \tag{7.6}$$

and

$$l_1 = d_1 - d_0 \tag{7.7}$$

is the thickness of the slab.

From Eq. (7.5), it follows that

$$\begin{aligned} \psi(d_0) &= M_2 S_{+1}^{-1}\Phi_0^{(+1)}(0), \\ \psi(d_1) &= 2\beta_1 e_1(l_1) S_{+1}^{-1}\Phi_0^{(+1)}(0), \end{aligned} \tag{7.8}$$

where we use the notation

$$\begin{aligned} S_{\sigma_0} = (\beta_1+\sigma_0\beta_0)(\beta_1+\beta_2) &- (\beta_1-\sigma_0\beta_0)(\beta_1-\beta_2) \\ &\times \exp(-2q_1l_1), \quad \sigma_0 = \pm 1, \end{aligned} \tag{7.9}$$

$$M_\mu = \beta_1 + \beta_\mu + (\beta_1-\beta_\mu)\exp(-2q_1l_1), \quad \mu = 0,2. \tag{7.10}$$

Substitution of Eqs. (3.2) and (7.8) into Eqs. (5.5), (5.8) ($n = 2$), and (7.4) gives Eqs. (5.14), (5.15) ($\mu = 1$), and (5.16) ($n = 2$) with

$$\begin{aligned} \alpha_0^{(-1)} &= a_0^{-1}[\psi(d_0) - \phi_0(0)] \\ &= 2\beta_0 M_2 S_{+1}^{-1} - 1 = -S_{-1}S_{+1}^{-1}, \\ \alpha_1^{(+1)} &= (a_0S_{+1})^{-1}(\beta_1+\beta_2)\Phi_0^{(+1)}(0) \\ &= 2\beta_0(\beta_1+\beta_2)S_{+1}^{-1}, \\ \alpha_1^{(-1)} &= (a_0S_{+1})^{-1}(\beta_1-\beta_2)e_1(l_1)\Phi_0^{(+1)}(0) \\ &= 2\beta_0(\beta_1-\beta_2)S_{+1}^{-1}\exp(if_1^{(+1)x}l_1), \\ \alpha_2^{(+1)} &= a_0^{-1}\psi(d_1) = 4\beta_0\beta_1S_{+1}^{-1}\exp(if_1^{(+1)x}l_1) \end{aligned} \tag{7.11}$$

[see also Eqs. (5.13) with $\mu = 1$ and (5.19) with $n = 2$]. In developing these results, we have made use of Eqs. (5.9), (5.10), (7.9), and (7.10). We have also taken into account the definitions of $\phi_0(x)$, $e_\mu(x)$, and $\Phi_0^{(+1)}(0)$ [see Eqs. (2.21), (3.4), and (4.22), respectively]. As before, we omit bars over all variables. Thus, β_μ with $\mu = 1,2$, is given by Eqs. (3.16) and (3.17) together with Eq. (3.40); β_0 is given by Eq. (4.14). The quantities q_1 and $f_1^{(+1)x}$ are expressed in terms of β_1 by Eqs. (3.8) and (3.14).

Since β_0 is always pure imaginary, and β_2 can be either real or pure imaginary, it is convenient to rewrite Eq. (7.9) in the form

$$S_{\pm 1} = \pm 2e^{-q_1l_1}S_\pm, \tag{7.12}$$

where

$$\begin{aligned} S_\pm &= S' \pm S'', \\ S' &= \beta_0(\beta_1\cosh q_1l_1 + \beta_2\sinh q_1l_1), \end{aligned} \tag{7.13}$$

$S'' = \beta_1(\beta_2\cosh q_1l_1 + \beta_1\sinh q_1l_1)$ if $E_2 < u_2$ (β_2 is real)

or

$$S_\pm = S_\pm^{(1)} + S_\pm^{(2)},$$
$$S_\pm^{(1)} = -is_* \beta_1(|\beta_0| \pm |\beta_2|)\cosh q_1 l_1, \quad (7.14)$$
$$S_\pm^{(2)} = (\pm\beta_1^2 - |\beta_0||\beta_2|)\sinh q_1 l_1 \quad \text{if } E_2 > u_2$$
$$(\beta_2 \text{ is imagnary}).$$

In order to derive Eqs. (7.14) from Eqs. (7.13), we rearrange somewhat the terms in Eqs. (7.13) and make use of Eqs. (3.17) and (3.40) with $\mu = 2$ [see also Eqs. (2.22) and (4.14)].

Taking into account Eqs. (3.16) and (3.17) with $\mu = 1$, one can readily see from Eqs. (7.13) and (7.14) that S' or $S_\pm^{(1)}$ is pure imaginary (real) and simultaneously, S'' or $S_\pm^{(2)}$ is real (pure imaginary) if β_1 is real (pure imaginary). Hence,

$$|S_\pm|^2 = |S'|^2 + |S''|^2 \quad \text{if } E_2 < u_2, \quad (7.15)$$
$$|S_\pm|^2 = |S_\pm^{(1)}|^2 + |S_\pm^{(2)}|^2 = |\beta_1|^2(|\beta_0|^2 + |\beta_2|^2) \times\cosh^2 q_1 l_1 + (|\beta_1|^4 + |\beta_0|^2|\beta_2|^2)|\sinh q_1 l_1|^2 \pm 2|\beta_0||\beta_1|^2|\beta_2| \quad \text{if } E_2 > u_2. \quad (7.16)$$

In developing Eq. (7.16), we have made use of the identity

$$\beta_1^2|\sinh q_1 l_1|^2 \equiv |\beta_1|^2 \sinh^2 q_1 l_1 \quad (q_1 = (\epsilon_1^{xx})^{-1}\beta_1), \quad (7.17)$$

which is valid if β_1 is real or pure imaginary.

Combination of Eqs. (7.11) and (7.12) gives

$$\alpha_0^{(-1)} = S_-/S_+, \quad \alpha_2^{(+1)} = 2\beta_0\beta_1 e^{ip_1 l_1}/S_+$$

(see also Eq. (3.7)). Therefore, Eqs. (6.4) and (6.5) become

$$R = |S_-|^2/|S_+|^2, \; T = 4|\beta_0||\beta_1|^2|\beta_2|/|S_+|^2. \quad (7.18)$$

In view of Eqs. (7.15) and (7.16), $R = 1$ if $E_2 < u_2$, and $R < 1$ if $E_2 > u_2$. Also, one can readily verify that R and T as given by Eqs. (7.18) subject to Eq. (7.16) satisfy Eq. (6.7).

If $\beta_0 = \beta_2$ (i.e., $\epsilon_0^{ik} = \epsilon_2^{ik}$ and $U_0 = U_2$), Eq. (7.16) becomes

$$|S_\pm|^2 = (|\beta_0|^2 + |\beta_1|^2)^2|\sinh q_1 l_1|^2 + 4\sigma_\pm|\beta_0|^2|\beta_1|^2, \quad \sigma_+ = 1, \; \sigma_- = 0, \quad (7.19)$$

where we have again made use of Eq. (7.17).

C. A compound half-space

If $|\beta_2| \to \infty$ (i.e., $U_2 \to \infty$ or $\epsilon_2^{xx} \to \infty$, $|q_2| < \infty$), Eqs. (7.11) subject to Eqs. (7.12) and (7.13) become

$$\alpha_0^{(-1)} = S_-^{(\infty)}/S_+^{(\infty)}, \; \alpha_1^{(-1)} = \beta_0 e^{q_1 l_1}/S_+^{(\infty)},$$
$$\alpha_1^{(+1)} = -\beta_0 e^{ip_1 l_1}/S_+^{(\infty)}, \quad \alpha_2^{(+1)} = 0, \quad (7.20)$$

where

$$S_\pm^{(\infty)} = \beta_0 \sinh q_1 l_1 \pm \beta_1 \cosh q_1 l_1. \quad (7.21)$$

Evidently, $S_+^{(\infty)} = \frac{1}{2}e^{q_1 l_1}M_0$, where M_0 is defined by Eq. (7.10).

In this case, Eqs. (5.14) and (5.15) ($\mu = 1$) give the scattering amplitudes in a medium consisting of a flat slab $d_0 < x < d_1$ with a rigid boundary $x = d_1$ and a semi-infinite space $x < d_0$ in contact ($\psi(x) = \alpha_2^{(+1)} = 0$ if $x > d_1$). In analogy with Eq. (7.15), one can readily see that

$$|S_\pm^{(\infty)}|^2 = |\beta_0|^2|\sinh q_1 l_1|^2 + |\beta_1|^2(\cosh q_1 l_1)^2, \quad (7.22)$$

and therefore $R = |S_-^{(\infty)}|^2/|S_+^{(\infty)}|^2 = 1$, as it should be expected.

VIII. RELATIONS BETWEEN THE SCATTERING AMPLITUDES AND THE TOTAL GREEN'S FUNCTION

A. The total Green's function

In Ref. 1, we derived the set of Dyson's equations for the total Green's function of a stratified medium in the $(x, \mathbf{f}_\parallel)$ representation, $D(x,x_0) \equiv D(x,x_0;\mathbf{f}_\parallel,\{w_\nu\})$. In analogy with Eqs. (4.17)–(4.19), those equations can be written

$$\theta(d_0 - x)D(x,x_0) - B_0^{(1)}(x - d_0,x_0) = \theta(d_0 - x_0)D_0(x - x_0), \quad \mu = 0, \quad (8.1)$$
$$s_\mu(x)D(x,x_0) + B_\mu^{(\mu-1)}(x - d_{\mu-1},x_0) - B_\mu^{(\mu+1)}(x - d_\mu,x_0) = s_\mu(x_0)D_\mu(x - x_0), \quad \mu = 1,2,\ldots,n-1, \quad (8.2)$$
$$\theta(x - d_{n-1})D(x,x_0) + B_n^{(n-1)}(x - d_{n-1},x_0) = \theta(x_0 - d_{n-1})D_n(x - x_0), \quad \mu = n, \quad (8.3)$$

where

$$B_\mu^{(\lambda)}(x - d_\nu,x_0) = D_\mu(x - d_\nu)P(d_\nu,x_0) + T_\mu(x - d_\nu)D(d_\nu,x_0),$$
$$\lambda = \nu = \mu - 1 \text{ or } \lambda = \mu + 1, \; \nu = \mu, \quad (8.4)$$
$$D(d_\nu,x_0) = D(d_\nu - 0,x_0) = D(d_\nu + 0,x_0), \quad (8.5)$$
$$P(d_\nu,x_0) = P(d_\nu - 0,x_0) = P(d_\nu + 0,x_0), \quad x_0 \neq d_\nu, \quad (8.6)$$

$$P(x,x_0) = -\epsilon_\mu^{xk}\partial_x^k D(x,x_0), \quad d_{\mu-1} < x < d_\mu, \quad -\infty < x_0 < \infty, \; \mu = 0,1,\ldots,n \quad (8.7)$$

[see Eqs. (2.12)–(2.17) of Ref. 1]. The quantities $D(d_\nu,x_0)$ and $P(d_\nu,x_0)$ are unknowns to be found in the course of solving the Dyson's equations [compare with $\psi(d_\nu)$ and $\Pi(d_\nu)$ in the case of Lippmann–Schwinger equations].

In the general case, the total Green's function of a stratified medium can be written as

$$D(x,x_0) = \sum_{\mu=0}^{n}\sum_{\mu_0=0}^{n} s_\mu(x)s_{\mu_0}(x_0)D_{\mu\mu_0}(x,x_0), \quad (8.8)$$

where

$$D_{\mu\mu_0}(x,x_0) = \delta_{\mu\mu_0}D_\mu(x - x_0) + A_{\mu\mu_0}(x,x_0), \quad (8.9)$$
$$A_{\mu\mu_0}(x,x_0) = \sum_{\nu=\mu-1,\mu}\;\sum_{\nu_0=\mu_0-1,\mu_0} D_\mu(x - d_\nu)K_{\mu\mu_0,\nu\nu_0}D_{\mu_0}(d_{\nu_0} - x_0), \quad (8.10)$$

while $K_{\mu\mu_0,\nu\nu_0} = K_{\mu\mu_0,\nu\nu_0}(\mathbf{f}_\parallel,\{w_\lambda\})$ is a matrix which depends neither on x nor on x_0. In each particular case, this matrix can be found in the course of solving the corresponding set of Dyson's equations.

From the reciprocity relation[1]

$$D(x,x_0;\mathbf{f}_\parallel,\{w_\lambda\}) = D(x_0,x;-\mathbf{f}_\parallel,\{w_\lambda\}), \quad (8.11)$$

if follows that

$$A_{\mu\mu_0}(x,x_0;\mathbf{f}_\parallel,\{w_\lambda\}) = A_{\mu_0\mu}(x_0,x;-\mathbf{f}_\parallel,\{w_\lambda\}) \quad (8.12)$$

and therefore

$$K_{\mu\mu_0,\nu\nu_0}(\mathbf{f}_\parallel,\{w_\lambda\}) = K_{\mu_0\mu,\nu_0\nu}(-\mathbf{f}_\parallel;\{w_\lambda\}). \quad (8.13)$$

In order to derive Eqs. (8.8)–(8.10), we insert Eq. (8.4) into Eqs. (8.1)–(8.3) and make use of Eqs. (4.9). Thus, we arrive at Eqs. (8.8) and (8.9) with

$$A_{\mu\mu_0}(x,x_0) = \sum_{\nu=\mu-1,\mu} D_\mu(x-d_\nu)L_{\mu\mu_0,\nu}(x_0), \tag{8.14}$$

where

$$L_{\mu\mu_0,\mu-1}(x_0) = -P(d_{\mu-1},x_0) - \beta_\mu D(d_{\mu-1},x_0), \tag{8.15}$$

$$L_{\mu\mu_0,\mu}(x_0) = P(d_\mu,x_0) - \beta_\mu D(d_\mu,x_0)$$

[cf. Eqs. (5.13)]. In writing these results, we have taken into account that μ and μ_0 in Eqs. (8.9) and (8.10) are indices of layers which contain x and x_0, respectively [see Eqs. (2.7) and (8.8)]. Therefore the right-hand sides of Eqs. (8.15) depend on μ_0 implicitly. Since

$$D_\mu(x;\mathbf{f}_\parallel,w_\mu) = D_\mu(-x;-\mathbf{f}_\parallel,w_\mu), \tag{8.16}$$

combination of Eqs. (8.12) and (8.14) immediately gives Eq. (8.10).

According to Eqs. (3.2), (3.4), and (3.7), $D_\mu(x)$ is continuous on the whole x-axis. From Eqs. (8.5), (8.8)–(8.10), it therefore follows that $D(x,x_0)$ is a continuous (and hence bounded) function of both x and x_0, i.e.,

$$D(x\pm\delta,x_0) = D(x,x_0\pm\delta) = D(x,x_0), \tag{8.17}$$

where δ is a positive infinitesimal. Equation (8.17) holds for $|x-x_0| \gg \delta$ as well as for $|x-x_0| \ll \delta$ (i.e., for $x = x_0$). In particular, x or x_0 or both can be equal to d_μ.

On the other hand, the quantity $P(x,x_0)$ as defined by Eq. (8.7) satisfies the relations

$$P(x-\delta,x_0) - P(x+\delta,x_0) = \begin{cases} 0 & \text{if } |x-x_0| \gg \delta, & (8.18) \\ 1 & \text{if } |x-x_0| \ll \delta. & (8.19) \end{cases}$$

This result can immediately be obtained by integrating both sides of Eq. (2.9) of Ref. 1 with respect to x between $x_1-\delta$ and $x_1+\delta$, where x_1 is any fixed point on the x-axis, and δ is, as before, a positive infinitesimal. Taking into account Eqs. (8.7) and (8.17) and replacing x_1 by x, we arrive at the desired result.

Thus, in calculating $P(x\pm\delta,x_0)$ as $x \to x_0$ and $\delta \to +0$, the order of the limiting transitions with respect to x and δ is essential.

It should also be noted that Eq. (8.6) is a particular case of Eq. (8.18), which corresponds to $|x-d_\mu| \ll \delta$ and $|x_0-d_\mu| \gg \delta$.

B. The scattering amplitudes

Setting $x_0 = d_0 - 0$ in Eqs. (8.1)–(8.3), we obtain

$$\theta(d_0-x)D(x,d_0-0) - B_0^{(1)}(x-d_0,d_0-0) = D_0(x-d_0), \quad \mu = 0, \tag{8.20}$$

$$s_\mu(x)D(x,d_0-0) + B_\mu^{(\mu-1)}(x-d_{\mu-1},d_0-0) - B_\mu^{(\mu+1)}(x-d_\mu,d_0-0) = 0, \quad \mu = 1,2,\ldots,n-1, \tag{8.21}$$

$$\theta(x-d_{n-1})D(x,d_0-0) + B_n^{(n-1)}(x-d_{n-1},d_0-0) = 0, \quad \mu = n. \tag{8.22}$$

Equation (8.20) and Eq. (8.21) with $\mu = 1$ involve the quantities $D(d_0,d_0-0)$ and $P(d_0,d_0-0)$ which should be understood as the limiting values

$$D(d_0,d_0-0) = \lim_{\delta_2\to+0}\lim_{\delta_1\to+0} D(d_0\pm\delta_1,d_0-\delta_2), \tag{8.23}$$

$$P(d_0,d_0-0) = \lim_{\delta_2\to+0}\lim_{\delta_1\to+0} P(d_0\pm\delta_1,d_0-\delta_2), \tag{8.24}$$

where δ_1 tends to zero first. This follows from the definitions of $D(d_\mu,x_0)$ and $P(d_\mu,x_0)$, Eqs. (8.5) and (8.6) $(|x-d_0| \ll |x_0-d_0|)$.

On account of the continuity of $D(x,x_0)$ (see Eq. (8.17)), we have

$$D(x,d_0-0) = D(x,d_0+0) = D(x,d_0) \tag{8.25}$$

and, in particular,

$$D(d_0,d_0,-0) = D(d_0,d_0+0) = D(d_0,d_0). \tag{8.26}$$

This means that the order of the limiting transitions in Eq. (8.23) can be changed. On the other hand, the same limiting transitions in Eq. (8.24) are not permutable because $P(x,x_0)$ is discontinuous at $x = x_0$. It should be noted that we do not need Eqs. (8.25) and (8.26) in the further considerations. However, these equations can be useful in specific calculations [see, for example, Eq. (8. 37) and the related discussion].

We observe that the sets of Eqs. (4.17)–(4.19) and (8.20)–(8.22) have the same "matrix of coefficients", and the unknowns which are involved in these equations, except for $\chi_0(x)$, satisfy the boundary conditions of the same kind [compare Eqs. (2.26) and (2.27) with Eqs. (8.5) and (8.6)]. However, the fact that $\chi_0(x)$ and $D(x,d_0-0)$ satisfy completely different boundary conditions at $x = d_0$ [see Eqs. (2.20)], (2.26) and (2.27) does not enable us to express the solution of Eqs. (4.17)–(4.19) directly in terms of $D(x,d_0-0)$, which is the solution of Eqs. (8.20)–(8.22).

This difficulty can be avoided by elminating Eqs. (4.17) and (8.20). Equation (4.17) is eliminated by the transformation considered in Sec. V. Likewise, in order to eliminate Eq. (8.20), we evaluate it at $x = d_0 + 0$, thus obtaining

$$P(d_0,d_0-0) = -1 - R_0^{(+1)}D(d_0,d_0-0), \tag{8.27}$$

where $R_0^{(+1)}$ is defined in Eqs. (4.9). Substitution of Eq. (8.27) back into Eq. (8.20) and into Eq. (8.21) with $\mu = 1$ gives, in analogy with Eqs. (5.5) and (5.6), that

$$\theta(d_0-x)D(x,d_0-0) = \theta(d_0-x)D_0(x-d_0)D_0^{-1}(0)D(d_0,d_0-0), \quad \mu = 0, \tag{8.28}$$

$$s_1(x)D(x,d_0-0) - D_1(x-d_1)P(d_1,d_0-0) - T_1(x-d_1) \times D(d_1,d_0-0) + \Theta_{01}^{(+1)}(x-d_0)D(d_0,d_0-0) = D_1(x-d_1), \quad \mu = 1, \tag{8.29}$$

where $\Theta_{01}^{(+1)}(x-d_0)$ is defined by Eqs. (5.9) and (5.10).

Let us now consider two reduced sets of functional equations, Lippmann–Schwinger and Dyson's. The first consists of Eqs. (5.6), (4.18) with $\mu = 2,3,\ldots,n-1$, and (4.19). The second consists of Eqs. (8.29), (8.21) with $\mu = 2,3,\ldots,n-1$, and (8.22). It is essential that these sets of equations involve only such unknown functions [viz., $\psi(x)$ and $D(x,d_0-0)$, respectively] which satisfy the boundary conditions of the same kind at all separation planes $x = d_\mu$, $\mu = 0,1,\ldots,n-1$.

We multiply Eqs. (8.21) with $\mu = 2,3,\ldots,n-1$, (8.22), and (8.29) by $\Phi_0^{(+1)}(0)$. Then putting

$$\psi(x) = -D(x,d_0-0)\Phi_0^{(+1)}(0) = -2a_0\beta_0 D(x,d_0-0),$$
$$x > d_0, \tag{8.30}$$

as well as

$$\psi(d_\mu) = -D(d_\mu,d_0-0)\Phi_0^{(+1)}(0)$$
$$= -2a_0\beta_0 D(d_\mu,d_0-0), \tag{8.31}$$
$$\Pi(d_\mu) = -P(d_\mu,d_0-0)\Phi_0^{(+1)}(0)$$
$$= -2a_0\beta_0 P(P_\mu,d_0-0) \tag{8.32}$$

for all $\mu = 0,1,...,n-1$, we arrive at the set of n functional equations, which exactly coincides with the reduced set of Lippmann–Schwinger equations [for $\Phi_0^{(+1)}(0)$, see Eqs. (4.22), (4.13), and (4.14)]. Since this set of equations has a unique solution, the quantity $\psi(x)$ as defined by Eq. (8.30) should be identified with the scattering amplitude, while the quantity $\Pi(d_\mu)$ as defined by Eq. (8.32) is identical with $\Pi(d_\mu)$ given by Eq. (2.27) and (2.28). Equation (8.31) is, obviously, the particular case of Eq. (8.30), which corresponds to $x = d_\mu$. It should also be noted that the reduced sets of Lippmann–Schwinger and Dyson's equations involve $\Pi(d_\mu)$ and $P(d_\mu,d_0-0)$ with all μ, except for $\mu = 0$. Nevertheless, Eq. (8.32) remains in force for $\mu = 0$ as well. This can be seen if we multiply Eq. (8.27) by $\Phi_0^{(+1)}(0)$ and compare the resulting with Eq. (5.3).

Inserting $\psi(d_0)$ as given by Eq. (8.31) into Eq. (5.5) and making use of Eq. (8.28), we obtain

$$\chi_0(x) = [D(x,d_0-0) - D_0(x-d_0)]D_0^{-1}(0)\phi_0(0)$$
$$= -a_0\{2\beta_0 D(x,d_0-0) + \exp[if_0^{(-1)x}(x-d_0)]\},$$
$$x < d_0, \tag{8.33}$$

where β_0 and $f_0^{(-1)x}$, are understood as $\bar{\beta}_0$ and $\bar{f}_0^{(-1)x}$, respectively [see Eqs. (4.14) and (5.17)].

It is convenient to write

$$\psi(x) = \sum_{\mu=1}^{n} s_\mu(x)\chi_\mu(x), \quad x > d_0, \tag{8.34}$$

where $s_\mu(x)$ are given by Eqs. (2.7). Thus, we introduce the notation

$$\psi(x) \equiv \chi_\mu(x), \quad d_{\mu-1} < x < d_\mu, \quad \mu = 1,2,...,n \tag{8.35}$$

[for $\mu = 0$, see Eq. (2.20)]. Then making use of Eqs. (8.8) and (8.9), we can rewrite Eqs. (8.30) and (8.33) in the form

$$\chi_\mu(x) = -2a_0\beta_0 A_{\mu 0}(x,d_0-0), \quad d_{\mu-1} < x < d_\mu,$$
$$\mu = 0,1,...,n. \tag{8.36}$$

Equations (8.30) and (8.33) or, equivalently, Eqs. (8.36) are the desired relations which express the scattering amplitudes in terms of the total Green's function. Combining these equations with the corresponding expressions for $D(x,x_0)$ of Ref. 1, we recover the results of Sec. VII.

In connection with Eqs. (8.30) and (8.33), we recall that the first step in the calculation of $D(x,x_0)$ consists of solving a set of linear algebraic equations in the unknowns $D(d_\mu,x_0)$ and $P(d_\mu,x_0)$, $\mu = 0,1,...,n-1$. These equations follow from Eqs. (8.1)–(8.3) as a result of any of the series of the substitutions mentioned in Sec. V. Having, in particular, found $D(d_0,x_0)$, we can make use of Eqs. (8.11) and (8.25) to write

$$D(x,d_0-0;\mathbf{f}_\parallel,\{w_\nu\}) = D(d_0,x;-\mathbf{f}_\parallel,\{w_\nu\}). \tag{8.37}$$

Thus, in order to calculate the scattering amplitudes by Eqs. (8.30) and (8.33), we need not calculate $D(x,x_0)$ for arbitrary x and x_0 and can confine ourselves to the first step in solving the Dyson's equations.

C. Space inversion

So far, we considered the stationary scattering problem in the case where the "bare" plane wave was given in the semi-infinite space $x < d_0$. If the "bare" plane wave is given in the semi-infinite space $x > d_{n-1}$, one should write, instead of Eq. (2.20), that

$$\psi(x) = \phi_n(x-d_{n-1}) + \chi_n(x), \quad x > d_{n-1}. \tag{8.38}$$

In this case, the "bare" wave $\phi_n(x)$ can be defined by substituting the subscript $\mu = n$ for $\mu = 0$ in Eqs. (2.21) which define $\phi_0(x)$. Also, we have now, by definition, that

$$s_* = \operatorname{sgn} v_n^x(\mathbf{f}_*), \tag{8.39}$$

in place of Eq. (2.22). Therefore the index s_* should be replaced by $-s_*$ in the criteria (2.24), (2.25), and in all related considerations. In particular, Eq. (3.40) turns into

$$s_{\gamma_\mu} = -s_*, \quad \mu = 0,1,...,n. \tag{8.40}$$

In analogy with what was done in Sec. IV, one can readily show that the set of Lippmann–Schwinger equations consists now of Eqs. (4.18) and the equations

$$\theta(d_0-x)\psi(x) - A_0^{(1)}(x-d_0) = 0, \quad \mu = 0, \tag{8.41}$$
$$\theta(x-d_{n-1})\chi_n(x) + A_n^{(n-1)}(x-d_{n-1}) = F_n(x-d_{n-1},0),$$
$$\mu = n, \tag{8.42}$$

where

$$F_n(x-d_{n-1},0) = \theta(d_{n-1}-x)D_n(x-d_{n-1})\Phi_n^{(-1)}(0), \tag{8.43}$$

$$\Phi_n^{(-1)}(0) = D_n^{-1}(0)F_n(-0,0) = D_n^{-1}(0)\phi_n(0) = -2a_n\beta_n \tag{8.44}$$

[compare with Eqs. (4.21) and (4.22)]. The quantities $A_\mu^{(\lambda)}(x-d_\nu)$ are, as before, defined by Eqs. (4.20). In order to make sure of the validity of Eqs. (8.43) and (8.44), we note that Eqs. (3.34) and (4.8)–(4.13) can be rewritten correctly by substituting the subscript n (or, in general, any μ) for the subscript 0. However, on account of Eq. (8.40), Eqs. (4.14) and (4.15) are replaced now by

$$\beta_n = \bar{\beta}_n = i\epsilon_n^{xk} f_*^k = \tfrac{1}{2} i v_n^x(\mathbf{f}_*),$$
$$F_n(s_x \times 0,x_0) = \theta(-x)\phi_n(x_0). \tag{8.45}$$

The set of Lippmann–Schwinger equations so obtained can be solved in a straightforward way as was described in Sec. V. Alternatively, the scattering amplitudes can be expressed in terms of the total Green's function in analogy with what was done in the previous subsection.

Following the latter method, we make the substitution $x_0 = d_{n-1} + 0$ in Eqs. (8.1)–(8.3). Thus we arrive at equations similar to Eqs. (8.21) as well as at the equations

$$\theta(d_0-x)D(x,d_{n-1}+0) - B_0^{(1)}(x,d_{n-1}+0) = 0,$$
$$\mu = 0, \tag{8.46}$$

$$\theta(x-d_{n-1})D(x,d_{n-1}+0) + B_n^{(n-1)}(x,d_{n-1}+0) = D_n(x-d_{n-1}),$$
$$\mu = n, \tag{8.47}$$

which appear instead of Eqs. (8.20) and (8.22), respectively. On eliminating Eqs. (8.42) and (8.47) (by the substitution $x = d_{n-1} - 0$), we proceed as in the previous subsection. The final results can be written as

$$\psi(x) = D(x,d_{n-1}+0)\Phi_n^{(-1)}(0) = -2a_n\beta_n D(x,d_{n-1}+0), \quad x<d_{n-1}, \tag{8.48}$$

$$\chi_n(x) = [D(x,d_{n-1}+0) - D_n(x-d_{n-1})]D_n^{-1}(0)\phi_n(0) = -a_n\{2\beta_n D(x,d_{n-1}+0) + \exp[if_n^{(+1)x}(x-d_{n-1})]\}, \quad x>d_{n-1}, \tag{8.49}$$

$$\psi(d_\mu) = D(d_\mu,d_{n-1}+0)\Phi_n^{(-1)}(0) = -2a_n\beta_n D(d_\mu,d_{n-1}+0), \tag{8.50}$$

$$\Pi(d_\mu) = P(d_\mu,d_{n-1}+0)\Phi_n^{(-1)}(0) = -2a_n\beta_n P(d_\mu,d_{n-1}+0) \tag{8.51}$$

$(\mu = 0,1,\ldots, n-1)$, where

$$f_n^{(+1)x} = -f_*^x - 2(\epsilon_n^{xx})^{-1}\epsilon_n^{xk}f_\parallel^k, \tag{8.52}$$

[compare with Eq. (5.17)].

In analogy with Eq. (8.34), we write

$$\psi(x) = \sum_{\mu=0}^{n-1} s_\mu(x)\chi_\mu(x), \quad x<d_{n-1}, \tag{8.53}$$

thus introducing the notation as given by Eq. (8.35) with $\mu = 0,1,\ldots,n-1$, [$\chi_n(x)$ is defined by Eq. (8.38)]. As a consequence, Eqs. (8.48) and (8.49) become

$$\chi_\mu(x) = -2a_n\beta_n A_{\mu n}(x,d_{n-1}+0), \quad \mu = 0,1,\ldots,n \tag{8.54}$$

[compare with Eq. (8.36)].

IX. DISCUSSION

A. Stationary states

Equations (2.20), (8.30), and (8.33), on the one hand, or Eqs. (8.38), (8.48), and (8.49), on the other hand, determine the stationary states of the field, which can be denoted as $\psi_0(x;\mathbf{f}_\parallel,\{E_\nu \pm i0\})$ or $\psi_n(x;\mathbf{f}_\parallel,\{E_\nu \pm i0\})$, respectively. The state $\psi_\mu(x,\mathbf{f}_\parallel,\{E_\nu + i0\})$ $(\mu = 0,n)$ is associated with a wave process which develops in the usual causal time sequence. In accordance with Eq. (2.12), the state $\psi_\mu(x;\mathbf{f}_\parallel,\{E_\nu - i0\})$ is then time-reversed to $\psi_\mu(x;-\mathbf{f}_\parallel;\{E_\nu + i0\})$. The corresponding quantity $D(x,x_0;\mathbf{f}_\parallel,\{E_\nu + i0\})$ or $D(x,x_0;\mathbf{f}_\parallel,\{E_\nu - i0\})$ should therefore be identified as the total retarded or advanced Green's function, respectively.

B. Localized states

We have studied the stationary states which have the asymptotic form of an incident ("bare") plane wave plus a reflected plane wave either in the semi-infinite space $x<d_0$ or in the semi-infinite space $x>d_{n-1}$. In addition, there can exist stationary states of the field, which penetrate neither of the two semi-infinite spaces. Such states can be called localized.

Putting formally that $\chi_0(x) = \psi(x)$ and $\phi_0(x) = 0$ in Eq. (4.17) or $\chi_n(x) = \psi(x)$ and $\phi_n(x) = 0$ in Eq. (8.42), we find that the complete set of Lippmann–Schwinger equations for localized states consists of Eqs. (8.41), (4.18), and (4.19). By any of the three series of substitutions mentioned in Sec. V, we reduce these equations to a set of linear homogeneous algebraic equations with respect to $\psi(d_\mu)$ and $\Pi(d_\mu)$, $\mu = 0,1,\ldots,n-1$. Since the number of equations is equal to the number of unknowns, we obtain in a regular manner the secular equation for E [see Eq. (2.11)]. If the secular equation has a solution, $E = E(\mathbf{f}_\parallel)$, we can then find the unknowns $\psi(d_\mu)$ and $\Pi(d_\mu)$ up to an arbitrary constant factor, a_0. Inserting $\psi(d_\mu)$ and $\Pi(d_\mu)$ so obtained back into the original functional equations, we find $\psi(x)$, also up to a a_0 which is a normalization coefficient.

It follows from the above that the energy (or the frequency) of a localized state is a pole of the scattering amplitudes and of the total Green's function, in agreement with the general theory. For example, in the case of a three-layer medium, the secular equation is $S_{+1} = 0$ or, equivalently, $S_+ = 0$ [see Eqs. (7.9), (7.11), (7.12), and (7.14)]. If $U_2 \to \infty$, the equation becomes $S_+^{(\infty)} = 0$, in accordance with Eqs. (7.20) and (7.21).

The wave function of a localized state, $\psi(x)$, can obviously be written in the form of Eqs. (5.14)–(5.16), but one should put $\chi_0(x) \equiv \psi(x)$ in Eq. (5.14). In this case, the quantities $f_0^{(-1)x} = \bar{f}_0^{(-1)x}$ and $f_n^{(+1)x} = \bar{f}_n^{(+1)x}$ must simultaneously be complex. In other words, β_0 and β_n must be real and therefore E must satisfy the inequality

$$E < \min\{u_0 + U_0, u_n + U_n\}, \tag{9.1}$$

in accordance with Eqs. (2.11) and (3.16). In view of Eq. (3.12), the inequality (9.1) means that the energy of a localized state is less than the minimum of the unperturbed energies $\mathscr{E}_\mu(\mathbf{f}) + U_\mu$ [see Eq. (2.15)] with $\mathbf{f}_\parallel = \text{const}$ in both outer layers $(\mu = 0,n)$.

If both boundaries $x = d_0$ and $x = d_{n-1}$ are rigid, we have to omit Eqs. (8.41) and (4.19) and to combine the first $(\mu = 0)$ and the last $(\mu = n)$ of Eq. (4.18) with the equations $\psi(d_0) = 0$ and $\psi(d_{n-1}) = 0$, respectively. Solution of the set of $n-1$ equations so obtained gives *all* possible state of the field in the stratified flat slab $d_0 < x < d_{n-}$ with rigid boundaries. The eigenenergies (or eigenfrequencies) of the field can alternatively be found by poles of the corresponding Green's function.[1]

C. The one-dimensional motion

Setting $\mathbf{f}_\parallel = 0$ and $\epsilon_\mu^{xx} = \hbar^2/2m_\mu$ in the resulting relations of this paper, we obtain the solution of the problem of the one-dimensional motion of a particle with a piecewise-constant mass in a piecewise constant potential. If, in addition, $m_\mu = m$, $\mu = 0,1,\ldots,n$, we arrive at the conventional quantum-mechanical problem.[5]

X. CONCLUSION

We formulated a chainlike set of functional equations of the Lippmann–Schwinger type for the scattering amplitude of a plane wave in a medium of anisotropic flat layers in contact. The angle of incidence, the thicknesses of the layers and the orientations of the crystallographic axes of individual layers with respect to the interfaces are arbitrary. The functional equations obtained include the equation of motion for the scattering amplitude and the boundary conditions at all separation planes. The boundary conditions are represented in the Lippmann–Schwinger equations by certain constants which are to be eliminated in the course of solving the equations. In the case of a system with n separation planes, there are $2n$ such constants. Half of them are the values of the wave function at the separation planes. The rest are the values, at the same points, of a quantity which is, like the wave function, *continuous* at the separation planes.

We suggested a simple algorithm to solve the set of Lippmann–Schwinger equations. This algorithm is based on the idea of successive elimination of unknown constants and consists of n steps for a medium with n interfaces. It ends at the nth step, thus giving the solution. Also, we described an alternative, straightforward method of solving the Lippmann–Schwinger equations. In the framework of this method, one should evaluate the Lippmann–Schwinger equation for each given layer at the points corresponding to the layer boundaries. This gives a chainlike set of $2n$ linear algebraic equations for the same number of unknown constants. One of the equations is inhomogeneous while the remaining are homogeneous. The inhomogeneous equation corresponds to the region $x < d_0 (\mu = 0)$ or $x > d_{n-1} (\mu = n)$, in which the incident plane wave is given. Having found the unknown, constants, we insert them back into the original functional equations and thus find the scattering amplitude within each layer.

As far as we know, there are no publications where the problem of propagation of a plane wave in an anisotropic stratified medium is considered in the general form within the framework of the conventional method. This is due to the enormous amount of calculations needed in attacking the problem by the conventional method. The advantages of the proposed approach over the conventional one is obvious when both methods are applied particularly in the case of an isotropic system.

We recall that the conventional method (see, e.g., Ref. 7) deals directly with the amplitudes of secondary (reflected and transmitted) plane waves. There are generally one reflected plane wave in the zeroth layer ($x < d_0, \mu = 0$), two secondary plane waves in each inner layer ($d_{\mu-1} < x < d_\mu$, $\mu = 1,2,...,n-1$) and one transmitted wave in the last layer ($x > d_{n-1}, \mu = n$), $2n$ waves altogether. Thus the number of unknown wave amplitudes is the same as the number of auxiliary constants involved in the Lippmann–Schwinger equations. Nevertheless, by whichever method the set of Lippmann–Schwinger equations is solved, the approach based on these equations has essential advantages over and is much simpler than the conventional approach owing to several factors.

(i) The Lippmann–Schwinger equations contain, from the very beginning, some "finished blocks" in the form of combinations of the "standard" Green's functions.

(ii) The constants involved in the Lippmann–Schwinger equations are to be *eliminated* in the course of solving the equations and therefore play a subsidiary role. In contrast to this, the constants dealt with in the conventional method are the sought after scalar amplitudes. These are to be *found* in the course of solving the corresponding equations.

(iii) The set of Lippmann–Schwinger equations has, with respect to the unknown constants, a chainlike form with a very simple "matrix of coefficients". This facilitates the problem of eliminating the unknown constants even if their elimination is achieved through finding them. Also, the constants involved in the Lippmann–Schwinger equations are particular values of the quantities which are continuous at the interfaces. This is in constrast to the conventional method, where constants do not satisfy any such condition.

(iv) The method of Lippmann–Schwinger equations enables us to find, in the general form, the scattering amplitude, i.e., the total field at each point, with allowance of all existing separation planes. At the same time, we established general relations to express the amplitudes of individual scattered waves in terms of constants involved in the Lippmann–Schwinger equations. Thus, on solving the Lippmann–Schwinger equations, the amplitudes of all secondary waves can also be found.

Our approach is illustrated by solving the free field problem for a compound infinite space consisting of two anisotropic half-spaces in contact and for an arbitrary anisotropic three-layer medium (a flat slab and two half-spaces). As particular cases of such systems, we also consider an anisotropic homogeneous half-space with a rigid boundary and a compound half-space consisting of a flat slab with a rigid boundary and a homogeneous half-space in contact.

In addition to the direct methods of solving the Lippmann–Schwinger equations, we made use of these equations to establish general relations expressing the scattering amplitude in terms of the total Green's function. The latter can be found by solving a chainlike set of Dyson's equations.[1] Although the Dyson's equations are solved by the same methods as the Lippmann–Schwinger equations, the relations obtained are useful in several respects. Firstly, they express the solution of the free field problem in terms of the solution of the forced field problem for a point source. Secondly, the total Green's function seems to be more fundamentally related to the dynamic properties of the system because it can be used not only for finding the scattering amplitudes but also in solving many other problems (see e.g., Ref. 8). Lastly, the relations established give another constructive method to find the scattering amplitude.

Besides the free fields which can propagate throughout the system in the form of waves travelling at all possible angles with respect to the separation planes, and which are therefore nonlocalized, we also considered briefly localized states. The field in a localized state exists within one or several neighboring layers and can propagate only in directions parallel to the interfaces. The localized states gives rise to poles in the scattering amplitude at the associated frequencies as well as giving poles in the total Green's function. This fact enables us to find the amplitudes of localized states, up to a normalization constant, and their energies (frequencies) from the same set of Lippmann–Schwinger equations.

[1]Ya. A. Iosilevskii, Phys. Rev. B **19**, 856 (1979).
[2]B. A. Lippmann and J. Schwinger, Phys. Rev. **79**, 469 (1950).
[3]L. I. Schiff, *Quantum Mechanics*, 3rd ed. (McGraw-Hill, New York, 1968).
[4]Ya. A. Iosilevskii, Ann. Phys. (N.Y.) **107**, 360 (1977).
[5]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Pergamon, Oxford, 1965).
[6]Ya. A. Iosilevskii, Phys. Lett. A **65**, 92 (1978).
[7]L. M. Brekhovskikh, *Waves in Layered Media* (Academic, New York, 1960).
[8]Ya. A. Iosilevskii, Phys. Rev. B **19**, 1933 (1979).

Some topics in pre-Hilbert space

G. Epifanio
Istituto di Fisica dell'Università di Palermo, Palermo, Italy
C. Trapani
Istituto di Matematica dell'Università di Palermo, Palermo, Italy

(Received 26 January 1981; accepted for publication 26 June 1981)

We examine some questions concerning pre-Hilbert spaces and operators defined in them. Some classical results, which hold true in Hilbert space, are extended, under particular conditions, to noncomplete space.

PACS numbers: 03.65.Bz, 03.65.Fd, 02.10.Sp

INTRODUCTION

It is well known that the operators in quantum mechanics are, generally, unbounded; as a consequence, one has to deal with quite complicated problems of domain. There is a situation, however, when a simplification is possible; this occurs when all the operators, involved in the description of the physical system have, together with their adjoints, a common invariant dense domain $\mathscr{D}$. (See, for instance, the simple case of the harmonic oscillator, where such a domain can be understood to be the set

$$\mathscr{D} = \{ p(x)e^{-x^2/2} \text{with } p(x) \text{ polynomial} \}$$

or, better, the space $\mathscr{S}$ of the Schwartz test functions. See also the definition of Wightman field, etc.)

In this case, it is interesting to consider noncomplete scalar product spaces and operators defined in them.

In previous papers we considered this problem already and we studied, particularly, an interesting class of operators, defined in a pre-Hilbert space, which is in some cases and from some point of view, the natural algebra of observables and of the corresponding operators (see references in Ref. 1).

Completeness of a space is, on the other hand, a very strong property and when it fails it is not possible to recover many results, even some of the simplest ones, true in Hilbert space and for the operators defined in it.

The aim of this paper is to study further some properties of a scalar product space $\mathscr{D}$ and of the operators defined in it.

In Sec. 1 we examine some properties of pre-Hilbert space relative to orthocomplementation; in Sec. 2 we extend, so far as it is possible, some classical results, true for operators in $\mathscr{H}$, to a particular class of operators in $\mathscr{D}$.

1. WEAKLY CONTINUOUS FORMS IN SCALAR PRODUCT SPACE

In the sequel, we will indicate by $\mathscr{D}$ a scalar product space (pre-Hilbert space) and by $\mathscr{H}$ a complete scalar product space (Hilbert space). Beside, we denote by $\sigma(\mathscr{D},\mathscr{D})$ the usual weak topology in $\mathscr{D}$ and by $\tau(\mathscr{D},\mathscr{D})$ the Mackey topology in $\mathscr{D}$.

Theorems concerning projection operators and orthocomplemented subspace in Hilbert space are well known. These theorems, however, are not always true if the completeness of the space fails; in fact, in contrast with what happens in $\mathscr{H}$, not all norm-closed subspaces of $\mathscr{D}$ are orthocomplemented in $\mathscr{D}$.

We report here, for the reader's convenience, only the following proposition, true in pre-Hilbert space.

Proposition 1.1. Let $\mathscr{D}$ be a pre-Hilbert space and M a subspace of $\mathscr{D}$. The following statements are equivalent:

(i) M is $\sigma(\mathscr{D},\mathscr{D})$-closed,

(ii) M is $\tau(\mathscr{D},\mathscr{D})$-closed,

(iii) $M^{\perp\perp} = M$.

The proof of (i)$\Leftrightarrow$(ii) can be carried out from Ref. 2, prop. 35.2, and that of (ii)$\Leftrightarrow$(iii) from Ref. 2, Chap. 35, Cor. 2.

Notice that in pre-Hilbert space a norm-closed or $\sigma(\mathscr{D},\mathscr{D})$-closed subspace M cannot be orthocomplemented and $M^{\perp\perp} = M$ is only a necessary condition in order that M be orthocomplemented, but if M is maximal we can prove the following theorem. (For the definition of maximality see Ref. 3, Chap. I, Sec. 5.)

Theorem 1.2. Let M be a proper maximal subspace of $\mathscr{D}$. For M one, and only one, of the following statements holds true:

(i) M is orthocomplemented,

(ii) $M^{\perp} = 0$.

Proof: If $M^{\perp} \neq 0$, $M \oplus M^{\perp}$ is isomorphic to a subspace of $\mathscr{D}$ which contains M properly; for the maximality of M, $M \oplus M^{\perp}$ is isomorphic to the whole space $\mathscr{D}$, i.e., M is orthocomplemented.

Conversely, it is obvious that, if M is a proper orthocomplemented subspace of $\mathscr{D}$, then $M^{\perp} \neq 0$.

Corollary 1.3. If M is a $\sigma(\mathscr{D},\mathscr{D})$-closed proper and maximal subspace of $\mathscr{D}$, then it is orthocomplemented.

Proof: If M is not orthocomplemented, by Theorem 1.2, $M^{\perp} = 0$ and then $M^{\perp\perp} = \mathscr{D}$; from Proposition 1.1, it follows that $M = \mathscr{D}$ and this is not possible.

Corollary 1.4. Let F be a linear form on $\mathscr{D}$. F is continuous in the $\sigma(\mathscr{D},\mathscr{D})$-topology if, and only if, Ker F is an orthocomplemented subspace of $\mathscr{D}$.

Proof: If $F = 0$ the thesis is trivial. If $F \neq 0$ and Ker F is orthocomplemented then, by Proposition 1.1, it is $\sigma(\mathscr{D},\mathscr{D})$-closed and hence F is $\sigma(\mathscr{D},\mathscr{D})$-continuous.

Conversely, if F is $\sigma(\mathscr{D},\mathscr{D})$-continuous, Ker F is a proper closed maximal subspace of $\mathscr{D}$ (see Ref. 3, Chap. I, Sec. 5) and then, by Corollary 1.3, Ker F is orthocomplemented.

Theorem 1.5. Let $\mathscr{D}$ be a scalar product space. If every maximal norm-closed subspace is orthocomplemented, then $\mathscr{D}$ is complete.

Proof: It is sufficient to prove that every bounded form

 0022-2488/82/010039-04$02.50

is $\sigma(\mathscr{D},\mathscr{D})$-continuous; in fact, in this case, the weak dual of $\mathscr{D}$ coincides with the strong dual, which is the norm completion $\hat{\mathscr{D}}$ of $\mathscr{D}$. Hence $\mathscr{D}$ is complete.

But every bounded form F has the kernel maximal and norm-closed, hence, by the hypothesis, orthocomplemented, and by Proposition 1.1, $\sigma(\mathscr{D},\mathscr{D})$-closed; then F is $\sigma(\mathscr{D},\mathscr{D})$-continuous.

Notice that the above theorem was proved with the stronger requirement that M be only norm-closed (see Ref. 4).

2. SOME PROPERTIES OF OPERATORS IN SCALAR PRODUCT SPACE

We now consider a class of operators in a pre-Hilbert space $\mathscr{D}$ which is, in particular, a $*$-algebra having some analogies with the algebra $B(\mathscr{H})$ of bounded operators in Hilbert space; further, it coincides with $B(\mathscr{H})$ when $\mathscr{D}$ is chosen to be complete. We call this algebra $C_{\mathscr{D}}$.

Let $\mathscr{D}$ be a scalar product space; we indicate by the symbol $C_{\mathscr{D}}$ the $*$-algebra of all linear operators in $\mathscr{D}$, which have adjoint in $\mathscr{D}$, or equivalently, the $*$-algebra of all $\sigma(\mathscr{D},\mathscr{D})$-continuous operators. We denote by $B_{\mathscr{D}}$ the subalgebra of bounded operators of $C_{\mathscr{D}}$.

The algebra $C_{\mathscr{D}}$ can be understood to be the set of all closable operators A in $\mathscr{H}$ having $\mathscr{D}$ as a dense common invariant domain and such that $A^*(\mathscr{D})\subseteq\mathscr{D}$. The involution in $C_{\mathscr{D}}$ is then defined by $A\to A^+$ with $A^+=A^*{}_{\restriction\mathscr{D}}$.

For the definitions and theorems concerning $C_{\mathscr{D}}$, see references cited in Ref. 1.

Some authors also indicate the algebra $C_{\mathscr{D}}$ by the symbol $L_+(\mathscr{D})$ (see, for instance, Ref. 5).

By $\mathscr{A}\leqslant C_{\mathscr{D}}$ we will mean that $\mathscr{A}$ is an involutive subalgebra with unity of $C_{\mathscr{D}}$.

If $\mathscr{B}\subset C_{\mathscr{D}}$, we will indicate with $\mathscr{B}'$ the weak commutant of $\mathscr{B}$, i.e., the set

$$\mathscr{B}'=\{B\in B(\mathscr{H}):(S^+\varphi,B\psi)=(\varphi,BS\psi)\forall S\in\mathscr{B},\forall\varphi,\psi\in\mathscr{D}\}.$$

The commutants of higher order are defined in the usual way in $B(\mathscr{H})$; for instance

$$\mathscr{B}''=\{C\in B(\mathscr{H}):BC=CB\ \ \forall B\in\mathscr{B}'\}.$$

If $\mathscr{B}\subseteq C_{\mathscr{D}}$, we will denote by the notation $[\mathscr{B}]$ the subalgebra of $C_{\mathscr{D}}$ generated by $\mathscr{B}$. When $\mathscr{B}=\{A\}$ we will write $[A]$ instead of $[\{A\}]$.

In the sequel we call $\mathscr{D}$-self-adjoint an operator T of $C_{\mathscr{D}}$ such that $T=T^+$.

Theorem 2.1. Let $T\in C_{\mathscr{D}}$ be a $\mathscr{D}$-self-adjoint operator. Suppose that the following conditions are satisfied:

(a) T has a unique self-adjoint extension $\hat{T}$ to $\mathscr{H}$;

(b) there exists an algebra $\mathscr{A}\leqslant C_{\mathscr{D}}$ such that $[T]''\subseteq\mathscr{A}'$ and $\mathscr{D}$ is $\mathscr{A}$-self-adjoint (see Ref. 1, def. 2).

Let $\{E_\lambda\}$ be the spectral family associated with $\hat{T}$ and $u:R\to R$ a measurable function, finite and determined almost everywhere with respect to $\{E_\lambda\}$ such that

$$\mathscr{D}\subseteq\{\varphi\in\mathscr{H}:\int_{-\infty}^{\infty}u^2(\lambda)\,d(E_\lambda\varphi,\varphi)<\infty\},\tag{1}$$

letting

$$(u(T)\varphi,\psi)=\int_{-\infty}^{\infty}u(\lambda)\,d(E_\lambda\varphi,\psi),\quad \varphi,\psi\in\mathscr{D};\tag{2}$$

then (2) defines an operator $u(T)\in C_{\mathscr{D}}$.

Proof: It is known that (see, for instance, Ref. 6, n.127)

$$(u(\hat{T})\varphi,\psi)=\int_{-\infty}^{\infty}u(\lambda)\,d(E_\lambda\varphi,\psi),\quad \varphi,\psi\in\mathscr{H}$$

defines an operator $u(\hat{T})$ in $\mathscr{H}$. We will prove that $u(\hat{T})_{\restriction\mathscr{D}}\in C_{\mathscr{D}}$ [(1) implies that $\mathscr{D}\subseteq D(u(\hat{T}))$]. We now show that

$$\forall A\in\mathscr{A},\quad (A^+\varphi,u(\hat{T})\psi)=(\varphi,u(\hat{T})A\psi),\quad \forall\varphi,\psi\in\mathscr{D}.$$

In fact we have

$$\begin{aligned}(u(\hat{T})\psi,A^+\varphi)&=\int_{-\infty}^{\infty}u(\lambda)\,d(E_\lambda\psi,A^+\varphi)\\&=\int_{-\infty}^{\infty}u(\lambda)\,d(E_\lambda A\psi,\varphi)=(u(\hat{T})A\psi,\varphi).\end{aligned}$$

We used here the fact that $E_\lambda\in[T]''\subseteq\mathscr{A}'$ (see Ref. 1, Theorem 13) By Lemma 12 of Ref. 1, $u(T)$ is invariant in $\mathscr{D}$.

Corollary 2.2. Let $T\in C_{\mathscr{D}}$ be a $\mathscr{D}$-self-adjoint positive operator of $C_{\mathscr{D}}$. If the conditions (a) and (b) of Theorem 2.1 are satisfied, then there exists a unique operator $H\in C_{\mathscr{D}}$, $\mathscr{D}$-self-adjoint and positive, such that $H^2=T$.

Proof: We know that, if T satisfies the conditions (a) and (b) of Theorem 2.1, then it is $\mathscr{D}$-spectral, i.e., the spectral family $\{E_\lambda\}$ associated with it, takes its values in the same algebra $C_{\mathscr{D}}$ (see Ref. 1, Theorem 13).

Let $\{E_\lambda\}$ be the spectral family associated with T. Since T is positive, $E_\lambda=0$ for $\lambda<0$ results.

We pose

$$D(H)=\left\{\varphi\in\mathscr{H}:\int_0^\infty\lambda\,d(E_\lambda\varphi,\varphi)<\infty\right\}.$$

We will show that $\mathscr{D}\subseteq D(H)$. We observe that $(\hat{T}+I)_{\restriction\mathscr{D}}\in C_{\mathscr{D}}$ and therefore, for $\psi\in\mathscr{D}$, we have

$$\psi\in\left\{\varphi\in\mathscr{H}:\int_0^\infty(\lambda+1)^2\,d(E_\lambda\varphi,\varphi)\right\}<\infty.$$

Hence for such a ψ

$$\int_0^\infty\lambda^2\,d(E_\lambda\psi,\psi)+2\int_0^\infty\lambda\,d(E_\lambda\psi,\psi)+\int_0^\infty d(E_\lambda\psi,\psi)<\infty;$$

that is

$$||T\psi||^2+2\int_0^\infty\lambda\,d(E_\lambda\psi,\psi)+||\psi||^2<\infty$$

and so

$$\int_0^\infty\lambda\,\mathrm{d}(E_\lambda\psi,\psi)<\infty.$$

For $\varphi\in\mathscr{D}$ we set

$$H\varphi=\int_0^\infty\lambda^{1/2}\,dE_\lambda\varphi.$$

By Theorem 2.1, $H\in C_{\mathscr{D}}$. From a direct calculation $H^2=T$ results. The positivity and uniqueness of the operator H follows easily from the positivity and uniqueness of the spectral family.

Because $C_{\mathscr{D}}$ is a $*$-algebra no difficulty arises for the Cartesian decomposition. On the contrary, the polar decomposition, to be handled carefully for closed operators in $\mathscr{H}$ (Von Neumann's theorem, see Ref. 7, Chap. IV, Sec. 21.1), becomes more difficult for the operators of $C_{\mathscr{D}}$ and requires very strong hypotheses.

The above theorem allows us to state the following propositions.

Theorem 2.3. Let T be an operator of $C_{\mathscr{D}}$ such that T^+T satisfies the conditions of Corollary 2.2. Then

$$T = UH,$$

where $H = (T^+T)^{1/2}$ and U is an isometry from $R(H)$ into $R(T)$.

Proof: For $H = (T^+T)^{1/2}$, we define $U: R(H) \to R(T)$ by

$$U(H\varphi) = T\varphi \quad \forall \varphi \in \mathscr{D}.$$

Furthermore, we have

$$\| U(H\varphi)\|^2 = \|T\varphi\|^2 = (T\varphi, T\varphi) = (T^+T\varphi, \varphi)$$

$$= (H^2\varphi, \varphi) = (H\varphi, H\varphi) = \|H\varphi\|^2.$$

From the fact that $\|T\varphi\| = \|H\varphi\|$ it follows that $H\varphi = 0$ if and only if $T\varphi = 0$ and so U is a well-defined operator. It is thus proved that U is an isometry from $R(H)$ in $R(T)$.

Example. If we consider the creation and annihilation operators of the harmonic oscillator, in the space $\mathscr{S}$ of the Schwartz test functions

$$b = 2^{-1/2}(q + ip),$$
$$b^+ = 2^{-1/2}(q - ip),$$

the operator $b^+b = \frac{1}{2}(q^2 + p^2)$ satisfies, as was proved in Example 2 in the Appendix of Ref. 1, the conditions required by Theorem 2.3, and then b, or b^+, admits the polar decomposition.

It is not possible, generally, to say that the polar decomposition, given in Theorem 2.3, is unique. However, the following propositions hold true.

Theorem 2.4. In the hypotheses of Theorem 2.3, if $R(H)$ and $R(T)$ are orthocomplemented in $\mathscr{D}$, then the polar decomposition of the operator T of $C_{\mathscr{D}}$ is unique.

Proof: In this case U can be understood to be a partial isometry from $R(H)$ in $R(T)$. In fact, we set $U\varphi = 0$ for $\varphi \in R(H)^\perp$ and define $U^+\varphi = U^{-1}\varphi$ for $\varphi \in R(T)$ and $U^+\varphi = 0$ for $\varphi \in R(T)^\perp$.

Then $U \in C_{\mathscr{D}}$ and we have $T^+ = HU^+$ and $T^+T = HU^+UH = HP_{R(H)}H = H^2$ and therefore H, and consequently U, are uniquely determined.

Theorem 2.5. In the hypotheses of Theorem 2.3, if the operator T is invertible in $\mathscr{D}$, then the polar decomposition $T = UH$ is unique and U is a unitary operator of $C_{\mathscr{D}}$.

Proof: If T is invertible, so are T^+ and T^+T. We have $R(T^+T) = R(T^+) = \mathscr{D}$ and further $R(T^+T) \subseteq R(T^+T)^{1/2} = R(H)$. It follows that $R(H) = \mathscr{D}$. The uniqueness can be deduced immediately.

From the fact that $R(H) = \mathscr{D}$, U is everywhere defined and it is invertible, because $R(T) = \mathscr{D}$. It is easy to see that U is unitary.

We conclude with some propositions concerning the Cayley transform.

The existence of a unitary operator V, called the Cayley transform, for any self-adjoint operator of $B(\mathscr{H})$ is well known. It is also known that a closed symmetric operator T defined in a Hilbert space $\mathscr{H}$, admits a Cayley transform which is generally isometric; V is unitary if and only if T is self-adjoint in $\mathscr{H}$ (see, for instance, Ref. 6, n. 121 and n. 123).

With rather natural hypotheses, one can prove the existence of a Cayley transform for some $\mathscr{D}$-self-adjoint operators of $C_{\mathscr{D}}$.

Definition 2.6. Let T be an operator of $C_{\mathscr{D}}$. We call a resolvent set of T the subset of the complex field $\mathbb{C}$,

$$\rho_{\mathscr{D}}(T) = \{\lambda \in \mathbb{C} : \exists (T - \lambda I)^{-1} \in B_{\mathscr{D}}\}.$$

We call the spectrum of T the set $\sigma_{\mathscr{D}}(T) = \mathbb{C} - \rho_{\mathscr{D}}(T)$.

We have already discussed in a previous paper the convenience of giving such a definition of resolvent set for operators of $C_{\mathscr{D}}$ (see Ref. 1).

Definition 2.7. Let T be a self-adjoint operator of $C_{\mathscr{D}}$ such that $i \in \rho_{\mathscr{D}}(T)$. We call a Cayley transform of T the operator

$$V = (T - iI)(T + iI)^{-1}.$$

From Definition 2.6 it follows immediately that $V \in C_{\mathscr{D}}$.

Theorem 2.8. The Cayley transform V of a $\mathscr{D}$-self-adjoint operator T such that $i \in \rho_{\mathscr{D}}(T)$ is a unitary operator in $\mathscr{D}$. Also

$$T = i(I + V)(I - V)^{-1}.$$

Proof: The fact that V is unitary is straightforward. The relation $T = i(I + V)(I - V)^{-1}$ can be verified with easy algebraic calculations in $C_{\mathscr{D}}$, after having shown that the operator $I - V$ has an inverse in $C_{\mathscr{D}}$; this in turn can be proven in close analogy with the case of self-adjoint operators in $\mathscr{H}$.

The existence of the Cayley transform V for some $\mathscr{D}$-self-adjoint operator of $C_{\mathscr{D}}$ has some interesting implications, when it is considered as an operator in $\hat{\mathscr{D}}$. Notice that the operator T considered as an operator in $\hat{\mathscr{D}}$ cannot be self-adjoint; in fact, in this case its closure in $\hat{\mathscr{D}}$ would agree with T, and therefore $\mathscr{D}$ would be complete and $C_{\mathscr{D}}$ would coincide with $B(\hat{\mathscr{H}})$ (see Ref. 5, Lemma 2.2). Hence T is, as an operator in $\hat{\mathscr{D}}$, symmetric and its Cayley transform in $\hat{\mathscr{D}}$, by the general theory, should be isometric but, not necessarily, unitary. The above theorems allow us to state the following.

Theorem 2.9. A symmetric operator S in a Hilbert space $\mathscr{H}$ such that for some $\mathscr{D}$, dense in $\mathscr{H}$, the operator $S_{\upharpoonright \mathscr{D}} = T$ belong to $C_{\mathscr{D}}$ and such that $i \in \rho_{\mathscr{D}}(T)$ has a unitary Cayley transform in $\mathscr{H}$ and therefore it is self-adjoint.

Proof: It is sufficient to prove that the Cayley transform of S in $\mathscr{H}$ is the continuous extension U to $\mathscr{H}$ of the Cayley transform V of T in $\mathscr{D}$ and that U is unitary in $\mathscr{H}$.

In fact, by the continuity of the scalar product and that of the operators V and V^+, the adjoint U^* of the operator U in $\mathscr{H}$ is the continuous extension to $\mathscr{H}$ of the operator V^+ and, further, U is unitary in $\mathscr{H}$. If W is the Cayley transform of the operator S in $\mathscr{H}$

$$W = (S - iI)(S + iI)^{-1}$$

from the fact that $T - iI$ and $(T + iI)^{-1}$ are invariant in $\mathscr{D}$,

$W_{\restriction \mathscr{S}} = V$ results, and by the principle of extension of identities we have $W = U$ in $\mathscr{H}$.

From the above theorem one can deduce that a $\mathscr{D}$-self-adjoint operator $T \in C_{\mathscr{S}}$ such that $i \in \rho_{\mathscr{S}}(T)$ admits a self-adjoint extension. Besides

Corollary 2.10. A $\mathscr{D}$-self-adjoint operator T such that $i \in \rho_{\mathscr{D}}(T)$ admits a unique self-adjoint extension to $\hat{\mathscr{D}}$.

We prove only the uniqueness.

Let S be another self-adjoint extension of T to $\hat{\mathscr{D}}$; to it a unitary Cayley transform U is associated in $\hat{\mathscr{D}}$.

If we call V the Cayley transform of T in $\mathscr{D}$, because $T = S_{\restriction \mathscr{S}}, U_{\restriction \mathscr{S}} = V$ results and therefore $U = \hat{V}$; this implies that $\hat{T} = S$.

This theorem was recovered, in another way, in Ref. 1 (Prop. 11).

[1]G. Epifanio and C. Trapani, "Some spectral properties in algebras of unbounded operators," J. Math. Phys. **22**, 974 (1981).

[2]F. Treves, *Topological Vector Spaces, Distributions and Kernels* (Academic, New York, 1967).

[3]J. Horvath, *Topological Vector Spaces and Distributions* (Addison-Wesley, Reading, Mass., 1966).

[4]C. Piron, "Axiomatique Quantique," Helv. Phys. Acta **37**, 439 (1964).

[5]G. Lassner, "Topological algebras of operators," Rep. Math. Phys. **3**, 279 (1972).

[6]F. Riesz and B. Sz. Nagy, *Leçons d'analyse fonctionnelle* (Gauthier-Villars, Paris, 1965).

[7]M. A. Naimark, *Normed Rings* (P. Noordhoff N. V., Groningen, The Netherlands, 1964).

Symmetry of time-dependent Schrödinger equations. II. Exact solutions for the equation $\{\partial_{xx} + 2i\partial_t - 2g_2(t)x^2 - 2g_1(t)x - 2g_0(t)\}\Psi(x,t) = 0$

D. Rodney Truax
Department of Chemistry, University of Calgary, Calgary, Alberta, T2N 1N4, Canada

(Received 21 May 1981; accepted for publication 21 August 1981)

The maximal kinematical algebra of the Schrödinger equation $\{\partial_{xx} + 2i\partial_t - 2g_2(t)x^2 - 2g_1(t)x - 2g_0(t)\}\Psi(x,t) = 0$ is known to be the Schrödinger algebra, $\mathscr{S}_1$. The kinematical symmetries are realized as first-order differential operators in the space and time variables. A subalgebra $\mathscr{G}$ of $\mathscr{S}_1$ is chosen and from $\mathscr{G}$ and its invariants a complete set of commuting observables are constructed. The solution space of the Schrödinger equation is identified with the appropriate irreducible representation space of $\mathscr{G}$. The wave functions, simultaneous eigenvectors of the compatible observables, are computed as explicit functions of space and time. The properties of a system with a potential $V(x,t) = g_2(t)x^2 + g_1(t)x + g_0(t)$ are discussed.

PACS numbers: 03.65.Fd, 02.20. + b

1. INTRODUCTION

For quantum-mechanical systems in one spatial dimension with a Hamiltonian of the form

$$\mathscr{H}(x,t) = -\tfrac{1}{2}\partial_{xx} + V(x,t), \tag{1.1}$$

the time-dependent Schrödinger equation can be solved exactly only for a limited number of potentials $V(x,t)$. Two such cases are the time-dependent harmonic oscillator[1] where

$$V(x,t) = \omega^2(t)x^2/2, \tag{1.2}$$

and the harmonic oscillator subject to a purely time-dependent force[2] for which

$$V(x,t) = \omega^2x^2/2 + f(t)x. \tag{1.3}$$

In Eqs. (1.2) and (1.3) respectively, $\omega(t)$ and $f(t)$ are arbitrary real functions of time while in (1.3) ω is a real constant. The time-dependent harmonic oscillator has been used as a model to describe the behavior of the charged particle in a time-varying magnetic field [1b] or in quantum electronics to study parametric amplification and quantum noise.[3] The second case, Eq. (1.3), serves as a model of a vibrating diatomic molecule in the presence of a temporally-fluctuating force. In the sequel we shall investigate the generalization of these two models to systems governed by the time-dependent interaction

$$V(x,t) = g_2(t)x^2 + g_1(t)x + g_0(t), \tag{1.4}$$

where the functions $g_i(t)$, $1 \leqslant i \leqslant 3$ are arbitrary, real, and depend only on time. Our main objective is to solve the Schrödinger equation,[4]

$$\begin{aligned} Q\Psi(x,t) &= (-2\mathscr{H} + 2i\partial_t)\Psi(x,t) = 0 \\ &= \{\partial_{xx} + 2i\partial_t - 2g_2(t)x^2 \\ &\quad - 2g_1(t)x - 2g_0(t)\}\Psi(x,t) = 0, \end{aligned} \tag{1.5}$$

with Hamiltonian (1.1) and potential (1.4) and examine some of the properties of its solutions. A remarkable feature of our analysis is the independence of the form of the solutions from the specific nature of the functions $g_i(t)$, $1 \leqslant i \leqslant 3$, in (1.5). Another is that a system described by such a Schrödinger equation remains in a sense quantized. These points and others will be discussed in further detail in later sections.

In our approach to solving (1.5) we shall construct a complete set of commuting observables and determine their simultaneous eigenvectors and eigenvalues.[5] The observables will be elements of the algebra of the constants of the motion and its invariants. The constants of the motion are members of a Lie algebra called the kinematical or space-time symmetry algebra for (1.5). In the first paper in this series (denoted by I),[6] time-dependent Schrödinger equations were classified according to their maximal kinematical algebras[7,8] in which the generators of the space-time symmetries were realized as differential operators in the variables x and t. In particular, we showed that the Schrödinger equation (1.5) had the maximal kinematical algebra $\mathscr{S}_1 = \mathrm{sl}(2,\mathbb{R})\Box w_1$, where $\mathscr{S}_1$ called the Schrödinger algebra, is the semidirect sum of the Lie algebra $\mathrm{sl}(2,\mathbb{R})$ of the real special linear group in two dimensions[8] and the Lie algebra w_1, the Heisenberg–Weyl algebra in one dimension.[7,9] In Sec. 2 we choose an appropriate subalgebra $\mathscr{G}$ of $\mathscr{S}_1$ and determine its Casimir operators.[8] Then from $\mathscr{G}$ we select one generator and with the Casimir operators obtain a maximal set of commuting observables. The remaining two generators in $\mathscr{G}$ act as ladder operators.

We outline in detail in Sec. 3 the different irreducible representations of $\mathscr{G}$ and demonstrate the relationship between the normed solution space of (1.5) and the irreducible-representation spaces of $\mathscr{G}$. Further we show that the imposition of a norm on the representation spaces restricts us to the irreducible representation of $\mathscr{G}$ corresponding to that one in which the diagonalized observable from $\mathscr{G}$ has a discrete spectrum bounded below. The basis of eigenvectors for this irreducible representation of $\mathscr{G}$ is then a basis for the solution space of (1.5). The ladder operators in $\mathscr{G}$ step from one eigenvalue to another and in Sec. 4 we use them to construct the eigenfunctions as explicit functions of the space and time variables x and t.

The system is quantized with respect to the observables we have chosen to be diagonal in this representation. Except in the case of the time-independent oscillator where the diagonalized generator is the time-translation operator, the Hamiltonian is not a constant of the motion. Therefore the eigenvectors of the compatible observables will not be eigenvectors of the Hamiltonian $\mathscr{H}(x,t)$ and the energy will have a time dependence. In Sec. 5 we express the Hamiltonian and other properties in terms of the generators of $\mathscr{G}$ and calculate their expectation values. We conclude the paper in that section with a discussion of the features of time-dependent systems described by a potential of the form (1.4).

2. THE SYMMETRY ALGEBRA $\mathscr{G}$

The generators of kinematical symmetries have the general form

$$L = A(x,t)\partial_t + B(x,t)\partial_x + C(x,t) \tag{2.1}$$

and they are symmetries of the Schrödinger equation (1.5) if they satisfy the operator relation

$$[Q, L] = \lambda(x,t)Q, \tag{2.2}$$

where $\lambda(x,t)$ is a function of the variables x, t and Q is the differential operator defined by Eqs. (1.5). The symmetry L transforms solutions of (1.5) into solutions.[6,7,9]

In I we have shown that the maximal kinematical algebra for the differential equation (1.5) is the Schrödinger algebra $\mathscr{S}_1 = \mathrm{sl}(2,\mathbb{R})\square w_1$. The generators[10]

$$\left.\begin{aligned} B_1 &= -\chi_1(t)\partial_x + i\dot{\chi}_1(t)x - i\mathscr{C}_1(t),\\ B_2 &= \chi_2(t)\partial_x - i\dot{\chi}_2(t)x + i\mathscr{C}_2(t),\\ E &= i \end{aligned}\right\} \tag{2.3}$$

form a basis for the Heisenberg albegra w_1. The functions of time $\chi_1(t)$ and $\chi_2(t)$ are two real, linearly independent, nontrivial solutions of the homogeneous second-order differential equation

$$\ddot{b} + 2g_2(t)b = 0. \tag{2.4}$$

The solutions χ_1 and χ_2 have constant Wronskian[6] and have been chosen so that $W(\chi_1,\chi_2) = \chi_1\dot{\chi}_2 - \dot{\chi}_1\chi_2 = 1$. The functions $\mathscr{C}_1(t)$ and $\mathscr{C}_2(t)$ are defined by

$$\mathscr{C}_\sigma(t) = \int^t g_1\chi_\sigma, \quad \sigma = 1, 2. \tag{2.5}$$

Because the Wronskian of χ_1 and χ_2 is unity, the generators (2.3) satisfy the commutation relations

$$[B_1, B_2] = E, \quad [B_1, E] = [B_2, E] = 0, \tag{2.6}$$

and furthermore

$$[Q, E] = [Q, B_\sigma] = 0, \quad \sigma = 1, 2. \tag{2.7}$$

The generators which form a basis for the sl(2, $\mathbb{R}$) algebra are

$$\begin{aligned} L_j = \varphi_j\partial_t + (\tfrac{1}{2}\dot{\varphi}_j x + \mathscr{A}_j)\partial_x - (i/4)\ddot{\varphi}_j x^2 - i\dot{\mathscr{A}}_j x + \tfrac{1}{4}\dot{\varphi}_j \\ + ig_0\varphi_j + i\mathscr{D}_j, \quad 1 \leqslant j \leqslant 3. \end{aligned} \tag{2.8}$$

The functions $\varphi_j(t)$, $1 \leqslant j \leqslant 3$, are functions of time only and are linearly independent, nontrivial solutions of the third order, homogeneous differential equation

$$\dddot{A} + 8g_2\dot{A} + 4\dot{g}_2 A = 0. \tag{2.9}$$

In I we showed that the functions φ_j which have constant Wronskian may be chosen as follows[6]:

$$\varphi_1 = \chi_1^2, \quad \varphi_2 = \chi_2^2, \quad \varphi_3 = 2\chi_1\chi_2, \tag{2.10}$$

in which case

$$\begin{aligned} &\mathscr{A}_1 = -\chi_1\mathscr{C}_1, \quad \mathscr{A}_2 = -\chi_2\mathscr{C}_2,\\ &\mathscr{A}_3 = -(\chi_1\mathscr{C}_2 + \chi_2\mathscr{C}_1), \end{aligned} \tag{2.11a}$$

$$\mathscr{D}_1 = -\tfrac{1}{2}\mathscr{C}_1^2, \quad \mathscr{D}_2 = -\tfrac{1}{2}\mathscr{C}_2^2, \quad \mathscr{D}_3 = -\mathscr{C}_1\mathscr{C}_2, \tag{2.11b}$$

where the χ_σ are the solutions of (2.4) and the $\mathscr{C}_\sigma$, $\sigma = 1, 2$, are defined by (2.5). The L_j have the commutation relations

$$[L_1, L_2] = L_3, \quad [L_3, L_1] = -2L_1, \quad [L_3, L_2] = 2L_2, \tag{2.12}$$

and furthermore satisfy

$$[Q, L_j] = \dot{\varphi}_j Q. \tag{2.13}$$

The commutation relations between the generators of the sl(2, $\mathbb{R}$) and w_1 algebras are[6]

$$\left.\begin{aligned} &[L_1, B_1] = 0, \quad &&[L_2, B_1] = B_2, \quad &&[L_3, B_1] = -B_1,\\ &[L_1, B_2] = -B_1, \quad &&[L_2, B_2] = 0, \quad &&[L_3, B_2] = B_2. \end{aligned}\right\} \tag{2.14}$$

Now, let $\mathscr{F}_Q$ denote the solution space of the Schrödinger equation (1.5), that is if $f(x,t) \in \mathscr{F}_Q$, then $Qf(x,t) = 0$. We shall require that the functions in $\mathscr{F}_Q$ be square integrable with time-independent, bounded norm[5,11]

$$0 \leqslant \int_{-\infty}^{+\infty} f^*(x,t)f(x,t)dx < \infty, \quad \text{for all } f(x,t) \in \mathscr{F}_Q. \tag{2.15}$$

The first inequality follows from the definition of a norm and equality holds only for the trivial solution $f(x,t) = 0$. Furthermore, let A be some linear operator defined on $\mathscr{F}_Q$. Then $\langle A \rangle$, the expectation value[5] of A, is

$$\langle A \rangle = \int_{-\infty}^{+\infty} f^*(x,t)Af(x,t)\,dx. \tag{2.16}$$

Although the operator A may not have an explicit time dependence, $\langle A \rangle$ may depend on time.

The general equation giving the time dependence of the mean value of a linear operator A is[5b]

$$i\frac{d}{dt}\langle A \rangle = i\left\langle \frac{\partial A}{\partial t} \right\rangle + \langle [A, H] \rangle. \tag{2.17}$$

If $d\langle A \rangle/dt = 0$, then A is said to be a constant of the motion, that is if $\langle A \rangle$ is the expectation value of a linear operator A at some point in time then this expectation value remains unchanged in the course of time. By (2.7) and (2.13), for $f(x,t) \in \mathscr{F}_Q$, we have $d\langle L_j \rangle/dt = 0$, $1 \leqslant j \leqslant 3$ and $d\langle B_\sigma \rangle/dt = d\langle E \rangle/dt = 0$, $\sigma = 1, 2$. So on the solution space $\mathscr{F}_Q$, all the generators of $\mathscr{S}_1$ are constants of the motion. Notice that the Hamiltonian (1.1), when it has an explicit time dependence, is not a constant of the motion.

Let us choose now a subalgebra of $\mathscr{S}_1$ consisting of the symmetries $\{L_3, B_1, B_2, E\}$ and denote this subalgebra by $\mathscr{G}$. The commutation relations of the generators of $\mathscr{G}$ are given by (2.6) and (2.14). The Casimir operators of $\mathscr{G}$ are[12]

$$C = B_2B_1 - L_3E \text{ and } E, \tag{2.18}$$

where $[C, L_3] = [C, B_1] = [C, B_2] = [C, E] = 0$. Since on $\mathscr{F}_Q$, the elements of $\mathscr{G}$ are constants of the motion, the Casimir operator C will be also. Using the appropriate expressions from (2.3) and (2.8) for the generators of $\mathscr{G}$ we can derive the relation

$$C = -\tfrac{1}{2}(\varphi_3 Q + i). \tag{2.19}$$

The motivation for our choice of the subalgebra $\mathscr{G}$ becomes clear. If f is any member of $\mathscr{F}_Q$, then $Cf = (-i/2)f$ and $Ef = if$ and so the Casimir operators act as multiples of the unit operator on elements of $\mathscr{F}_Q$. This implies a direct relationship between the solution space $\mathscr{F}_Q$ of (1.5) and the irreducible representation spaces $\mathscr{V}$ of the Lie algebra $\mathscr{G}$ as we shall show more rigorously in Sec. 3.

Before we construct the irreducible representation spaces $\mathscr{V}$ it is necessary to choose the compatible observables we shall simultaneously diagonalize. We pick the Casimir operators C and E and the generator L_3 from $\mathscr{G}$. Now to be observables these linear operators must be made Hermitian with respect to the norm (2.15). Furthermore, it would be convenient to have the remaining generators in $\mathscr{G}$ Hermitian conjugates. One way to do this is to define new solutions to (2.4) by taking the linear combinations of the real functions χ_1 and χ_2,

$$\xi(t) = (2^{-1/2})(\chi_1(t) + i\chi_2(t)),$$
$$\xi^*(t) = (2^{-1/2})(\chi_1(t) - i\chi_2(t)), \tag{2.20}$$

that is, we take the complexification of the solution space of (2.4). These new solutions have the Wronskian

$$W(\xi, \xi^*) = \xi\dot{\xi}^* - \dot{\xi}\xi^* = -i. \tag{2.21}$$

If we repeat the analysis in I, then we have

$$\varphi_1 = \xi^2, \quad \varphi_2 = (\xi^*)^2, \quad \varphi_3 = 2\xi\xi^* \tag{2.22}$$

and

$$\mathscr{A}_1 = -\xi\mathscr{C}, \qquad \mathscr{A}_2 = -\xi^*\mathscr{C}^*, \tag{2.23a}$$
$$\mathscr{A}_3 = -(\xi\mathscr{C}^* + \xi^*\mathscr{C}),$$
$$\mathscr{D}_1 = -\tfrac{1}{2}\mathscr{C}^2, \quad \mathscr{D}_2 = -\tfrac{1}{2}(\mathscr{C}^*)^2, \quad \mathscr{D}_3 = -\mathscr{C}\mathscr{C}^*, \tag{2.23b}$$

where

$$\mathscr{C} = \int^t g_1\xi, \quad \mathscr{C}^* = \int^t g_1\xi^*. \tag{2.24}$$

Note that φ_3, $\mathscr{A}_3$, and $\mathscr{D}_3$ are all real functions and the new operator L_3 defined by (2.22) and (2.23) is skew-Hermitian. Therefore, if we take $M_3 = iL_3$, M_3 will be Hermitian. The generators of $\mathscr{G}$ will be of the form

$$M_3 = i\{\varphi\partial_t + (\tfrac{1}{2}\dot{\varphi}x + \mathscr{A})\partial_x - (i/4)\ddot{\varphi}x^2 - i\dot{\mathscr{A}}x + \tfrac{1}{4}\dot{\varphi} + ig_0\varphi + i\dot{\mathscr{D}}\}, \tag{2.25a}$$
$$J_+ = -\xi^*\partial_x + i\dot{\xi}^*x - i\mathscr{C}^*, \tag{2.25b}$$
$$J_- = \xi\partial_x - i\dot{\xi}x + i\mathscr{C}, \tag{2.25c}$$
$$E = 1, \tag{2.25d}$$

where we have dropped the subscript 3 from the time-dependent functions φ_3, $\mathscr{A}_3$, and $\mathscr{D}_3$ in (2.25a) for future convenience. These linear operators satisfy the commutation relations

$$[J_-, J_+] = E, \quad [M_3, J_+] = J_+, \quad [M_3, J_-] = -J_-. \tag{2.26}$$

The Casimir operator C is

$$C = J_+J_- - M_3E = -\tfrac{1}{2}(\varphi Q + 1). \tag{2.27}$$

Both C and E are Hermitian operators; the linear operators J_+ and J_- are Hermitian conjugates, i.e.,

$$J_-^\dagger = J_+. \tag{2.28}$$

The set of commuting observables we shall simultaneously diagonalize are the invariants C and E and the generator M_3 from $\mathscr{G}$. The Hermitian conjugates J_+ and J_- will act as ladder operators, stepping the eigenvalues of M_3.

The time independence of the norm (2.15) will be guaranteed by requiring the Hamiltonian to be Hermitian,[5b] hence the requirement that the $g_i(t)$ be real-valued functions of time in (1.4) and (1.5).

3. REPRESENTATIONS OF $\mathscr{G}$

The irreducible representations of $\mathscr{G}$ have been computed by Miller[12] and we reproduce briefly his analysis here to provide a framework for the remaining work. Also, we shall see that because of the imposition of the norm (2.15) on $\mathscr{F}_Q$ and our identification of $\mathscr{F}_Q$ with the irreducible representation spaces of the Lie algebra $\mathscr{G}$, not all of the irreducible representations are pertinent.

We proceed by outlining pointwise the computation of irreducible representations of $\mathscr{G}$.

(i) From Sec. 2 we have that $\mathscr{G}$ consists of the generators $\{M_3, J_+, J_-, E\}$ where these are realized as differential operators in space and time variables in (2.25). Their commutation relations are given by Eqs. (2.26). The Casimir operators are

$$C = J_+J_- - M_3E, \quad E = 1. \tag{3.1}$$

The operators C, E, and M_3 are Hermitian operators in our Hilbert space and so must have real eigenvalues; since they commute we can define a set of states which are simultaneously eigenstates of these three observables.

If $\mathscr{V}$ is a representation space then it is sufficient[12] to consider only those irreducible representations in which M_3 has nondegenerate eigenvalues and in which the representation space $\mathscr{V}$ has a countable basis consisting of eigenvectors of M_3.

(ii) If S is the spectrum of M_3, then S is countable and there exists a basis for $\mathscr{V}$ of vectors f_s such that

$$M_3 f_s = s f_s, \quad \text{for all} \quad s \in S. \tag{3.2}$$

Choosing any $s \in S$,

(a) $[M_3, J_+] = J_+ \Rightarrow M_3 J_+ f_s = (s+1)J_+ f_s$. So either $J_+ f_s = \theta_{s+1} f_{s+1}$, where θ_{s+1} is a nonzero constant for $s+1 \in S$, or $J_+ f_s = 0$.

(b) $[M_3, J_-] = -J_- \Rightarrow M_3 J_- f_s = (s-1)J_- f_s$. So either $J_- f_s = \eta_s f_{s-1}$, where η_s is a nonzero constant for $s-1$ in S, or $J_- f_s = 0$.

(c) $[E, M_3] = 0 \Rightarrow E f_s = \mu_s f_s$, for some real constant μ_s.

(d) $[C, M_3] = 0 \Rightarrow C f_s = \lambda_s f_s$ for some real constant λ_s.

Since the representation is irreducible the spectrum of S is connected, i.e., if s lies in S, then

$$S = \{s + n : n \in \mathbb{Z} \quad \text{such that } n_1 < n < n_2\}, \tag{3.3}$$

where $\mathbb{Z}$ is the set of integers and we do not rule out the possibility that $n_1 = -\infty$ and $n_2 = +\infty$.

(iii) For any s in S such that $s+1$ lies in S,

(a) $[E, M_+] = 0 \Rightarrow [E, M_+] f_s = \theta_{s+1}(\mu_{s+1} - \mu_s) f_{s+1} = 0 \Rightarrow \mu_s = \mu_{s+1} = \mu$, and so $E = \mu I$, where I is the unit operator on $\mathscr{V}$. In our case, because of (3.1), $\mu = 1$.

(b) $[C, M_+] = 0 \Rightarrow [C, M_+] f_s = \theta_{s+1}(\lambda_{s+1} - \lambda_s) f_{s+1} = 0 \Rightarrow \lambda_{s+1} = \lambda_s = \lambda$, and $C = \lambda I$ on $\mathscr{V}$, where λ is a real number.

If C is given by (3.1), then the action of C on f_s, for $s \in S$ yields the relation

$$\theta_s \eta_s = \lambda + s\mu = \mu(\alpha + s), \tag{3.4}$$

which is valid for all s or $s-1$ in S and for convenience we set $\lambda = \mu\alpha$. The number α will of course be real.

(iv) Since each $f_s \in \mathscr{V}$ for $s \in S$ is simultaneously an eigenvector of C, E, and M_3 with eigenvalues λ, μ, and s, respectively, we could label f_s with all three eigenvalues. However, since λ and μ are the same for each $s \in M_3$, the extra labels are in a sense redundant and have been omitted.

We can always define a new basis $\{f'_s\}$ for $\mathscr{V}$ by means of the set of nonzero constants $\{\gamma_s : s \in S\}$ such that $f'_s = \gamma_s f_s$ for each s in S. In this new basis, we have

$$\left.\begin{array}{lll} Cf'_s = \lambda f'_s, & Ef'_s = \mu f'_s, & M_3 f'_s = s f'_s, \\ J_+ f'_s = \theta'_{s+1} f'_{s+1}, & J_- f'_s = \eta'_s f'_{s-1}, & \end{array}\right\} \tag{3.5}$$

where

$$\theta'_s = \frac{\gamma_{s-1}}{\gamma_s}\theta_s, \quad \eta'_s = \frac{\gamma_s}{\gamma_{s-1}}\eta_s,$$

assuming $\theta_{s+1} = 0$ if $s+1$ is not in S and $\eta_s = 0$ if $s-1$ does not lie in S. The constants θ'_s, η'_s must satisfy $\theta'_s \eta'_s = \mu(\alpha + s)$ for each s in S. So, for all $s \in S$ such that $s - 1 \in S$, it is possible to select the constants η_s arbitrarily and define the constants θ_s by (3.4). Thus the irreducible representation of $\mathscr{G}$ is uniquely determined by the eigenvalues λ and μ of C and E, respectively, and the spectrum S of M_3. The constants θ_s and η_s are not unique and may be selected arbitrarily subject only to (3.4).

(v) Every representation of $\mathscr{G}$ which satisfies (i) above and for which $E \neq 0$, is isomorphic to one of the following representations:

(a) The representations $R(\alpha, s_0, \mu)$ defined for all real α, s_0, and μ such that $\mu \neq 0$, $0 < s_0 < 1$, and $\alpha + s_0$ is not an integer. The spectrum $S = \{s_0 + n : n \in \mathbb{Z}\}$.

(b) The representations $\uparrow_{\alpha,\mu}$ defined for all real α and μ such that $\mu \neq 0$. The spectrum $S = \{-\alpha + n : n \in \mathbb{Z}_0^+\}$ where $\mathbb{Z}_0^+$ is the set of positive integers including zero. For this case $\alpha + s \in \mathbb{Z}_0^+$.

For (a) and (b) there exists a basis for $\mathscr{V}$ consisting of vectors f_s defined for each $s \in S$ such that

$$\left.\begin{array}{lll} f_s = \mu\alpha f_s, & Ef_s = \mu f_s, & M_3 f_s = s f_s, \\ J_+ f_s = \mu f_{s+1}, & J_- f_s = (s+\alpha) J_{s-1}. & \end{array}\right\} \tag{3.6}$$

(c) The representation $\downarrow_{\alpha,\mu}$ defined for all real α and μ such that $\mu \neq 0$ and $s + \alpha$ is a negative integer. The spectrum $S = \{-\alpha - 1 - n : n \in \mathbb{Z}_0^+\}$. For each representation there is a basis for $\mathscr{V}$ of vectors f_s defined for all s in S such that

$$\left.\begin{array}{lll} Cf_s = \mu\alpha f_s, & Ef_s = \mu f_s, & M_3 f_s = s f_s, \\ J_+ f_s = (\alpha + s + 1) f_{s+1}, & J_- f_s = \mu f_{s-1}. & \end{array}\right\} \tag{3.7}$$

The representations $\uparrow_{\alpha,\mu}$ are bounded below; the representations $\downarrow_{\alpha,\mu}$ are bounded above, and $R(\alpha, s_0, \mu)$ is unbounded.

Now we shall explore more thoroughly the connection between the representation space $\mathscr{V}$ of $\mathscr{G}$ and the solution space $\mathscr{F}_Q$ of the time-dependent Schrödinger equation (1.5). Let f be any function in $\mathscr{F}_Q$. Then $Qf = 0$ and we have

$$Cf = -\tfrac{1}{2}(\varphi Q + 1) f = -\tfrac{1}{2} f \text{ and } Ef = f.$$

Since $\mu = 1$, $\lambda = \mu\alpha = -\frac{1}{2}$ and f is an element of the vector space, the representation space $\mathscr{V}$. Conversely, let f_s be any basis vector in $\mathscr{V}$. Then

$$Cf_s = -\tfrac{1}{2}(\varphi Q + 1) f_s = \mu\alpha f_s, \quad Ef_s = \mu f_s = f_s,$$

and f_s is in $\mathscr{F}_{Q'}$. for some Q', where

$$Q' = \partial_{xx} + 2i\partial_t - 2g_2(t)x^2 - 2g_1(t)x - 2g'_0(t)$$

and

$$g'_0(t) = g_0(t) - (2\alpha + 1)/2\varphi.$$

In particular, if $\alpha = -\frac{1}{2}$, then f_s is in $\mathscr{F}_Q$, i.e., $g'_0(t) = g_0(t)$. This implies that we can use the structure of the Lie algebra $\mathscr{G}$, a symmetry algebra of the Schrödinger equation (1.5), and the irreducible representations of $\mathscr{G}$ to define the solution space $\mathscr{F}_Q$.

Now, the requirement that our space $\mathscr{F}_Q$ be normed imposes restrictions upon which of the three irreducible representations $R(-\frac{1}{2}, s_0, 1)$, $\uparrow_{-\frac{1}{2},1}$, and $\downarrow_{-\frac{1}{2},1}$ are acceptable in defining the solution space $\mathscr{F}_Q$. Since

$$0 \leqslant \int_{-\infty}^{+\infty} f_s^* f_s \, dx < \infty, \text{ for any } s \text{ in } S, f_s \text{ in } \mathscr{V}, \tag{3.8}$$

then for s and $s-1$ in S,

$$\int_{-\infty}^{+\infty} f_s^* J_+ J_- f_s \, dx = \int_{-\infty}^{+\infty} (J_- f_s)^* (J_+ f_s) \, dx = \eta_s^* \eta_s \int_{-\infty}^{+\infty} f_{s-1} f_{s-1} \, dx > 0, \tag{3.9}$$

where the first equality follows from (2.28), the second from (ii)(b) above, and the inequality because of (3.8). Alternatively

$$\int_{-\infty}^{+\infty} f_s^* J_+ J_- f_s \, dx = \theta_s \eta_s \int_{\infty}^{\infty} f_s^* f_s dx \geqslant 0,$$

by (ii)(a), (ii)(b), and (3.9). Because of equations (3.4) and (3.8),

$$\mu(\alpha + s) = (\alpha + s) \geqslant 0. \tag{3.10}$$

Since $(\alpha + s) \geqslant 0$, we can eliminate the representations $\downarrow_{-\frac{1}{2},1}$ and $R(-\frac{1}{2}, s_0, 1)$ from further consideration at this point since for both of them $(\alpha + s) < 0$ for certain choices of s. This leaves only the irreducible representation $\uparrow_{-\frac{1}{2},1}$, which is bounded below.

Since $\alpha = -\frac{1}{2}$, Eq. (3.11) implies that the eigenvalues of M_3 satisfy the inequality $s \geqslant \frac{1}{2}$, and the spectrum S of M_3 is $\{n + \frac{1}{2} : n \in \mathbb{Z}_0^+\}$. Then by (3.6), the irreducible representation space $\mathscr{V}$ and the solutions space $\mathscr{F}_Q$ of (1.5) are spanned by the set of eigenvectors satisfying the identities

$$\left.\begin{array}{ll} Cf_{n+1/2} = -\frac{1}{2} f_{n+1/2}, & Ef_{n+1/2} = f_{n+1/2}, \\ M_3 f_{n+1/2} = (n + \frac{1}{2}) f_{n+1/2}, & \\ J_+ f_{n+1/2} = f_{n+3/2}, & J_- f_{n+1/2} = n f_{n-1/2}. \end{array}\right\} \tag{3.11}$$

This choice of basis is not the most convenient and so making use of (iv) we can choose a new set of normalized eigenvectors relabelled with the quantum number n rather than the eigenvalue $n+\frac{1}{2}$. We denote the new normalized eigenvectors by h_n. Let $h'_n = f_{n+1/2}$ and assume that $h_0 = h'_0$ is normalized. Then $h'_n = J^n_+ h_0$. We choose a set of real constants $\{\gamma_n : n\in\mathbb{Z}_0^+\}$ such that $h_n = \gamma_n h'_n$ and,

$$\int_{-\infty}^{+\infty} h^*_n h_n dx = 1. \tag{3.12}$$

Using (2.28) and the commutation relation $[J_-, J_+] = E$ we find that $\gamma_n = (n!)^{-1/2}$. In the new basis, the relations (3.11) become

$$\left.\begin{aligned} &Ch_n = -\tfrac{1}{4}h_n, \qquad &&Eh_n = h_n,\\ &M_3 h_n = (n+\tfrac{1}{2})h_n &&\\ &J_+h_n = (n+1)^{1/2}h_{n+1}, \qquad &&J_-h_n = n^{1/2}h_{n-1}.\end{aligned}\right\} \tag{3.13}$$

Since M_3 is Hermitian, the eigenvectors h_n at the same time are orthogonal, that is,

$$\int_{-\infty}^{+\infty} h^*_m(x,t)h_n(x,t)dx = 0, \quad m\neq n. \tag{3.14}$$

Another important feature is that any time-dependent system which can be described by Eq. (1.5) is quantized; the quantum numbers arise naturally when we diagonalize the constant of the motion M_3 a generator of a space-time symmetry of (1.5). We shall discuss these points further in the concluding Sec. 5 and only mention them out of interest here.

We remark that the representation space $\uparrow_{-1/2,1}$ of solutions $\mathscr{F}_Q$ defined by (3.13) is equivalent to that obtained by Lewis and Reisenfeld [1b] for the restricted case $V(x,t) = \omega^2(t)x^2/2$.

4. EXPLICIT SOLUTIONS OF THE SCHRÖDINGER EQUATION $Qf = 0$

For many properties which can be expresses as polynomials in momentum and position operators we can rewrite them as polynomials in the ladder operators $J_\pm$ and M_3 and use the relations (3.13) and (2.16) to calculate expectation values. However, in other cases, and especially when we wish to explore the behavior of the system at different times, it would be advantageous to know the wave functions as explicit functions of the space and time variables. In this section we shall make use of the realization (2.25) for the generators of $\mathscr{G}$ and the properties (3.13) to obtain h_n as a function of x and t. We can do this regardless of the form of the time-dependent functions $g_i(t)$, $1\leqslant i\leqslant 3$ in (1.5) by making good use of the identities (2.22), (2.23), (2.24), and the Wronskian (2.21).

We begin by solving the eigenvalue problem

$$M_3 h_n = (n+\tfrac{1}{2})h_n.$$

Substituting (2.25a) for M_3 yields a first order partial differential equation for $h_n(x,t)$:

$$\varphi h_{n,t} + (\tfrac{1}{2}\dot{\varphi}x + \mathscr{A})h_{n,x} + (-(i/4)\ddot{\varphi}x^2 - i\dot{\mathscr{A}}x + \tfrac{1}{4}\dot{\varphi} + ig_0\varphi + i\mathscr{D} + i(n+\tfrac{1}{2}))h_n = 0, \tag{4.1}$$

which may be solved by integrating the subsidiary equations

$$\frac{dt}{\varphi} = \frac{d_x}{(-\frac{1}{2}\dot{\varphi}x + \mathscr{A})} = -\frac{dh_n}{(-(i/4)\ddot{\varphi}x^2 - i\dot{\mathscr{A}}x + \frac{1}{4}\dot{\varphi} + ig_0\varphi + i\mathscr{D} + i(n+\frac{1}{2}))h_n}. \tag{4.2}$$

Details of this calculation are supplied in Appendix B. Hence

$$h_n(x,t) = \varphi^{-1/2}\exp i\{x^2\dot{\varphi}/4\varphi + x\mathscr{A}/\varphi - \Lambda - G_0 - (n+\tfrac{1}{2})\Phi\}a_n\Big(\frac{x}{\varphi^{1/2}} - \mathscr{B}\Big), \tag{4.3}$$

where the function a_n is, for the moment, an arbitrary function of its argument, the new variable $x/\varphi^{1/2} - \mathscr{B}$. Recall that φ and $\mathscr{A}$ are defined by (2.22), where $\varphi = \varphi_3$ and by (2.23a), where $\mathscr{A} = \mathscr{A}_3$ and

$$\mathscr{B} = \int^t \frac{\mathscr{A}}{\varphi^{3/2}}, \tag{4.4}$$

$$\Phi = \int^t \frac{1}{\varphi}, \tag{4.5}$$

$$G_0 = \int^t g_0, \tag{4.6}$$

$$\Lambda = \int^t \frac{\mathscr{A}^2}{\varphi^2} + \int^t \frac{\mathscr{D}}{\varphi}, \tag{4.7}$$

where $\mathscr{D}$ is given by (2.23b) with $\mathscr{D} = \mathscr{D}_3$. Note that all the quantities in (4.4) through (4.7) are real-valued functions of time.

Let $\xi = x/\varphi^{1/2} - \mathscr{B}$ and define the operators

$$Z = \frac{1}{\sqrt{2}}\Big(\frac{d}{d\xi} + \xi\Big), \quad Z^+ = \frac{1}{\sqrt{2}}\Big(-\frac{d}{d\xi} + \xi\Big), \tag{4.8}$$

which satisfy the commutation relations

$$[Z, Z^+] = E. \tag{4.9}$$

Then we have the relationships

$$\begin{aligned} J_+h_n(x,t) &= \varphi^{-1/4} \times \exp i\Big\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \Big(n+\frac{1}{2}\Big)\Phi\Big\} \times \frac{\xi^{*1/2}}{\xi}Z^+a_n \\ &= \varphi^{-1/4} \times \exp i\Big\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \Big(n+\frac{3}{2}\Big)\Phi\Big\} \times Z^+a_n, \end{aligned} \tag{4.10}$$

where we have used the identity (see Appendix B)

$$\Phi = (i/2)\ln(\xi^*/\xi). \tag{4.11}$$

For $J_-h_n(x,t)$ we obtain

$$\begin{aligned} J_+h_n(x,t) &= \varphi^{-1/4} \times \exp i\Big\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - (n+\tfrac{1}{2})\Phi\Big\}\frac{\xi^{1/2}}{\xi^*}Za_n \\ &= \varphi^{-1/4} \end{aligned}$$

$$\times \exp i\left\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - (n - \tfrac{1}{2})\Phi\right\} Za_n. \tag{4.12}$$

We have made use of (4.11) in arriving at Eq. (4.12).

Now, from (3.13) we have $J_- h_0 = 0$ which implies $Za_0(\zeta) = 0$. The latter condition gives the following ordinary differential equation for a_0:

$$\frac{da_0}{d\zeta} + \zeta a_0 = 0$$

which has the solution

$$a_0 = Ne^{-\zeta^2/2}, \tag{4.13}$$

where N is a real constant. We can fix N by normalizing $h_0(x, t)$:

$$\int_{-\infty}^{+\infty} h_0^*(x, t) h_0(x, t) dx$$

$$= N^2 \int_{-\infty}^{+\infty} \frac{dx}{\varphi^{1/2}} e^{-\zeta^2} = N^2 \int_{-\infty}^{+\infty} d\zeta e^{-\zeta^2} = N^2 \pi^{1/2} = 1.$$

Thus $N = \pi^{-1/4}$ and the eigenfunction for $n = 0$ is

$$h_0(x, t) = (\pi\varphi)^{-1/4} \exp i\left\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \frac{1}{2}\Phi\right\} \times e^{-(x/\varphi^{1/2} - \mathscr{B})^2/2}. \tag{4.14}$$

We obtain $h_n(x, t)$ by repeated application of J_+ to $h_0(x, t)$. Since $h_n(x, t) = (n!)^{-1/2} J_+^n h_0(x, t)$,

$$h_n(x, t) = (n!)^{-1/2}\varphi^{-1/4} \times \exp i\left\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \left(n + \frac{1}{2}\right)\Phi\right\} \times (Z^+)^n a_0(\zeta)$$

$$= \varphi^{-1/4} \times \exp i\left\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \left(n + \frac{1}{2}\right)\Phi\right\} \times a_n(\zeta),$$

where we have again made use of (4.11). Thus we get the relationship

$$a_n(\zeta) = (n!)^{-1/2}(Z^+)^n a_0(\zeta)$$

$$= \pi^{-1/4}(n!)^{-1/2} 2^{-n/2}(-)^n \left(\frac{d}{d\zeta} - \zeta\right)^n e^{-\zeta^2/2} \tag{4.15}$$

$$= \pi^{-1/4}(n!)^{-1/2} 2^{-n/2} H_n(\zeta) e^{-\zeta^2/2}. \tag{4.16}$$

Equation (4.15) follows from the definition (4.8) of Z^+. Together Eqs. (4.15) and (4.16) imply that

$$H_n(\zeta) = (-)^n e^{\zeta^2/2}\left(\frac{d}{d\zeta} - \zeta\right)^n e^{-\zeta^2/2},$$

$$= (-)^n e^{\zeta^2} \frac{d^n}{d\zeta^n} e^{-\zeta^2}, \tag{4.17}$$

where we have used the operator identity

$$\left(\frac{d}{d\zeta} - \zeta\right)^n = e^{\zeta^2/2} \frac{d^n}{d\zeta^n} e^{\zeta^2/2}.$$

Equation (4.17) is the Rodrigues formula[13] for the Hermite functions, a class of orthogonal polynomials. Thus the wave function $h_n(x, t)$ has the form

$$h_n(x, t) = (\pi\varphi)^{-1/4}(n!)^{-1/2} 2^{-n/2} \times \exp i\left\{\frac{x^2\dot{\varphi}}{4\varphi} + \frac{x\mathscr{A}}{\varphi} - \Lambda - G_0 - \left(n + \frac{1}{2}\right)\Phi\right\} \times H_n\left(\frac{x}{\varphi^{1/2}} - \mathscr{B}\right) e^{-(x/\varphi^{1/2} - \mathscr{B})^2/2}. \tag{4.18}$$

These functions form a complete set of orthonormal solutions to the Schrödinger equation (1.5).

We remark here that if we evaluate the relation

$$J_+ J_- h_n = n h_n$$

we get the second-order ordinary differential equation for $a_n(\zeta)$

$$\frac{d^2 a_n}{d\zeta^2} + (2n + 1 - \zeta^2) a_n = 0,$$

which is the parabolic cylinder equation[9,13] which has normalized solutions (4.16) with $H_n(\zeta)$ defined by (4.17). That (4.18) is indeed a solution to the Schrödinger equation (1.5) may be confirmed by substitution.

We mention too that solving the first-order equation (4.1) is equivalent to finding a separable coordinate system[9] for the time-dependent Schrödinger equation (1.5). The new variables $(x/\varphi^{1/2} - \mathscr{B}, t)$ permit R separation[9] of (1.5) yielding the solutions (4.18).

5. DISCUSSION

In nonrelativistic quantum mechanics, time is a parameter and the equations of motion and their solutions have a parametric time dependence.[5] In particular, the Schrödinger equation

$$\{\partial_{xx} + 2i\partial_t - 2V(x, t)\}\Psi_\alpha(x, t) = 0 \tag{5.1}$$

gives the evolution of the wave function or state vector $\Psi_\alpha(x, t)$ which corresponds to the state α at any time. The set of solutions $\Psi_\alpha(x, t)$ of (5.1) form a Hilbert space; each state vector $\Psi_\alpha(x, t)$ corresponding to a state α has a definite direction in the Hilbert space of solutions at each point in time.[5] The relative orientations of the state vectors may change then during the evolution of the system.

For nonconservative systems, $V(x, t)$ is time dependent and the directions of the state vectors in Hilbert space shift with time. The possibility exists then for transitions between states. The probability that the system, in state α at time t_1, will be in state β at time t_2 is given by the square of the modulus of the probability amplitude, $|I_{\alpha\beta}(t_1, t_2)|^2$, where the probability amplitude is given by[5]

$$I_{\alpha\beta}(t_1, t_2) = \int_{-\infty}^{+\infty} \Psi_\alpha^*(x, t_1)\Psi_\beta(x, t_2)\, dx. \tag{5.2}$$

On the other hand, for conservative systems, the Hamiltonian is a constant of the motion and the system will be in a particular energy eigenstate, $\Psi_\alpha(x, t)$, where[5]

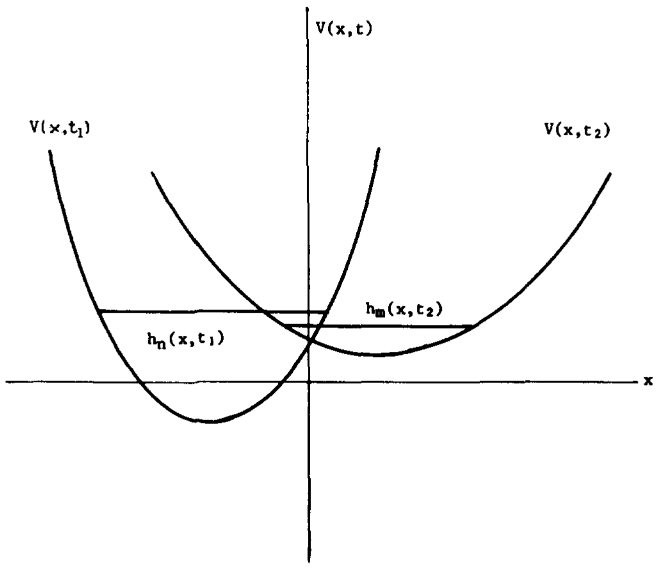


FIG. 1. Schematic diagram of the potential function $V(x,t) = g_2(t)x^2 + g_1(t)x + g_0(t)$ at two different times t_1 and t_2. The actual surface is three dimensional. We show only a two-dimensional projection.

$$\Psi_\alpha(x,t) = e^{i\mathscr{E}_\alpha t}\psi_\alpha(x).$$

If the Hamiltonian is Hermitian, the $\psi_\alpha(x)$ may always be chosen to be orthogonal[5] and

$$I_{\alpha\beta}(t_1, t_2) = 0, \quad \alpha \neq \beta.$$

So for conserved systems, the direction of the state vectors in Hilbert space is fixed and no transitions between energy eigenstates occurs.

In the specific case where the potential has the form (1.4), then the evolution of a state is governed by the coefficients $g_i(t)$ in the potential. In Fig. 1 we show schematically, the appearance of the potential at two different times t_1 and t_2. Both the position of the minimum and the curvature may vary, and the state vectors, the solutions of (1.5) will reflect this alteration in the potential. Now these solutions, $h_n(x,t)$ are given by (4.18) and are simultaneous eigenvectors of the complete set of commuting observables C, E, and M_3 which are symmetries of the Schrödinger equation (1.5). The state vectors $h_n(x,t)$ have, in general, a complicated time-dependence. They are generally not energy eigenstates since the Hamiltonian is not a constant of the motion. Since the probability amplitude

$$I_{nm}(t_1, t_2) = \int_{-\infty}^{+\infty} h_n^*(x, t_1) h_m(x, t_2) dx$$

does not vanish [see Appendix B(3)], transitions between states can occur. Why this is so, is suggested by the potential energy curves in Fig. 1.

For the time-independent harmonic oscillator [see Appendix A(1)], the symmetry M_3 is proportional to $-i\partial_t$ the energy operator. The Hamiltonian is a constant of the motion and the energy is conserved. Since M_3 is the diagonalized generator from the Lie symmetry algebra $\mathscr{G}$ and so the energy is quantized with

$$\mathscr{E}_n = (n + \tfrac{1}{2})\omega$$

and the system is in a definite energy eigenstate. Furthermore, the system remains in that state since the transition probability vanishes due to the orthogonality of the Hermite polynomials. We emphasize that it is the generator M_3 which is quantized in both the time-dependent and time-independent cases and only in the latter does M_3 correspond to the energy operator. It is in this sense that both cases, the time-dependent and the time-independent oscillators, are quantized.

Both sets of solutions, $\{h_n(x,t): n = 0, 1, ...\}$ for the time-dependent oscillator and $\{\Psi_n(x,t): n = 0, 1, ...\}$ for the harmonic oscillator (Appendix A1), form complete sets by hypothesis. We can expand the h_n in terms of the Ψ_n according to

$$h_n(x,t) = \sum_{m=0}^{\infty} c_{nm}(t)\Psi_m(x,t), \tag{5.3}$$

where the $\Psi_l(x,t)$ are given by (A9). The coefficients c_{nm} are time dependent,

$$c_{nm}(t) = \int_{-\infty}^{+\infty} \Psi_m^*(x,t) h_n(x,t) dx \tag{5.4}$$

and are evaluated in Appendix B(4), Eq. (B18). In form, they bear a close resemblance to the probability amplitudes $I_{nm}(t_1, t_2)$. In fact, $|c_{nm}(t)|^2$ is the probability that, at time t, the system under the influence of (1.4), will be found in the oscillator energy eigenstate Ψ_m with energy $\mathscr{E}_m = (m + \frac{1}{2})\omega$.

Now we shall turn our attention to the computation of other properties for systems with potentials (1.4). In particular those properties which can be expressed as polynomials of position and momentum operators are of interest.

We can express the momentum and position operators in terms of the raising and lowering operators J_+ and J_-,

$$x = \xi J_+ + \xi^* J_- + i(\xi\mathscr{C}^* - \xi^*\mathscr{C}), \tag{5.5a}$$

$$p_x = i\partial_x = \dot\xi J_+ + \dot\xi^* J_- + i(\dot\xi\mathscr{C}^* - \dot\xi^*\mathscr{C}), \tag{5.5b}$$

where we have used the definitions (2.25) and the Wronskian (2.21). With the help of (3.13) and the orthonormality of the eigenvectors $h_n(x,t)$ we have

$$\langle x \rangle = i(\xi\mathscr{C}^* - \xi^*\mathscr{C}), \tag{5.6a}$$

$$\langle p_x \rangle = i(\dot\xi\mathscr{C}^* - \dot\xi^*\mathscr{C}), \tag{5.6b}$$

their expectation values. Consistency with Ehrenfest's Theorem[5] can now be established. Taking the time derivative of $\langle x \rangle$ and using the definitions (2.24) for $\mathscr{C}$ and $\mathscr{C}^*$ we obtain

$$\frac{d\langle x \rangle}{dt} = i(\dot\xi\mathscr{C}^* - \dot\xi^*\mathscr{C}) = \langle p_x \rangle.$$

Secondly, the force is the derivative of the momentum

$$\frac{d\langle p_x \rangle}{dt} = 2ig_2(\xi^*\mathscr{C} - \xi\mathscr{C}^*) - g_1(t) = -\left\langle \frac{\partial V}{\partial x} \right\rangle.$$

This means that the expectation values $\langle x \rangle$ and $\langle p_x \rangle$ are good representations of the corresponding classical variables. Finally, because of the commutation relation $[J_-, J_+] = E$ and the Wronskian (2.21), we have the Heisenberg uncertainty relation

$$\Delta x \cdot \Delta p_x \geqslant \tfrac{1}{2}.$$

The Hamiltonian

$$\mathscr{H} = -\tfrac{1}{2}\partial_{xx} + g_2(t)x^2 + g_1(t)x + g_0(t)$$

corresponds to the total energy of the system[5] and can be expressed in terms of the ladder operators through (5.5). In this realization we get

$$\begin{aligned}\mathscr{H} = {} & \tfrac{1}{4}(\ddot{\varphi}_1 + 8g_2\varphi_1)J_+^2 + \tfrac{1}{4}(\ddot{\varphi}_2 + 8g_2\varphi_2)J_-^2 + \tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi)(J_+J_- + \tfrac{1}{2}) + i\{\tfrac{1}{2}(\ddot{\varphi}_1 + 8g_2\varphi_1)\mathscr{C}^* - \tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi)\mathscr{C} - ig_1\xi\}J_+ \\ & + i\{\tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi)\mathscr{C}^* - \tfrac{1}{2}(\ddot{\varphi}_2 + 8g_2\varphi_2)\mathscr{C} - ig_1\xi^*\}J_- + \tfrac{1}{2}(\ddot{\varphi}_1 + 8g_2\varphi_1)\mathscr{D}_1 + \tfrac{1}{2}(\ddot{\varphi}_2 + 8g_2\varphi_2)\mathscr{D}_2 - \tfrac{1}{8}(\ddot{\varphi} + 8g_2\varphi)\mathscr{D} \\ & + ig_1(\xi\mathscr{C}^* - \xi^*\mathscr{C}) + g_0. \end{aligned} \tag{5.7}$$

Clearly the Hamiltonian is not diagonal in the representation $\uparrow_{-\frac{1}{2},1}$ of $\mathscr{G}$, which is consistent with the fact that the Hamiltonian is not a constant of the motion. However we can calculate the average energy,

$$\begin{aligned}\mathscr{E}_n = \langle\mathscr{H}\rangle = \int_{-\infty}^{+\infty} h_n^*(x,t)\mathscr{H}h_n(x,t)dx = {} & \tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi)(n + \tfrac{1}{2}) + \tfrac{1}{2}(\ddot{\varphi}_1 + 8g_2\varphi_1)\mathscr{D}_1 \\ & + \tfrac{1}{2}(\ddot{\varphi}_2 + 8g_2\varphi_2)\mathscr{D}_2 - \tfrac{1}{8}(\ddot{\varphi} + 8g_2\varphi)\mathscr{D} + ig_1(\xi\mathscr{C}^* - \xi^*\mathscr{C}) + g_0, \end{aligned} \tag{5.8}$$

where we have used (5.7) and the properties (3.13) of the orthonormal eigenvectors of $\uparrow_{-\frac{1}{2},1}$. It is interesting that the average energy $\mathscr{E}$ depends upon both the quantum number n and the time t in such a way that the "level" separations $\Delta\mathscr{E} = \mathscr{E}_{n+1} - \mathscr{E}_n = \frac{1}{4}(\ddot{\varphi} + 8g_2\varphi)$ are equally spaced at each instant[14] but the spacings themselves vary with time. Also, we note that $\Delta\mathscr{E}$ is always a positive quantity since $\frac{1}{4}(\ddot{\varphi} + 8g_2\varphi) = \dot{\xi}\dot{\xi}^*$ which is positive definite.

Finally, the special cases with potentials (1.3) and (1.2) are worked out in Appendices A(2) and A(3), respectively. They have been included for completeness.

APPENDIX A

(1) Time-independent harmonic oscillator. In this case, the functions $g_2(t) = \omega^2/2$ and $g_1(t) = g_0(t) = 0$, where ω is a real positive constant. We choose the real solutions of (2.4) to be

$$\chi_1(t) = \frac{1}{\sqrt{\omega}}\cos\omega t, \quad \chi_2(t) = \frac{1}{\sqrt{\omega}}\sin\omega t,$$
$$W(\chi_1,\chi_2) = 1. \tag{A1}$$

According to the definitions (2.20)

$$\xi(t) = \frac{1}{(2\omega)^{1/2}}e^{i\omega t} \quad\text{and}\quad \xi^*(t) = \frac{1}{(2\omega)^{1/2}}e^{-i\omega t} \tag{A2}$$

with the Wronskian

$$W(\xi,\xi^*) = -i. \tag{A3}$$

Hence

$$\varphi = \varphi_3 = \frac{1}{\omega} \quad\text{and}\quad \Phi = \int^t \frac{1}{\varphi} = \omega t. \tag{A4}$$

Since $g_1 = 0$,

$$\mathscr{C} = \mathscr{A}_j = 0, \quad 1 \leqslant j \leqslant 3 \tag{A5}$$

and the Lie algebra $\mathscr{G}$ is realized by the differential operators

$$M_3 = \frac{i}{\omega}\partial_t, \; J_+ = \frac{1}{(2\omega)^{1/2}}e^{-i\omega t}(-\partial_x + \omega x),$$
$$J_- = \frac{1}{(2\omega)^{1/2}}e^{i\omega t}(\partial_x + \omega x), \qquad E = 1. \tag{A6}$$

Note that the diagonalized operator M_3 is a multiple of the energy operator $-i\partial_t$. Because of this, the Hamiltonian

$$\mathscr{H} = \tfrac{1}{2}\partial_{xx} + (\omega^2/2)x^2 \tag{A7}$$

is a constant of the motion and the energy is quantized and is a constant having the value [see Eq. (5.8)]

$$\mathscr{E}_n = \langle\mathscr{H}\rangle = (n + \tfrac{1}{2})\omega. \tag{A8}$$

The solutions for the harmonic oscillator Schrödinger equation can be obtained from (4.18)

$$\begin{aligned}\Psi_n(x,t) = {} & (n!)^{-1/2}\left(\frac{\pi}{\omega}\right)^{-1/4}2^{-n/2} \\ & \times\exp\{-i\mathscr{E}_n t\}H_n(\omega^{1/2}x)e^{-\omega x^2/2}, \\ & n = 0, 1, \cdots. \end{aligned} \tag{A9}$$

The functions $h_n(x,t)$ are eigenfunctions of both M_3 and the Hamiltonian (A7). As a consequence of (A7) and the form of (1.5), the Schrödinger equation is separable in the Cartesian coordinate system (x,t).

(2) The Harmonic oscillator subject to a time-dependent force.[2] Now $g_2(t) = \omega^2/2$ as in A(1) above and $g_1(t)$ is an arbitrary real function. We take $g_0(t) = 0$. Thus the potential has the form (1.3). We can use the values for the functions ξ and ξ^* we found in (A2) and so φ and Φ are given by (A4). From (2.24) we obtain

$$\mathscr{C} = \frac{1}{(2\omega)^{1/2}}\int^t g_1e^{i\omega s}, \quad \mathscr{C}^* = \frac{1}{(2\omega)^{1/2}}\int^t g_1e^{-i\omega s}. \tag{A10}$$

According to the Eqs. (2.23a) we have

$$\mathscr{A}_1 = \frac{1}{2\omega}e^{i\omega t}\int^t g_1e^{i\omega s}, \quad \mathscr{A}_2 = \frac{1}{2\omega}e^{-i\omega t}\int^t g_1e^{-i\omega s},$$
$$\mathscr{A} = \frac{-1}{2\omega}\left(e^{i\omega t}\int^t g_1e^{-i\omega s} + e^{-i\omega t}\int^t g_1e^{i\omega s}\right). \tag{A11}$$

The $\mathscr{D}$ functions are given by (2.23b).

The generators of the symmetry algebra, $\mathscr{G}$ have the form

$$\left.\begin{aligned} & M_3 = i\{(1/\omega)\partial_t + \mathscr{A}\partial_x - i\mathscr{A}x + i\mathscr{D}\}, \\ & J_+ = \frac{-1}{(2\omega)^{1/2}}\left\{e^{-i\omega t}(\partial_x - \omega x) + i\int^t g_1e^{-i\omega s}\right\}, \\ & J_- = \frac{1}{(2\omega)^{1/2}}\left\{e^{i\omega t}(\partial_x + \omega x) + i\int^t g_1e^{i\omega s}\right\}, \\ & E = 1. \end{aligned}\right\} \tag{A12}$$

The simultaneous eigenvectors of C, E, and M_3 are given by (4.18)

$$h_n(x,t) = (n!)^{-1/2}(\pi/\omega)^{-1/4}2^{-n/2} \times \exp i\{x\omega\mathscr{A} - \Lambda - (n+\tfrac{1}{2})\omega t\} \times H_n(\omega^{1/2}x - \mathscr{B})e^{-(\omega^{1/2}x-\mathscr{B})^2/2}, \tag{A13}$$

where

$$\left.\begin{aligned} \mathscr{B} &= \omega^{3/2}\int^t \mathscr{A}, \\ \Lambda &= \omega^2\int^t \mathscr{A}^2 + \omega\int^t \mathscr{D}. \end{aligned}\right\} \tag{A14}$$

Form (5.8), the average energy is given by

$$\mathscr{E}_n(t) = \omega(n+\tfrac{1}{2}) - \tfrac{1}{2}\omega\mathscr{D} + (g_1/\omega)\mathscr{B}, \tag{A15}$$

where $\mathscr{D} = \mathscr{D}_3$ is defined by (2.23b).

The probability amplitude can be computed by a procedure analogous to that in B(3).[15]

$$I_{nm}(1,2) = AMB \times \sum_{j=0}^{m}\sum_{r=0}^{(n,j)}\sum_{s=0}^{\left[\frac{n+j-2r}{2}\right]} \frac{(-)^r 2^{m-j+r}\varphi^{m-j}(n+j-2r)!i^{n+j}}{(m-j)!(n-r)!(n+j-2r-2s)!r!s!} \times H_{n+j-2r-2s}(i\beta/\sqrt{2}), \tag{A16}$$

where

$$\begin{aligned} A &= (n!m!)^{1/2}2^{-(n+m)/2}, \\ M &= \exp i\{\Lambda(1) - \Lambda(2) + (n+\tfrac{1}{2})\omega t_1 - (m+\tfrac{1}{2})\omega t_2 \\ &\quad + \tfrac{1}{2}(\mathscr{B}(2) - \mathscr{B}(1))(\mathscr{A}(1) - \mathscr{A}(2))\}, \\ B &= \exp\{-\beta\beta^*/4\}, \\ \beta &= (\mathscr{B}(1) - \mathscr{B}(2)) + i\omega^{1/2}(\mathscr{A}(1) - \mathscr{A}(2)), \\ \sigma &= \mathscr{B}(1) - \mathscr{B}(2). \end{aligned}$$

The functions Λ, $\mathscr{B}$, and $\mathscr{A} = \mathscr{A}_3$ are given by (A14) and (A11).

(3) The time-dependent harmonic oscillator.[1] For this case $g(t) = g_0(t) = 0$ and $g_2(t)$ is an arbitrary differentiable function and the potential is given by (1.2). Since g_2 has not been specified the functions ξ and ξ^* are defined in terms of the χ_1 and χ_2, real solutions to (2.20). However, some simplification results since

$$\mathscr{C} = \mathscr{C}^* = 0. \tag{A17}$$

Therefore, $\mathscr{D}_j = \mathscr{A}_j = 0$, $1 \leqslant j \leqslant 3$, and the generators of the Lie algebra $\mathscr{G}$ have the form

$$\begin{aligned} M_3 &= i\{\varphi\partial_t + \tfrac{1}{2}\dot{\varphi}x\partial_x - (i/4)\ddot{\varphi}x^2 + \tfrac{1}{4}\dot{\varphi}\}, \\ J_+ &= -\xi^*\partial_x + i\dot{\xi}^*x, \\ J_- &= \xi\partial_x - i\dot{\xi}^*x, \\ E &= 1. \end{aligned} \tag{A18}$$

The basis vectors for $\uparrow_{-\frac{1}{2},1}$, simultaneous eigenvectors of the observables C, E, and M are given by a modified form of (4.18),

$$h_n(x,t) = (n!)^{-1/2}(\pi\varphi)^{-1/4}2^{-n/2} \times \exp i\{x^2\dot{\varphi}/4\varphi - (n+\tfrac{1}{2})\Phi\} \times H_n(x/\varphi^{1/2})\exp(-x^2/\varphi), \tag{A19}$$

were Φ is defined by (4.5) with (4.11).

The average energy may be obtained form (5.8) and is given by

$$\mathscr{E}_n(t) = \tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi)(n+\tfrac{1}{2}). \tag{A20}$$

Since $\tfrac{1}{4}(\ddot{\varphi} + 8g_2\varphi) = \dot{\xi}\dot{\xi}^*$, the energy $\mathscr{E}_n(t)$ is always positive.

The probability amplitudes

$$I_{nm}(1,2) = \int_{-\infty}^{+\infty} h_n^*(x,t_1)h_m(x,t_2)dx$$

may be obtained as follows using the wave function (A19).

$$I_{nm}(1,2) = AMB\sum_{k=0}^{[m/2]}\frac{(-)^k m!}{(m-2k)!}\mathscr{P}_{mk}(\rho)\int_{\infty}^{\infty}dy \times \exp\{-2\gamma^2y^2\}H_n(y)H_{m-2k}(y), \tag{A21}$$

where

$$\begin{aligned} y &= x/\varphi(1)^{1/2}, \\ A &= \pi^{-1/2}(n!m!)^{-(n+m)/2}, \\ M &= \exp i\{(n+\tfrac{1}{2})\Phi(1) - (m+\tfrac{1}{2})\Phi(2)\}, \\ B &= \rho^{1/2} = [\varphi(1)/\varphi(2)]^{1/4}, \\ 2\gamma^2 &= \frac{1}{2}\left(1 + \frac{\varphi(1)}{\varphi(2)}\right) + \frac{i}{2}\left[\frac{\varphi(1)\dot{\varphi}(2)}{\varphi(2)} - \dot{\varphi}(1)\right]. \end{aligned}$$

Note that $\mathrm{Re}\gamma^2 = \tfrac{1}{4}(1 + \varphi(1)/\varphi(2)) > 0$ since φ is positive definite. The function $\mathscr{P}_{mk}(\rho)$, a polynomial in ρ is given by (B15) in Appendix B. The integral in (A21) is a standard integral[16] and so we obtain

$$I_{nm}(1,2) = A'MB\sum_{k=0}^{m/2}\frac{(-)^k m!}{(m-2k)!}\mathscr{P}_{mk}(\rho)2^{-(2k+1)/2}\gamma^{-m-n+2k-1}(1-2\gamma^2)^{\frac{m+n-2k}{2}}\Gamma\left(\frac{m+n-2k+1}{2}\right) \times {}_2F_1\left(-m,\ -n+2k;\ \frac{1-n-m+2k}{2};\ \frac{\gamma^2}{2\gamma^2-1}\right),$$

where $\mathrm{Re}\gamma^2 > 0$ and $m+n$ even. The function ${}_2F_1(a,b;c;z)$ is a hypergeometric function.[13,16] The coefficient A' is defined by $A' = (\pi n!m!)^{-1/2}$.

APPENDIX B

(1) We wish to integrate the subsidiary conditions (4.2) in order to obtain the wave function (4.3), solution to the partial differential equation (4.1). The conditions (4.2) are,

$$\frac{dt}{\varphi} = \frac{dx}{\tfrac{1}{2}\dot{\varphi}x + \mathscr{A}} = -\frac{dh_n}{(-(i/4)\ddot{\varphi}x^2 - i\mathscr{A}x + \tfrac{1}{4}\dot{\varphi} + ig_0\varphi + i\mathscr{D} + i(n+\tfrac{1}{2}))h_n}. \tag{B1}$$

By appropriately rearranging the first equality we obtain the inhomogeneous differential equation,

$$\frac{dx}{dt} - \frac{1}{2}\frac{\dot{\varphi}}{\varphi}x = \frac{\mathscr{A}}{\varphi},$$

which when integrated yields a solution

$$\frac{x}{\varphi^{1/2}} - \mathscr{B} = C_1, \tag{B2}$$

where C_1 is an arbitrary constant and

$$\mathscr{B}(t) = \int^t \frac{\mathscr{A}}{\varphi^{3/2}}. \tag{B3}$$

Taking the first and third terms we can recast them into the differential equation

$$\left\{\frac{-i}{4}\frac{\ddot{\varphi}}{\varphi}x^2 - i\frac{\dot{\mathscr{A}}}{\varphi}x + \frac{1}{4}\frac{\dot{\varphi}}{\varphi} + ig_0 + i\frac{\mathscr{D}}{\varphi} + i\left(n+\frac{1}{2}\right)\frac{1}{\varphi}\right\}dt = \frac{-dh_n}{h_n} \tag{B4}$$

for h_n. We can eliminate the x dependence in (B4) by substituting Eq. (B2). Integrating the resulting expression and substituting (B2) for the constant C_1 we obtain

$$h_n \exp\{-ix^2\dot{\varphi}/4\varphi + ix\Lambda_1 + i\Lambda + iG_0 + i(n+\tfrac{1}{2})\Phi\} = C_2 \tag{B5}$$

where C_2 is an arbitrary constant. G_0 and Φ are defined by (4.5) and (4.6), respectively, and

$$\Lambda_1 = \left(\dot{\varphi}\mathscr{B}/2 - \frac{1}{2}\int^t \ddot{\varphi}\mathscr{B} - \int^t (\dot{\mathscr{A}}/\varphi^{1/2})\right)/\varphi^{1/2},$$

$$\Lambda = \left(-\frac{1}{4}\dot{\varphi}\mathscr{B}^2 - \frac{1}{4}\int^t \ddot{\varphi}\mathscr{B}^2 + \mathscr{B}\int^t \dot{\mathscr{A}}/\varphi^{1/2} - \int^t \dot{\mathscr{A}}\mathscr{B}/\varphi^{1/2}\right) + (\mathscr{B}/2)\int^t \ddot{\varphi}\mathscr{B} + \int^t \mathscr{D}/\varphi\Big).$$

Both Λ_1 and Λ may be simplified. For Λ_1 we obtain after integrating by parts and some algebra

$$\Lambda_1 = (1/\varphi)\left[\left[\int^t \left(\frac{1}{2}\dot{\varphi}\mathscr{A} - \varphi\dot{\mathscr{A}}\right)/\varphi^{3/2}\right]\right]. \tag{B6}$$

By repeatedly integrating by parts we can reduce Λ to

$$\Lambda = -\int^t (\mathscr{A}/\varphi^{3/2})\int^s \left(\frac{1}{2}\dot{\varphi}\mathscr{A} - \varphi\dot{\mathscr{A}}\right)/\varphi^{3/2} + \int^t \mathscr{D}/\varphi. \tag{B7}$$

Again integrating by parts the common term in (B6) and (B7) we have

$$\int^t \left(\frac{1}{2}\dot{\varphi}\mathscr{A} - \varphi\dot{\mathscr{A}}\right)/\varphi^{3/2} = -\mathscr{A}/\varphi^{1/2}. \tag{B8}$$

Substituting (B8) into (B6) and (B7) yields

$$\Lambda_1 = -\mathscr{A}/\varphi, \quad \Lambda = \int^t \mathscr{A}^2/\varphi^2 + \int^t \mathscr{D}/\varphi. \tag{B9}$$

Hence (B5) becomes upon substitution for Λ_1

$$h_n \exp\{-ix^2\dot{\varphi}/4\varphi - ix\mathscr{A}/\varphi + i\Lambda + iG_0 + i(n+\tfrac{1}{2})\Phi\} = C_2. \tag{B10}$$

Now (B2) and (B10) are two functionally independent solutions of the system (B1) and so we have a general integral,[17]

$$h_n\varphi^{1/4}\exp\{-x^2\dot{\varphi}/4\varphi - ix\dot{\mathscr{A}}/\varphi + i\Lambda + iG_0 + i(n+\tfrac{1}{2})\Phi\} = a_n(x\varphi^{1/2} - \mathscr{B}),$$

or

$$h_n(x,t) = \varphi^{-1/4}\exp\{ix^2\dot{\varphi}/4\varphi + ix\dot{\mathscr{A}}/\varphi - i\Lambda - iG_0 - i(n+\tfrac{1}{2})\Phi\} \times a_n(x/\varphi^{1/2} - \mathscr{B}). \tag{B11}$$

(2) There are a number of useful relationships which simplify many calculations:

$$\frac{\xi^*\mathscr{C} - \xi\mathscr{C}^*}{\varphi^{1/2}} = i\mathscr{B}, \tag{B12}$$

$$\Phi = \int^t \frac{1}{\varphi} = \frac{i}{2}\ln\left(\frac{\xi^*}{\xi}\right). \tag{B13}$$

In (B12), ξ is given by (2.20) and $\mathscr{C}$ by (2.24). Equation (B12) is obtained by integrating by parts and using the Wronskian, $W(\xi,\xi^*) = -i$.

The second relation follows from the definition of φ and the Wronskian, $W(\xi,\xi^*) = -i$:

$$\int^t \frac{1}{\varphi} = \frac{1}{2}\int^t \frac{1}{\xi\xi^*} = \frac{i}{2}\int^t \frac{\xi\dot{\xi}^* - \dot{\xi}\xi^*}{\xi\xi^*} = \frac{i}{2}\ln\frac{\xi^*}{\xi}.$$

(3) Using the specific form of the wave functions (4.18) we can compute the probability amplitudes

$$\begin{aligned} I_{nm}(t_1,t_2) &= I_{nm}(1,2) = \int_{-\infty}^{+\infty} h_n^*(x,t_1)h_m(x,t_2)dx \\ &= AMB\int_{-\infty}^{+\infty} dx\exp\{-\alpha' x^2 - \beta' x\} \\ &\quad \times H_n(x/\varphi^{1/2}(1) - \mathscr{B}(1)) \\ &\quad \times H_m(x/\varphi^{1/2}(2) - \mathscr{B}(2)), \end{aligned}$$

where

$$A = \pi^{-1/2}(n!m!)^{-1/2}2^{-(n+m)/2},$$

$$M = \exp i\{\Lambda(1) - \Lambda(2) + G_0(1) - G_0(2) + (n+\tfrac{1}{2})\Phi(1) - (m+\tfrac{1}{2})\Phi(2)\},$$

$$B = (\varphi(1)\varphi(2))^{-1/4}\exp\{-(\mathscr{B}^2(1) + \mathscr{B}^2(2))/2\},$$

$$\alpha' = \frac{1}{2}\left[\frac{1}{\varphi(1)} + \frac{1}{\varphi(2)}\right] + \frac{i}{4}\left[\frac{\dot{\varphi}(2)}{\varphi(2)} - \frac{\dot{\varphi}(1)}{\varphi(1)}\right],$$

$$\beta' = i\left(\frac{\mathscr{A}(1)}{\varphi(1)} - \frac{\mathscr{A}(2)}{\varphi(2)}\right) - \left(\frac{\mathscr{B}(1)}{\varphi(1)^{1/2}} + \frac{\mathscr{B}(2)}{\varphi(2)^{1/2}}\right).$$

Note that β' is complex since all the quantities, φ, $\mathscr{A}$, $\mathscr{B}$ are real. Furthermore, $\mathrm{Re}\alpha' > 0$ since $\varphi = 2\xi\xi^*$ is real and positive. We now perform a variable transformation

$$y = x/\varphi^{1/2}(1) - \mathscr{B}(1),$$

which gives

$$I_{nm}(1,2) = AMB'\int_{-\infty}^{+\infty} dy\exp(-\alpha y^2 - \beta y) \times H_n(y)H_m(\rho y + \sigma),$$

where

$$\alpha = \alpha'\varphi(1),$$

$$\beta = \beta'\varphi^{1/2}(1) + 2\alpha\mathscr{B}(1),$$

$$\rho = [\varphi(1)/\varphi(2)]^{1/2},$$
$$\sigma = \mathscr{B}(1)\rho - \mathscr{B}(2),$$
$$B' = \rho^{1/2} \exp[-(\tfrac{1}{2} - \alpha)\mathscr{B}^2(1) - \mathscr{B}(1)\beta - \tfrac{1}{2}\mathscr{B}^2(2)].$$

Employing identities for Hermite polynomials[13] we obtain

$$I_{nm}(1,2) = AMB' \times \sum_{j=0}^{m} \sum_{k=0}^{[j/2]} \sum_{r=0}^{(n,j-2k)} \frac{m!n!2^{m-j+r}\mathscr{P}_{jk}(\rho)\sigma^{m-j}}{(m-j)!(n-r)!(j-2k-r)!r!} \times \int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 + y) H_{n+j-2k-2r}(y), \tag{B14}$$

where $(n, j-2k)$ indicates the upper limit to the sum is the smaller of the two numbers. The function $\mathscr{P}_{jk}(\rho)$ is a polynomial in the variable ρ:

$$\mathscr{P}_{jk}(\rho) = \sum_{l=0}^{k} \frac{\rho^{j-2l}}{l!(k-l)!}.$$

To solve the integral in (B14) we note that

$$\int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 - \beta y) H_l(y) = H_l\left(\frac{-\partial}{\partial\beta}\right) \int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 - \beta y).$$

The integral in the expression on the right[16] is

$$\int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 - \beta y) = \left(\frac{\pi}{\alpha}\right)^{1/2} \exp\left(\frac{\beta^2}{4\alpha}\right), \qquad \operatorname{Re}\alpha > 0$$

whence

$$\int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 - \beta y) H_l(y) = \left(\frac{\pi}{\alpha}\right)^{1/2} \sum_{k=0}^{[l/2]} \frac{(-)^{l+k} l! 2^{l-2k}}{k!(l-2k)!} \left(\frac{\partial}{\partial\beta}\right)^{l-2k} \exp\left(\frac{\beta^2}{4\alpha}\right).$$

If we use the Rodrigues formula for Hermite polynomials; then

$$\int_{-\infty}^{+\infty} dy \exp(-\alpha y^2 - \beta y) H_l(y) = \left(\frac{\pi}{\alpha}\right)^{1/2} \sum_{k=0}^{[l/2]} \frac{l! i^l}{k!(l-2k)!} \left(\frac{1}{\alpha}\right)^{(l-2k)/2} \times H_{l-2k}\left(\frac{i\beta}{2\sqrt{\alpha}}\right) e^{\beta^2/4\alpha}.$$

Therefore

$$I_{nm}(1,2) = A'MB'' \sum_{j=0} \sum_{k=0}^{[j/2]} \sum_{r=0}^{(n,j-2k)} \sum_{s=0}^{[(n+j-2k-2r)/2]} \frac{i^{n+j}(-)^r 2^{m-j+r}(n+j-2k-2r)!\sigma^{m-j}\mathscr{P}_{jk}(\rho)}{(m-j)!(n-r)!(n+j-2k-2r-2s)!(j-2k-r)!r!s!} \times \left(\frac{1}{\alpha}\right)^{(n+j-2k-2r-2s)/2} H_{n+j-2k-2r-2s}\left(\frac{i\beta}{2\sqrt{\alpha}}\right), \tag{B15}$$

where

$$A' = [n!m!2^{-(n+m)}]^{1/2}, \quad B'' = \frac{\rho^{1/2}}{\alpha} \exp[\beta^2/4\alpha - (\tfrac{1}{2} - \alpha)\mathscr{B}^2(1) - \mathscr{B}(1)\beta - \tfrac{1}{2}\mathscr{B}^2(2)]. \tag{B16}$$

(4) If we expand the functions (4.18) in terms of the time-dependent harmonic oscillator as in (5.3), then the coefficients

$$C_{nl}(t) = \int_{-\infty}^{+\infty} \Psi_l^*(x,t) h_n(x,t) dx, \tag{B17}$$

where the solutions to the time-independent oscillator are given by (A9) in Appendix A(1). The integral (B17) may be evaluated in the same fashion as the probability amplitudes computed in (B3) above. We quote only the result here.

$$C_{nm}(t) = AMB \sum_{j=0}^{m} \sum_{k=0}^{[j/2]} \sum_{r=0}^{(n,j-2k)} \sum_{s=0}^{[(n+j-2k-2r)/2]} \frac{(-)^{n+j+r} 2^{m-j+r} i^{n+j} (n+j-2k-2r)! \mathscr{B}^{m-j} \mathscr{P}_{jk}(\rho)}{(m-j)!(n-r)!(j-2k-r)!(n+j-2k-2r-2s)!r!s!} \times \left(\frac{1}{\alpha}\right)^{(n+j-2k-2r-s)/2} H_{n+j-2k-2r-2s}\left(\frac{i\beta}{2\sqrt{\alpha}}\right) e^{\beta^2/4\alpha}, \tag{B18}$$

where

$$A = (n!m!)^{1/2} 2^{-(n+m)/2},$$
$$\alpha = \frac{1}{2}\left(1 + \frac{1}{\omega\varphi}\right) - \frac{i\dot{\varphi}}{4\omega},$$
$$\beta = \frac{-\mathscr{B}}{(\omega\varphi)^{1/2}} - \frac{i\mathscr{A}}{\omega^{1/2}\varphi},$$
$$\rho = \left(\frac{1}{\omega\varphi}\right)^{1/2},$$
$$B = \left(\frac{1}{\alpha}\right)^{1/2} \left(\frac{\omega}{\varphi}\right)^{1/4} \exp(-\mathscr{B}^2/2).$$

ACKNOWLEDGMENTS

The author wishes to acknowledge Dr. R. Paul and Dr. H. Laue for several helpful discussions, and the Natural Sciences and Engineering Research Council for partial support of this work.

[1] H. R. Lewis, Jr., J. Math. Phys. **9**, 1976 (1968); H. R. Lewis, Jr. and W. B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969); B. Remand and E. S. Hernandez, Physica A **103**, 55 (1980); S. Solimeno, P. DiPorto, and B. Crosignani, J. Math. Phys. **10**, 1922 (1969).

[2] E. H. Kerner, Can. J. Phys. **36**, 371 (1958); I. L. Thomas, Chem. Phys. Lett. **70**, 413 (1980).

[3]W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York, 1964).
[4]We recover the usual time-dependent Schrödinger equation if we change variables: $x = (\sqrt{m/k})q$ and $t = \tau/\hbar$.
[5](a) P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. (Oxford U. P., London, 1958); (b) A. Messiah, *Quantum Mechanics* 3rd ed. (Wiley, New York, 1961), Vol. I and II.
[6]D. R. Truax, J. Math Phys. **22**, 1959 (1981).
[7]For time-dependent potentials in three dimensions, see C. P. Boyer, Helv. Phys. Acta **47**, 589 (1974).
[8]For examples, see W. Miller, Jr., *Symmetry Groups and their Applications* (Academic, New York, 1972) or R. Gilmore, *Lie Groups, Lie Algebras and Some of their Applications* (Wiley, New York, 1974),.
[9]W. Miller, Jr., *Symmetry and Separation of Variables* (Addison-Wesley, Reading, Mass., 1977); K. B. Wolf, The Heisenberg-Weyl Ring in Quantum Mechanics in *Group Theory and its Applications*, edited by E. Loebl (Academic, New York, 1975), Vol. 3.
[10]$\dot{f}(t) = df(t)/dt$ and $i^2 = -1$.
[11]T. F. Jordon, *Linear Operators for Quantum Mechanics* (Wiley, New York, 1969), Ref. 5.
[12]W. Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968).
[13]W. Magnus, F. Oberhattinger, R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*, 3rd ed. (Springer, New York, 1966).
[14]This agrees with Lewis' observation for the special case with potential (1.2). See Ref. 1.
[15]The symbol (n, m) means the smaller of the two numbers n and m; $[n/2]$ means the largest integer less than $n/2$.
[16]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965).
[17]E. C. Zachmanoglou and D. W. Thoe, *Introduction to Partial Differential Equations with Applications* (Williams and Wilkens, Baltimore, Md., 1976).

Statistics with number operator CC^*

Steven Robbins
Division of Mathematics, Computer Science, and Systems Design, The University of Texas at San Antonio, San Antonio, Texas 78285

(Received 25 March 1980; accepted for publication 28 August 1980)

The number operator a^*a, where a is an annihilation operator, plays a fundamental role in the statistics of bosons and fermions. However, it is possible for other statistics to have the same number operator. We have previously shown that for one degree of freedom there is a type of statistics having this number operator corresponding to each p which is a positive integer or ∞. Fermions are obtained when $p = 1$ and bosons are obtained when $p = \infty$. No state can have more than p particles. In order to treat many degrees of freedom it is necessary to first consider the case of one degree of freedom differently. We show here that for more degrees of freedom a similar situation occurs but for each case other than bosons and fermions there is a positive integer q, such that no state can have more than q particles, even when the number of degrees of freedom is infinite. Thus these statistics are probably not physically realizable except in an approximate way.

PACS numbers: 03.65.Fd, 02.30.Tb

1. ONE DEGREE OF FREEDOM

This work is an extension of work done for one degree of freedom (Ref. 1). In order to extend to many degrees of freedom, the earlier work must be looked at differently.

In keeping with the notation introduced in the next section, C will represent a creation operator and its adjoint C^* will be the corresponding annihilation operator. The statement that $n = CC^*$ is a number operator can be expressed by the commutation relation

$$[n,C] = C.$$

Since C and n may be unbounded operators this relation may only be satisfied on a certain dense domain. This domain should contain the vacuum vector v which should have zero particles and thus satisfy $C^*v = 0$. The vacuum should be essentially unique; that is, there should be only one zero-particle state. The domain should also include the result of creating and then annihilating any number of particles from the vacuum. These are the hypotheses of Theorem 1. We start with a definition.

Definition: Let $\mathscr{A}$ be a collection of operators on a Hilbert space K. Poly $(\mathscr{A})$ will be the set of all polynomials formed from operators in $\mathscr{A}$. $\mathscr{D}^{\infty}(\mathscr{A})$ will denote the set of all vectors in K which are in the domain of all elements of poly $(\mathscr{A})$. If $v \in \mathscr{D}^{\infty}(\mathscr{A})$ then $K_v(\mathscr{A}) = \{Bv: B \in \text{poly}(\mathscr{A})\}$. If $K_v(\mathscr{A})$ is a dense subset of K then v will be called a cyclic vector of $\mathscr{A}$.

In the following if p is a positive integer we set $N_p = \{0,1,2,3,...p\}$ and let N_∞ denote the nonnegative integers.

Theorem 1: Suppose C is a closed, densely-defined operator on a complex Hilbert space K, $\mathscr{A} = \text{poly}(\{C,C^*\})$, $n = CC^*$, there is a unique (up to scalar multiple) unit vector $v \in K$ such that $C^*v = 0$, this v is a cyclic vector for $\mathscr{A}$, and, for $w \in K_v(\mathscr{A})$,

$$[n,C]w = Cw. \tag{1.1}$$

Then for some p, the set

$$\{v_k = (k!)^{-1/2}C^k v : k \in N_p\}$$

is an orthonormal basis for K and

$$C^*v_k = k^{1/2}\, v_{k-1}, \; k \in N_p, \quad k \neq 0.$$

Remark: It is an easy computation to show that all nonnegative integral values of p are possible and give operators, C, which satisfy the hypotheses of the theorem. The case $p = 0$ corresponds to $C = 0$ identically and so to avoid trivialities in the following we assume that C is not identically zero. If $p = 1$, C is just the fermion creation operator (in this case K is two dimensional) and $p = \infty$ gives bosons.

Proof of Theorem 1: Let $K' = K_v(\mathscr{A})$. If $w,u \in K'$, Eq. (1.1) gives

$$\langle [n,C]u,w\rangle = \langle Cu,w\rangle,$$
$$\langle u, -[n,C^*]w\rangle = \langle u,C^*w\rangle,$$
$$0 = \langle u,[n,C^*]w + C^*w\rangle.$$

Since this holds for all $u \in K'$ and K' is dense, if $w \in K'$,

$$[n,C^*]w = -C^*w. \tag{1.2}$$

From Eq. (1.1) and (1.2) and $nv = 0$ it follows that if a monomial in C and C^* acts on v, the result is an eigenvector of n (if it is nonzero) with eigenvalue equal to the difference between the number of creators and annihilators. In particular if the number of annihilators exceeds the number of creators the result is zero since n is a nonnegative operator. Also, if the number of annihilators equals the number of creators the result is proportional to v since v is the essentially unique vector such that $nv = 0$.

Lemma 1: If $C^k v \neq 0$ then

$$C^*C^k v = kC^{k-1}v \tag{1.3}$$

and

$$\|C^k v\|^2 = k! \tag{1.4}$$

Proof of Lemma 1: $C^{*k}C^k v = \alpha v$, where $\alpha = \|C^k v\|^2$. Successively applying C to this gives

 0022-2488/82/010055-09$02.50

$$CC^{*k}C^kv = \alpha Cv,$$
$$nC^{*k-1}C^kv = \alpha Cv,$$
$$C^{*k-1}C^kv = \alpha Cv,$$
$$CC^{*k-1}C^kv = \alpha C^2v,$$
$$nC^{*k-2}C^kv = \alpha C^2v,$$
$$2C^{*k-2}C^kv = \alpha C^2v.$$

Continuing, we get

$$(k-1)!C^*C^kv = \alpha C^{k-1}v\ , \tag{1.5}$$

and

$$k!C^kv = \alpha C^kv. \tag{1.6}$$

Since $C^kv \neq 0$, Eq. (1.6) gives $\alpha = k!$ and Eq. (1.5) reduces to Eq. (1.3).

Equation (1.3) implies that all annihilation operators can be eliminated from any element of $\mathscr{A}$ when it acts on v and so the set

$$\{C^kv: k = 0,1,2,...\}$$

spans K. This is an orthogonal set since

$$nC^kv = kC^kv$$

and n is self-adjoint. This completes the proof of Theorem 1.

Remark: The condition that n is a number operator can be stated in a mathematically more rigorous way than Eq. (1.1) which must explicitly give a domain to make sense when C is unbounded. Equation (1.1) is formally equivalent to

$$e^{itn}Ce^{-itn} = e^{it}C$$

for real values of t. Since e^{itn} is unitary when n is self-adjoint this equation is meaningful even when C and n are unbounded. This enables us to state that n is a number operator when we don't have a particular domain in mind. This is the situation when the vacuum vector is not explicitly given. In fact, when the number of degrees of freedom is finite a vacuum vector must necessarily exist and thus does not have to appear in the hypotheses of the theorem. First we must make a definition which will allow us to replace the condition in Theorem 1 that $K_v(\mathscr{A})$ is dense in K by a condition that does not involve the vacuum. This irreducibility condition states that no nontrivial subspace of K is invariant under both C and C^* but is applicable even when C is unbounded so that the domain of C need not contain the entire subspace.

Definition: Let C be a closed densely-defined operator on a Hilbert space K. Let M be a closed subspace of K and let P be the projection onto M. We say that M (or P) reduces C if $CP \supset PC$. Such a subspace is called nontrivial if it is neither the zero subspace nor all of K.

If follows from the general theory that if M reduces C it also reduces C^*.

Theorem 2: Suppose C is a closed, densely-defined operator on a complex Hilbert space K such that no nontrivial subspace of K reduces C and the operator $n = CC^*$ satisfies

$$e^{itn}Ce^{-itn} = e^{it}C \tag{1.7}$$

for all real values of t. Then C satisfies the hypotheses of Theorem 1.

Proof of Theorem 2: Let $n = \int \lambda E(d\lambda)$ be the spectral resolution of n.

Lemma 2: If Δ is the closed interval $[0,\beta]$ and $E(\Delta)w = w$ then

$$E[0,\beta - 1])\,C^*w = C^*w; \tag{1.8}$$

$C^*w \in \mathrm{Dom}(n^k)$ for each k and

$$(n+1)^kC^*w = C^*n^kw. \tag{1.9}$$

In this lemma we interpret a closed interval $[0,\alpha]$ with $\alpha < 0$ to be empty.

Proof of Lemma 2: Since $E(\Delta)w = w$, $w \in \mathrm{Dom}(C^*)$. The adjoint of Eq. (1.7) gives

$$e^{itn}C^*e^{-itn} = e^{-it}C^*, \tag{1.10}$$

which implies that

$$\frac{e^{itn}-1}{it}\,C^*w = C^*\,\frac{e^{it(n-1)}-1}{it}\,w. \tag{1.11}$$

Now $(e^{it(n-1)} - 1)w \in E(\Delta)K$ and C^* is bounded by β on $E(\Delta)K$ so as $t \to 0$, the right side of Eq. (1.11) approaches $C^*(n-1)w$ and so $C^* \in \mathrm{Dom}(n)$ and

$$nC^*w = C^*(n-1)w$$

or

$$(n+1)C^*w = C^*nw.$$

This establishes (1.9) for $k = 1$. Arguing by induction, if $(n+1)^kC^*w = C^*n^kw$ then since $n^kw \in E(\Delta)K$, $(n+1)^kC^*w \in \mathrm{Dom}(n)$ and

$$(n+1)C^*n^kw = C^*n^{k+1}w,$$
$$(n+1)^{k+1}C^*w = C^*n^{k+1}w.$$

This establishes Eq. (1.9). Equation (1.8) will follow from the following general fact about self-adjoint operators.

Lemma 3: Let T be a self-adjoint operator with spectral resolution $T = \int \lambda E(d\lambda)$. Suppose $E([0,\infty))\,x = x$ and for each positive integer k, $x \in \mathrm{Dom}(T^k)$. Let $b = \lim_{k\to\infty} \|T^kx\|^{1/k}$ and $a = b - \lim_{k\to\infty} \|(b-T)^kx\|^{1/k}$. Then $[a,b]$ is the smallest closed interval such that $E([a,b])x = x$. The proof of Lemma 3 is straightforward and is similar to the proof that on a finite measure space the L_p norms of a function approach the L_∞ norm. Here we need only the calculation of b.

$$\begin{aligned}\|(n+1)^kC^*w\|^2 &= \|C^*n^kw\|^2\\ &= \langle CC^*n^kw,n^kw\rangle\\ &= \|n^{k+1/2}w\|^2.\end{aligned}$$

Thus

$$\|(n+1)^kC^*w\| = \|n^kn^{1/2}w\|.$$

Since $E(\Delta)n^{1/2}w = n^{1/2}w$, if

$$b = \lim_{k\to\infty} \|n^kn^{1/2}w\|^{1/k},$$

then $b \leqslant \beta$. Thus

$$\lim_{k\to\infty} \|(n+1)^kC^*w\|^{1/k} \leqslant \beta$$

so $E([0,\beta - 1])C^*w = C^*w$. This completes the proof of Lemma 2.

Combining Eq. (1.7) and (1.10) gives

$$e^{itn} C^* C e^{-itn} = C^* C$$

and so C^*C commutes with n. Let $\bar{n} = C^*C$ and let $\bar{n} = \int \lambda F(d\lambda)$ be the spectral resolution of $\bar{n}$.

Lemma 4: If $nw = \alpha w$ then $w \in \mathrm{Dom}(\bar{n})$ and $nCw = (\alpha + 1)Cw$. Also, there exist w_1 and w_2 such that

$$w = w_1 + w_2,$$

$$nw_1 = \alpha w_1, \; nw_2 = \alpha w_2,$$

$$\bar{n}w_1 = (\alpha + 1)w_1, \; \bar{n}w_2 = 0.$$

Proof of Lemma 4: Note that if $nu = \alpha u$ and $u \in \mathrm{Dom}(C)$, then

$$\frac{e^{itn} - 1}{it} Cu = C \frac{e^{it(n+1)} - 1}{it} u = \frac{e^{it(\alpha + 1)} - 1}{it} Cu.$$

As $t \to 0$ the right side approaches $(\alpha + 1)\, Cu$, so

$$nCu = (\alpha + 1)Cu.$$

Now suppose $b > 0$ and $F((b - \epsilon, b])w \neq 0$ for $0 < \epsilon < b$. Let ϵ be given, $0 < \epsilon < b$ and let u be a unit vector parallel to $F((b - \epsilon, b])w$. Since n and $\bar{n}$ commute, $nu = \alpha u$ and since $u \in \mathrm{Dom}(\bar{n})$, $u \in \mathrm{Dom}(C)$. Let $x = (b - \bar{n})u$. Then $\|x\| \leqslant \epsilon$ and

$$\|Cx\|^2 = \langle \bar{n}x, x\rangle \leqslant b\,\|x\|^2 \leqslant b\epsilon^2,$$

$$\|Cu\|^2 = \langle \bar{n}u, u\rangle \geqslant (b - \epsilon)\|u\|^2 = b - \epsilon,$$

$$nCu = C\bar{n}u = bCu - Cx,$$

$$(\alpha + 1)Cu = bCu - Cx,$$

$$Cx = (b - \alpha - 1)Cu,$$

$$b\epsilon^2 \geqslant (b - \alpha - 1)^2 (b - \epsilon),$$

which is a contradiction if ϵ is smaller than both $\frac{1}{2}|b - \alpha - 1|$ and $\frac{1}{2}b$. Thus $b = \alpha + 1$. This shows that $F(\Delta)w \neq 0$ only when Δ contains 0 or $\alpha + 1$. The lemma follows with $w_2 = F(\{0\})w$ and $w_1 = F(\{\alpha + 1\})w$.

Let Δ be any bounded interval such that $E(\Delta) \neq 0$. Then for some $w \neq 0$, $E(\Delta)w = w$. Applying Lemma 2 repeatedly gives that for some nonnegative integer k, $C^{*k}w \neq 0$ but $C^{*k+1}w = 0$. Let v be a unit vector parallel to $C^{*k}w$. Then $C^*v = 0$ and so $nv = 0$.

Lemma 5: If $C^*v = 0$ then for each k, $0 \leqslant k < \infty$, $v \in \mathrm{Dom}(C^k)$ and either

$$\bar{n}C^k v = 0\,,$$

or

$$\bar{n}C^k v = (k + 1)C^k v.$$

Proof of Lemma 5: Let $P = F(\{0\})$. The proof is by induction.

$k = 0$: By Lemma 4, since $nv = 0$, $v \in \mathrm{Dom}(\bar{n})$ and $v = v_1 + v_2$, where $v_1 = Pv$, $Cv_1 = 0$ and $\bar{n}v_2 = v_2$. But $v_1 = 0$ since otherwise $nv_1 = 0$ (since n commutes with $\bar{n}$) and $\bar{n}v_1 = 0$ so the one-dimensional space spanned by v_1 would reduce C; and therefore be all of K making C indentically zero. Thus $v = v_2$ and $\bar{n}v = v$.

Next assume that the lemma is true for 0, 1, 2, 3, ..., $k - 1$. If for some $j < k$, $\bar{n}C^j v = 0$, then $C^{j+1}v = 0$ since

$$\|C^{j+1}v\|^2 = \langle \bar{n}C^j v, C^j v\rangle.$$

Thus we may assume that

$$\bar{n}C^j v = (j + 1)C^j v, \quad \text{for } 1 \leqslant j < k.$$

Since $nC^k v = kC^k v$, $C^k v \in \mathrm{Dom}(\bar{n})$ and $w_k = PC^k v$ satisfies $Cw_k = 0$, $nw_k = kw_k$. If $w_k = 0$ then $\bar{n}C^k v = (k + 1)C^k v$ and we are done. Assume $w_k \neq 0$. By Lemma 2, $w_k \in \mathrm{Dom}(C^{*j})$ for all j. Let $w_{k-j} = C^{*j}w_k$. Then $nw_{k-j} = (k - j)w_{k-j}$ for $j \leqslant k$,

and $w_{k-j} \neq 0$ when $j \leqslant k$ since

$$\begin{aligned}\|w_{k-j}\|^2 &= \langle C^{*j}w_k, C^{*j}w_k\rangle \\ &= \langle nC^{*j-1}w_k, C^{*j-1}w_k\rangle \\ &= (k - j + 1)\|C^{*j-1}w_k\|^2 \\ &= (k - j + 1)\|w_{k-j+1}\|^2.\end{aligned}$$

Thus we have

$$Cw_j = CC^*w_{j+1} = (j + 1)w_{j+1}, \quad 0 \leqslant j < k,$$

$$Cw_k = 0,$$

$$C^*w_0 = 0,$$

$$C^*w_j = w_{j-1}, \quad 1 \leqslant j \leqslant k.$$

These equations show that the space spanned by $\{w_0, w_1, w_2, \ldots, w_k\}$ reduces C. Since $w_k \neq 0$ this must be all of K. Since $n(1 - P)C^k v = k(1 - P)C^k v$, $(1 - P)C^k v$ is orthogonal to each of $w_0, w_1, \ldots, \omega_{k-1}$ and since it is also orthogonal to $C^k v$, it is zero. Thus $PC^k v = C^k v$ and $\bar{n}C^k v = (k + 1)C^k v$.

If p is the smallest integer such that $C^{p+1}v = 0$, where we set $p = \infty$ if $C^k v$ is never 0, then the span of the set

$$\{C^k v : k \in N_p\}$$

reduces C and thus is all of K. Thus v is a cyclic vector for $\mathscr{A} = \mathrm{poly}(\{C, C^*\})$. All vectors w with $nw = 0$ are parallel to v since $nw = 0$ implies that w is orthogonal to $C^k v$ for $k \geqslant 1$. Lastly, Lemma 4 implies that if $w \in K_v(\mathscr{A})$ then

$$[n, C]w = Cw.$$

This completes the proof of Theorem 2.

2. MANY DEGREES OF FREEDOM

In order to set a mathematical framework, suppose we have d annihilation operators $a_1, a_2, \ldots, a_d$ and corresponding creation operators $a_1^*, a_2^*, \ldots, a_d^*$. Let H be a complex d-dimensional Hilbert space with orthonormal basis $\{e_1, e_2, \ldots, e_j\}$. If $z \in H$ and $z = \Sigma \alpha_i e_i$ we define $C(z) = \Sigma \alpha_i a_i^*$. Then $C(z)$ is a creation operator and formally

$$C(z + y) = C(z) + C(y),$$

$$C(\alpha z) = \alpha C(z),$$

for $z, y \in H$ and α complex. When there is an essentially unique vacuum vector, H is isomorphic to the 1-particle vector states and is referred to as the single particle Hilbert space. The single particle space need not be finite-dimensional and in the physically important case H is a separable infinite-dimensional Hilbert space.

The key structure is that of a Fock space in which there is a vacuum vector from which all other states can be built. This is given in the following definition.

Definition: An irreducible clothed quantum structure over H is a collection $\{H, C, K, v\}$ where H and K are complex Hilbert spaces, C is a map from H into the set of closed densely defined operators on K and v is a unit vector in K

such that if

$$\mathscr{A} = \{C(z), C^*(z): z \in H\},$$

then $v \in D_\infty(\mathscr{A})$, v is cyclic for $\mathscr{A}$, v is the only vector in K up to scalar multiples such that $C^*(z)v = 0$ for all $z \in H$ and for $w \in K_v(\mathscr{A})$, $z, y \in H$ and complex α we have

$$C(z + y)w = C(z)w + C(y)w,$$

$$C(\alpha z)w = \alpha C(z)w.$$

K' will be used to denote $K_v(\mathscr{A})$.

The equation $C^*(z)v = 0$ indicates that particles cannot be annihilated from v, that is, v has no particles. Such a vector is called a vacuum. The condition that an operator $n(z)$ is a number operator for (the state) z is expressed formally by the commutation relation

$$[n(z), C(y)] = \langle y, z\rangle C(z).$$

This states that $n(z)$ commutes with $C(y)$ when y is orthogonal to z and it reduces to the commutation relation (1.1) when y and z are equal unit vectors. We will require that this formal relation is satisfied on K'.

Our main result is that if $C(z)C^*(z)$ is a number operator for z then the cases other than bosons and fermions give statistics with essentially a finite number of particles. For a finite number of degrees of freedom the fermion Fock space also has a finite number of particles but when the dimension of the single particle space is increased the number of possible particles also increases. This is not true in the other cases. To make this precise requires a few more definitions.

If $\{H, C, K, v\}$ and $\{\hat{H}, \hat{C}, \hat{K}, \hat{v}\}$ are two irreducible clothed quantum structures then $\{\hat{H}, \hat{C}, \hat{K}, \hat{v}\}$ is said to be an extension of $\{H, C, K, v\}$ if $H \subset \hat{H}$, $K \subset \hat{K}$, $v = \hat{v}$ and for all $z \in H$, $C(z) \subset \hat{C}(z)$. To express the notion of the number of particles we will use the total number operator. A self-adjoint operator N on K will be called a total number operator for $\{H, C, K, v\}$ if $Nv = 0$ and for $w \in K'$ and $z \in H$,

$$[N, C(z)]w = C(z)w.$$

We denote by K_m the closed subspace of K generated by those elements of K' which have the form $C(z_1)C(z_2)\cdots C(z_m)v$. Such an element will be called elementary. K_0 is just the one-dimensional space spanned by v. From the definition of N, each nonzero element of K_m is an eigenvector of N with eigenvalue m so that K_m is a space of m-particle states. If $\mathscr{B}$ is an orthonormal basis for H, $K_m(\mathscr{B})$ will denote the (not necessarily closed) linear space generated by the elementary elements of K_m which involve only creation operators $C(z)$ with $z \in \mathscr{B}$.

Theorem 3: Suppose $\{H, C, K, v\}$ is an irreducible clothed quantum structure over H such that if $z, y \in H$, $n(z) = C(z)C^*(z)$ and $w \in K'$, then

$$[n(z), C(y)]w = \langle y, z\rangle C(z)w. \tag{2.1}$$

Then $\{H, C, K, v\}$ has a total number operator N and if $\{e_\alpha : \alpha \in I\}$ is an orthonormal basis for H then $\Sigma_\alpha n(e_\alpha w$ converges to Nw for all $w \in K'$. Also, one of the following holds:

i) $\{H, C, K, v\}$ is equivalent to the free boson field,

ii) $\{H, C, K, v\}$ is equivalent to the free fermion field,

iii) $\|N\| < \infty$ and if $\{\hat{H}, \hat{C}, \hat{K}, \hat{v}\}$ is an extension of $\{H, C, K, v\}$ satisfying Eq. (2.1) for all $z, y \in \hat{H}$ and $w \in \hat{K}'$ and $\hat{N}$ is the total number operator for $\{\hat{H}, \hat{C}, \hat{K}, \hat{v}\}$, then $\|\hat{N}\| = \|N\|$.

Proof of Theorem 3: Let $\mathscr{B} = \{e_\alpha : \alpha \in I\}$ be an orthonormal basis for H and define $\mathscr{C}(\mathscr{B}) = \{C(e_\alpha): \alpha \in I\}$ and $\mathscr{C}^*(\mathscr{B}) = \{C^*(e_\alpha): \alpha \in I\}$. As in the proof of Theorem 1, the adjoint of Eq. (2.1),

$$[n(z), C^*(y)w] = -\langle z, y\rangle C^*(z)w, \quad \text{for } w \in K', \ y \in H \tag{2.2}$$

also is satisfied and together with the hypothesis that $C^*(z)v = 0$ for all $z \in H$ it follows that if $E_i \in \mathscr{C}(\mathscr{B}) \cup \mathscr{C}^*(\mathscr{B})$, $i = 1, 2, \ldots, k$ and $D \in \mathscr{C}(\mathscr{B})$ then $E_1E_2 \ldots E_k v$ (if it is not zero) is an eigenvector of DD^* with eigenvalue equal to the number of the E_i's which are equal to D, minus the number which are equal to D^*. Thus the eigenvalue is an integer.

Lemma 6: Suppose $D_i \in \mathscr{C}(\mathscr{B})$, $i = 1, 2, \ldots, k$. If $1 \leqslant j \leqslant k$, $D_jD_{j+1} \ldots D_k v \neq 0$ and α_j is the cardinality of the set

$$\{D_i: \ j \leqslant i \leqslant k \ \text{ and } \ D_i = D_j\},$$

then

$$D_j^*D_jD_{j+1}\cdots D_k v = \alpha_j D_{j+1}\cdots D_k v, \tag{2.3}$$

and if $D_1D_2\cdots D_k v \neq 0$,

$$\|D_1D_2\cdots D_k v\|^2 = \prod_{i=1}^{k} \alpha_i.$$

Furthermore, if σ is a permutation on k elements, there is a scalar, β, with $|\beta| = 1$ such that

$$D_1D_2\cdots D_k v = \beta D_{\sigma 1} D_{\sigma 2}\cdots D_{\sigma k} v. \tag{2.4}$$

Proof of Lemma 6: Since for any $z \in H$,

$$C^*(z)D_k^*D_{k-1}^*\cdots D_1^*D_1D_2\cdots D_k v = 0,$$

there is a scalar α such that

$$D_k^*D_{k-1}^*\cdots D_1^*D_1D_2\cdots D_k v = \alpha v,$$

$$D_kD_k^*D_{k-1}^*\cdots D_1^*D_1D_2\cdots D_k v = \alpha D_k v,$$

$$D_{k-1}^*\cdots D_1^*D_1D_2\cdots D_k v = \alpha D_k v.$$

Applying $D_{k-1}, D_{k-2}, \cdots D_2$ similarly, we get positive integers $\beta_{k-1}, \beta_{k-2}, \cdots \beta_2$ such that

$$\beta_{k-1}\beta_{k-2}\cdots\beta_2 D_1^*D_1D_2, \cdots, D_k v = \alpha D_2D_3, \cdots, D_k v.$$

Thus $\alpha_1 = \alpha/\beta_{k-1}\beta_{k-2}\cdots\beta_2$. Applying D_1 to

$$D_1^*D_1D_2\cdots D_k v = \alpha_1 D_2D_3\cdots D_k v,$$

we get

$$D_1D_1^*D_1D_2\cdots D_k v = \alpha_1 D_1D_2\cdots D_k v.$$

Since $D_1D_2\cdots D_k v$, if it is not zero, is an eigenvector of $D_1D_1^*$ with eigenvalue equal to the number of the D_i's which equal D_1, α_1 has the value stated in the lemma if $D_1D_2\cdots D_k v \neq 0$. A similar argument applies to α_j. The statement about the norm of $D_1D_2\cdots D_k v$ follows from

$$\begin{aligned}\|D_1D_2\cdots D_k v\|^2 &= \langle D_1^*D_1D_2\cdots D_k v, D_2\cdots D_k v\rangle \\ &= \alpha_1\|D_2D_3\cdots D_k v\|^2.\end{aligned}$$

Now let $w = D_1D_2\cdots D_k v$ and $u = D_{\sigma 1}D_{\sigma 2}\cdots D_{\sigma k} v$. There is a scalar γ, such that

$$D_{\sigma k}^*D_{\sigma(k-1)}^*\cdots D_{\sigma 2}^*D_{\sigma 1}^* w = \gamma v.$$

As above, applying $D_{\sigma k}, D_{\sigma(k-1)}, \cdots, D_{\sigma 2}D_{\sigma 1}$ we obtain positive integers $\gamma_k, \gamma_{k-1}, \cdots \gamma_1$ such that

$$\gamma_k \gamma_{k-1} \cdots \gamma_1 w = \gamma D_{\sigma 1} D_{\sigma 2} \cdots D_{\sigma k} v = \gamma u.$$

Thus, $w = \beta u$ where $\beta = \gamma / \gamma_k \cdots \gamma_1$. This shows that if $u = 0$, then $w = 0$ and the converse follows from a similar argument. If both w and u are zero then $w = \beta u$ with $\beta = 1$ and the proof is complete. If $w \neq 0$ then the definition of the α_i's above shows that

$$\|D_1 D_2 \cdots D_k v\|^2 = \prod_{\alpha \in I} (\lambda_\alpha !),$$

where λ_α is the number of the D_i's which are equal to $C(e_\alpha)$. Note that all but a finite number of terms in the product are 1 so it is well-defined. This formula shows that $\|D_1 D_2 \cdots D_k v\|$ is independent of the order of the terms so $\|w\| = \|u\|$ and thus $|\beta| = 1$. This completes the proof of Lemma 6.

The subspaces $K_m, m = 0,1,2,3,\ldots$ are mutually orthogonal. If $w \in K'$, there is an orthonormal basis, $\mathscr{B}$, of H such that w can be expressed as a linear combination of a finite number of terms in the form $E_1 E_2 \ldots E_k v$, where $E_i \in \mathscr{C}(\mathscr{B}) \cup \mathscr{C}^*(\mathscr{B})$, but Eqs. (2.3) and (2.4) imply that those $E_i \in \mathscr{C}^*(\mathscr{B})$ can be eliminated. Thus $K = \oplus_{m=0}^{\infty} K_m$.
$K = \oplus_{m=0}^{\infty} K_m$.

Lemma 7: If z is a unit vector in H then on K_m, $C(z)C^*(z)$ is bounded by m and $C^*(z)C(z)$ is bounded by $m+1$.

Proof of Lemma 7: Any elementary vector in K_m can be written as a finite linear combination of vectors in the form $w = D_1 D_2 \ldots D_m v$ with each $D_i \in \mathscr{C}(\mathscr{B})$, where $\mathscr{B}$ is an orthonormal basis containing z. Each of these is an eigenvector of $C(z)C^*(z)$ with eigenvalue less than or equal to m. Thus $C(z)C^*(z)$ is bounded by m on K_m, and K_m is spanned by eigenvectors of $C(z)C^*(z)$. Since by Lemma 6, each such eigenvector can be written as a sum of eigenvectors of both $C(z)C^*(z)$ and $C^*(z)C(z)$, these also span K_m. Suppose

$$C(z)C^*(z)w = \gamma w$$

and

$$C^*(z)C(z)w = \beta w,$$

thus

$$C(z)C^*(z)C(z)w = \beta C(z)w,$$

$$(\gamma + 1)C(z)w = \beta C(z)w.$$

So either $C(z)w = 0$ in which case $\beta = 0$ or $C(z)w \neq 0$ in which case $\beta = \gamma + 1 \leqslant m + 1$. Thus $C^*(z)C(z)$ is bounded by $m+1$ on K_m.

Let $\mathscr{B}$ be an orthonormal basis for H and define

$$r = \min\{j : D_1 D_2 \cdots D_j v = 0 \text{ for some } D_i \in \mathscr{C}(\mathscr{B}), 1 \leqslant i \leqslant j\},$$

$$s = \min\{j : D_1 D_2 \cdots D_j v = 0 \text{ for some distinct } D_i \in \mathscr{C}(\mathscr{B}), 1 \leqslant i \leqslant j\},$$

where we take $r = \infty$ or $s = \infty$ if the set is empty. Clearly $s \geqslant r$. We will show that if $r = s = \infty$ we get bosons and if $r = 2$, $s = \infty$ we get fermions. If $2 < r < \infty$, then $C(z_1)C(z_2)\cdots C(z_r)v = 0$ for all $z_i \in H$ whereas if $r = 2$ and $s < \infty$ then $C(z_1)C(z_2)\cdots C(z_s)v = 0$ for all $z_i \in H$.

Suppose $r = s = \infty$, that is, suppose $D_1 D_2 \cdots D_j v \neq 0$ and every choice of $D_i \in \mathscr{C}(\mathscr{B}), 1 \leqslant i \leqslant j$. Then this is true for every orthonormal basis as can be seen from the following argument. Suppose $\mathscr{B}' = \{e'_\alpha : \alpha \in I\}$ is another orthonormal basis for H. Let $D'_i = C(e'_{\alpha_i})$ and $D_i = C(e_{\alpha_i}), 1 \leqslant i \leqslant j$. Let

$$H' = \operatorname{span}\{e_{\alpha_1}, e_{\alpha_2}, \ldots, e_{\alpha_j}, e'_{\alpha_1}, e'_{\alpha_2}, \ldots, e'_{\alpha_j}\}.$$

If U is a unitary operator on H' then by Lemma 6,

$$\|C(Ue_{\alpha_1})C(Ue_{\alpha_2})\cdots C(Ue_{\alpha_j})v\|^2$$

is an integer and since the unitary group of H' is connected, it is independent of U. First taking U to be the identity and then a unitary operator such that $Ue_{\alpha_i} = e'_{\alpha_i}$, we see that

$$\|D_1 D_2 \cdots D_j v\|^2 = \|D'_1 D'_2 \cdots D'_j v\|^2.$$

Now suppose that z is any unit vector in H and let $\mathscr{B}$ be an orthonormal basis containing z. If w is an elementary vector of $K_m(\mathscr{B})$ then for some scalar β,

$$C(z)C^*(z)w = \beta w.$$

From Lemma 6 since $C(z)w \neq 0$ by assumption,

$$C^*(z)C(z)w = (\beta + 1)w.$$

Thus, for $w \in K_m(\mathscr{B})$,

$$[C^*(z), C(z)]w = w. \tag{2.5}$$

Since by Lemma 7, $[C^*(z), C(z)]$ is bounded on K_m and since $K_m(B)$ is a dense subset of K_m, Eq. (2.5) holds for all $w \in K_m$ and thus all $w \in K'$. Polarization of Eq. (2.5) shows that for $w \in K'$,

$$[C^*(z), C(y)]w = \langle y, z \rangle w.$$

This gives the boson field.

Next suppose $r = 2$, $s = \infty$ and that $D_1^2 v = 0$. Then for all $z \in H$, $C(z)^2 v = 0$ since $\|C(z)v\|^2$ is an integer when z is a unit vector and as above the connectedness of the unitary group implies that it is independent of z. Similarly, since $s = \infty$, $D_1 D_2 \cdots D_j v \neq 0$ for distinct $D_i \in \mathscr{C}(\mathscr{B})$ and this holds for any orthonormal basis $\mathscr{B}$. Also, since $\|D_1 D_2 \cdots D_j v\|$ is independent of the order of the terms, this is zero unless the D_i's are distinct.

Suppose z is a unit vector in H and $\mathscr{B}$ is an orthonormal basis containing z. Let w be an elementary vector in $K_m(\mathscr{B}), w = D_1 D_2 \cdots D_m v$, with $w \neq 0$ so that the D_i's are distinct. If $C(z)$ is distinct from all of the D_i's then $C^*(z)w = 0$ and $C^*(z)C(z)w = w$. If $C(z)$ is equal to one of the D_i's, then $C(z)C^*(z)w = w$ and $C(z)w = 0$. In either case

$$[C^*(z), C(z)]_+ w = w.$$

This holds for all $w \in K_m(\mathscr{B})$ and thus for all $w \in K'$ and polarization gives that for $w \in K'$, $z, y \in H$,

$$[C^*(z), C(y)]_+ w = \langle y, z \rangle w,$$

and this gives fermion field.

We will next show that if $r > 2, E_i \in \mathscr{C}(\mathscr{B})$ and $E_1 E_2 \cdots E_r v = 0$, then if $D \in \mathscr{C}(\mathscr{B})$, we have $E_1 D E_3 \cdots E_r v = 0$. Since the order of the creation operators does not affect the norm this implies that each of the E_i's can be replaced by arbitrary elements $D_i \in \mathscr{C}(\mathscr{B})$ and that $D_1 D_2 \cdots D_r v = 0$.

Polarization of Eq. (2.1) gives that for $x, y, z \in H$ and $w \in K'$,

$$[C(x)C^*(z), C(y)]w = \langle y, z \rangle C(x)w. \tag{2.6}$$

If $E_1 \neq E_2$ then this gives

$$[DE_2^* E_1]w = 0.$$

Thus,

$$DE_2^* E_1(E_2E_3 \cdots E_r v) = E_1 DE_2^*(E_2E_3 \cdots E_r v)$$

and if $E_1E_2 \cdots E_r v = 0$, then

$$E_1 DE_2^* E_2 E_3 \cdots E_r v = 0.$$

By Lemma 6,

$$E_2^* E_2 E_3 \cdots E_r v = \alpha E_3 \cdots E_r v,$$

where $\alpha \neq 0$ since $E_2E_3 \cdots E_r v \neq 0$ by the definition of r. Thus $E_1 DE_3 \cdots E_r v = 0$.

If $E_1 = E_2$, then rearrange the order of the E_i's, if possible, so that E_2 is unmoved but the first E_i is not equal to E_2. The above argument can then be used. If no such rearrangement is possible, then all of the E_i's are equal and it is necessary to show that if $E^r v = 0$ then $DE^{r-1}v = 0$. We will assume that $D \neq E$, for otherwise there is nothing to prove. From Eq. (2.6), for $w \in K'$,

$$[DE^*, E]w = Dw,$$

so

$$DE^*E(E^{r-1}v) - EDE^*(E^{r-1}v) = DE^{r-1}v,$$

$$-EDE^*E^{r-1}v = DE^{r-1}v,$$

$$-(r-1)EDE^{r-2}v = DE^{r-1}v.$$

Assume for the sake of contradiction that $DE^{r-1}v \neq 0$ so that $EDE^{r-2}v \neq 0$. From Lemma 6 there exists a scalar α such that

$$E^*DE^{r-1}v = \alpha DE^{r-2}v,$$

$$EE^*DE^{r-1}v = \alpha EDE^{r-2}v,$$

$$(r-1)DE^{r-1}v = \alpha EDE^{r-2}v$$

$$-(r-1)^2 EDE^{r-2}v = \alpha EDE^{r-2}v,$$

so $\alpha = -(r-1)^2$ and

$$E^*DE^{r-1}v = -(r-1)^2 DE^{r-2}v.$$

Thus

$$\langle E^*DE^{r-1}v, DE^{r-2}v\rangle = \langle DE^{r-1}v, EDE^{r-2}v\rangle,$$

$$-(r-1)^2\|DE^{r-2}v\|^2 = -\|DE^{r-1}v\|^2/(r-1),$$

$$(r-1)^3\|DE^{r-2}v\|^2 = \|DE^{r-1}v\|^2.$$

On the other hand,

$$\|DE^{r-1}v\|^2 = \langle D^*DE^{r-1}v, E^{r-1}v\rangle,$$

$$\|DE^{r-1}v\|^2 = \|E^{r-1}v\|^2,$$

since $D^*DE^{r-1}v = E^{r-1}v$ because $DE^{r-1}v \neq 0$. Similarly,

$$\|DE^{r-2}v\|^2 = \|E^{r-2}v\|^2.$$

Lastly,

$$\|E^{r-1}v\|^2 = \langle E^*E^{r-1}v, E^{r-2}v\rangle$$
$$= (r-1)\|E^{r-2}v\|^2.$$

Thus,

$$\|E^{r-1}v\|^2 = \|DE^{r-1}v\|^2 = (r-1)^3\|DE^{r-2}v\|^2$$
$$= (r-1)^3\|E^{r-2}v\|^2$$

and so $(r-1)^3 = (r-1)$ which contradicts $r > 2$. Thus we have $DE^{r-1}v = 0$.

We have shown that for $r > 2$, $D_1D_2 \cdots D_r v = 0$ for all $D_i \in \mathscr{C}(\mathscr{B})$. The same argument applies if $\{\hat{H}, \hat{C}, \hat{K}, \hat{v}\}$ is an extension of $\{H, C, K, v\}$. If $\hat{\mathscr{B}}$ is an orthonormal basis for $\hat{H}$, with $\mathscr{B} \subset \hat{\mathscr{B}}$, then $D_1D_2 \cdots D_r v = 0$ for $D_i \in \mathscr{C}(\hat{\mathscr{B}})$. If $r = 2$ and $s < \infty$, then a similar argument shows that $D_1D_2 \cdots D_s v = 0$ for all $D_i \in \mathscr{C}(\mathscr{B})$ or $D_i \in \mathscr{C}(\hat{\mathscr{B}})$.

The existence of a total number operator, N, satisfying the hypotheses of the theorem is well known for bosons and fermions. For the other cases there is an integer q such that $D_1D_2 \cdots D_{q+1} v = 0$ for all $D_i \in \mathscr{C}(\mathscr{B})$. In fact, from the connectedness of the unitary group, q is independent of the basis used and

$$C(z_1)C(z_2) \cdots C(z_{q+1})v = 0,$$

for all $z_1, z_2, \ldots, z_{q+1} \in H$ and so $K_j = \{0\}$ if $j > q$. The same is true for any extension of $\{H, C, K, v\}$. Thus,

$$K = K_0 \oplus K_1 \oplus \cdots \oplus K_q$$

since the spaces K_j are mutually orthogonal and the operators $C(z)$ are bounded. Define N as the operator which has the value j on K_j, $0 \leqslant j \leqslant q$. Then N is self-adjoint, $Nv = 0$, and if $w \in K_j$, $C(y)w \in K_{j+1}$ so

$$[N, C(y)]w = C(y)w.$$

Let $\mathscr{B} = \{e_\alpha : \alpha \in I\}$ be an orthonormal basis for H and suppose $w \in K_j$. Let $\epsilon > 0$ be given. We can write $w = w_1 + w_2$, where $w_1 \in K_j(\mathscr{B})$ and $\|w_2\| < \epsilon/2j$. Let J be a finite subset of I such that w_1 is a finite linear combination of terms of the form $D_1D_2 \cdots D_j v$ with $D_i \in \{C(e_\alpha) : \alpha \in J\}$, $1 \leqslant i \leqslant j$. If J' is finite set containing J, then

$$\sum_{\alpha \in J'} n(e_\alpha)w_1 = jw_1$$

and

$$\|\sum_{\alpha \in J'} n(e_\alpha)w_2\| \leqslant j\|w_2\|,$$

so

$$\|\sum_{\alpha \in J'} n(e_\alpha)w - Nw\| \leqslant \|\sum_{\alpha \in J'} n(e_\alpha)w_1 - Nw\| + \|\sum_{\alpha \in J'} n(e_\alpha)w_2\|$$
$$\leqslant \|jw_1 - jw\| + j\|w_2\|$$
$$\leqslant 2j\|w_2\| < \epsilon.$$

Thus $\Sigma_\alpha n(e_\alpha)w$ converges to Nw. This completes the proof of Theorem 3.

As in Theorem 2, when H is finite-dimensional the existence of the vacuum vector need not be assumed. Equation (2.1) can be replaced by a formally equivalent relation, (2.7), which does not have an explicit domain condition. When no convenient domain is given, it is necessary to replace the linearity conditions on $C(\cdot)$ by ones which do not refer to any particular domain and the irreducibility condition must also be modified.

Definition: A quantum structure over H is a collection $\{H, C, K\}$, where H and K are complex Hilbert spaces and C is a function from H into the set of closed densely defined operators of K such that for all $z, y \in H$ and nonzero complex numbers α,

$$C(z+y)\supset C(z)+C(y),$$

and

$$C(\alpha z)=\alpha C(z).$$

A quantum structure is called irreducible if no nontrivial subspace of K simultaneously reduces all $C(z)$ for $z\in H$.

There are at least two distinct methods for generalizing the boson and fermion fields. One method, generalizing the commutation and anticommutation relations, leads to parabosons and parafermions. For these, as for bosons and fermions, irreducible systems have unique vacuums. [See Ref. 2, Theorems 3–6]. The method of generalization discussed here in terms of the number operator does not have this property.

Even though a vacuum vector must exist when H is finite-dimensional, it need not be unique even when $\{H, C, K\}$ is irreducible. An example of this for which $n(z)=C(z)C^*(z)$ is the number operator is given in the next section.

Let

$$V=\{w\in K: C^*(z)w=0, \text{ for all } z\in H\}.$$

V is called the vacuum space of K. We must assume that V does not have dimension greater than one.

Theorem 4: Suppose $\{H, C, K\}$ is an ireducible quantum structure whose vacuum space, V, has dimension zero or one, such that if z is a unit vector of H and P is the projection onto the one-dimensional space spanned by z, then $n(z)=C(z)C^*(z)$ satisfies

$$e^{itn(z)}C(y)e^{-itn(z)}=C(e^{itP}y) \tag{2.7}$$

for all sufficiently small real values of t. If v is a unit vector in V, then $\{H, C, K, v\}$ satisfies the hypotheses of Theorem 3. If H is finite-dimensional then V contains a unit vector.

Proof of Theorem 4: First assume that H is finite-dimensional and let $\mathscr{B}=\{e_i: 1\leqslant i\leqslant d\}$ be an orthonormal basis for H, let $C_i=C(e_i)$ and $n_i=n(e_i)$. Then from Eq. (2.7), $\{n_i: 1\leqslant i\leqslant d\}$ is a set of mutually commuting self-adjoint operators on K. Let $n_i=\int\lambda E_i(d\lambda)$ be the spectral resolutions. There is a nonzero vector $w\in K$ and a bounded interval $\Delta=[0,\beta]$ such that $E_i(\Delta)w=w$ for $1\leqslant i\leqslant d$. Thus $w\in\mathrm{Dom}(C_i^*)$. By using the method of Lemma 2, it follows that $C_i^*w\in\mathrm{Dom}(n_j^k)$,

$$(n_j+\delta_{ij})^k C_i^* w = C_i^* n_j^k w,$$

and

$$E_j([0,\beta-\delta_{ij}])C_i^*w=C_i^*w.$$

By applying C_i^* the appropriate number of times we obtain a unit vector $v\in K$ such that $C_i^*v=0$ for $1\leqslant i\leqslant d$. Thus $C^*(z)v=0$ for all $z\in H$, and V contains a unit vector when H is finite-dimensional.

Now suppose H is arbitrary and that v is a unit vector of V. Let $\mathscr{B}=\{e_\alpha: \alpha\in I\}$ be an orthonormal basis for H and define $C_\alpha=C(e_\alpha)$, $n_\alpha=n(e_\alpha)$ and $\bar{n}_\alpha=C^*(e_\alpha)C(e_\alpha)$. Analogously to Lemma 4, we have that if $n_\alpha w=\gamma w$, then $w\in\mathrm{Dom}(\bar{n}_\alpha)$ and

$$n_\alpha C_\beta w=(\gamma+\delta_{\alpha\beta})C_\beta w. \tag{2.8}$$

If $v\in V$, then $n_\alpha v=0$ so $v\in\mathrm{Dom}(C_\alpha)$ and by induction every elementary vector in $K(\mathscr{B})$ is an eigenvector of each n_α and thus in the domain of each n_α and $\bar{n}_\alpha$. Since this is true for any basis, $v\in D^\infty(\mathscr{A})$.

Let z be a fixed unit vector in H. From Eq. (2.7) it follows that $U=e^{2\pi i n(z)}$ commutes with each $C(y)$. Thus, each spectral projection of U reduces $C(y)$ so the irreducibility implies that U is a scalar. Since $n(z)v=0$, U is the identity and thus the spectrum of $n(z)$ contains only nonnegative integers. Let

$$K_m(z)=\{w\in K: n(z)w=mw\}$$

so that $K=\oplus_{m=0}^\infty K_m(z)$. From Eq. (2.8) it follows that $C(z)$ maps $K_m(z)$ into $K_{m+1}(z)$ and $C(z)$ is bounded on $K_m(z)$. Let $K'_m(z)$ be the closure of $K_m(z)\cap K'$. Then $C(z)$ maps $K'_m(z)$ into $K'_{m+1}(z)$.

Suppose $w\in K'$. For some orthonormal basis $\mathscr{B}$ containing z, $w\in K'(\mathscr{B})$. Each elementary vector in $K'(\mathscr{B})$ is in some $K'_m(z)$ and so $w\in\oplus_{m=0}^\infty K'_m(z)$. Thus, the closure of K' is $K=\oplus_{m=0}^\infty K'_m(z)$ and the following lemma shows that this subspace reduces $C(z)$.

Lemma 8: Suppose C is a closed densely defined operator on a complex Hilbert space K, K_i are mutually orthogonal closed subspaces of K such that $K=\oplus_{i=1}^\infty K_i$, $K_i\subset\mathrm{Dom}(C)\cap\mathrm{Dom}(C^*)$ and $CK_i\subset K_{i+1}$. Let Q_i be the projection onto K_i. Then

a) $w\in\mathrm{Dom}(C)$ if and only if ΣCQ_iw converges

and

b) $w\in\mathrm{Dom}(C^*)$ if and only if ΣC^*Q_iw converges.

Furthermore, suppose that for each i, M_i is a closed subspace of K_i such that $CM_i\subset M_{i+1}$. Let $M=\oplus M_i$ and P be the projection onto M. Then $PC\subset CP$ and $PC^*\subset C^*P$.

Proof of Lemma 8: If $u\in K_i$ and $w\in K_j$, then $\langle C^*u,w\rangle=\langle u,Cw\rangle$. This is zero unless $i=j+1$ so $C^*K_i\subset K_{i-1}$. Now suppose only that $w\in\mathrm{Dom}(C)$ and $u\in K$.

$$\begin{aligned}\langle Q_{i+1}Cw,u\rangle &= \langle Cw,Q_{i+1}u\rangle=\langle w,C^*Q_{i+1}u\rangle\\ &=\langle w,Q_iC^*Q_{i+1}u\rangle=\langle Q_iw,C^*Q_{i+1}u\rangle\\ &=\langle CQ_iw,Q_{i+1}u\rangle=\langle CQ_iw,u\rangle.\end{aligned}$$

Thus $Q_{i+1}Cw=CQ_iw$ so ΣCQ_iw converges. If $\Sigma\, CQ_iw$ congress, since ΣQ_iw converges to w and C is closed, $w\in\mathrm{Dom}(C)$ and $Cw=\Sigma CQ_iw$. This gives a), and b) is done similarly.

Now suppose $CM_i\subset M_{i+1}$ so that as above, $C^*M_i\subset M_{i-1}$. Let P_i be the projection onto M_i. Let $L_i=M_i^\perp\cap K_i$ If $w\in L_i$, then $Cw\in K_{i+1}$. If $u\in M_{i+1}$, $\langle Cw,u\rangle=\langle w,C^*u\rangle=0$, so $Cw\in L_{i+1}$ and thus $CL_i\subset L_{i+1}$ and similarly $C^*L_i\subset L_{i-1}$.

Next suppose only that $w\in\mathrm{Dom}(C)$. Then ΣCQ_iw converges. Since

$$CQ_iw=CP_iw+C(Q_i-P_i)w,$$

and $CP_iw\in M_{i+1}$ while $C(Q_i-P_i)w\in L_{i+1}$,

$$\|CQ_iw\|^2=\|CP_iw\|^2+\|C(Q_i-P_i)w\|^2.$$

Thus $\|CP_iw\|\leqslant\|CQ_iw\|$ and so ΣCP_iw converges and $Pw\in\mathrm{Dom}(C)$. Since $P_{i+1}Cw=CP_iw$,

$$CPw=\Sigma CQ_iPw=\Sigma CP_iw=\Sigma P_{i+1}Cw=PCw.$$

Thus $PC \subset CP$ and similarly or by general theory $PC^* \subset C^*P$. This completes the proof of Lemma 8.

Thus the closure of K' reduces each $C(z)$ and so by the irreducibility, K' is a dense subset of K and v is cyclic for $\mathscr{A}$. From Eq. (2.8) and the linearity of $C(\cdot)$, Eq. (2.1) is satisfied for $w \in K_m(z)$ and thus for $w \in K'$. The proof of Theorem 4 is now complete.

3. EXAMPLES

As already noted in Theorem 3, $r = \infty, s = \infty$ corresponds to bosons and $r = 2, s = \infty$ corresponds to fermions. All examples except for bosons have $C(z)$ as a bounded operator. The simplest example has $r = s = 2$ but this has only one-particle states so it is of little interest. We first give a simple finite-dimensional example to show that for more than one degree of freedom there are statistics other than bosons and fermions which satisfy the hypotheses of Theorem 3.

Example 1: Let H be a three-dimensional Hilbert space with orthonormal basis $\{e_1, e_2, e_3\}$. Let K be a seven-dimensional Hilbert space with orthonormal basis $\{v, u_1, u_2, u_3, w_{12}, w_{13}, w_{23}\}$. If $z = \alpha e_1 + \beta e_2 + \gamma e_3$, let $C(z)$ be the operator on K defined by

$$C(z)v = \alpha u_1 + \beta u_2 + \gamma u_3,$$
$$C(z)u_1 = -\beta w_{12} - \gamma w_{13},$$
$$C(z)u_2 = \alpha w_{12} - \gamma w_{23},$$
$$C(z)u_3 = \alpha w_{13} + \beta w_{23},$$
$$C(z)w_{12} = C(z)w_{13} = C(z)w_{23} = 0.$$

If $n(z) = C(z)C^*(z)$ and $x,y \in H$ the following relations are satisfied:

$$[n(z),C(y)] = \langle y,z \rangle C(z), \tag{3.1}$$
$$C(z)^2 = 0, \tag{3.2}$$
$$C(z)C(y) = -C(y)C(z), \tag{3.3}$$
$$C(z)C(y)C(x) = 0. \tag{3.4}$$

Thus, $C(z)$ behaves like a fermion creation operator until we get to states with three different particles (which do not exist). We next extend Example 1 to the case in which H is infinite-dimensional.

Example 2: Let H be a separable Hilbert space with orthonormal basis $\mathscr{B} = \{e_j : 1 \leqslant j < \infty\}$. Let K be the separable Hilbert space with orthonormal basis

$$\{v, u_j, w_{ij} : 1 \leqslant i < j < \infty\}.$$

Define bounded operators C_k, $1 \leqslant k < \infty$ on K by

$$C_k v = u_k,$$
$$C_k u_j = \begin{cases} w_{kj} & \text{if} \quad k < j, \\ -w_{jk} & \text{if} \quad k > j, \\ 0 & \text{if} \quad k = j, \end{cases}$$
$$C_k w_{ij} = 0.$$

Let $K'(\mathscr{B})$ be the subset of K consisting of finite linear combinations of the given basis vectors of K. If $z = \Sigma \alpha_k e_k$ and $w \in K'(\mathscr{B})$, then $\Sigma \alpha_k C_k w$ converges. Let $C(z)$ be the operator with domain $K'(\mathscr{B})$ such that $C(z)w = \Sigma \alpha_k C_k w$. $C(z)$ is bounded with bound $\|z\|$ and so can be extended to all of K. If $n(z) = C(z)C^*(z)$, then Eq. (3.1)–(3.4) are satisfied on $K'(\mathscr{B})$ and therefore on all of K. Examples 1 and 2 correspond to $r = 2, s = 3$ in Theorem 3.

Example 3: Let $H, \mathscr{B}$ and z be as in Example 1 and let K be a 10-dimensional Hilbert space with orthonormal basis $\{v, u_1, u_2, u_3, w_{12}, w_{13}, w_{23}, w_{11}, w_{22}, w_{33}\}$. Let $C(z)$ be defined by

$$C(z)v = \alpha u_1 + \beta u_2 + \gamma u_3,$$
$$C(z)u_1 = \beta w_{12} + \gamma w_{13} + \alpha(\sqrt{2})w_{11},$$
$$C(z)u_2 = \alpha w_{12} + \gamma w_{23} + \beta(\sqrt{2})w_{22},$$
$$C(z)u_3 = \alpha w_{13} + \beta w_{23} + \gamma(\sqrt{2})w_{33},$$
$$C(z)w_{ij} = 0 \text{ if } 1 \leqslant i \leqslant j \leqslant 3.$$

Equations (3.1) and (3.4) are still satisfied but (3.2) and (3.3) are not. This example can easily be extended to the case in which $\dim(H) = d$ with $3 \leqslant d \leqslant \infty$.

Example 4: We can obtain an example of $r = 3, s = \infty$ by considering the subspace $\tilde{H}$ of H spanned by e_1 and e_2 and the subspace $\tilde{K}$ of K spanned by $\{v, u_1, u_2, w_{12}, w_{11}, w_{22}\}$. $\{H, C, K\}$ of Example 3 is then an extension of the irreducible clothed quantum structure thus obtained.

For parabosons and parafermions a theorem similar to Theorem 4 is obtainable without having to assume the uniqueness of the vacuum [Ref. 2, Theorems 5, 6]. The next example will show that this extra assumption actually is necessary. Let r' be defined by

$$r' = \min\{k : C_1 C_2 \cdots C_k v = 0 \text{ for some } C_i \in \mathscr{C}(\mathscr{B}) \text{ and some } v \in V\}.$$

Note that the definition of r' differs from that of r in Theorem 3 only in that we allow any vacuum v instead of a particular one. The following example has $r' = 2$ but is a mixture of $r = 2$ and $r = 3$ without being decomposable into a direct sum.

Example 5: Let H be a two-dimensional Hilbert space with orthonormal basis $\{e_1, e_2\}$. Let K be a 12-dimensional Hilbert space with orthonormal basis

$$\{v, v', u_1, u_1', u_2, u_2', w_{12}, w_{12}', w_{11}, w_{22}, x_{12}, x_{21}\}.$$

The vacuum space V will be spanned by v and v', with v having $r = 3$ and v' having $r = 2$. The unprimed vectors are built up from v while the primed vectors are built from v'. Let $z = \alpha e_1 + \beta e_2$ and define $C(z)$ by

$$C(z)v = \alpha u_1 + \beta u_2,$$
$$C(z)v' = \alpha u_1' + \beta u_2',$$
$$C(z)u_1 = \alpha(\sqrt{2})w_{11} + \beta w_{12},$$
$$C(z)u_1' = -\beta w_{12}',$$
$$C(z)u_2 = w\alpha_{12} + \beta(\sqrt{2})w_{22},$$
$$C(z)u_2' = \alpha w_{12}',$$
$$C(z)w_{12} = -\tfrac{1}{2}\alpha(\sqrt{2})x_{21} - \tfrac{1}{2}\beta(\sqrt{2})x_{12},$$
$$C(z)w_{12}' = \tfrac{1}{2}\alpha(\sqrt{6})x_{21} - \tfrac{1}{2}\beta(\sqrt{6})x_{12},$$
$$C(z)w_{11} = \beta x_{21},$$
$$C(z)w_{22} = \alpha x_{12},$$
$$C(z)x_{12} = 0,$$
$$C(z)x_{21} = 0.$$

Thus, $C(z)C(z)v' = 0$, but $C(z)C(y)v \neq 0$. Equation (3.1) is satisfied as can be seen by a tedious calculation which is simplified by the fact that $n(e_1)$ and $n(e_2)$ are diagonal.

This is in fact irreducible. Since any nontrivial invariant subspace of K must contain a vacuum vector, it is sufficient to show that if $\bar{v} \in V$ and $\bar{v} \neq 0$ then no proper subspace containing $\bar{v}$ is invariant under all $C(z)$ and $C^*(z)$. To do this it suffices to show that both v and v' can be obtained from $\bar{v}$ by suitable application of creation and annihilation operators. Let $C_1 = C(e_1)$ and $C_2 = C(e_2)$ and assume $\bar{v} = Av + Bv'$. Then $C_1^{*2} C_1^2 \bar{v} = 2Av$ so if $A \neq 0$, v can be obtained while if $B \neq 0$, v' can be obtained from $Bv' = \bar{v} - \frac{1}{2} C_1^{*2} C_1^2 \bar{v}$. It remains to be shown that v can be recovered from v' and that v' can be recovered from v. These follow from

$$C_1^* C_2^* C_1^* C_1^2 C_2 v = \tfrac{1}{2} v + \tfrac{1}{2}(\sqrt{3}) v'$$

and

$$C_1^* C_2^* C_1^* C_1^2 C_2 v' = -\tfrac{1}{2}(\sqrt{3}) v - (3/2) v'.$$

[1] S. Robbins, "A Generalization of the Canonical Commutation and Anti-Commutation Relations," Proc. Amer. Math. Soc. **71** 85–88 (1978).

[2] S. Robbins. "A Uniform Approach to Field Quantization," J. Funct. Anal. **29** 23–36 (1978).

Ground state energy bounds for potentials $|x|^\nu$

R. E. Crandall and Mary Hall Reno [a)]
Department of Physics, Reed College, Portland, Oregon 97202

(Received 16 July 1980; accepted for publication 11 November 1980)

A theory is developed from which both upper and lower analytic bounds on Schrödinger eigenvalues can be obtained. We propose a recursion algorithm with which ground energies for certain potentials can be rigorously bounded to arbitrary precision. These analytic and numerical methods, together with existing techniques, are applied to the ground state problem for power potentials $|x|^\nu$, $\nu > 0$.

PACS numbers: 03.65.Ge, 02.30.Hq, 02.60.Lj

I. INTRODUCTION

For many potentials the one-dimensional Schrödinger equation can be solved only with approximation methods. The well-known WKB method[1,2] was introduced as a semi classical theory, and as such gives relatively weak estimates for ground state energies. The method of Rayleigh[3] and Ritz[4] is effective for ground states but provides only upper bounds. These bounds have nevertheless enjoyed wide success in the domain of atomic and chemical physics.[5,6] Recently, Barnsley[7] extended the work of Barta[8] and of Duffin[9] to develop a method for obtaining lower bounds on eigenvalues. Both the Rayleigh–Ritz and Barnsley methods require carefully chosen test functions for good precision. We shall describe a theoretical approach, valid for certain potential functions, which yields both upper and lower bounds without the use of test functions.

Recent interest in the specific problem of power potentials $V(x) = |x|^\nu$ has been stirred by Turschner[10] who claimed a remarkable closed formula for all bound-state eigenvalues. Subsequently, Crowley and Hill[11] showed that the formula is incorrect, but that the Turschner approach may be a new approximation scheme of considerable power. Since numerical counterexamples figure strongly in the Crowley–Hill rebuttal, the present authors attempted to work out means by which rigorous bounds on eigenvalues can be computed to arbitrary precision. Such bounds, it was felt, could then be used to efficiently test old and new approximation methods.

The Schrödinger equation is taken to be (in units for which $\hbar^2 = 2m$)

$$-\psi''(x) + V(x)\psi(x) = E\psi(x). \tag{1}$$

When bound states exist we denote by $E_n^{(V)}$ the nth bound state egenvalue for the potential V, with $n = 0,1,2,\cdots$. In what follows we shall always assume that V is a member of a class of potentials denoted by M. This is the class of all symmetric, nonnegative, unbounded potential functions which vanish at the origin and possess the following growth property for $x > 0$:

$$\frac{d^2}{dx^2}\log V < 0 < \frac{d}{dx}\log V,$$

where the relevant derivatives are assumed continuous on $(0,\infty)$. It is evident from the growth condition that any $V \in M$ is strictly increasing, and V'/V strictly decreasing, on $(0,\infty)$.

Every power potential $V(x) = |x|^\nu$ for $\nu > 0$ is in the class M. The class also contains functions not of polynomial growth, for example the potential $V(x) = \sinh^\nu|x|$ is in the class M whenever $\nu > 0$.

In Sec. II we initiate the theoretical development by establishing theorems which pertain to a nonlinear equivalent of the Schrödinger equation (1). Particular attention is given to the case $E = E_0^{(V)}$, the ground state energy. In Sec. III we describe an algorithm for computing rigorous bounds on $E_0^{(V)}$, and tabulate numerical results for various power potentials. In Sec. IV analytic upper and lower bounds are derived for $E_0^{(|x|^\nu)}$, $\nu > 2$, using the methods of Sec. II. For $\nu < 2$, analytic bounds are easier to generate from the Rayleigh–Ritz and Barnsley methods. Such bounds are derived in Sec. V. Finally, the problem of estimating higher states is discussed in Sec. VI.

We first observe that ground energies $E_0^{(|x|^\nu)}$ for some ν can be given exactly using standard techniques.[12] The simple harmonic oscillator and absolute-linear well cases, $\nu = 2,1$ respectively, are given by

$$E_0^{(x^2)} = 1,$$

$$E_0^{(|x|)} = w_1 = 1.018\,792\,97\ldots,$$

where w_1 is the first positive zero of the Airy derivative $\mathrm{Ai}'(-z)$.[13] We also expect on somewhat intuitive grounds that

$$\lim_{\nu\to 0^+} E_0^{(|x|^\nu)} = 1,$$

$$\lim_{\nu\to\infty} E_0^{(|x|^\nu)} = \pi^2/4.$$

The former limit is that of a progressively thinner potential well of essentially unit height, and the latter is the infinite square well limit. We shall eventually be able to prove that both limits are correct, as the relevant upper and lower analytic bounds will converge to the values indicated.

II. THE TRAJECTORY EQUATIONS

Central to the present treatment is a certain transformation of the Schrödinger equation (1). For a state ψ of energy $E > 0$ (not necessarily bound) and even parity we assign

[a)] Work performed in connection with thesis research.

$$\gamma_E(z) = \tan^{-1}\left[\frac{-\psi'(z/\sqrt{E})}{\sqrt{E}\,\psi(z/\sqrt{E})}\right]. \quad (2)$$

The trajectory γ_E will satisfy the transformed Schrödinger equation

$$\gamma_E'(z) = 1 - \left[V(z/\sqrt{E})/E\right]\cos^2\gamma_E(z), \quad (3)$$

together with the boundary condition

$$\gamma_E(0) = 0. \quad (4)$$

Odd-parity states can be handled in the same way, except that the boundary condition is then replaced by $\gamma_E(0) = -\pi/2$. When E is so large that an even-parity state of energy E possesses zero-crossings, we allow $\tan^{-1}$ to pass continuously through any required number of intercepts $(2n+1)\pi/2$. Equivalent to (3) for even-parity states is the integral equation

$$\gamma_E(z) = z - (1/E)\int_0^z V\left(\frac{u}{\sqrt{E}}\right)\cos^2\gamma_E(u)\,du. \quad (5)$$

For any given $V\in M$, we next consider the collection of all E-indexed trajectories defined by (3) and (4). This collection possesses a remarkable topological property, namely, no two distinct trajectories can intersect for positive z. This behavior stands in sharp contrast to that of the E-indexed even-parity wave functions whose graphs interlace in a complicated manner.

To prove the nonintersection property, we first investigate the behavior of any pair $\psi_E(x),\psi_F(x)$ of even-parity (nontrivial) wavefunctions, where the energies satisfy $E>F>0$. It is a result of standard Sturm-Liouville theory[14,15] that if x_n,x_{n+1} are consecutive zeros of $\psi_F(x)$, then $\psi_E(x)$ possesses a zero in the open interval (x_n,x_{n+1}). It is a matter of simple combinatorics to show from this that ψ_E and ψ_F cannot have a common jth positive zero. Define $G(z) = \gamma_E(z) - \gamma_F(z)$ and assume for some $z_1>0$ that $G(z_1) = 0$. Then from (3) and the monotonicity of V, we have either $G'(z_1)>0$ or $\cos^2\gamma_E(z_1) = 0$. The latter alternative is ruled out since it implies, by definition (2) and our extension of the $\tan^{-1}$ function, that γ_E,γ_F have a common jth zero. Thus G has positive slope at z_1 and at any other of its positive zeros. Such a function cannot be positive at any point in the interval $(0,z_1)$. However, the integrand in Eq. (5) is $V(u/\sqrt{E})[1-O(\gamma_E^2)]$ for small γ_E, so the monotonicity of V forces $\gamma_E>\gamma_F$, that is, $G>0$, on some open interval $(0,\delta)$ of the z axis. This contradiction stems from the assumption that a positive zero z_1 of G exists. Thus G is positive definite on $(0,\infty)$ and we conclude:

Theorem 1: Let $E>F>0$. Then, for all positive z, $\gamma_E(z)>\gamma_F(z)$.

The nonintersection property embodied in Theorem 1 is depicted in Fig. 1 for the quartic potential $V(x)=x^4$. It can be shown that for any $V\in M$, each trajectory is asymptotic to some multiple $n\pi/2$, where n is one of the integers $-1, 1, 3, 5,\cdots$. Even-parity bound state energies are precisely those E for which the relevant asymptote is approached from below. These remarks also hold for odd-parity states subject to suitable modification of condition (4).

It is now apparent that the monotonicity restriction on the members of the class M enables us, by way of Theorem 1,

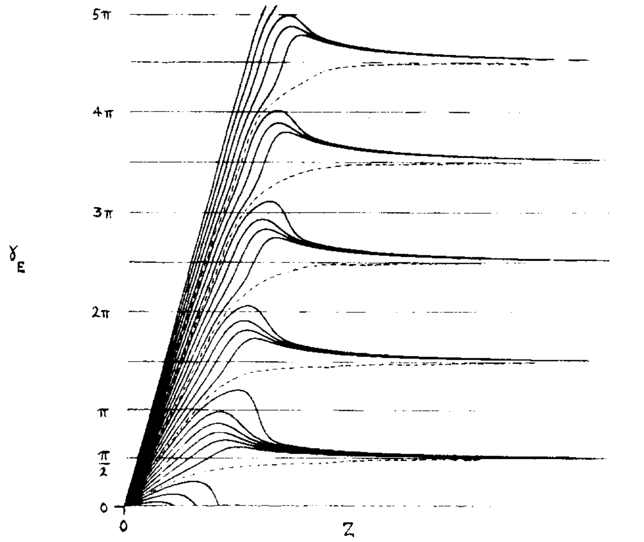


FIG. 1. Computer-generated plot of trajectories $\gamma_E(z)$ for the quartic potential x^4. Each trajectory corresponds to a different E. The manually inserted dashed curves are even-parity bound state trajectories.

to bound eigenvalues by the method of bounding trajectories themselves. We shall presently focus our attention on the ground state trajectory γ_{E_0}. Through a series of theorems it will be shown that this trajectory is monotone increasing, convex downward, and asymptotic to $\pi/2$.

Since the ground state wavefunction $\psi_0(x)$ will have no zeros,[14] it follows that $|\gamma_{E_0}(z)|<\pi/2$ for all real z. Using standard techniques[15] it is straightforward to show that for positive x, $\psi_0'(x)$ has itself no zeros. These observations and Eq. (2) dictate the possible range of γ_{E_0}.

Theorem 2: For $z>0$, $0<\gamma_{E_0}(z)<\pi/2$.

This theorem in turn places restrictions on the derivative $\gamma_{E_0}'(z)$. We can show that this derivative is positive for all positive z. Note that, on the basis of (3), $\gamma_{E_0}'(0)=1$, so it is enough to show that γ_{E_0}' has no positive zeros. The second derivative is

$$\gamma_{E_0}''(z) = (1-\gamma_{E_0}')\left(2\gamma_{E_0}'\tan\gamma_{E_0} - \frac{V'(z/\sqrt{E_0})}{\sqrt{E_0}\,V(z/\sqrt{E_0})}\right), \quad (6)$$

valid for positive z. Now assume that $\gamma_{E_0}'(z_1)=0$ for some $z_1>0$. Since $\log V$ has positive derivative, we have from (6) that $\gamma_{E_0}''(z_1)<0$, hence for some $z_2>z_1$, $\gamma_{E_0}'(z_2)$ is negative. This in turn implies, on the basis of (3) and the monotonicity properties of V and of $\cos^2$, that γ_{E_0} must have some zero $z_3>z_2$. This contradicts Theorem 2, so the assumption that z_1 exists is untenable. This establishes the following:

Theorem 3: $\gamma_{E_0}(z)$ is monotone increasing for positive z.

We have shown with Theorems 2 and 3 that the ground state trajectory is monotone increasing and bounded. From the fact that V diverges we infer by Eq. (3) that γ_{E_0} has the limit $\pi/2$ as $z\to\infty$. A concise summary of the ground state trajectory is the following:

Theorem 4: $\sup_{z>0}\gamma_{E_0}(z) = \pi/2, \quad \inf_{z>0}\gamma_{E_0}(z) = 0.$

Theorems 1 and 4 can be combined to give Theorem 5.

Theorem 5: If for some positive $z,\gamma_E(z)>\pi/2$, then

$E > E_0^{(V)}$. If for some positive $z, \gamma_E(z) < 0$, then $E < E_0^{(V)}$.

This result means that if for some E the trajectory γ_E can be rigorously (either by numerics or by analysis) bounded below by some function that eventually exceeds $\pi/2$, then E is an upper bound on $E_0^{(V)}$. A lower bound is obtained in a similar fashion. The behavior of the trajectories for energies E near the ground state energy is schematized in Fig. 2.

In order to develop an effective computational algorithm we must know a little more about the behavior of the derivative γ'_{E_0}. It can be shown from (6) that the ground state trajectory is in fact concave downward. Observe first that from (3), $\gamma'_{E_0}(z) < 1$ for $z > 0$. Thus any positive zero, say z_1, of γ_{E_0} must be a zero of the second factor in Eq. (6). But V'/V is monotone decreasing, so the existence of such a z_1 implies that for all $z > z_1, \gamma'_{E_0}(z) \geqslant \gamma'_{E_0}(z_1)$. This contradicts Theorem 2, and we have established the following:

Theorem 6: $\gamma'_{E_0}(z)$ is monotone decreasing for positive z.

Corollary: $V(z/\sqrt{E_0}) \cos^2\gamma_{E_0}(z)$ is monotone increasing for $z > 0$.

The corollary follows directly from the theorem and Eq. (3).

III. RECURSION ALGORITHM

Numerical estimates on the integral $(t > 0)$,

$$I(z,t) = \int_z^{z+t} V(u/\sqrt{E_0}) \cos^2\gamma_{E_0}(u) du, \tag{7}$$

will prove useful for obtaining bounds on $E_0^{(V)}$. From the corollary to Theorem 6 we have

$$I(z,t) \geqslant t V(z/\sqrt{E_0}) \cos^2\gamma_{E_0}(z). \tag{8}$$

From Theorem 3 and the monotonicity of V we infer

$$I(z,t) \leqslant t V((z+t)/\sqrt{E_0}) \cos^2\gamma_{E_0}(z). \tag{9}$$

Now from the integral trajectory equation (5) we write

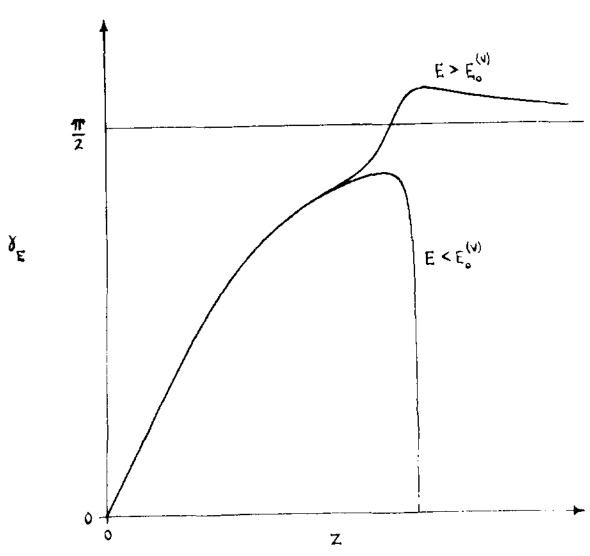


FIG. 2. Plot of two trajectories for energies near the ground energy, showing the inequalities resulting from Theorem 5. The actual potential used for this plot is $V(x) = x^4$. The two energies differ from the true ground energy by only $\pm 0.01\%$.

$$\gamma_{E_0}(z+t) = \gamma_E(z) + t - I(z,t)/E_0. \tag{10}$$

From (8), (9), and (10) it is evident that the ground state trajectory can be bounded with appropriate recursion relations. Theorem 5 gives a test on the appropriate sequences of real numbers to be computed, and we have the following:

Theorem 7: Choose $E, t > 0$ and define sequences G_k, H_k by

$$G_0 = H_0 = 0,$$

$$G_{k+1} = G_k + t - tE^{-1} V(kt/\sqrt{E}) \cos^2 G_k,$$

$$H_{k+1} = H_k + t - tE^{-1} V((k+1)t/\sqrt{E}) \cos^2 H_k.$$

If some H_n exceeds $\pi/2$, then $E > E_0^{(V)}$. If some G_n is negative, then $E < E_0^{(V)}$.

One feature of Theorem 7 is its validity for any initial increment t. If the chosen E value happens to lie very close to the true ground energy, then only for very small t will one of the inequalities $H_n > \pi/2, G_n < 0$ be true for some n.

In actual machine implentation, the function $\cos^2$ must be itself rigorously bounded, preferably with rational bounds. The t, E^2 can be chosen rational, and the recursion relations can be iterated with integer arithmetic.

Values for various bounds on $E_0^{(|x|^\nu)}$ are tabulated in Table I. Some potentials have more precise bounds simply because more machine time was allocated to them. For the purpose of testing the analytic methods of the next sections, more bounds are plotted in Fig. 3.

The exact result for $\nu = 1$ is the Airy zero w_1 as discussed in Sec. I. Table I entries for the quartic $(V = x^4)$ potential are consistent with an independent, nonrigorous estimate submitted to the authors by M. A. Penk, who computed

$$E_0^{(x^4)} \approx 1.060\,363\,090\,484\,1820...$$

This number is plausibly (though not yet provably) correct to 14 decimals, since Penk's numbers for known cases were that accurate.

In Fig. 3 there is apparently an absolute minimum at $\nu \approx 1.8$. This is consistent with standard perturbation theory as applied to the oscillator ground state $\psi_0(x) = N \exp(-x^2/2)$. The derivate

$$\frac{\partial E_0^{(|x|^\nu)}}{\partial \nu}\Big|_{\nu=2} = \int_{-\infty}^{\infty} \psi_0^2(x) x^2 \log x \, dx$$
$$= (1/2\sqrt{\pi}) \Gamma'(3/2)$$

being positive, is consistent with Fig. 3.

IV. ANALYTIC BOUNDS FOR $V = |x|^\nu$, $\nu > 2$

The recursion algorithm embodied in Theorem 7 is applicable only to isolated values of ν in the power potential problem. However, continuous bounds on $E_0^{(|x|^\nu)}$ can be obtained from a more detailed analysis of the ground state trajectory. Let $\nu > 2$ and define a function $h(z)$ by

$$\tan^2\gamma_{E_0}(z) = z^\nu/p + h(z), \tag{11}$$

where $p = [E_0^{(|x|^\nu)}]^{1+\nu/2}$. Then h satisfies the differential equation [here and elsewhere the (1/2)-power denotes positive root]

TABLE I. Bounds on ground energies $E_0^{(|x|^\nu)}$, various ν, obtained with the algorithm of Theorem 7. All implied inequalities are strict ones, except for the solvable cases $\nu = 1,2$.

ν	Lower bound	Upper bound
0.05	1.0498	1.0508
0.1	1.0687	1.0689
0.2	1.0798	1.08
0.22	1.080 078	1.080 127
0.24	1.080 029	1.080 078
0.3	1.0771	1.0773
0.5	1.0595	1.0597
1	w_1	w_1
1.5	1.0011	1.0013
1.7	0.9991	0.9993
1.8	0.9989	0.9991
1.9	0.9992	0.9994
2	1	1
2.1	1.0009	1.002
3	1.022	1.024
4	1.060 3618	1.060 3624
5	1.102	1.1026
6	1.144	1.146
7	1.186	1.187
8	1.225	1.227
9	1.263	1.264
10	1.298	1.3
16	1.472	1.474
32	1.743	1.744
64	1.9819	1.9825
128	2.291	2.292
256	2.333	2.334
1024	2.439	2.44

$$h' = (1 + h)(z^\nu/p + h)^{1/2} - \nu z^{\nu-1}/p, \qquad (12)$$

subject to the condition $h(0) = 0$. It is easy to show from Theorem 3 that $h(z) > -1$ for all positive z. These observations will now be used to show that h is positive on $(0,\infty)$. First, since $\nu > 2$, $\lim_{\nu\to 0^+} h^2/z^2 = \lim_{\nu\to 0^+} [\gamma_{E_0}/z]^2 = 1$, so h is positive on some interval $(0,\epsilon)$. Let $F(z) = (z^\nu/p)^{1/2} - \nu z^{\nu-1}/p$. If h has any positive zeros, then there must be one such, say z_1 with $h'(z_1) = F(z_1) \leqslant 0$. But as $\nu > 2$, F has itself just one positive zero to the right of which F is negative. Thus for all $z > z_1$, $h(z) \leqslant 0$, and furthermore (12) thus dictates that $h'(z) \leqslant F(z)$ on (z_1, ∞). But this contradicts the fact that $h(z) > -1$ for positive z. Therefore h is positive on $(0,\infty)$ and we have the following:

Theorem 8: Let $V(x) = |x|^\nu$, $\nu > 2$, and $p = [E_0^{(|x|^\nu)}]^{1+\nu/2}$. Then for all positive z, $\tan^2\gamma_{E_0} > z^\nu/p$.

This theorem amounts to a bound on the ground state trajectory sufficiently tight to establish the upper bound:

$$E_0^{(|x|^\nu)} < \left[\frac{\nu}{2}\sin\frac{\pi}{\nu}\right]^{2\nu/(\nu+2)}, \quad \nu > 2. \qquad (13)$$

The argument runs as follows. The result of Theorem 8 can be paraphased

$$\cos^2\gamma_{E_0} < \frac{1}{1 + z^\nu/p}, \quad z > 0 \qquad (14)$$

which together with Theorem 4 and the integral equation (5) gives

$$\frac{\pi}{2} > \int_0^\infty \frac{dz}{1 + z^\nu/p}. \qquad (15)$$

The integral can be evaluated and the result is (13).

The bound (13) is exact in both limits $\nu \to 2^+$ and $\nu \to \infty$. It is evident from (13) that for all $\nu > 2$,

$$E_0^{(|x|^\nu)} < \pi^2/4. \qquad (16)$$

It will be seen in Sec. V that (16) also holds for $0 < \nu < 2$ so that the infinite square well ground energy $\pi^2/4$ is an absolute upper bound for the power potential problem, positive ν. It should be remarked that even though the infinite square well is in some sense a geometrically extreme case, the abso-

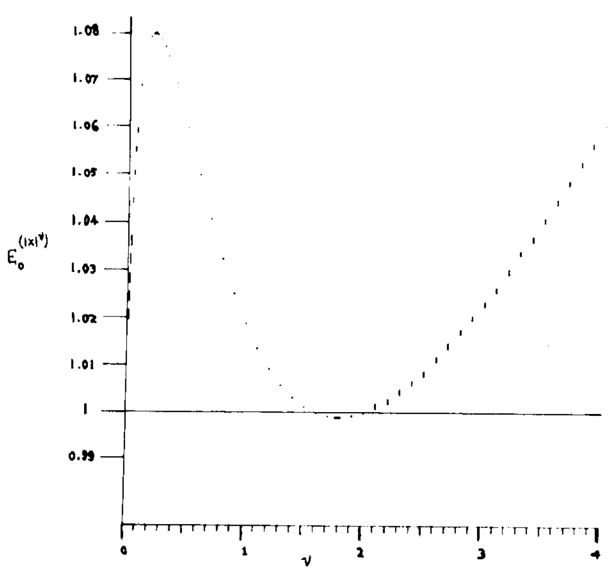


FIG. 3. Plot of rigorous upper and lower bounds on ground energies for $V = |x|^\nu$. Error bars appear for various ν. In some cases, such as $\nu = 4$, the bars appear as dots (see Table I).

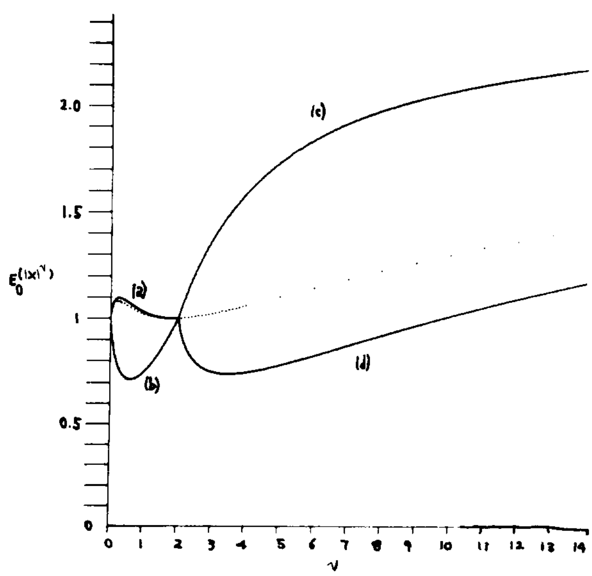


FIG. 4. Continuous bounds on $E_0^{(|x|^\nu)}$. (a) is a Rayleigh–Ritz bound [Eq. (22)]; (b) is a Barnsley bound [Eq. (24)]. The bounds (c) and (d) arise from the trajectory theory [Eqs. (13) and (18), respectively]. All four bounding segments are exact a their endpoints, which read, from left to right, $\nu \to 0^+$, $\nu = 2$, $\nu \to \infty$. The dotted curve represents the exact ground energy.

lute bound $\pi^2/4$ is not obvious, since any positive power of $|x|$ is sometimes less and sometimes greater than any other positive power.

A lower bound for $\nu > 2$ can also be derived on the basis of Theorem 8. The theorem can be paraphrased

$$E_0^{(|x|^\nu)} > [z^\nu \cot^2 \gamma_{E_0}(z)]^{2/(\nu+2)} \tag{17}$$

for all $z > 0$. It is clear from Eq. (5) that $\gamma_{E_0}(z)$ is less than z for $z > 0$. A simple lower bound can thus be written (see Fig. 4):

$$E_0^{(|x|^\nu)} \geqslant \sup_{0<z<\pi/2} [z^\nu \cot^2 z]^{2/(\nu+2)}, \tag{18}$$

It is straightforward to show that this bound is correct in the limits $\nu \to 2^+$ and $\nu \to \infty$. It is possible to obtain better bounds by strengthening the inequality $\gamma_{E_0}(z) \leqslant z$. For example, the factor $\cot^2 z$ in (18) can be replaced with

$$\cot^2\left(z - \frac{1}{p}\int_0^z u^\nu \cos^2 u \, du\right), \tag{19}$$

where use has been made again of Eq. (5). Though the number p appears implicitly in (19), it can be replaced itself with any good upper bound such as that arising from (16).

In the spirit of completeness we now turn to the problem of establishing continuous bounds for the region $0 < \nu < 2$.

V. EXISTING METHODS FOR $\nu < 2$

Owing to the failure of Theorem 8 for $\nu < 2$, the methods of the last section cannot be applied directly for these small ν. Existing techniques give reasonable bounds over the finite interval $0 < \nu < 2$. We include these here in order to complete our search for continuous upper and lower bounds for all positive ν.

A Rayleigh–Ritz bound on $E_0^{(V)}$ can be obtained from the inequality

$$E_0^{(V)} \leqslant \int_{-\infty}^{\infty} \psi_t H \psi_t \, dx, \tag{20}$$

where $\psi_t(x)$ is a normalized, real test function and H is the Hamiltonian operator $-d^2/dx^2 + V$. For $V(x) = |x|^\nu$, a particularly effective choice of test function is

$$\psi_t(x) = N \exp\left(-\sqrt{2\nu}\,\frac{|x|^{1+\nu/2}}{\nu+2}\right), \tag{21}$$

where N is a ν-dependent normalization factor. The upper bound from (20) is

$$E_0^{(|x|^\nu)} \leqslant \left[\frac{\nu(\nu+2)^\nu}{2^{2\nu+1}}\right]^{2/(\nu+2)} \frac{\Gamma(\nu/(\nu+2))}{\Gamma(2/(\nu+2))} \tag{22}$$

This bound is exact for $\nu = 2$ and in the limit $\nu \to 0^+$. For large ν, however, the bound is weak, owing to the failure of (21) to well-approximate the infinite square well (cosine) ground state. The bound (22) is plotted in Fig. 4.

Barnsley's method can be used to find lower bounds for small ν as follows. Let ψ_t satisfy the criteria (a) ψ_t is positive-definite, symmetric, (b) ψ_t is twice-differentiable. Then Barnsley's theorems state effectively that[7]

$$E_0^{(V)} \geqslant \inf_{x>0} \frac{H\psi_t}{\psi_t}. \tag{23}$$

The test function (21) gives a bound when $0 < \nu < 2$ as

$$E_0^{(|x|^\nu)} \geqslant \left[\frac{\nu}{8}\right]^{\nu/(\nu+2)} \left(1 + \frac{\nu}{2}\right). \tag{24}$$

This bound is exact in both limits $\nu \to 0^+$ and $\nu \to 2^-$, as indicated in Fig. 4.

In attempting to apply the Barnsley method to cases $\nu > 2$, we found that an extension of said method greatly improves its applicability. In many instances a test function ψ_t closely approximates the true eigenstate over some finite region, but the overall infimum (23) is weak, or even trivial (negative). Success is more likely if the domain of the infimum can be collapsed to a finite interval. Assume, for example, that there happens to exist a point $y > 0$ at which ψ_t' is negative but

$$\frac{d^2}{dx^2}\log\psi_t,$$

vanishes. Consider the function ϕ_t defined on the nonnegative real axis by

$$\phi_t(x) = \begin{cases} \psi_t(x), & x \leqslant y, \\ \psi_t(y)\exp\{[\psi_t'(y)/\psi_t(y)](x-y)\}, & x > y, \end{cases} \tag{25}$$

and summarized for negative x by $\phi_t(x) = \phi_t(-x)$. Then it is easy to show that ϕ_t itself satisfies the Barnsley criteria, and moreover that

$$E_0^{(V)} \geqslant \inf_{0<x<y} \frac{H\phi_t}{\phi_t}, \tag{26}$$

so that the infimum need only be computed over a finite region. A good example of this extended method is provided by the quartic potential $V = x^4$. Choose

$$\psi_t(x) = \exp(-x^2/2 - x^4/12 + x^6/90). \tag{27}$$

This sort of test function arises naturally if we attempt to transform the Schrödinger equation (1) by $\psi_t(x) = \exp[-\int_0^x g(u)\,du]$ and perform a Taylor expansion on g to several terms. Note that the standard Barnsley infimum for the function (27) is negative infinity. However, there is a fortuitous choice:

$$y = \{\tfrac{3}{2}[1 + (\tfrac{7}{3})^{1/2}]\}^{1/2}$$

such that the test function (25) gives the bound

$$E_0^{(x^4)} > 1.$$

From Table I we have the value 1.060...for the energy, showing that this extended method can yield tight bounds.

It is difficult to produce analytic bounds for $\nu > 2$ using even the extended Barnsley method. It was for this reason that the trajectory method was developed.

VI. CONNECTION WITH WKB THEORY

The WKB approximation of the nth energy eigenvalue $E_n^{(V)}$ is taken to be that solution E of[16]

$$(n + 1/2)\pi/2 = \int_0^{x_c} (E - V(u))^{1/2} du, \tag{28}$$

where x_c satisfies $V(x_c) = E$. For power potentials $V = |x|^\nu$, E can be obtained in closed form.[11] The resulting expression is relatively poor for the ground energy ($n = 0$), and it can be shown by methodical comparisons with Fig. 3 that the WKB value is generally too large for $v < 2$ but too small for $v > 2$. Furthermore, the WKB ground energy does not have the proper limit as $v \to \infty$. Nevertheless, for large n the WKB approximation will be asymptotically valid, and serve as a test of any alternate approximation theory for high quantum numbers.

We shall now establish a connection between the trajectory theory and WKB theory. Let $\gamma_{E_n}(z)$ be the trajectory for the nth bound-state energy, n even (see Fig. 1). Denote by z_j the z coordinates of the intercepts

$$\gamma_{E_n}(z_j) = j\pi/2, \quad j = 0,1,\dots,n.$$

For comparison with the WKB integral (28) we also set $x_j = z_j/\sqrt{E_n}$. We invoke a key integral identity

$$\frac{\pi}{2}\frac{1}{(1-D)^{1/2}} = \int_0^{\pi/2} \frac{d\gamma}{1 - D\cos^2\gamma}, \tag{29}$$

valid for $D < 1$, and integrate over any one of the n regions between lines $\gamma = \pi j/2$, $\gamma = \pi(j+1)/2$, $j < n$, to get the inequality ($E = E_n$):

$$[E - V(x_{j+1})]^{1/2} < \frac{\pi}{2(x_{j+1} - x_j)} < [E - V(x_j)]^{1/2}, \tag{30}$$

where use has been made of (3) and the monotonicity property of V. If we denote $\Delta_j = x_{j+1} - x_j$ and perform the appropriate summation over regions, we obtain

$$\sum_{j=0}^{n-1} [E - V(x_{j+1})]^{1/2}\Delta x_j < \frac{n\pi}{2} < \sum_{j=0}^{n-1} [E - V(x_j)]^{1/2}\Delta x_j. \tag{31}$$

This is the difference analog of (28) and presumably has applications to the problem of bounding higher energy eigenvalues for $V \in M$.

The natural approximation implied by (31) is that solution E of

$$n\pi/2 = \int_0^{x_n} [E - V(u)]^{1/2} du. \tag{32}$$

Equation (32) and inequalities (31) do not explicitly involve the classical turning point x_c. Instead, knowledge of the intercepts $x_j = z_j/\sqrt{E}$ is required; in particular it would be of interest to determine the last intercept x_n, corresponding to the last right-hand critical point of the nth bound-state wave function. We shall presently sketch an argument which justifies the asymptotic equivalence of (32) and the WKB integral (28).

Observe the last segment of the trajectory, defined by $n\pi/2 < \gamma_E < \pi(n+1)/2$, which did not figure into the inequalities (31), is just a rigid translate of the ground state trajectory for the potential

$$W(x) = V(x + x_n) - V(x_n), \quad x > 0, \tag{33}$$

It follows from the trajectory equation (3) that x_c and x_n are related implicitly by

$$V(x_c) = V(x_n) + E_0^{(W)}. \tag{34}$$

This equation is exact but generally extremely difficult to solve for x_n. It is possible, however, to use (34) to give an estimate of the integral

$$J = \int_{x_n}^{x_c} [E - V(u)]^{1/2} du. \tag{35}$$

It is evident that the WKB integral and the integral of (32) differ by the magnitude of $J - \pi/4$. From (34) and the monotonicity properties of V it is possible to show that

$$J < \tfrac{2}{3}A\,[V'(x_n)/V'(x_c)]^{3/2}[V(x_c)/V(x_n)]^2, \tag{36}$$

where

$$A = [E_0^{(W)}/V'^{2/3}(x_n)]^{3/2}. \tag{37}$$

If we now estimate $W(x) \sim xV'(x_n)$ for large n, finite $x > 0$, we have from the scaling properties of the absolute-linear potential that $A \sim w_1^{3/2}$, where w_1 is the Airy zero discussed in Sec. I. We also infer from (34) that $V(x_n) \sim V(x_c)$, so that

$$J \sim \tfrac{2}{3}w_1^{3/2} \approx 0.69. \tag{38}$$

Evidently the approximation (32) can be written for large n

$$n\pi/2 + J = \int_0^{x_c} [E - V(u)]^{1/2} du. \tag{39}$$

It is interesting that J is so close to $\pi/4$, the equivalent WKB value in (28). The last step in achieving (39), namely that $A \sim w_1^{3/2}$, is heuristic since we do not yet have a rigorous procedure for bounding $E_0^{(W)}$. However, for $V \in M$ it is possible to prove $J = o(n)$, which establishes the asymptotic equivalence of (32) and (28).

It would be fruitful to extend the trajectory method so that rigorous error terms on the WKB approximation could be obtained. Likewise useful would be a refinement of the algorithm of Theorem 7 for application to higher states. The main obstacle to such a refinement is that the integral $I(z,t)$ of Eq. (7) cannot be easily bounded by simple terms when $n > 0$.

Finally, the present methods should be extended to higher-dimensional cases, and also to a wider class of potentials. Such a class might include, for example, the quarkonium potentials of recent interest.[17,18]

ACKNOWLEDGMENTS

The authors are indebted to T. Wieting and R. Mayer for their numerous contributions to the theoretical development. Many of the key ideas arose in the discussions with N. A. Wheeler.

[1] G. Wentzel, Z. Physik **38**, 518 (1926); H. A. Kramers, Z. Physik **39**, 828 (1926); L. Brillouin, C.R. Acad. Sci. Paris **183**, 24 (1926).
[2] R. E. Langer, Phys. Rev. **51**, 669 (1937).
[3] Lord Rayleigh, *Theory of Sound* (Macmillan, London, 1937), 2nd rev. ed., Vol. 1.
[4] W. Ritz, J. Reine Agnew. Math. **135**, 1 (1908).
[5] C. L. Pekeris, Phys. Rev. **112**, 1649 (1958).
[6] E. Hylleraas, Z. Physik **54**, 347 (1929).
[7] M. F. Barnsley, J. Phys. A **11**, 55 (1978).
[8] J. Barta, C. R. Acad. Sci. Paris **204**, 472 (1937).
[9] R. J. Duffin, Phys. Rev. **71**, 827 (1947).
[10] H. Turschner, J. Phys. A Gen. Phys. **12**, 451 (1979).
[11] B. J. B. Crowley and T. F. Hill, J. Phys. A Gen. Phys. **12**, L223 (1979).

[12] H. and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge U. P., Cambridge, 1956), 3rd ed.
[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York 1964).
[14] W. E. Milne, Trans. Am. Math. Soc. **30**, 797 (1929).
[15] J. L. Powell and B. Crasemann, *Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1961).
[16] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970).
[17] E. C. Poggio and H. J. Schnitzer, Phys. Rev. D **19**, 1557 (1979).
[18] N. Barik and S. N. Jena, Phys. Rev. D **21**, 2647 (1980).

The three magnon bound-state equation in one dimension

Chanchal K. Majumdar
Physics Department, Calcutta University, Calcutta, 700009, India
S. Banerjee
S. N. Bose Institute of Physical Sciences, Calcutta University, 92, Acharya Prafulla Chandra Road, Calcutta 700009, India

(Received 9 September 1980; accepted for publication 21 November 1980)

The mathematical analogy of the three magnon bound-state equation with other momentum-space integral equations is studied. It is shown that a variable transformation, similar to Wick's transformation in the Bethe–Salpeter equation, leads to alternative methods of complete analytic solutions of this equation.

PACS numbers: 03.65.Ge, 03.65.Db, 75.90. + w

I. INTRODUCTION

The solution[1] of the three magnon bound-state equation in one dimension provides an interesting example of a solvable nontrivial three-body problem in quantum mechanics, where the homogeneous Faddeev equations have been shown to be analytically tractable. It is worth while to explore its connection with other equations arising in physical problems and to analyze the mathematical features that may be of more general applicability. In this paper we shall point out its similarity and differences with two other well-studied equations in momentum space—the hydrogen atom[2] and the Bethe–Salpeter[3] (BS) equations. In particular, we shall show that transformation analogous to Wick's transformation for the BS equations can be used to obtain alternative methods of solution of the three magnon bound-state equation.

The equation for the three magnon bound state is

$$\psi(p_1) = \frac{2\cos^2\frac{1}{2}p_1}{\alpha - \cos(K-p_1)}\left(1-\frac{f}{d}\right)^{-1}\frac{1}{\pi}\int_{-\pi}^{\pi}\frac{[\sigma\cos\frac{1}{2}p_1 - \cos(K-\frac{1}{2}p_1-p_2)]\cos(K-p_1-\frac{1}{2}p_2)}{\frac{1}{2}[-\alpha+\cos(K-p_1)+\cos(K-p_2)+\cos(K-p_1-p_2)]}\psi(p_2)\,dp_2 \quad (1)$$

with

$$\alpha = a/\sigma, \quad (2)$$

$$a = 3 - E, \quad (3)$$

$$f = \alpha - \cos(K-p_1) - 2\sigma\cos^2\tfrac{1}{2}p_1, \quad (4)$$

$$d = \{[\alpha - \cos(K-p_1)]^2 - 4\cos^2\tfrac{1}{2}p_1\}^{1/2}. \quad (5)$$

E is the eigenvalue, σ is the longitudinal anisotropy parameter, and K is the momentum of the center of mass of the three-body system. Equation (1) is an integral equation in momentum space, linear in the wavefunction but highly nonlinear in the eigenvalue E.

Consider now the hydrogen atom equation

$$\Phi(\mathbf{p}) = \frac{me^2}{\pi\hbar(p^2+a^2)}\int\frac{d\mathbf{p}'}{|\mathbf{p}-\mathbf{p}'|^2}\Phi(\mathbf{p}'), \quad (6)$$

with the eigenvalue E given by $a^2 = 2m|E|$. Let us also take the BS equation, as in Wick's paper,[3]

$$\Phi(p) = \frac{\lambda}{\pi^2(p^2+m_a^2)(p^2+m_b^2)}\int\frac{d^4k}{(p-k)^2+K^2}\Phi(k). \quad (7)$$

Let Φ be a function of p^2 only. On carrying out angular integration and substituting

$$p^2 = s, \qquad \Phi(p) = u(s), \quad (8)$$

we get

$$u(s) = \frac{2\lambda}{(s+m_a^2)(s+m_b^2)} \times\int_0^\infty\frac{tu(t)\,dt}{\{s+t+K^2+[(s+t+K^2)^2-4st]^{1/2}\}}. \quad (9)$$

Note that all these Eqs. (1), (6), and (7) have some factors in front of the integral—these depend on the variable of the left hand side, and arise from free particle motion $(1-\cos p_1)$, $p^2/2m$, and (p^2+m^2) for the three cases, respectively. There is, however, an important difference. In (6) and (7), the nonseparable part of the kernel arises from the interaction potential. In (1) the two-particle t-matrix gives only separable terms—the nonseparability comes from the three-particle Green's function. However, the two-particle t-matrix gives the complicated branch cut [the square root term d of Eq. (5)] in Eq. (1).

II. WICK'S TRANSFORMATION

For Eq. (9) Wick introduced the transformation

$$x = f(s), \qquad y = f(t),$$

$$x \equiv f(s) = \int_0^s\frac{ds'}{(s'+m_a^2)(s'+m_b^2)}. \quad (10)$$

We express s and t in terms of x and y, and writing $s^{1/2}(s+m_a^2)(s+m_b^2)u(s) = v(x)$, we get the simpler equation from (9):

$$v(x) = \lambda\int_0^e K(x,y)v(y)\,dy, \quad (11)$$

with $e = f(\infty)$. To see the efficacy of the transformation, consider the s-wave solutions of the hydrogen atom; $\Phi(\mathbf{p})$ is then a function of the magnitude p only. Measuring momenta in units of a and carrying out the angular integration, we get

$$p\Phi(p) = \frac{2me^2}{\pi a\hbar(p^2+1)} \int_0^\infty dp' \ln\left|\frac{p+p'}{p-p'}\right| p'\Phi(p'). \quad (12)$$

Now, similar to (10), we take the variable

$$\frac{x}{2} = \int_0^p \frac{dp'}{p'^2+1}$$

or

$$p = \tan\tfrac{1}{2}x. \quad (13)$$

Putting $u(x) = p(p^2+1)\Phi(p)$, we obtain

$$u(x) = \frac{2me^2}{\pi a\hbar} \int_0^\pi dy\, \tfrac{1}{2} \ln\left|\frac{\sin\frac{1}{2}(x+y)}{\sin\frac{1}{2}(x-y)}\right| u(y). \quad (14)$$

Now[4]

$$\frac{1}{2}\ln\left|\frac{\sin\frac{1}{2}(x+y)}{\sin\frac{1}{2}(x-y)}\right| = \sum_{n=1}^\infty \frac{\sin nx \sin ny}{n}, \quad (15)$$

and the eigenvalues are $2n/\pi$, so that the hydrogen atom eigenvalues $E = -me^4/2\hbar^2n^2$ are reproduced. Let us note that here Wick's method is one way of generating symmetric kernels out of polar kernels. In (15), the completeness of the eigenfunctions is obvious. In the problem of the three magnon bound states we shall not find such complete resolution.

Rationalizing the denominator in front of the integral in (1), we write

$$\psi(p_1) = d(d+f)/\{2\cos\tfrac{1}{2}p_1[\alpha - \cos(K-p_1)][\alpha\sigma - 1 - \tfrac{1}{2}\sigma^2) - \cos p_1(\sigma\cos K + \tfrac{1}{2}\sigma^2) - \sigma\sin K \sin p_1]\}$$
$$\times \frac{1}{\pi}\int_{-\pi}^{\pi} dp_2 \left(\frac{f}{\alpha - \cos(K-p_1) - \cos(K-p_2) - \cos(K-p_1-p_2)} - 1\right)\cos(K - p_1 - \tfrac{1}{2}p_2)\psi(p_2). \quad (16)$$

Several transformations are suggested. For example, we take the term $\alpha - \cos(K-p_1)$, and put

$$x = -\pi + (\alpha^2-1)^{1/2} \times \int_{-\pi}^{p_1} \frac{dp_1'}{\alpha - \cos K \cos p_1' - \sin K \sin p_1'}. \quad (17)$$

The factors are chosen so that the inversion gives the relatively simple expression

$$\tan\tfrac{1}{2}p_1 = [\sin K + (\alpha^2-1)^{1/2}\tan\tfrac{1}{2}x]/(\alpha + \cos K). \quad (18)$$

Similarly,

$$\tan\tfrac{1}{2}p_2 = [\sin K + (\alpha^2-1)^{1/2}\tan\tfrac{1}{2}y]/(\alpha + \cos K). \quad (19)$$

A second transformation can be worked out with the other factor in the denominator in front of the integral:

$$x = -\pi + u\int_{-\pi}^{p_1} [(\alpha\sigma - 1 - \tfrac{1}{2}\sigma^2) - (\sigma t + \tfrac{1}{2}\sigma^2)\cos p_1' - \sigma\sin K \sin p_1']^{-1}\, dp_1'$$

or

$$\tan\tfrac{1}{2}p_1 = (\sigma\sin K + u\tan\tfrac{1}{2}x)/(\alpha\sigma - 1 + \sigma t), \quad (20)$$

with

$$t = \cos K \quad (21)$$

and

$$u = [\alpha^2\sigma^2 - \alpha(2\sigma + \sigma^3) + 1 - \sigma^3 t]^{1/2}. \quad (22)$$

One could have considered both factors together. The difference between (18) and (20) is only marginal; the anisotropy parameter also introduces nothing new in principle. So it will be enough to illustrate the details of the solution for (18) and the isotropic case $\sigma = 1$.

III. SOLUTION OF (1) WITH (18) AND (19)

Introduce in (18) and (19)

$$r = \tan\tfrac{1}{2}x, \qquad s = \tan\tfrac{1}{2}y. \quad (23)$$

Equation (1) becomes with (3), (21), and (23),

$$\psi(r) = -\frac{4(a+t)}{\pi(a^2-1)^{1/2}h(r)^{3/2}}\left(1 - \frac{f}{g}\right)^{-1}\int_{-\infty}^{\infty} ds \times \frac{ps - q}{(s^2+1)h(s)^{1/2}}\left(1 - \frac{fh(s)}{(a^2-1)h(r)(s-z)(s-\bar{z})}\right)\psi(s), \quad (24)$$

where

$$f = (a^2-1)(r^2+1) - 2(a+t), \quad (25)$$

$$g = [(a^2-1)^2r^4 + 2(a^2-1)(a^2-3)r^2 - 8\sin K(a^2-1)^{1/2}r + a^4 - ba^2 - 8at - 3]^{1/2}, \quad (26)$$

$$p = (a^2-1)^{3/2}[-r^2\sin K - 2(a^2-1)^{-1/2} \times (at+1)r + \sin K], \quad (27)$$

$$q = (a^2-1)(at+1)r^2 - 2(a^2-1)^{3/2}\sin K r - (a^3t - 3a^2 + 3at + \cos 2K), \quad (28)$$

$$h(r) = (a^2-1)r^2 + 2\sin K(a^2-1)^{1/2}r + a^2 + 2at + 1, \quad (29)$$

$$z = [p + i(a^2-1)^{1/2}(a+t)g]/(a^2-1)h(r), \qquad \bar{z} = \text{c.c. of } z. \quad (30)$$

Now

$$g^2 - f^2 = 4(a^2-1)(a-1+t)(r^2+mr+n), \quad (31)$$

with

$$m = -2\sin K(a^2-1)^{-1/2}(a-1+t)^{-1}, \quad (32)$$

$$n = \frac{a^3 + a^2(t-2) - a(4t+1) - (t^2+t+1)}{(a^2-1)(a-1+t)}. \quad (33)$$

The subsequent algebraic manipulations are analogous to those of Majumdar and Bose.[1] The Ansatz for the wavefunction ψ is

$$\psi(r) = \frac{c_0 + c_1r + c_2r^2 + c_3r^3 + c_4r^4 + c_5r^5}{h(r)^{3/2}(r^2+mr+n)}. \quad (34)$$

Substituting it into (24), we get a rational function on the left. From the pole $s = z$ on the right, we extract a term that is rational and free of the branch cut of g. Insisting on the equality of the rational parts of the two sides, we get the eigenvalue and a set of linear equations for the coefficients c_i.

All other terms of the integration must vanish; hence, we get another set of linear equations for c_i. Not all these equations are linearly independent, but we get just the right number to solve for the six quantities c_i. The results are

$$a = (8 + t)/3, \tag{35}$$

$$c_5 = A(t+5)^{1/2}(t+11)^{1/2}\sin K(4t+5)^2[(t+5)^{1/2}(128t^4 + 2164t^3 + 11\,100t^2 + 19\,639t + 10\,709)] - (t+11)^{1/2}(128t^4 + 1780t^3 + 8067t^2 + 13\,492t + 7151)],$$

$$c_4 = -A(t-1)(4t+5)[(t+5)^{1/2}(2560t^6 + 51\,840t^5 + 347\,124t^4 + 1\,016\,548t^3 + 1\,534\,959t^2 + 1\,175\,334t + 359\,359) - (t+11)^{1/2}(2560t^6 + 44\,160t^5 + 266\,076t^4 + 740\,083t^3 + 1\,082\,343t^2 + 809\,991t + 243\,433)],$$

$$c_3 = A(t+5)^{-1/2}(t+11)^{-1/2}2\sin K[(t+5)^{1/2}(-2048t^8 + 24\,064t^7 + 1\,332\,928t^6 + 14\,069\,404t^5 + 66\,886\,852t^4 + 168\,314\,677t^3 + 227\,622\,199t^2 + 155\,584\,555t + 42\,160\,321) - (t+11)^{1/2}(-2048t^8 + 30\,208t^7 + 1\,204\,384t^6 + 11\,301\,004t^5 + 50\,143\,141t^4 + 120\,486\,538t^3 + 157\,977\,328t^2 + 105\,734\,128t + 28\,233\,571)],$$

$$c_2 = -A(t+5)^{-1}(t+11)^{-1}2(t-1)[(t+5)^{1/2}(14\,336t^9 + 600\,384t^8 + 9\,723\,216t^7 + 79\,283\,580t^6 + 354\,774\,792t^5 + 897\,546\,549t^4 + 1\,290\,306\,438t^3 + 1\,033\,154\,364t^2 + 424\,177\,050t + 68\,413\,675) - (t+11)^{1/2}(14\,336t^9 + 557\,376t^8 + 8\,331\,648t^7 + 63\,187\,632t^6 + 267\,109\,809t^5 + 649\,176\,495t^4 + 909\,213\,114t^3 + 716\,461\,674t^2 + 291\,673\,347t + 46\,960\,765)],$$

$$c_1 = A(t+5)^{-3/2}(t+11)^{-3/2}\sin K[(t+5)^{1/2}(-6\,144t^{10} - 122\,688t^9 + 2\,784\,672t^8 + 94\,038\,804t^7 + 986\,518\,116t^6 + 5\,092\,599\,411t^5 + 14\,503\,363\,857t^4 + 24\,086\,488\,374t^3 + 23\,695\,111\,818t^2 + 12\,970\,403\,331t + 3\,055\,184\,865) - (t+11)^{1/2}(-6\,144t^{10} - 104\,256t^9 + 2\,984\,688t^8 + 83\,835\,180t^7 + 799\,925\,679t^6 + 3\,870\,894\,744t^5 + 10\,562\,562\,981t^4 + 17\,048\,275\,416t^3 + 16\,425\,705\,609t^2 + 8\,842\,960\,068t + 2\,054\,419\,155)],$$

$$c_0 = -A(t+5)^{-2}(t+11)^{-2}(t-1)[(t+5)^{1/2}(18\,432t^{11} + 1\,111\,680t^{10} + 25\,932\,240t^9 + 313\,740\,324t^8 + 2\,199\,146\,976t^7 + 9\,408\,284\,919t^6 + 25\,253\,056\,098t^5 + 43\,162\,846\,677t^4 + 46\,639\,268\,748t^3 + 30\,281\,197\,905t^2 + 10\,368\,780\,642t + 1\,319\,344\,191) - (t+11)^{1/2}(18\,432t^{11} + 1\,056\,384t^{10} + 23\,118\,480t^9 + 262\,396\,584t^8 + 1\,738\,051\,047t^7 + 7\,101\,809\,919t^6 + 18\,409\,978\,593t^5 + 30\,659\,930\,703t^4 + 32\,480\,524\,341t^3 + 20\,774\,293\,569t^2 + 7\,039\,664\,307t + 892\,063\,881)]. \tag{36}$$

A is an arbitrary normalization constant.

IV. DISCUSSION

The details of the transformation (20) are similar. With the anisotropy parameter σ the method works just as well, the eigenvalue condition being[5]

$$\alpha = (\sigma^3 t + 8)/\sigma(4 - \sigma^2). \tag{37}$$

Wick's transformation has thus provided alternative ways of solving the three magnon bound-state equation. However, the setting up of the solution involves an Ansatz, Eq. (34), and therefore, we have not found any method of proving rigorously that all the solutions of the equation can be found by the transformation. The existence of the two-body branch cut and its cancellation from the three-body kernel distinguishes this equation from the hydrogen atom or BS equation.

[1]C. K. Majumdar and I. Bose, J. Math. Phys. **19**, 2187 (1978).

[2]M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 330 (1966).

[3]G. C. Wick, Phys. Rev. **96**, 1124 (1954).

[4]R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience, New York, 1953), Vol. 1, p. 153.

[5]I. G. Gochev, Zh. Eksp. Teor. Fiz. **61**, 1674 (1971) [Sov. Phys. JETP **34**, 892 (1972)].

On the structure of Coulomb-type scattering amplitudes

F. Gesztesy
Institut für Theoretische Physik, Universität Graz, A-8010 Graz, Austria

(Received 3 September 1980; accepted for publication 13 November 1980)

On the basis of the Gell-Mann–Goldberger formula for Møller operators we introduce a natural splitting of the total scattering operator S into the pure Coulomb scattering operator S^{c} plus a remainder

$$S = S^{c} - 2\pi i T^{sc},$$

implying a decomposition of the total scattering amplitude f into the Coulomb scattering amplitude f^{c} and a remaining part f^{sc}

$$f(k,\omega,\omega') = f^{c}(k,\omega,\omega') + f^{sc}(k,\omega,\omega').$$

Concerning continuity properties, etc., of f^{sc}, close similarities between f^{sc} and ordinary short-range amplitudes are proved. In particular, we introduce transition operators $t^{sc}(z)$ and show how to obtain f^{sc} by an appropriate on-shell limit, thereby avoiding the notion of so-called Coulomb transition operators and the difficulties associated with them. Possible extensions of this approach to charged three-particle systems are also sketched.

PACS numbers: 03.65.Nk

I. INTRODUCTION

Since two-body off-shell data are an important input in the many-body problem for charged particles, much interest has been devoted to the off-shell behavior of two-body Coulomb scattering quantities like Jost functions and transition operators and to the corresponding on-shell limits. van Haeringen,[1] following a renormalization approach due to Zorbas,[2] introduced Coulombian asymptotic states $|\mathbf{k}_\epsilon \pm\rangle$,

$$|\mathbf{k}_\epsilon \pm\rangle = \Gamma(1 \mp i\gamma/2H_0^{1/2})^{-1}(2H_0/\epsilon)^{\pm i\gamma/2H_0^{1/2}}|\mathbf{k}\rangle \equiv A_{\epsilon\pm}(H_0)|\mathbf{k}\rangle, \tag{1.1}$$

where $H_0 = -\Delta$, and $|\mathbf{k}\rangle$ abbreviates the plane-wave state with momentum $\mathbf{k}\in R^3$ ($\hbar = 2m = 1$). With the help of the usual definitions

$$(z - H_c)^{-1} = (z - H_0)^{-1} + (z - H_0)^{-1}T^{c}(z)(z - H_0)^{-1},$$
$$(z - H)^{-1} = (z - H_0)^{-1} + (z - H_0)^{-1}T(z)(z - H_0)^{-1}, \tag{1.2}$$

and

$$T(z) = T^{c}(z) + T^{sc}(z), \tag{1.3}$$

where

$$H_c = H_0 + \gamma/|\mathbf{x}|, \quad H = H_c + gV(\mathbf{x}), \quad g\in\mathscr{R}$$

[V a suitable short-range potential (cf. Sec. II)], he was able to perform an on-shell limit of the form

$$\lim_{\epsilon\to 0_+} \langle\mathbf{k}|A^{*}_{\epsilon-}\, T^{c}(k^2 + i\epsilon)A_{\epsilon+}|\mathbf{k}'\rangle|_{|\mathbf{k}| = |\mathbf{k}'| \equiv k}$$
$$= \lim_{\epsilon\to 0_+} \langle\mathbf{k}_\epsilon - |T^{c}(k^2 + i\epsilon)|\mathbf{k}'_\epsilon +\rangle|_{|\mathbf{k}| = |\mathbf{k}'| = k}$$
$$= -(1/2\pi^2)f^{c}(k,\omega,\omega'), \quad (\mathbf{k} = k,\omega). \tag{1.4}$$

Here $f^{c}(k,\omega,\omega')$ denotes the Coulomb scattering amplitude (cf. Sec. III), and similar relations hold for $T(z)$ and $T^{sc}(z)$. If plane waves $|\mathbf{k}\rangle$ instead of $|\mathbf{k}_\epsilon \pm\rangle$ are used in (1.4), the on-shell limit does not exist in the ordinary sense and yields zero in the distributional sense.[1] However, besides technical complications [$T(z)$ are unbounded operators and thus domain questions arise], there are conceptional difficulties associated with $T^{c}(z)$ and $T(z)$ which sometimes culminated in statements like "Coulomb scattering violates unitarity," etc. (see the discussion in Ref. 3 and the references cited therein). In order to illuminate some of these problems it suffices to treat the pure Coulomb case in a somewhat heuristic way (cf. Sec. III for precise statements):

From Dollard's time-dependent definition[4] of the Coulomb scattering operator S^{c} and the formula

$$e^{-2iH_c t} = \frac{1}{2\pi i}\int_\Gamma dz e^{-2izt}(z - H_c)^{-1},$$

(Γ a suitable path in the complex plane) one immediately gets

$$S^{c} = \text{s} - \lim_{t\to\infty} U^{*}_{D}(t)e^{-2iH_c t}U^{*}_{D}(t)$$
$$= \text{s} - \lim_{t\to\infty}\{\exp[i\gamma \ln(4H_0 t)/H_0^{1/2}]$$
$$+ \frac{1}{2\pi i}\int_\Gamma dz e^{-2izt}U^{*}_{D}(t)(z - H_0)^{-1}$$
$$\times T^{c}(z)(z - H_0)^{-1}U^{*}_{D}(t)\}, \tag{1.5}$$

where $U_D(t)$ denotes the modified free evolution operator according to Dollard,[4]

$$U_D(t) = \exp[-iH_0 t - i\gamma \ln(4H_0 t)/2H_0^{1/2}]$$
$$\equiv e^{-iH_D(t)}, \quad t > 0.$$

If $\gamma = 0$, i.e., no Coulomb interaction exists, then (1.5) leads to $S = 1$ or if H instead of H_c is used, to the well-known formula $S = 1 - 2\pi i\, T^{s}$, where T^{s} denotes the short-range transition operator. But for $\gamma\neq 0$ the first term in (1.5), $\exp[i\gamma \ln(4H_0 t)/H_0^{1/2}]$, converges weakly to zero as $t\to\infty$ and thus has no strong limit. Since S^{c} obviously exists, the second term in (1.5) also has no strong limit as t tends to infinity. In fact a part of the second term in (1.5) must cancel the oscillating term $\exp[i\gamma \ln(4H_0 t)/H_0^{1/2}]$ in order to yield S^{c} in the limit $t\to\infty$.

From these remarks we thus conclude that the $T(z)$-

 0022-2488/82/010074-09$02.50

operator approach based on (1.2) and (1.3), although extremely useful in short-range ($\gamma = 0$) calculations, is no longer an appropriate tool if Coulomb interactions are present. This statement may be further confirmed by the following observation of Gibson and Chandler[5]: In the short-range case ($\gamma = 0$) the scattering operator may be expressed in terms of spectral integrals over the resolvent $(z - H)^{-1}$ and thus is directly relate to $T^s(z)$, whereas the S operator for Coulomb-type interactions involves spectral integration over a complex power of the resolvent[5] and not the resolvent itself—a fact which is easily proved by looking at the associated Møller operators (see Prop. 3.2) and is in accordance with relativistic investigations.[6] Another point which also shows the inadequacy of (1.2) for long-range interactions consists in the failure of (1.4) when restricted to partial waves. It was recently proved by van Haeringen[7] that if the partial wave projected Coulomb transition operator and Coulombian asymptotic states are inserted into (1.4), the limit $\epsilon \to 0_+$ gives $T_l^c(k) = (i/\pi k)e^{2i\delta_l^c(k)}$ plus a nonconverging oscillating term [$\delta_l^c(k) = \arg\Gamma(l + 1 + i\gamma/2k)$ are the usual Coulomb phase shifts].

Having reviewed some of the difficulties associated with Coulomb-type $T(z)$ operators based on (1.2), we now turn to the approach to be discussed in this paper. This approach relies on the Gell-Mann–Goldberger formula[8] or more precisely, on the chain rule for wave operators (see Sec. II):

$$\Omega_{\pm}(H,H_D) = \Omega_{\pm}(H,H_c)\Omega_{\pm}(H_c,H_D), \tag{1.6}$$

which suggests H_c (instead of H_0) as "unperturbed" Hamiltonian and thus results in the definition[9] (cf. Sec. III)

$$(z - H)^{-1} = (z - H_c)^{-1} + (z - H_c)^{-1}|V|^{1/2}t^{sc}(z)|V|^{1/2}(z - H_c)^{-1} \tag{1.7}$$

instead of (1.2) and (1.3). The main advantage of this definition lies in the fact that it implies a natural splitting of the total scattering operator S into the Coulomb scattering operator S^c plus a remainder denoted by $-2\pi i T^{sc}$,

$$S = S^c - 2\pi i T^{sc}, \tag{1.8}$$

and analogously for the total scattering amplitude

$$f(k,\omega,\omega') = f^c(k,\omega,\omega') + f^{sc}(k,\omega,\omega'). \tag{1.9}$$

In Sec. II we describe some relevant properties of the resolvents of H and H_c and discuss the spectral and scattering theory associated with H. In Sec. III we study $t^{sc}(z)$ and $f^{sc}(k,\omega,\omega')$ and show their close similarity to the corresponding short-range ($\gamma = 0$) quantities $t^s(z)$ and $f^s(k,\omega,\omega')$. In particular the on-shell $T^{sc}(k)$ operator [whose kernel is given by $f^{sc}(k,\omega,\omega')$] is trace-class and continuous in trace-norm, and $f^{sc}(k,\omega,\omega')$ is uniformly continuous in all variables if k varies in compact intervals under appropriate conditions on V (Theorem 3.2). Similarly the on-shell limit $\epsilon \to 0_+$ of $t^{sc}(k^2 + i\epsilon)$ (according to the Gell-Mann–Goldberger formula between Coulomb wavefunctions) immediately yields $f_{sc}(k,\omega,\omega')$ or the corresponding partial wave amplitude (Theorem 3.3).

Quite recently the basic idea underlying the above formalism (namely to separate out the pure Coulomb interaction) has been applied to the three-body problem of charged particles by Merkuriev[10] and to the N-body problem with repulsive Coulomb forces by Chandler and Gibson.[11] At the end of Sec. III we indicate how modified Faddeev equations for three-body transition operators $t_{ij}(z)$ may be obtained. These equations avoid the notion of Coulomb-transition operators for two-particle subsystems and only contain two-particle $t_i(z)$ operators of the type (1.7).

II. SPECTRAL AND SCATTERING PROPERTIES OF H

In the Hilbert space $\mathscr{L}^2(\mathscr{R}^3)$ we introduce the Coulomb Hamiltonian H_c:

$$H_c = H_0 + \gamma V_c, \quad D(H_c) = D(H_0), \quad V_c(\mathbf{x}) = 1/|\mathbf{x}|, \quad \gamma \in \mathscr{R}, \tag{2.1}$$

where H_0 denotes the usual self-adjoint realization of $-\Delta$ in $\mathscr{L}^2(\mathscr{R}^3)$. In addition to H_c we introduce the total Hamiltonian H as the form sum of H_c and gV,

$$H = H_c + gV, \quad g \in \mathscr{R}, \tag{2.2}$$

where the short-range potential V belongs to the Rollnik class[12] R, i.e.,

$$\int_{\mathscr{R}^6} d^3x\, d^3x' |V(\mathbf{x})V(\mathbf{x}')|/|\mathbf{x} - \mathbf{x}'|^2 < \infty.$$

Since various properties of $(z - H)^{-1}$ are basic to the whole subject treated in this paper, we summarize them in

Proposition 2.1: Let $V \in R$.

(a) $(H_c - z)^{-1}$ is a Carleman type operator with kernel

$$G_c(\mathbf{x},\mathbf{x}',\gamma,z) = \frac{\Gamma(1 + i\gamma/2\sqrt{z})}{4\pi|\mathbf{x} - \mathbf{x}'|}\left[\left(\frac{d}{d\alpha} - \frac{d}{d\beta}\right)\mathscr{M}_{-i\gamma/2\sqrt{z};1/2}(\alpha)\mathscr{W}_{-i\gamma/2\sqrt{z};1/2}(\beta)\right]\Big|_{\substack{\alpha = -i\sqrt{z}x_- \\ \beta = -i\sqrt{z}x_+}} \tag{2.3}$$

$$x_{\pm} = |\mathbf{x}| + |\mathbf{x}'| \pm |\mathbf{x} - \mathbf{x}'|, \quad 0 < \arg z < 2\pi, \quad \operatorname{Im}\sqrt{z} > 0,$$

$$z \neq -\gamma^2/4n^2, \quad n = 1,2,3,\cdots \text{ if } \gamma < 0,$$

[here $\mathscr{M}_{k;\mu}(\xi)$ and $\mathscr{W}_{k;\mu}(\xi)$ denote Whittaker functions[13]], and $|V|^{1/2}(z - H_c)^{-1}V^{1/2}$ is Hilbert–Schmidt for all $z \in \rho(H_c)$. [We recall $V(\mathbf{x})^{1/2} = V(\mathbf{x})|V(\mathbf{x})|^{-1/2}$.]

(b) Let $z \in \rho(H) \cap \rho(H_c)$ and in addition $V \in \mathscr{L}^1(\mathscr{R}^3)$ then

$$(z - H)^{-1} - (z - H_c)^{-1} = g(z - H_c)^{-1}V^{1/2}[1 - g|V|^{1/2}(z - H_c)^{-1}V^{/2}]^{-1} \times |V|^{1/2}(z - H_c)^{-1} \tag{2.4}$$

is trace-class.

(c) If $z \in \rho(H)$, then $|V|^{1/2}(z-H)^{-1}V^{1/2}$ is Hilbert–Schmidt; if furthermore, $V \in \mathscr{L}^1(\mathscr{R}^3)$, then $|V|^{1/2}(z-H)^{-1}$ and $(z-H)^{-1}V^{1/2}$ are Hilbert–Schmidt too.

(d) If $z = k^2 \pm i\epsilon$, $\epsilon, k > 0$, then in the limit $\epsilon \to 0_+$, $|V|^{1/2}(H_c - k^2 \pm i0)^{-1}V^{1/2}$ with kernel $G_c^{(\mp)}(\mathbf{x},\mathbf{x}',\gamma,k^2) = \lim_{\epsilon \to 0_+} G_c(\mathbf{x},\mathbf{x}',\gamma,k^2 \pm i\epsilon)$ are Hilbert–Schmidt. In particular there are constants $c_\gamma(k)$ such that

$$4\pi|\mathbf{x}-\mathbf{x}'||G_c^{(\mp)}(\mathbf{x},\mathbf{x}',\gamma,k^2)| \leqslant c_\gamma(k), \quad k>0, \tag{2.5}$$

where

$$\lim_{\gamma \to 0} c_\gamma(k) = \lim_{k \to \infty} c_\gamma(k) = 1. \tag{2.6}$$

Proof: Formula (2.3) which is due to Hostler[14] proves (a). For the proof of (b) we note that (2.4) certainly holds for z negative and large enough. By analytic continuation (2.4) holds for $z \in \rho(H) \cap \rho(H_c)$. If $V \in \mathscr{L}^1(\mathscr{R}^3)$, then

$$|V|^{1/2}(z-H_c)^{-1} = |V|^{1/2}(z-H_0)^{-1}[(z-H_0)(z-H_c)^{-1}]$$

[and similarly $(z-H_c)^{-1}V^{1/2}$] is Hilbert–Schmidt which proves (b). (c) is a simple consequence of (b), and (d) follows from (2.5), which in turn is implied by (2.3).

Next we give a short description of the nonpositive spectrum of H. Since for $\gamma<0$ there are obviously infinitely many negative eigenvalues of H, we concentrate our attention on the case $\gamma \geqslant 0$. By $N(gV,\gamma)$ we denote the number of bound states of H with bound state energy less than or equal to zero. As ususal we adopt the notation $V_\pm(\mathbf{x}) = [|V(\mathbf{x})| \pm V(\mathbf{x})]/2$ and exclude the trivial case where $V_-(\mathbf{x}) = 0$ a.e. Then we have

Proposition 2.2: Let $V \in R$. Then the number of nonpositive eigenvalues of H is bounded by

$$\begin{aligned} N(gV,\gamma) &< \frac{1}{4\pi^2}\int_{\mathscr{R}^6} d^3x\, d^3x' \frac{x_+}{x_-}\big[I_1(\sqrt{2\gamma x_-})K_1(\sqrt{2\gamma x_+})\big]^2 \frac{V_-(\mathbf{x})V_-(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^2} \\ &\leqslant \frac{1}{16\pi^2}\int_{\mathscr{R}^6} d^3x\, d^3x' \frac{V_-(\mathbf{x})V_-(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^2} < \infty, \quad \gamma \geqslant 0, \end{aligned} \tag{2.7}$$

[here $I_\beta(z)$ and $K_\beta(z)$ denote the modified Bessel functions[13] of order β].

Proof: Following the proof of Proposition 2 in Ref. 15 step by step (using the Hilbert–Schmidt norm instead of the trace-norm) and noting

$$\lim_{\lambda \to 0} G_c(\mathbf{x},\mathbf{x}',\gamma,\lambda) = \frac{1}{2\pi|\mathbf{x}-\mathbf{x}'|}\Big(\frac{x_+}{x_-}\Big)^{1/2} I_1(\sqrt{2\gamma x_-})K_1(\sqrt{2\gamma x_+}),$$

we arrive at (2.7). The finiteness of the right-hand side of (2.7) simply follows from monotonic descrease of $yK_1(y)$ $\{(d/dy)[yK_1(y)] = -yK_0(y) < 0$ for all $y>0\}$ and the bound[16] $I_1(y)K_1(y) \leqslant 1/2$.

In order to treat the positive part of the spectrum and the scattering theory associated with H we introduce Lippmann–Schwinger type equations of the form

$$\begin{aligned} \Phi^{(\mp)}(\mathbf{k},\mathbf{x}) = \Phi_c^{(\mp)}(\mathbf{k},\mathbf{x}) - g\int_{\mathscr{R}^3} d^3x' |V(\mathbf{x})|^{1/2} G_c^{(\mp)}(\mathbf{x},\mathbf{x}',\gamma,k^2) \\ \times V(\mathbf{x}')^{1/2}\Phi^{(\mp)}(\mathbf{k},\mathbf{x}'), \quad \mathbf{k},\mathbf{x} \in \mathscr{R}^3, \quad k = |\mathbf{k}| > 0 \end{aligned} \tag{2.8}$$

[we suppress the γ and g dependence of $\Phi^{(\mp)}(\mathbf{k},\mathbf{x})$], where

$$\begin{aligned} \Phi_c^{(-)}(\mathbf{k},\mathbf{x}) = |V(\mathbf{x})|^{1/2}(2\pi)^{-3/2}e^{-\pi\gamma/4k}\Gamma(1+i\gamma/2k) \\ \times e^{i\mathbf{k}\cdot\mathbf{x}}\, {}_1F_1(-i\gamma/2k;1;i(kr-\mathbf{k}\cdot\mathbf{x})), \end{aligned} \tag{2.9}$$

[${}_1F_1(a;b;z)$ denotes the regular confluent hypergeometric function[13]] and

$$\begin{aligned} \Phi^{(+)}(\mathbf{k},\mathbf{x}) &= \overline{\Phi^{(-)}(-\mathbf{k},\mathbf{x})}, \\ G^{(+)}(\mathbf{x},\mathbf{x}',\gamma,k^2) &= \overline{G^{(-)}(\mathbf{x},\mathbf{x}',\gamma,k^2)}. \end{aligned} \tag{2.10}$$

Let us denote by $\mathscr{E}$ the set of $k^2 \in [0,\infty)$ such that the homogeneous equation associated with (2.8) possesses a nontrivial solution, including the point $k^2=0$. Then we have

Proposition 2.3: Let $V \in \mathscr{L}^1(\mathscr{R}^3) \cap R$. Then $\mathscr{E}$ is a compact subset of $[0,\infty)$ of Lebesgue measure zero.

Proof: Using (2.6) and the Riemann–Lebesgue lemma, one proves

$$\lim_{k \to \infty} \||V|^{1/2}(H_c - k^2 \pm i0)^{-1}V^{1/2}\| = 0, \tag{2.11}$$

which proves the boundedness of $\mathscr{E}$. To prove that $\mathscr{E}$ is closed and of measure zero, one needs an improvement of the analytic Fredholm theorem (involving statements about the distribution of zeros of an holomorphic function on the boundary).[12,17,19] For $\gamma \geqslant 0$ the usual proof[12] carries through without any change. For $\gamma<0$ one has to take care of the bound states of H_c on the negative real axis.

Incidentally, relation (2.11) shows that under the conditions of Proposition 2.3 the Boren series (Taylor series in g) for $\Phi^{(\mp)}(\mathbf{k},\mathbf{x})$, obtained by iterating (2.8), converges for k sufficiently high (see also Ref. 15). Thus, if $V \in \mathscr{L}^1(\mathscr{R}^3) \cap R$ and $k^2 \notin \mathscr{E}$, Eq. (2.8) is uniquely solvable in $\mathscr{L}^2(\mathscr{R}^3)$.

With the help of Dollard's modified free evolution operator[4]

$$U_D(t) = e^{-iH_D(t)}, \quad H_D(t) = H_0 t + \epsilon(t)\gamma \ln(4H_0|t|)/2H_0^{1/2}, \tag{2.12}$$

we finally state

Proposition 2.4: Let $V \in \mathscr{L}^1(\mathscr{R}^3) \cap R$. Then

(a) $$\sigma_{\mathrm{ess}}(H) = \sigma_{\mathrm{ac}}(H) = [0,\infty), \tag{2.13}$$

$$\sigma_P(H) \cap [0,\infty) \subset \mathscr{E}, \quad \sigma_{\mathrm{sc}}(H) \subset \mathscr{E}. \tag{2.14}$$

(b) The Møller operators

$$\Omega_\pm(H,H_D) = \mathrm{s}-\lim_{t \to \pm\infty} e^{iHt}U_D(t)$$

exist and are complete,

$$\mathscr{R}(\Omega_\pm(H,H_D)) = \mathscr{H}_{\mathrm{ac}}(H).$$

In particular the scattering operator $S = \Omega_+(H,H_D)^*\Omega_-(H,H_D)$ is unitary.

(c) $(\Omega_\pm(H,H_c)f)(\mathbf{x})$

$$= s - \lim_{\substack{R\to\infty\\ \delta\to 0_+}} \int_{M_{R,\delta}} d^3k \Psi^{(\pm)}(\mathbf{k},\mathbf{x}) \hat{f}_c^{(\pm)}(\mathbf{k}), \tag{2.15}$$

$(\Omega_\pm(H,H_D)f)(\mathbf{x})$

$$= s - \lim_{\substack{R\to\infty\\ \delta\to 0_+}} \int_{M_{R,\delta}} d^3k \Psi^{(\pm)}(\mathbf{k},\mathbf{x}) \tilde{f}(\mathbf{k}), \quad f\in\mathscr{L}^2(\mathscr{R}^3) \tag{2.16}$$

where

$$\Phi^{(\pm)}(\mathbf{k},\mathbf{x}) = |V(\mathbf{x})|^{1/2}\Psi^{(\pm)}(\mathbf{k},\mathbf{x}), \tag{2.17}$$

$$M_{R,\delta} = \{\mathbf{k}\in\mathscr{R}^3 \,|\, |\mathbf{k}|\leqslant R,\ \operatorname{dist}(|\mathbf{k}|^2,\mathscr{E})\geqslant\delta\},$$

and

$$\tilde{f}(\mathbf{k}) = s - \lim_{R\to\infty} (2\pi)^{-3/2} \int_{|\mathbf{x}|\leqslant R} d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}),$$

$$\hat{f}_c^{(\pm)}(\mathbf{k}) = s - \lim_{R\to\infty} \int_{|\mathbf{x}|\leqslant R} d^3x\ \overline{\Psi_c^{(\pm)}(\mathbf{k},\mathbf{x})} f(\mathbf{x}).$$

Proof: Eq. (2.13) and (b) immediately follow from, Prop. 2.1(b) and the chain rule (see also Ref. 12, p.111). Equations (2.14) and (c) are proved in an analogous (but tedious) manner as in the short-range case $\gamma = 0$[12,18] (see also Refs. 19).

Remark 2.1: (a) In realistic situations one expects $\mathscr{E} = \{0\}$. Actually the absence of positive energy bound states, although physically plausible, usually involves additional regularity assumptions on V (cf. Ref. 20 for an extensive discussion). The absence of a singular continuous part in the spectrum of H was recently discussed by Enss[21] using geometric methods (for other results in this direction see Refs. 18 and 20). If in addition V is spherically symmetric then $\sigma_{sc}(H) = \sigma_P(H)\cap(0,\infty) = \varnothing$ by results of Weidmann.[22]

(b) If $e^{c|x|}V(\mathbf{x})\in R$ for some $c > 0$ then with the help of asymptotic expansions for the Whittaker functions one proves that $|V|^{1/2}(z - H_c)^{-1}V^{1/2}$ can be analytically continued in a neighborhood of the positive real axis and remains Hilbert–Schmidt there. An application of the analytic Fredholm theorem then shows that $\mathscr{E}$ is discrete [and thus $\sigma_{sc}(H) = \varnothing$].

III. THE STRUCTURE OF COULOMB-TYPE SCATTERING AMPLITUDES

We start with an appropriate definition of $t^{sc}(z)$ (avoiding unbounded operators),

$$\begin{aligned} t^{sc}(z) &= g(\operatorname{sgn}V)[1 - g|V|^{1/2}(z - H_c)^{-1}V^{1/2}]^{-1} \\ &= g(\operatorname{sgn}V)[1 + g|V|^{1/2}(z - H)^{-1}V^{1/2}], \\ &\qquad z\in\rho(H_c)\cap\rho(H). \end{aligned} \tag{3.1}$$

Then $t^{sc}(z)$ fulfills the equation

$$(z - H)^{-1} = (z - H_c)^{-1} + (z - H_c)^{-1}|V|^{1/2}t^{sc}(z)|V|^{1/2}(z - H_c)^{-1}. \tag{3.2}$$

Let $\Delta = [a,b]$, $b > a > 0$ be a compact interval on the real line and suppose $\Delta\cap\mathscr{E} = \varnothing$. Then the scattering operator S_Δ restricted to the energy interval Δ is defined by

$$S_\Delta = s - \lim_{t\to\infty} U_D^*(t)e^{-2iHt}E_\Delta U_D^*(t), \tag{3.3}$$

where E_Δ represents the spectral projection of H associated with Δ. Eq. (3.2) and the formula

$$e^{-2iHt}E_\Delta = s - \lim_{\epsilon\to 0_+} \frac{1}{2\pi i}\left(\int_{a-i\epsilon}^{b-i\epsilon} + \int_{b+i\epsilon}^{a+i\epsilon}\right) dz\ e^{-2izt}(z - H)^{-1}, \tag{3.4}$$

then imply

$$S_\Delta = S_\Delta^c - 2\pi i T_\Delta^{sc}, \tag{3.5}$$

where T_Δ^{sc} is defined by

$$T_\Delta^{sc} = s - \lim_{\substack{t\to\infty\\ \epsilon\to 0_+}} \frac{1}{4\pi^2}\left(\int_{a-i\epsilon}^{b-i\epsilon} + \int_{b+i\epsilon}^{a+i\epsilon}\right) dz\ e^{-2izt}U_D^*(t) \times (z - H_c)^{-1}|V|^{1/2}t^{sc}(z)|V|^{1/2}(z-H_c)^{-1}U_D^*(t). \tag{3.6}$$

In order to get an explicit expression for T_Δ^{sc}, we turn to the time-independent approach and state

Theorem 3.1: Assume $V\in\mathscr{L}^1(\mathscr{R}^3)\cap R$, let Δ be as above, and suppose $\tilde{\Phi}$, $\tilde{\Psi}\in C^\infty_0(\mathscr{R}^3)$, supp $\tilde{\Phi}\cap$ supp $\tilde{\Psi} = \varnothing$. Then

$$(\Psi,S_\Delta\Phi) = \frac{i}{4\pi}\int_\Delta k^2\,dk^2 \int d\omega \int d\omega' f(k,\omega,\omega')\ \overline{\tilde{\Psi}(k,\omega)}\tilde{\Phi}(k,\omega'), \tag{3.7}$$

where the scattering amplitude $f(k,\omega,\omega')$ may be written as

$$f(k,\omega,\omega') = f^c(k,\omega,\omega') + f^{sc}(k,\omega,\omega'), \quad k^2\notin\mathscr{E}. \tag{3.8}$$

$$f^c(k,\omega,\omega') = -\lim_{\epsilon\to 0_+} \frac{\Gamma(1 + i\gamma/2k)}{\Gamma(1 - i\gamma/2k)} 2^{1+i\gamma/2k}(4k^2)^{\epsilon-1}[1 - \cos\Theta]^{\epsilon-1-i\gamma/2k}, \quad \Theta = \sphericalangle(\omega,\omega'), \tag{3.9}$$

$$f^{sc}(k,\omega,\omega') = -2\pi^2 g\int_{\mathscr{R}^3} d^3x\ \overline{\Psi_c^{(+)}(k,\omega,\mathbf{x})}\ V(\mathbf{x})\Psi^{(-)}(k,\omega',\mathbf{x}), \quad k^2\notin\mathscr{E}. \tag{3.10}$$

Proof: We first note that the representation of S_Δ^c in terms of (3.9) has been derived by various authors.[5,23–25] The proof that $f^{sc}(k,\omega,\omega')$ is related to T_Δ^{sc} by (3.7) parallels the arguments given in Ref. 12, p. 143 and Ref. 18, p. 107 for the short-range case ($\gamma = 0$).

Remark 3.1: (a) That S^c (and thus S) is totally connected (i.e., S^c contains no "no scattering" part) and $S^c(\mathbf{k},\mathbf{k}')$ is more singular than $\delta(\mathbf{k} - \mathbf{k}')$, was proved by Herbst.[25]

(b) Since (3.9) defines a tempered distribution on $\mathscr{S}(\mathscr{R}^6)$ it is not hard to see that Φ, $\Psi\in\mathscr{S}(\mathscr{R}^3)$ (without disjoint supports in momentum space) suffices in (3.7). Instead of $\Delta = [a,b]$ and $\Delta\cap\mathscr{E} = 0$ one can use intervals of the type $[a,\infty)$, $a > 0$ if the integral over $f(k,\omega,\omega')$ in (3.7) is interpreted as

$$\lim_{\delta\to 0_+}\int_{N_{a,\delta}} k^2\,dk^2\int d\omega\int d\omega' f(k,\omega,\omega')\,\overline{\tilde\Psi(k,\omega)}\tilde\Phi(k,\omega'),$$

$$N_{a,\delta}=\{k^2\geqslant a\,|\,\mathrm{dist}(k^2,\mathscr{E})\geqslant\delta\},$$

and $\Phi,\Psi\in\mathscr{S}(\mathscr{R}^3)$.[12]

(c) Formal expansion of $\Psi_c^{(-)}(\mathbf{k}',\mathbf{x})$ and $G_c^{(-)}(\mathbf{x},\mathbf{x}',\gamma,k'^2)$ for $\mathbf{x}'$ fixed, $|\mathbf{k}'|=k>0$, $|\mathbf{x}|=r\to\infty$ in (2.8) finally yields after some calculations

$$\Psi^{(-)}(k,\omega',\mathbf{x})\sim(2\pi)^{-3/2}\Big\{e^{i[\mathbf{k}'\cdot\mathbf{x}+\gamma\ln(kr-\mathbf{k}'\cdot\mathbf{x})/2k]}+[f^c(k,\omega,\omega')+f^{sc}(k,\omega,\omega')]\frac{e^{i[kr-\gamma\ln(2kr)/2k]}}{r}\Big\}\quad \omega=\mathbf{x}/r,\quad k>0,\quad \omega\neq\omega'. \tag{3.11}$$

According to Remark 3.1(a) we define the T and T^c operator by

$$S=-2\pi iT,\quad S^c=-2\pi iT^c, \tag{3.12}$$

in order to get

$$T=T^c+T^{sc}. \tag{3.13}$$

In analogy to the short-range case ($\gamma=0$) we introduce in the Hilbert space $\mathscr{L}^2(S^{(2)})$ ($S^{(2)}$ the unit sphere in $\mathscr{R}^3$) the on-shell T^{sc} operator $T^{sc}(k)$ by

$$(T^{sc}(k)\Phi)(\omega)=-\frac{1}{2\pi^2}\int d\omega' f^{sc}(k,\omega,\omega')\Phi(\omega'),\quad \Phi\in\mathscr{L}^2(S^{(2)}). \tag{3.14}$$

In order to circumvent the forward singularity of $f^c(k,\omega,\omega')$ we define

$$(P_{\Omega_2}T^c(k)P_{\Omega_1}\Phi)(\omega)=-\frac{1}{2\pi^2}\int_{\Omega_1}d\omega'\chi_{\Omega_2}(\omega)f^c(k,\omega,\omega')\Phi(\omega')$$

$$\Phi\in\mathscr{L}^2(S^{(2)}),\quad \chi_{\Omega_2}(\omega)=\begin{cases}1, & \omega\in\Omega_2\\ 0, & \omega\notin\Omega_2\end{cases}, \tag{3.15}$$

where Ω_1, and Ω_2 are disjoint subsets of $S^{(2)}$ separated by a positive distance, and P_{Ω_i}, $i=1,2$ denote the projections onto Ω_i. The total on-shell T operator $T(k)$ is defined analogously.

With these definitions we are ready to state

Theorem 3.2: Let $V\in\mathscr{L}^1(\mathscr{R}^3)\cap R$. Then

(a) If $k^2\notin\mathscr{E}$, $T^{sc}(k)$ as an operator in $\mathscr{L}^2(S^{(2)})$, is trace-class and continuous in trace-norm. In particular $f^{sc}(k,\omega,\omega')$ is uniformly continuous (with respect to all variables) whenever k^2 varies in compact interval not intersecting $\mathscr{E}$.

(b) Let C_i, $i=1,2$, be closed cones with vertices at the origin. Suppose $C_1\cap C_2=\{0\}$ and define $\Omega_i=C_i\cap S^{(2)}$, $i=1,2$. Then the scattering cross section for scattering into C_2 from an initial state having momentum support in C_1,

$$\sigma(k,C_1\to C_2)=(1/4\pi)\int_{\Omega_2}d\omega\int_{\Omega_1}d\omega'|f(k,\omega,\omega')|^2, \tag{3.16}$$

is finite and continuous in k whenever $k^2\notin\mathscr{E}$.

Proof: (a) We introduce operators $A_V(k):\mathscr{L}^2(S^{(2)})\to\mathscr{L}^2(\mathscr{R}^3)$ and $B_V(k):\mathscr{L}^2(\mathscr{R}^3)\to\mathscr{L}^2(S^{(2)})$ by

$$(A_V(k)\Phi)(\mathbf{x})=\int_{S^{(2)}}d\omega\,V^{1/2}(\mathbf{x})\Psi_c^{(-)}(k,\omega,\mathbf{x})\Phi(\omega),\quad \Phi\in\mathscr{L}^2(S^{(2)}),$$

$$(B_V(k)\Psi)(\omega)=\int_{\mathscr{R}^3}d^3x\,V^{1/2}(\mathbf{x})\Psi_c^{(+)}(k,\omega,\mathbf{x})\Psi(\mathbf{x}),\quad \Psi\in\mathscr{L}^2(\mathscr{R}^3), \tag{3.17}$$

and note

$$T^{sc}(k)=gB_V(k)\times[1-g|V|^{1/2}(H_c-k^2+i0)^{-1}V^{1/2}]^{-1}A_{|V|}(k),\quad k^2\notin\mathscr{E}. \tag{3.18}$$

From $|V|^{1/2}\in\mathscr{L}^2(\mathscr{R}^3)$, $A_{|V|}(k)$ and $B_V(k)$ are Hilbert–Schmidt which proves that $T^{sc}(k)$ is trace-class in $\mathscr{L}^2(S^{(2)})$. Since $\||V|^{1/2}(H_c-k^2+i0)^{-1}V^{1/2}\|$ is continuous in k and $A_{|V|}(k)$ and $B_V(k)$ are continuous in k in Hilbert–Schmidt norm, $T^{sc}(k)$ is continuous in trace-norm for $k^2\notin\mathscr{E}$. The continuity of

$$f^{sc}(k,\omega,\omega')=-2\pi^2g(\Phi_c^{(+)}(k,\omega,\cdot),\mathrm{sgn}\,V\Phi^{(-)}(k,\omega',\cdot)),\quad k^2\notin\mathscr{E},$$

follows from the fact that $\Phi_c^{(\pm)}(\mathbf{k},\mathbf{x})$ and $\Phi^{(\pm)}(\mathbf{k},\mathbf{x})$ are strongly continuous in k for $k^2\notin\mathscr{E}$.

(b) The finiteness of $\sigma(k,C_1\to C_2)$ follows from $C_1\cap C_2=\{0\}$; its continuity in k is clear from (a).

Remark 3.2: (a) The finiteness of scattering cross sections between nonintersecting cones for short-range potentials ($|V(\mathbf{x})|\underset{|\mathbf{x}|\to\infty}{\sim}|\mathbf{x}|^{-\alpha}$, $\alpha>1$), has been discussed in detail by Amrein and Pearson.[26] For previous results including long-range potentials see Agmon.[27]

(b) In order to work entirely in Hilbert space we assumed $V\in\mathscr{L}^1(\mathscr{R}^3)\cap R$. Following the methods employed in Davies,[28] one can extend the above continuity results for potentials V obeying $\int_{\mathscr{R}^3}d^3x'\,|V(\mathbf{x}')|/|\mathbf{x}-\mathbf{x}'|^{1\pm\epsilon}\leqslant C(1+|\mathbf{x}|)^{-\epsilon}$ for some $\epsilon>0$, $\mathbf{x}\in\mathscr{R}^3$, by working in suitable Banach spaces.

If in addition $V(\mathbf{x})$ is spherically symmetric, $\delta_l(k)$, the total phase shift associated with the angular momentum subspace indexed by l, can be split up into

$$\delta_l(k)=\delta_l^c(k)+\delta_l^{sc}(k),\quad l=0,1,2,\cdots,\ k>0, \tag{3.19}$$

where

$$\delta_l^c(k)=\arg\Gamma(l+1+i\gamma/2k)$$

denote the Coulomb phase shifts. In this case $T^{sc}(k)$ may be decomposed into

$$T^{sc}(k)=\bigoplus_{l=0}^{\infty}\bigoplus_{m=-l}^{l}T_l^{sc}(k),\quad T_l^{sc}(k)=(i/\pi k)e^{2i\delta_l^c(k)}(e^{2i\delta_l^{sc}(k)}-1), \tag{3.20}$$

and

$$\|T^{sc}(k)\|_p = \frac{2}{\pi k}\Big[\sum_{l=0}^{\infty}(2l+1)|\sin\delta_l^{sc}(k)|^p\Big]^{1/p},$$
$$p\geqslant 1. \tag{3.21}$$

In particular

$$|f^{sc}(k,\omega,\omega')| \leqslant \frac{1}{\pi k}\sum_{l=0}^{\infty}(2l+1)|e^{2i\delta_l^{sc}(k)}-1|\,\frac{1}{4\pi}$$
$$= \frac{1}{4\pi}\|T^{sc}(k)\|_1. \tag{3.22}$$

For a detailed description of the high-energy behavior of $\delta_l^{sc}(k)$ see Ref. 29.

Since there is no obvious physical reason that $\mathscr{E}\neq\{0\}$ (and thus $[1+g|V|^{1/2}(H_c-k^2\pm i0)^{-1}V^{1/2}]^{-1}$ exists for $k>0$), it is hoped that $g|V|^{1/2}(H_c-k^2\pm i0)^{-1}V^{1/2}$ has no eigenvalue -1 if k^2 varies in $(0,\infty)$. But the operator[30]

$$R(k^2) = g|V|^{1/2}[\mathrm{Re}(H_c-k^2\pm i0)^{-1}]V^{1/2}$$

clearly has an eigenvalue -1 for certain values of k^2. In fact these k^2 correspond to maxima in $\|T^{sc}(k)\|_2$, which are associated with resonance phenomena. For example, if $V(\mathbf{x})$ is spherically symmetric, then an eigenvalue -1 of $R(k^2)$ occurs at those values of k^2 where $e^{2i\delta_l^{sc}(k)}=-1$ (and thus $|T_l^{sc}(k)|$ has a maximum) for some l. In the short-range case ($\gamma=0$) these values of k^2 are responsible for peaks in the partial wave cross section, i.e., for resonances (large time delay) and have been investigated by Rollnik[31] (see also Ref. 32). More precisely, we have

Proposition 3.1: Let $V\epsilon\mathscr{L}^1(\mathscr{R}^3)\cap R$ and suppose $\int_0^R dr\, r|V(r)|<\infty$ for some $R>0$. Then

$$\{1+g|V|^{1/2}[\mathrm{Re}(H_c-k^2\pm i0)^{-1}]V^{1/2}\}^{-1}$$

does not exist {or equivalently the operator $R(k^2)=g|V|^{1/2}[\mathrm{Re}(H_c-k^2\pm i0)^{-1}]V^{1/2}$ has eigenvalue -1} precisely at those values of $k^2\epsilon(0,\infty)$ where

$$e^{2i\delta_l^{sc}(k)}=-1 \quad\text{for some } l.$$

Proof: After separation of variables the eigenvalue equation

$$(R(k^2))\Phi(\mathbf{x}) = -\Phi(\mathbf{k},\mathbf{x}),\quad k>0,\quad \Phi\epsilon\mathscr{L}^2(\mathscr{R}^3) \tag{3.23}$$

reduces to

$$\phi_l(k,r) = -g\int_0^{\infty}dr'|V(r)|^{1/2}\mathrm{Re}\ \hat{g}_l^c(k,\gamma,r,r')V^{1/2}(r')\phi_l(k,r'),\quad k>0,\quad \phi_l\in\mathscr{L}^2(0,\infty), \tag{3.24}$$

where

$$\Phi(\mathbf{k},\mathbf{x}) = \frac{1}{kr}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\phi_l(k,r)\,\overline{Y_{l,m}(\omega_{\mathbf{k}})}\,Y_{l,m}(\omega_{\mathbf{x}}), \tag{3.25}$$

$$G_c^{(\mp)}(\mathbf{x},\mathbf{x}',\gamma,k^2) = \frac{1}{rr'}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\hat{g}_l^c(\mp k,\gamma,r,r')\,\overline{Y_{l,m}(\omega_{\mathbf{x}})}\,Y_{l,m}(\omega_{\mathbf{x}'}),\qquad |\mathbf{x}|=r,\quad |\mathbf{x}'|=r'. \tag{3.26}$$

[See Ref. 15 for an explicit representation of $\hat{g}_l^c(\pm k,\gamma,r,r')$ in terms of confluent hypergeometric functions.] On the other hand, if one expands

$$\Phi^{(-)}(\mathbf{k},\mathbf{x}) = \frac{1}{kr}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}i^l e^{i\delta_l^c(k)}\frac{1}{\overline{\mathscr{F}_l(k,\gamma,g)}}\phi_l^{(-)}(k,r)\,\overline{Y_{l,m}(\omega_{\mathbf{k}})}\,Y_{l,m}(\omega_{\mathbf{x}}), \tag{3.27}$$

$$\Phi_c^{(-)}(\mathbf{k},\mathbf{x}) = \frac{1}{kr}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}i^l e^{i\delta_l^c(k)}\phi_l^{(-)c}(k,r)\,\overline{Y_{l,m}(\omega_{\mathbf{k}})}\,Y_{l,m}(\omega_{\mathbf{x}}), \tag{3.28}$$

where $\mathscr{F}_l(k)$ denotes the Jost function as introduced in Ref. 15, one obtains from (2.8)

$$\phi_l^{(-)}(k,r) = \mathscr{F}_l(k)\phi_l^{(-)c}(k,r) - g\int_0^{\infty}dr'\,|V(r)|^{1/2}\hat{g}_l^c(-k,\gamma,r,r')V(r')^{1/2}\phi_l^{(-)}(k,r'). \tag{3.29}$$

Under our hypothesis on $V(r)$, (3.29) is solved uniquely by iteration for all $k>0$.[15] Since $\phi_l^{(-)}$, and $\phi_l^{(-)c}$ are real, we finally get

$$\phi_l^{(-)}(k,r) = [\mathrm{Re}\mathscr{F}_l(k)]\phi_l^{(-)c}(k,r) - g\int_0^{\infty}dr'\,|V(r)|^{1/2}\mathrm{Re}\hat{g}_l^c(k,\gamma,r,r')V(r')^{1/2}\phi_l^{(-)}(k,r'). \tag{3.30}$$

Comparison of (3.30) and (3.24) shows that $g|V|^{1/2}[\mathrm{Re}(H_c-k_0^2\pm i0)^{-1}]V^{1/2}$ has eigenvalue -1 if and only if $\mathrm{Re}\mathscr{F}_l(k_0)=0$ for some l. But since[15]

$$e^{2i\delta_l^{sc}(k)} = \overline{\mathscr{F}_l(k,\gamma,g)}/\mathscr{F}_l(k,\gamma,g),\quad k>0,\quad l=0,1,2,\cdots, \tag{3.31}$$

this is equivalent to

$$e^{2i\delta_l^{sc}(k_0)} = -1 \text{ for some } l\ .$$

Remark 3.3: (a) At least in the spherically symmetric case discussed in Proposition 3.1 we infer from (3.29) and the fact that $\mathscr{F}_l(k)\neq 0$ for $k>0$ that $\mathscr{E}=\{0\}$. The condition $rV(r)\in\mathscr{L}^1([0,R])$ for some $R>0$ may be relaxed to $V\in\mathscr{W}$, the so called $\mathscr{W}$ class[33] [i.e., $V\in\mathscr{L}^1([R,\infty))$ for all $R>0$ and $W(r)=-\int_r^{\infty}dr'\ V(r')$ fulfills $W\in\mathscr{L}^1(0,\infty)$] allowing for potentials that are singular and oscillating near the orgin.[29]

(b) Under the conditions of Proposition 3.1 the operator

$$R(E) = g|V|^{1/2}[\mathrm{Re}(H_c-E\pm i0)^{-1}]V^{1/2} \tag{3.32}$$

has remarkable properties: If for some $E_0<0$ $R(E_0)$ has an eigenvalue -1, then E_0 is an eigenvalue of H. On the other hand, if $R(E_1)$ has an eigenvalue -1 for some $E_1>0$, then E_1 corresponds to a maximum of $T_l^{sc}(k)$ at $k^2=E_1$ for some l. In analogy to the short-range case ($\gamma=0$) these maxima give rise to peaks in $\|T^{sc}(k)\|_2^2$ (which for $\gamma=0$ is proportional to the total scattering cross section averaged over all initial directions[17] $\bar{\sigma}(k)=\pi^3\|T^s(k)\|_2^2$).

(c) Since $\lim_{|E|\to\infty}\|R(E)\|=0$ [cf. Eq. (2.11)] there exists no eigenvalue -1 of $R(E)$ if $|E|$ is sufficiently large. For $E<0$ this simply means that H is bounded from below, whereas for $E>0$ this is connected with the fact that Born expansions for $\Phi(\mathbf{k},\mathbf{x})$ and $f^{sc}(k,\omega,\omega')$ [by iterating Eq. (2.8)] converge for k^2 sufficiently high (cf. Ref. 15 for estimates on the radius of convergence of various Born expansions in the spherically symmetric case). Next we turn to the on-shell limit of $t^{sc}(z)$:

Theorem 3.3: (a) Let $V\in\mathscr{L}^1(\mathscr{R}^3)\cap R$; then

$$\lim_{\epsilon\to 0_+}[\Phi_c^{(+)}(k,\omega,\cdot),t^{sc}(k^2+i\epsilon)\Phi_c^{(-)}(k,\omega',\cdot)] = -(1/2\pi^2)f^{sc}(k,\omega,\omega'),\quad k^2\notin\mathscr{E}. \tag{3.33}$$

(b) If in addition $V(\mathbf{x})$ is spherically symmetric and $rV(r)\in\mathscr{L}^1([0,R])$ for some $R>0$ [cf. Remark 3.3a)], then

$$\lim_{\epsilon\to 0_+}(\phi_l^{(-)c}(k,.),t_l^{sc}(k^2+i\epsilon)\phi_l^{(-)c}(k,.))=T_l^{sc}(k) = (i/\pi k)e^{2i\delta_l^c(k)}(e^{2i\delta_l^{sc}(k)}-1),\quad k>0, \tag{3.34}$$

where

$$t^{sc}(z)=\bigoplus_{l=0}^{\infty}\bigoplus_{m=-l}^{l}U^{-1}t_l^{sc}(z)U\otimes 1 \quad \text{in } \mathscr{L}^2(\mathscr{R}^3)=\mathscr{L}^2((0,\infty);r^2dr)\otimes\mathscr{L}^2(S^{(2)}), \tag{3.35}$$

and

$$U:\begin{cases}\mathscr{L}^2((0,\infty);r^2dr)\longleftrightarrow\mathscr{L}^2((0,\infty);dr)\\ g(r)\longleftrightarrow h(r)=rg(r)\end{cases}$$

Proof: (a) Formula (3.33) immediately follows from the strong continuity of $t^{sc}(k^2+i\epsilon)$ as $\epsilon\to 0_+$ and from Eq. (2.8):

$$(t^{sc}(k^2+i0)\Phi_c^{(-)})(k,\omega',\mathbf{x})=g((\operatorname{sgn}V)\times[1+g|V|^{1/2}(H_c-k^2-i0)^{-1}V^{1/2}]^{-1}\Phi_c^{(-)})(k,\omega',\mathbf{x}) = g(\operatorname{sgn}V(\mathbf{x}))\Phi^{(-)}(k,\omega',\mathbf{x}),\quad k^2\notin\mathscr{E}, \tag{3.36}$$

and thus

$$\lim_{\epsilon\to 0_+}(\Phi_c^{(+)}(k,\omega,\cdot),t^{sc}(k^2+i\epsilon)\Phi_c^{(-)}(k,\omega',\cdot)) = g\int_{\mathscr{R}^3}d^3x\ \overline{\Phi_c^{(+)}(k,\omega,\mathbf{x})}\operatorname{sgn}V(\mathbf{x})\Phi^{(-)}(k,\omega',\mathbf{x}) = -(1/2\pi^2)f^{sc}(k,\omega,\omega'),\quad k^2\notin\mathscr{E}.$$

In order to prove (b), we note

$$t_l^{sc}(z)=g(\operatorname{sgn}V)[1+g|V|^{1/2}(h_l^c-z)^{-1}V^{1/2}]^{-1}, \tag{3.37}$$

where h_l^c denotes the Friedrichs extension of $\dot{h}_l^c$,[15,29]

$$\dot{h}_l^c=-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}+\frac{\gamma}{r},\quad D(\dot{h}_l^c)=C_0^\infty(0,\infty),$$

and

$$\{[1+g|V|^{1/2}(h_l^c-k^2-i0)^{-1}V^{1/2}]^{-1}\mathscr{F}_l(k)\phi_l^{(-)c}\}(r) = \phi_l^{(-)}(k,r) \tag{3.38}$$

by (3.29). Finally, with the help of

$$\frac{g\pi}{ik\mathscr{F}_l(k)}\int_0^\infty dr\,\phi_l^{(-)c}(k,r)\operatorname{sgn}V(r)\phi_l^{(-)}(k,r)=(e^{2i\delta_l^{sc}(k)}-1), \tag{3.39}$$

[cf. Eq. (3.44) in Ref. 15] (3.34) follows.

Remark 3.4: We emphasize once again the naturalness of $t^{sc}(z)$. The on-shell limit may be performed without any complications (also in the partial wave subspaces). In Eq. (3.10) we only introduced the on-shell scattering amplitude $f^{sc}(k,\omega,\omega')$. The definition of a corresponding half-shell amplitude and corresponding half-shell limits are clearly obvious.

As mentioned in Sec. I we finally discuss the different connection between short- or long-range wave operators and corresponding resolvents:

Proposition 3.2: Let $V\in\mathscr{L}^1(\mathscr{R}^3)\cap R$. Then

$$\begin{aligned}\Omega_-(H,H_0)&=\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_{\mathscr{R}}dE(\lambda)(H_0-\lambda+i\epsilon)^{-1}\\ &=-\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_0^\infty(H-\lambda-i\epsilon)^{-1}dE_0(\lambda)\quad\text{if }\gamma=0,\end{aligned} \tag{3.40}$$

$$\begin{aligned}\Omega_-(H,H_c)&=\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_{\mathscr{R}}dE(\lambda)(H_c-\lambda+i\epsilon)^{-1}E^{ac}(H_c)\\ &=-\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_0^\infty(H-\lambda-i\epsilon)^{-1}dE_c(\lambda),\end{aligned} \tag{3.41}$$

$$\begin{aligned}\Omega_-(H_c,H_D)&=\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_{\mathscr{R}}dE_c(\lambda)(4H_0)^{i\gamma/2H_0^{1/2}}e^{-\pi\gamma/4H_0^{1/2}}\\ &\quad\times\Gamma(1+i\gamma/2H_0^{1/2})(H_0-\lambda+i\epsilon)^{-1-i\gamma/2H_0^{1/2}}\\ &=-\text{s}-\lim_{\epsilon\to 0_+}i\epsilon\int_0^\infty(4\lambda)^{i\gamma/2\lambda^{1/2}}e^{\pi\gamma/4\lambda^{1/2}}\\ &\quad\times\Gamma(1+i\gamma/2\lambda^{1/2})(H_c-\lambda-i\epsilon)^{-1-i\gamma/2\lambda^{1/2}}dE_0(\lambda),\end{aligned} \tag{3.42}$$

where $E(\lambda)$, $E_c(\lambda)$, $E_0(\lambda)$ abbreviate the spectral projections of H, H_c, and H_0, and $E^{ac}(H_c)$ denotes the projector onto the absolutely continuous spectral subspace of H_c.

Proof: It suffices to prove the first relation in (3.42): The Bochner integral

$$\Omega_{\epsilon,-}(H_c,H_D) = \epsilon\int_{-\infty}^{0} \exp(\epsilon t)\exp(iH_c t)\exp\{-i[H_0 t - \gamma\ln(-4H_0 t)/2H_0^{1/2}]\}\,dt \tag{3.43}$$

clearly exists, and

$$\text{s} - \lim_{\epsilon\to 0_+} \Omega_{\epsilon,-}(H_c,H_D) = \Omega_-(H_c,H_D). \tag{3.44}$$

Thus

$$\Omega_{\epsilon,-}(H_c,H_D) = \epsilon\int_{-\infty}^{0} \exp(\epsilon t)\int dE_c(\lambda)\exp(i\lambda t)\exp\{-i[H_0 t - \gamma\ln(-4H_0 t)/2H_0^{1/2}]\}\,dt \tag{3.45}$$

and the boundedness properties of the integrand immediately allow the interchange of the t and λ integration.[5] The first part of (3.42) then follows by computation of a simple Γ function[13] integral in (3.45). With these results in mind we conclude the two-body case by the following

Remark 3.5: Whenever two Hamiltonians H_1, H_2 differ by a short-range interaction V [as in (3.40) and (3.41)], the corresponding MØller operators $\Omega_\pm(H_2,H_1)$ are directly related to the resolvents of these Hamiltonians in the usual way. Since in this case the resovents may be related to a $t^{1,2}(z)$ operator by an equation of the type (3.2) there is a direct connection between $\Omega_\pm(H_2,H_1)$ and $t^{1,2}(z)$. As we proved in Theorems 3.1–3.3 and Propositions 3.1 and 3.2 the $t^{sc}(z)$ operator indeed has almost all properties of an ordinary short-range $t^s(z)$ operator. In addition to those properties we also note that the partial-wave expansion of $f^{sc}(k,\omega,\omega')$ converges in the ordinary sense and no distribution techniques[34] or generalized summation procedures like Abel[35] or Pade summation[36] [which have to be introduced for $f^c(k,\omega,\omega')$] must be applied. On the other hand Eq. (3.42) indicates a phenomenon discussed by Gibson and Chandler.[5] If two Hamiltonians H_1, H_2 differ by a Coulomb potential then the associated modified wave operators are connected with a complex power of the resolvents of H_1 and H_2 and not with the resolvents itself. This shows that the conventional approach based on (1.2), which contains the resolvents only linearly, is not sophisticated enough and thus leads to unpleasant properties (e.g., in partial wave subspaces) as described in Sec. I.

We finally conclude with some remarks concerning the generalization of this approach to more than two particles. We note that with quite similar ideas in mind, Chandler and Gibson[11] recently investigated the N-body problem including repulsive Coulomb potentials, and Merkuriev[10] discussed the three-body problem of charged particles. Here we consider the three-body problem and sketch how modified Faddeev equations may be derived: We introduce in the three-particle center of mass system the usual pair of coordinates $\{\mathbf{x}_i,\mathbf{y}_i\}$

$$\mathbf{x}_1 = \left(\frac{2m_2m_3}{m_2+m_3}\right)^{1/2}(\mathbf{x}^{(1)} - \mathbf{x}^{(2)}),$$
$$\mathbf{y}_1 = \left(\frac{2m_1(m_2+m_3)}{m_1+m_2+m_3}\right)^{1/2}\left(\frac{m_2\mathbf{x}^{(2)}+m_3\mathbf{x}^{(3)}}{m_2+m_3} - \mathbf{x}^{(1)}\right), \tag{3.46}$$

where $\mathbf{x}^{(i)}$ and m_i are the positions and masses of the particles. The other pairs $\{\mathbf{x}_i,\mathbf{y}_i\}$, $i=2,3$, are obtained by cyclic permutation. Let $V_i(\mathbf{x}_i)\in\mathscr{L}^1(\mathscr{R}^3)\cap R$ and define

$$\tilde{H}_0 = -\Delta_{\mathbf{x}_i} - \Delta_{\mathbf{y}_i},\quad \tilde{H}_c = \tilde{H}_0 + \sum_{i=1}^{3}\frac{\gamma_i}{|\mathbf{x}_i|},\quad \gamma_i\in\mathscr{R},$$
$$\tilde{H}_i = \tilde{H}_c + g_iV_i,\quad g_i\in\mathscr{R},\quad \tilde{H} = \tilde{H}_c + \tilde{V}$$

(in the sense of forms),

$$\tilde{V} = \sum_{i=1}^{3} g_iV_i. \tag{3.47}$$

Next we define

$$t_i(z) = g_i(\operatorname{sgn}V_i)[1 + g_i|V_i|^{1/2}(z-\tilde{H}_i)^{-1}V_i^{1/2}],$$
$$t_{ij}(z) = g_i(\operatorname{sgn}V_i)[\delta_{ij}\ 1 + g_j|V_i|^{1/2}(z-\tilde{H})^{-1}V_j^{1/2}],\quad i,j=1,2,3, \tag{3.48}$$

and note that

$$(z-\tilde{H}_i)^{-1} = (z-\tilde{H}_c)^{-1} + (z-\tilde{H}_c)^{-1}|V_i|^{1/2}t_i(z)|V_i|^{1/2}(z-\tilde{H}_c)^{-1},\quad i=1,2,3,$$
$$(z-\tilde{H})^{-1} = (z-\tilde{H}_c)^{-1} + (z-\tilde{H}_c)^{-1}\sum_{i,j=1}^{3}|V_i|^{1/2}t_{ij}(z)|V_j|^{1/2}(z-\tilde{H}_c)^{-1}. \tag{3.49}$$

The modified Faddeev equation finally reads

$$t_{ij}(z) = \delta_{ij}t_i(z) + t_i(z)\sum_{k\neq i}|V_i|^{1/2}(z-\tilde{H}_c)^{-1}|V_k|^{1/2}t_{kj}(z). \tag{3.50}$$

We emphasize that (3.50) contains no objects like two-particle subsystem Coulomb transition operators. The only two-particle input consists in $t_i(z)$ operators of the type (3.1). In contrast to the two-body situation, the kernel of $(z - \tilde{H}_c)^{-1}$ is not known in closed form. For this reason one has to develop appropriate approxiations for $(z - \tilde{H}_c)^{-1}$, e.g., the eikonal type approximations discussed in Refs. 10.

ACKNOWLEDGMENTS

I would like to thank B. Thaller and W. Plessas for many helpful suggestions.

[1]H. van Haeringen, J. Math. Phys. **17**, 995 (1976). For a survey of this and related work see his thesis: "The Coulomb Potential in Quantum Mechanics and Related Topics" (Free University, Amsterdam, 1978).

[2]J. Zorbas, J. Math. Phys. **17**, 498 (1976); **18**, 1112 (1977); **19**, 177, 2426 (1978); **20**, 6 (1979).

[3]J. C. Y. Chen and A. C. Chen, in *Advances in Atomic and Molecular Physics*, edited by D. R. Bates and I. Esterman (Academic, New York, 1972), Vol. 8.

[4]J. D. Dollard, J. Math. Phys. **5**, 729 (1964); Rocky Mountain J. Math. **1**, 5 (1971).

[5]A. G. Gibson and C. Chandler, J. Math. Phys. **15**, 1366 (1974).

[6]D. Zwanziger, Phys. Rev. D **11**, 3481, 3504 (1975).

[7]H. van Haeringen, "*Long Range Potentials in Quantum Mechanics*" (University of Groningen, 1979).

[8]M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91**, 398 (1953); J. Zorbas, Rep. Math. Phys. **9**, 309 (1976); Gy. Bencze, Lett. Nuovo Cimento **17**, 91 (1976); Gy. Bencze, G. Cattapan, and V. Vanzani, Lett. Nuovo Cimento **20**, 248 (1977).

[9]The use of $V^{1/2}t^{sc}(z)V^{1/2}$ instead of some $\tilde{T}^{sc}(z)$ in (1.7) simply avoids unbounded operators.

[10]S. P. Merkuriev, Sov. J. Nucl. Phys. **24**, 150 (1976); Theoret. and Math. Phys. **32**, 680 (1977); **38**, 134 (1979); "On the three-body Coulomb scattering problem", Preprint FUB/HEP 2/80, Free University, Berlin, (1980).

[11]C. Chandler and A. G. Gibson, "A two-Hilbert-space formulation of multichannel scattering theory", lecture presented at the Conference on Mathematical Methods and Applications of Scattering Theory, Washington (1979).

[12]B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton University, Princeton, N.J., 1971).

[13]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

[14]L. Hostler, J. Math. Phys. **8**, 642 (1967).

[15]F. Gesztesy and B. Thaller, "Born expansions for Coulomb-type interactions," J. Phys. A **14** (1980).

[16]S. Klarsfeld, Nuovo Cimento A **48**, 1059 (1967).

[17]W. O. Amrein, J. M. Jauch, and K. B. Sinha, *Scattering Theory in Quantum Mechanics* (Benjamin, Reading, Mass., 1977).

[18]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. III, Scattering Theory* (Academic, New York, 1979).

[19]D. Marchesin and M. L. O'Carroll, J. Math. Phys. **13**, 982 (1972); J. C. Guillot, Indiana, Univ. Math. J. **25**, 1105 (1976).

[20]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV, Analysis of Operators* (Academic, New York, 1978).

[21]V. Enss, Ann. Phys. **119**, 117 (1979); Comm. Math. Phys. **65**, 151 (1979).

[22]J. Weidmann, Math. Z. **98**, 268 (1967).

[23]A. M. Veselova, Theoret. and Math. Phys. **13**, 1200 (1972); H. Grosse, H. Narnhofer, and W. Thirring, in *Proceedings of the Adriatic Summer Meeting on Particle Physics*, edited by M. Mertinis, S. Pallua, and N. Zovko (Rovinj, 1973), p. 22; H. Grosse, H. R. Grümm, H. Narnhofer, and W. Thirring, Acta Phys. Austr. **40**, 97 (1974); Z. Bajzer, Z. Phys. A **278**, 97 (1976).

[24]W. Thirring, *Lehrbuch der Mathematischen Physik*, Vol. 3, Quantenmechanik von Atomen and Molekülen (Springer-Verlag, Wien, 1979).

[25]I. W. Herbst, Comm. Math. Phys. **35**, 181 (1974).

[26]W. O. Amrein and D. B. Pearson, Ann. Inst. H. Poincaŕe A **30**, 89 (1979).

[27]S. Agmon, Goulaouic–Schwartz Seminar, Ecole Polytechnique, Palaiseau (1978).

[28]E. B. Davis, Arch. Rat. Mech. Anal. **63**, 261 (1977).

[29]F. Gesztesy, W. Plessas, and B. Thaller, J. Phys. A **13**, 2659 (1980).

[30]If T is a bounded operator we abbreviate $\mathrm{Re}T = (T + T^*)/2$.

[31]H. Rollnik, Z. Phys. **145**, 639 (1956).

[32]M. Ciafaloni and P. Menotti, Nuovo Cimento **35**, 160 (1965).

[33]M. L. Baeteman and K. Chadan, Ann. Inst. H. Poincaŕe A **24**, 1 (1976); O. Brander and K. Chadan, *A Tauberian theorem in quantum mechanical inverse scattering theory*, Preprint, LPTHE 80/14, Université de Paris-Sud, Orsay, 1980.

[34]J. R. Taylor, Nuovo Cimento B **23**, 313 (1974); M. D. Semon and J. R. Taylor, Nuovo Cimento A **26**, 48 (1975).

[35]F. Gesztesy and C. B. Lang, J. Math. Phys. **22**, 312 (1981).

[36]A. K. Common and T. W. Stacy, J. Phys. A **11**, 275 (1978); C. R. Garibotti and F. F. Grinstein, J. Math. Phys. **20**, 141 (1979).

The equivalence of the Feshbach and J-matrix methods

Hashim A. Yamani
Saudi Arabian National Center for Science and Technology, P.O. Box 6086, Riyadh, Saudi Arabia

(Received 2 April 1981; accepted for publication 28 August 1981)

It is shown, by taking P to be the projection operator on the subspace of function space in which the potential is truncated, that the exact solution of the scattering problem for the truncated potential using the Feshbach formalism is identical to the J-matrix solution.

PACS numbers: 03.65.Nk

I. INTRODUCTION

The J-matrix method[1-3] has been introduced to solve a model scattering problem exactly in L^2-function space. The model is defined by truncating the infinite-dimensional matrix representation of the given short range potential to a finite representation. The resulting Hamiltonian, composed of the exact zeroth-order Hamiltonian H_0 plus the truncated potential, is then solved exactly by finding the representation of its eigenvector in the space. By demanding that the eigenvector behaves asymptotically as a linear combination of the eigenvector solutions of H_0, one is able to write a closed-form expression for the tangent of the phase shift caused by the truncated potential.

Heller and Yamani[1] compared at length the J-matrix method to both the R-matrix method and L^2 Fredholm technique pointing out computational as well as formal similarities. Later, Broad[4] in his analysis of the quadrature resulting from diagonalizing a scattering Hamiltonian in a finite L^2 basis, proved the equivalence of the J-matrix and the Fredholm equivalent quadrature methods. In this paper, we propose to show the formal equivalence of the J-matrix and the Feshbach[5] methods.

The hint to the equivalence comes from the truncation procedure used in the J-matrix method which formally results from an application of a projection operator on a subspace spanned by the first N members of the basis set used. It is, therefore, expected that when the Feshbach P and Q projection operators are properly defined and the Feshbach equation solved exactly, a result identical to the J-matrix is obtained, thereby establishing the equivalence of the two methods.

In Sec. II, the J-matrix procedure is summarized and the main result written down. In Sec. III, the Feshbach equations set up, and the needed results regarding the abbreviated Green's function quoted, leaving the details to the appendix. Finally, in Sec. IV, the equivalence of the two methods is established

II. THE J-MATRIX METHOD

Since the results of the method have already been detailed elsewhere[1,3] only an outline of the steps leading to the main result is given.

A. The basis

A convenient choice for the basis set spanning the L^2 space is either the Slater set,

$$\phi_n(r) = \langle r|\phi_n\rangle = \zeta^{l+1} e^{-\zeta/2} L_n^{2l+1}(\zeta),\ n = 0, 1, 2\ldots, \quad (1)$$

or the Oscillator set,

$$\phi_n(r) = \langle r|\phi_n\rangle = \zeta^{l+1} e^{-\zeta^2/2} L_n^{l+1/2}(\zeta^2),\ n = 0, 1, 2\ldots, \quad (2)$$

where $\zeta = \lambda r$ and λ is a free scaling parameter and L_n^α is the Laguerre polynomial. Both basis sets are complete, although the Slater set is not orthogonal. The set $\{|\bar{\phi}_n\rangle\}$ is defined as the orthogonal complement to the basis; i.e.,

$$\langle\phi_n|\bar{\phi}_m\rangle = \langle\bar{\phi}_n|\phi_m\rangle = \delta_{n,m}. \quad (3)$$

The reason this choice of basis is convenient is the fact that it renders the matrix representation of the J operator,[3,6]

$$J = H_0 - E = \frac{-1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} - E, \quad (4)$$

tridiagonal. This leads to a three-term recursion relation among the coefficients $\{z_n\}_{n=0}^{\infty}$ of the representation of its eigenvector in the basis

$$J_{n,n-1} z_{n-1} + J_{n,n} z_n + J_{n,n+1} z_{n+1} = 0, \quad (5)$$

where $J_{n,m} = \langle\phi_n|J|\phi_m\rangle$. This equation has two basic solutions, $z_n = s_n$ and $z_n = c_n$ such that the functions

$$S(r) = \langle r|S\rangle = \sum_{n=0}^{\infty} s_n \phi_n(r) \quad (6)$$

and

$$C(r) = \langle r|C\rangle = \sum_{n=0}^{\infty} c_n \phi_n(r) \quad (7)$$

behave asymptotically sinelike and cosinelike, respectively,[6] i.e.,

$$S(r) \underset{r\to\infty}{\sim} \sin(kr - \pi l/2), \quad (8)$$

$$C(r) \underset{r\to\infty}{\sim} \cos(kr - \pi l/2). \quad (9)$$

Furthermore, Heller[7] showed that the J-matrix can be inverted, thereby giving the matrix representation of the Green's function; e.g.,

$$G_{n,m}^{(+)} = \langle\bar{\phi}_n|[J^{(+)}]^{-1}|\bar{\phi}_m\rangle = (-2/k) s_{n_<}(c_{n_>} + i s_{n_>}), \quad (10)$$

where $n_<$ $(n_>)$ is the lesser (greater) of the two numbers n, m.

B. The phase shift

When a short range potential V is given, the J-matrix method solves the scattering problem exactly for a model potential $\bar{V}$ whose matrix elements are identical to those of V in the finite $N\times N$ block and vanish outside it; i.e.,

$$\bar{V}_{nm} = \begin{cases} V_{nm}, & 0 \leqslant n, m \leqslant N-1, \\ 0 & \text{otherwise}. \end{cases}$$

 0022-2488/82/010083-04$02.50

It has been argued[1,3] that the eigenvector solution has the form

$$|\Psi\rangle = \sum_{n=0}^{N-1} R_n|\phi_n\rangle + |S\rangle + t\,|C\rangle, \tag{11}$$

where t is identified as the tangent of the phase-shift caused by the model potential. The projection of the null vector

$$(J+\overline{V})|\Psi\rangle = 0$$

on each member of the basis results in enough conditions to solve for the unknowns, $\{R_n, t\}$. The result of interest is that for t,

$$t = -\frac{s_{N-1} + r(E)J_{N-1,N}s_N}{c_{N-1} + r(E)J_{N-1,N}c_N}. \tag{12}$$

Here $r(E)$ is the $(N-1, N-1)$ matrix element of the Green's matrix which inverts the $N\times N$ representation of the full matrix $(J+\overline{V})$.

The goal for the rest of the paper is to show that the result (12) can be obtained by using the Feshbach method.

III. THE FESHBACH METHOD

A. The projection operators

The truncation procedure used in the J-matrix method has the effect of using an operator which projects on to the subspace U_N spanned by the first N members of the basis set. More precisely the P operator is defined as

$$P = \sum_{n=0}^{N-1} |\phi_n\rangle\langle\bar{\phi}_n|. \tag{13}$$

It is clear that P is idempotent and its range is the subspace U_N. Thus, P is a projection operator. It is also clear that the adjoint operator

$$P^\dagger = \sum_{n=0}^{N-1} |\bar{\phi}_n\rangle\langle\phi_n| \tag{14}$$

is a projection operator in the dual space. With P, the projection operators $Q = 1-P$ and hence $Q^\dagger = 1 - P^\dagger$ are defined and have obvious meanings. The model potential can now be simply written as

$$\overline{V} = P^\dagger V P. \tag{15}$$

B. The Feshbach equations

Associated with the definition of P and Q is a natural division of L^2 space into "inner" and "outer" spaces, using the Feshbach language. The wavevector $|\Psi\rangle$ which solves the Schrödinger equation

$$(H_0 + P^\dagger VP - E)|\Psi\rangle = 0 \tag{16}$$

can be divided into two parts, $P|\Psi\rangle$ and $Q|\Psi\rangle$, which satisfy the coupled Feshbach equations

$$[P^\dagger(J+V)P]P|\Psi\rangle + [P^\dagger JQ]Q|\Psi\rangle = 0 \tag{17}$$

and

$$[Q^\dagger JP]P|\Psi\rangle + [Q^\dagger JQ]Q|\Psi\rangle = 0. \tag{18}$$

As is usually done in the Feshbach method,[8] Eqs. (17) and (18) are solved for $Q|\Psi\rangle$ with the requirement that its asymptotic value be identical to that of $|\Psi\rangle$ itself; i.e., that $Q|\Psi\rangle$ contains the correct phase shift for the true scattering problem caused by $\overline{V}$. For example, if an outgoing wave boundary condition is imposed, then $Q|\Psi\rangle$, as well as $|\Psi\rangle$, will have the following asymptotic form:

$$\Psi_Q(r) = \langle r|Q|\Psi\rangle \underset{r\to\infty}{\sim} e^{i\delta_F}\sin(kr - l\pi/2 + \delta_F), \tag{19}$$

where δ_F is the Feshbach phase shift which is to be compared with the J-matrix phase shift of Eq. (12).

It is noted that, due to the tridiagonal nature of the J matrix, the operator $(P^\dagger JQ)$ has the simple form

$$(P^\dagger JQ) = |\bar{\phi}_{N-1}\rangle J_{N-1,N}\rangle\bar{\phi}_N| = (Q^\dagger JP)^\dagger. \tag{20}$$

Therefore, Eq. (17) can be written as

$$P|\Psi\rangle = [P^\dagger(J+V)P]^{-1}|\bar{\phi}_{N-1}\rangle J_{N-1,N}\langle\bar{\phi}_N|Q|\Psi\rangle. \tag{21}$$

Consequently, Eq. (18) can be written as

$$(Q^\dagger JQ + V_{\text{opt}})Q|\Psi\rangle = 0, \tag{22}$$

where the Feshbach optical potential V_{opt} is given by

$$V_{\text{opt}} = -|\bar{\phi}_N\rangle J_{N,N-1}\langle\bar{\phi}_{N-1}|[P^\dagger(J+V)P]^{-1}|\bar{\phi}_{N-1}\rangle \times J_{N-1,N}\langle\bar{\phi}_N|. \tag{23}$$

It is noted that the matrix element $\langle\bar{\phi}_{N-1}|[P^\dagger(J+V)P]^{-1}|\bar{\phi}_{N-1}\rangle$ is just the $r(E)$ used in the J-matrix method. Now, Eq. (22) can be solved for $Q|\Psi\rangle$:

$$Q|\Psi\rangle = |\chi\rangle - [Q^\dagger JQ^{-1}]V_{\text{opt}}Q|\Psi\rangle, \tag{24}$$

where $|\chi\rangle$ satisfies the conditions:

(i) $P|\chi\rangle = 0$,

(ii) $\{Q^\dagger JQ\}|\chi\rangle = 0$,

(iii) $\langle r|\chi\rangle \underset{r\to\infty}{\sim} \langle r|S\rangle$.

It is clear that the vector $|\chi\rangle$ is different from $Q|S\rangle$ since it can be easily shown that

$$[Q^\dagger JQ]Q|S\rangle = -|\bar{\phi}_N\rangle J_{N,N-1}s_{N-1} \neq 0.$$

In fact, the choice

$$|\chi\rangle = Q|S\rangle + [Q^\dagger JQ]^{-1}|\bar{\phi}_N\rangle J_{N,N-1}s_{N-1} \tag{25}$$

satisfies the stated condition, and is hence the desired vector. Therefore Eq. (24) can be written explicitly as

$$Q|\Psi\rangle = Q|S\rangle + [Q^\dagger JQ]^{-1}|\bar{\phi}_N\rangle J_{N,N-1} \times\{s_{N-1} + r(E)J_{N-1,N}\langle\bar{\phi}_N|Q|\Psi\rangle\}. \tag{26}$$

The appendix analyzes in details the inverse of $(Q^\dagger JQ)$ which has been called the abbreviated[9] Green's function $\overline{G}$. Thus, with an outgoing wave boundary condition built in, the matrix elements needed for the solution of Eq. (26) are

$$\overline{G}^{(+)}_{n,N} = \frac{-1}{J_{N-1,N}}\,\frac{c_n + is_n}{c_{N-1} + is_{N-1}}, \quad n \geqslant N. \tag{A14}$$

Now it becomes easy to use Eq. (26) to solve for $\langle\bar{\phi}_N|Q|\Psi^{(+)}\rangle$ and to insert the result back into the right-hand side of Eq. (26). The final result is

$$Q|\Psi^{(+)}\rangle = Q|S\rangle - \overline{G^{(+)}}|\bar{\phi}_N\rangle J_{N,N-1}(c_{N-1} + is_{N-1}) \times b/(a - ib), \tag{27}$$

where

$$a = c_{N-1} + r(E)J_{N-1,N}c_N \tag{28a}$$

and

$$-b = s_{N-1} + r(E)J_{N-1,N}s_N. \qquad (28b)$$

Equation (27) is the desired explicit solution of the Feshbach equations with outgoing wave boundary condition.

IV. THE EQUIVALENCE

In order to compare the Feshbach and the J-matrix phase-shifts, Eq. (27) is needed to examine the asymptotic form of $Q|\Psi^{(+)}\rangle$ in light of condition (19). More easily, the coefficients of $Q|\Psi^{(+)}\rangle$ will be found first and $Q|\Psi\rangle$ will be reconstructed again. If $Q|\Psi^{(+)}\rangle$ is written as

$$Q|\Psi^{(+)}\rangle = \sum_{n=N}^{\infty} d_n^{(+)}|\phi_n\rangle, \qquad (29)$$

then it is easily seen that

$$d_n^{(+)} = s_n - \overline{G_{n,N}^{(+)}}J_{N,N-1}(c_{N-1} + s_{N-1}) \times b/(a-ib) \quad n \geqslant N \qquad (30)$$

which, with the use of Eq. (A14) becomes

$$d_n^{(+)} = (as_n + bc_n)/(a-ib). \qquad (31)$$

Consequently $Q|\Psi^{(+)}\rangle$ can be reconstructed again as

$$Q|\Psi^{(+)}\rangle = [a/(a-ib)]Q|S\rangle + [b/(a-ib)]Q|C\rangle.$$

Thus, asymptotically $Q|\Psi^{(+)}\rangle$ behaves as

$$\Psi_Q^{(+)}(r) = \langle r|Q|\Psi^{(+)}\rangle \underset{r\to\infty}{\sim} [a/(a-ib)]S(r) + [b/(a-ib)]C(r) \quad (32)$$

or more explicitly,

$$\Psi_Q^{(+)}(r) \underset{r\to\infty}{\sim} [a/(a-ib)]\sin(kr - \pi l/2) + [b/(a-ib)] \times\cos(kr - \pi l/2). \qquad (33)$$

In reaching this result, Eqs. (8) and (9) have been used. The asymptotic behavior exhibited by Eq. (33) is consistent with that required by Eq. (19) provide δ_F is connected to the quantities a and b via the relation

$$\tan\delta_F = b/a. \qquad (34)$$

With a and b given by Eqs. (28a) and (28b), it is easily seen that the Feshbach phase shift (34) is identical to the J-matrix phase shift (12). This completes the proof of equivalence.

V. DISCUSSION

It is clear from the analysis that "folding-in" the physics of the P-part of the space into the Q-part in terms of an optical potential $V_{\rm opt}$, as well as subsequent steps in the analysis, are exact. Unlike the usual procedure, no approximation is made in the optical potential to be able to solve the Schrödinger equation in the Q-part of the space [Eq. (22)].

Since the J-matrix and Feshbach methods solve the *model* problem *exactly*, the equivalence is, therefore, not suprising. The benefit of the previous analysis has been to show the precise sense in which the J-matrix divides the L^2-function space into "inner" and "outer" parts. Also, in course of the proof of equivalence, more light is shed on the representation of the various abbreviated Green's functions in the basis chosen.

ACKNOWLEDGMENT

The author thanks E. J. Heller for suggesting the problem and for helpful discussions. The initial phase of this work was carried out while the author was at the Physics Department, University of Petroleum and Minerals, Dhahran, Saudi Arabia.

APPENDIX: THE ABBREVIATED GREEN'S FUNCTIONS

The Green's matrix as given by Eq. (10) is the inverse of the J-matrix; i.e.,

$$JG = GJ = 1. \qquad (A1)$$

Here, what is of interest is the inverse to the abbreviated[9] J-matrix $Q^\dagger JQ$. With the help of the projection operators P and Q (A1) can be written in more details as

$$\begin{pmatrix} P^\dagger JP & P^\dagger JQ \\ Q^\dagger JP & Q^\dagger JQ \end{pmatrix}\begin{pmatrix} PGP^\dagger & PGQ^\dagger \\ QGP^\dagger & QGQ \end{pmatrix} = \begin{pmatrix} P^\dagger 1P^\dagger & 0 \\ 0 & Q^\dagger 1Q^\dagger \end{pmatrix}. \qquad (A2)$$

Equation (A2) is actually four equations in one, namely,

$$(P^\dagger JP)(PGP^\dagger) + (P^\dagger JQ)(QGP^\dagger) = P^\dagger, \qquad (A3)$$

$$(P^\dagger JP)(PGQ^\dagger) + (P^\dagger JQ)(QGP^\dagger) = 0, \qquad (A4)$$

$$(Q^\dagger JP)(PGP^\dagger) + (Q^\dagger JQ)(QGP^\dagger) = 0, \qquad (A5)$$

$$(Q^\dagger JP)(PGQ^\dagger) + (Q^\dagger JQ)(QGQ^\dagger) = Q^\dagger. \qquad (A6)$$

Equations (A5) and (A6) can be solved to yield the relation

$$(Q^\dagger JQ)(QGQ^\dagger - QGP^\dagger(PGP^\dagger)^{-1}PGQ^\dagger) = Q^\dagger. \quad (A7)$$

Therefore, the inverse of $(Q^\dagger JQ)$, which is what has been called $\overline{G}$ in the text, is given by

$$\overline{G} = (QGQ^\dagger - QGP^\dagger(PGP^\dagger)^{-1}PGQ^\dagger). \qquad (A8)$$

In order to find $\overline{G}$ explicitly, $(PGP^\dagger)^{-1}$ has to be found. From Eq. (A3), it is clear that

$$(PGP^\dagger)^{-1} = F(P^\dagger JP), \qquad (A9)$$

where F satisfies the relation $DF = P^\dagger$ and $D = P^\dagger - P^\dagger JQ\, QGP^\dagger$. Due to the simplicity of the operator $P^\dagger JQ$, operator F can be easily obtained,

$$F = P^\dagger + (P^\dagger JQ)(QGP^\dagger)/\Delta, \qquad (A10)$$

where

$$\Delta = 1 - J_{N-1,N}G_{N,N-1}.$$

Consequently, Eq. (A8) for $\overline{G}$ reduces to

$$\overline{G} = QGQ^\dagger + QG|\bar{\phi}_{N-1}\rangle J_{N-1,N}\langle\bar{\phi}_N|GQ^\dagger/\Delta. \quad (A11)$$

More explicitly, if p and q are integers greater than or equal to N, then

$$\overline{G}_{pq} = G_{pq} + G_{p,N-1}J_{N-1,N}G_{N,q}/\Delta. \qquad (A12)$$

It is stressed that the matrix elements of $\overline{G}$ are indeed symmetric in p and q. In particular, if $q = N$, the outgoing-wave abbreviated Green's function is

$$\overline{G_{pN}^{(+)}} = \overline{G_{Np}^{(+)}} = G_{pN}^{(+)} + G_{N,N-1}^{(+)}J_{N-1,N}G_{Np}^{(+)}/\Delta^{(+)} = G_{p,N}^{(+)}/(1 - J_{N-1,N}G_{N,N-1}^{(+)}), \quad p \geqslant N. \qquad (A13)$$

Since[7]

$$J_{N-1,N}(s_N c_{N-1} - s_{N-1} c_N) = k/2,$$

and with the help of Eq. (10), (A13) reduces to

$$\overline{G^{(+)}_{pN}} = \frac{-1}{J_{N-1,N}} \frac{(c_p + is_p)}{(c_{N-1} + is_{N-1})}, \quad p \geqslant N \qquad \text{(A14)}$$

which has been quoted in the text.

As a by-product of the above analysis, the inverse $\widetilde{G}$ to the abbreviated J-matrix $(P^{\dagger}JP)$ may be found. Starting from the relations (A3)–(A6) and following a similar procedure as outlined above, an analogous expression $\widetilde{G}$ may be obtained.

$$\widetilde{G} = PGP^{\dagger} + PG\,|\overline{\phi}_N\rangle J_{N,N-1}\langle \phi_{N-1}|GP^{\dagger}/\Delta.$$

Explicitly, if i and j are integers less or equal to $N-1$, then

$$\widetilde{G}_{ij} = G_{ij} + G_{i,N} J_{N,N-1} G_{N-1,j}/\Delta.$$

In particular, $\widetilde{G}_{N-1,N-1} = -s_{N-1}/(J_{N-1,N} s_N)$.

Therefore, when $V=0$, $r(E) = \widetilde{G}_{N-1,N-1}$. Thus Eq. (12) implies that $\tan\delta = 0$. This result is of course expected, yet never proved before.

[1]E. J. Heller and H. A. Yamani, Phys. Rev. A **9**, 1201 (1974).

[2]E. J. Heller and H. A. Yamani, Phys. Rev. A **9**, 1209 (1974).

[3]H. A. Yamani and L. Fishman, J. Math. Phys. **16**, 410 (1975).

[4]J. T. Broad, Phys. Rev. A **18**, 1012 (1978).

[5]H. Feshbach, Ann. Phys. (N.Y.) **19**, 287 (1962).

[6]The case when H_0 contains an $1/r$ Coulomb term can be treated similarly in the Laguerre basis; see Ref. 3.

[7]E. J. Heller, Phys. Rev. A **12**, 1222 (1975).

[8]S. Geltman, *Topics in Atomic Scattering Theory* (Academic, New York, 1969), pp. 144.

[9]The term "P-abbreviated matrix B" has been used by N.I. Akhiezer [*The Classical Moment Problem* (Hanfer, New York, 1965)] to indicate the original matrix B less the P rows and columns. In our case $(Q^{\dagger}JQ)$ would be called the N-abbreviated matrix J.

The asymptotic form of the continuum wavefunctions and redundant poles in the Heisenberg condition

Paul Terry
Department of Physics, The University of Texas at Austin, Austin, Texas 78712

(Received 15 August 1980; accepted for publication 26 November 1980)

The derivation of the Heisenberg condition is re-examined to show why it is not an identity for potentials possessing redundant poles. Consideration of several such potentials for which exact solutions are known reveals that, in the process of taking an asymptotic limit, the usual derivation of the Heisenberg condition improperly neglects a set of terms. These terms are just those necessary to make the Heisenberg condition an identity; more importantly, it is demonstrated that these terms, providing information on the redundant poles, i.e., the sum of the residues of $S(k)$ at the redundant poles, come from the asymptotic expansion of the continuum wavefunction. By this we are able to give the details of the nature of the asymptotic expansion of the continuum wavefunction and the information contained therein.

PACS numbers: 03.65.Nk, 11.20.Fm

1. INTRODUCTION

The analytic S-matrix of potential scattering theory is generally credited with incorporating information on all of the bound-state solutions corresponding to a given potential. Specifically, the bound-state energies correspond to poles of the S-matrix $S(k)$, with the energy of the state being given by the value of k for which the pole occurs.

In addition to these poles, it has long been known that the S-matrix may possess other poles which are dynamical (i.e., they disappear when the potential is cut off at large distances) but correspond to no bound states of the potential. Several authors have investigated these poles[1-3] and have shown in particular that the states corresponding to these poles do not contribute to the completeness relation. This is contrasted with the fact that the bound states do contribute to the completeness relation and must be included in any set of states in order for it to be a complete set.

The fact that states associated with redundant poles do not contribute to the completeness relation, whereas bound states do, has some interesting consequences. Some of these consequences have been investigated and detailed by Biswas *et al.*[3] in relation to the analogous concept of shadow states in quantum field theory. In particular, they demonstrated how the roles of redundant poles and bound-state poles can be interchanged, both with local potentials and separable nonlocal potentials. It was shown that phase equivalent systems can provide cases where two distinct potentials having an identical S-matrix can be such that one has a bound state while the other does not. Thus, the unique pole applying to both theories is a bound-state pole in one theory and a redundant pole in the other. In the case of a separable potential, it was shown how in a theory containing two redundant poles, a larger theory may be constructed in which the two redundant poles are associated with bound states, and thus cease to be redundant. They also solved the inverse problem, showing how in a theory with two poles, a reduced theory may be constructed in which both poles have become redundant.

The fact that the roles of bound-state poles and redundant poles may be interchanged carries implications on the nature of the continuum wavefunctions. This is manifested in the resolution of a seeming paradox given by Biswas *et al.*[3] They consider a general condition on the S-matrix obtained by Heisenberg[4] from the completeness condition

$$\sum_n u_n^*(r)u_n(r') + \int_0^\infty dk\, u_k^*(r)u_k(r') = \delta(r-r'). \tag{1.1}$$

Since Eq. (1.1) is valid for all values of r and r' Heisenberg replaced the discrete wavefunction $u_n(r)$ and the continuum wave function $u_k(r)$ by their asymptotic expressions:

$$u_n(r) \sim C_n(2\pi)^{-1/2}\exp(-|k_n|r), \tag{1.2}$$

$$(r\to\infty)$$

$$u_k(r) \sim (2/\pi)^{1/2}\sin(kr+\delta(k)). \tag{1.3}$$

The resulting condition

$$\int_{-\infty}^{\infty} dk\, S(k)e^{ik(r+r')} = \sum_n |C_n|^2 e^{-|k_n|(r+r')} \tag{1.4}$$

contains on the right-hand side only the contribution from the bound states, the states corresponding to the redundant poles making no contribution. However, for S-matrices containing redundant poles the left-hand side is evaluated as

$$\int_{-\infty}^{\infty} dk\, S(k)e^{ik(r+r')} = \sum_n |C_n|^2 e^{-|k_n|(r+r')} - \sum_r \left(\frac{dS}{dk}\right)^{-1}\Bigg|_{k=i|k_r|} e^{-|k_r|(r+r')}, \tag{1.5}$$

where the second term on the right-hand side gives the residue of $S(k)$ at the redundant poles.

The fact that S-matrices containing redundant poles do not satisfy the Heisenberg condition suggests that the asymptotic expressions (1.2) and (1.3) used in deriving Eq. (1.4) do not give the correct asymptotic form of the completeness relation. In a paper by Nelson *et al.*,[5] the effect of discarding the terms not present in the asymptotic expressions (1.2) and (1.3) is discussed in its relation to the essential character of the completeness relation as an identity. They show that by retaining in the derivation the terms normally discarded, an identity is maintained throughout. This exercise is performed, however, without considering asymptotic expres-

 0022-2488/82/010087-05$02.50

sions, and the identity in its final form, obtained after some algebraic manipulations, necessarily appears as an obvious equality. They are thus led to the conclusion that the correct derivation of the Heisenberg condition yields a tautology. The manipulations of Nelson *et al.* do not, however, consitute a derivation of the Heisenberg condition. The Heisenberg condition is an asymptotic statement and its correct derivation requires asymptotic analysis, i.e., in a derivation valid to a given order, it is necessary to ensure that all terms of that order are retained and that all terms retained contribute in that order. When this is done the resulting identity is far from a tautology because it makes a statement concerning the physical or mathematical content of the order considered. Indeed, as we shall see, the amended Heisenberg condition, now rigorously an identity, contains contributions arising from more than one order in the expansion of the wavefunction and thus reveals structure in the asymptotic series while clarifying the content of the S-matrix and its meaning.

Thus, we will demonstrate that the correct asymptotic form of the Heisenberg condition must include an additional set of terms, these terms being just those represented in equation (1.5) as the sum of the residues of $S(k)\exp(ik(r+r'))$ at the redundant poles. They come from the term $\int_0^\infty dk\, u_k^*(r)u_k(r')$ in the completeness relation when for the continuum wavefunction $u_k(r)$ we substitute not Eq. (1.3) but Eq. (1.3) with the addition of the next term in the asymptotic expansion of $u_k(r)$. This next term in the asymptotic expansion of $u_k(r)$ is subdominant to the leading contribution given by Eq. (1.3). For this reason it is not usually included in statements on the asymptotic form of the scattering wavefunction and was not retained in the above derivation of the Heisenberg condition. However, when redundant poles are present these terms are exponentially damped and of the same order as the right-hand side of Eq. (1.4). Hence, this term makes a contribution to $\int_0^\infty dk\, u_k^*(r)u_k(r')$ of the same order as $S(k)\exp(ik(r+r'))$ evaluated at the bound states and redundant poles. So while the states associated with the redundant poles can make no contribution to the completeness relation in the sum over states $\Sigma_n u_n^*(r)u_n(r')$, a knowledge of the redundant poles is nevertheless present in the completeness relation. This knowledge is contained in the integral over the continuum $\int_0^\infty dk\, u_k^*(r)u_k(r')$ in the leading order correction term to the dominant asymptotic form of the continuum wavefunction $u_k(r)$.

This may be stated in terms of Eq. (1.5) which we now take as the correct form of the Heisenberg condition for the potentials with redundant poles: The integral over the S-matrix comes from the dominant asymptotic form of the continuum wavefunction [Eq. (1.3)], the sum over bound states comes directly from the sum over bound states in the completeness relation, and the sum of residues of $S(k)\exp(ik(r+r'))$ at the redundant poles comes from the leading correction term to the dominant asymptotic form of the continuum wavefunction.

From a different point of view, we observe that given an S-matrix and an *a priori* way of distinguishing the bound-state poles we have information contained in the leading correction term to the dominant form of the asymptotic continuum wavefunction. This occurs even though the S-matrix itself is formed only from the dominant form of the asymptotic wavefunction.

In the remainder of this paper we will establish the above statements by considering several potentials for which exact solutions are known. These are potentials previously studied by Ma,[1] Ter Haar,[2] Biswas *et al.*,[3] Bargmann,[6] Bhattacharjie and Sudarshan[7] and Nelson *et al.*,[5] and are illustrative of a variety of phenomena possible from different redundant pole "spectra." In each case we will consider the completeness relation with the exact wavefunctions for zero angular momentum, carefully performing the integral over the continuum and taking the asymptotic limit. This paper is organized as follows. In Sec. 2 we analyze the case of the exponential potential first studied by Ma. Phase equivalent systems are examined in Sec. 3 and 4, taking examples from Bargmann. In Sec. 3 we consider the potentials of the linear type and in Sec. 4 we address ourselves to the potentials of the quadratic type. In Sec. 5 we give the conclusions.

2. THE EXPONENTIAL POTENTIAL

The first potential we consider is the rather simple potential of Ma, referred to as the exponential potential

$$V(r) = -V_0\exp(-\alpha r).$$

We seek solutions of the transformed radial wave equation for zero angular momentum, $\phi_k''(r) + k^2\phi_k(r) = V(r)\phi_k(r)$, where $\phi(r) = r\cdot\psi(r)$ and $\psi(r)$ is the radial wavefunction. With $V(r)$ as given above a simple transformation of the wave equation yields a Bessel equation. The complete solution for $\phi_k(r)$ up to a multiplicative factor is

$$\begin{aligned}\phi_k(r) = \left(2i\left|\frac{\Gamma(i\eta+1)}{(A/2)^{i\eta}}J_{i\eta}(A)\right|\right)^{-1}\{|\Gamma(i\eta+1)|^2\\ \times[J_{i\eta}(A)J_{-i\eta}(Ae^{-\alpha r/2})\\ - J_{-i\eta}(A)J_{i\eta}(Ae^{-\alpha r/2})]\},\end{aligned} \tag{2.1}$$

where we have written the solution in terms of the linearly independent Jost functions $f(k,r)$ and $f(-k,r)$ constructed so that $\lim_{r\to\infty}\exp(ikr)f(k,r) = 1$ and $\lim_{r\to\infty}\exp(-ikr)\times f(-k,r) = 1$. Here

$$f(k,r) = \Gamma(i\eta+1)J_{i\eta}(A\exp(-\alpha r/2)),$$

where J_ν is the Bessel function of order ν, Γ is the gamma function, A is a constant, and η equals $2k/\alpha$.

The series expansion of the Bessel functions gives the asymptotic expansion of $\phi_k(r)$, which we use in the completeness relation to obtain the asymptotic expansion of $\int_0^\infty dk\,\phi_k^*(r)\phi_k(r')$. This integral is given exactly by the series

$$\begin{aligned}\int_{-\infty}^{\infty} dk\,\phi_k^*(r)\phi_k(r')\\ = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\,\{e^{ik(r-r')}[1 + a_1(k,r) + a_1(-k,r')\\ + a_1(k,r)a_1(-k,r')] - e^{ik(r+r')}S(k)\\ \times[1 + a_1(k,r) + a_1(k,r') + a_1(k,r)a_1(k,r')]\},\end{aligned} \tag{2.2}$$

where

$$a_1(k,r) = \Gamma(-i\eta+1)\sum_{j=1}^{\infty}\frac{(-1)^j(A/2)^{2j}}{j!\Gamma(-i\eta+j+1)}e^{-\alpha rj} \tag{2.3}$$

and $S(k)$ is the S-matrix given by

$$S(k) = \frac{J_{i\eta}(A)\Gamma(i\eta+1)}{J_{-i\eta}(A)\Gamma(-i\eta+1)}\left(\frac{A}{2}\right)^{-2i\eta}. \tag{2.4}$$

The imaginary values of η for which $J_{-i\eta}(A) = 0$ comprise the bound-state poles. The poles of $\Gamma(i\eta+1)$ (occurring also for η imaginary) are the redundant poles. We note that there are an infinite number of redundant poles evenly spaced along the positive imaginary η axis beginning at $\eta = i$.

We now consider the dominant contribution to equation (2.2) in the asymptotic limit. Because of the presence of factors of $\exp(-\alpha rj)$, $\exp(-\alpha r'l)$, and $\exp(-\alpha(rj + r'l))$ in all terms containing a_1, these terms are subdominant to $\exp(ik(r-r'))\cdot 1$ and $\exp(ik(r+r'))\cdot S(k)$ as $r, r'\to\infty$. Dropping these subdominant terms at this point in the derivation constituties the use of the asymptotic expressions given in Eqs. (1.2) and (1.3). On the other hand, if all terms in the asymptotic expansion of $\int dk\, \phi^*\phi$ are retained until after the integration over k, we find that additional terms must be included in the dominant contribution.

Considering Eq. (2.2), we obtain $\delta(r-r')$ for the integral over the first term. The remaining terms are evaluated by contour integration. When we take $r, r'\to\infty$ subsequent to doing the integration, we will do it in such a way that $r > r'$; this enables the contour to be closed above for all terms with vanishing contribution on the closing semicircle. The only poles lie on the imaginary k axis and are contained in either the factor $1/J_{-i\eta}(A)$ or $\Gamma(i\eta+1)$. There are six terms in Eq. (2.2) containing either or both of these factors; they are $a_1(-k,r')$, $a_1(k,r)a_1(-k,r')$, $S(k)$, $S(k)a_1(k,r)$, $S(k)a_1(k,r')$, and $S(k)a_1(k,r)a_1(k,r')$.

We begin by computing $\int S(k)\exp(ik(r+r'))\,dk$. At the redundant poles the residue of $S(k)\exp(ik(r+r'))$ is given by

$$\operatorname{Res}\{e^{ik(r+r')}S(k);n\} = -e^{-(\alpha n/2)(r+r')}[(-1)^n/n!](A/2)^{2n}, \tag{2.5}$$

where n is the n^{th} redundant pole of $S(k)$ corresponding to $\eta = in (n = 1, 2, 3\cdots)$. We then compute the residues at the redundant poles of the other terms $a_1(-k,r')$, $a_1(k,r)a_1(-k,r')$, etc. to compare asymptotically with Eq. (2.5). For $a_1(-k,r')$ we find that

$$\operatorname{Res}\{e^{ik(r-r')}a_1(-k,r');n\} = e^{-(n\alpha/2)(r-r')}\sum_{j=1}^{\infty}\frac{(-1)^j(A/2)^{2j}e^{-\alpha r'j}}{j!\Gamma(-n+j+1)}.$$

We consider the contribution to this result from each j. At values of $j < n$ the contribution is zero due to the factor $[\Gamma(-n+j+1)]^{-1}$, at $j = n$ we have

$$\operatorname{Res}\Big\{e^{ik(r-r')}\underset{(n\text{th term})}{a_1}(-k,r');n\Big\} = e^{-(n\alpha/2)(r+r')}\frac{(-1)^n(A/2)^n}{n!}, \tag{2.6}$$

and at values of $j > n$ each contribution is smaller by a factor of $\exp(-r'(j-n))$ than the contribution for $j = n$. But the dominant contribution $j = n$ given by Eq. (2.6) is exactly equal and opposite to the residue of $S(k)\exp(ik(r+r'))$ at the nth redundant pole. These terms must be included in the dominant contribution and are the additional terms of Eq. (1.5) which were not present in Eq. (1.4). In looking again at Eq. (2.2) we observe that even though $a_1(\mp k, r')$ is subdominant to 1 in the asymptotic expansion of $\phi_k(r)$, $\exp(ik(r-r'))a_1(-k,r')$ at selected values of k and j is of the same order as $\exp(ik(r+r'))S(k)$ at those values of k. It just so happens that those critical values of k coincide with the poles and that the residues are equal and opposite.

If we now compute the residues of the remaining terms in Eq. (2.2), we find that they are all subdominant to the residues of $\exp(ik(r+r'))S(k)$ at the bound-state poles and redundant poles. We therefore find that the completeness relation, Eq. (1.1), leads to Eq. (1.5) and not Eq. (1.4). The leading terms in the asymptotic expansion of the continuum wavefunction give the integral over the S-matrix while the leading correction terms give the redundant pole information of Eq. (1.5). The exact nature of the redundant pole information thus retained in the continuum wavefunction is summarized by writing the asymptotic expansion of $f(k,r)$ as

$$f(k,r) = e^{-ikr}\Big[1 + \sum_{j=1}^{\infty}\mathscr{D}_j(k)e^{-\gamma rj}\Big]. \tag{2.7}$$

We then have that the residue of $S(k)$ at the nth redundant pole $[k = k(n)]$ gives the residue of the nth coefficient at that pole,

$$\operatorname{Res}\{S(k);k(n)\} = \operatorname{Re}\{\mathscr{D}_n(k);k(n)\}, \tag{2.8}$$

with γ given as

$$\gamma = -2ik(n)/n. \tag{2.9}$$

Thus amended the Heisenberg condition is satisfied by the wavefunctions of the exponential potential.

3. THE POTENTIAL OF LINEAR TYPE

The next potential we will examine is one of the phase equivalent families of potentials studied by Bargmann:

$$V(r) = -2\beta\lambda^2[e^{-\lambda r}/(\beta e^{-\lambda r}+1)^2]. \tag{3.1}$$

The radial wavefunction is

$$\phi_k(r) = \frac{1}{2i}\left|\frac{2k - i\lambda}{2k + i\lambda(\beta-1)/(\beta+1)}\right|\Big\{\frac{[-2k + i\lambda(\beta-1)/(\beta+1)][2k + i\lambda(\beta e^{-\lambda r}-1)/(\beta e^{-\lambda r}+1)]}{(2k+i\lambda)(2k-i\lambda)}e^{-ikr} - \frac{[2k + i\lambda(\beta-1)/(\beta+1)][-2k + i\lambda(\beta e^{-\lambda r}-1)/(\beta e^{-\lambda r}+1)]e^{ikr}}{(2k-i\lambda)(2k+i\lambda)}\Big\}. \tag{3.2}$$

Forming the quantity $\int_0^\infty dk\,\phi_k^*(r)\phi_k(r')$ and expanding the integrand for $r, r'\to\infty$, we obtain

$$\int_0^\infty dk\,\phi_k^*(r)\phi_k(r') = \int_{-\infty}^{\infty} dk\,\{e^{ik(r-r')}[1 + a_2(k,r) + a_2(-k,r') + a_2(k,r)a_2(-k,r')] - e^{ik(r+r')}S(k) \times[1 + a_2(-k,r) + a_2(-k,r') + a_2(-k,r)a_2(-k,r')\}], \tag{3.3}$$

where

$$a_2(k,r) \sim i\lambda\,(-2\beta e^{-\lambda r} + 2\beta^2 e^{-2\lambda r} - 2\beta^3 e^{-3\lambda r} + \cdots)/(2k + i\lambda) \quad (r \to \infty) \tag{3.4}$$

and

$$S(k) = \frac{[2k + i\lambda(\beta - 1)/(\beta + 1)](2k + i\lambda)}{[2k - i\lambda(\beta - 1)/(\beta + 1)](2k - i\lambda)}. \tag{3.5}$$

As before, we note that the terms $\exp(ik(r - r'))\cdot 1$ and $\exp(ik(r + r'))S(k)$ appear to represent the dominant contributions to Eq. (3.3). Retaining only these terms constitutes the use of the asymptotic forms given in Eqs. (1.2) and (1.3). Here we retain all terms, however, and then integrate. The terms $a_2(-k, r')$, $a_2(k, r)a_2(-k, r')$, $S(k)$, $S(k)a_2(-k, r)$, $S(k)a_2(-k, r')$, and $S(k)a_2(-k, r)a_2(-k, r')$ all possess poles and contribute to the integral. The pole at $k = \frac{1}{2}i\lambda(\beta - 1)/(\beta + 1)$ is the bound-state pole while the pole at $k = \frac{1}{2}i\lambda$ is the redundant pole. We find the residue at $k = \frac{1}{2}i\lambda$ of each of the above terms: for $S(k)$ we have

$$\mathrm{Res}\{S(k)e^{ik(r + r')};\ k = i\lambda/2\} = 2i\lambda\beta e^{-\frac{1}{2}\lambda(r + r')}, \tag{3.6}$$

and for $\exp(ik(r - r'))a_2(-k, r')$ we obtain

$$\mathrm{Res}\{e^{ik(r - r')}a_2(-k, r');\ k = i\lambda/2)\} = 2i\lambda\beta e^{-\frac{1}{2}\lambda(r + r')} + O(e^{-\frac{1}{2}\lambda(r + 3r')}). \tag{3.7}$$

The leading contribution to the residue of $\exp(ik(r - r'))a_2(-k, r')$ at the redundant pole is equal and opposite to that of $\exp(ik(r + r'))S(k)$ and so provides the additional term of Eq. (1.5). Evaluation of all remaining residues shows that they are subdominant as $r, r' \to \infty$.

4. THE POTENTIALS OF QUADRATIC TYPE

In this section we examine four additional phase equivalent families of potentials due to Bargmann,

$$V_1(r) = \frac{\rho\sigma[4\rho\sigma + (\rho - \sigma)^2\cosh((\rho + \sigma)r - 2\theta) - (\rho + \sigma)^2\cosh(\rho - \sigma)r]}{[\sigma\sinh(\rho r - \theta) - \rho\sinh(\sigma r - \theta)]^2}, \tag{4.1}$$

$$V_2(r) = \frac{\rho\sigma[4\rho\sigma + (\rho - \sigma)^2\cosh(\rho + \sigma)r - (\rho + \sigma)^2\cosh((\rho - \sigma)r + 2\phi)]}{[\sigma\sinh(\rho r + \phi) - \rho\sinh(\sigma r - \phi)]^2}, \tag{4.2}$$

$$V_3(r) = \frac{-2(\rho/\sigma)(\rho + \sigma)^2 e^{-(\rho + \sigma)r}}{1 + (\rho/\sigma)e^{-(\rho + \sigma)r}}, \tag{4.3}$$

$$V_4(r) = \frac{-\rho\sigma[4\rho\sigma + (\rho - \sigma)^2\cosh(\rho + \sigma)r + (\rho + \sigma)^2\cosh((\rho - \sigma)r - 2\phi)]}{[\sigma\cosh(\rho r - \phi) + \rho\cosh(\sigma r + \phi)]^2}. \tag{4.4}$$

The solutions of all of these potentials are related and can be parametrically represented as a single solution with parameters α and β subject to the requirement that either $\alpha = -1$ or $\beta = 1$ and that

$$\left.\begin{array}{l}\text{(a) when } \alpha = -1, \quad \beta > 1\\ \text{(b) when } \beta = 1, \quad \alpha > -1\end{array}\right\}. \tag{4.5}$$

Case (a) gives the potential V_1 with $\beta = \exp(2\theta)$ $(\theta > 0)$, and Case (b) gives potentials V_2, V_3, and V_4 with $\alpha = -\exp(-2\phi)$ $(\phi > 0)$, $\alpha = 0$, and $\alpha = \exp(2\phi)$ $(-\infty < \phi < \infty)$, respectively.

The scattering solutions $\phi_k(k)$ are complicated functions of exponentials of r:

$$\phi_k(r) = \left[\frac{-\chi(-k, 0)\chi(k, r)}{\chi(-k, \infty)\chi(k, \infty)}e^{-ikr} + \frac{\chi(k, 0)\chi(-k, r)}{\chi(k, \infty)\chi(-k, \infty)}e^{ikr}\right]\left|\frac{\chi(k, \infty)}{\chi(k, 0)}\right|, \tag{4.6}$$

with

$$\begin{aligned}\chi(k, r) = 4k^2 - \{&4ikp\sigma(e^{\rho r} - \alpha\beta e^{-\rho r} + \alpha e^{\sigma r} - \beta e^{-\sigma r})\\ &- (\sigma^2 - \rho^2)[\sigma(e^{\rho r} + \alpha\beta e^{-\rho r}) - \rho(\alpha e^{\sigma r} + \beta e^{-\sigma r})]\}\\ &\times[\sigma(e^{\rho r} + \alpha\beta e^{-\rho r}) + \rho(\alpha e^{\sigma r} + \beta e^{-\sigma r})]^{-1}.\end{aligned} \tag{4.7}$$

The S-matrix is

$$\begin{aligned}S(k) = {}&[2k - i(\sigma - \rho)][2k + i(\rho + \sigma)]/[2k + i(\sigma - \rho)]\\ &\times[2k - i(\rho + \sigma)].\end{aligned} \tag{4.8}$$

We note that $S(k)$ is independent of the parameters α and β. However, the spectrum of bound states, curiously, is not independent of α and β in the following sense: for $\alpha = -1$ the pole $2k - i(\rho + \sigma) = 0$ corresponds to the bound state and $2k + i(\sigma - \rho) = 0$ corresponds to a redundant pole. For the other possibility, $\beta = 1$, $2k - i(\rho + \sigma) = 0$ now corresponds to a redundant pole, and $2k + i(\sigma - \rho) = 0$ the bound state. For the two possible cases, the bound-state spectrum and redundant pole spectrum interchange. This is the same type of situation studied by Biswas *et al.*

We proceed exactly as we have done in Secs. 2 and 3 obtaining the asymptotic expansion of $\int dk\,\phi_k^*(r)\phi_k(r')$. The expansion of $\phi^*\phi$ is tedious and involves considerable algebraic manipulation and the lengthy result is not reproduced here. There are two singular points in the expression, the bound-state pole and the redundant pole. The integral over $\phi^*\phi$ is readily performed by taking the residue of each term possessing one or the other or both poles. The leading order in the asymptotic expansion of $\int\phi^*\phi\,dk$ consists of four terms: $\delta(r - r')$ and $\int S(k)dk$ from the leading contribution to $\phi_k(r)$ and two terms from the leading correction. The exact form of the Heisenberg condition for this problem is found to be

$$\begin{aligned}\int dk\,S(k)e^{ikR} = {}&\delta_{\alpha,-1}|C_1|^2 e^{-|k_1|R} + \delta_{\beta,1}|C_2|^2 e^{-|k_2|R}\\ &+ \pi\rho(\rho + \sigma)(\beta/\sigma)e^{-|k_1|R} + \pi\rho(\rho - \sigma)(\alpha/\sigma)e^{-|k_2|R},\end{aligned} \tag{4.9}$$

where $k_1 = \frac{1}{2}(\rho + \sigma)$, $k_2 = \frac{1}{2}(\rho - \sigma)$, and $S(k)$ is given by Eq.

(4.8).

The first two terms on the right-hand side of Eq. (4.9) are the contribution ot the Heisenberg condition from the sum over bound states. For this potential there is but one bound state; two terms are written, but of the two terms, only one or the other actually occurs. When $\alpha = -1$, $|C_1|^2\exp(-|k_1|R)$ is the contribution from the sum over bound states, and when $\beta = 1$ ($\alpha \neq -1$ necessarily) $|C_2|^2\exp(-|k_2|R)$ is this contribution. In order to include the two possibilities in one equation both possibilities are written and the Kronecker delta is used to eliminate the term not applicable to the given parameter values.

The second two terms on the right-hand side of Eq. (4.9) comprise the contribution to the Heisenberg condition coming from the leading correction terms to the dominant asymptotic form of $\phi_k(r)$. We have seen in previous cases that this contribution gives the sum of the residues at the redundant poles. As with the bound-state contribution both poles are represented since either pole can be the redundant pole, depending on the values of the parameters. We note that these two terms enter Eq. (4.9) in an essentially different way from that in which the bound-state terms enter. While only one or the other of the bound-state terms is actually present both of the second two terms of Eq. (4.9) are present. Since there is only one redundant pole, only one of the terms gives the residue of $S(k)\exp(ik(r+r'))$ at the redundant pole, the other term constitutes an anomalous contribution to the absolute squared value of the asymptotic amplitude of the bound-state wavefunction. The anomaly of this contribution comes by way of contrast with cases where the roles of bound-state pole and redundant pole are not interchangeable. In such cases, among which are the potentials of Secs. 2 and 3, the anomalous contribution does not occur and we have

$$2\pi i \operatorname{Res}\{S(\kappa)\exp(i\kappa R);\ \kappa = ik_l\} = |C_l|^2\exp(-|k_l|R),$$

where $\kappa = ik_l$ is one of the bound-state poles and C_l is the leading amplitude of the bound-state wave function in the asymptotic limit as given in Eq. (1.2). For potentials with the anomalous contribution this equation is modified so that $2\pi i \operatorname{Res}\{S(\kappa)\exp(ikR);\ \kappa = ik_l\} = (|C_l|^2 + \Delta) \times\exp(-|k_l|R)$, where Δ is the anomalous contribution.

For the potential just studied we have that

$$2\pi i \operatorname{Res}\{S(\kappa);\ \kappa = ik_l\} = -\pi\rho(\rho+\sigma)/\sigma,$$

$$\Delta = \pi\rho(\rho+\sigma)\beta/\sigma,$$

for $\alpha = -1$. Therefore, in the formula for the bound-state amplitude C_1,

$$|C_1|^2 = \pi\rho(\rho+\sigma)(\beta-1)/\sigma \quad (\alpha = -1), \tag{4.10}$$

the important β dependence is coming entirely from the anomalous contribution and not from the S-matrix. A similar statement with regard to α dependence holds for $\beta = 1$, where now $|C_2|^2$ is given by

$$|C_2|^2 = \pi\rho(\rho-\sigma)(\alpha+1)/\sigma \quad (\alpha = 1). \tag{4.11}$$

5. CONCLUSIONS

In this paper we have given further consideration to the matter of redundant poles showing that the existence of redundant poles in an S-matrix has implications on the nature of the continuum wavefunction. The modified Heisenberg condition has been the ideal vehicle for obtaining the details of these implications, for a careful derivation of the Heisenberg condition for a variety of potentials with exact solutions shows that the usual statement of this condition incorrectly ignores a set of terms giving information on the redundant poles. This set of terms comes precisely from the leading correction terms in the asymptotic expansion of the continuum wavefunction, while the leading terms give the S-matrix itself. Since the sum of the residues of the S-matrix at the redundant poles is proportional to the new terms arising from second order, a relationship between the terms of the leading order and the next lower order is thus implied.

ACKNOWLEDGMENTS

The outline of this study was proposed by Professor E. C. G. Sudarshan in his course lectures at Texas where he had conjectured all the results. I am grateful to him for suggesting this problem to me and for continued guidance. I thank Professor Chinn-Chann Chiang for a critical reading of the manuscript.

[1] S. T. Ma, Phys. Rev. **69**, 668 (1946); **71**, 195 (1947).
[2] D. Ter Haar, Physica **12**, 501 (1946).
[3] S. N. Biswas, T. Pradhan, and E. C. G. Sudarshan, Nucl. Phys. B **50**, 269 (1972).
[4] W. Heisenberg, Z. Naturforsch. **1**, 608 (1946).
[5] C. A. Nelson, A. K. Rajagopal, and C. S. Shastry, J. Math. Phys. **12**, 737 (1971).
[6] V. Bargmann, Rev. Mod. Phys. **21**, 488 (1949).
[7] A. Bhattacharjie and E. C. G. Sudarshan, Nuovo Cimento **25**, 864 (1962).

The quantum pendulum in the WKBJ approximation

R. N. Kesarwani
Department of Mathematics, University of Ottawa, Ottawa, Ontario K1N 9B4, Canada
Y. P. Varshni
Department of Physics, University of Ottawa, Ottawa, Ontario K1N 9B4, Canada

(Received 5 May 1981; accepted for publication 11 August 1981)

Eigenvalues are determined for the plane pendulum problem by the WKBJ method in one- and four-term approximations, and the results are compared. It is found that at high quantum numbers, the four-term WKBJ approximation can yield eigenvalues of eight-significant-figure accuracy, but for low quantum numbers the results continue to be poor.

PACS numbers: 03.65.Sq

I. INTRODUCTION

The problem of the plane pendulum, that is, of the motion of a masspoint constrained to move in a circle and acted on by a uniform field, was first discussed from the standpoint of quantum mechanics by Condon.[1] A more detailed discussion was given by Pradhan and Khare.[2] The Schrödinger equation for the plane pendulum can be recast[1,2] into a Mathieu equation, the characteristic values for which are known.[3,4] Khare[5] has solved the pendulum problem in the one-term WKBJ approximation and has shown that the eigenvalues are given by solutions of two equations which involve elliptic integrals and are thus different from the exact spectrum. Khare,[5] however, did not obtain any numerical results for the eigenvalues to show the degree of disagreement. Garbaczewski[6] has derived the conditions under which the quantum pendulum becomes equivalent to the elementary spin 1/2.

The improvement in the eigenvalues by the inclusion of higher-order terms in the WKBJ approximation has been examined for the Lennard-Jones potential[7–9] and the anharmonic oscillator.[10,11] The behavior of the potential-energy expression for the plane pendulum is very different from that of the two above-mentioned potentials. Also, above a certain energy, there are no turning points in the plane pendulum problem. Thus it was of special interest to examine the effect of higher-order terms in the WKBJ aproximation on the eigenvalues of the plane pendulum. In the present paper we have considered the four-term WKBJ approximation. In the course of this investigation an interesting phenomenon for the one-term WKBJ approximation was also discovered, that is, for certain combinations of the quantum number and a characteristic parameter of the pendulum, it is not possible to determine the eigenvalue. We discuss this unusual situation more fully in Sec. III.

II. THEORY

For a plane pendulum, the potential energy V due to the earth's gravitational field is

$$V(\theta) = mgl(1 - \cos\theta), \tag{1}$$

where m is the mass of the particle, l the radius of the circle in which it is constrained to move, and θ is the angle between the downward vertical and the radius vector of the particle measured from the center of the circle.

The Schrödinger equation for the plane pendulum can be written as

$$\frac{\hbar^2}{2ml^2}\frac{d^2\psi}{d\theta^2} + [\epsilon - mgl(1-\cos\theta)]\,\psi = 0\,. \tag{2}$$

If we set

$$k = [\epsilon/(2mgl)]^{1/2}$$

and

$$\beta = 16m^2gl^3/\hbar^2\,,$$

Eq. (2) reduces to

$$\frac{d^2\psi}{d\theta^2} + \frac{\beta}{4}[k^2 - \sin^2(\theta/2)]\,\psi = 0\,, \tag{3}$$

which is the Mathieu equation.[3]

We shall find it convenient to express the energy in reduced units, $\epsilon^* = \epsilon/(2mgl)$. It is known[2] that if $0 < \epsilon^* < 1$, the motion of the pendulum is oscillatory between the turning points $-\theta_0$ and θ_0, where $\sin(\theta_0/2) = k$; whereas if $1 < \epsilon^*$, there are no turning points and the motion of the pendulum is rotatory. For $\epsilon^* = 1$, the motion is nonperiodic and the pendulum swings up to the position $\theta = \pi$ and remains there forever.

Until now, the effect of higher-order terms in the WKBJ approximation on the eigenvalues has been examined for only such potentials as do have two turning points; the effect of such terms on the pendulum eigenvalues is thus of special interest.

A. Case of oscillatory motion, $0 < \epsilon^* < 1$

In the WKBJ approximation, the quantization condition up to four nonzero terms[9,12–15] is

$$n + \tfrac{1}{2} = I_1 + I_2 + I_3 + I_4\,, \tag{4}$$

where

$$I_1 = \frac{(2ml^2)^{1/2}/\hbar}{\pi}\int_{-\theta_0}^{\theta_0}(\epsilon - V)^{1/2}\,d\theta\,,$$

$$I_2 = -\frac{\hbar/(2ml^2)^{1/2}}{24\pi}\frac{d}{d\epsilon}\int_{-\theta_0}^{\theta_0}V''(\epsilon - V)^{-1/2}\,d\theta\,,$$

$$I_3 = \frac{[\hbar/(2ml^2)^{1/2}]^3}{2880\pi}\frac{d^3}{d\epsilon^3}\int_{-\theta_0}^{\theta_0}(7V''^2 - 5V'V''')\times(\epsilon - V)^{-1/2}\,d\theta\,,$$

 0022-2488/82/010092-04$02.50

$$I_4 = -\frac{[\hbar/(2ml^2)^{1/2}]^5}{725760\pi} \times\left[\frac{d^5}{d\epsilon^5}\int_{-\theta_0}^{\theta_0}(93V''^3 - 224V'V''V''' + 35V'^2V'''')\times(\epsilon - V)^{-1/2}\,d\theta + 216\frac{d^4}{d\epsilon^4}\int_{-\theta_0}^{\theta_0}V'''^2(\epsilon - V)^{-1/2}\,d\theta\right].$$

It is possible to evaluate the I's in a closed form in terms of elliptic integrals $K(k)$ and $E(k)$. The expressions for the I's can be written in a compact form in terms of $A(k,n)$ and $B(k,n)$ defined by

$$A(k,n) = (1-k^2)^{-n} - (k^2)^{-n} \qquad n = 1,2,\cdots$$

and

$$B(k,n) = (1-k^2)^{-n} + (k^2)^{-n} \qquad n = 1,2,\cdots.$$

The final expressions are as follows.

$$I_1 = \frac{2\beta^{1/2}}{\pi}[E(k) - (1-k^2)K(k)],$$

$$I_2 = \frac{\beta^{-1/2}}{12\pi}\{2K(k) + A(k,1)[E(k) - (1-k^2)K(k)]\},$$

$$I_3 = \frac{\beta^{-3/2}}{2880\pi}\{[28A(k,2) + 8A(k,1)]K(k) + [56B(k,3) + 23B(k,2) + 30B(k,1)]\times[(E(k) - (1-k^2)K(k)]\},$$

$$I_4 = \frac{\beta^{-5/2}}{161280\pi}\{[1984B(k,4) - 404B(k,3) + 190B(k,2) + 124B(k,1)]K(k) + [3968A(k,5) - 312A(k,4) + 589A(k,3) + 461A(k,2)]\times[E(k) - (1-k^2)K(k)]\}.$$

B. Case of rotatory motion, $1 < \epsilon^*$

In this case $\epsilon - V$ is always positive so that there are no turning points. The quantization condition takes the form[1,16]

$$n = J_1 + J_2 + J_3 + J_4, \tag{5}$$

where

$$J_1 = \frac{(2ml^2)^{1/2}/\hbar}{\pi}\int_0^{2\pi}(\epsilon - V)^{1/2}\,d\theta,$$

$$J_2 = -\frac{\hbar/(2ml^2)^{1/2}}{24\pi}\frac{d}{d\epsilon}\int_0^{2\pi}V''(\epsilon - V)^{-1/2}\,d\theta,$$

$$J_3 = \frac{[\hbar/(2ml^2)^{1/2}]^3}{2880\pi}\frac{d^3}{d\epsilon^3}\int_0^{2\pi}(7V''^2 - 5V'V''')\times(\epsilon - V)^{-1/2}\,d\theta,$$

$$J_4 = -\frac{[\hbar/(2ml^2)^{1/2}]^5}{725760\pi}\times\left[\frac{d^5}{d\epsilon^5}\int_0^{2\pi}(93V''^3 - 224V'V''V''' + 35V'^2V'''')\times(\epsilon - V)^{-1/2}\,d\theta + 216\frac{d^4}{d\epsilon^4}\int_0^{2\pi}V'''^2(\epsilon - V)^{-1/2}\,d\theta\right].$$

The J's can be evaluated in a closed form in terms of elliptic integrals. The final expressions are given below. For the sake of convenience, we write $\lambda = k^{-1} = (2mgl/\epsilon)^{1/2}$. Then

$$J_1 = \frac{2\beta^{1/2}}{\pi\lambda}E(\lambda),$$

$$J_2 = \frac{\lambda\beta^{-1/2}}{12\pi}\left\{\lambda^2K(\lambda) - \left(1 + \frac{1}{1-\lambda^2}\right)[E(\lambda) - (1-\lambda^2)K(\lambda)]\right\},$$

$$J_3 = \frac{\lambda\beta^{-3/2}}{2880\pi}\left\{\left[-56\lambda^6 + 5\lambda^4 - 15\lambda^2 - 43 + \frac{71}{1-\lambda^2} - \frac{28}{(1-\lambda^2)^2}\right]K(\lambda) + \left[56\lambda^4 + 5\lambda^2 + 30 - \frac{109}{1-\lambda^2} + \frac{135}{(1-\lambda^2)^2} - \frac{56}{(1-\lambda^2)^3}\right][E(\lambda) - (1-\lambda^2)K(\lambda)]\right\},$$

$$J_4 = \frac{\lambda\beta^{-5/2}}{161280\pi}\left\{\left[3968\lambda^{10} - 2296\lambda^8 + 497\lambda^6 + 62\lambda^4 - 337\lambda^2 - 2812 + \frac{9347}{1-\lambda^2} - \frac{12579}{(1-\lambda^2)^2} + \frac{8028}{(1-\lambda^2)^3} - \frac{1984}{(1-\lambda^2)^4}\right]K(\lambda) + \left[-3968\lambda^8 + 312\lambda^6 - 589\lambda^4 - 461\lambda^2 - \frac{4706}{1-\lambda^2} + \frac{16575}{(1-\lambda^2)^2} - \frac{23461}{(1-\lambda^2)^3} + \frac{15560}{(1-\lambda^2)^4} - \frac{3968}{(1-\lambda^2)^5}\right]\times[E(\lambda) - (1-\lambda^2)K(\lambda)]\right\}.$$

III. RESULTS AND DISCUSSION

The energy eigenvalues were calculated by solving (4) and (5). The elliptic integrals $K(k)$ and $E(k)$ can be expanded as series in powers of k^2. The number of terms which need be retained depends on the accuracy desired and on the eigenvalue. The root is determined by the Newton–Raphson method. The eigenvalues obtained from the one-term WKBJ approximation shall be denoted by $\epsilon^{*(1)}$ and those obtained from the four-term WKBJ approximation, by $\epsilon^{*(4)}$. The exact eigenvalues can be calculated from the characteristic values of the Mathieu equation[3,4] and shall be represented by ϵ_0^*.

TABLE I. Reduced energy eigenvalues for the potential (1). The symbols are explained in the text.

β	n	$\epsilon^{*(1)}$	$[\epsilon^{*(1)}-\epsilon_0^*]\times 10^7$	$\epsilon^{*(4)}$	$[\epsilon^{*(4)}-\epsilon_0^*]\times 10^7$	ϵ_0^*
2	0	0.6373109	1981937	0.5273797[a]	882625	0.4391172
	1	1.0690723	−1643111	1.2480202[a]	146368	1.2333834
	2	2.5157024	−347479	2.5216292	−288211	2.5504503
	3	5.0069512	−18523	5.0078601	−9434	5.0088035
	4	8.5039074	−2749	8.5041793	−30	8.5041823
	5	13.0025003	−1044	13.0026092	45	13.0026047
	6	18.5017362	−497	18.5017881	22	18.5017859
	7	25.0012756	−265	25.0013034	13	25.0013021
	8	32.5009766	−155	32.5009928	7	32.5009921
	9	41.0007716	−97	41.0007817	4	41.0007813
	10	50.5006250	−63	50.5006316	3	50.5006313
	11	61.0005165	−43	61.0005210	2	61.0005208
	12	72.5004340	−31	72.5004372	1	72.5004371
	13	85.0003698	−22	85.0003721	1	85.0003720
	14	98.5003189	−16	98.5003206	1	98.5003205
20	0	0.2171686	71709	0.2102475	2498	0.2099977
	1	0.6084539	155445	0.5986790	57696	0.5929094
	2	0.9235989	511434	0.8904405[a]	179850	0.8724555
	3	1.0295279	−479137	1.0522161[b]	−252255	1.0774416
	4	1.3403747	−144544	1.3446179	−102112	1.3548291
	5	1.7753244	−21742	1.7765226	−9760	1.7774986
	6	2.3174676	−5774	2.3179954	−496	2.3180450
	7	2.9627970	−2757	2.9630718	−9	2.9630727
	8	3.7097844	−1577	3.7099426	5	3.7099421
	9	4.5577253	−974	4.5578231	4	4.5578227
	10	5.5062549	−636	5.5063187	2	5.5063185
	11	6.5551680	−433	6.5552115	2	6.5552113
	12	7.7043419	−305	7.7043725	1	7.7043724
	13	8.9536992	−221	8.9537214	1	8.9537213
	14	10.3031894	−164	10.3032059	1	10.3032058
100	0	0.0987338	13016	0.0974322	0	0.0974322
	1	0.2882828	14318	0.2868512	2	0.2868510
	2	0.4664083	16299	0.4647845	61	0.4647784
	3	0.6316976	20568	0.6297531	1123	0.6296408
	4	0.7819512	38988	0.7793211	12687	0.7780524
	5	0.9130049	125030	0.9077444[a]	72425	0.9005019
	6			0.9897508[b]	−71	0.9897579
	7	1.0608818	−144651	1.0669397[a]	−84072	1.0753469
	8	1.1915580	−36827	1.1929874	−22533	1.1952407
	9	1.3498412	−9288	1.3505016	−2684	1.3507700
	10	1.5318977	−4043	1.5322795	−225	1.5323020
	11	1.7361850	−2451	1.7364287	−14	1.7364301
	12	1.9619116	−1653	1.9620769	0	1.9620769
	13	2.2086201	−1170	2.2087372	1	2.2087371
	14	2.4760262	−854	2.4761117	1	2.4761116

[a]Three-term WKBJ value.
[b]Two-term WKBJ value.

Calculations were carried out for three different values of β, namely, 2, 20, and 100, for $n = 0$ to $n = 14$. The results are shown in Table I. Columns 3 and 5 show the reduced energies obtained from the one- and four-term WKBJ approximations respectively. The exact eigenvalues are shown in column 7. The differences $\epsilon^{*(1)} - \epsilon_0^*$ and $\epsilon^{*(4)} - \epsilon_0^*$ are shown in columns 4 and 6, respectively. We may note here one point. The series expansion from which (4) is obtained is, in general, semiconvergent.[17,18] Consequently, situations can arise in which $|I_{j+1}/I_j|$ is greater than 1. In such a case it would be appropriate to take terms only up to and including I_j on the right-hand side of (4). Similar remarks are applicable to (5). In actual practice, calculations were carried out in stages for one-, two-, three- and four-term WKBJ approximations. A few cases were encountered for which $|I_3/I_2|$ or $|I_4/I_3|$ was greater than 1. Such cases are identified by superscripts a and b in Table I. All other eigenvalues are from the four-term WKBJ approximation.

It was found that when ϵ^* is very close to 1, it is not possible to calculate ϵ^* from the one-term WKBJ method, and for this reason there is no entry for $n = 6$, $\beta = 100$ for one-term WKBJ in Table I. The cause of this unusual situation can be understood by referring to Fig. 1, where we have plotted I_1 (for $k^2 < 1$) and J_1 (for $k^2 > 1$) versus k^2 for the case $\beta = 100$. An eigenvalue for the case $n = 6$ can be obtained from (4) only if at some stage I_1 becomes equal to 6.5, but we see in Fig. 1 that I_1 remains always below 6.5. Similarly, an eigenvalue for $n = 6$ can be obtained from (5) only if at some stage $J_1 = 6$, but the minimum value of J_1 is 6.367. We note

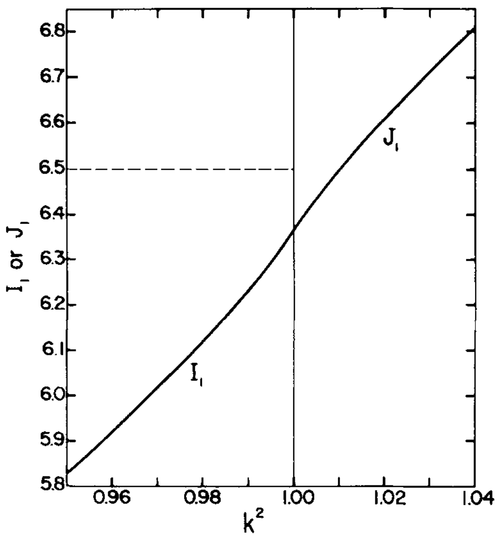


FIG. 1. I_1 and J_1 as functions of k^2 in the vicinity of $k^2 = 1$.

that $I_1(k^2 \to 1) = J_1(1/k^2 \to 1)$. For each value of n, there is a narrow range of β values where this phenomenon occurs. Some other combinations of such (n, β) values are as follows: (1, 3.6), (2, 12.2), (3, 26), (4, 44.5), and (5, 68). If higher-order terms are taken in (4) and in (5), this problem is eliminated, and eigenvalues can be calculated close to $\epsilon^* = 1$.

It will be noticed from Table I that the four-term WKBJ results show an improvement over the one-term WKBJ results in all cases, the accuracy being improved by between one and four additional significant figures. It is of some interest to compare these results with those obtained for the Lennard–Jones potential[7,9] where it was found that the four-term WKBJ results improve upon the one-term WKBJ results by seven or eight additional significant figures. Clearly the improvement obtained in the eigenvalues by taking higher-order terms in the WKBJ approximation is very much potential-dependent.

The results shown in Table I indicate that for a given β, errors appear to increase as ϵ^* increases toward 1 and then they decrease fairly rapidly. At high quantum numbers, four-term WKBJ results show an eight-significant-figure accuracy. A comparison of the one- and four-term results shows that the relative improvement is greater as β becomes larger. This result is somewhat surprising in view of the fact that the higher-order terms in (4) and (5) involve inverse powers of β. It will also be noticed in Table I that for a given quantum number, the eigenvalue decreases as β increases.

In conclusion we find that the inclusion of higher-order terms in applying the WKBJ method to the quantum pendulum problem can lead to eigenvalues of good accuracy (eight significant figures) at high quantum numbers, but the results at low quantum numbers remain poor.

ACKNOWLEDGMENT

The work was supported in part by a research grant from the Natural Sciences and Engineering Research Council of Canada to one of the authors (Y.P.V.).

[1] E. U. Condon, Phys. Rev. **31**, 891 (1928).
[2] T. Pradhan and A. Khare, Am. J. Phys. **41**, 59 (1973).
[3] N. W. McLachlan, *Theory and Application of Mathieu Functions* (Oxford, 1947).
[4] *Tables Relating to Mathieu Functions*, Natn. Bureau Stand. Appl. Math. Series No. 59 (U.S. Govt. Printing Office, Washington, DC, 1967).
[5] A. Khare, Lett. Nuovo Cimento **18**, 346 (1977).
[6] P. Garbaczewski, Phys. Lett. A **71**, 9 (1979).
[7] R. N. Kesarwani and Y. P. Varshni, Can. J. Phys. **56**, 1488 (1978).
[8] S. M. Kirschner and R. J. LeRoy, J. Chem. Phys. **68**, 3139 (1978).
[9] R. N. Kesarwani and Y.P. Varshni, Can. J. Phys. **58**, 363 (1980).
[10] P. O. Fröman, F. Karlsson, and M. Lakshmanan, Phys. Rev. D **20**, 3435 (1979).
[11] R. N. Kesarwani and Y. P. Varshni, J. Math. Phys. **22**, 1983 (1981), and the references given therein.
[12] J. L. Dunham, Phys. Rev. **41**, 713, 721 (1932).
[13] C. M. Bender, K. Olaussen, and P. S. Wang, Phys. Rev. D **16**, 1740 (1977).
[14] R. N. Kesarwani and Y. P. Varshni, J. Math. Phys. **21**, 90 (1980); **21**, 2852 (1980).
[15] F. T. Hioe, E. W. Montroll, and M. Yamawaki, in *Perspectives in Statistical Physics*, edited by H. J. Raveche (North-Holland, Amsterdam, 1981).
[16] H. Jeffreys, Proc. London Math. Soc. **23**, 437 (1925).
[17] G. D. Birkhoff, Bull. Am. Math. Soc. **39**, 696 (1933).
[18] E. C. Kemble, *The Fundamental Principles of Quantum Mechanics* (Dover, New York, 1958); this is a reprint of the 1937 edition published by McGraw-Hill.

A new mathematical formulation of accelerated observers in general relativity. I

D. G. Retzloff
Department of Chemical Engineering, University of Missouri-Columbia, Columbia, Missouri 65211
B. DeFacio
Department of Physics, University of Missouri-Columbia, Columbia, Missouri 65211 and Applied Mathematical Sciences, Ames Laboratory USDOE, and Iowa State University, Ames, Iowa 50011
P. W. Dennis
BDM Corporation, Redondo Beach, California 90278

(Received 25 August 1980; accepted for publication 14 November 1980)

Invariant methods of modern differential geometry are used to formulate exact closed form expressions for the coordinate velocity and coordinate acceleration of a geodesic particle in the tangent space of a general relativistic accelerating rotating observer. The observation of a general vector field is shown to be definable in two ways from presymmetry and covariance arguments. Our results for the parallel translation definition of observation are shown to subsume existing work in both special and general relativity on accelerated observers.

PACS numbers: 04.20.Cv

I. INTRODUCTION

Synge[1] considered observations by a general relativistic accelerated observer in his discussion of stellar aberration. By using Fermi coordinates[2] he obtains a first order equation for the stellar aberration. A natural proper reference frame for a general relativistic accelerating rotating observer (GRARO) was introduced by Misner, Thorne, and Wheeler (MTW)[3] who suitably extended the Fermi normal coordinates discussed by Manasse and Misner[4] to obtain a local set of coordinates for describing the GRARO. They obtained the first-order expansion for the metric and the connection coefficients along the GRARO world line which give rise to both the rotation and the acceleration of the observer. Burghardt[5] utilized a covariant projection formalism to obtain a decomposition of the Einstein equations of motion for a rotating system of observers. Using a dyadic formalism Estabrook and Wahlquist[6] obtained an equation for the acceleration near a general world line. Within the framework of special relativity Li and Ni[7] employed the MTW coordinate system to obtain an expansion of the metric for an accelerated observer. They also have done extensive work on determining the metric expansions for an accelerating observer in both special and general relativity. At about the same time DeFacio, Dennis, and Retzloff,[8] using a coordinate-free approach, derived an exact closed form expression for the coordinate acceleration of a geodesic particle relative to a noninertial observer which they showed reduced to the second-order expansion of Ni and Zimmerman[9] for the case of special relativity.

In the setting of general relativity Mashhoon[10] in his extensive treatment of tidal processes derived a second-order expression for the metric and a first-order expansion of the connection coefficients in Fermi Coordinates for a general relativistic nonrotating accelerating observer. Ni and Zimmerman,[9] using the MTW coordinates, extended these results to a GRARO to obtain a second-order expansion of both the coordinate velocity and coordinate acceleration of a geodesic particle in the frame of the GRARO. These latter results were extended to third order by Li and Ni,[11] who in the process derived a third-order metric expansion and a second-order expansion of the connection coefficients. Coupling effects between gravitation and special relativity were first exhibited in the third-order terms obtained by Li and Ni.

The general relativity calculations thus far have been formulated in terms of various coordinate systems eminently suited for computational purposes. The MTW coordinate frame is the most natural GRARO frame for obtaining expansions of the coordinate velocity and coordinate acceleration of a freely falling particle in the GRARO frame. However, the coordinate velocity and coordinate acceleration have not been defined to date in a coordinate-free invariant manner and the coordinate velocity is currently determined from the coordinate acceleration by integration in contrast to the pedagogical approach of defining the acceleration as a derivative of the velocity. Furthermore, in the coordinate formulation it is not abundantly clear where the domains and ranges of the associated vector field mappings reside on their respective manifolds. This is of fundamental importance in general relativity where vectors associated with particular observations by the GRARO must live in the tangent space of the GRARO.[8]

In this paper we will formulate as coordinate-free invariant objects in the tangent space of the GRARO the coordinate velocity and coordinate acceleration of a freely falling particle employing the approach of DeFacio, Dennis, and Retzloff[8] extended to general relativity. The coordinate acceleration will be identified as the derivative of the coordinate velocity. Hence the coordinate velocity will be obtained independent of the coordinate acceleration. The manifold defined by the GRARO and the geodesic particle will be explicitly identified. Using these results and the definitions of "presymmetry", "spacetime," and "observations of spacetime" as given by DeFacio, Dennis, and Retzloff,[8] closed form expressions for both the coordinate velocity and coordinate acceleration will be obtained. These expressions will be shown to agree with the corrected results of Li and Ni[11] to the order of their expansion.

The organization of this paper is as follows. In Sec. II we define in a coordinate-free manner the observation of a geodesic particle by a GRARO. Section III contains the derivation of the closed form expressions for the coordinate velocity and coordinate acceleration. In Sec. IV the connection coefficients on the manifold defined by the GRARO and the geodesic particle are obtained and our results are shown to subsume the third-order expansion of Li and Ni[11] as well as the exact results of DeFacio, Dennis, and Retzloff[8] for special relativity. Our conclusions are given in Sec. V.

II. GRARO OBSERVATIONS OF A GEODESIC PARTICLE

In the notation of DeFacio, Dennis, and Retzloff[8] the basic description of the GRARO observation of a geodesic particle is depicted in Fig. 1. The manifold M determined by this description is given by the map

$$\alpha : R \times R \rightarrow M, \quad \alpha(t,s) = \exp_{\gamma(t)} s\mathbf{r}(t). \tag{1}$$

From this map α, the world line and velocity of the GRARO $(\gamma(t), u)$ and the geodesic particle $(\gamma_1(t), V)$ are easily found to be

$$\begin{aligned} &\gamma(t) = \alpha(t,0), \quad u = \alpha_*(\partial/\partial t)_{|s=0,t=t} = \nabla_{d/dt}\gamma(t), \\ &\gamma_1(\lambda) = \alpha(t,1), \\ &V = \alpha_*(\partial/\partial t)_{|s=1,t=t} = \nabla_{d/d\lambda}, \gamma_1(\lambda) = \Gamma \nabla_{d/dt}\gamma_1(t), \end{aligned} \tag{2}$$

where $\Gamma = dt/d\lambda$ and ∇ is the usual pseudo-Riemannian connection of general relativity on the manifold.

In general presymmetry[8] it has been shown that for a general spacetime (M, g) any event q is identified as $q \in M$. The observation of this event $X \in TM$ is found by lifting M into TM with the map

$$\underset{M}{\overset{TM}{\uparrow}} \exp^{-1}, \quad \text{i.e.,} \quad X_p = \exp_p^{-1} q. \tag{3}$$

Because the tangent space of the GRARO is flat, the "natural" connection for observations in the GRARO frame is the pseudoflat connection ∇' defined by

$$\begin{aligned} &\nabla'_{d/dt} B = \nabla_{d/dt} B - \langle \Omega, B \rangle, \\ &\Omega = \mathbf{a} \otimes u - u \otimes \mathbf{a} + \Omega_s(\omega, u), \\ &\langle \Omega_s(\omega,u), u \rangle = 0, \\ &\langle \Omega_s(\omega,u), B \rangle = \omega \times B, \end{aligned} \tag{4}$$

where $\mathbf{a}$ is the acceleration of the GRARO. The coordinate velocity W and coordinate acceleration A obtain their simplest forms when expressed in terms of this natural GRARO connection as

$$\begin{aligned} &W = \nabla'_{d/dt} \alpha_* (\partial/\partial s)|_{s=0,t=t}, \\ &A = \nabla'_{d/dt} W = \nabla'_{d/dt} \nabla'_{d/dt} \alpha_* (\partial/\partial s)|_{s=0,t=t}. \end{aligned} \tag{5}$$

Clearly by definition $W, A \in T_{\gamma(t)} M$ and hence "live" in the frame of the GRARO.

To conclude our discussion of GRARO observations we give the following definitions of the observation of a general spacetime event by a GRARO which arises from general presymmetry consideration.[8,12,13]

Definition 1: For an arbitrary event $q \in M$ and observation of that event $X_q \in T_q M$ by an observer at q, the observation of event q by a GRARO at p is determined by the commuting diagram

$$\begin{tikzcd} T_v T_p M \arrow[rr, "\exp_{p_{*v}}"] & & T_q M \arrow[dd, "\exp_q"] \\ v \in T_p M \arrow[u, "|J_v^{-1}"] \arrow[urr, "\exp_{p_{*v}} \circ J_v"'] \arrow[d, "\exp_p"] & & \\ M \arrow[rr, "\exp_p v"'] & & M \end{tikzcd},$$

with J_v being the usual projection map. The map $\exp_{p_*} J$ preserves vector equations as well as equivalence classes of state preparation and observation procedures. Thus it satisfies the covariance and presymmetry requirements for defining the observation of a spacetime event. However, the map is only a radial isometry. The consequences of this definition of observation are presented in Part II. A second definition of the observation of an event that also satisfies the requirements of covariance and presymmetry is the following:

Definition 2: For an arbitrary event $q \in M$ and observation of that event $X_q \in T_q M$ by an observer at q, the observation of event q by a GRARO at p is determined by the parallel translation map τ_{qp} with

$$\tau_{qp} : T_q M \rightarrow T_p M.$$

For our purpose q is given by

$$q = \gamma_2(s_0(t), t) = \exp_{\gamma(t)} \mathbf{r}(t). \tag{6}$$

The map τ is an isometry and hence preserves the magnitude of the observed vector field. A direct application of Definition 2 is the calculation of the velocity $\tilde{V}$ of the geodesic particle as seen in the GRARO frame. The result will be used in this paper to obtain closed form expressions for W and A

An alternative method of formulating the observation of a particle by a rotating accelerating observer based on the

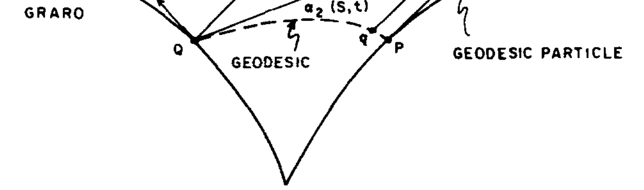


FIG. 1. The manifold defined by the GRARO and the geodesic particle. The symbols have the following definitions; $\gamma(t)$—world line of the GRARO; $\gamma_1(\lambda) = \gamma_1(t)$—world line of the geodesic particle; λ—affine parameter of the geodesic world line; t—nonaffine parameter; $\gamma_2(s,t)$—unique geodesic passing through q and p, s—affine parameter of γ_2; $\mathbf{r}(t) = n(t)|r(t)|$—position of geodesic particle relative to the GRARO; $n(t)$—direction cosines of $\mathbf{r}(t)$ in GRARO frames; u—velocity of GRARO; V—velocity of geodesic particle; $\tilde{V}$—parallel translation of $v \in T_{\gamma_1(\lambda)} M$ to $T_{\gamma(t)} M$ along unique geodesic $\gamma_2(s,t)$; $\hat{V}(t,s)$—parallel translation of $V \in T_{\gamma_1(\lambda)} M$ to $T_q M$ along $\gamma_2(t,s)$.

cotangent bundle of the manifold has been developed by DeFacio and Retzloff[14] in their treatment of noninertial frames in special relativity. Their approach has the advantge of providing a natural setting in which the velocity vector field of the geodesic particle is a Killing vector field and "cotangent geodesics" are described by the vanishing of the Lie derivative of the associated velocity form. This method is readily extended to a general relativistic setting and will be the subject of a future paper on this problem.

We now consider the results of defining an observation of a spacetime event by a GRARO via Definition 2.

III. THE COORDINATE VELOCITY AND COORDINATE ACCELERATION

From Definition 2 we know that $\tilde{V}$ is the solution at $s=0$ of

$$\nabla_{d/ds}\hat{V}=0\langle=\rangle\ \frac{d\hat{V}^\mu}{ds}=-\Gamma^\mu{}_{\alpha\beta}(t,s)\hat{V}^\beta\frac{d\gamma_2^\alpha}{ds},$$

$$\mu,\alpha,\beta=0,1,2,3,\quad \hat{V}=V \text{ at } s=s_0(t). \tag{7}$$

To proceed further, we use the MTW coordinate system to give M a coordinate chart and note that

$$\begin{aligned}
&\gamma(t)=\{t,0,0,0\},\quad \gamma_1(t)=\exp_{\gamma(t)}\mathbf{r}(t)=\{t,r^1(t),r^2(t),r^3(t)\},\\
&\gamma_2(t,s)=\exp_{\gamma(t)}sn(t)=\{t,n^1(t)s,n^2(t)s,n^3(t)s\},\\
&\mathbf{r}(t)=r^i(t)e_{i|_{\gamma(t)}}=s_0(t)n^i(t)e_{i|_{\gamma(t)}},\quad i=1,2,3,\\
&n(t)=n^i(t)e_{i|_{\gamma(t)}},\quad i=1,2,3,\quad \langle n(t),n(t)\rangle=1,\\
&W=W(t)=W^i(t)e_{i|_{\gamma(t)}},\quad i=1,2,3,\\
&V=\nabla_{d/d\lambda}\exp_{\gamma(t)}\mathbf{r}(t)=\Gamma\nabla_{d/dt}\exp_{\gamma(t)}\mathbf{r}(t)\\
&\quad=\Gamma e_{0|_{\gamma_1(t)}}+\Gamma W^i(t)e_{i|_{\gamma_1(t)}},\quad i=1,2,3,\\
&\quad=\Gamma\{1,W^1(t),W^2(t),W^3(t)\}|_{\gamma_1(t)}.
\end{aligned} \tag{8}$$

Because (7) is a linear matrix differential equation its solution can be written in terms of the matrixant $\Omega(s,t)$ as[15]

$$\hat{V}=\Omega(s,t)\tilde{V}\quad\text{or}\quad V=\Omega(s_0(t),t)\tilde{V}, \tag{9}$$

where

$$\hat{V}=\begin{pmatrix}\hat{V}^0\\ \hat{V}^1\\ \hat{V}^2\\ \hat{V}^3\end{pmatrix},\quad \tilde{V}=\begin{pmatrix}\tilde{V}^0\\ \tilde{V}^1\\ \tilde{V}^2\\ \tilde{V}^3\end{pmatrix},\quad \mathrm{V}=\Gamma\begin{pmatrix}1\\ W^1(t)\\ W^2(t)\\ W^3(t)\end{pmatrix},$$

$$\begin{aligned}
\Omega(s,t)&=I+\sum_{n=1}^{\infty}\prod_{j=0}^{n}\int_0^{s_j}B(s_j,t)ds_j,\quad s_0=s,\\
&=P\exp\left[\int_0^s B(s_1,t)ds_1\right],
\end{aligned} \tag{10}$$

$$B(s,t)=(b_{\alpha\beta}),\quad b_{\alpha\beta}=-\Gamma^\alpha{}_{j\beta}(s,t)n^j(t),$$

$$\alpha,\beta=0,1,2,3,\quad j=1,2,3,$$

and P is the Dyson chronological-order operator for the s_i. Substituting (10) into (9), we obtain an expression for the coordinate velocity of the form

$$\overline{W}\equiv\begin{pmatrix}1\\ W^1(t)\\ W^2(t)\\ W^3(t)\end{pmatrix}_{|_{\gamma(t)}}=\Gamma^{-1}\Omega(s_0(t),t)\begin{pmatrix}\tilde{V}^0\\ \tilde{V}^1\\ \tilde{V}^2\\ \tilde{V}^3\end{pmatrix}_{|_{\gamma(t)}}, \tag{11}$$

which in component form is

$$W^i(t)\tilde{e}_{i|_{\gamma(t)}}=\Gamma^{-1}\Omega^i{}_\alpha(s_0(t),t)\tilde{V}^\alpha\tilde{e}_{i|_{\gamma(t)}}, \tag{12}$$

$$\tilde{V}=\tilde{V}^\alpha e_{\alpha|_{\gamma(t)}},\quad i=1,2,3,\quad \alpha=0,1,2,3,$$

$$\tilde{V}=\tilde{V}^\alpha e_{\alpha|_{\gamma(t)}},\quad i=1,2,3,\quad \alpha=0,1,2,3,$$

where $\tilde{e}_{i|_{\gamma(t)}}$ are the basis vectors obtained from $e_{i|_{\gamma(t)}}$ by parallel transport. The physical interpretation of (12) [or (11)] is that W^i represent the components of the coordinate velocity for the basis vectors $\tilde{e}_{i|_{\gamma(t)}}$ while the geodesic particle velocity $\tilde{V}$ is given in the natural basis vectors $e_{i|_{\gamma(t)}}=\partial/\partial x^i|_{\gamma(t)}$, $x^i=\{t,r^1,r^2,r^3\}$. Thus (12) represents the transformation law from the tilde set of basis vectors at $\gamma(t)$ to the natural set of basis vectors at $\gamma(t)$. By writing (12) as a vector equation we can obtain the coordinate velocity in terms of the natural basis vectors at $\gamma(t)$.

The coordinate acceleration is obtained from $\overline{W}$ as

$$\begin{aligned}
A=\nabla'_{d/dt}\overline{W}=\rangle&\begin{pmatrix}0\\ A^1\\ A^2\\ A^3\end{pmatrix}_{|_{\gamma(t)}}\\
&=\Gamma^{-1}\left(\frac{d\Omega(s_0(t),t)}{dt}-\Gamma^{-1}\Omega(s_0(t),t)\frac{d\Gamma}{dt}\right)\begin{pmatrix}\tilde{V}^0\\ \tilde{V}^1\\ \tilde{V}^2\\ \tilde{V}^3\end{pmatrix}_{|_{\gamma(t)}}\\
&\quad+\Gamma^{-1}\Omega(s_0(t),t)\nabla'_{d/dt}\begin{pmatrix}\tilde{V}^0\\ \tilde{V}^1\\ \tilde{V}^2\\ \tilde{V}^3\end{pmatrix}_{|_{\gamma(t)}}.
\end{aligned} \tag{13}$$

We must express $\nabla'_{d/dt}\tilde{V}$ in terms of $\overline{W}$ and $\Omega(s_0(t),t)$ to complete the calculation of the coordinate acceleration. To accomplish this we consider the equation of the geodesic particle for the nonaffine parameter t which is

$$\begin{aligned}
&\nabla_{d/dt}[\Gamma\nabla_{d/dt}\gamma_1(t)]=0=\rangle\nabla_{d/dt}\nabla_{d/dt}\gamma_1(t)\\
&\quad=-\Gamma^{-1}\frac{d\Gamma}{dt}\nabla_{d/dt}\gamma_1(t).
\end{aligned} \tag{14}$$

In component form this reduces to

$$\frac{dV^\mu}{dt}=-\Gamma^\mu{}_{\alpha\beta}(t,s_0(t))\Gamma\overline{W}^\alpha\overline{W}^\beta,\quad V=\Gamma\nabla_{d/dt}\gamma_1(t). \tag{15}$$

From (9) we have

$$\begin{aligned}
\nabla'_{d/dt}\tilde{V}&=\nabla'_{d/dt}[\Omega^{-1}(s_0(t),t)V]\\
&=\frac{d\Omega^{-1}(s_0(t),t)}{dt}\Gamma\overline{W}+\Omega^{-1}(s_0(t),t)D,
\end{aligned} \tag{16}$$

where

$$D=-\Gamma\begin{pmatrix}\Gamma^0{}_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta\\ \Gamma^1{}_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta\\ \Gamma^2{}_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta\\ \Gamma^3{}_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta\end{pmatrix}_{|_{\gamma(t)}}. \tag{17}$$

Combining (11), (13), (16), and (17), we obtain the desired expression for the coordinate acceleration

$$\begin{pmatrix}0\\ A^1\\ A^2\\ A^3\end{pmatrix}_{|_{\gamma(t)}}=\Gamma^{-1}\left(-\frac{d\Gamma}{dt}\overline{W}+D\right)_{|_{\gamma(t)}}. \tag{18}$$

The remarks concerning the natural and parallel translated basis vectors at $\gamma(t)$ made previously also apply to (18).

Next we will show that (11) and (18) include the results of DeFacio, Dennis, and Retzloff,[8] and Li and Ni[11] as special cases.

IV. THE SPECIFIC CALCULATION OF A AND W USING THE CONNECTION COEFFICIENTS ON M

Although (11) and (18) are exact closed form formulas for the coordinate velocity and coordinate acceleration of a geodesic particle, these quantities are normally expressed in terms of the curvature tensor, 4-rotation and acceleration of the GRARO.[7–11] To do this we must write the connection coefficients Γ and $d\Gamma/dt$ in terms of these latter quantities. The connection coefficients on M are obtained in terms of the appropriate quantities from a Taylor series expansion of the $\Gamma^{\mu}{}_{\alpha\beta}$ about the world line of the GRARO where they are known from the work of MTW. Thus on M we have

$$\begin{aligned}\Gamma^{\mu}{}_{\alpha\beta}(t,s) = \Gamma^{\mu}{}_{\alpha\beta}(t,0) + \Gamma^{\mu}{}_{\alpha\beta,l}(t,0)n^{l}(t)s \\ + \Gamma^{\mu}{}_{\alpha\beta,li}(t,0)n^{l}(t)n^{i}(t)s^{2}/2! + \cdots,\end{aligned} \tag{19}$$

$\mu,\alpha,\beta = 0,1,2,3, \quad l,i, \text{ etc.} = 1,2,3.$

Equations (15) and (22) require the $\Gamma^{\mu}{}_{\alpha\beta}$ along $\gamma_1(t)$ which from (12) and (23) are

$$\begin{aligned}\Gamma^{\mu}{}_{\alpha\beta}(t,s_0(t)) = \Gamma^{\mu}{}_{\alpha\beta}(t,0) + \Gamma^{\mu}{}_{\alpha\beta,l}(t,0)r^{l}(t) \\ + \Gamma^{\mu}{}_{\alpha\beta,li}(t,0)r^{l}(t)r^{i}(t)/2! + \cdots.\end{aligned} \tag{20}$$

The derivatives of $\Gamma^{\mu}{}_{\alpha\beta}$ along $\gamma(t)$ occurring in (19) and (20) are obtained by solving the Jacobi field equations associated with the following two separate one-parameter family of geodesics:

$$\alpha(s,n^{i}) = \exp_{\gamma(t)} sn^{i}e_{i}, \quad t \text{ fixed}, \tag{21}$$

$$\beta(t,s) = \exp_{\gamma(t)} sn^{i}(t)e_{i}. \tag{22}$$

The one-parameter family of geodesics given by (21) is depicted in Fig. 2 and is identical to the reference frame of the GRARO described by MTW.[3] From (21) we have

$$\begin{aligned}&\alpha_{*}(\partial/\partial s) = n^{i}e_{i} \equiv n, \quad \nabla_{n}n = 0,\\ &\alpha_{*}(\partial/\partial n^{i}) = se_{i} \equiv \mu^{i}e_{i}, \quad i = 1,2,3 \ (i \text{ not summed}),\\ &[n,\mu] = \alpha_{*}[\partial/\partial s, \ \partial/\partial n] = 0,\end{aligned} \tag{23}$$

where μ is the Jacobi field and n is the tangent vector field. From the definition of the Riemannian curvature tensor we obtain

$$R(n,\mu)\mu = \nabla_{n}\nabla_{\mu}\mu - \nabla_{\mu}\nabla_{n}\mu - \nabla_{[n,\mu]}\mu = 0, \tag{24}$$

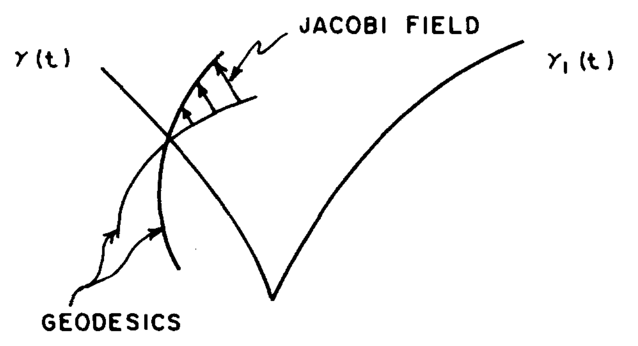


FIG. 2. The l-parameter family of geodesics determined by $\alpha(s,n^{i}) = \exp_{\gamma(t)} sn^{i}e_{i}$.

which combined with (23) gives the Jacobi field equation

$$\nabla_{n}\nabla_{n}\mu + R(n,\mu)\mu = 0. \tag{25}$$

In component form (25) is

$$\begin{aligned}\frac{d^{2}\mu^{\alpha}}{ds^{2}} + 2\frac{d\mu^{\beta}}{ds}\Gamma^{\alpha}{}_{\beta\xi}(s,t)n^{\xi} + R^{\alpha}{}_{\beta\xi\sigma}\mu^{\beta}n^{\xi}n^{\sigma} \\ + \mu^{\beta}n^{\xi}n^{\sigma}(\Gamma^{\alpha}{}_{\beta\xi,\sigma} + \Gamma^{\tau}{}_{\beta\xi}\Gamma^{\alpha}{}_{\tau\sigma} - \Gamma^{\alpha}{}_{\beta\tau}\Gamma^{\tau}{}_{\xi\sigma}) = 0.\end{aligned} \tag{26}$$

This is Eq. (10) of Ni and Zimmerman[9] and Eq. (14) of Li and Ni.[11] Substituting (19) and (23) into (26), we obtain

$$\begin{aligned}&2\sum_{m=0}^{\infty}\frac{s^{m}}{m!}[(n^{l}(t)D_{l})^{m}\Gamma^{\alpha}{}_{jk}](t,0)n^{k}(t) + R^{\alpha}{}_{jkp}sn^{k}(t)n^{p}(t)\\ &+ sn^{p}(t)n^{k}(t)\Bigg[\sum_{m=0}^{\infty}\frac{s^{m}}{m!}[(n^{l}(t)D_{l})^{m}\Gamma^{\alpha}{}_{jp,k}](t,0)\\ &+ \sum_{m=0}^{\infty}\frac{s^{m}}{m!}[(n^{l}(t)D_{l})^{m}\Gamma^{\tau}{}_{jp}](t,0)\sum_{q=0}^{\infty}\frac{s^{q}}{q!}[(n^{i}(t)D_{i})^{q}\Gamma^{\alpha}{}_{\tau k}](t,0)\\ &- \sum_{m=0}^{\infty}\frac{s^{m}}{m!}[(n^{l}(t)D_{l})^{m}\Gamma^{\alpha}{}_{j\tau}](t,0)\sum_{q=0}^{\infty}\frac{s^{q}}{q!}[(n^{i}(t)D_{i})^{q}\Gamma^{\tau}{}_{pk}](t,0)\Bigg]\\ &= 0,\\ &\tau,\alpha = 0,1,2,3, \quad i,l,m,j,k,p = 1,2,3,\\ &D_{i} = \frac{\partial}{\partial x^{i}}, \quad x^{i} = sn^{i}(t).\end{aligned} \tag{27}$$

Equating powers of s through s^{2} in (27) yields the results of Ni and Zimmerman[9] as well as those of Li and Ni for the first and second derivatives of $\Gamma^{\alpha}{}_{ij}(i,j = 1,2,3)$. Higher derivatives of $\Gamma^{\alpha}{}_{ij}$ are found by continuing the process. Because only the first and second derivatives of $\Gamma^{\alpha}{}_{\beta\gamma}$ are needed to compare our general relativity formulation with the work of Li and Ni[11] and the computation of the higher derivatives of the connection coefficients provide no new insight into the basic physics of the problem, we will not compute these higher derivatives in this paper. We also refer to the work of Li and Ni[11] for the tabulated formulas for the first and second derivatives of $\Gamma^{\alpha}{}_{ij}$. However, in the case of special relativity we require the derivatives to be determined to all orders in order to compare with the exact result of DeFacio, Dennis, and Retzloff.[8] The results for special relativity are obtained from the above computational process by setting $R^{\alpha}{}_{\beta\gamma\sigma} = 0$. This gives the values in Table I for $\Gamma^{\alpha}{}_{ij}$. The remaining connection coefficients listed in Table I are found using (22) which defines the one-parameter family of geodesics shown in Fig. 3. Following a procedure identical to that of (23) and (24), we obtain the Jacobi field equation

$$\begin{aligned}&\nabla_{U}\nabla_{U}N + R(U,N)N = 0,\\ &U = \beta_{*}(\partial/\partial s) = n^{i}(t)e_{i},\\ &N = \beta_{*}(\partial/\partial t) = e_{0},\end{aligned} \tag{28}$$

which reduces to the identity

$$R^{\mu}{}_{\alpha0\beta} = \Gamma^{\mu}{}_{\alpha\beta,0} - \Gamma^{\mu}{}_{\alpha0,\beta} + \Gamma^{\xi}{}_{\alpha\beta}\Gamma^{\mu}{}_{\xi0} - \Gamma^{\xi}{}_{\alpha0}\Gamma^{\mu}{}_{\xi\beta}. \tag{29}$$

Combining (29) with differentiation of the known connection coefficients along $\gamma(t)$ yields the remaining derivatives of the $\Gamma^{\alpha}{}_{\beta\sigma}$[9,11] which agree with the second order results of Li and Ni[11] in general relativity and produce the remaining entries in Table I in the case of special relativity.

TABLE I. Connection coefficients on the M in special relativity $\Gamma^{\alpha}{}_{\beta\sigma}(t,s)$, $\sigma = 0$, $\Gamma^{\alpha}{}_{ij}(s,t) = 0, \alpha = 0,1,2,3,\ i,j = 1,2,3.$

$\alpha \backslash \beta$	0	1
0	$\dfrac{a\cdot sn(t) + a\cdot\omega\times sn(t)}{1 + a\cdot sn(t)}$	$\dfrac{a_1}{1 + a\cdot sn(t)}$
1	$a^1(1 + a\cdot sn(t)) + [\omega\times(\omega\times sn(t))]^1 + [\dot\omega\times sn(t)]^1 - \dfrac{[\omega\times sn(t)]^1(\dot a\cdot sn(t) + a\cdot\omega\times sn(t))}{1 + a\cdot sn(t)}$	$a_1(\omega\times n(t))^1 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$
2	$a^2(1 + a\cdot sn(t)) + [\omega\times(\omega\times sn(t))]^2 + [\dot\omega\times sn(t)]^2 - \dfrac{[\omega\times sn(t)]^2(\dot a\cdot sn(t) + a\cdot\omega\times sn(t))}{1 + a\cdot sn(t)}$	$-\omega^k\epsilon^2_{k1} + a_1(\omega\times n(t))^2 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$
3	$a^3(1 + a\cdot sn(t)) + [\omega\times(\omega\times sn(t))]^3 + [\dot\omega\times sn(t)]^3 - \dfrac{[\omega\times sn(t)]^3(\dot a\cdot sn(t) + a\cdot\omega\times sn(t))}{1 + a\cdot sn(t)}$	$-\omega^k\epsilon^3_{k1} + a_1(\omega\times n(t))^3 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$

$\alpha \backslash \beta$	2	3
0	$\dfrac{a_2}{1 + a\cdot sn(t)}$	$\dfrac{a_3}{1 + a\cdot sn(t)}$
1	$-\omega^k\epsilon^1_{k2} + a_2(\omega\times n(t))^1 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$	$-\omega^k\epsilon^1_{k3} + a_3(\omega\times n(t))^1 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$
2	$a_2(\omega\times n(t))^2 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$	$-\omega^k\epsilon^2_{k3} + a_3(\omega\times n(t))^2 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$
3	$-\omega^k\epsilon^3_{k2} + a_2(\omega\times n(t))^3 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$	$a_3(\omega\times n(t))^3 \sum_{m=1}^{\infty}(-a\cdot n(t))^{m-1}s^m$

We are now ready to express the coordinate velocity and coordinate acceleration in terms of the curvature tensor, the 4-rotation and the acceleration of the observer. The case of special relativity will be treated first. Using Table I the B matrix in (10) becomes

$$B(s,t) = \begin{pmatrix} -a\cdot n(t)/[1 + a\cdot sn(t)] & 0 & 0 & 0 \\ (\omega\times n(t))^1/[1 + a\cdot sn(t)] & 0 & 0 & 0 \\ (\omega\times n(t))^2/[1 + a\cdot sn(t)] & 0 & 0 & 0 \\ (\omega\times n(t))^3/[1 + a\cdot sn(t)] & 0 & 0 & 0 \end{pmatrix}. \tag{30}$$

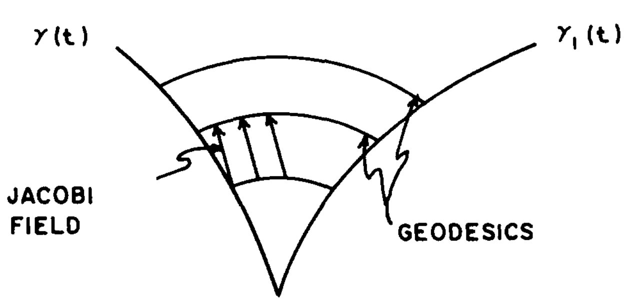


FIG. 3. The l-parameter family of geodesics given by $\beta(t,s) = \exp_{\gamma(t)} sn^i(t)e_i$.

From (30) and (10) the matrizant is calculted as

$$\Omega(s_0(t),t) = \begin{pmatrix} 1/(1 + a\cdot r) & 0 & 0 & 0 \\ -(\omega\times r)^1/(1 + a\cdot r) & 1 & 0 & 0 \\ -(\omega\times r)^2/(1 + a\cdot r) & 0 & 1 & 0 \\ -(\omega\times r)^3/(1 + a\cdot r) & 0 & 0 & 1 \end{pmatrix}. \tag{31}$$

Combining (11) and (31), we obtain the following result for the velocity of the geodesic particle as seen by the observer:

$$\begin{pmatrix} \tilde V^0 \\ \tilde V^1 \\ \tilde V^2 \\ \tilde V^3 \end{pmatrix} = \Omega^{-1}(s_0(t),t) \begin{pmatrix} \Gamma \\ \Gamma W^1(t) \\ \Gamma W^2(t) \\ \Gamma W^3(t) \end{pmatrix} = \begin{pmatrix} 1 + a\cdot r & 0 & 0 & 0 \\ (\omega\times r)^1 & 1 & 0 & 0 \\ (\omega\times r)^2 & 0 & 1 & 0 \\ (\omega\times r)^3 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Gamma \\ \Gamma\dot r^1(t) \\ \Gamma\dot r^2(t) \\ \Gamma\dot r^3(t) \end{pmatrix}, \tag{32}$$

which can be written in vector notation as

$$\tilde V = \Gamma[u(1 + \mathbf{a}\cdot\mathbf{r}) + \boldsymbol\omega\times\mathbf{r} + \dot{\mathbf{r}}], \quad u = e_0. \tag{33}$$

This last result is identical to Eq. (4.8) obtained by DeFacio, Dennis, and Retzloff.[8] To calculate the coordinate acceleration, we note that from Table I and (15) for $\mu = 0$ we obtain

$$\frac{\dot{\Gamma}}{\Gamma} = \frac{d\Gamma/dt}{\Gamma} = -\Gamma^0{}_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta = -\frac{(\dot{\mathbf{a}}\cdot\mathbf{r} + 2\mathbf{a}\cdot\dot{\mathbf{r}} + \mathbf{a}\cdot\boldsymbol{\omega}\times\mathbf{r})}{1+\mathbf{a}\cdot\mathbf{r}}. \tag{34}$$

The definition of D in (17) gives

$$\Gamma^{-1}D = \begin{pmatrix} \Gamma^0_{\sigma\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta \\ -a^1(1+\mathbf{a}\cdot\mathbf{r}) - [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})]^1 - \dot{\omega}_1 r^1 - 2(\boldsymbol{\omega}\times\mathbf{r})^1 + \dfrac{(\boldsymbol{\omega}\times\mathbf{r})^1}{1+\mathbf{a}\cdot\mathbf{r}}[\dot{\mathbf{a}}\cdot\mathbf{r} + 2\mathbf{a}\cdot\dot{\mathbf{r}} + \mathbf{a}\cdot\boldsymbol{\omega}\times\mathbf{r}] \\ -a^2(1+\mathbf{a}\cdot\mathbf{r}) - [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})]^2 - \dot{\omega}_2 r^2 - 2(\boldsymbol{\omega}\times\mathbf{r})^2 + \dfrac{(\dot{\boldsymbol{\omega}}\times\mathbf{r})^2}{1+\mathbf{a}\cdot\mathbf{r}}[\dot{\mathbf{a}}\cdot\mathbf{r} + 2\mathbf{a}\cdot\dot{\mathbf{r}} + \mathbf{a}\cdot\boldsymbol{\omega}\times\mathbf{r}] \\ -a^3(1+\mathbf{a}\cdot\mathbf{r}) - [\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})]^3 - \dot{\omega}_3 r^3 - 2(\boldsymbol{\omega}\times\mathbf{r})^3 + \dfrac{(\boldsymbol{\omega}\times\mathbf{r})^3}{1+\mathbf{a}\cdot\mathbf{r}}[\dot{\mathbf{a}}\cdot\mathbf{r} + 2\mathbf{a}\cdot\dot{\mathbf{r}} + \mathbf{a}\cdot\boldsymbol{\omega}\times\mathbf{r}] \end{pmatrix}. \tag{35}$$

Substituting (35), (34), and (11) into (18) gives in compact vector notation

$$A = \ddot{\mathbf{r}} = -\mathbf{a}(1+\mathbf{a}\cdot\mathbf{r}) - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) - \dot{\boldsymbol{\omega}}\cdot\mathbf{r} - 2(\boldsymbol{\omega}\times\dot{\mathbf{r}}) + \frac{(\dot{\mathbf{r}} + \boldsymbol{\omega}\times\mathbf{r})}{1+\mathbf{a}\cdot\mathbf{r}}(\dot{\mathbf{a}}\cdot\mathbf{r} + 2\mathbf{a}\cdot\dot{\mathbf{r}} + \mathbf{a}\cdot\boldsymbol{\omega}\times\mathbf{r}). \tag{36}$$

This latter result agrees exactly with equation (4.14) of DeFacio, Dennis, and Retzloff.[8]

In a general relativistic framework the calculations begin with the Taylor series expansion of the $\Gamma^\mu{}_{\alpha\beta}$ as given by (19) from which the B matrix in (10) can be written as

$$\begin{aligned} B(s,t) &= A(0,t) + A_i(0,t)n^i(t)s + A_{ij}(0,t)n^i(t)n^j(t)s^2/2! + \cdots, \\ a_{\mu\nu}(0,t) &= -n^k(t)\Gamma^\mu_{k\nu}(0,t), \\ (a_{\mu\nu})_i(0,t) &= -n^k(t)\Gamma^\mu_{k\nu,i}(t,0), \\ (a_{\mu\nu})_{ij}(0,t) &= -n^k(t)\Gamma^\mu_{k\nu,ij}(t,0), \quad \text{etc.} \end{aligned} \tag{37}$$

The A matrices for the first three terms of (37) are

$$A(0,t) = \begin{pmatrix} -a\cdot n(t) & 0 & 0 & 0 \\ -[\omega\times n(t)]^1 & 0 & 0 & 0 \\ -[\omega\times n(t)]^2 & 0 & 0 & 0 \\ -[\omega\times n(t)]^3 & 0 & 0 & 0 \end{pmatrix}$$

$$A_j(0,t)n^j(t) = \begin{pmatrix} [a\cdot n(t)]^2 - R^0{}_{i0j}n^i(t)n^j(t) & \frac{1}{3}R^0{}_{i1k}n^i(t)n^k(t) & \frac{1}{3}R^0{}_{i2k}n^i(t)n^k(t) & \frac{1}{3}R^0{}_{i3k}n^i(t)n^k(t) \\ [a\cdot n(t)][\omega\times n(t)]^1 - R^1{}_{ik0}n^i(t)n^k(t) & \frac{1}{3}R^1{}_{i1k}n^i(t)n^k(t) & \frac{1}{3}R^1{}_{i2k}n^i(t)n^k(t) & \frac{1}{3}R^1{}_{i3k}n^i(t)n^k(t) \\ [a\cdot n(t)][\omega\times n(t)]^2 - R^2{}_{ik0}n^i(t)n^k(t) & \frac{1}{3}R^2{}_{i1k}n^i(t)n^k(t) & \frac{1}{3}R^2{}_{i2k}n^i(t)n^k(t) & \frac{1}{3}R^2{}_{i3k}n^i(t)n^k(t) \\ [a\cdot n(t)][\omega\times n(t)]^3 - R^3{}_{ik0}n^i(t)n^k(t) & \frac{1}{3}R^3{}_{i1k}n^i(t)n^k(t) & \frac{1}{3}R^3{}_{i2k}n^i(t)n^k(t) & \frac{1}{3}R^3{}_{i3k}n^i(t)n^k(t) \end{pmatrix}$$

$$A_{ij}(0,t)n^in^j = \left\{ \begin{array}{ll} \{R_{0l0j;1}n^ln^jn^i + 2(a\cdot n)R_{0i0j}n^in^j & -\big[\frac{1}{12}P_{l1}(5R^0{}_{li1;j} - R^0{}_{ilj;1})n^i \\ -2(a\cdot n)^3 + \frac{2}{3}(\omega\times n)^p(R_{0lip} + R_{0pil})n^in^l\} & \quad -\frac{2}{3}(a\cdot r)P_{l1}R^0{}_{lj1}\big]n^ln^j \\ \{R_{0i1l;j}n^in^ln^j + \frac{2}{3}(a\cdot n)R_{01jl}n^jn^l & -\big[\frac{1}{12}P_{l1}(5R^1{}_{li1;j} - R^1{}_{ilj;1})n^i \\ +2(\omega\times n)^1R_{0l0j}n^ln^j + \frac{2}{3}(\omega\times n)^k(R_{kj1l} + R_{k1jl})n^jn^l & \quad -\frac{2}{3}(\omega\times n)^1P_{l1}R^0{}_{lj1}\big]n^ln^j \\ \quad -2(\omega\times n)^1(a\cdot n)^2\} & \\ \{R_{0i2l;j}n^in^ln^j + \frac{2}{3}(a\cdot n)R_{02jl}n^jn^l & -\big[\frac{1}{12}P_{l1}(5R^2{}_{li1;j} - R^2{}_{ilj;1})n^i \\ +2(\omega\times n)^2R_{0l0j}n^ln^j + \frac{2}{3}(\omega\times n)^k(R_{kj2l} + R_{k2jl})n^jn^l & \quad -\frac{2}{3}(\omega\times n)^2P_{l1}R^0{}_{lj1}n^ln^j\big] \\ \quad -2(\omega\times n)^2(a\cdot n)^2\} & \\ \{R_{0i3l;j}n^in^ln^j + \frac{2}{3}(a\cdot n)R_{03jl}n^jn^l & -\big[\frac{1}{12}P_{l1}(5R^3{}_{li1;j} - R^3{}_{ilj;1})n^i \\ +2(\omega\times n)^3R_{0l0j}n^ln^j + \frac{2}{3}(\omega\times n)^k(R_{kj3l} + R_{k3jl})n^jn^l & \quad -\frac{2}{3}(\omega\times n)^3P_{l1}R^0{}_{lj1}\big]n^ln^j \\ \quad -2(\omega\times n)^3(a\cdot n)^2\} & \end{array} \right.$$

$$
\left.\begin{array}{ll}
-\left[\tfrac{1}{12}P_{l2}(5R^{0}{}_{li2;j}-R^{0}{}_{ilj;2})n^{i} & -\left[\tfrac{1}{12}P_{l3}(5R^{0}{}_{li3;j}-R^{0}{}_{ilj;3})n^{i}\right. \\
\quad -\tfrac{2}{3}(a\cdot n)P_{l2}R^{0}{}_{lj2}\right]n^{l}n^{j} & \quad -\tfrac{2}{3}(a\cdot n)P_{l3}R^{0}{}_{lj3}\left.\right]n^{l}n^{j} \\
 & \\
-\left[\tfrac{1}{12}P_{l2}(5R^{1}{}_{li2;j}-R^{1}{}_{ilj;2})n^{i}\right. & -\left[\tfrac{1}{12}P_{l3}(5R^{1}{}_{li3;j}-R^{1}{}_{ilj;3})n^{i}\right. \\
\quad -\tfrac{2}{3}(\omega\times n)^{1}P_{l2}R^{0}{}_{lj2}\left.\right]n^{l}n^{j} & \quad -\tfrac{2}{3}(\omega\times n)^{1}P_{l3}R^{0}{}_{lj3}\left.\right]n^{l}n^{j} \\
 & \\
-\left[\tfrac{1}{12}P_{l2}(5R^{2}{}_{li2;j}-R^{2}{}_{ilj;2})n^{i}\right. & -\left[\tfrac{1}{12}P_{l3}(5R^{2}{}_{li3;j}-R^{2}{}_{ilj;3})n^{i}\right. \\
\quad \tfrac{2}{3}(\omega\times n)^{2}P_{l2}R^{0}{}_{lj2}\left.\right]n^{l}n^{j} & \quad -\tfrac{2}{3}(\omega\times n)^{2}P_{l3}R^{0}{}_{lj3}\left.\right]n^{l}n^{j} \\
 & \\
-\left[\tfrac{1}{12}P_{l2}(5R^{3}{}_{li2;j}-R^{3}{}_{ilj;2})n^{i}\right. & -\left[\tfrac{1}{12}P_{l3}(5R^{3}{}_{li3;j}-R^{3}{}_{ilj;3})n^{i}\right. \\
\quad -\tfrac{2}{3}(\omega\times n)^{3}P_{l2}R^{0}{}_{lj2}\left.\right]n^{l}n^{j} & \quad -\tfrac{2}{3}(\omega\times n)^{3}P_{l3}R^{0}{}_{lj3}\left.\right]n^{l}n^{j}
\end{array}\right] \tag{38}
$$

By substituting (38) and (37) into (10), calculating the matrizant $\Omega(s_0(t),t)$ to third order in r and using this result in (11), we obtain the following third-order expression for the coordinate velocity.

$$
\begin{pmatrix} \Gamma \\ \Gamma\dot{r}^{1} \\ \Gamma\dot{r}^{2} \\ \Gamma\dot{r}^{3} \end{pmatrix} =
$$

$$
\left(\begin{array}{l}
1-a\cdot r+(a\cdot r)^{2}-(a\cdot r)^{3}-\tfrac{1}{2}R_{0i0j}r^{i}r^{j}(1-\tfrac{5}{3}a\cdot r)+(1/3!)R_{0i0j,k}r^{i}r^{j}r^{k} \\
-(\omega\times r)^{1}[1-a\cdot r+(a\cdot r)^{2}]-\tfrac{1}{2}R_{1ij0}r^{i}r^{j}(1-\tfrac{2}{3}a\cdot r)+\tfrac{1}{2}(\omega\times r)^{1}R_{0i0j}r^{i}r^{j}+(1/3!)R_{0i1j,k}r^{i}r^{j}r^{k} \\
-(\omega\times r)^{2}[1-a\cdot r+(a\cdot r)^{2}]-\tfrac{1}{2}R_{2ij0}r^{i}r^{j}(1-\tfrac{2}{3}a\cdot r)+\tfrac{1}{2}(\omega\times r)^{2}R_{0i0j}r^{i}r^{j}+(1/3!)R_{0i2jk}r^{i}r^{j}r^{k} \\
-(\omega\times r)^{3}[1-a\cdot r+(a\cdot r)^{2}]-\tfrac{1}{2}R_{3ij0}r^{i}r^{j}(1-\tfrac{2}{3}a\cdot r)+\tfrac{1}{2}(\omega\times r)^{3}R_{0i0j}r^{i}r^{j}+(1/3!)R_{0i3j,k}r^{i}r^{j}r^{k}
\end{array}\right.
$$

$$
\begin{array}{l}
\tfrac{1}{6}R^{0}{}_{i1j}r^{i}r^{j}(1-a\cdot r)-\tfrac{1}{72}P_{i1}(5R^{0}{}_{ik1;l}-R^{0}{}_{kil;1})r^{i}r^{k}r^{l} \\
1+\tfrac{1}{6}R^{1}{}_{i1j}r^{i}r^{j}-(1/3!)(\omega\times r)^{1}R^{0}{}_{i1j}r^{i}r^{j}-\tfrac{1}{72}P_{i1}(5R^{1}{}_{ik1;l}-R^{1}{}_{kil;1})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{2}{}_{i1j}r^{i}r^{j}-(1/3!)(\omega\times r)^{2}R^{0}{}_{i1j}r^{i}r^{j}-\tfrac{1}{72}P_{i1}(5R^{2}{}_{ik1;l}-R^{2}{}_{kil;1})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{3}{}_{i1j}r^{i}r^{j}-(1/3!)(\omega\times r)^{3}R^{0}{}_{i1j}r^{i}r^{j}-\tfrac{1}{72}P_{i1}(5R^{3}{}_{ik1;l}-R^{3}{}_{kil;1})r^{i}r^{k}r^{l}
\end{array}
$$

$$
\begin{array}{l}
\tfrac{1}{6}R^{0}{}_{i2j}r^{i}r^{j}(1-a\cdot r)-\tfrac{1}{72}P_{i2}(5R^{0}{}_{ik2;l}-R^{0}{}_{kil;2})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{1}{}_{i2j}r^{i}r^{j}-(1/3!)(\omega\times r)^{1}R^{0}{}_{i2j}r^{i}r^{j}-\tfrac{1}{72}P_{i2}(5R^{1}{}_{ik2;l}-R^{1}{}_{kil;2})r^{i}r^{k}r^{l} \\
1+\tfrac{1}{6}R^{2}{}_{i2j}r^{i}r^{j}-(1/3!)(\omega\times r)^{2}R^{0}{}_{i2j}r^{i}r^{j}-\tfrac{1}{72}P_{i2}(5R^{2}{}_{ik2;l}-R^{2}{}_{kil;2})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{3}{}_{i2j}r^{i}r^{j}-(1/3!)(\omega\times r)^{3}R^{0}{}_{i2j}r^{i}r^{j}-\tfrac{1}{72}P_{i2}(5R^{3}{}_{ik2;l}-R^{3}{}_{kil;2})r^{i}r^{k}r^{l}
\end{array}
$$

$$
\left.\begin{array}{l}
\tfrac{1}{6}R^{0}{}_{i3j}r^{i}r^{j}(1-a\cdot r)-\tfrac{1}{72}P_{i3}(5R^{0}{}_{ik3;l}-R^{0}{}_{kil;3})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{1}{}_{i3j}r^{i}r^{j}-(1/3!)(\omega\times r)^{1}R^{0}{}_{i3j}r^{i}r^{j}-\tfrac{1}{72}P_{i3}(5R^{1}{}_{ik3;l}-R^{1}{}_{kil;3})r^{i}r^{k}r^{l} \\
\tfrac{1}{6}R^{2}{}_{i3j}r^{i}r^{j}-(1/3!)(\omega\times r)^{2}R^{0}{}_{i3j}r^{i}r^{j}-\tfrac{1}{72}P_{i3}(5R^{2}{}_{ik3;l}-R^{2}{}_{kil;3})r^{i}r^{k}r^{l} \\
1+\tfrac{1}{6}R^{3}{}_{i3j}r^{i}r^{j}-(1/3!)(\omega\times r)^{3}R^{0}{}_{i3j}r^{i}r^{j}-\tfrac{1}{72}P_{i3}(5R^{3}{}_{ik3;l}-R^{3}{}_{kil;3})r^{i}r^{k}r^{l}.
\end{array}\right] \cdot \begin{pmatrix} \widetilde{V}^{0} \\ \widetilde{V}^{1} \\ \widetilde{V}^{2} \\ \widetilde{V}^{3} \end{pmatrix}_{|\gamma(t)}, \tag{39}
$$

where

$$
\widetilde{V}_{|\gamma(t)} = \begin{pmatrix} \widetilde{V}^{0} \\ \widetilde{V}^{1} \\ \widetilde{V}^{2} \\ \widetilde{V}^{3} \end{pmatrix}_{|\gamma(t)} \tag{40}
$$

is the velocity of the geodesic particle and W, the coordinate velocity, is related to the left-hand side of (39) by

$$
\overline{W} = \begin{pmatrix} 1 \\ W^{1}(t) \\ W^{2}(t) \\ W^{3}(t) \end{pmatrix}_{|\chi(t)} = \Gamma^{-1} \begin{pmatrix} \Gamma \\ \Gamma\dot{r}^{1} \\ \Gamma\dot{r}^{2} \\ \Gamma\dot{r}^{3} \end{pmatrix}_{|\chi(t)}. \tag{41}
$$

It is immediately apparent that (39) reduces to the third order expansion of (32) for the case of special relativity. To compare this result with the work of Li and Ni[11] we must obtain an equation relating Γ to the physical variables of the problem. This is accomplished using the e_0-equation of (18), or equivalently (14), which reduces to

$$
\begin{aligned}
\dot{\Gamma}/\Gamma = & -\Gamma^{0}{}_{\alpha\beta}(t,s_0(t))\overline{W}_\alpha \overline{W}_\beta \\
= & -2(a\cdot\dot{r})[1-a\cdot r+(a\cdot r)^2]-2a\cdot\omega\times r[1-a\cdot r] \\
& -b\cdot r[1-a\cdot r]+\tfrac{2}{3}R^{0}{}_{ijk}\dot{r}^i\dot{r}^j r^k[1-a\cdot r] \\
& -2R_{0i0j}\dot{r}^i r^j\left[1-\tfrac{2}{3}a\cdot r\right] \\
& -\left[\tfrac{1}{2}R_{0i0j;0}+\tfrac{1}{3}a^k R_{0jki}-\tfrac{2}{3}(a\cdot\dot{r})R_{0i0j}\right]r^i r^j \\
& +\tfrac{2}{3}(\omega\times r)^k(R_{0ijk}+R_{0kji})\dot{r}^i r^j \\
& +\tfrac{1}{3}(R_{0i0k;j}+R_{0k0i;j}+R_{0k0j;i})\dot{r}^i r^j r^k \\
& -(2/4!)(5R^{0}{}_{ikj;l}-R^{0}{}_{kil;j})\dot{r}^i\dot{r}^j r^k r^l.
\end{aligned}
\tag{42}
$$

The solution to (42) is

$$
\begin{aligned}
\Gamma = & \,\Gamma(t=0)\exp\left\{-\int_0^t \Gamma^{0}{}_{\alpha\beta}(t,s_0(t))\overline{W}^\alpha\overline{W}^\beta dt\right\} \\
= & \,\Gamma(t=0)\left\{1-\int_0^t \Gamma^{0}{}_{\alpha\beta}(t,s_0(t))\overline{W}^\alpha\overline{W}^\beta dt \right. \\
& \left. +\frac{1}{2!}\left[\int_0^t \Gamma^{0}{}_{\alpha\beta}(t,s_0(t))\overline{W}^\alpha\overline{W}^\beta dt\right]^2+\cdots\right\}.
\end{aligned}
\tag{43}
$$

Integrating only the terms involving a first derivative with respect to t, we obtain

$$
\begin{aligned}
\Gamma = & \,\Gamma(t=0)\left\{1-\left\{3a\cdot r\left[1-a\cdot r+\tfrac{2}{3}(a\cdot r)^2\right]\right.\right. \\
& +R_{0i0j}r^i r^j[3-2a\cdot r] \\
& -\tfrac{2}{3}(\omega\times r)^k r_{0ijk}r^i r^j-\tfrac{1}{3}(R_{0i0k;j} \\
& +R_{0k0i;j}+R_{0k0j;i})r^i r^j r^k \\
& \left.\left.-(2/4!)(5R^{0}{}_{ikj;l}-R^{0}{}_{kil;j})\dot{r}^i r^j r^k r^l+\cdots\right\}+(1/2!)\{\ \}^2-\cdots\right\}.
\end{aligned}
\tag{44}
$$

To second order in r, (44) becomes

$$
\Gamma=\Gamma(t=0)\left\{1-3a\cdot r\left[1+\tfrac{1}{2}a\cdot r\right]+3R_{0i0j}r^i r^j+O(r^3)\right\}. \tag{45}
$$

Li and Ni[10] implicitly treat the case $\Gamma(t=0)=1$ and consider only the first-order representation of Γ for which (45) reduces to

$$
\Gamma=1-3a\cdot r+O(r^2). \tag{46}
$$

If, however, we integrate only the exact terms in (43) the first-order representation of Γ is

$$
\Gamma=1-a\cdot r+O(r^2). \tag{47}
$$

A comparison of (47) with equation (33) of Li and Ni[11] shows that they are the same. This establishes that the results of Li and Ni contain only the exact terms for the representation of Γ.

Now from (39) we have

$$
\frac{\widetilde{V}^0}{\Gamma}=\frac{1-\left[\tfrac{1}{6}R^{0}{}_{ipj}r^i r^j(1-a\cdot r)-\tfrac{1}{72}P_{ip}(5R^{0}{}_{ikp;l}-R^{0}{}_{kil;p})r^i r^k r^l\right]\widetilde{V}^p/\Gamma}{1-a\cdot r+(a\cdot r)^2-(a\cdot r)^3-\tfrac{1}{2}R_{0i0j}r^i r^j(1-\tfrac{5}{3}a\cdot r)+(1/3!)R_{0i0j;k}r^i r^j r^k},
$$

$$
\begin{aligned}
W^i=\dot{r}^i=(\widetilde{V}^0/\Gamma)\{ & -(\omega\times r)^i[1-a\cdot r+(a\cdot r)^2] \\
& -\tfrac{1}{2}R_{ijk0}r^j r^k(1-\tfrac{2}{3}a\cdot r)+\tfrac{1}{2}(\omega\times r)^i R_{0j0k}r^j r^k \\
& +(1/3!)R_{0jik;l}r^j r^k r^l\}+(\widetilde{V}^k/\Gamma)\{\delta^i_k \\
& +\tfrac{1}{6}R^{i}{}_{jkl}r^j r^l-(1/3!)(\omega\times r)^i R^{0}{}_{jkl}r^j r^l \\
& -\tfrac{1}{72}P_{lk}(5R^{i}{}_{lpk;m}-R^{i}{}_{plm;k})r^l r^p r^m\},
\end{aligned}
\tag{48}
$$

which when combined with (47) and the usual expansion of the denominator of (48) yields

$$
\begin{aligned}
W^i=\dot{r}^i= & -(\omega\times r)^i[1+a\cdot r-(a\cdot r)^2] \\
& -\tfrac{1}{2}R_{ijk0}r^j r^k(1+\tfrac{1}{3}a\cdot r) \\
& +\widetilde{V}^p\left[(1+a\cdot r)\delta^i_p+\tfrac{1}{6}R^{i}{}_{jpl}r^j r^l(1+a\cdot r)\right. \\
& \left.-\tfrac{1}{72}P_{lp}(5R^{i}{}_{lkp;m}-R^{i}{}_{klm;p})r^l r^k r^m\right]+O(r^4).
\end{aligned}
\tag{49}
$$

A comparison of (49) with Eq. (37) of Li and Ni[11] shows that the two equations do not agree. This is indicative of the problems that are encountered in trying to obtain the coordinate velocity by integrating the coordinate acceleration as per Li and Ni. Basically, Li and Ni's equations (27) and (28) are incomplete because they consider only those terms that give rise to exact differentials and do not recognize that the coordinate acceleration is given in the $\tilde{e}_{|\mu|}$ basis while $\widetilde{V}$ is stated in terms of the $e_{|\mu|}$ basis. The resulting mixing of the components of $\widetilde{V}$ in the expression for $\overline{W}$ in (49) does not occur in special relativity because the underlying manifold is flat and the two basis sets $\tilde{e}_{|\mu|}$ and $e_{|\mu|}$ are the same to within scale factors. We can also write (49) as the vector equation

$$
\begin{aligned}
W= & -(\omega\times r)[1+a\cdot r-(a\cdot r)^2]-\tfrac{1}{2}(1+\tfrac{1}{3}a\cdot r)R(r,e_0)r \\
& +\widetilde{V}(1+a\cdot r)+\tfrac{1}{6}(1+a\cdot r)R(V,r)r \\
& -\tfrac{1}{72}[\nabla_r 5R(r,V)r-\nabla_v R(r,r)r],
\end{aligned}
\tag{50}
$$

which allows us to obtain the coordinate acceleration in terms of either the $\tilde{e}_{|\mu|}$ or $e_{|\mu|}$ basis set.

In order to calculate the coordinate acceleration, we first note that the D matrix is

$$
\begin{aligned}
-D^0/\Gamma= & \,2(a\cdot\dot{r})[1-a\cdot r+(a\cdot r)^2] \\
& +a\cdot\omega\times r[1-a\cdot r]+(\dot{a}\cdot r)[1-a\cdot r] \\
& -\tfrac{2}{3}R^{0}{}_{ijk}\dot{r}^i\dot{r}^j r^k[1-a\cdot r]+2R_{0j0i}\dot{r}^j r^i\left[1-\tfrac{2}{3}a\cdot r\right] \\
& +\left[\tfrac{1}{2}R_{0i0j;0}+\tfrac{1}{3}a^k R_{0jki}-\tfrac{2}{3}(a\cdot\dot{r})R_{0i0j}\right]r^i r^j \\
& -\tfrac{2}{3}(\omega\times r)^k(R_{0ijk}+R_{0kji})\dot{r}^i r^j \\
& -\tfrac{1}{3}(R_{0i0k;j}+R_{0k0i;j}+R_{0k0j;i})\dot{r}^i r^j r^k \\
& +(2/4!)(5R^{0}{}_{ikj;l}-R^{0}{}_{kil;j})\dot{r}^i\dot{r}^j r^k r^l,
\end{aligned}
\tag{51}
$$

$$
\begin{aligned}
-D^i/\Gamma= & \,a^i[1+a\cdot r]+2(\omega\times\dot{r})^i \\
& -2(\omega\times r)^i(a\cdot r)[1-a\cdot r]+(\eta\times r)^i \\
& +\omega^i(\omega\cdot r)-\delta^i_j r^j(\omega^l)^2-(b\cdot r)(\omega\times r)^i-2a\cdot\omega\times r(\omega\times r)^i \\
& +[R_{0i0j}(1+2a\cdot r)-2(\omega\times r)^k R_{0jik}]r^j+[-\tfrac{1}{2}R_{0j0k,i} \\
& +R_{0i0k,j}-\tfrac{1}{3}a^l R_{iklj}+\tfrac{2}{3}(a\cdot\dot{r})R_{0kij}]r^j r^k \\
& +\left[2R_{ijk0}-\tfrac{2}{3}(a\cdot r)(R_{0kij}+R_{0ikj})-2(\omega\times r)^i R_{0j0k}\right. \\
& \left.-\tfrac{2}{3}(\omega\times r)^p(R_{pkij}+R_{pikj})\right]\dot{r}^j r^k \\
& +\left[-\tfrac{2}{3}R^{i}{}_{jkl}-\tfrac{2}{3}(\omega\times r)^i R^{0}{}_{jlk}\right]\dot{r}^j\dot{r}^k r^l \\
& +\left[-\tfrac{1}{3}R_{0jik;l}-R_{0kij;l}+\tfrac{1}{3}R_{0kil;j}\right]\dot{r}^j\dot{r}^k r^l \\
& +(2/4!)(5R^{i}{}_{jlk;p}-R^{i}{}_{ljp;k})\dot{r}^j\dot{r}^k r^l r^p.
\end{aligned}
$$

Combining (18), (42), and (51), we obtain the coordinate acceleration as given by

$$
\begin{aligned}
A^i{}_{|_{\chi(t)}} = & - a^i(1 + a\cdot r) - [\omega\times(\omega\times r)]^i - (\eta\times r)^i - 2(\omega\times W)^i + 2(a\cdot W)(\omega\times r)^i \\
& + W^i[b\cdot r + 2a\cdot(\omega\times r) + 2a\cdot W(1 - a\cdot r)] + (b\cdot r)(\omega\times r)^i + 2a\cdot\omega\times r(\omega\times r)^i \\
& - 2a\cdot W a\cdot r(\omega\times r)^i - W^i a\cdot r b\cdot r + 2\omega^i a\cdot W(a\cdot r)^2 - 2W^i(a\cdot r)a\cdot\omega\times r \\
& - R_{0i0j}r^j - 2R_{ijk0}r^k W^j + \tfrac{2}{3}R_{ijkl}W^j W^k r^l + 2R_{0j0k}W^i W^j r^k \\
& + \tfrac{2}{3}R_{0jkl}W^i W^j W^k r^l + \tfrac{1}{3}a^k R_{ijkp}r^j r^p + 2(\omega\times r)^p R_{0jip}r^j \\
& - 2(a\cdot r)R_{0i0j}r^j + \tfrac{2}{3}(a\cdot r)(R_{0jik} + R_{0ijk})W^k r^j - \tfrac{2}{3}a\cdot W R_{0jip}r^j r^p \\
& + 2(\omega\times r)^i R_{0j0k}W^j r^k + \tfrac{2}{3}(\omega\times r)^k(R_{klij} + R_{kilj})W^j r^l \\
& + \tfrac{2}{3}(\omega\times r)^i R_{0kjp}W^j W^k r^p + \tfrac{1}{3}a^k R_{0pkl}W^i r^l r^p \\
& - \tfrac{4}{3}(a\cdot r)R_{0j0l}W^i W^j r^l - \tfrac{2}{3}a\cdot W R_{0j0k}W^i r^j r^k + \tfrac{2}{3}(\omega\times r)^k(R_{0jkl} + R_{0kjl})W^i W^j r^l \\
& - \tfrac{2}{3}(a\cdot r)R_{0kjp}W^i W^j W^k r^p - \tfrac{1}{2}(R_{ijk0;0} + R_{i0j0;k})r^j r^k \\
& + \tfrac{1}{3}R_{ijkl;0}r^j r^l W^k - R_{ijl0;p}r^l r^p W^j + \tfrac{1}{12}(5R_{ikjl;p} + R_{iljp;k})W^j W^k r^l r^p \\
& + \tfrac{1}{2}R_{0l0p;0}r^l r^p W^i + R_{0j0k;l}r^k r^l W^i W^j \\
& + \tfrac{1}{3}R_{0ljk;0}r^l r^k W^i W^j + \tfrac{1}{12}(5R_{0kjl;p} + R_{0ljp;k})r^l r^p W^i W^j W^k + O(r^3).
\end{aligned}
\tag{52}
$$

Equation (52) is identical to equation (21) of Li and Ni.[11] This is no surprise as (18) is the parallel translation of (14) and hence the same as equation (20) of Li and Ni. The proof of this statement is the subject of the next theorem.

Theorem 1: Equation (18) is equal to the parallel translation of (14).

Proof: Writing (18) out in component form we have

$$
\begin{aligned}
(\dot{\Gamma}/\Gamma)\tilde{e}_{0|_{\gamma(t)}} &= -\Gamma^0{}_{\alpha\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta\tilde{e}_{0|_{\gamma(t)}} \\
[A^i + \Gamma^i{}_{\alpha\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta]\tilde{e}_{i|_{\gamma(t)}} &= -\frac{\dot{\Gamma}}{\Gamma}\overline{W}^i\tilde{e}_{i|_{\gamma(t)}}.
\end{aligned}
\tag{53}
$$

Parallel translating (53) to $\gamma_1(t)$ we obtain

$$
\begin{aligned}
\frac{\dot{\Gamma}}{\Gamma e_{0|_{\gamma_1(t)}}} &= -\Gamma^0{}_{\alpha\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta e_{0|_{\gamma_1(t)}} \\
[A^i + \Gamma^i{}_{\alpha\beta}(t,s_0(t))\overline{W}^\sigma\overline{W}^\beta]e_{i|_{\gamma_1(t)}} &= -\frac{\dot{\Gamma}}{\Gamma}\overline{W}^i e_{i|_{\gamma(t)}}.
\end{aligned}
\tag{54}
$$

which can be rewritten in vector notation as

$$
\nabla_{d/dt}\nabla_{d/dt}\gamma_1(t) = -\frac{\dot{\Gamma}}{\Gamma}\nabla_{d/dt}\gamma_1(t). \qquad \text{Q.E.D.} \tag{55}
$$

V. CONCLUSION

In this paper we have used Ekstein's presymmetry and covariance to obtain two distinct definitions of the observation of a vector field by a general observer, and modern differential geometry to define a natural connection for the general observer. This led to an invariant definition of the coordinate velocity and coordinate acceleration of a geodesic particle and yielded closed form expressions for these quantities. We have examined the consequences of our definition of observation based on the mapping of tangent spaces to tangent spaces via parallel translation. Our results subsume the exact special relativity treatment of DeFacio, Dennis, and Retzloff as well as the general relativity work of Ni and Zimmerman and Li and Ni. We have clearly identified the tangent spaces of the relevant vector fields and in so doing have shown that the expressions of Ni and Zimmerman and Li and Ni for the coordinate velocity and Γ are incomplete to the order stated in their work. The procedure for obtaining higher-order coordinate representations of A and W are clearly stated.

ACKNOWLEDGMENTS

All authors thank J. K. Beem, S. Harris, P. Ehrlich, and C. Ahlbrandt of the University of Missouri Mathematics Department for valuable conversations and comments.

[1] J. L. Synge, *Relativity: The General Theory* (North-Holland, Amsterdam, 1960).
[2] E. Fermi, Atti R. Accad. Linceri Rend. Cl. Sci. Fis. Mat. Nat. **31**, 21 (1922); **31**, 51 (1922).
[3] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), henceforth cited as MTW.
[4] F. K. Manasse and C. W. Misner, J. Math. Phys. **4**, 735 (1963).
[5] R. Burghardt, Acta Phys. Austriaca **48**, 181 (1978).
[6] F. B. Estabrook and H. D. Wahlquist, J. Math. Phys. **5**, 1629 (1964).
[7] W. -Q. Li and W. -T. Ni, Chin. J. Phys. **16**, 223 (1978); **15**, 51 (1977); **16**, 214 (1978); **16**, 228 (1978).
[8] B. DeFacio, P. W. Dennis, and D. G. Retzloff, Phys. Rev. D **18**, 2813 (1978); Phys. Rev. D **20**, 570 (1979).
[9] W. -T. Ni and M. Zimmerman, Phy. Rev. D **17**, 1473 (1978).
[10] B. Mashhoon, Astrophys. J. **216**, 591 (1977).
[11] W. -Q. Li and W. -T. Ni, J. Math. Phys. **20**, 1473 (1979).
[12] H. Ekstein, Phy. Rev. **184**, 1315 (1969); Phys. Rev. D **1**, 1851(E) (1970); Y. Avishai and H. Ekstein, Found. Phys. **2**, 257 (1972); Commun. Math. Phys. **37**, 193 (1974); Phy. Rev. D **7**, 983 (1973).
[13] H. P. W. Gottlieb, J. Phys. A **9**, 417 (1976).
[14] B. DeFacio and D. G. Retzloff, J. Math. Phys. **21**, 751 (1980).
[15] W. T. Reid, *Ordinary Differential Equations* (Wiley, New York, 1971).

A new mathematical formulation of accelerated observers in general relativity. II

D. G. Retzloff
Department of Chemical Engineering, University of Missouri-Columbia, Columbia, Missouri 65211

B. DeFacio
Department of Physics, University of Missouri-Columbia, Columbia, Missouri 65211, Applied Mathematical Sciences, Ames Laboratory, USDOE, and Iowa State University, Ames, Iowa 50011

P. W. Dennis
BDM Corporation, Redondo Beach, California 90278

(Received 25 August 1980; accepted for publication 14 November 1980)

The observation of a general vector field based on $\exp_*$ is employed to obtain formulas for the coordinate velocity and coordinate acceleration of a geodesic particle. Our results are shown to reduce to those based on a parallel transport definition of observation in special relativity. In general relativity the difference between the expressions for the coordinate velocity and coordinate acceleration derived from the two definitions of observation is given in terms of the Riemann curvature tensor.

PACS numbers: 04.20.Cv

I. INTRODUCTION

In a previous paper[1] (hereafter referred to as I) we gave two distinct definitions of the observation of a general spacetime event by a general relativistic accelerating rotating observer (GRARO) which are consistent with covariance and presymmetry[2,3] considerations. In I we developed the consequences of the second definition in which the observation of an event at $q \in M$ by a GRARO at $p \in M$ is given by the parallel translation map τ_{pq}. The purpose of the current paper is to investigate the consequences of the first definition wherein the observation of an event at q by a GRARO at p is given by $\exp_*$. The analysis will focus on the observation of a geodesic particle by a GRARO. It will be shown that for special relativity the two definitions of observation of a spacetime event give identical results, while by contrast, the difference in the results of the two definitions of observation in general relativity is measured by the Riemann curvature tensor.

The organization of this paper is as follows. In Sec. II we show that $\exp_*$ is the solution of an appropriate Jacobi field equation and prove that the difference between $\exp_*$ and τ_{pq} for the observation of a geodesic particle by a GRARO is determined by the Riemann curvature tensor. Section III contains our calculations of the coordinate velocity and coordinate acceleration. Our conclusions are given in Sec. IV.

II. THE JACOBI FIELD EQUATION AND exp.

From Definition 1 of I the observation of the velocity of the geodesic particle by the GRARO is given by

$$\dot{\gamma}_{r(t)} \widetilde{V} = \exp^{-1}_{\gamma(t)*_{r(t)}} V, \tag{1}$$

or

$$V = \exp_{\gamma(t)*_{r(t)}} \dot{\gamma}_{r(t)} \widetilde{V}$$

as shown in Fig. 1. Using (1), we state and prove the following theorem.

Theorem 1: The velocity vector $\widetilde{V}$ is the covariant derivative with respect to $d\gamma_2/ds$ of the Jacobi field associated with the map $\alpha(s, l) \rightarrow \exp_{\gamma(t)} s(r(t) + l\widetilde{V}(t))$ for fixed t.

Proof: Consider the map

$$\alpha(s, l) = \exp_{\gamma(t)} s(r(t) + l\widetilde{V}(t)), \tag{2}$$

and the associated Jacobi field along $s \rightarrow \exp_{\gamma(t)} r(t)$ given by

$$J(s) = \alpha_*(\partial/\partial l)\big|_{l=0}, \tag{3}$$

with

$$J(0) = 0, \quad J'(0) = \nabla_{d/ds} J\big|_{s=0} = \widetilde{V}. \tag{4}$$

This Jacobi Field is depicted in Fig. 2 and satisfies the equation

$$\nabla_{d/ds} \nabla_{d/ds} J = R\,(d/ds, J)d/ds. \tag{5}$$

Now from a consideration of the canonical isomorphism

$$\dot{\Upsilon}_{r(t)}: T_p M \rightarrow T_{r(t)} T_p M \tag{6}$$

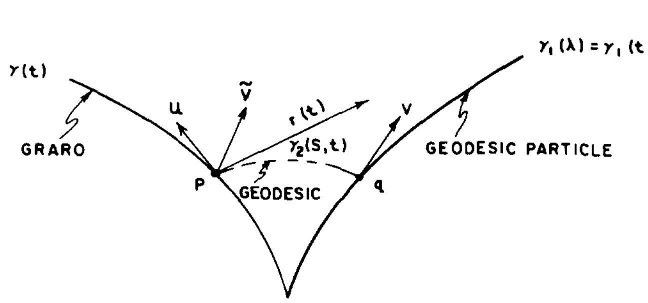


FIG. 1. The manifold defined by the GRARO and the geodesic particle. The symbols are defined as follows $\gamma(t)$ —world line of the GRARO; $\gamma_1(\lambda) = \gamma_1(t)$ —world line of the geodesic particle; λ —affine parameter of the geodesic world line; t —nonaffine parameter; $\gamma_2(s, t)$ —unique geodesic passing through p and q, s —affine parameter of γ_2; $r(t)$ position of geodesic particle relative to the GRARO; u —velocity of GRARO; V —velocity of the geodesic particle; $\widetilde{V}$ —velocity of geodesic particle as seen by the GRARO.

 0022-2488/82/010105-04$02.50

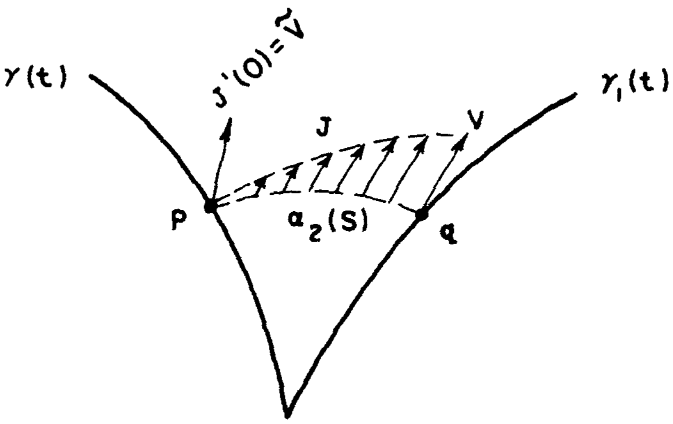


FIG. 2. The Jacobi field associated with the velocity vectors $\widetilde{V}$ and V. Here $\gamma_2(s)$ is given by $\gamma_2(s) = \gamma_2(t, s) = \{t, n^1(t)s, n^2(t)s, n^3(t)s\}$ for fixed t.

given by

$$\dot{\Upsilon}_{r(t)}: \frac{d}{dm}(r(t) + mJ)_{|_{m=0}} \to J, \tag{7}$$

and the map

$$\phi: T_{\gamma(t)}M \to T_{\gamma_1(t)}M \tag{8}$$

determined by

$$\phi = \exp_{\gamma(t)*_{r(t)}} \circ \dot{\Upsilon}_{r(t)},$$

we have

$$\begin{aligned} \phi J &= \exp_{\gamma(t)*_{r(t)}}\left[\frac{d}{dm}(r(t) + mJ)_{|_{m=0}}\right] \\ &= \frac{d}{dm}\exp_{\gamma(t)}(r(t) + mJ)_{|_{m=0}} = J(1) \\ &= \alpha_*(\partial/\partial l)_{|_{l=0,s=1}} = \exp_{\gamma(t)*_{r(t)}} \widetilde{V} = V. \end{aligned} \tag{9}$$

At a general point along $\gamma_2(s)$ we obtain

$$J(s) = \alpha_*(\partial/\partial l)_{|_{l=0,s=s}} = \exp_{\gamma(t)*_{sr(t)}} S\widetilde{V}. \tag{10}$$

Hence

$$J(0) = 0 \quad J'(0) = \nabla_{d/ds} J_{|_{s=0}} = \widetilde{V}, \tag{11}$$

$$J(1) = \exp_{\gamma(t)*_{r(t)}} \widetilde{V} = V.$$

Q.E.D.

It follows from the theorem that in special relativity $\exp_{\gamma(t)*_{r(t)}} \widetilde{V}$ is equivalent to parallel transport and hence Definition 1 and Definition 2 of I for an observation are identical for this case. The proof of this statement is the content of the following corollary.

Corollary. In special relativity $V = \exp_{\gamma(t)*_{r(t)}} \widetilde{V}$ is identical to the parallel translation of $\widetilde{V}$, i.e., $V = \tau_{\gamma(t), \exp_{\gamma(t)} r(t)} \widetilde{V}$.

Proof. For special relativity $R(d/ds, J)d/ds = 0$. Hence (5) reduces to

$$\nabla_{d/ds} \nabla_{d/ds} J = 0. \tag{12}$$

Let $\tilde{e}_{i|_s}$ be a set of basis vectors along $\gamma_2(s)$ obtained from the natural basis vectors $e_{i|_{\gamma_1(t)}}$ at $\gamma_1(t)$ by parallel transport. Furthermore, let

$$J(s) = J^\alpha \tilde{e}_{\alpha|_s} \quad \widetilde{V} = \widetilde{\widetilde{V}}{}^\nu \tilde{e}_{\nu|_{s=0}}, \; V = V^\alpha \tilde{e}_{\alpha|_{s=1}}. \tag{13}$$

With these basis vectors (12) reduces to

$$d^2 J^\alpha / ds = 0, \tag{14}$$

with

$$J^\alpha(0) = 0, \quad \frac{dJ^\alpha}{ds}(0) = \widetilde{\widetilde{V}}{}^\alpha, \tag{15}$$

and

$$J^\alpha(1) = V^\alpha. \tag{16}$$

The solution of (14) satisfying the boundary conditions (15) is

$$J^\alpha(s) = \widetilde{\widetilde{V}}{}^\alpha s. \tag{17}$$

Combining (17) and (16) we have

$$V = \widetilde{\widetilde{V}}{}^\alpha, \tag{18}$$

which are the equations for the parallel translation of V along $\gamma_2(s)$ to $\gamma(t)$. Q.E.D.

We now consider the calculation of the coordinate velocity W and the coordinate acceleration A.

III. SPECIFIC CALCULATIONS OF *A* AND *W*

To calculate the coordinate velocity W we must first solve (5). To this end let e_α represent the natural coordinate basis vectors along the geodesic $\gamma_2(s)$ based on the MTW coordinate system given in I, and $\tilde{e}_\alpha$ be the parallel translated basis vectors along $\gamma_2(s)$ that are identical to $e_{\alpha|_{\gamma_1(t)}}$ at $\gamma_1(t)$ (see Figure 1). Then the tangent vector to $\gamma_2(s)$, T_{γ_2}, at any point s can be written as

$$T_{\gamma_2} \equiv \frac{d}{ds} = \frac{d\gamma_2}{ds} = \tilde{n}^\alpha(t)\tilde{e}_\alpha, \tag{19}$$

with t fixed. Thus we have the following proposition.

Proposition. The $\tilde{n}^\alpha(t)$ given in (19) are independent of s.

Proof. In the MTW coordinate system $\gamma_2(s)$ is written in I as

$$\begin{aligned} \gamma_2(s) &= \{x^0(s), x^1(s), x^2(s), x^3(s)\} \\ &= \{t, n^1(t)s, n^2(t)s, n^3(t)s\}. \end{aligned} \tag{20}$$

T_{γ_2} is given by

$$T_{\gamma_2} = \dot{\gamma}_2(s) = \tilde{n}^\alpha(t)\tilde{e}_\alpha = \dot{x}^\beta_{\tilde{e}_\beta}, \tag{21}$$

and

$$\begin{aligned} 0 = \nabla_{T_{\gamma_2}} T_{\gamma_2} &= \nabla_{\tilde{n}^\beta(t) e_\beta} \tilde{n}^\alpha(t) e_\alpha \\ &= \tilde{n}^\beta(t)\left[\nabla_{\tilde{e}_\beta} \tilde{n}^\alpha(t)\right]\tilde{e}_\alpha. \end{aligned} \tag{22}$$

This implies

$$\begin{aligned} 0 &= \tilde{n}^\beta(t)\nabla_{\tilde{e}_\beta}\tilde{n}^\alpha(t) = \tilde{n}^\beta(t)\tilde{e}_\beta \tilde{n}^\alpha(t) \\ &= \dot{x}^\beta e_\beta \tilde{n}^\alpha(t) = \frac{d\tilde{n}^\alpha(t)}{ds}. \end{aligned} \tag{23}$$

Hence $\tilde{n}^\alpha(t)$ is constant with respect to s. Q.E.D. We now write (5) as

$$\frac{d^2 J^\beta(s)}{ds^2}\tilde{e}_\beta = \tilde{e}_\alpha \widetilde{R}^\alpha{}_{\sigma\gamma\xi}(t, s)\tilde{n}^\sigma(t)\tilde{n}^\gamma(t) J^\xi(s), \tag{24}$$

or

$$\frac{d^2 J^\beta(s)}{ds^2} = \widetilde{R}^\beta{}_{\sigma\gamma\xi}(t, s)\tilde{n}^\sigma(t)\tilde{n}^\gamma(t) J^\xi(s), \quad \beta = 0, 1, 2, 3, \tag{25}$$

with t fixed. Here $\widetilde{R}^\beta{}_{\sigma\gamma\xi}(t, s)$ denotes the Riemann curvature components with respect to the $\tilde{e}_\alpha$ basis set. To proceed further, we expand $\widetilde{R}^\beta{}_{\sigma\gamma\xi}(t, s)$ in a Taylor series, i.e.,

$$\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(t,s)=\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(t,0)+\frac{d}{ds}\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(t,0)s$$
$$+\frac{d^2}{ds^2}\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(t,0)\frac{s^2}{2!}+\cdots=\sum_{j=0}^{\infty}\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(j)\frac{s^j}{j!}. \tag{26}$$

Then

$$\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(t,s)\tilde{n}^{\sigma}(t)\tilde{n}^{\gamma}(t)=\sum_{j=0}^{\infty}\frac{s^j}{j!}\tilde{R}^{\beta}{}_{\sigma\gamma\xi}(j)\tilde{n}^{\sigma}(t)\tilde{n}^{\gamma}(t)$$
$$\equiv\sum_{j=0}^{\infty}\tilde{R}^{\beta}{}_{\xi}(j)\frac{s^j}{j!}. \tag{27}$$

With (27) we can now write (25) as

$$\frac{dz}{dt}=Nz, \tag{28}$$

where

$$Z^T=\{J^0(s),J^1(s),J^2(s),J^3(s),\dot{J}^0(s),\dot{J}^1(s),\dot{J}^2(s),\dot{J}^3(s)\},$$
$$\dot{J}(s)\equiv dJ(s)/ds, \tag{29}$$

and

$$N=\begin{pmatrix}\sum_j\tilde{R}^0{}_0(j)s^j/j! & \sum_j\tilde{R}^0{}_1(j)s^j/j! & \sum_j\tilde{R}^0{}_2(j)s^j/j! & \sum_j\tilde{R}^0{}_3(j)s^j/j!\\ \sum_j\tilde{R}^1{}_0(j)s^j/j! & \sum_j\tilde{R}^1{}_1(j)s^j/j! & \sum_j\tilde{R}^1{}_2(j)s^j/j! & \sum_j\tilde{R}^1{}_3(j)s^j/j!\\ \sum_j\tilde{R}^2{}_0(j)s^j/j! & \sum_j\tilde{R}^2{}_1(j)s^j/j! & \sum_j\tilde{R}^2{}_2(j)s^j/j! & \sum_j\tilde{R}^2{}_3(j)s^j/j!\\ \sum_j\tilde{R}^3{}_0(j)s^j/j! & \sum_j\tilde{R}^3{}_1(j)s^j/j! & \sum_j\tilde{R}^3{}_2(j)s^j/j! & \sum_j\tilde{R}^3{}_3(j)s^j/j!\end{pmatrix}. \tag{30}$$

The initial condition for (28) is

$$Z^T(0)=(0,0,0,0,\tilde{\tilde{V}}^0,\tilde{\tilde{V}}^1,\tilde{\tilde{V}}^2,\tilde{\tilde{V}}^3). \tag{31}$$

Using the matrizant method[4] as in I to solve (28), we can write the solution to (5) to fourth order in s as

$$\bar{W}=\Gamma^{-1}(I+B)\tilde{\tilde{V}}, \tag{32}$$

where $\bar{W}$ and $\tilde{\tilde{V}}\equiv\tilde{V}$ are given with respect to the $\tilde{e}_\alpha$ basis set, the spatial components of $\bar{W}$ are the components of the coordinate velocity [see (41) of I] and

$$B=\begin{pmatrix}\tilde{R}^0{}_0(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^0{}_0(1)\frac{s_0^3(t)}{3!} & \tilde{R}^0{}_1(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^0{}_1(1)\frac{s_0^3(t)}{3!} & \tilde{R}^0{}_2(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^0{}_2(1)\frac{s_0^3(t)}{3!} & \tilde{R}^0{}_3(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^0{}_3(1)\frac{s_0^3(t)}{3!}\\ \tilde{R}^1{}_0(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^1{}_0(1)\frac{s_0^3(t)}{3!} & \tilde{R}^1{}_1(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^1{}_1(1)\frac{s_0^3(t)}{3!} & \tilde{R}^1{}_2(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^1{}_2(1)\frac{s_0^3(t)}{3!} & \tilde{R}^1{}_3(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^1{}_3(1)\frac{s_0^3(t)}{3!}\\ \tilde{R}^2{}_0(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^2{}_0(1)\frac{s_0^3(t)}{3!} & \tilde{R}^2{}_1(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^2{}_1(1)\frac{s_0^3(t)}{3!} & \tilde{R}^2{}_2(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^2{}_2(1)\frac{s_0^3(t)}{3!} & \tilde{R}^2{}_3(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^2{}_3(1)\frac{s_0^3(t)}{3!}\\ \tilde{R}^3{}_0(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^3{}_0(1)\frac{s_0^3(t)}{3!} & \tilde{R}^3{}_1(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^3{}_1(1)\frac{s_0^3(t)}{3!} & \tilde{R}^3{}_2(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^3{}_2(1)\frac{s_0^3(t)}{3!} & \tilde{R}^3{}_3(0)\frac{s_0^2(t)}{2!}+2\tilde{R}^3{}_3(1)\frac{s_0^3(t)}{3!}\end{pmatrix} \tag{33}$$

Using the notation that $\tilde{\tilde{V}}$ is the representation of $\tilde{V}$ in the e_α basis set and $\tilde{V}$ is the representation of $\tilde{V}$ in the $\tilde{e}_\alpha$ basis set, we have from I that

$$\tilde{\tilde{V}}=\Omega(s_0(t),t)\tilde{V}, \tag{34}$$

where $\Omega(s_0(t),t)$ is given by (10) and (39) of I. We further note that

$$\bar{W}_{\tau_{pq}}=\Gamma^{-1}I\tilde{\tilde{V}}=\Gamma^{-1}\tilde{\tilde{V}}=\Gamma^{-1}\Omega(s_0(t),t)\tilde{V} \tag{35}$$

is the result we obtained from the parallel transport definition of observation in I [see (39) of I]. Thus the coordinate velocity based on the $\exp_{\gamma(t)*}$ definition of GRARO observation $\bar{W}_{\exp_{\gamma(t)*}}$ can be written in terms of the coordinate velocity obtained from the parallel transport definition of GRARO observation $\bar{W}_{\tau_{pq}}$, plus a correction as follows:

$$\bar{W}_{exp_{\gamma(t)*}}=\bar{W}_{\tau_{pq}}+\Gamma^{-1}B\tilde{\tilde{V}}. \tag{36}$$

If we let $\tilde{B}$ be the B matrix expressed with respect to the e_α basis set and note that $\bar{W}_{\exp_{\gamma(t)*}}$ and $\bar{W}_{\tau_{pq}}$ are to be expressed in the $\tilde{e}_\alpha$ basis set, we obtain (36) in a form given by (37) that can be compared directly with (39) of I, i.e.,

$$\bar{W}_{exp_{\gamma(t)*}}=\bar{W}_{\tau_{pq}}+\Gamma^{-1}\tilde{B}\Omega(s_0(t),t)\tilde{V}, \tag{37}$$

where $\bar{W}_{\tau_{pq}}$ is given by (39) of I;

$$\bar{W}^T_{\exp_{\gamma(t)*}}=(1,\dot{r}^1,\dot{r}^2,\dot{r}^3)_{|\gamma(t)}, \tag{38}$$

and

$$\tilde{B}=\begin{pmatrix}R^0{}_{ij0}(t,0)\frac{r^ir^j}{2!}+R^0{}_{ij0,k}(t,0)\frac{r^ir^jr^k}{3!} & R^0{}_{ij1}(t,0)\frac{r^ir^j}{2!}+R^0{}_{ij1,k}(t,0)\frac{r^ir^jr^k}{3!}\\ R^1{}_{ij0}(t,0)\frac{r^ir^j}{2!}+R^1{}_{ij0,k}(t,0)\frac{r^ir^jr^k}{3!} & R^1{}_{ij1}(t,0)\frac{r^ir^j}{2!}+R^1{}_{ij1,k}(t,0)\frac{r^ir^jr^k}{3!}\\ R^2{}_{ij0}(t,0)\frac{r^ir^j}{2!}+R^2{}_{ij0,k}(t,0)\frac{r^ir^jr^k}{3!} & R^2{}_{ij1}(t,0)\frac{r^ir^j}{2!}+R^2{}_{ij1,k}(t,0)\frac{r^ir^jr^k}{3!}\\ R^3{}_{ij0}(t,0)\frac{r^ir^j}{2!}+R^3{}_{ij0,k}(t,0)\frac{r^ir^jr^k}{3!} & R^3{}_{ij1}(t,0)\frac{r^ir^j}{2!}+R^3{}_{ij1,k}(t,0)\frac{r^ir^jr^k}{3!}\end{pmatrix}$$

$$\begin{pmatrix} R^{0}{}_{ij2}(t,0)\frac{r^i r^j}{2!} + R^{0}{}_{ij2,k}(t,0)\frac{r^i r^j r^k}{3!} & R^{0}{}_{ij3}(t,0)\frac{r^i r^j}{2!} + R^{0}{}_{ij3,k}(t,0)\frac{r^i r^j r^k}{3!} \\ R^{1}{}_{ij2}(t,0)\frac{r^i r^j}{2!} + R^{1}{}_{ij2,k}(t,0)\frac{r^i r^j r^k}{3!} & R^{1}{}_{ij3}(t,0)\frac{r^i r^j}{2!} + R^{1}{}_{ij3,k}(t,0)\frac{r^i r^j r^k}{3!} \\ R^{2}{}_{ij2}(t,0)\frac{r^i r^j}{2!} + R^{2}{}_{ij2,k}(t,0)\frac{r^i r^j r^k}{3!} & R^{2}{}_{ij3}(t,0)\frac{r^i r^j}{2!} + R^{2}{}_{ij3,k}(t,0)\frac{r^i r^j r^k}{3!} \\ R^{3}{}_{ij2}(t,0)\frac{r^i r^j}{2!} + R^{3}{}_{ij2,k}(t,0)\frac{r^i r^j r^k}{3!} & R^{3}{}_{ij3}(t,0)\frac{r^i r^j}{2!} + R^{3}{}_{ij3,k}(t,0)\frac{r^i r^j r^k}{3!} \end{pmatrix} \tag{39}$$

to fourth order and $j = 1, 2, 3$. The $R^{\alpha}{}_{ij\xi}$ are the Riemann curvature components in the $e_{\alpha}|_{\gamma(t)}$ basis. It is immediately apparent from (39) that the correction is of the order of the Riemann curvature coefficients.

The general expression for the coordinate acceleration is given by (18) of I. The coordinate acceleration obtained from the exp. definition of observation, $A_{\text{exp.}}$, can be written in terms of coordinate acceleration derived from the parallel transport definition of observation, $A_{\tau_{pq}}$, and a correction as

$$A_{\text{exp.}} \equiv \begin{pmatrix} 0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix}_{\substack{|\chi(t) \\ \text{exp.}}} = A_{\tau_{pq}} - \Gamma^{-2}\frac{d\Gamma}{dt}\tilde{B}\Omega(s_0(t), t)\tilde{V} + C, \tag{40}$$

where

$$A_{\tau_{pq}} = \Gamma^{-1}\Big(-\frac{d\Gamma}{dt}\,\overline{W}_{\tau_{pq}} + D\Big)_{|\chi(t)}, \tag{41}$$

$$C = -\Gamma^{-1}\begin{pmatrix} \Gamma^{0}{}_{\sigma\beta}(t, s_0(t))[\tilde{B}(s_0(t), t)\tilde{V}]^{\sigma}[\tilde{B}\Omega(s_0(t), t)\tilde{V}]^{\beta} \\ \Gamma^{1}{}_{\sigma\beta}(t, s_0(t))[\tilde{B}(s_0(t), t)\tilde{V}]^{\sigma}[\tilde{B}\Omega(s_0(t), t)\tilde{V}]^{\beta} \\ \Gamma^{2}{}_{\sigma\beta}(t, s_0(t))[\tilde{B}(s_0(t), t)\tilde{V}]^{\sigma}[\tilde{B}\Omega(s_0(t), t)\tilde{V}]^{\beta} \\ \Gamma^{3}{}_{\sigma\beta}(t, s_0(t))[\tilde{B}(s_0(t), t)\tilde{V}]^{\sigma}[\tilde{B}\Omega(s_0(t), t)\tilde{V}]^{\beta} \end{pmatrix}_{|\chi(t)}, \tag{42}$$

$\Gamma^{\mu}{}_{\sigma\beta}(t, s_0(t))$ is defined by (20) in I and $[\tilde{B}\Omega(s_0(t), t)\tilde{V}]^{\sigma}$ is the σ-component of the vector expression within the brackets. The correction terms for the coordinate acceleration are also of the order of magnitude of the Riemann curvature coefficients.

IV. CONCLUSIONS

In this paper we have used our definition of observation based on the mapping of tangent spaces into tangent spaces via exp. to obtain expressions for the coordinate velocity and coordinate acceleration. The results are stated in terms of the corresponding expression obtained from our parallel transport definition of observation and a correction term that is a function of the Riemann curvature tensor. Thus the difference between the formulas for the coordinate velocity and coordinate acceleration obtained from the two definitions of observation will be numerically small in most situations. This difference arises principally from the fact that exp. is only a radial isometry while the velocity vector of the geodesic particle is normal to the geodesic $\gamma_2(s)$. Finally, exp. is not defined through conjugate points. However, this presents no basic difficulty when null geodesic photons are employed to make the observations.

ACKNOWLEDGMENTS

All authors thank J. K. Beem, S. Harris, P. Ehrlich, and C. Ahlbrandt of the University of Missouri Mathematics Department for valuable conversations and comments.

[1]D. G. Retzloff, B. DeFacio, and P. W. Dennis, J. Math. Phys. **23**, 96 (1982).
[2]Y. Avishai and H. Ekstein, Commun. Math. Phys. **37**, 193 (1974).
[3]Avishai and Ekstein used acceleration covariance together with presymmetry to deduce the equivalence principle of general relativity.
[4]W. T. Reid, *Ordinary Differential Equations* (Wiley, New York, 1971).

Vacuum handles carrying angular momentum; electrovac handles carrying net charge [a)]

John L. Friedman and Steven Mayer
Department of Physics, University of Wisconsin, Milwaukee, Wisconsin 53201

(Received 9 February 1981; accepted for publication 24 July 1981)

Nonsimply-connected spacetimes can have locally defined Killing vectors that are globally defined only up to sign (pseudovectors). We show the existence of asymptotically flat vacuum spacetimes which are axisymmetric (have a rotational Killing pseudovector), are topologically trivial outside a spatially compact region, and which nevertheless have nonzero angular momentum. An analogous construction establishes the existence of source-free Einstein–Maxwell spacetimes which are topologically trivial outside a spatially compact region and which nevertheless carry nonzero net electric charge. The existence of such spacetimes leads to a new variant of the combined positive energy-cosmic-censorship conjecture: Given an asymptotically flat vacuum or electrovac initial data set which is axisymmetric and geodesically complete, the asymptotic mass, charge Q, and angular momentum J satisfy $m \geqslant [Q^2 + (J/m)^2]^{1/2}$.

PACS numbers: 04.20.Cv

I. INTRODUCTION

In a nonorientable spacetime, the duality invariance of the electromagnetic field is broken, because there is no globally defined, totally antisymmetric tensor. A remarkable consequence, pointed out by Sorkin,[1] is that the topological model for charge (a nontrivial spatial topology threaded by an electromagnetic field) predicts the nonexistence of magnetic charge if one assumes that:

(a) Spacetime admits a Cauchy surface M which is compact inside a neighborhood of spatial infinity.

(b) All prime factors (nontrivial topological structures) of M are microscopic. That is, handles do not join measurably distant points of M (points with, say, larger than nuclear separation), nor, by (a), do they bridge two asymptotically flat regions.

(c) The electromagnetic field is described by *some* antisymmetric tensor $f_{\alpha\beta}$ satisfying $\nabla_\alpha f^{\alpha\beta} = 0$, $\nabla_{[\alpha} f_{\beta\gamma]} = 0$.

A priori, one does not know which combination, $f \sin\alpha + {}^*f \cos\alpha$, to identify with the usual electromagnetic field tensor F, defined on the simply connected part $\hat{M}$ of M ($\hat{M}$ is obtained by excluding from M the interior of spheres containing its prime factors). However, the observed nonexistence of magnetic charge is the statement that for one choice of α the net flux $\oint_\sigma F_{\alpha\beta} dS^{\alpha\beta}$ vanishes through any sphere $\sigma \subset \hat{M}$.[2] If σ encloses an orientable factor of M, Stokes' theorem implies that both $\oint_\sigma f_{\alpha\beta} dS^{\alpha\beta}$ and $\oint_\sigma^* f_{\alpha\beta} dS^{\alpha\beta}$ vanish; there is no net charge of any kind. If the interior of σ is nonorientable, then an extension of Stokes' theorem[3] implies $\oint_\sigma^* f_{\alpha\beta} dS^{\alpha\beta}$ is still zero; but $\oint_\sigma f_{\alpha\beta} dS^{\alpha\beta}$ does *not* in general vanish. Thus the model conforms to experience: There is a choice ($\alpha = 0$) of α for which $\oint_\sigma F_{\alpha\beta} dS^{\alpha\beta}$ through any (larger-than-nuclear) sphere σ vanishes, and one can detect only electric charges until it becomes possible to probe the microscopic topology, to measure, for example, the flux through just one of the entrances to a handle.

There is a formal resemblance between the definition of charge as flux through a two surface and the definition of the angular momentum of an axisymmetric spacetime N with rotational Killing vector ϕ^α. The angular momentum J associated with ϕ^α may also be written as a surface integral, namely

$$J = \frac{1}{8\pi} \oint_{\sigma_\infty} \nabla^\alpha \phi^\beta \, dS_{\alpha\beta}, \tag{1}$$

where σ_∞ denotes a sphere at spatial infinity; if N is a vacuum spacetime with spatially compact interior—i.e., satisfying assumption (a) above—then J vanishes (even when N is nonorientable). We will find, however, that J can be nonzero if N has a somewhat weaker symmetry—a rotational Killing vector defined only up to sign. In Sec. II, below, we will establish the existence of vacuum spacetimes which have compact interior, are axisymmetric in this weakened sense, and whose angular momentum is nevertheless nonzero.

In Sec. III, we similarly show that nonorientable topologies carrying net charge arise as solutions to the Einstein–Maxwell equations. The proofs in Secs. II and III show the existence of asymptotically flat vacuum (or electrovac) initial data; that the Cauchy development will remain asymptotically flat is known for finite time evolutions.[4] In Sec. IV, on the basis of the cosmic censorship hypothesis and Gannon's[5] singularity theorem for topologically nontrivial spacetimes, we conjecture that any asymptotically flat, axisymmetric, electrovac spacetime must satisfy $m > (e^2 + J^2/m^2)^{1/2}$, where e, J, and m are the asymptotic charge, angular momentum, and mass, respectively.

Spacetime indices will be Greek, and spatial indices Latin. Our signature is $- + + +$, and we set $\nabla_{[a} \nabla_{b]} v^c = \frac{1}{2} R^c{}_{dab} v^d$ and $R_{ab} = R^c{}_{acb}$. Our notation for integrals has already been mentioned in footnote 2.

The portion of this work dealing with angular momentum in axisymmetric vacuum spacetimes is also discussed in S. Mayer's Ph.D. thesis.[6]

[a)]Work supported in part by the National Science Foundation under grant PHY 79 06657.

 0022-2488/82/010109-07$02.50

II. VACUUM HANDLES CARRYING ANGULAR MOMENTUM

Let N be an asymptotically flat spacetime with a rotational Killing vector ϕ^{α} defined at least on a neighborhood of spatial infinity. A spacelike hypersurface M of N will be said to have compact interior if it is complete and has only one asymptotic region—if, that is, it can be compactified by adjoining a single point at infinity. If N is a vacuum spacetime with topology $\mathbb{R}\times M$, where M has compact interior, and if ϕ^{α} is globally defined, then the angular momentum associated with ϕ^{α} vanishes

$$8\pi J=\int_{\sigma_{\infty}}\nabla^{\alpha}\phi^{\beta}\,dS_{\alpha\beta}=\int_{M}\nabla_{\alpha}\nabla^{\alpha}\phi^{\beta}\,dS_{\beta}=0. \qquad (2)$$

Stokes' theorem has been used to integrate by parts,[7] and the final equality follows from the Killing identity

$$\nabla_{\beta}\nabla^{\beta}\phi^{\alpha}=R^{\alpha}{}_{\beta}\phi^{\beta}, \qquad (3)$$

and from the vacuum field equation $R^{\alpha}{}_{\beta}=0$.

A nonsimply-connected spacetime may, however, have a weaker symmetry, which will be called a Killing pseudovector field.

Definition: A pseudovector field on N is an assignment to each point $p\in N$ of a pair, $\{\phi^{\alpha}(p),-\phi^{\alpha}(p)\}$, of vectors at p such that on any simply connected submanifold of M, ϕ^{α} is itself a smooth vector field. (Pseudotensor fields are analogously defined.) A pseudovector field is thus a vector field up to sign, and it gives rise to a true vector field on the universal covering space $\overline{N}$ of N. A Killing pseudovector ϕ^{α} is a pseudovector field that Lie derives the metric

$$\nabla_{(\alpha}\phi_{\beta)}=0. \qquad (4)$$

For example, if one constructs a Möbius strip by identifying left and right edges of a rectangle, a constant vector field parallel to those edges becomes a Killing pseudovector of the Möbius strip.

In what follows, the word axisymmetric will, for brevity, be retained to describe a spacetime having a rotational Killing pseudovector. The remainder of this section will be devoted to proving the existence of axisymmetric vacuum spacetimes whose spacelike hypersurfaces have compact interior and which nevertheless have nonzero angular momentum. Specifically, it will be shown that on some three-manifold M with compact interior, one can find axisymmetric, asymptotically flat initial data which satisfies the vacuum constraint equations and has nonzero angular momentum. We are then guaranteed a finite time evolution: a vacuum spacetime N in which M is isometrically embedded and which exhibits the axisymmetry and nonzero angular momentum of the initial data. This time development will also be asymptotically flat at least for some finite evolution.[4] Note that in a neighborhood of spatial infinity, M will have a true Killing vector and Eq. (1) will therefore continue to provide a well defined angular momentum.

A vacuum initial data set is a triple (M, g_{ab}, p_{ab}), where M is a three manifold, g_{ab} a positive definite metric on M, and p_{ab} a symmetric tensor on M (which will be the extrinsic curvature of M in the spacetime N evolved from the data), satisfying

$$D_b(p^{ab}-g^{ab}p)=0, \qquad (5a)$$

$$R-p_{ab}p^{ab}-p^2=0, \qquad (5b)$$

where D_a is the covariant derivative associated with g_{ab}, and $p\equiv p^{a}{}_{a}$.

The set (M, g_{ab}, p_{ab}) will be called axisymmetric if there is a rotational Killing pseudovector, ϕ^{a} on M which Lie derives both g_{ab} and p_{ab}. The angular momentum J can be written in terms of the initial data g_{ab}, p_{ab} as follows. Suppose that M is an axisymmetric spacelike hypersurface of an axisymmetric vacuum spacetime N and let t_{α} be the unit normal to M. Let σ_{∞} be any sphere in M enclosing all nontrivial topology [any sphere for which $(M-\operatorname{int}\sigma_{\infty})\simeq(\mathbb{R}^3-\text{a ball})$]. If n_{α} is the unit normal to σ_{∞},

$$\begin{aligned}8\pi J&=\int_{\sigma_{\infty}}\nabla_{\alpha}\phi_{\beta}t^{\alpha}n^{\beta}dS,\\&=\int_{\sigma_{\infty}}\phi^{\alpha}\nabla_{\alpha}t_{\beta}n^{\beta}dS,\\&=\int_{\sigma_{\infty}}p_{\alpha\beta}\phi^{\alpha}n^{\beta}dS,\end{aligned}$$

where $\pounds_{\phi}t^{\alpha}=0$ was used in the second equality. In this last form, J is expressed solely in terms of tensors on the three-manifold M (tensors orthogonal in all indices to n_{α}). Thus if ϕ_a, n_b, p_{ab}, and g_{ab} denote the pullbacks to M of $\phi_{\alpha}, n_{\alpha}, p_{\alpha\beta}$, and $g_{\alpha\beta}$, we have

$$8\pi J=\int_S p_{ab}\phi^{a}n^{b}\,dS. \qquad (6)$$

Finally, we recapitulate a definition of asymptotically flat initial data.[8] Let M be a compact manifold and let $\widetilde{M}=M-\{P\}$, P a point of M. Then an initial data set $(\widetilde{M}, \tilde{g}_{ab}, \tilde{p}_{ab})$ is asymptotically flat at spatial infinity if there exists a scalar field Ω and a metric g_{ab} on M such that

(i) Ω is smooth on $\widetilde{M}$ and C^2 at P, g_{ab} is smooth on $\widetilde{M}$ and C^0 at P, $g_{ab}=\Omega^2\tilde{g}_{ab}$ on $\widetilde{M}$, and

$$\Omega=0,\quad D_a\Omega=0,\quad D_aD_b\Omega=2g_{ab}\ \text{at}\ P,$$

where D_a is the covariant derivative associated with g_{ab};

(ii) $p_{ab}\equiv\Omega\tilde{p}_{ab}$ is bounded in a neighborhood of P;

(iii) the tensors $\Omega^{1/2}\widetilde{E}_{ab}$ and $\Omega^{1/2}\widetilde{B}_{ab}$ admit regular, direction-dependent limits at P, where (raising indices with $\tilde{g}^{ab}$ and using the covariant derivative $\widetilde{D}_a$ of $\tilde{g}_{ab}$),

$$\widetilde{E}_{ab}=\widetilde{R}_{ab}-\tilde{p}_{am}\tilde{p}_{bn}\tilde{g}^{mn}+\tilde{p}\tilde{p}_{ab},$$

$$\widetilde{B}_{ab}=\tilde{\epsilon}_{mn(a}\widetilde{D}^{m}\tilde{p}^{n}{}_{b)}$$

are parts of the (spacetime) Weyl tensor.

In our existence theorems we will begin with smooth fields g_{ab} and p_{ab} on the compact manifold M and will then construct a conformally related vacuum initial data set on $\widetilde{M}$, with

$$\tilde{g}_{ab}=\Psi^4 g_{ab}, \qquad (7)$$

$$\tilde{p}_{ab}=\Psi^2 p_{ab}. \qquad (8)$$

(Ψ is easier to use than $\Omega\equiv\Psi^{-2}$). Let r be the geodesic distance from P with respect to the metric g_{ab} on the compact manifold M. Then to prove asymptotic flatness of $(\widetilde{M}, \tilde{g}_{ab}, \tilde{p}_{ab})$ it suffices to show that (a) Ψ is smooth and positive on $\widetilde{M}$, (b) near P, Ψ has the form $\Psi=1/r+\phi$, where ϕ is continuous at P, and where (c) $D_a\,\phi$ and $r^2\,D_aD_b\phi$ have regular direction dependent limits at P.

In less esoteric terms, conditions (a)–(c) imply the existence of a chart $\tilde{x}^i$ on $\widetilde{M}$ (take $\tilde{x}^i = x^i/r^2$, where x^i are geodesic coordinates of g_{ab} about P) in which the components of $\tilde{g}_{ab}$ and $\tilde{p}_{ab}$ are smooth on $\widetilde{M}$ and have the asymptotic behavior

$$\tilde{g}_{ij} = \delta_{ij} + O(\tilde{r}^{-1}),\ \partial_k\tilde{g}_{ij} = O(\tilde{r}^{-2}),\ \partial_k\partial_l g_{ij} = O(\tilde{r}^{-3}),$$

$$\tilde{p}_{ij} = O(\tilde{r}^{-2}),\ \partial_k\tilde{p}_{ij} = O(\tilde{r}^{-3}).$$

The point P of M, of course, plays the role of the "point at infinity" of $\widetilde{M}$.

Theorem 1: On some manifolds $\widetilde{M}$ with compact interior, there exist axisymmetric vacuum initial data sets ($\widetilde{M}$, $\tilde{g}_{ab}$,$\tilde{p}_{ab}$) which are asymptotically flat and have nonzero angular momentum.

Proof: We first establish three lemmas.

Lemma 1: Let M be a compact three-manifold with positive definite metric g_{ab} whose scalar curvature is everywhere positive. Then there is a conformally related, asymptotically flat initial data set $(\widetilde{M}, \tilde{g}_{ab}, 0)$, where $\widetilde{M} = M - \{P\}$, some $P\in M$. (This result was suggested by Geroch.[9] Its implication—the fact that to any compact three-manifold M admitting a metric with positive scalar curvature R corresponds an asymptotically flat space $(\widetilde{M}, g_{ab})$ with vanishing R—is mentioned by Schoen and Yau.)[10]

Proof: Since $p_{ab} = 0$, the momentum constraint (5a) is automatically satisfied. Under a conformal transformation of the form (7), the Ricci scalar becomes

$$\widetilde{R} = -8\Psi^{-5}(D^2 - \tfrac{1}{8}R)\Psi, \tag{9}$$

whence the Hamiltonian constraint (5b) will be satisfied if

$$\Theta\Psi \equiv (-D^2 + \tfrac{1}{8}R)\Psi = 0 \tag{10}$$

on $\widetilde{M}$. If, following Geroch,[9] we set $\Theta\Psi = \delta_p$ on M, with δ_p the covariant δ function at P, then we will see below that $\Psi > \epsilon > 0$ everywhere (thus $\tilde{g}_{ab}$ will be positive definite) and that asymptotic flatness of $(\widetilde{M}, \tilde{g}_{ab})$ will be guaranteed as well. Now $\delta_p\in H_{-2}(M)$, where $H_n(M)$ is the nth Sobolev space.[11] Because R is positive definite, Θ is positive definite and $\ker\Theta = 0$. Because Θ is symmetric $(\Theta' = \Theta)$ the index of Θ,

$$\mathrm{ind}\Theta = \dim\ker\Theta - \dim\ker\Theta',$$

vanishes as well. Then by a theorem due to Seeley[12] Θ is an isomorphism from $H_k(M)$ to $H_{k-2}(M)$ for all integers k. In particular, there is a Ψ in $H_0(M)$ for which $\Theta\Psi = \delta_p$. Furthermore, because Θ is C^∞ and elliptic on M, Ψ is C^∞ on $\widetilde{M}$ by the local smoothness of solutions to elliptic equations.[13]

To prove (b) and (c), pick Riemann normal coordinates x^i about P [so that $x^i(P) = 0$] and write Ψ in the form

$$\begin{aligned}\Psi = {} & \frac{1}{r} + \frac{1}{12}\Big(R_{ij} - \frac{1}{4}\delta_{ij}R\Big)(P)\frac{x^ix^j}{r} \\ & + \frac{1}{24}\Big(\nabla_iR_{jk} - \nabla_lR^l{}_i\delta_{jk} - \frac{1}{4}\nabla_iR\delta_{jk}\Big)\frac{x^ix^jx^k}{r} \\ & + \psi,\end{aligned} \tag{11}$$

where $r = \Big[\sum_i (x^i)^2\Big]^{1/2}$. Then, using

$$\nabla^2(1/r) = \delta^3(x) = \delta_p(x),$$

we find that $\Theta\Psi = \delta_p$ has the form

$$\Theta\psi = A_{ij}\frac{x^ix^j}{r} + B_{ijkl}\frac{x^ix^jx^kx^l}{r^3} + C_{ijklmn}\frac{x^i\ldots x^n}{r^5} + D, \tag{12}$$

where A_{ij},B_{ijkl},C_{ijklmn}, and D are smooth. The right-hand side of (12) is C^α,$\alpha > 0$ and all the coefficients in Θ are smooth. Thus ψ is C^2 and Eq. (11) implies that the requirements (b) and (c) for asymptotic flatness at spatial infinity are satisfied.

Finally, $\Psi > \epsilon > 0$ follows by noting first that Ψ is positive in a neighborhood U of P because $\Psi \sim 1/r$. Then Ψ cannot be negative on the compact set $\overline{M - U}$, for if it were, it would have a minimum at some point q and there $R\Psi < 0$, $-D^2\Psi \leqslant 0$, contradicting (10). It now follows from (10) that if $\Psi = 0$ at some q, all derivatives of Ψ vanish there as well, and by a theorem of Aronszajn,[14] Ψ would then have to vanish everywhere, contradicting $\Psi > 0$ on N. Thus $\Psi > \epsilon > 0$ everywhere. □

Lemma 2: Let A and B be Banach spaces and T_λ a family of continuously differentiable maps of a neighborhood V of $a_0\in A$ into B. Let $T'_0(a_0)$ be an isomorphism of A onto B and suppose λ continuously parameterizes T_λ and T'_λ. Then for sufficiently small λ, there is an open set $U\subset V$ such that the restriction of T_λ to U is a homeomorphism onto an open neighborhood of $T_0(a_0)$ in B.

Proof: The proof is virtually identical to that of the implicit function theorem. For sufficiently small $\lambda\in\mathbb{R}$, $b\in B$, T'_λ is an isomorphism and the map

$$S(a) = -(T'_\lambda)^{-1}[T_\lambda(a + a_0) - T_0(a_0) - b] + a \tag{13}$$

is a contraction in a neighborhood N of $O\in A$. Hence S has a unique fixed point, an $a\in N$ for which $S(a) = a$. This a satisfies $T_\lambda(a + a_0) = T_0(a_0) + b$. Thus $a + a_0$ is the unique solution in the neighborhood $a_0 + N$ of a_0 to the equation $T_\lambda(a) = T_0(a_0) + b$.

Lemma 3: Let M, g_{ab} be as in Lemma 1 and let p_{ab} be a smooth tensor field on M with $D_bp^{ab} = 0$, $p^a{}_a = 0$. Then for sufficiently small real λ, there is an asymptotically flat initial data set $(\widetilde{M}, \tilde{g}_{ab}(\lambda), \lambda\tilde{p}_{ab})$, where $\widetilde{M}$ and $\tilde{g}_{ab}(\lambda) = \Psi^4 g_{ab}$ are as in Lemma 1, and where

$$\tilde{p}_{ab} = \Psi^{-2}p_{ab}. \tag{14}$$

Proof: Because $\widetilde{D}_b\tilde{p}_a{}^b = \Psi^{-6}D_bp_a{}^b$, the momentum constraint is automatically satisfied. As in Lemma 1, we need only show the existence of a Ψ_λ satisfying conditions (a)–(c) and for which $\widetilde{R}$ satisfies the constraint (5b) on $\widetilde{M}$:

$$\widetilde{R} - \tilde{p}_{ab}\tilde{p}^{ab} = 0,$$

or

$$\Theta_\lambda\Psi_\lambda \equiv (D^2 - \tfrac{1}{8}R)\Psi_\lambda + \tfrac{1}{8}\lambda^2p^{ab}p_{ab}\Psi_\lambda^{-7} = 0. \tag{15}$$

Again we seek a solution to

$$\Theta_\lambda\Psi_\lambda = \delta_P$$

on M. If we write $\Psi_\lambda = \Psi_0 + \psi_\lambda$, with Ψ_0 the (unique) solution to $\Theta_0\Psi_0 = \delta_p$, Eq. (16) becomes an elliptic equation for ψ_λ, namely

$$\theta_\lambda\psi_\lambda \equiv \Theta_0\psi_\lambda + (\lambda^2/8)p^{ab}p_{ab}(\Psi_0 + \psi_\lambda)^{-7} = 0.$$

For ψ_λ in $H_2(M)$ with $\|\psi_\lambda\|_2$ sufficiently small, $\Psi_0 + \psi_\lambda$ is bounded away from 0. Therefore θ_λ maps to $H_0(M)$ a neighborhood of 0 in $H_2(M)$. Moreover, θ_λ is continuously differentiable at ψ_λ for ψ_λ near 0 in $H_2(M)$, with derivative the linear map $\theta'_\lambda: H_2(M)\to H_0(M)$ given by

$$\theta'_\lambda: \phi\to[\Theta_0 - \tfrac{7}{8}\lambda^2 p^{ab}p_{ab}(\Psi_0+\psi_\lambda)^{-8}]\phi.$$

For $\lambda = 0, \theta_\lambda = \Theta_0$ is an isomorphism from $H_2(M)$ to $H_0(M)$, whence, by Lemma 2, $\theta_\lambda\psi_\lambda = 0$ has a solution $\psi_\lambda\in H_2(M)$ for λ sufficiently small. By the construction of the proof of Lemma 2, ψ_λ will be close to 0 in $H_2(M)$ and so $\Psi_\lambda = \Psi_0 + \psi_\lambda$ will be positive on $\widetilde{M}$ for small λ. By an argument analogous to that used in the proof of Lemma 1, Ψ_λ will behave near P like $1/r + \phi$, ϕ satisfying (c). Finally, to show that Ψ_λ is smooth on $\widetilde{M}$, we proceed by induction on the differentiability index k. Let Ω be any smooth submanifold with $\overline{\Omega}$ compact in $\widetilde{M}$ and suppose ψ_λ is in $H_k(\overline{\Omega})$. Then because $\Psi_0 + \psi_\lambda > \epsilon > 0$ on $\overline{\Omega}$ [and $\Psi_0\in C^\infty(\overline{\Omega})$], $p^{ab}p_{ab}(\Psi_0+\psi_\lambda)^{-7}\in H_k(\overline{\Omega})$, $k\geqslant 2$.[15] Because ψ_λ satisfies on $\overline{\Omega}$ the elliptic equation

$$\Theta_0\psi_\lambda = \lambda^2 p_{ab}p^{ab}(\Psi_0+\psi_\lambda)^{-7},$$

with right-hand side in $H_k(\overline{\Omega})$, $\psi_\lambda\in H_{k+2}(\overline{\Omega})$. By construction, ψ_λ is in $H_2(\overline{\Omega})$, whence $\psi_\lambda\in H_k(\overline{\Omega})$, all k; that is, $\psi_\lambda\in C^\infty(\overline{\Omega})$. Finally, since Ω was arbitrary we conclude that $\Psi_\lambda = \Psi_0 + \psi_\lambda$ is in $C^\infty(\widetilde{M})$. □

To prove Theorem 1, we will pick a particular compact manifold M, a metric g_{ab}, with a rotational Killing pseudovector and with positive scalar curvature R. We will then find an axisymmetric divergence-free tensor p_{ab} that yields nonzero angular momentum on a conformally related asymptotically flat manifold.

Let $\mathscr{C}$ be the cylinder of constant curvature $R = 2/a^2$, constructed as a product $S^2\times I$ of the metric sphere with the closed interval $[-1,1]$. Let ϕ^a be a rotational Killing vector of $\mathscr{C}$ and let $\phi\in[0,2\pi]$ parameterize its orbits. M is constructed by identifying the boundary spheres $\sigma_- = S^2\times\{-1\}$ and $\sigma_+ = S^2\times\{1\}$ after an inversion. In terms of the natural cylindrical coordinates (θ,ϕ,z), the identification is (see Fig. 1)

$$(\theta,\phi,-1)\to(\pi-\theta,2\pi-\phi,1). \tag{16}$$

Then the induced metric g_{ab} of M is smooth with constant curvature $R = 2/a^2$ and Killing pseudovector $\pm\phi^a$.

Let

$$p_{ab} = \omega_{(a}\phi_{b)}, \tag{17}$$

where ω_a is a pseudovector on M Lie derived by ϕ^a,

$$\pounds_\phi\omega_a = 0. \tag{18}$$

The projection of the momentum constraint (5a) orthogonal to ϕ^a,

$$K^a{}_c D_b p^{bc} = 0,$$

where $K^a{}_b = \delta^a{}_b - \phi^a\phi_b(\phi^c\phi_c)^{-1}$, can be written in the form

$$K^a{}_c\omega^b D_a(\phi^d\phi_d)D_b\phi = 0, \tag{19}$$

where Killing's equation $(D_{(a}\phi_{b)} = 0)$ and the identity

$$\phi_a = \phi^b\phi_b D_a\phi$$

have been used (ϕ is a pseudoscalar on M). The projection of (5a) along ϕ^a takes the form

$$D_a(\omega^a\phi_b\phi^b) = 0, \tag{20}$$

after Killing's equation and (18) are used. Then, by choosing

$$\omega^a = \epsilon^{abc}D_b\phi\chi_c(\phi_d\phi^d)^{-1}, \tag{21}$$

where χ_c is a curl-free vector field $(D_{[a}\chi_{b]} = 0)$, both projections (19) and (20) of Eq. (5a) will be satisfied. The pseudovector ω^a will be axisymmetric and smooth if $\pounds_\phi\chi_a = 0$ and if χ_a vanishes sufficiently fast as $\phi_a\phi^a\to 0$, i.e., near the axis of symmetry (we find such a χ_a below).

Given such a χ_a, Lemma 3 now guarantees the existence of an asymptotically flat initial data set $(\widetilde{M},\tilde{g}_{ab},\tilde{p}_{ab})$, with $\tilde{p}_{ab} = \Psi^{-2}p_{ab}$, $\tilde{g}_{ab} = \Psi^4 g_{ab}$, for a scalar Ψ defined on $\widetilde{M} = M - \{P\}$. By choosing P to lie on the axis of symmetry, we make the conformal factor Ψ axisymmetric.[16] Consequently $(\widetilde{M},\tilde{g}_{ab},\tilde{p}_{ab})$ will be axisymmetric with Killing pseudovector ϕ^a.

The corresponding angular momentum J is conformally invariant:

$$8\pi J = \int_{\sigma_\infty}\tilde{p}_a{}^b\phi^a\, d\tilde{S}_b = \int_{\sigma_\infty} p_a{}^b\phi^a\, dS_b,$$

where σ_∞ is any sphere in $\widetilde{M}$ enclosing the handle (equivalently, as seen from the other side, σ_∞ is any sphere enclosing P). To evaluate J, we will use the construction of M from the cylinder $\mathscr{C}$ with boundary spheres σ_+ and σ_- identified. Let A_+, B_+, A_-, B_-, be the points of σ_+ and σ_- that lie on the symmetry axis [so that in the natural cylindrical coordinates, A_+, is $(z = +1, \theta = 0)$, B_+ is $(z = +1, \theta = \pi)$, A_- is $(z = -1, \theta = \pi)$, and B_- is $(z = -1, \theta = 0)$]. In M, A_+ and A_- (and B_+ and B_-) are identified (see Fig. 1).

Claim: Let χ be any scalar on $\mathscr{C} - \{P\}$ with $D_c\chi = \chi_c$,

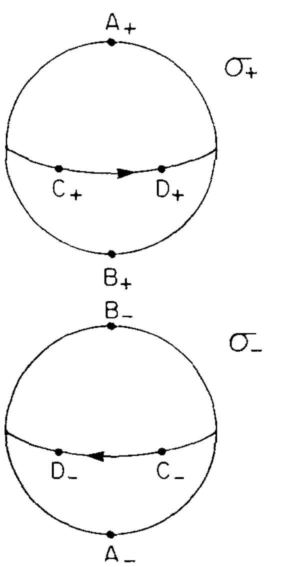


FIG. 1. Boundary spheres $\sigma_\pm$ of the cylinder $S^2\times[-1,1]$ are represented here. The manifold M of Sec. II is constructed by identifying σ_+ and σ_- in such a way that points labeled by the same letter (e.g., A_+ and A_-) are identified.

where χ_c is the vector field introduced in Eq. (21), and suppose that χ is constant in a neighborhood of each piece of the symmetry axis. Then

$$J = \tfrac{1}{8}[\chi(A_-) - \chi(B_-) + \chi(A_+) - \chi(B_+)]. \qquad (22)$$

To prove the claim, notice that Stokes' theorem implies

$$J = \int_{\sigma_\infty} p^{ab}\phi_a \, dS_b = -\int_{\sigma_+ + \sigma_-} p^{ab}\phi_a \, dS_b, \qquad (23)$$

after Killing's equation and (5a) have been used to eliminate the volume integral. Here $dS_b = n_b dS$, where n_b is the outward normal, $\pm \nabla_a z$, on $\sigma_\pm$; the minus sign in Eq. (23) comes from the fact that the outward normal to a sphere at infinity used to define J is an inward normal when σ_∞ is regarded as a sphere enclosing P in $S^2 \times I$. Equations (20) and (21) now imply

$$\begin{aligned} 8\pi J &= -\int_{\sigma_+ + \sigma_-} \tfrac{1}{2}[\phi^a \epsilon^{bcd} D_c \phi D_d \chi (\phi_m \phi^m)^{-1}]\phi_a \, dS_b, \\ &= -\Big(\int_{\sigma_+} - \int_{\sigma_-}\Big) D_{[a}\phi D_{b]}\chi \, dS^{ab}, \\ &= \Big(\int_{\sigma_+} - \int_{\sigma_-}\Big) D_{[a}(\chi D_{b]}\phi) \, dS^{ab}. \end{aligned} \qquad (24)$$

We can restrict the integral in (24) to parts of σ_+ and σ_-, where $D_c\chi \neq 0$: let $\hat\sigma_+$ and $\hat\sigma_-$ be spheres σ_+ and σ_- with small disks about the axis removed. Then

$$\begin{aligned} 8\pi J &= \Big(\int_{\hat\sigma_+} - \int_{\hat\sigma_-}\Big) D_{[a}(\chi D_{b]}\phi) \, dS^{ab}, \\ &= \frac{1}{2}\Big(\int_{\partial\hat\sigma_+} - \int_{\partial\hat\sigma_-}\Big) \chi D_a \phi \, dl^a; \end{aligned}$$

$$J = \tfrac{1}{8}[\chi(A_+) - \chi(B_+) + \chi(A_-) - \chi(B_-)],$$

as claimed.

To conclude the proof of our main theorem, we need only show that we can choose a χ on $C - \{P\}$ such that $D_c\chi$ is smooth on $\tilde M$ and J is nonzero. But this is easy: Choose a constant $j \neq 0$, and a function $f(\theta)$ with $f(\theta) = 0$ near $\theta = 0$ and $f(\theta) = -(1/2\pi)j$ near $\theta = \pi$; let χ be smooth with

$$\chi(\theta,\phi,z) = \begin{cases} f(\theta) \text{ in a neighborhood of } \sigma_+, \\ f(\theta) + \dfrac{1}{2\pi} j \text{ in a neighborhood of } \sigma_-, \end{cases}$$

and with χ constant on a neighborhood of each piece of the symmetry axis. Then $D_c\chi$ is smooth on $\tilde M$ and $J = j$.□

Expressions of the form (24) are valid under more general circumstances. For stationary, axisymmetric vacuum spacetimes with Killing vectors ϕ^a and t^a (e.g., the exterior of a rotating star, or the Kerr geometry), one can define a scalar t for which $\nabla_a t$ is in the t^a–ϕ^a plane, $t^a\nabla_a t = 1$, and $\phi^a \nabla_a t = 0$. On a $t = $ const surface, p_{ab} and ω_a have the forms (17) and (21), respectively, with

$$\chi_a = D_a \chi = \phi\cdot\phi(-g^{mn}\nabla_m t \nabla_n t)^{-1/2} \epsilon_{abc} D^b \Big(\frac{t\cdot\phi}{\phi\cdot\phi}\Big)\phi^c.$$

χ is constant along the upper and lower symmetry axes and has the asymptotic form

$$\chi \sim 2J\cos\theta(3 - \cos^2\theta) + O(1/r),$$

and

$$J = \tfrac{1}{8}[\chi(\theta = 0) - \chi(\theta = \pi)].$$

III. ELECTROVAC HANDLES CARRYING NET CHARGE

An initial data set for the Einstein–Maxwell equations is a quintuple $(M, g_{ab}, p_{ab}, E^a, B^a)$, where E^a and B^a are divergence-free vector fields,

$$D_a E^a = D_a B^a = 0, \qquad (25)$$

and where

$$D_b(p^{ab} - g^{ab}p) = 2\epsilon^{abc}E_b B_c, \qquad (26)$$

$$R - p^{ab}p_{ab} + p^2 = 2(E^2 + B^2). \qquad (27)$$

On a nonorientable spacetime exactly one combination, $E^a \sin\alpha + B^a \cos\alpha$, is a true vector. If we call the vector B^a, then E^a is a pseudovector (an axial vector, in this case), and by the generalization of Stokes' theorem, Eq. (25) implies that $\oint B^a \, dS_a = 0$ for any sphere σ enclosing a prime factor of M. In general, however $\oint_\sigma E^a \, dS_a \neq 0$ when σ encloses a nonorientable prime factor; we want to show that such electric fields with asymptotic charge can arise as solutions to the Einstein–Maxwell equations on spaces with compact spatial interior. We take $B^a = 0$ and $p_{ab} = 0$ and prove

Theorem 2: On some manifolds M with compact interior there exist asymptotically flat Einstein–Maxwell initial data sets $(\tilde M, \tilde g_{ab}, \tilde E^a)$ which have nonzero asymptotic charge.

To the previous definition of asymptotic flatness we add the requirement that $\tilde E_a$, regarded as a tensor field on M, have a direction-dependent limit at P; that is, $\tilde E_a \sim 1/\tilde r^2$ on $\tilde M$.

The proof of Theorem 2 is essentially the following

Lemma 4: Let M be a compact three-manifold (not necessarily orientable) with positive definite metric g_{ab} whose scalar curvature is everywhere positive. Let E^a be an axial vector field on $\tilde M = M - \{P\}$, some $P \in M$, with $D_a E^a = 0$ on $\tilde M$. Suppose that in a geodesic chart $\{x^i\}$ about P,

$$E^i = (x^i/r^3)(1 + A^i{}_{jk} x^j x^k), \qquad (28)$$

some smooth $A^i{}_{jk}$. Then for sufficiently small real e, there is an asymptotically flat initial data set $(\tilde M, \tilde g_{ab}, e\tilde E^a)$, where $\tilde M = M - \{P\}$, some $P \in M$, where $\tilde g_{ab} = \Psi_e^4 g_{ab}$, and where $\tilde E_a = \Psi_e^{-2} E_a$.

Proof of Lemma 4: Because $D_b \tilde E^b = \Psi_e^{-6} D_b E^b$, $\tilde E^a$ is divergence-free. We need only show the existence of a Ψ_e satisfying

$$\Theta_e \Psi_e \equiv (D^2 - \tfrac{1}{8}R)\Psi_e + \tfrac{1}{4}e^2 g^{ab}E_a E_b \Psi_e^{-3} = \delta_P, \qquad (29)$$

where Ψ_e is smooth and positive on $\tilde M$ and has the form $1/r + \phi$ near P, some ϕ satisfying condition (c).

Just to show that $\Theta_0 : H_r \to H_{r-2}$ is an isomorphism requires some discussion, because the Fredholm alternative seems to be established in the literature only for orientable manifolds. Let $\hat M \xrightarrow{\pi} M$ be the orientable, two-sheeted cover of M. Because π is a local isomorphism, Θ_0 determines an elliptic operator $\hat\Theta_0$ on $\hat M$. $\hat\Theta_0$ is again symmetric and positive definite, so $\ker \hat\Theta_0 = 0$, $\operatorname{ind} \hat\Theta_0 = 0$, and thus $\hat\Theta_0$ is an isomorphism $H_r(\hat M) \to H_{r-2}(\hat M)$. Now the pullback to $\hat M$ of any $f \in H_{r-2}(\hat{M})$ is a function $f = \hat{f} \circ \pi$ invariant under the involution $I: \hat{M} \rightarrow \hat{M}$ (I interchanges the inverse images $\{p_1, p_2\}$ of each $p \in M$). Since $\hat{\Theta}_0$ is also invariant under I, if $\hat{\Psi}$ satisfies $\hat{\Theta}_0 \hat{\Psi} = \hat{f}$ so does $\hat{\Psi} \circ I$. But $\hat{\Theta}_0$ an isomorphism means $\hat{\Psi}$ is unique: $\hat{\Psi}_0 \circ I = \hat{\Psi}_0$. Thus, there is a (unique) $\Psi \in H_r(M)$ with $\hat{\Psi} = \Psi \circ \pi$, and Θ_0 is an isomorphism.

Let

$$\Psi_e = \Psi_0 - (e^2/8) r + \psi_e, \tag{30}$$

where Ψ_0 is the unique solution to

$$\Theta_0 \Psi_0 = \delta_P \tag{31}$$

and r is any smooth scalar on M that agrees near P with $r = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ (defined in terms of the geodesic chart $\{x^i\}$). The symbols $f_0, \ldots, f_8$ will denote scalars on M that are smooth on $\widetilde{M}$ and C^α at P. From (28) and (30), and using $\Psi_0 = 1/r + f_0$, we find

$$E^2 \Psi_e^{-3} = \frac{1}{r} + f_1 + r \frac{f_3 \psi_e + f_4 \psi_e^2 + f_5 \psi_e^3}{(1 + r f_6 + r \psi_e)^3}.$$

Then, using $\Theta_0 r = 2/r + f_7$, we have

$$\begin{aligned} \theta_e \psi_e &\equiv \Theta_e \Psi_e - \delta_P \\ &= \Theta_0 \psi_e + \frac{1}{4} e^2 r \frac{f_3 \psi_e + f_4 \psi_e^2 + f_5 \psi_e^3}{(1 + r f_6 + r \psi_e)^3} + e^2 f_8 \end{aligned}$$

and (29) becomes $\theta_e \psi_e = 0$. Because $f_i \in C^0(M)$, θ_e maps a neighborhood of 0 in $H_2(M)$ to $H_0(M)$. Because Θ_0 is an isomorphism, $\theta'_e|_{\psi_e = 0}: H_2(M) \rightarrow H_0(M)$ is an isomorphism for e sufficiently small. By Lemma 2, θ_e is an isomorphism from a neighborhood of $0 \in H_2(M)$ to a neighborhood of $0 \in H_0(M)$. Then for small e there is a ψ_e in $H_2(M)$ satisfying $\theta \psi_e = 0$, and ψ_e is smooth on $\widetilde{M}$ (as in the proof of Lemma 3). Thus the conformal factor $\Psi_e = \Psi_0 - (e^2/8) r + \psi_e$ is smooth on $\widetilde{M}$. Because $\psi_e \in H_2(M) \subset C^\alpha(M)$, $\theta_0 \psi_e \in C^\alpha(M)$; hence $\psi_e \in C^{2+\alpha}(M)$ and $\Psi_e = 1/r + \phi$, ϕ satisfying (c). Finally, since $\Psi_0 > \epsilon > 0$ on $\widetilde{M}$, so also is Ψ_e, for e sufficiently small. (Note that with this Ψ_e, and with E_a defined by (28), $\widetilde{E}_a = \Psi_e^{-2} E_a$ has a regular, direction-dependent limit at P, as required.)□

Proof of Theorem 2: We again construct M from the cylinder $S^2 \times I$ with its natural metric, but this time we make the nonorientable identification of the boundary spheres $\sigma_\pm$

$$(\theta, \phi, -1) \rightarrow (\pi - \theta, \phi, 1).$$

We now define a vector field E^a on $S^2 \times I - \{P\}$, where P is the point $z = 0$, $\theta = 0$. Let α be any smooth axisymmetric scalar on $S^2 \times I - \{P\}$ satisfying

$$\alpha = \begin{cases} ez(z^2 + a^2\theta^2)^{-1/2}, & \text{in a neighborhood of } P \\ \pm e(1 + \cos\theta), & \text{in neighborhoods of } \sigma_\pm \\ \pm e/2, & \theta = 0, z \gtrless 0 \\ 0, & \theta = \pi, \end{cases}$$

and $\nabla \alpha = 0(\sin^2\theta)$ near the symmetry axes ($\theta = 0, \pi$); (it is easy to verify that such α's exist). Then $E^a = \epsilon^{abc} D_b \phi D_c \alpha$ is smooth and divergence-free on $S^2 \times I - \{P\}$, and its flux through a sphere σ_∞ enclosing P is

$$\int_{\sigma_\infty} E^a \, dS_a = (\alpha_{\text{top}} - \alpha_{\text{bottom}}) 2\pi = 4\pi e, \tag{32}$$

while

$$\int_{\sigma_\pm} E^a dS_a = 2\pi e,$$

where the outward normal $\pm \nabla_a z$ is taken in each case. Moreover, $E^a = \pm (e/2a^2) \nabla^a z$ near $\sigma_\pm$, so E^a is a smooth axial vector field on $\hat{M} = M - \{P\}$. Finally, in a geodesic chart about P, E^i has the form (28).

Lemma 4 now guarantees an asymptotically flat initial data set $(\hat{M}, \tilde{g}_{ab}, \widetilde{E}_a)$, and we need only check that $\int_{\sigma_\infty} \widetilde{E}^a d\widetilde{S}_a \neq 0$. But, like angular momentum, the charge is conformally invariant, $\oint \widetilde{E}^a \, d\widetilde{S}_a = \oint E^a dS_a$, whence, by (32),

$$Q \equiv \frac{-1}{4\pi} \oint \widetilde{E}^a \, d\widetilde{S}_a = -e.$$

(The charge seen at ∞ is $-e$ because flux emerging from P is flux coming in from ∞.)□

IV. A CONJECTURE

A theorem due to Gannon[6] shows that any initial data sets of the type constructed here (asymptotically flat with nontrivial topology) evolve to geodesically incomplete spacetimes. If the cosmic censorship hypothesis is true, the ultimate fate of such initial data will be a set of charged, rotating black holes, and these black holes will satisfy the inequalities

$$m_i^2 \geqslant Q_i^2 + J_i^2/m_i^2. \tag{33}$$

The time evolution of axisymmetric electrovac initial data with angular momentum J and net charge Q conserves both J and Q, while the total mass m can only decrease. By (33), the final mass will then satisfy

$$m_f = \sum m_i \geqslant \sum (Q_i^2 + J_i^2/m_i^2)^{1/2} \geqslant (Q^2 + J^2/m^2)^{1/2}.$$

We therefore conjecture that all complete, axisymmetric, asymptotically flat electrovac initial data sets with net charge Q, angular momentum J, and mass m, satisfy $m \geqslant (Q^2 + J^2/m^2)^{1/2}$.

An analogous conjecture for initial data sets with trapped surfaces was first suggested by Penrose and has been discussed by several subsequent authors.[17]

ACKNOWLEDGMENTS

We would like to thank R. Sorkin for helpful discussions and M. Cantor for pointing out an error in an earlier draft of this paper and for comments that improved the clarity of our presentation.

[1] R. Sorkin, J. Phys. A **10**, 717 (1978); **12**, 403 (1979).

[2] The integral of a two-form $\omega_{\alpha\beta}$ over a two-surface S will be written $\int_S \omega_{\alpha\beta} dS^{\alpha\beta}$. Let N be a four-manifold with volume form $\epsilon_{\alpha\beta\gamma\delta}$; let $M \subset N$ be a three-manifold with induced volume form $\epsilon_{\alpha\beta\gamma}$; and let $S \subset M$ be a two-manifold with induced volume form $\epsilon_{\alpha\beta}$. We will write

$$dS = \epsilon_{\alpha\beta} dS^{\alpha\beta}, \; dS_\alpha = \epsilon_{\alpha\beta\gamma} dS^{\beta\gamma}, \text{ and } dS_{\alpha\beta} = \tfrac{1}{2} \epsilon_{\alpha\beta\gamma\delta} dS^{\gamma\delta}.$$

Finally, the line integral of a one-form ω_α along a curve c will be written $\int_c \omega_\alpha dl^\alpha$.

[3] If $f_{\alpha\beta}$ is an antisymmetric tensor, then $\oint_\sigma \epsilon_{\alpha\beta}{}^{\gamma\sigma} f_{\gamma\delta} dS^{\alpha\beta} = 0$, where $\epsilon_{\alpha\beta\gamma\delta}$ is defined in a neighborhood of any sphere σ enclosing only prime factors of M (see Ref. 1).

[4] Y. Choquet-Bruhat and D. Christodoulou, Acta. Math. **146**, 129 (1981); D. Christodoulou, J. Math. Pures Appl. **60**, 99 (1981).

[5] D. Gannon, J. Math. Phys. **16**, 2364 (1975).

[6]S. Mayer, Ph.D. thesis, University of Wisconsin, Milwaukee, Wisconsin 1979.

[7]Stokes' theorem in this form is valid even when M is nonorientable as long as ϕ^α is a true (i.e., polar) vector field.[1]

[8]R. Geroch, in *Asymptotic Structure of Space-Time*, edited by P. Esposito and L. Witten (Plenum, New York, 1976); A. Ashtekhar, in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980), Vol. 2.

[9]R. Geroch, "The positive-mass conjecture," in *Theoretical Principles in Astrophysics and Relativity*, edited by N. R. Lebovitz (Univ. of Chicago, Chicago, 1978).

[10]R. Schoen and S.-T. Yau, Phys. Rev. Lett. **43**, 1457 (1979).

[11]See R. S. Palais, *Seminar on the Atiyah–Singer Index Theorem* (Princeton U. P., Princeton, N.J., 1965). In a chart $U \rightarrow M$, $U \subset \mathbb{R}^3$, $H_n(M)$ agrees with the restriction to U of $H_n(\mathbb{R}^3)$, and $H_n(\mathbb{R}^3)$ is the completion of C_0^∞ functions in the norm $\|f\|_n = \int |\hat{f}(\xi)|^2(1+\xi^2)^{n/2}\, d^3\xi$, where $\hat{f}$ is the Fourier transform of f.

[12]R. T. Seeley, Trans. Am. Math. Soc. **117**, 167 (1965) (Theorem 8.3). See also Ref. 10, Chap. XI, Theorem 12.

[13]A. Friedman, *Partial Differential Equations* (Holt, Rinehart and Winston, New York, 1969).

[14]N. Aronszajn, J. Math. (Paris) **36**, 235 (1957).

[15]Multiplication maps $H_1 \times H_1 \rightarrow H_0$ and $H_k \times H_k \rightarrow H_k$, for $k \geqslant 2$, when the manifold is three dimensional. See R. S. Palais, *Foundations of Global Nonlinear Analysis* (Benjamin, New York, 1968). From this and $\Psi > \epsilon > 0$ it follows that $\Psi \in H_k(\bar{\Omega}) \Rightarrow \Psi^{-8} \in H_k(\bar{\Omega})$, $k \geqslant 2$.

[16]If R_α is a rotation by α about the symmetry axis, then, since Θ and δ_p are axisymmetric, $\Theta(R_\alpha \Psi) = \delta_p$. Since Θ is an isomorphism, $R_\alpha \Psi = \Psi$.

[17]R. Penrose, Ann. N. Y. Acad. Sci. **224**, 125 (1973); G. Gibbons, Comm. Math. Phys. **27**, 87 (1972); P. S. Jang and R. M. Wald, J. Math. Phys. **18**, 41 (1977).

The form of Killing vectors in expanding $\mathcal{H}\mathcal{H}$ spaces

Stephanie A. Sonnleitner and J. D. Finley, III
Department of Physics and Astronomy, The University of New Mexico, Albuquerque, New Mexico 87131

(Received 21 May 1981; accepted for publication 1 August 1981)

The Killing vector structure of those spaces of complexified general relativity known as expanding hyperheavens is investigated using the methods of spinor calculus. The Killing equations for all left-algebraically degenerate Einstein vacuum spaces are completely integrated. Using the available gauge freedom, the resulting homothetic and isometric Killing vectors are classified in an invariant way according to Petrov–Penrose type. A total of four distinct kinds of isometric Killing vectors and three distinct kinds of homothetic Killing vectors are found. A master Killing vector equation is found which gives the form that the Lie derivative of the metric potential function W must take in order that it admit a given Killing vector.

PACS numbers: 04.20.Cv, 04.20.Me

1. INTRODUCTION

The most general algebraically degenerate solutions of the complex Einstein's vacuum field equations are called $\mathcal{H}\mathcal{H}$ spaces.[1] These spaces have a curvature tensor with a self-dual part that is algebraically degenerate (i.e., it possesses a multiple Debever–Penrose vector), while its anti-self-dual part is completely arbitrary. This degeneration is geometrically characterized by the existence of a 2-parameter congruence of totally null 2-surfaces which foliate the four-dimensional manifold under consideration.[2] On the other hand, these spaces are also characterized algebraically by the possession of a Hertz-like potential function W. This function determines the local metric structure of the space and is itself subject to a single (nonlinear) partial differential constraint,[3,4] analogous to the wave equation that Hertz's potential must satisfy.

The general class of $\mathcal{H}\mathcal{H}$ spaces may be divided into a number of different subclasses. It is clear that all real-valued (Minkowski signature), algebraically-degenerate Einstein spaces are special cases of them. This is of course the principal reason for our interest in these complex solutions, although quantum gravity theory may in fact have other uses for complex-valued solutions as well as those sections with Euclidean signature. However, staying on the complex level, it is useful to subdivide by the special properties of the null 2-surfaces. In Ref. 3, all $\mathcal{H}\mathcal{H}$ spaces were divided into expanding (case II), plane (case I), and left-flat ($\mathcal{H}$ spaces). The geometrical meaning of this division has its basis in the behavior of the variation over a given 2-surface of the vectors normal to the surface. In particular, the normal components of the covariant derivatives along the surface of the normal vectors form a 1-form referred to as the (complex) *expansion* of the leaves of the congruence. The general case is then simply that in which the expansion is not zero. On the other hand, if the expansion is zero, then the covariant derivatives of the normal vectors lie totally tangential to the 2-surface in question. Therefore, it is possible, in that case, to choose the normal directions to all be parallel. It is then reasonable to refer to such a situation as "plane." The case of $\mathcal{H}$ spaces is simply that in which the self-dual part of the curvature actually vanishes. In this case there is more than one (independent and distinct) 2-parameter family of such 2-surfaces. (A more technical discussion of this division is given in Ref. 1.)

The fact that these spaces are describable by scalar solutions of a single partial differential equation makes it possible to study various properties of them without necessarily having the general solution of the equation. Since this is such a large class, it is reasonable to work toward a better understanding of its structure by studying various additional structures, such as symmetry properties. In the case of $\mathcal{H}$ spaces it was shown some time ago[5] that the requirement of a Killing structure on the space as well could lead to a considerable simplification in the equations which must be solved. More explicitly the coexistence of the "Hertz" structure typified by the existence of the function $\overline{W}$, and the Killing structure given by the existence of (at least) one Killing vector gives rise to a coexistence relation referred to as a master equation. (Put quite differently, the set of 10 Killing's equations can be completely integrated, leaving only as a residual consistency condition a single first-order partial differential equation for $\overline{W}$, in an analogous manner to the integration of the 10 vacuum Einstein equations down to the constraint equation which $\overline{W}$ itself satisfies.) In any set of coordinates adapted to the existence of the special null congruence this equation gives the form the Killing vector must have, modulo some functions which are constant on any given leaf of the congruence, and the form the Lie derivative (in the direction of the Killing vector) of $\overline{W}$ must have, again modulo some set of functions constant on each leaf.

The gauge functions which appear in the master equation and the description of the Killing vector can be considered as determined by these equations if $\overline{W}$ is given in a specific set of coordinates. On the other hand, more generally these functions characterize the available gauge freedom in the choice of suitable coordinates. By considering all gauge transformations which preserve an appropriate choice of tetrad and coordinates, it was shown in Ref. 5 that the quotient of all the modulus functions in the description of the Killing vectors by all the functions in the description of the gauge group is a five-dimensional set. That is, that any single Killing vector in a given $\mathcal{H}$ space can always have appropriate coordinates chosen which are adapted for it in such a way that it takes on one of only five distinct simple forms (without any arbitrary functions). In that case, insertion of each of these forms into the equation which $\overline{W}$ must satisfy causes a

 0022-2488/82/010116-07$02.50

simplification of that equation which allows the determination of the general solution in four of the five cases.[6]

In the case of plane $\mathscr{H}\mathscr{H}$ spaces the integration of Killing's equations has also been performed in general giving an appropriate master equation[7] in that case, although the insertion of this information back into the hyperheavenly equation has yet to be done. In contrast to the exposition given in Ref. 5, the discussion of the plane $\mathscr{H}\mathscr{H}$ case was done in 2-spinorial coordinates that are particularly well suited for calculations related to this *2-parameter* congruence of *2-surfaces*, which were first introduced in Refs. 4 and 8. However, the general case of the integration of Killing's equations in an expanding $\mathscr{H}\mathscr{H}$ space has been somewhat resistant. It is the purpose of this article to give the derivation of a master equation (for the Lie derivative of $\overline{W}$) in the expanding case and to determine the allowed forms which these Killing vectors may take when expressed in an optimal gauge, so as to eliminate all arbitrary functions from their expression.

To describe how this occurs and some uses to which it may be put, we proceed in Sec. 2 to give a brief introduction to the technical description of expanding $\mathscr{H}\mathscr{H}$ spaces in a spinorial notation, while sending the reader to Refs. 1 (or 4) for more detail. In Sec. 3 we then outline the solution of Killing's equations and their attendant integrability conditions, while Sec. 4 is concerned with finding the optimal gauge mentioned above. Finally, in Sec. 5, we present two simple examples of the use of these results.

2. EXPANDING $\mathscr{H}\mathscr{H}$ SPACES

A general $\mathscr{H}\mathscr{H}$ space is most efficiently described by a pair of coordinates conceived of as a 2-spinor [but enjoying somewhat more general transformation properties than the usual SL(2, C)] $p^{\dot{A}}$ which are coordinates along any given leaf of the congruence, and another pair of coordinates given as a 2-spinor $q_{\dot{B}}$ which are parameters which label the various members of the congruence. The 2-surfaces in the congruence are then the integral surfaces of the 2-form

$$\Sigma = \tfrac{1}{2} dq_{\dot{A}} \wedge dq^{\dot{A}}. \tag{2.1}$$

In general, the (nonzero) expansion of the congruence picks out a special direction on any given leaf, which we specify by the constant spinor $J_{\dot{A}}$, with the expansion 1-form being given (up to proportionality) by $J_{\dot{A}}\, dp^{\dot{A}}$.[4] As has already been indicated, the structure of the space is essentially given by a potential function $\overline{W} \equiv \overline{W}(p^{\dot{A}}, q_{\dot{B}})$ that must satisfy the hyperheavenly equation[4]

$$\begin{aligned} &\tfrac{1}{2}\phi^4(\partial^{\dot{A}}\phi^{-2}\partial^{\dot{B}}\overline{W})(\partial_{\dot{A}}\phi^{-2}\partial_{\dot{B}}\overline{W}) + \phi^{-1}\partial^{\dot{A}}W_{,\dot{A}} \\ &\quad - \mu\phi^4\partial_\phi\phi^{-1}\partial_\phi\phi^{-1}\overline{W} + (\eta/\tau^2)K^{(\dot{A}}J^{\dot{B})}p_{\dot{A}}\mu_{,\dot{B}} \\ &\qquad = N_{\dot{A}}p^{\dot{A}} + \gamma, \end{aligned} \tag{2.2}$$

where μ, γ, and $N_{\dot{A}}$ are any functions of the $q_{\dot{B}}$ only that permit the equation to be satisfied. In the statement of the hyperheavenly equation we have used both the coordinates $p^{\dot{A}}$ and a version adapted to the direction of the expansion by

$$\phi \equiv J_{\dot{A}}p^{\dot{A}} + \kappa, \quad \eta \equiv K^{\dot{A}}p_{\dot{A}}, \quad K^{[\dot{A}}J^{\dot{B}]} = \tau\epsilon^{\dot{A}\dot{B}}, \tag{2.3}$$

where κ and τ are constants and $K^{\dot{A}}$ is complementary to $J_{\dot{B}}$ in the sense that the pair form a constant, local spinor basis, and we have used the abbreviations

$$\partial_{\dot{A}}\overline{W} \equiv \partial\overline{W}/\partial p^{\dot{A}}, \quad \overline{W}_{,\dot{A}} \equiv \partial\overline{W}/\partial q^{\dot{A}}, \quad \partial_\phi\overline{W} \equiv \partial\overline{W}/\partial\phi \tag{2.4}$$

as convenient abbreviations for partial derivative operations that will recur many times.

We prefer to describe the metric by a conventional null tetrad

$$ds^2 = 2e^1 \underset{s}{\otimes} e^2 + 2e^3 \underset{s}{\otimes} e^4, \tag{2.5}$$

which adapts itself in the usual way to a spinorial treatment via the usual symbols

$$g^{A\dot{B}} = \sqrt{2}\begin{pmatrix} e^4, & e^2 \\ e^1, & e^3 \end{pmatrix} \equiv \sqrt{2}\begin{pmatrix} E^{\dot{B}} \\ -e^{\dot{B}} \end{pmatrix}, \quad \begin{matrix} A = 1 \\ A = 2' \end{matrix} \tag{2.6a}$$

and

$$ds^2 = -\tfrac{1}{2}g^{A\dot{B}} \underset{s}{\otimes} g_{A\dot{B}} = 2E^{\dot{B}} \underset{s}{\otimes} e_{\dot{B}}. \tag{2.6b}$$

The 2-spinorial basis 1-forms are simply related to the differentials of the spinorial coordinates:

$$e_{\dot{A}} = \phi^{-2}dq_{\dot{A}}, \quad E^{\dot{A}} = Q^{\dot{A}\dot{B}}dq_{\dot{B}} - dp^{\dot{A}}. \tag{2.7}$$

The essential solution of Einstein's vacuum equations for $\mathscr{H}\mathscr{H}$ spaces[3,4] is then given by Eq. (2.2) and the following equation which expresses the metric in terms of $\overline{W}$:

$$Q^{\dot{A}\dot{B}} = -\partial^{(\dot{A}}\phi^4\partial^{\dot{B})}\phi^{-3}\overline{W} + (\mu/\tau^2)\phi^3K^{\dot{A}}K^{\dot{B}}. \tag{2.8}$$

As a convenient summary of some of the essential characteristics of these spaces we list here the components of the conformal tensor in this tetrad,[4] from which we see how the various possible complex Petrov types are determined by W, μ, $N_{\dot{A}}$, and γ:

$$\begin{aligned} &C^{(5)} = 0 = C^{(4)}; \quad C^{(3)} = -2\mu\phi^3; \\ &C^{(2)} = 2\phi^5\{N_{\dot{A}}J^{\dot{A}} - [p^{\dot{A}} + (\kappa/2\tau)K^{\dot{A}}]\mu_{,\dot{A}}\}; \\ &C^{(1)} \\ &= 2\phi^7\{\phi[(\mu/\tau)\phi^2K^{\dot{A}} - \partial/\partial q_{\dot{A}}][N_{\dot{A}} + (1/2\tau)p_{\dot{A}}K^{\dot{R}}\mu_{,\dot{R}}] \\ &\quad + J^{\dot{B}}(N_{\dot{A}}p^{\dot{A}} + \gamma + 3\mu\overline{W})_{,\dot{B}} - (\eta/2\tau)J^{\dot{A}}p^{\dot{B}}\mu_{,\dot{A}\dot{B}} \\ &\quad + [2N_{\dot{A}}J^{\dot{A}}J_{\dot{B}} - \mu_{,\dot{C}}(J^{\dot{C}}p_{\dot{B}} + p^{\dot{C}}J_{\dot{B}} + \kappa K_{\dot{B}}J^{\dot{C}}/\tau)]\partial^{\dot{B}}\overline{W}\}; \\ &C_{\dot{A}\dot{B}\dot{C}\dot{D}} = \phi^3\partial_{\dot{A}}\partial_{\dot{B}}\partial_{\dot{C}}\partial_{\dot{D}}(\overline{W} - \tfrac{3}{2}\mu\phi^2\eta^2/\tau^2). \end{aligned} \tag{2.9}$$

We can see that $\mu \neq 0$ is necessary for left-Petrov-type II or D while, when $\mu = 0$, the nonvanishing of $N_{\dot{A}}J^{\dot{A}} \equiv 2\nu$ distinguishes between left-Petrov-type III and N. In this last instance N degenerates to a left-conformally-flat space (a right $\mathscr{H}$ space) when $J^{\dot{B}}\gamma_{,\dot{B}}$ vanishes as well as μ and $N_{\dot{A}}J^{\dot{A}}$. (The quantity $K^{\dot{A}}N_{\dot{A}} \equiv 2\xi$ can always be gauged away if desired. See Ref. 1 for a more complete discussion of optimal gauges for the form of the $C^{(i)}$.)

We are looking for solutions of Killing's equations for the general case of homothetic Killing vectors:

$$K_{(\mu;\nu)} = \chi_0 g_{\mu\nu}, \tag{2.10}$$

where χ_0 is a constant, covariant derivatives are indicated by a semicolon, and K_μ indicates the covariant components of our Killing vector. The case $\chi_0 = 0$ corresponds to the usual case of a pure Killing vector—an isometry—while the case where χ varies over the manifold—a conformal symmetry—is known[5] to be too restrictive to allow very general Petrov types. Since, however, we have couched even our coordinates for the space in spinorial terms, it is desirable and convenient to rephrase Killing's equations in a spinorial language as well.

In this language, Killing's equations take the form[10]

$$E_{RS}{}^{\dot A\dot B} \equiv \nabla_{(R}{}^{(\dot A}K_{S)}{}^{\dot B)} = 0, \tag{2.11}$$

where the Killing spinor $K_S{}^{\dot B}$ is related to the Killing 1-form by

$$K = -\tfrac{1}{2}g^{A\dot B}K_{A\dot B} = \not k_{\dot B}E^{\dot B} + \not\chi^{\dot B}e_{\dot B}. \tag{2.12}$$

Assuming that χ_0 is a constant, the integration conditions for these equations may be written as

$$\begin{aligned} L_{RST}{}^{\dot A} &\equiv \nabla_R{}^{\dot A}l_{ST} + 2K_U{}^{\dot A}C^U{}_{RST} = 0,\\ M_{RSTU} &\equiv K_{P\dot A}\nabla^{P\dot A}C_{RSTU} - 4\chi_0 C_{RSTU} + 4l_{V(R}C^V{}_{STU)}\\ &= 0, \end{aligned} \tag{2.13}$$

with similar equations for objects with dotted indices, and where l_{ST} and $l^{\dot S\dot T}$ are symmetric spinors such that

$$\nabla_R{}^{\dot A}K_S{}^{\dot B} = l_{RS}\epsilon^{\dot A\dot B} + l^{\dot A\dot B}\epsilon_{RS} - 2\epsilon_{RS}\epsilon^{\dot A\dot B}\chi_0, \tag{2.14a}$$

or

$$\begin{aligned} l_{RS} &= \tfrac{1}{2}\epsilon_{\dot A\dot B}\nabla_{(R}{}^{\dot A}K_{S)}{}^{\dot B}, \quad l^{\dot A\dot B} = \tfrac{1}{2}\epsilon^{RS}\nabla_R{}^{(\dot A}K_S{}^{\dot B)},\\ \chi_0 &= -\tfrac{1}{8}\epsilon_{\dot A\dot B}\epsilon^{RS}\nabla_R{}^{\dot A}K_S{}^{\dot B}. \end{aligned} \tag{2.14b}$$

(Note that $\nabla_R{}^{\dot A} = g_R{}^{\dot A\mu}\nabla_\mu$ are the spinorial components of the covariant derivative.)

We next give the results of the complete integration of Eqs. (2.11), as well as a complete, simple classification of these results. Details[10] of these calculations are presented in the following two sections. As stated above, these results are restricted to the case where χ_0 is constant.

We find that the Lie derivative of the metric potential function $\overline{W}$ must take the following form to admit a given Killing vector K:

$$\mathscr{L}_K\overline{W} = 2\omega\overline{W} + P, \tag{2.15}$$

where P is a fourth-order polynomial in $p^{\dot A}$,

$$\begin{aligned} P \equiv{}& (1/2\tau)p_{\dot D}K_{\dot C}\delta^{\dot C,\dot D} - (\kappa/2\tau^2)p_{\dot A}K_{\dot C}K_{\dot D}\delta^{\dot D,\dot C,\dot A}\\ &+ (\eta\phi/2\tau^3)(K_{\dot E}K_{\dot D}J_{\dot C} + 2J_{\dot E}K_{\dot D}K_{\dot C})\delta^{\dot C,\dot D,\dot E}\\ &+ (\phi^2/2\tau^3)K_{\dot E}K_{\dot C}K_{\dot D}\delta^{\dot D,\dot C,\dot E}\\ &+ (\eta^2/2\tau^3)J_{\dot E}J_{\dot C}K_{\dot D}\delta^{\dot D,\dot C,\dot E} - (\phi^3/3\tau)K_{\dot C}\Delta^{\dot C}\\ &\cdot (\mu\phi^3\eta/2\tau^3)K_{\dot C}K_{\dot D}\delta^{\dot D,\dot C} + \iota, \end{aligned} \tag{2.16}$$

and $\omega = 3\chi_0 - \delta_{\dot A}{}^{,\dot A}$ is functionally the same form found for plane hyperheavens.[7] The term corresponding to $2\alpha_0\Lambda$ for plane-spaces was not carried through in these calculations since, as before, it restricts us to left-flat spaces. The Killing vector (in a coordinate basis), which corresponds to Eq. (2.10) of Ref. 7, is given by

$$\begin{aligned} K = \delta_{\dot A}\frac{\partial}{\partial q_{\dot A}} &+ (p_{\dot B}\delta^{\dot B,\dot A} - 2\chi_0 p^{\dot A}\\ &- \frac{2}{\tau}J_{\dot B}K_{\dot C}\delta^{\dot C,\dot B}p^{\dot A} + \epsilon^{\dot A})\partial_{\dot A}. \end{aligned} \tag{2.17}$$

These two equations, (2.16) and (2.17), contain four apparently arbitrary functions of $q_{\dot C}$, namely $\delta^{\dot A}$, $\epsilon^{\dot A}$, $\Delta^{\dot A}$, and ι (7 degrees of freedom). In this form, a general solution is not feasible, but when the integration conditions [$L_{ABC}{}^{\dot A}$ and M_{ABCD}, Eq. (2.13)] are taken into consideration these results become more tractable. The final classification is accomplished by treating each Petrov–Penrose type separately as well as dividing the problem into the purely homothetic and the isometric cases. This results in one distinct kind of purely homothetic Killing vector per type and from two to four distinct kinds of isometric Killing vectors, depending on type (see Table I). All of the Killing vectors are then free of arbitrary functions of $q_{\dot C}$ and are quite simple in form, while the complete generality of the results is preserved.

TABLE I. Classification of Killing vectors.

Killing vector	Left–Petrov types allowed	Master equation $\mathscr{L}_K W =$
Isometric		
∂_w	II, D, III, N	0
∂_t	II, D, III	0
∂_η	II, D	0
∂_η	N	$(1/3\tau)\phi^3 e(w,t)$
$w\partial_t + \phi\partial_\eta$	II, D	$(1/2\tau^2)\mu_0\phi^3\eta + (1/6\tau)\phi^3\nu$
$w\partial_t + \phi\partial_\eta$	III	0
Homothetic		
$\chi_0\{2t\partial_t - \phi\partial_\phi + \eta\partial_\eta\}$	II, D	0
$\chi_0\{t\partial_t - \phi\partial_\phi\}$	III	$-\chi_0 W + (1/3\tau)\phi^3 e(w,t)$
$\chi_0\{\phi\partial_\phi + \eta\partial_\eta\}$	N	$2\chi_0 W$

3. INTEGRATION OF KILLING'S EQUATIONS

We are now ready to begin integrating Killing's equations. Starting with the simplest triple, $E_{11}{}^{\dot A\dot B} = \phi^{-2}\partial^{(\dot A}\phi^2\not k^{\dot B)} = 0$, we immediately find that $\not k^{\dot B} = \phi^{-2}\{\alpha p^{\dot B} + \delta^{\dot B}\}$, where α and $\delta^{\dot B}$ are functions of $q_{\dot C}$ only. Consideration of the M_{ABCD} in Eq. (2.13) leads easily to the conclusion that either $\alpha = 0$ or the spaces in question are left-flat. Since the latter are $\mathscr{H}$-spaces, and their Killing vectors are completely discussed in Ref. 6, we hereafter restrict ourselves to the case $\alpha = 0$, so that

$$\not k^{\dot B} = \phi^{-2}\delta^{\dot B}. \tag{3.1}$$

The first integration condition $L_{111}{}^{\dot A}$, from Eq. (2.13), is identically zero, while one part of $L_{211}{}^{\dot A}$ yields the condition $J_{\dot A}J_{\dot B}\delta^{\dot B,\dot A} = 0$ and the other part gives a compact expression for l_{12}:

$$l_{12} = -(\mu\phi^2/\tau)K_{\dot A}\delta^{\dot A} + 2\phi^{-1}\tau\delta_{\dot A}J^{\dot A}\overline{W}_\eta + (1/\tau)K_{\dot A}J_{\dot B}\delta^{\dot B,\dot A}. \tag{3.2}$$

The second triple $E_{12}{}^{\dot A\dot B}$ can be cast in the form

$$E_{12}{}^{\dot A\dot B} = 0 = \partial^{(\dot A}\{\not\chi^{\dot B)} - \delta_{\dot D}Q^{\dot B)\dot D} - \delta_{\dot C}{}^{,\dot B)}p^{\dot C}\}. \tag{3.3}$$

This is of the form $\partial^{(\dot A}T^{\dot B)}$, which has the obvious solution $T^{\dot B} = \xi p^{\dot B} - \epsilon^{\dot B}$, where ξ and $\epsilon^{\dot B}$ are independent of $p^{\dot C}$. Extracting $\not k^{\dot B}$ from the above solution gives us an expression for the other half of the Killing spinor $K_A{}^{\dot B}$:

$$\not\chi^{\dot B} = \delta_{\dot A}Q^{\dot A\dot B} - M^{\dot B}, \tag{3.4}$$

where, for ease of further calculation, we have defined

$$M^{\dot{B}} = p_A \delta^{\dot{A},\dot{B}} - \xi p^{\dot{B}} + \epsilon^{\dot{B}}. \tag{3.5}$$

We now have expressions for the Killing spinor $K_A{}^{\dot{A}}$ in terms of five arbitrary functions of $q_{\dot{C}}$, namely $\delta^{\dot{A}}$, $\epsilon^{\dot{A}}$, and ξ. The third triple will yield a master equation for $\overline{W}$, while remaining integration conditions and the as yet unused singlet equation will impose restrictions on the arbitrary functions in the Killing spinor. It is convenient to generate a few more simplifying conditions before proceeding to the final integration. Comparison of the direct calculation of l_{12} with the expression generated from $L_{211}{}^{\dot{A}}$ yields

$$-\kappa\xi + (\kappa/\tau)J_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}} = J_{\dot{A}}\epsilon^{\dot{A}}. \tag{3.6}$$

The expression for χ_0 generates a relation between χ_0, considered as a given quantity ξ, and $\delta^{\dot{A}}$, viz, $4\chi_0 = 2\xi - (4/\tau)J_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}}$. Then the condition represented by $M_{1122} = 0$ produces a further simplifying constraint, given by

$$\mu(4\chi_0 + (3/\tau)J_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}}) = \delta_{\dot{A}}\mu^{,\dot{A}}. \tag{3.7}$$

Returning now to Killing's equations and utilizing the information gained to this point, the last triple becomes (with $\eth^{\dot{A}} = \phi^2\{\partial/\partial q_{\dot{A}} + Q^{\dot{A}\dot{B}}\partial_{\dot{B}}\}$:

$$\begin{aligned} E_{22}{}^{\dot{A}\dot{B}} = 0 = {} & \eth^{(\dot{A}}\delta_{\dot{D}}Q^{\dot{B})\dot{D}} - \eth^{(\dot{A}}M^{\dot{B})} + \tfrac{1}{2}\phi^2\delta_{\dot{C}}Q^{\dot{D}\dot{C}}\partial^{(\dot{A}}Q_{\dot{D}}{}^{\dot{B})} \\ & - \tfrac{1}{2}\phi^2 M^{\dot{D}}\partial^{(\dot{A}}Q_{\dot{D}}{}^{\dot{B})} - \tfrac{1}{2}\phi^2\delta_{\dot{C}}Q^{\dot{D}\dot{C}}\partial_{\dot{D}}Q^{\dot{A}\dot{B}} \\ & + \tfrac{1}{2}\phi^2 M^{\dot{D}}\partial_{\dot{D}}Q^{\dot{A}\dot{B}} \\ & - \tfrac{1}{2}\phi^2\delta_{\dot{D}}Q^{\dot{D}(\dot{A}}\partial_{\dot{C}}Q^{\dot{B})\dot{C}} + \tfrac{1}{2}\phi^2 M^{(\dot{A}}\partial_{\dot{C}}Q^{\dot{B})\dot{C}} \\ & + \delta^{(\dot{A}}\eth_{\dot{C}}Q^{\dot{B})\dot{C}}. \end{aligned} \tag{3.8}$$

With judicious combination of terms, Eq. (3.8) reduces to

$$\begin{aligned} E_{22}{}^{\dot{A}\dot{B}} = 0 = {} & Q^{\dot{D}(\dot{B}}\delta_{\dot{D}}{}^{,\dot{A})} + M^{\dot{R}}\partial_{\dot{R}}Q^{\dot{A}\dot{B}} + \delta_{\dot{R}}Q^{\dot{A}\dot{B},\dot{R}} \\ & - M^{(\dot{A},\dot{B})} - Q^{\dot{C}(\dot{A}}\partial_{\dot{C}}M^{\dot{B})}. \end{aligned} \tag{3.9}$$

At this point we introduce[4] a potentialization $A^{\dot{B}}$ of $Q^{\dot{A}\dot{B}}$ having the form $Q^{\dot{A}\dot{B}} = \phi^3\partial^{(\dot{A}}A^{\dot{B})}$ which causes (3.9) to take the following form:

$$\begin{aligned} 0 = {} & \partial^{(\dot{A}}\{\delta_{\dot{R}}A^{\dot{B}),\dot{R}} + \delta_{\dot{D}}{}^{,\dot{B})}A^{\dot{D}} - (1/\tau)A^{\dot{B})}J_{\dot{C}}K_{\dot{D}}\delta^{\dot{D},\dot{C}} \\ & - 4\chi_0 A^{\dot{B})} + M^{|\dot{R}|}\partial_{\dot{R}}A^{\dot{B})}\} \\ & - \{\partial_{\dot{R}}A^{(\dot{B}}\}\partial^{\dot{A})}M^{\dot{R}} + \partial^{\dot{D}}A^{(\dot{A}}\delta_{\dot{D}}{}^{,\dot{B})} - \phi^{-3}M^{(\dot{B},\dot{A})}, \\ \equiv {} & \partial^{(\dot{A}}R^{\dot{B})} + H^{(\dot{A}\dot{B})}. \end{aligned} \tag{3.10}$$

The first term $\partial^{(\dot{A}}R^{\dot{B})}$ is already in the desired, easily integrable form, while the second term may be cast in this form after introduction of a potentialization of $A^{\dot{A}}$ in terms of $\overline{W}$, namely, $A^{\dot{A}} = -\phi^{-2}\partial^{\dot{A}}\overline{W} + (\mu/\tau^2)\eta K^{\dot{A}}$. After this substitution and application of earlier integration conditions, Eqs. (3.10) takes the form $\partial^{(\dot{A}}T^{\dot{B})} = 0$ which imples $T^{\dot{B}} = \lambda p^{\dot{B}} + \Delta^{\dot{B}}$. This new equation can then be written as $\partial^{\dot{B}}S + P^{\dot{B}} = 0$. Operating with $\partial_{\dot{B}}$ yields the condition $\partial_{\dot{B}}P = 0$, leading to two new conditions between the integration variables $\delta^{\dot{A}}$, $\epsilon^{\dot{A}}$, λ, and $\Delta^{\dot{A}}$. When these conditions are applied, the final equation reads $\partial^{\dot{B}}\overline{N} = 0$ so that we may immediately write $\overline{N} = \iota(q_{\dot{C}})$. This last equation is the master equation for $\overline{W}$, and, after appropriate sorting of terms, takes the form of Eq. (2.15). Up to this point we have generated five integration conditions, these being

$$J_{\dot{A}}J_{\dot{B}}\delta^{\dot{B},\dot{A}} = 0 \quad (\text{from } L_{211}{}^{\dot{A}} = 0);$$

$$J_{\dot{A}}\epsilon^{\dot{A}} = -\kappa\{2\chi_0 + (1/\tau)J_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}}\}$$

(from $L_{211}{}^{\dot{A}} = 0$ and χ_0);

$$\mu\{4\chi_0 + (3/\tau)J_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}}\} = \delta_{\dot{A}}\mu^{,\dot{A}} \quad (\text{from } M_{1122} = 0); \tag{3.11}$$

$$\lambda = -(\mu/2\tau^2)K_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}} \quad (\text{from } E_{22}{}^{\dot{A}\dot{B}} = 0);$$

$$(\mu/\tau^2)\kappa K_{\dot{A}}K_{\dot{B}}\delta^{\dot{B},\dot{A}} - (2\mu/\tau)K_{\dot{R}}\epsilon^{\dot{R}} = 2J_{\dot{B}}\Delta^{\dot{B}}$$

(from $E_{22}{}^{\dot{A}\dot{B}} = 0$).

Consideration of the remaining integration conditions (M_{1122}, $L_{122}{}^{\dot{A}}$, $L_{212}{}^{\dot{A}}$, $L_{222}{}^{\dot{A}}$ and M_{2222}) gives rise to three more conditions on the variables $\delta^{\dot{A}}$, $\epsilon^{\dot{A}}$, $\Delta^{\dot{A}}$, and ι. These conditions will lead to further simplification of the Killing vector and master equations, as well as aiding in the classification of the Killing vectors by eliminating the arbitrariness in the coefficients of Eq. (2.17). They are

$$\begin{aligned} 0 = {} & 3\mu\epsilon_{\dot{A}}{}^{,\dot{A}} + 2\epsilon_{\dot{A}}\mu^{,\dot{A}} + (\kappa/\tau)(\delta_A K_{\dot{B}}\mu^{,\dot{B}})^{,\dot{A}} + (1/\tau)J_{\dot{A}}\epsilon^{\dot{A}}K_{\dot{B}}\mu^{,\dot{B}} \\ & + 2N^{\dot{C}}J_{\dot{C}}(\delta_{\dot{A}}{}^{,\dot{A}} + (1/\kappa)J_{\dot{A}}\epsilon^{\dot{A}}) + 2\delta_{\dot{A}}J_{\dot{B}}N^{\dot{B},\dot{A}} \\ & \qquad (\text{from } M_{1222} = 0); \\ 0 = {} & 4\tau^2\gamma K_{\dot{A}}J_{\dot{B}}\delta^{\dot{B},\dot{A}} + 2\tau^3\delta_{\dot{A}}\gamma^{\dot{A}} + 6\mu\tau^3\iota + 3\mu\kappa^2 K_{\dot{A}}K_{\dot{B}}K_{\dot{C}}\delta^{\dot{C},\dot{B},\dot{A}} \\ & - K_{\dot{A}}K_{\dot{B}}K_{\dot{C}}J_{\dot{D}}\delta^{\dot{D},\dot{C},\dot{B},\dot{A}} + 2\tau^2 J_{\dot{A}}N^{\dot{A}}\{K_{\dot{B}}\epsilon^{\dot{B}} - \kappa K_{\dot{B}}K_{\dot{C}}\delta^{\dot{C},\dot{B}}\} \\ & + 2\kappa\tau^2\delta_{\dot{A}}K_{\dot{B}}N^{\dot{B},\dot{A}} - 3\mu\kappa\tau K_{\dot{A}}K_{\dot{B}}\epsilon^{\dot{B},\dot{A}} + \kappa\tau K_{\dot{A}}\mu^{,\dot{A}}K^{\dot{B}}\epsilon_{\dot{B}} \\ & + \kappa^2\tau K_{\dot{A}}\mu^{,\dot{A}}K_{\dot{B}}K_{\dot{C}}\delta^{\dot{C},\dot{B}} \quad (\text{from } L_{222}{}^{\dot{A}} = 0); \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} 0 = {} & 2\mu\tau J_{\dot{A}}K_{\dot{B}}\Delta^{\dot{B},\dot{A}} + 2\mu N_{\dot{A}}J^{\dot{A}}K_{\dot{B}}K_{\dot{C}}\delta^{\dot{C},\dot{B}} + 2\mu\tau\delta_{\dot{A}}K_{\dot{B}}N^{\dot{B},\dot{A}} \\ & - 4\mu K^{\dot{A}}N_{\dot{A}}K_{\dot{B}}J_{\dot{C}}\delta^{\dot{C},\dot{B}} + \mu\tau K_{\dot{A}}\mu^{,\dot{A}}K_{\dot{B}}\epsilon^{\dot{B}} \\ & \qquad (\text{from } M_{2222} = 0). \end{aligned}$$

This completes the task of integrating Killing's equations $E_{AB}{}^{\dot{A}\dot{B}}$ consistent with the constraints $L_{ABC}{}^{\dot{A}}$ and M_{ABCD}.

4. CLASSIFICATION OF KILLING'S VECTORS

It is convenient at this point to introduce explicit variables t and w for the $q^{\dot{A}}$ and to decompose the spinor integration variables $\delta^{\dot{A}}$, $\Delta^{\dot{A}}$, and $\epsilon^{\dot{A}}$ into parts parallel to $J^{\dot{A}}$ and $K^{\dot{A}}$. Following Ref. 1, we introduce the following choice of $q^{\dot{A}}$:

$$w \equiv J_{\dot{A}}q^{\dot{A}}, \quad t \equiv K^{\dot{A}}q_{\dot{A}}, \tag{4.1}$$

which implies $\tau\partial_t = J_{\dot{A}}(\partial/\partial q_{\dot{A}})$ and $\tau\partial_w = K_{\dot{A}}(\partial/\partial q_{\dot{A}})$. Upon setting $\kappa = 0$ (the case $\kappa \neq 0$ is only relevant to plane hyperheavens which have been dealt with elsewhere[7]) and taking account of the first and fourth conditions of Eq. (3.11), $\delta^{\dot{A}}$, $\Delta^{\dot{A}}$, and $\epsilon^{\dot{A}}$ take the form

$$\begin{aligned} \delta^{\dot{A}} &= (1/\tau)a(w)K^{\dot{A}} + (1/\tau)b(t, w)J^{\dot{A}}, \\ \epsilon^{\dot{A}} &= (1/\tau)c(t, w)J^{\dot{A}}, \\ \Delta^{\dot{A}} &= (1/\tau)e(t, w)J^{\dot{A}} + (\mu/\tau^2)e(t, w)K^{\dot{A}}. \end{aligned} \tag{4.2}$$

With these definitions, the relevant integration conditions become

$$\begin{aligned} & 4\mu\chi_0 - 3\mu b_t = a\mu_w + b\mu_t, \\ & 3\mu c_t + 2c\mu_t + \tau\nu(a_w + 2b_t - 2\chi_0) + \tau a\nu_w + \tau b\nu_t = 0, \\ & 4\gamma a_w + 2a\gamma_w + 2b\gamma_t + 6\mu\iota - a_{www} - \nu c = 0, \\ & -2\mu e_t + \mu\nu b_w - \mu a\xi_w - \mu b\xi_t - 4\mu\xi a_w - \mu c\mu_w = 0. \end{aligned} \tag{4.3}$$

Further progress is facilitated by breaking the investigation down by Petrov type and also handling the isometric and homothetic cases separately. We will also utilize the transformation properties of the $q^{\dot{A}}$ to simplify our calculations. Thus, we wish to determine the conditions for those transformations of the $q_{\dot{A}}$ which allow both $e_{\dot{A}}$ and $E^{\dot{A}}$ to transform in a form-invariant way. Define $q'_{\dot{A}} \equiv q'_{\dot{A}}(q_{\dot{B}})$, where the transformation matrix is given by $D_{\dot{R}}{}^{\dot{A}} \equiv (\partial q'_{\dot{R}}/\partial q_{\dot{A}})$ with determinant $\Delta \neq 0$, i.e., locally invertible. By first transforming $e_{\dot{A}}$ and demanding form invariance for $g = 2E^{\dot{A}} \underset{s}{\otimes} e_{\dot{A}}$, we are led to the transformation equations for $e_{\dot{A}}$, $E^{\dot{A}}$, and $p^{\dot{A}}$, which are

$$
\begin{aligned}
e'_{\dot{R}} &= \lambda D_{\dot{R}}{}^{\dot{A}} e_{\dot{A}}, \\
E'^{\dot{R}} &= (1/\lambda)(D^{-1}{}_{\dot{A}}{}^{\dot{R}} E^{\dot{A}} - h\phi^2 e^{\dot{R}}), \qquad (4.4) \\
p'^{\dot{R}} &= (1/\lambda) D^{-1}{}_{\dot{A}}{}^{\dot{R}} p^{\dot{A}} + \sigma^{\dot{R}},
\end{aligned}
$$

where $\lambda = \lambda(w,t)$, $h = h(w,t)$, $\zeta = \zeta(w,t)$ and where ζ and λ are related by $\zeta_t = \lambda^{-1/2} \neq 0, \infty$. The quantities $D_{\dot{R}}{}^{\dot{A}}$, $D^{-1}{}_{\dot{K}}{}^{\dot{A}}$, $J'_{\dot{R}}$ and $K'_{\dot{R}}$ are determined by the requirement that $J'_{\dot{K}}$, $K'_{\dot{R}}$, κ' be constant in the new coordinates, and are given by

$$
\begin{aligned}
D_{\dot{R}}{}^{\dot{A}} &= (1/\tau) J'_{\dot{R}} \partial\zeta/\partial q_{\dot{A}} - (j/\tau') K'_{\dot{R}} J^{\dot{A}}, \\
D^{-1}{}_{\dot{A}}{}^{\dot{R}} &= (1/\tau' j)\{ -a_0 K_{\dot{A}} J'^{\dot{R}} + a_0 \lambda^{1/2} \zeta_w J_{\dot{A}} J'^{\dot{R}} \\
&\qquad + j\lambda^{1/2} J_{\dot{A}} K'^{\dot{R}} \}, \qquad (4.5) \\
J'_{\dot{R}} &= \lambda^{1/2} D_{\dot{R}}{}^{\dot{A}} J_{\dot{A}}, \\
K'_{\dot{R}} &= (a_0/j) D_{\dot{R}}{}^{\dot{A}} \left(K_{\dot{A}} - \frac{\zeta_w}{\zeta_t} J_{\dot{A}} \right),
\end{aligned}
$$

where $\tau' \equiv a_0 \tau$, $J'_{\dot{R}}$, $K'_{\dot{R}}$ are constant additional degrees of freedom for the transformation if $J_{\dot{R}}$ and $K_{\dot{R}}$ are, $g = g(w)$, and j denotes dg/dw. Finally, $\sigma'^{\dot{R}}$ is

$$
\sigma'^{\dot{R}} = sJ'^{\dot{R}} + (\kappa\zeta_t - \kappa')K'^{\dot{R}}/\tau', \qquad (4.6)
$$

where $s = s(w,t)$ and $\kappa' = \lambda^{-1/2}\kappa - J'_{\dot{R}}\sigma'^{\dot{R}}$ ensure that the transformed $\kappa(\kappa')$ is also constant. With the above, one finally finds the transformation relation for $q_{\dot{C}}$, namely,

$$
q'_{\dot{C}} = (\zeta/\tau) J'_{\dot{C}} + (g/\tau) K'_{\dot{C}}. \qquad (4.7)
$$

These transformation equations then give transformed versions of ϕ and η,

$$
\phi' = \lambda^{-1/2}\phi, \quad \eta' = (a_0/\lambda j) D_{\dot{R}}{}^{\dot{A}} \left(K_{\dot{A}} - \frac{\zeta_w}{\zeta_t} J_{\dot{A}} \right). \qquad (4.8)
$$

The independent variables w and t obey the transformation equations $w' = g(w)$ and $t' = a_0 \zeta(t,w)$. Considering the transformation of $C^{(3)}$ and the hyperheavenly equation, we obtain the equations for μ', ν', ξ', and γ', where μ, ν, ξ, and γ are the arbitrary functions of $q_{\dot{B}}$ only that enter into the hyperheavenly equation (2.2). The relations are

$$
\begin{aligned}
\mu' &= \lambda^{3/2}\mu, \\
\nu' &= (\lambda/a_0 j)\{ \nu - 2\lambda js\mu_t - 3\mu j(\lambda s)_t \}, \\
\xi' &= (\lambda^{1/2}/j^2)\{ \xi + \lambda^{1/2}\zeta_w \nu + 2\tau\theta_t \\
&\qquad + \lambda^{1/2} js[\lambda(\zeta_w\mu_t - \mu_w\zeta_t) + \tfrac{3}{2}\mu(\zeta_w\lambda_t - \lambda_w\zeta_t)] \}, \qquad (4.9) \\
\gamma' &= (1/j^2)\{ \gamma + 3\mu v + (1/2j)\partial^2 j/\partial w^2 - (3/4j^2)(\partial j/\partial w)^2 \\
&\qquad + \tfrac{1}{2}\lambda j\tau s v + \tfrac{1}{2}\tau\lambda j^2 s^2 (\tfrac{3}{2}\mu\lambda_t + \lambda\mu_t) \\
&\qquad + \tfrac{1}{2}\lambda j\tau s[2\lambda js\mu_t + 3\mu j(\lambda s)_t] \},
\end{aligned}
$$

where $\theta = \omega(w,t)$ and $v = v(w,t)$ are new degrees of freedom, $j = dg/dw$, and we have chosen $\kappa = 0 = \kappa'$. Lastly, before beginning the simplification of the Killing vector, we need transformed expressions for the functions which appear in it, namely a, b, and c. The appropriate relations are

$$
\begin{aligned}
a' &= ja, \\
b' &= (a_0/\lambda^{1/2})[b + (\zeta_w/\zeta_t)a] = a_0(\zeta_t b + \zeta_w a), \qquad (4.10) \\
c' &= (a_0/\lambda j)\{ c + s\lambda j\tau[a_w - 2b_t \\
&\qquad + 2\chi_0 + a(\ln \lambda js)_w + b(\ln \lambda s)_t] \},
\end{aligned}
$$

where these equalities are determined by considering the transformation of the Killing vector. The quantities e and ι, which appear in the master equation, transform as

$$
\begin{aligned}
e' &= a_0\lambda^{3/2}\left\{ e + \frac{3\mu}{2\tau}\frac{\zeta_w}{\zeta_t}c - \frac{3\mu}{2}s\lambda^{3/2} j\left(\partial_w - \frac{\zeta_w}{\zeta_t}\right)\lambda^{1/2} \right. \\
&\qquad \left. \times\left(b + \frac{\zeta_w}{\zeta_t}a\right)\Big] - 2\tau(a_w - \chi_0)\theta - \tau b\theta_t - \tau a\theta_w \right\}, \qquad (4.11) \\
\iota' &= (1/\lambda^{3/2} j^2)\iota + \tfrac{1}{2}\lambda^{1/2}\tau s^2\{ \partial_t \lambda^{1/2}\partial_t[\lambda^{-1/2}(b + \zeta_w/\zeta_t)a] \} \\
&\qquad + \tfrac{1}{2}s\lambda^{1/2}\partial_t(1/\lambda j)\{ c + s\lambda j\tau[a_w - 2b_t + 2\chi_0 + a(\ln \lambda js)_w \\
&\qquad + b(\ln \lambda js)_t] \}) + (4\chi_0 - 3b_t + 2a_w)v + \tfrac{1}{2}\lambda jcs_t \\
&\qquad + av_w + bv_t.
\end{aligned}
$$

It turns out to be possible to always gauge μ to a constant and ι to zero, so we will work with the set of constraints simplified in this way (as well as setting $\kappa = 0$). The Eqs. (4.3) then become

$$
\begin{aligned}
&4\mu_0\chi_0 - 3\mu_0 b_t = 0, \\
&3\mu_0 e_t + \tau v(a_w + 2b_t + 2\chi_0) + \tau a v_w + \tau b v_t = 0, \qquad (4.12) \\
&4\gamma a_w + 2a\gamma_w + 2b\gamma_t - a_{www} - vc = 0, \\
&\mu_0[2e_t + vb_w - a\xi_w - b\xi_t - 4\xi a_w] = 0.
\end{aligned}
$$

We now divide the remaining computations according to Petrov type and according to the nature of χ_0 (0 or constant). Further, for the isometric Killing vectors, we will investigate the cases $a \neq 0$, $a = 0 \neq b$, and $a = b = 0 \neq c$ separately.

Case 1: Petrov–Penrose types [II] $\otimes$ [Any] and [D] $\otimes$ [Any] with $\chi_0 = 0$. In this case we have $\mu_0 \neq 0$. For the first situation, $a \neq 0$, it is possible to gauge away b, c, e, v, ξ, and γ, leaving a very simple form for the Killing vector and master equation, namely, $K = \partial_w$ and $\mathscr{L}_K \overline{W} = \overline{W}_w = 0$. In the case $a = 0 \neq b$, we find two distinct solutions, one with b

a constant and the other where $b = w$. The first choice leads simply to $K = \partial_t$, $\overline{W}_t = 0$ with $\xi = \gamma = c = e = 0$, $\nu = \nu(w)$. The other possibility, $b = w$, leads to $\xi = \gamma = c = 0$, but $e = \frac{1}{2}\nu t$ along with $K = w\partial_t + \phi\partial_\eta$ and $\mathscr{L}_K\overline{W} = \frac{1}{2}\mu_0\phi^3\eta/\tau^2 + (1/6\tau)\phi^3\nu$. Finally, if $a = b = 0 \neq c$, then we find $\xi = \nu = \gamma = e = 0$ and $K = \partial_\eta$, $\overline{W}_\eta = 0$.

Case 2: Petrov–Penrose type [III] $\otimes$ [Any] with $\chi_0 = 0$. Type III implies $\mu = 0 \neq \nu$, and we are always able to obtain $\nu = \nu_0$. For the initial conditions $a \neq 0$, we are able to obtain $\gamma = c = 0$, but we may not have both e and ξ gauged to zero. For the simplicity of the master equation we choose to eliminate e and get $K = \partial_w$, $\overline{W}_w = 0$ but $\xi \neq 0$. The second possibility, $a = 0 \neq b$, again splits into two cases, $b = w$ and $b = b_0$. The first yields $K = w\partial_t + \phi\partial_\eta$, $\mathscr{L}_K\overline{W} = 0$, $\gamma = c = e = 0$, $\xi \neq 0$, while the second gives $K = \partial_w$, $\overline{W}_w = 0$, and $\gamma = c = e = \xi = 0$. The case $a = b = 0 \neq c$ is not allowed for type [III] since one of the conditions states $\nu c = 0$ and we require $\nu \neq 0$.

Case 3: Petrov–Penrose type $[N] \otimes$ [Any] with $\chi_0 = 0$. Type N requires $\mu = \nu = 0$, $\gamma_t \neq 0$. For the condition $a \neq 0$, we find $e = 0$, $\xi \neq 0$, $K = \partial_w$, and $\overline{W}_w = 0$. One of the integration conditions states that $b\gamma_t = 0$, so the case $a = 0 \neq b$ is not allowed. Finally, for $a = b = 0 \neq c$, we obtain $e \neq 0 \neq \gamma$, $K = \partial_\eta$, and $\mathscr{L}_K W = (1/3\tau)\phi^3 e(w, t)$.

Case 4: The purely homothetic Killing vectors for all algebraically degenerate Petrov–Penrose types. For this case we have $\chi_0 \neq 0$ and choose $a = c = 0$. The general case is then a linear combination of the homothetic and isometric Killing vectors. For types II and D we must have $b = \frac{4}{3}\chi_0 t$ but we may eliminate ν, ξ, γ, and e and obtain $K = \chi_0\{2t\partial_t - \phi\partial_\phi + \eta\partial\eta\}$ and $\mathscr{L}_K\overline{W} = 0$. For type III, we must have $b = \chi_0 t$; we may gauge ξ and γ to zero, and are left with $K = \chi_0\{t\partial_t - \phi\partial\phi\}$, $\mathscr{L}_K\overline{W} = -\chi_0\overline{W} + (1/3\tau)\phi^3 e(w, t)$. Finally, for type N, we may have $a = b = c = 0$ leading to $e = 0$, $\xi = \xi_0$, $K = \chi_0\{\phi\partial\phi + \eta\partial\eta\}$, and $\mathscr{L}_K\overline{W} = 2\chi_0\overline{W}$.

We present the above results in tabular form (see Table I).

From the table we see that there are four possible distinct isometric Killing vectors for types II and D, three for type III, and only two for type N. Each type has one distinct kind of purely homothetic Killing vector allowed.

5. APPLICATIONS

The use of the above results will be demonstrated for two simple cases in order to give an idea of their application. First we will find the Killing vectors for the Schwarzschild metric and, secondly, we will find the metrics corresponding to a given Killing vector.

We obtain the Schwarzschild Killing vectors by taking the limit of the general type D Plebański–Demiański[11] solution to the Kerr metric, and then from Kerr to Schwarzschild. At the same time we perform similar limits on the complex extension of the PD metric, giving us the form of the metric function $Q^{\dot{A}\dot{B}}$ for the Schwarzschild case.

The PD metric can be brought to the Kerr form by applying the limits $(p, q) \to l^{-1}(p', q')$; $\tau \to l\tau'$, $\sigma \to l^3\sigma'$, $m \to l^{-3}m'$, $\epsilon \to l^{-2}\epsilon'$, $\gamma \to l^{-4}\gamma'$, and $n = e = g = \lambda = 0$ by making the identifications $p = -a\cos\theta$ and $q = r$. With this

$$ds^2 = (1 - pq)^{-2}\{(p^2 + q^2)\mathscr{P}^{-1}dp^2 + \mathscr{P}(p^2 + q^2)^{-1} \times (d\tau + q^2 d\sigma) + (p^2 + q^2)\mathscr{Q}^{-1}dq^2 - \mathscr{Q}(p^2 + q^2)^{-1} \times (d\tau - p^2 d\sigma)^2\} \tag{5.1}$$

becomes

$$ds^2 = -dt^2 + (r^2 + a^2)\sin^2\theta\, d\varphi^2 + (2mr/\Sigma) \times (dt - a\sin^2\theta\, d\varphi)^2 + \Sigma(\Delta^{-1}dr^2 + d\theta^2), \tag{5.2}$$

where $\Delta = r^2 + a^2 - 2mr$, $\Sigma = r^2 + a^2\cos^2\theta$, and $\mathscr{P}$, $\mathscr{Q}$ are structure functions involving p and q, respectively. The relationship between PD coordinates and Plebański–Robinson (PR) coordinates[12] is given by

$$x = -(q + ip)^{-1}, \quad y = pq(q + ip)^{-1},$$
$$u = \tau + \int \mathscr{Q}^{-1}q^2\, dq - 2\int \mathscr{P}^{-1}p^2\, dp,$$
$$v = i\sigma - i\int \mathscr{Q}^{-1}\, dq + \int \mathscr{P}^{-1}\, dp. \tag{5.3}$$

$$\mathbf{P} = \tfrac{1}{2}x^2(\mathscr{P} - \mathscr{Q})(p^2 + q^2)^{-1},$$
$$\mathbf{Q} = \tfrac{1}{2}x^2(p^4\mathscr{Q}^{-1} - q^4\mathscr{P})(p^2 + q^2)^{-1},$$
$$\mathbf{R} = -\tfrac{1}{2}ix^2(p^2\mathscr{Q} + q^2\mathscr{P})(p^2 + q^2)^{-1}, \tag{5.4}$$

where $Q^{\dot{A}\dot{B}} = \left(\begin{smallmatrix}\mathbf{Q}\mathbf{R}\\ \mathbf{R}\mathbf{P}\end{smallmatrix}\right)$, is the metric. Taking the limit as in PD to Kerr then yields the Kerr metric with the following identificiations:

$$\mathbf{P} \to \tfrac{1}{2}\phi^2\Sigma^{-1}(\Sigma - 2mr),$$
$$\mathbf{Q} \to -\tfrac{1}{2}\phi^2\Sigma^{-1}(r^4a^2\sin^2\theta - a^4\Delta\cos^4\theta),$$
$$\mathbf{R} \to -\tfrac{1}{2}i\phi^2\Sigma^{-1}\{r^2a^2\sin^2\theta + a^2\Delta\cos^2\theta\}, \tag{5.5}$$

$$\phi \to x = -(r - ia\cos\theta)^{-1}, \quad \eta \to -y = -xar\cos\theta,$$
$$du = -dt_{\text{Kerr}} + ad\varphi + \Delta^{-1}r^2dr - ia\cos^2\theta\csc\theta d\theta,$$
$$dv = ia^{-1}d\varphi - i\Delta^{-1}\, dr + a^{-1}\csc\theta\, d\theta.$$

It is clear that simply allowing a to go to zero does not yield a satisfactory limit for v and also causes $\mathbf{Q}$ and $\mathbf{R}$ to vanish. However, if we write out the metric and take the limit as $a \to 0$, we can identify choices for x, y, u, and v which will yield the Schwarzschild metric.

The following choices give the proper form:

$$dx = \phi^2\, dr, \quad dy = \sin\theta\, d\theta, \quad du = dt + \Delta^{-1}r^2\, dr,$$
$$dv = i\, d\varphi + \csc\theta\, d\theta, \quad \mathbf{P} = -\tfrac{1}{2}r^{-2}(1 - 2m/r),$$
$$\mathbf{Q} = -\tfrac{1}{2}\sin^2\theta, \quad \mathbf{R} = 0. \tag{5.6}$$

Passing over to x, y, w, t using the relations (5.3) and the definition of w and t in Sec. 4, one obtains

$$x = \phi = -r^{-1}, \quad -y = \cos\theta = \eta,$$
$$dw = -i\, d\varphi + (1 - \eta^2)^{-1}\, d\eta,$$
$$dt = -dt_S + \phi^{-2}(1 + 2m\phi)^{-1}\, d\phi. \tag{5.7}$$

Having the metric, we may find $\overline{W}$ from the relation $Q^{\dot{A}\dot{B}} = -2J^{(\dot{A}}\partial^{\dot{B})}\overline{W} - \partial^{\dot{A}}\partial^{\dot{B}}\overline{W} + \frac{1}{2}\mu\phi^3K^{\dot{A}}K^{\dot{B}}$, yielding

$$\overline{W} = \tfrac{1}{4}\phi(\eta^2 - 1). \tag{5.8}$$

Substituting this $\overline{W}$ into the hyperheavenly equation (2.2) gives the conditions

$$\gamma = \tfrac{1}{4}, \quad \xi = \nu = 0. \tag{5.9}$$

Insertion of $\overline{W}$ into the master equation (2.15) and use of the integration condition then gives the following forms for a, b, c, e, χ_0:

$$a = a_0 + c_0 e^{w} + d_0 e^{-w}, \quad b = b_0,$$
$$c = a_0 - a, \quad e = 0 = \chi_0. \tag{5.10}$$

Taken together, we have a four parameter group (a_0, b_0, c_0, d_0), with Killing vectors

$$K = a_0\partial_w + b_0\partial_t + c_0 e^{w}\{\partial_w + (1+\eta)\partial_\eta\} + d_0 e^{-w}\{\partial_w + (1+\eta)\partial_\eta\}. \tag{5.11}$$

The timelike vector is $b_0 \neq 0$, while the other three parameters generate the usual rotation group SO(3).

As a second example, we will demonstrate how a class of metrics is generated by a Killing vector; in this case we choose $K = \partial_\eta$ for types II and D, where $\nu = \xi = \gamma = 0$. When this information is inserted in the hyperheavenly equation we find a differential equation for $\overline{W}$:

$$\tau \overline{W}_{\phi t} = \mu[\phi^3 \overline{W}_{\phi\phi} - 3\phi^2 \overline{W}_\phi + 3\phi \overline{W}]. \tag{5.12}$$

Using $u = (\mu/\tau)t$ and $\overline{W} = T(u)X(\phi)$ to separate variables yields

$$T_u = -2\beta T \quad \text{and} \quad X_{\phi\phi}\phi^3 + (2\beta - 3\phi^2)X_\phi + 3\phi X = 0. \tag{5.13}$$

The general solution is then given by $\overline{W}$ such that

$$\overline{W}_\phi = \int_0^1 ds((1-s)/s)^{1/2} F(w) - \int_0^\infty ds\,((1+s)/s)^{1/2} G(v),$$
$$w = 2u - s/\phi^2, \quad v = 2u + s/\phi^2, \tag{5.14}$$

where F and G are arbitrary, sufficiently smooth functions of one complex variable, except that G must be such that the integrals converge.

We next calculate the Riemann curvature corresponding to this W, with $J_{\dot A}$ chosen such that $\phi = x = -p^2$ and $\eta = -y = p^1$. The components are

$$C_{\dot 2\dot 2\dot 2\dot 2} = \phi^3 \overline{W}_{\phi\phi\phi\phi}, \quad C_{\dot 2\dot 2\dot 2\dot 1} = C_{\dot 2\dot 1\dot 1\dot 1} = C_{\dot 1\dot 1\dot 1\dot 1} = 0, \quad C_{\dot 2\dot 2\dot 1\dot 1} = -(\mu_0/\tau^2)\phi^3; C^{(5)} = C^{(4)} = C^{(2)} = 0,$$
$$C^{(3)} = -2\mu_0\phi^3, \quad C^{(1)} = 6\mu_0\phi^7\tau\overline{W}_t, \tag{5.15}$$

where

$$\overline{W}_{\phi\phi\phi\phi} = -\phi^{-3}\int_0^1 [s(1-s)]^{-1/2}\{12(1-2s)F + 9s(3-4s)F_s + 2s^2(5-6s)F_{ss}\}\,ds$$
$$+\phi^{-3}\int_0^\infty [s(1+s)]^{-1/2}\{12(1+2s)G + 9s(3+4s)G_s + 2s^2(5+6s)G_{ss}\}\,ds. \tag{5.16}$$

6. CONCLUSIONS

In this paper we have determined the Killing structure of the expanding hyperheavens of Plebański and Robinson. We have also deduced an invariant classification of these Killing vectors and have determined a master equation which gives the form that the Lie derivative must have in order to admit a given Killing vector.

This work, combined with the earlier work of Finley and Plebański, completes the determination of the Killing structure of all one-sided algebraically-degenerate, complexified, Einstein vacuum space–times. These results then allow one to determine the symmetries (Killing vectors) of the metric of any hyperheaven and, conversely, to determine all metrics which allow a given Killing vector. In addition, the present work once again demonstrates the power of the spinor approach.

No attempt has been made to determine the real cross sections on these metrics in a systematic way, nor has an attempt been made to find all metrics corresponding to each of the Killing vectors presented here.

[1]C. P. Boyer, J. D. Finley, III, and J. F. Plebański, "Complex General Relativity, $\mathscr{H}$ and $\mathscr{H}\mathscr{H}$ Spaces—A Survey of One Approach," *Einstein Memorial Volume, Society for General Relativity and Gravitation* (Plenum, New York, 1979).

[2]J. F. Plebański and S. Hacyan, J. Math. Phys. **16**, 2403 (1975).

[3]J. F. Plebański and I. Robinson, Phys. Rev. Lett. **37**, 493 (1976).

[4]J. D. Finley, III and J. F. Plebański, J. Math. Phys. **17**, 2207 (1976).

[5]J. D. Finley, III and J. F. Plebański, J. Math. Phys. **17**, 585 (1976).

[6]J. D. Finley, III and J. F. Plebański, J. Math. Phys. **20**, 1938 (1979).

[7]J. D. Finley, III and J. F. Plebański, J. Math. Phys. **19**, 760 (1978).

[8]C. P. Boyer and J. F. Plebański, J. Math. Phys. **18**, 1022 (1977).

[9]S. Hacyan and J. F. Plebański, J. Math. Phys. **17**, 2203 (1976).

[10]S. A. Sonnleitner, Ph. D. Dissertation, University of New Mexico, 1980.

[11]J. F. Plebański and M. Demiański, Ann. Phys. (N.Y.) **98**, 98 (1976).

[12]A. Garcia and J. F. Plebański, Nuovo Cimento B **40**, 224 (1977).

An exceptional type D shearing twisting electrovac with λ

Alberto García D. and Jerzy F. Plebański[a)]
Centro de Investigación y Estudios Avanzados del I.P.N., Apartado Postal 14-740, México 14, D. F., México

(Received 15 September 1980; accepted for publication 19 December 1980)

A new electrovac with λ type D solution and with both Debever–Penrose vectors aligned along the real eigenvectors of the electromagnetic field is presented. The principal null directions are shearing and twisting. The existence of this solution, endowed with an O(2,1,$\mathbb{R}$) symmetry, requires $\lambda < 0$.

PACS numbers: 04.20.Jb

1. INTRODUCTION

This work is a sequel of a previous paper (Plebański and Hacyan[1]) dealing with exceptional electrovac with λ type D metrics; it corrects a serious error in that paper. The quoted paper belongs to a sequence of articles which had the objective of determining *all* type D electrovac solutions with λ which have the principal null directions aligned along the real eigenvectors of the electromagnetic field. With the null tetrad members e^3 and e^4 (we use the same notation as in Refs. 1–3) oriented along principal null directions, the conformal curvature has only the component $C^{(3)} \neq 0$; then, with the invariant (complex) of the electromagnetic field defined by

$$\mathscr{F} := \tfrac{1}{4} f_{\mu\nu} f^{\mu\nu} + \tfrac{i}{4} f_{\mu\nu}, \quad \check{f}^{\mu\nu} = : -\tfrac{1}{2}(E + i\check{B})^2, \qquad \check{f}^{\mu\nu} := (i/2\sqrt{-g})\epsilon^{\mu\nu\rho\sigma} f_{\rho\sigma}, \tag{1.1}$$

the invariant

$$I := (E^2 + \check{B}^2)^2 - (\tfrac{3}{2}C^{(3)})^2 \tag{1.2}$$

plays a crucial role in the problem considered. If $I \neq 0$, the Goldberg–Sachs[4] theorem applies and e^3 and e^4 must be geodesic and shearless. Moreover, by a theorem of Hughston *et al.*[5] the solution admits two commuting Killing vectors. The solutions of this type with complex expansion $Z \neq 0$ are completely described—modulo contractions—in the presence of λ in terms of the Plebański–Demiański[6] metrics. The case $Z = 0$ has also been covered completely in the recent work of Plebański.[2] The *exceptional* electrovac solutions with λ arise when $I = 0$, with two subpossibilities:

$$E_{(+)} : E^2 + \check{B}^2 = \tfrac{3}{2}C^{(3)} \neq 0, \quad E_{(-)} : E^2 + \check{B}^2 = -\tfrac{3}{2}C^{(3)} \neq 0. \tag{1.3}$$

In the two subcases $E_{(+)}$ and $E_{(-)}$, the Goldberg–Sachs and Hughston theorems do not apply. In the paper by Plebański and Hacyan[1] the branch $E_{(+)}$ was correctly integrated; it contains the Bertotti–Robinson[7,8] solution, with both e^3 and e^4 geodesic and shearless, together with an exceptional solution given in a chart $\{x^\mu\} = \{\xi, \bar{\xi}, u, v\}$ by

$$g = 2e^1 \otimes e^2 + 2e^3 \otimes e^4,$$
$$\omega := \tfrac{1}{2}(f_{\mu\nu} + \check{f}_{\mu\nu}) dx^\mu \wedge dx^\nu = (E + i\check{B})(e^1 \wedge e^2 + e^3 \wedge e^4), \tag{1.4}$$

where

$$e^1 = d\xi, \quad e^2 = d\bar{\xi}, \quad e^4 = dv,$$
$$e^3 = du + \left[\lambda u^2 + \xi\, \overline{F(v)} + \bar{\xi} F(v)\right] dv,$$
$$E + i\check{B} = \text{const}, \quad \tfrac{3}{2}C^{(3)} = E^2 + \check{B}^2 = -\lambda > 0, \tag{1.5}$$

with $F = F(v) = A(v) + iB(v)$ containing two arbitrary (real) functions of one variable. With $F \neq 0$, the principal null direction e^3 *is not* geodesic, and with $F(v)$ being a general enough function, the solution does not admit any Killing vector whatsoever.

In Sec. 3 of Ref. 1 the subcase $E_{(-)}$ was also treated. The logical chain up to (3.38) of Ref. 1 is correct, but the argument given after (3.38), which attempts to show that a nontrivial solution within the branch $E_{(-)}$ does not exist, is wrong. In the subsequent formulae after (3.38) in $zd\bar{z}$-$\bar{z}dz$-dv, the factor "i" at dv was missed. The present paper integrates correctly the branch $E_{(-)}$ (in a different tetrad gauge from that of the formulae of Sec. 3 of Ref. 1) determining the most general form of the nontrivial solution which exists when $\lambda < 0$. This solution, with both principal null directions geodesic but shearing and twisting and nonexpanding, is of interest for reasons of completeness within the category of D-type electrovac solutions with λ which have principal null directions aligned along real eigenvectors of the electromagnetic field. The solution is also of some interest, *per se*, exhibiting rather unusual singularities of the congruences of the principal geodesic null directions, which are characterized by infinite values of the shear and the twist.

2. THE DIFFERENTIAL PROBLEM AND ITS INTEGRAL

The differential problem of the exceptional branch $E_{(-)}$ states in the null tetrad formalism (notation is the same as that which was used in Refs. 1–3) can be summarized as follows. The tetrad e^a and connection 1-forms $\Gamma_{ab} = \Gamma_{[ab]}$ must satisfy the first structure equations

$$de^a + \Gamma^a{}_b \wedge e^b = 0, \tag{2.1}$$

while the second Cartan structure equations with built in Einstein equations $G_{\mu\nu} = 8\pi E_{\mu\nu} + \lambda g_{\mu\nu}$ amount to

$$\begin{aligned} d\Gamma_{42} + \Gamma_{42} \wedge (\Gamma_{12} + \Gamma_{34}) &= -2\mu e^3 \wedge e^1, \\ d\Gamma_{31} + (\Gamma_{12} + \Gamma_{34}) \wedge \Gamma_{31} &= -2\mu e^4 \wedge e^2, \\ d(\Gamma_{12} + \Gamma_{34}) + 2\Gamma_{34} \wedge \Gamma_{31} &= -(10\mu + 2\lambda) e^1 \wedge e^2 \\ &\quad + 2\mu e^3 \wedge e^4, \end{aligned} \tag{2.2}$$

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

where, with electromagnetic field given by (1.4),

$$\mu = -\tfrac{1}{6}(\lambda - E^2 - \check{B}^2), \tag{2.3}$$

and the Maxwell–Faraday equations take the form

$$d \ln(E + i\check{B})^{1/2} + \Gamma_{314}e^1 + \Gamma_{423}e^2 - \Gamma_{312}e^3 - \Gamma_{421}e^4 = 0. \tag{2.4}$$

The Bianchi identities require, with the $E_{(-)}$ condition from (1.3) assumed, that e^3 and e^4 be geodesic; that is,

$$\Gamma_{424} = 0 = \Gamma_{313}. \tag{2.5}$$

This granted, the Bianchi identities reduce to

$$dC^{(3)} = 6C^{(3)}(\Gamma_{314}e^1 + \Gamma_{423}e^2). \tag{2.6}$$

Within this formulation of the $E_{(-)}$ branch differential problem, we are still left with the freedom of the phase and boost gauge of the null tetrad; i.e.,

$$e'^1 = e^{i\phi}e^1, \quad e'^2 = e^{-i\phi}e^2, \quad e'^3 = e^{\chi}e^3, \quad e'^4 = e^{-\chi}e^4, \tag{2.7}$$

where the functions (real) ϕ and χ are arbitrary. The connections transform under this gauge correspondingly (see Ref. 1).

Now, the $E_{(-)}$ condition from (1.3) demands that $C^{(3)}$ is real. Hence, consistency of (2.6) requires

$$\Gamma_{314} = \Gamma_{413}, \quad \Gamma_{423} = \Gamma_{324}. \tag{2.8}$$

Moreover, combining (2.5) and (2.4), one easily infers that

$$\Gamma_{423} = 0 = \Gamma_{314}, \quad \Gamma_{421} + \Gamma_{412} = 0 = \Gamma_{312} + \Gamma_{321}, \tag{2.9}$$

and, therefore, $C^{(3)}$ and $E^2 + \check{B}^2$ are constants. Hence μ is also a constant. Consequently, setting

$$E + iB = :(E^2 + \check{B}^2)^{1/2}e^{-2i\psi}, \tag{2.10}$$

we infer that (2.4) reduces to

$$-id\psi = \Gamma_{312}e^3 + \Gamma_{421}e^4. \tag{2.11}$$

In the next step, employing (2.1), one deduces from (2.2) the further necessary algebraic conditions

$$(\Gamma_{421})^2 + \Gamma_{422}\overline{\Gamma}_{422} = 0 = (\Gamma_{312})^2 + \Gamma_{311}\overline{\Gamma}_{311}, \tag{2.12}$$

$$2\Gamma_{421}\Gamma_{312} - \Gamma_{422}\Gamma_{311} - \overline{\Gamma}_{422}\overline{\Gamma}_{311} + 4\mu = 0.$$

If $\mu = 0 \rightarrow E^2 + \check{B}^2 = \lambda = -\tfrac{3}{2}C^{(3)}$, one can select a tetrad gauge (2.7) such that $\Gamma_{42} = 0 = \Gamma_{31}$. In the subcase the solution reduces to the Bertotti–Robinson solution with e^3 and e^4 geodesic and shearless [see comments after (3.17) of Ref. 1]. The nontrivial branch of the problem considered arises when $\mu \neq 0$. With $\mu \neq 0$, denoting $\epsilon = \text{sign}(\mu)$, $\epsilon^2 = 1 \rightarrow \epsilon\mu > 0$, and employing the freedom of (2.7) gauge, one can always select the tetrad so that

$$\Gamma_{42} = \frac{i\sqrt{\epsilon\mu}}{\sin\nu}(e^1 + e^{i\nu}e^2), \quad \Gamma_{31} = \frac{i\epsilon\sqrt{\epsilon\mu}}{\sin\nu}(e^{i\nu}e^1 + e^2). \tag{2.13}$$

With this tetrad gauge (2.12) reduce to identities and (2.11) takes the form

$$-d\psi = (\sqrt{\epsilon\mu}/\sin\nu)(\epsilon e^3 + e^4). \tag{2.14}$$

Using (2.13) in (2.2)—accompanied by (2.1) and (2.14)—one extracts then the complete information contained in the second structure equations. One easily finds that consistency of (2.2) requires

$$\mu = -\lambda/2 \rightarrow -\tfrac{3}{2}C^{(3)} = E^2 + \check{B}^2 = -2\lambda \rightarrow \lambda < 0, \tag{2.15}$$

so that within the nontrivial $E_{(-)}$ branch, $\lambda < 0$ and $\epsilon = 1$. This granted and integrating the remaining information contained in (2.2) and (2.14), one finds that

$$\begin{aligned}
\Gamma_{42} &= \frac{i\sqrt{(-\lambda/2)}}{\sin\nu}(e^1 + e^{i\nu}e^2),\\
\Gamma_{31} &= \frac{i\sqrt{(-\lambda/2)}}{\sin\nu}(e^{i\nu}e^1 + e^2),\\
\Gamma_{34} &= \sqrt{(-\lambda/2)}\cos\nu(e^3 - e^4),\\
\Gamma_{12} &= -(i/2)\sqrt{(-\lambda/2)}\sin\nu\,(e^3 - e^4),\\
\Psi &= \cot\nu + \Psi_0,
\end{aligned} \tag{2.16}$$

with Ψ_0 being constant, (2.14) assumes now the equivalent form

$$d\nu = \sqrt{(-\lambda/2)}\sin\nu(e^3 + e^4). \tag{2.17}$$

Given (2.16) and (2.17), the equations modulo (2.1) are now identities. It remains thus to integrate the first structure equations (2.1), which if we introduce the 1-forms

$$\begin{aligned}
\Gamma_1 &:= \sqrt{(-\lambda/2)}\cot(\nu/2)(e^1 + e^2) = :\Gamma^1,\\
\Gamma_2 &:= -i\sqrt{(-\lambda/2)}\tan(\nu/2)(e^1 - e^2) = :\Gamma^2,\\
\Gamma_3 &:= \tfrac{1}{2}\sqrt{(-\lambda/2)}\sin\nu(e^3 - e^4) = : -\Gamma^3,
\end{aligned} \tag{2.18}$$

are equivalent to (2.17) and

$$d\Gamma_a + \tfrac{1}{2}\epsilon_{abc}\Gamma^b \wedge \Gamma^c = 0, \quad a,b,\cdots = 1,2,3, \tag{2.19}$$

or explicitly

$$\begin{aligned}
&d\Gamma_1 - \Gamma_2 \wedge \Gamma_3 = 0, \quad d\Gamma_2 - \Gamma_3 \wedge \Gamma_1 = 0,\\
&d\Gamma_3 + \Gamma_1 \wedge \Gamma_2 = 0.
\end{aligned} \tag{2.20}$$

We observe also that (2.17) and the definitions (2.18) imply

$$0 \neq -i\tfrac{1}{2}\lambda^2 e^1 \wedge e^2 \wedge e^3 \wedge e^4 = \Gamma_1 \wedge \Gamma_2 \wedge \Gamma_3 \wedge d(\cot\nu). \tag{2.21}$$

The real Γ_a's can now be interpreted as the connections of the group O(2,1,ℝ) and there are many obvious manners of expressing the Γ_a's in terms of three independent group parameters, say p_1, p_2, p_3 which together with ν can therefore be considered as the chart $\{x^\mu\} = \{p_1, p_2, p_3, \nu\}$. Some explicit parametrizations of the Γ_a's will be discussed below. At this moment, however, it is convenient to summarize the result obtained in terms of abstract Γ_a's and the coordinate ν, assuming $\Gamma_1 \wedge \Gamma_2 \wedge \Gamma_3 \wedge d(\cot\nu) \neq 0$. The metric g and the electromagnetic field ω from (1.4), with $\lambda < 0$, have the form

$$\begin{aligned}
(-\lambda)g &= \tan^2(\nu/2)\Gamma_1 \otimes \Gamma_1 + \cot^2(\nu/2)\Gamma_2 \otimes \Gamma_2\\
&\quad + \frac{d\nu}{\sin\nu} \otimes \frac{d\nu}{\sin\nu} - \frac{4}{\sin^2\nu}\Gamma_3 \otimes \Gamma_3,\\
\sqrt{(-\lambda/2)}\omega &= ie^{-2i\psi_0}d(e^{-2i\cot\nu}\Gamma_3),
\end{aligned} \tag{2.22}$$

the invariants of the electromagnetic field being

$$E + i\check{B} = 2\sqrt{(-\lambda/2)}e^{-2i\psi_0}e^{-2i\cot\nu}. \tag{2.23}$$

According to (2.16), the principal null directions e^3 and e^4 of this type D solution of the electrovac equations with λ are geodesic and possess the common complex expansion Z

and shear S given by

$$-\Gamma_{421} = -\Gamma_{312} = -i\sqrt{(-\lambda/2)}/\sin\nu = :Z,$$
$$\Gamma_{422} = \Gamma_{311} = \sqrt{(-\lambda/2)}(i\cot\nu - 1) = :S. \tag{2.24}$$

With Z pure imaginary, the congruences e^3 and e^4 are divergenceless but twisting; for $\nu = n\pi, n = \cdots -1,0,1,\cdots, Z$ and S are singular.

The solution described above contains only the constants $\lambda < 0$ and ψ_0, the last corresponding to the remaining freedom of the (constant) duality rotations of the electromagnetic field. Observe that if one uses in place of ν the equivalent variable z defined by

$$\tan(\nu/2) = :e^z \rightarrow d\nu/\sin\nu = dz, \tag{2.25}$$

the basic formulae which describe our solution assume the slightly simpler form

$$\begin{aligned}(-\lambda)\cdot g &= e^{2z}\Gamma_1 \otimes \Gamma_1 + e^{-2z}\Gamma_2 \otimes \Gamma_2 \\ &\quad + dz \otimes dz - (2\cosh z)^2 \Gamma_3 \otimes \Gamma_3, \\ \sqrt{(-\lambda/2)}\cdot\omega &= ie^{-2i\psi_0} d\,(e^{-2i\sinh z}\Gamma_3), \\ E + i\check{B} &= 2\sqrt{-(\lambda/2)}e^{-2i\psi_0}\cdot e^{-2i\sinh z}, \\ Z &= -i\sqrt{(-\lambda/2)}\cdot\cosh z, \\ S &= \sqrt{(-\lambda/2)}(i\sinh z - 1),\end{aligned} \tag{2.26}$$

with the values $z = \pm\infty$ being singular.

3. PARAMETRIZATIONS Γ_a's AND SYMMETRIES OF THE $E_{(-)}$ SOLUTION

Let k_A and l_A be a pair of real spinors normalized so that

$$k^A l_A \equiv k_1 l_2 - k_2 l_1 = 1. \tag{3.1}$$

Then

$$\Gamma_1 = l^A dk_A + k^A dl_A \equiv 2k^A dl_A \equiv 2l^A dk_A,$$
$$\Gamma_2 = k^A dk_A - l^A dl_A, \quad \Gamma_3 = k^A dk_A + l^A dk_A, \tag{3.2}$$

is the most general analytic form of the SL(2,$\mathbb{R}$) [equivalently, O(2,1,$\mathbb{R}$)] connections which satisfy (2.20). Parametrizing k_A and l_A in terms of real u and complex f according to

$$e^{i\pi/4}k_A + e^{-i\pi/4}l_A = [e^{iu/2}/(i - f\bar{f})^{1/2}] \times [(1 - if)\delta_A^1 - i(1 + if)\delta_A^2], \tag{3.3}$$

so that (3.1) reduces to an identity, the connections are

$$\Gamma_1 + i\Gamma_2 = 2\frac{e^{iu}}{1 - f\bar{f}}df, \quad \Gamma_3 = du - i\frac{fd\bar{f} - \bar{f}df}{1 - f\bar{f}}. \tag{3.4}$$

With the Γ_a's so understood, (2.26) gives our solution in terms of the chart $\{x^\mu\} = \{u,f,\bar{f},z\}$. Equivalently, using the chart $\{x^\mu\} = \{u,g,\bar{g},z\}$, where

$$f = :ge^{-iu}, \quad \bar{f} = :\bar{g}e^{iu}, \tag{3.5}$$

we have

$$\Gamma_1 + i\Gamma_2 = 2\frac{dg - igdu}{1 - g\bar{g}}, \quad \Gamma_3 = \frac{1 + g\bar{g}}{1 - g\bar{g}}du - i\frac{gd\bar{g} - \bar{g}dg}{1 - g\bar{g}}. \tag{3.6}$$

Working in the last chart, one easily finds that our $E_{(-)}$ solution always has three Killing vectors; namely,

$$\underset{1}{k} = \partial_u, \quad \underset{2}{k} = e^{iu}(\bar{g}\partial_u - i[1 - g\bar{g}]\partial_g),$$
$$\underset{2}{\bar{k}} = e^{-iu}(g\partial_u + i[1 - g\bar{g}]\partial_{\bar{g}}), \tag{3.7}$$

which close in the SL(2,$\mathscr{R}$) algebra

$$[\underset{1}{k},\underset{2}{k}] = i\underset{2}{k}, \; [\underset{1}{k},\underset{2}{\bar{k}}] = -i\underset{2}{\bar{k}}, \; [\underset{2}{k},\underset{2}{\bar{k}}] = -2i\underset{1}{k}. \tag{3.8}$$

Therefore, the $E_{(-)}$ solution always has a three parameter group of symmetries.

4. CONCLUSIONS

The results of this paper correct the conclusions of Ref. 1 in the form of the following statement: the electrovac with λ type D solutions with principal null directions aligned along real eigenvectors of the electromagnetic field are exhausted by the solutions from two classes; the regular (general) class with principal null directions being geodesic and shearless, consisting of the solutions of Refs. 6 and 3 including those of Refs. 7 and 8; and the exceptional class, characterized by the condition $I = (E^2 + B^2)^2 - (\frac{3}{2}C^{(3)})^2 = 0$, where the Goldberg–Sachs theorem does not apply. Nontrivial exceptional solutions exist if and only if $\lambda < 0$, and are exhausted by the solutions of type $E_{(+)}$ with one of the principal null directions being nongeodesic, and type $E_{(-)}$ characterized by principal null directions being geodesic, nonexpanding but shearing and twisting. The Robinson–Bertotti solution is a trivial solution with $I = 0$.

[1] J. Plebański and S. Hacyan, J. Math. Phys. **20**, 1004 (1979).
[2] J. Plebański, J. Math. Phys. **20**, 1946 (1979).
[3] G. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1942 (1969).
[4] J. Goldberg and R. Sachs, Acta Phys. Pol. Suppl. **22**, 13 (1962).
[5] L. P. Hughston, R. Penrose, and P. Sommers, Commun. Math. Phys. **27**, 303 (1972).
[6] J. Plebański and M. Demiański, Ann. Phys. **90**, 280 (1975).
[7] I. Robinson, Bull. Acad. Pol. Sci. Ser. Sci. Math. Astron. Phys. **7**, 35 (1959).
[8] B. Bertotti, Phys. Rev. **116**, 1331 (1959).

A Schwarzschild-like interior solution for charged spheres

K. D. Krori
Mathematical Physics Forum, Cotton College, Gauhati-781001, Assam, India
B. B. Paul
Nowgong College, Nowgong, Assam, India

(Received 3 September 1980; accepted for publication 31 October 1980)

A Schwarzschild-like interior solution for charged spheres is presented in this paper. The solution is regular everywhere.

PACS numbers: 04.40. + c, 04.20.Jb

I. INTRODUCTION

The study of charged matter distribution in general relativity has attracted wide attention recently. De and Ray–Choudhuri[1] have shown that a charged dust distribution in equilibrium will have the absolute value of the charge to mass ratio as unity in relativistic units, i.e., $c = G = 1$. Efinger,[2] Bailyn and Eimeral[3], and Nduka[4] have obtained some solutions of static spherical distributions which are not free from singularity at the origin. On the other hand, Kyle and Martin,[5] Wilson,[6] Kramer, and Neugebauer,[7] Krori and Barua,[8] and Chakraborty and De[9] have presented solutions for charged fluid spheres which are very regular. But no solution has so far been derived such that it may reduce to the Schwarzschild interior solution in the absence of charge. In this paper we present a new solution for charged spheres which has the interesting feature that it reduces to the Schwarzchild interior solution for a fluid sphere in the absence of charge. But it possesses the peculiarity that in the presence of charge, it does not represent a fluid distribution because it sustains tangential stress. The solution, however, is regular everywhere and is free from any singularity.

II. DERIVATION OF THE SOLUTION

We take the interior metric in the form

$$ds^2 = e^{\nu} dt^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2) - e^{\lambda} dr^2 \tag{1}$$

where $\lambda = \lambda(r)$ and $\nu = \nu(r)$. The field equations are

$$8\pi p_r - E = e^{-\lambda}\left(\frac{\nu_1}{r} + \frac{1}{r^2}\right) - \frac{1}{r^2}, \tag{2}$$

$$8\pi p_\theta + E = 8\pi p_\phi + E = e^{-\lambda}\left(\frac{\nu_{11}}{2} - \frac{\lambda_1 \nu_1}{4} + \frac{\nu_1^2}{4} + \frac{\nu_1 - \lambda_1}{2r}\right), \tag{3}$$

$$8\pi\rho + E = e^{-\lambda}\left(\frac{\lambda_1}{r} - \frac{1}{r^2}\right) + \frac{1}{r^2}, \tag{4}$$

where p_r is the radial pressure, p_θ and p_ϕ are the tangential stresses, $E = -F_{01}F^{01}$,F_{01} being the electric field, and the suffix 1 indicates differentiation with respect to r. Maxwell's equation is give by

$$\frac{d}{dr}(\sqrt{-g}\,F^{01}) = 4\pi\sigma v^0\sqrt{-g}, \tag{5}$$

where σ is charge density and $V^{\mu} = \delta^{\mu}_{0}/\sqrt{g_{00}}$.

Now since there are four equations and seven variables let us assume

$$e^{\nu} = (A - B\sqrt{x})^2, \tag{6}$$

$$e^{\lambda} = 1/x, \tag{7}$$

and

$$E = a^2r^2\sqrt{x}/(A - B\sqrt{x}), \tag{8}$$

where A,B, and a are constants and

$$X = 1 - Kr^2 + Qr^4, \tag{9}$$

K and Q being constants.

With Eq. (1) the Eqs. (2)–(5) give

$$8\pi p_r = \frac{K(3B\sqrt{x} - A) + Qr^2(A - 5B\sqrt{x}) + a^2r^2\sqrt{x}}{A - B\sqrt{x}}, \tag{10}$$

$$8\pi p_\theta = 8\pi p_\phi = \frac{K(3B\sqrt{x} - A) + 2Qr^2(A - 5B\sqrt{x}) - a^2r^2\sqrt{x}}{A - B\sqrt{x}}, \tag{11}$$

$$8\pi\rho = 3K - 5Qr^2 - \frac{a^2r^2\sqrt{x}}{A - B\sqrt{x}}, \tag{12}$$

$$4\pi\sigma = [3a - (a/4)r(\lambda_1 + \nu_1)]e^{-(3\lambda + \nu)/4}. \tag{13}$$

Now at $r = 0$ (center)we have from (10)–(13)

$$8\pi(p_r)_0 = 8\pi(p_\theta)_0 = 8\pi(p_\phi)_0 = \frac{K(3B - A)}{(A - B)}, \tag{14}$$

$$8\pi\rho_0 = 3K, \tag{15}$$

$$8\pi\sigma_0 = 6a/\sqrt{A - B}, \tag{16}$$

and Eq. (8) gives

$$E = 0. \tag{17}$$

At the exterior since $(e^{\nu}\cdot e^{\lambda}) = 1$, we have

$$(A - B\sqrt{x_1})^2/x_1 = 1, \tag{18}$$

where x_1, is the value of x at the boundary. With

$$A = \tfrac{3}{2}\sqrt{x_1} \text{ and } B = \tfrac{1}{2} \tag{19}$$

Eq. (18) is satisfied. Again for e^{ν} to be continuous at $r = r_1$ we have

$$(A - B\sqrt{x_1})^2 = 1 - 2m/r_1 + e^2/r_1^2, \tag{20}$$

where m and e are respectively the total mass and charge of the sphere.

From Eq. (19) and (20) we have

 0022-2488/82/010126-02$02.50

$$x_1 = 1 - 2m/r_1 + e^2/r_1^2. \tag{21}$$

With Eq. (19) we have for the value of p_r at the boundary

$$8\pi(p_r)_1 = (-Q + a^2)r^2. \tag{22}$$

since $(p_r)_1$ is zero at the boundary we have

$$Q = a^2 \tag{23}$$

The value of ρ at the boundary using (19) is given by

$$8\pi\rho_1 = 3K - 6a^2r_1^2. \tag{24}$$

For ρ_1 to be positive at the boundary we have

$$r_1 < (K/2a^2)^{1/2}. \tag{25}$$

The continuity of E at the boundary $r = r_1$ gives

$$E_1 = \frac{a^2r_1^2\sqrt{x_1}}{A - B\sqrt{x_1}} = \frac{e^2}{r_1^4}. \tag{26}$$

Thus from (18) and (26) we have

$$e = ar_1^3. \tag{27}$$

From $A = \frac{3}{2}\sqrt{x_1}$ we have

$$r_1^2 = \frac{K - [K^2 - 4Q(1 - \frac{4}{9}A^2)]^{1/2}}{2Q}. \tag{28}$$

The value of r_1 given by (28) must satisfy the condition (25). This is possible if $A < \frac{3}{2}$. Again $\rho_0 \geqslant 3(p_r)_0$ gives $A \geqslant 1$. Thus A lies between 1 and $\frac{3}{2}$. From (28) we have

$$K = Qr_1^2 + (1/r_1^2)(1 - \frac{4}{9}A^2). \tag{29}$$

Now from (21), using K from (29) and e from (27), we have

$$m = \tfrac{1}{2}[Qr_1^5 + (1 - \tfrac{4}{9}A^2)r_1], \tag{30}$$

which is the geometrical mass of the charged sphere. It appears that the charge makes contribution to this mass.

ACKNOWLEDGMENTS

The authors are grateful to the government of Assam for providing facilities to carry out this piece of work at Cotton College, Gauhati-781001. One of them (B.B.P) is thankful to U.G.C., New Delhi, for financial assistance.

[1] U.K. De and A. K. Raychaudhuri, Proc. R. Soc. London, Ser. **A 303**, 97–101, (1968).

[2] H. J. Efinger, Z. Phys. **188**, 31–37 (1965).

[3] M. Bailyn and D. Eimeral, Phys. Rev. D **5**, 1897–1907 (1972).

[4] A. Nduka, Acta. Polon B **8**, 75–79 (1977).

[5] C. F. Kyle, and A. W. Martin, Nuovo Cimento **50**, 583–604 (1967).

[6] S. J. Wilson, Can. J. Phys. **47**, 2401–04 (1967).

[7] D. Kramer and G. Neugebauer, Ann. Phys. (Leipzig) **27**, 129 (1971).

[8] K. D. Krori and J. Barua, J. Phys. A **8**, 508–11 (1975).

[9] P. C. Chakraborty and U.K.De, Indian J. Pure Appl. Math. **10**, 608–16, (1979).

Jordan–Kaluza–Klein type unified theories of gauge and gravity fields

T. Bradfield and R. Kantowski
Department of Physics and Astronomy, University of Oklahoma, Norman, Oklahoma 73019

(Received 15 April 1980; accepted for publication 17 June 1980)

We investigate the generalized Jordan–Kaluza–Klein type scalar tensor theories of gravity with gauge fields present for the purpose of restricting the spacetime dependence of the scalar fields. These scalars are essential in building Lagrangians for fields with internal degrees of freedom. By one rather simple consistency restriction on covariant differentiation we are able to show that the scalar fields must be spacetime constants.

PACS numbers: 04.50. + h

I. INTRODUCTION

The fiber bundle structure of the generalized Kaluza–Klein theory was described by Trautman and its extension to contain scalar fields (generalized Jordan Theory) was presented by Cho and Freund.[1–3] The geometry used here will be almost identical with that used in these references. Spacetime, M, is assumed to be endowed with a metric g_{st} and to be the base space of a principal G bundle P, for some gauge group G. The gauge field is assumed given by a connection form ω on this bundle. If we denote the action of G on P by $\Psi_a:P\rightarrow P$ for $a\in G$ and the projection of P onto M by π, then $\Psi^*_a\omega = \mathrm{ad}'_{a^{-1}}\circ\omega$ and $\pi\circ\Psi_a = \pi$. To proceed further and do physics we must provide equations of motion for, and interactions between, the two given fields g_{st} and ω as well as their physical interpretation, e.g., we need to supply a Lagrangian. We will follow the previous references and construct a metric on P and from it a Ricci scalar which will serve as a Lagrangian.[1–4] In the next section we try to make clear the procedure for building certain tensors on P from tensors on M and from tensor fields of type ad′ on the bundle. The purpose is to enlarge the stage for events from the four dimensions of spacetime to include the N internal degrees of freedom associated with the gauge group. In Sec. III we define a metric g_P on P and relate its covariant derivatives to those of g_{st} on M and those of ω. In Sec. IV we go on to show that the spacetime scalars that appear in g_P are constants. In Sec. V the geodesics of g_P are projected onto M and seen to be classical particle trajectories. We conclude by giving the Lagrangian for the gravity and gauge fields.

II. LIFTING TENSOR FIELDS FROM THE BASE SPACE AND SLIDING OVER ADJOINT FIELDS FROM THE LIE ALGEBRA

The purpose of this section is to show that two types of physical fields—those defined as tensors on the base space (spacetime) e.g., stress–energy–momentum, and those defined as tensors of type ad′, e.g., standard Higgs fields, can both be uniquely associated with group invariant horizontal or vertical tensors defined on P (e.g., in TP or TP^*, etc.). This procedure also allows the introduction of mixed type fields, i.e., those with horizontal and vertical components, e.g., conserved currents.

First we lift spacetime tensors. Any form on M can be lifted onto P by the pullback of π. Given a connection form ω, any vector field v on M can be uniquely lifted as a horizontal vector field h on P. If h = lift of v from TM_x to TP_x, then $\omega(h) = 0$ and $\pi_* h = v$. Given a coordinate chart, x^c, on M we denote the lifts of $\partial/\partial x^c$ and dx^c by h_c and θ^c respectively. Any tensor field t on M can be lifted horizontally as a tensor field T on P,

$$t(x) = t^{a\cdots}{}_{b\cdots}(x)\frac{\partial}{\partial x^a}\otimes\cdots\otimes dx^b\otimes\cdots,$$

$$T(p) = t^{a\cdots}{}_{b\cdots}(\pi(p))h_a\otimes\cdots\otimes\theta^b\otimes\cdots.$$

These horizontal fields are invariant under the action of the group, i.e., $T(\Psi_a p) = \Psi_a T(p)$ where vector components transform by Ψ_{a^*} and dual components transform by $\Psi^*_{a^{-1}}$.

Now for adjoint fields.[5] Let $A':P\rightarrow G'$ be a G'-valued field on P of type ad′, i.e., $A'(\Psi_a p) = \mathrm{ad}'_{a^{-1}}A'(p)\in G'$, where G' is the Lie algebra of left invariant vector fields on the gauge group G and where ad′ is the adjoint action of the group on G'. We wish to "slide" the field A' back from G' into the vertical tangent space of P. This we can do because of the Killing or fundamental map Σ of $G'\rightarrow TP$; i.e., to each $L\in G'$ there corresponds a unique vertical vector field l on P defined by $l(p)=$ pushover of $L(e)$ by the map $\Psi_{(\,)}p:G\rightarrow P$. The unique vertical field A on P can then be defined by $A(p) = \Sigma(A'(p))|_{\mathrm{at}\,p}$. It is easy to see that the ad′ invariance of A' implies the right invariance of A, i.e., $\Psi_{a^*}A(p) = A(\Psi_a p)$. To "slide" a form field $f':P\rightarrow G'^*$ of type ad′ over to the cotangent space TP^* as a vertical form we must make use of the connection form $\omega:TP\rightarrow G'$. Define $f:P\rightarrow TP^*$ by $f_p(V_p) = f'_p[\omega(V_p)]$. To clarify the notation, ad′ can be thought of as an action of G on both G' and G'^* (the vector space of left invariant forms), $\mathrm{ad}'_a:G'\rightarrow G', \mathrm{ad}'_a:G'^*\rightarrow G'^*$, defined by

$$\mathrm{ad}'_a(L) = K, \quad \text{where} \quad K(e) = \mathrm{ad}_{a^*}L(e),$$

$$\mathrm{ad}'_a(\Phi) = \Theta, \quad \text{where} \quad \Theta(e) = \mathrm{ad}^*_{a^{-1}}\Phi(e).$$

It then follows that f, the slide over of f', is invariant under the action of the group on P, i.e., $f(\Psi_a p) = \Psi^*_{a^{-1}}f(p)$. Given any tensor of type ad′, $t':P\rightarrow G'\otimes\cdots G'^*\otimes\cdots$, it is now straightforward to build a vertical tensor field on P which is invariant under the action of the group. It is somewhat more instructive to make the general construction by using a basis on P constructed from a basis of G' and TM. Let L_α

 0022-2488/82/010128-04$02.50

$[\alpha = 1,\cdots,N]$ be a basis of G' and Φ^α its dual basis of left invariant forms, and let l_α be the vertical Killing basis on P associated with L_α, i.e., $l_\alpha = \Sigma(L_\alpha)$. As a basis of the tangent space TP we take (h_a, l_α) and its dual basis becomes (θ^a, ϕ^α), where θ^a is the lift of dx^a and ϕ^α is the slide of Φ^α. If $T'(p) = T^{\alpha\cdots}{}_{\beta\ldots}(p) L_\alpha \otimes \cdots \otimes \Phi^\beta \otimes \cdots$ is a tensor field of type ad$'$ on P, then associated with it is a vertical tensor T on P given by

$$T(p) = T^{\alpha\cdots}{}_{\beta\ldots}(p) l_\alpha(p) \otimes \cdots \otimes \phi^\beta(p) \otimes \cdots$$

Because T' was of type ad$'$, T will be invariant under the action of the group on the bundle.

Finally, tensors of mixed type can be defined on P. For example the gauge covariant derivative DT' of a tensor field T' of type ad$'$, is a 1-form (horizontal at that) on P with values in the same tensor product space of G' as T' is. This form can now be "slid" back to P as a mixed tensor with one horizontal form index and the rest vertical indices. As an example if $T': P \to G'$, i.e., $T'(p) = T^\alpha(p) L_\alpha$, then $DT' = h_a(T^\alpha(p)) \theta^a \otimes L_\alpha$. This tensor is "slid" back to P as the mixed invariant tensor $DT = h_a(T^\alpha(p)) \theta^a(p) \otimes l_\alpha(p)$.

III. A METRIC FOR P AND COMPATIBILITY OF COVARIANT DERIVATIVES

If we wish to derive equations of motion and conservation laws from a Lagrangian we must be able to construct scalar fields from fields of type ad$'$. The simplest way to do this is to assume the theory contains a nondegenerate two-index-symmetric-tensor field $g'_G = g_{\alpha\beta}(P) \Phi^\alpha \otimes \Phi^\beta$ of type ad$'$, i.e., a metric field of type ad$'$. When this field is "slid" back to the bundle, it becomes a metric on the vertical part of TP,

$$g_G(p) = g_{\alpha\beta}(p) \phi^\alpha(p) \otimes \phi^\beta(p). \tag{3.1}$$

At this point it is possible to define a metric on the bundle space P by

$$g_P = g_G + g_{\rm st}, \tag{3.2}$$

where $g_{\rm st}$ is the horizontal lift of the spacetime metric. This metric is invariant under the action of the group and is identical with the one used by Cho and Freund, and, at this point, more general than the one proposed by Trautman.[1-3] We intend to use the Ricci scalar of the bundle metric as a Lagrangian as many authors suggest.[1-10] At this point it would be appropriate to call this theory a generalized Jordan theory because $g_G(p)$ will not be a constant on any cross section.[6-8]

However, at this point we are in possession of three types of covariant differentiation, one from the gauge connection, one from the spacetime metric, and the other from the bundle metric. We call them D, Δ, and ∇ respectively. Since we have defined a straightforward procedure for relating physical fields (tensors on spacetime and fields on type ad$'$) to fields on P, it is tempting to ask that appropriately comparable covariant derivatives agree. By this admittedly ad hoc restriction, we mean that if $t_{\rm st}$ and $T'_{\rm ad}$ are two such fields and we wish to compute their change in some spacetime direction v_x, then we would compute $\Delta_{v_x}(t_{\rm st})$ and $D_{h_p}(T'_{\rm ad})$, where h_p is the horizontal lift of v_x (the latter derivative usually being evaluated only on some cross section). Because these fields are also defined on P we could compute $\nabla_{h_p}(T_{\rm st})$ and $\nabla_{h_p}(T_{\rm ad})$. These two sets of derivatives can be compared by (1) lifting the tensor $\Delta(t_{\rm st})$ from M up the fiber to p and comparing it to the horizontal part of $\nabla(T_{\rm st})$ and by (2) sliding the G' valued form $D(T'_{\rm ad})$ of type ad$'$ back to the vertical tangent space at p by the procedure defined in the last section and comparing it to the vertical parts of $\nabla(T_{\rm ad})$. We find that a necessary and sufficient condition for these two pairs of covariant derivatives to agree is that the gauge covariant derivative of the group metric field vanish, i.e., $Dg'_G = 0$. This restriction on the metric reduces it from being a set of spacetime scalars to an invariant element of the theory.[5]

Necessity is relatively easy to show once it is observed that the vertical part of $\nabla_{h_p}(g_G)$ vanishes. This directly implies that $Dg'_G = 0$ for consistency. Proof of sufficiency requires more details; we simply write out the connection symbols for ∇ in the basis $X_A = (h_a, l_\alpha)$, where $\nabla_{X_A} X_B = \Gamma^C_{BA} X_C$.

$$\begin{aligned}
\Gamma^\alpha_{\beta\gamma} &= -\tfrac{1}{2}[2g^{\alpha\sigma}_G C^\mu_{\sigma(\beta} g_{G\gamma)\mu} + C^\alpha_{\beta\gamma}],\\
&\Gamma^\alpha_{ab} = \tfrac{1}{2}F^\alpha_{ab}, \qquad \Gamma^c_{ab}(p) = \begin{Bmatrix} c \\ ab \end{Bmatrix}(\pi(p)),\\
&\Gamma^b_{a\alpha}(p) = \Gamma^b_{\alpha a}(p) = \tfrac{1}{2}F^\sigma_{ac} g^{cb}_{\rm st}(\pi(p)) g_{G\sigma\alpha},\\
\Gamma^\alpha_{a\beta} &= \Gamma^\alpha_{\beta a} = \tfrac{1}{2}g^{\alpha\sigma}_G h_a(g_{G\beta\sigma}),\\
\Gamma^b_{\gamma\alpha(p)} &= -\tfrac{1}{2}g^{bc}_{\rm st}(\pi(p)) h_c(g_{G\alpha\gamma}),
\end{aligned} \tag{3.3}$$

where the various functions on P are defined by

$$\begin{aligned}
&[l_\alpha, l_\beta] = C^\gamma_{\alpha\beta} l_\gamma,\\
&[h_a, h_b] = -F_{ab}(p) = -F^\alpha_{ab}(p) l_\alpha(p),\\
&g_P(p) = g_{{\rm st}ab}(\pi(p)) \theta^a(p) \otimes \theta^b(p)\\
&\qquad + g_{G\alpha\beta}(p) \phi^\alpha(p) \otimes \phi^\beta(p),\\
&\pi_* h_a = \frac{\partial}{\partial x^a}, \qquad \Delta_{\partial/\partial x^a}\left(\frac{\partial}{\partial x^b}\right) = \begin{Bmatrix} c \\ ba \end{Bmatrix} \frac{\partial}{\partial x^c}.
\end{aligned} \tag{3.4}$$

It is straightforward to see that the horizontal part of $\nabla(t^{a\cdots}_{{\rm st}\ b\ldots}(\pi p) h_a \otimes \cdots \otimes \theta^b \otimes \cdots)$ is identical with the horizontal lift of $\Delta(t^{a\cdots}_{{\rm st}\ b\ldots}(x) \partial/\partial x^a \otimes \cdots \otimes dx^b \otimes \cdots)$ without any further restriction on the bundle metric. However, when we compare the vertical parts of $\nabla_{h_a}(T'^{\alpha\cdots}_{\rm ad}(p)_{\beta\ldots} l_\alpha \otimes \cdots \otimes \phi^\beta \otimes \cdots)$ with those of the "slide" over of $D_{h_a}(T'^{\alpha\cdots}_{\rm ad}{}_{\beta\ldots}(p)) L_\alpha \otimes \cdots \otimes \Phi^\beta \otimes \cdots$ we see that $\nabla_{h_a}(l_\alpha)$ and $\nabla_{h_a}(\phi^\beta)$ can have no vertical parts, i.e., $h_a(g_{G\alpha\beta}) = 0$ is not only necessary but also sufficient. Since this is the component version of $Dg'_G = 0$ we conclude that a necessary and sufficient condition for compatability of all three covariant derivatives, spacetime, gauge, and bundle metric, is that the group metric field be gauge covariantly constant.

IV. SPECIAL GAUGES, $g_{G\alpha\beta} = $ CONST

In this section we show that $Dg'_G = 0$ is a necessary and sufficient condition for finding a gauge (a local cross section of P) where the components of g_G are constants when restricted to that cross section. Because of the absence of the spacetime scalars it is more appropriate to call this theory the generalized Kaluza–Klein theory, rather than the generalized Jordan or Brans–Dicke theory.[6,9,10,11] We prove the conjecture in two steps. First, the necessary and sufficient conditions for $h_a g_{G^{\alpha\beta}} = 0$ to have solutions is that the integrability conditions $[h_a, h_b] g_{\alpha\beta} = 0$ be satisfied or equivalently,

$$2F^{\sigma}_{ab}(p) g_{G^{\pi\alpha}}(p) C^{\pi}_{\beta)\sigma} = 0. \tag{4.1}$$

Secondly, we show that Eq. (4.1) is the necessary and sufficient condition for finding a gauge transformation, $b: U \to G$ (U is some open neighborhood of $x \in M$) which would produce a local coordinate chart on P, $[p = (x,a)]$, in which $(\partial/\partial x^c) g_{G^{\alpha\beta}}(x,a) = 0$ These coordinates are defined by $x = \pi(p)$ and $p = \Psi_a(s(x))$ where s is some local cross section, $s: U \to P$, such that $\pi \circ s = \mathrm{id}_U$. If we define the matrix $(\mathrm{ad}'_a)^{\beta}_{\alpha}$ by

$$\mathrm{ad}'_a(L_\alpha) = L_\beta (\mathrm{ad}'_a)^{\beta}_{\alpha} \tag{4.2}$$

it follows that in the (x,a) coordinate system

$$\begin{aligned} &l_\alpha(p) = L_\alpha(a), \\ &h_c(p) = \frac{\partial}{\partial x^c} - (\mathrm{ad}'_{a^{-1}})^{\beta}_{\alpha} A^{\alpha}_{c}(x) L_\beta(a), \\ &\theta^c(p) = dx^c, \\ &\phi^\alpha(p) = \Phi^\alpha(a) + (\mathrm{ad}'_{a^{-1}})^{\alpha}_{\gamma} A^{\gamma}_{b}(x) dx^b, \\ &g_{G^{\alpha\beta}}(p) = g_{\gamma\delta}(x) (\mathrm{ad}'_a)^{\gamma}_{\alpha} (\mathrm{ad}'_a)^{\delta}_{\beta}, \end{aligned} \tag{4.3}$$

and

$$F_{cb}(p) = (\mathrm{ad}'_{a^{-1}})^{\beta}_{\alpha} F^{\alpha}_{cb}(x) L_\beta(a),$$

where $F^{\alpha}_{cb}(x) = (\partial/\partial x^c) A^{\alpha}_{b}(x) - (\partial/\partial x^b) A^{\alpha}_{c}(x) + A^{\gamma}_{c}(x) A^{\delta}_{b}(x) C^{\alpha}_{\gamma\delta}$, and $g_{\alpha\beta}(x) = g_{G^{\alpha\beta}}(p = s(x))$. These can be compared with equations given by Cho and Freund.[3] The vanishing gauge covariant derivative, $Dg'_G = 0$, becomes

$$\frac{\partial}{\partial x^c} g_{\alpha\beta}(x) + 2A^{\sigma}_{c}(x) g_{\pi(\alpha}(x) C^{\pi}_{\beta)\sigma} = 0, \tag{4.4}$$

and its integrability condition Eq. (4.1) becomes

$$2F^{\sigma}_{cb}(x) g_{\pi(\alpha}(x) C^{\pi}_{\beta)\sigma} = 0. \tag{4.5}$$

A gauge transformation can be looked at as a simple coordinate transformation

$$\begin{aligned} &\underset{\sim}{x} = x, \\ &\underset{\sim}{a} = b(x) \cdot a \end{aligned} \tag{4.6}$$

given by a mapping $b: U \to G$. Under such a transformation $g_{\alpha\beta}(x)$ changes by

$$g_{\alpha\beta}(x) = g_{\gamma\delta}(x) (\mathrm{ad}'_{b^{-1}}(x))^{\gamma}_{\alpha} (\mathrm{ad}'_{b^{-1}}(x))^{\delta}_{\beta}. \tag{4.7}$$

By putting $g_{\alpha\beta}(x) =$ const, Eq. (4.7) becomes an algebraic set of equations for $b(x)$ which can be turned into a set of differential equations by taking its exterior derivative.

$$b^* \phi^\delta g_{\gamma(\alpha}(x) C^{\gamma}_{\beta)\delta} = -\tfrac{1}{2} d g_{\alpha\beta}(x), \tag{4.8}$$

where $b^* \phi^\delta = dx^c (\partial b^\pi / \partial x^c) \phi^\delta (\partial/\partial b^\pi)$ in some coordinate system on G. This equation obviously cannot be solved for an arbitrary $g_{\alpha\beta}(x)$ at any point x, however, from Eq. (4.4) we can rewrite Eq. (4.8) as

$$(b^* \phi^\delta - A^\delta) g_{\gamma(\alpha}(x) C^{\gamma}_{\beta)\delta} = 0, \tag{4.9}$$

where $A^\delta(x) \equiv A^{\delta}_{c}(x) dx^c$ and where $g_{\alpha\beta}(x)$ and $A^\delta(x)$ are restricted by Eq. (4.4). Equation (4.9) can be solved at any point $x = x_0$ and its exterior derivative vanishes identically provided Eq. (4.4) is satisfied. This can be easily checked by using $d\phi^\delta = -\tfrac{1}{2} C^{\delta}_{\rho\omega} \phi^\rho \wedge \phi^\omega$.

In one of these special gauges we have

$$g_{\alpha\beta}(x) = \text{const}, \qquad A^{\delta}_{c}(x) g_{\gamma(\alpha} C^{\gamma}_{\beta)\delta} = 0 \tag{4.10}$$

$$\text{as well as} \qquad F^{\delta}_{cd}(x) g_{\gamma(\alpha} C^{\gamma}_{\beta)\delta} = 0.$$

The remaining gauge freedom is restricted by

$$b^* \phi^\delta g_{\gamma(\alpha} C^{\gamma}_{\beta)\delta} = 0. \tag{4.11}$$

If the metric $g_{G^{\alpha\beta}}(p)$ is invariant then $g_{G^{\gamma(\alpha}} C^{\gamma}_{\beta)\delta} = 0$ and no algebraic restrictions are placed on the vector potentials by Eq. (4.10). Since g'_G is a field of type ad$'$ it follows that $l_\gamma(g_{G^{\alpha\beta}}) = -2g_{G^{\pi(\alpha}}(p) C^{\delta}_{\beta)\gamma}$, from which we conclude that an ad$'$ invariant metric must be constant up the fiber as well as along the cross section. This case contains the semisimple groups and their "Killing" metrics as Trautman proposed and Cho elaborated.[1,2] It, however, contains many more groups which possess invariant metrics but whose Killing forms are degenerate. One interesting solvable four dimensional case is given in Sec. V. See Patera *et al.* for invariants of low dimensional Lie algebras.[12] The other and most frequent case is that for which $g_{\alpha\beta}$ is not invariant and consequently Eqs. (4.10) and (4.11) place restrictions on the combined set of gauge field, vector potentials, group metric, and gauge transformations. Examples of this case include the nonunimodular groups ($C^{\alpha}_{\alpha\beta} \neq 0$) for which no invariant metrics exists. As a consequence of Eq. (4.10), however, for all cases the metric $g'_G(p)$ must be invariant under the holonomy group of the connection.[13,14]

V. GEODESICS AND LAGRANGIANS FOR THE GENERALIZED KALUZA–KLEIN THEORIES

We start by giving the geodesic equations for the bundle metric defined in Sec. III. It is found that up to a global gauge transformation (action of G on P by Ψ), the geodesics are uniquely determined by a spacetime particle trajectory and by the "gauge charge" of the particle. The gauge charge being a form of type ad$'$, $q' = q_\alpha \Phi^\alpha$, whose components q_α are constant along each geodesic in P. If we use an affine parameter λ for the geodesic through point $p \in P$ and write the tangent vector $V(\lambda) = v^a h_a + q^\alpha l_\alpha$, we have the following geodesic equations,

$$\frac{d}{d\lambda} v^a + v^b v^c \begin{Bmatrix} a \\ bc \end{Bmatrix} + v^b q_\alpha F^{\alpha}_{bc}\, g^{ca}_{\mathrm{st}} = 0$$

and

$$\frac{d}{d\lambda} q^\alpha - g^{\alpha\sigma}_{G}\, C^{\mu}_{\sigma\beta} q^\beta q_\mu = 0, \tag{5.1}$$

where $q_\alpha \equiv g_{G^{\alpha\beta}} q^\beta$. The second equation implies $dq_\alpha / d\lambda = 0$ producing the conserved gauge charge, and the first tells how it must transform under the action of Ψ_a in order that the geodesic projects onto a fixed particle trajectory, i.e., $q_\alpha F^{\alpha}_{bc}$ has to be constant up the fiber.

From Eq. (5.1) it follows that $q^\alpha q_\alpha$ (and hence $v^a v_a$ remains constant along a geodesic. We conclude that the spacetime distance along the curve projected by π can be used for λ in Eq. (5.1). The conclusion is that the projected geodesic gives a spacetime curve $x^a(s)$ with tangent vector $v(s) = v^a(s)\partial/\partial x^a$ satisfying Eq. (5.1) which is the equation for a classical point particle with gauge charge q_α (at least from the E&M analogy).[10]

For a Lagrangian we use the Ricci scalar of the bundle metric,

$$\mathscr{L} = (-g_p)^{1/2} R_{4+N}, \tag{5.2}$$

where $(-g_p)^{1/2} = (-g_{st} g_G)^{1/2}$,

and

$$R_{4+N} = R_{st} - \tfrac{1}{4} g_{G\alpha\beta} F^\alpha_{ab} F^\beta_{cd}\, g^{ac}_{st} g^{bd}_{st} - g^{\alpha\beta}_G C^{\,\gamma}_{\alpha\gamma} C^{\,\delta}_{\beta\delta} - \tfrac{1}{2} g^{\alpha\beta}_G C^{\,\gamma}_{\alpha\delta} C^{\,\delta}_{\beta\gamma} - \tfrac{1}{4} g^{\alpha\beta}_G g^{\gamma\delta}_G C^{\,\sigma}_{\alpha\gamma} C^{\,\rho}_{\beta\delta} g_{G\sigma\rho}.$$

This can be compared with Cho and Freund by leaving out their cosmological constant term as well as the kinetic energy of the group metric, and by adding $-g^{\alpha\beta}_G C^{\,\gamma}_{\alpha\gamma} C^{\,\delta}_{\beta\delta}$ which is present for nonunimodular group.

The volumns V_p over which R_{4+N} is to be stabilized are given by the gauges permitted in Sec. IV and submanifolds H of G of finite volume as measured by $(g'_G)^{1/2} \Phi^1(a) \wedge \cdots \wedge \Phi^N(a)$.

$$I = \int_{V_p} (-g_p)^{1/2} R_{4+N}\, \theta^0 \wedge \cdots \wedge \theta^3 \wedge \phi^1 \wedge \cdots \wedge \phi^N, \tag{5.3}$$

where $V_p \equiv \{\Psi_a(x){:}x\in U,\ a\in H\}$. In coordinates adapted to a gauge this becomes

$$I = I_{st} \cdot I_G,$$

where

$$I_{st} = \int_U (-g_{st}(x)\, g(x))^{1/2} R_{4+N}(x) dx^0 \wedge \cdots \wedge dx^3,$$

$$I_G = \int_H \det(\mathrm{ad}'_a) \Phi^1(a) \wedge \cdots \wedge \Phi^N(a). \tag{5.4}$$

In the above $g_{st}(x) = \det(g_{st}(x)_{ab})$, $g(x) = \det(g_G(p)_{\alpha\beta})$ at $p = s(x)$, and $\det(\mathrm{ad}'_a) = \det[(\mathrm{ad}'_a)^\alpha_\beta]$. Up to a constant factor which depends on H, I is given by

$$I = \int_U (-g_{st}(x))^{1/2} R_{4+N}(x) dx^0 \wedge \cdots \wedge dx^3, \tag{5.5}$$

when evaluated in one of the special gauges $[g_{\alpha\beta}(x) = \text{const}]$.

As can be seen from Eq. (5.2) most examples of this theory will have an undesirable cosmological constant term, however, some will not. We conclude by giving the smallest dimensional nonabelian group with an invariant metric without a cosmological term. It is the four-dimensional group whose solvable algebra is given by

$$[L_4, L_\alpha] = 0, [L_1, L_2] = L_4,$$
$$[L_2, L_3] = L_1, \qquad [L_3, L_1] = L_2, \tag{5.6}$$

and classified as U3I2 by MacCallum.[15] The invariant metric is the Lorentz metric,

$$g_{\alpha\beta}(x) = g\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

and R_{4+N} simplifies to

$$R_{4+N} = R_{st}(x) - \tfrac{1}{4} g_{\alpha\beta} F^\alpha_{ab}(x) F^\beta_{cd}(x) g^{ac}_{st}(x) g^{bd}_{st}(x).$$

A somewhat different approach is used by Kopczyński to remove the cosmological term.[16]

[1] A. Trautman, The Application of Fibre Bundles in Physics, Notes on Lectures given at King's College, London, September (1967); A. Trautman, Rep. Math. Phys. **1**, 29 (1970); R. Kerner, Ann. Inst. Henri. Poincaré **9**, 143 (1968).
[2] Y. M. Cho, J. Math. Phys. **16**, 2029 (1975); L. N. Chang, K. I. Macrae, and F. Mansouri, Phys. Rev. D **13**, 235 (1976).
[3] Y. Mo. Cho and P. G. O. Freund, Phys. Rev. D **16**, 1711 (1975).
[4] B. DeWitt, in *Relativity, Groups and Topology*, edited by B. DeWitt and C. DeWitt (Gordon and Breach, New York, 1964), p. 725.
[5] A Trautman, *Elementary Introduction to Fiber Bundles and Gauge Fields* (Warsaw, Poland, December 1978); A. Trautman, Czech. J. Phys. B **29**, 107 (1979).
[6] P. Jordan, Z. Phys. **157**, 112 (1959).
[7] P. G. Bergmann, Ann. Math. **49**, 255 (1948).
[8] D. R. Brill, Int. School of Phys. Enrico Fermi, Varenna (1961).
[9] T. Kaluza, Situngsher. Preuss. Akad. Wiss. Phys. Math. Kl. **2**, 966 (1921).
[10] O. Klein, Z. Phys. **37**, 895 (1929).
[11] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961).
[12] J. Patera, R. T. Sharp, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **17**, 986 (1976).
[13] H. G. Loos, J. Math. Phys. **8**, 2114 (1967).
[14] M. Daniel and G. M. Viallet, Rev. Mod. Phys. **52**, 175 (1980).
[15] M. A. H. MacCallum, "On the Classification of the Real Four-Dimensional Lie Algebras," preprint (1980).
[16] W. Kopczyński, "A Fiber Bundle Description of Coupled Gravitational and Gauge Fields," in *Proceedings of the Conference on Differential Geometrical Methods in Mathematical Physics, Salamanca* 1979, Lecture Notes in Mathematics (Springer-Verlag, Berlin, 1980).

Stochastic quantization of the linearized gravitational field

Mark Davidson
Department of Physics, San Jose State University, San Jose, California 95192

(Received 25 August 1980; accepted for publication 31 October 1980)

Stochastic field equations for linearized gravity are presented. The theory is compared with the usual quantum field theory and questions of Lorentz covariance are discussed. The classical radiation approximation is also presented.

PACS numbers: 04.60. + n, 11.10.Np

I. INTRODUCTION

Quantum systems which are described by a Schrödinger equation allow a stochastic interpretation. The Fenyes–Nelson model[1–3] and its generalization[4–6] provide such a stochastic description of quantum mechanics. Several field theories have been considered in this context,[7–11] and the results have so far been of some interest. In the present paper, the generalized Fenyes–Nelson model is applied to the weak field approximation of Einstein's general theory of relativity, the so-called linearized gravitational field.[12–14]

It is worthwhile to attempt a stochastic model of quantum gravity for several reasons. First, since the gravitational field in the classical theory has the interpretation of a metric tensor, it is disturbing that in the usual quantum formulation this geometric interpretation is completely obliterated because the field and its derivatives become abstract operators on a rigged Hilbert space. Because of this difficulty it has become fashionable to play down the geometric role of gravity when dealing with quantization, as in the formulation of Weinberg.[14] A stochastic formulation of the quantized gravitational field may revitalize the geometrical interpretation of quantum gravity.

Second, the stochastic formulation of quantum mechanics, at least in spirit, is a progeny of Einstein's profound discomfiture with the complementarity vision of Bohr. Since Einstein's views were based at least in part on considerations of the general theory of relativity, it seems fitting to pursue a stochastic model of quantum gravity.

Third, the great successes which gauge theories have had in the theory of elementary particles suggest that fundamental efforts such as the stochastic reformulation of quantum mechanics be concentrated in this general area.

The linear theory of gravity has been selected for analysis because it avoids the extremely difficult problem of divergences in the full theory, and it is sufficiently simple to allow, perhaps, the beginning of a probabalistic geometric interpretation of quantum gravity.

In Sec. II the linearized Einstein field equations are briefly recounted. In Sec. III the usual quantization procedure is outlined. In IV the method of stochastic quantization is applied to the linear theory, and in V the curious random classical radiation approximation is presented.

II. THE LINEAR FIELD EQUATIONS

The Einstein field equations in the linear approximation may be written

$$-\bar{h}_{\mu\nu,\alpha}{}^{\alpha} - n_{\mu\nu}\bar{h}_{\alpha\beta},{}^{\alpha\beta} + \bar{h}_{\mu\alpha},{}^{\alpha}{}_{\nu} + \bar{h}_{\nu\alpha},{}^{\alpha}{}_{\mu} = 16\pi G T_{\mu\nu}, \tag{1}$$

where the notation of Ref. 13 is followed except that the gravitational constant G is not set to unity. Indices are raised and lowered via the Minkowski metric $n_{\mu\nu}$. The full metric tensor is related to the above fields by

$$g_{\mu\nu} = n_{\mu\nu} + h_{\mu\nu}, \tag{2}$$

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \tfrac{1}{2}n_{\mu\nu}\bar{h}^{\alpha}{}_{\alpha}, \tag{3}$$

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \tfrac{1}{2}n_{\mu\nu}h^{\alpha}{}_{\alpha}, \tag{4}$$

where $T_{\mu\nu}$ is the energy–momentum tensor of all nongravitational sources.

Once the linear theory is solved, classically, the nonlinear corrections of the full theory may be included as a perturbation by introducing a suitable term in the energy–momentum tensor (Ref. 14, p. 165). This procedure has not worked so far in the quantum theory because of the celebrated divergences introduced by the nonlinear terms.

Weinberg has suggested that general relativity be treated as an ordinary field theory with the linear theory as the starting point. This will be the approach taken here, although one of the primary motives for seeking a stochastic model of gravitation is to restore a geometric interpretation to the quantum theory.

The linear theory possesses a gauge invariance similar to electromagnetism (Refs. 13, p. 439; 14, p. 254). The field equations are left invariant under the transformation

$$h_{\mu\nu} \rightarrow h_{\mu\nu} - \partial\mathscr{E}_{\mu}/\partial x^{\nu} - \partial\mathscr{E}_{\nu}/\partial x^{\mu}, \tag{5}$$

where $\mathscr{E}^{\mu}$ is an arbitrary 4-vector field. All observables are independent of the gauge, and so gauge conditions may be imposed by fiat in order to facilitate solution. We shall work in the Lorentz gauge where one requires

$$\bar{h}_{\mu}{}^{\alpha}{}_{,\alpha} = 0. \tag{6}$$

The field equations become in this gauge

$$-\partial_{\alpha}\partial^{\alpha}\bar{h}_{\mu\nu} = 16\pi G T_{\mu\nu}. \tag{7}$$

The solutions to Eq. (7) may be written as a sum of a retarded solution plus a free field solution,

$$\bar{h}_{\mu\nu} = \bar{h}^{\mathrm{R}}_{\mu\nu} + \bar{h}^{\mathrm{in}}_{\mu\nu}, \tag{8}$$

where

$$\bar{h}^{\mathrm{R}}_{\mu\nu} = \int 4GT_{\mu\nu}(\mathbf{x}',t - |\mathbf{x} - \mathbf{x}'|)/|\mathbf{x} - x'|\, d^3x' \tag{9}$$

and

$$\partial_{\alpha}\partial^{\alpha}\bar{h}^{\mathrm{in}}_{\mu\nu} = 0. \tag{10}$$

The gauge condition [Eq. (6)] does not determine the gauge completely. Any additional transformation such that

$$\partial_{\alpha}\partial^{\alpha}\mathscr{E}_{\mu} = 0, \tag{11}$$

 0022-2488/82/010132-06$02.50

will preserve the Lorentz condition. This gauge transformation of the second kind can be used to make the free part of the field equation both transverse and traceless (Ref. 13, p. 946). No generality is lost, therefore, by requiring

$$\bar{h}^{\text{in}}_{\text{o}\nu} = 0, \tag{12}$$

$$\bar{h}^{\text{in}}_{ii} = 0. \tag{13}$$

Periodic boundary conditions in the spatial coordinates will be imposed as an aid to quantization. If the length associated with this periodicity is L, then wave numbers in a Fourier decompositon are restricted to have components which are integral multiples of $2\pi/L$. Although the energy momentum tensor must also have this periodicity for consistency, no generality is really lost since L will be taken to infinity eventually. With these conditions, the Fourier decomposition of the free part of the gravitational field is

$$h^{\text{in}}_{ij} = (1/\sqrt{2}L^{3/2})\sum_{\lambda,\mathbf{k}} \epsilon_{ij}(\lambda,\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}Q(\lambda,\mathbf{k},t), \tag{14}$$

where we have dropped the distinction between $\bar{h}^{\text{in}}$ and h^{in}, since in the transverse traceless gauge they are the same. λ denotes the polarization state and it can take on one of two values. The ϵ's must satisfy the conditions

$$\epsilon_{ij} = \epsilon_{ji}, \quad k_i\epsilon_{ij} = 0, \quad \epsilon_{ij}\epsilon_{ij} = 1, \quad \epsilon_{ii} = 0,$$
$$\epsilon_{ij}(\lambda)\epsilon_{ij}(\lambda') = \delta_{\lambda,\lambda'}, \quad \epsilon_{ij}(\lambda, -\mathbf{k}) = \epsilon^*_{ij}(\lambda,\mathbf{k}). \tag{15}$$

The polarization tensors can be chosen real, and, with $\mathbf{k}$ in the z direction, they can be taken as

$$\epsilon_{ij}(+) = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}, \tag{16}$$

$$\epsilon_{ij}(\times) = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, \tag{17}$$

for the two independent states of polarization denoted by $+$ and $\times$. The following expression is often useful:

$$\sum_{\lambda} \epsilon_{ij}(\lambda)\epsilon_{kl}(\lambda) = \tfrac{1}{2}(-\delta^{\text{tr}}_{ij}\delta^{\text{tr}}_{kl} + \delta^{\text{tr}}_{ik}\delta^{\text{tr}}_{jl} + \delta^{\text{tr}}_{il}\delta^{\text{tr}}_{jk}), \tag{18}$$

where

$$\delta^{\text{tr}}_{ij} = \delta_{ij} - k_ik_j/\mathbf{k}^2. \tag{19}$$

The equations of motion for the field require that the Q's satisfy

$$\ddot{Q} + \mathbf{k}^2Q = 0, \tag{20}$$

so that a Lagrangian may be chosen of the form

$$L = f\sum_{\lambda,\mathbf{k}}(\tfrac{1}{4}|\dot{Q}|^2 - \tfrac{1}{4}\mathbf{k}^2|Q|^2), \tag{21}$$

where f is an as yet undertermined parameter. In the sum over k each Q appears twice because reality requires:

$$Q(\lambda, -\mathbf{k}) = Q^*(\lambda,\mathbf{k}). \tag{22}$$

The conjugate momenta are

$$\frac{\partial L}{\partial\,\text{Re}\dot{Q}} = f\,\text{Re}\dot{Q} = \text{Re}P, \qquad \frac{\partial L}{\partial\,\text{Im}\dot{Q}} = f\,\text{Im}\dot{Q} = \text{Im}P, \tag{23}$$

and the Hamiltonian may be written as

$$H = \sum_{\lambda,\mathbf{k}}(|P|^2/4f + \tfrac{1}{4}\mathbf{k}^2f|Q|^2)$$
$$= f\sum_{\lambda,\mathbf{k}}(\tfrac{1}{4}|\dot{Q}|^2 + \tfrac{1}{4}\mathbf{k}^2|Q|^2). \tag{24}$$

This must now be compared to the actual energy of the gravitational field to set the scale for quantization.

The physical energy to be associated with a free gravitational field has been calculated (Ref. 13, p. 955). It is given by the 00 component of the stress energy tensor:

$$T_{\mu\nu} = \frac{1}{32\pi G}\int_\Omega(\partial_\mu\bar{h}^{\text{in}}_{ij})(\partial_\nu\bar{h}^{\text{in}}_{ij})d^3x, \tag{25}$$

where the integration is taken over a three-dimensional cube (Ω) of edge length L. One finds

$$T_{00} = (1/32\pi G)\int_\Omega(\dot{\bar{h}}^{\text{in}}_{ij})^2d^3x$$
$$= (1/32\pi G)\sum_{\lambda,\mathbf{k}}(\tfrac{1}{4}|\dot{Q}|^2 + \tfrac{1}{4}\mathbf{k}^2|Q|^2), \tag{26}$$

so that we may make the identification

$$f = 1/32\pi G, \tag{27}$$

and therefore,

$$P = (1/32\pi G)\dot{Q}. \tag{28}$$

III. QUANTIZATION

Quantization is achieved by imposing commutation rules as follows at equal times:

$$(1/32\pi G)[\text{Re}\dot{Q}(\lambda,\mathbf{k}),\text{Re}Q(\lambda',\mathbf{k}')]$$
$$= -i\hbar\delta_{\lambda\lambda'}(\delta_{\mathbf{k},\mathbf{k}'} + \delta_{\mathbf{k},-\mathbf{k}'}), \tag{29}$$

$$(1/32\pi G)[\text{Im}\dot{Q}(\lambda,\mathbf{k}),\text{Im}Q(\lambda',\mathbf{k}')]$$
$$= -i\hbar\delta_{\lambda\lambda'}(\delta_{\mathbf{k},\mathbf{k}'} - \delta_{\mathbf{k},-\mathbf{k}'}). \tag{30}$$

In terms of the fields, the equal time commutation rules become

$$[\dot{\bar{h}}_{ij}(\mathbf{x}),\bar{h}_{kl}(y)] = -i\hbar\delta_{ijkl}(\mathbf{x}-\mathbf{y}), \tag{31}$$

where

$$\delta_{ijkl}(\mathbf{x}-\mathbf{y}) = \int\frac{d^3k}{(2\pi)^3}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}\delta_{ijkl}(\mathbf{k}), \tag{32}$$

and where

$$\delta_{ijkl}(\mathbf{k}) = \sum_{\lambda}\epsilon_{ij}(\lambda,\mathbf{k})\epsilon_{kl}(\lambda,\mathbf{k}) \tag{33}$$

is given by Eq. (18). The dynamical relation which fixes the quantum theory is

$$\ddot{Q} = -\mathbf{k}^2Q. \tag{34}$$

The field theory is equivalent to uncoupled harmonic oscillators. In the ground state one finds

$$\langle 0|\bar{h}^{\text{in}}_{ij}(x)\bar{h}^{\text{in}}_{kl}(y)|0\rangle$$
$$= (32\pi G)i\hbar\int\frac{d^4k}{(2\pi)^4}\frac{\delta_{ijkl}(\mathbf{k})e^{ik^\mu(x_\mu - y_\mu)}}{-k^\mu k_\mu + i\epsilon}, \quad t_x > t_y. \tag{35}$$

Using the rule of Gaussian combinatorics,

$$\langle 0|T\phi_1(x_1)\times\cdots\times\phi_n(x_n)|0\rangle$$
$$= \sum_\pi \langle 0|T\phi_1(x_1)\phi_2(x_2)|0\rangle\times\cdots\times\langle 0|T\phi_{n-1}(x_{n-1})\phi_n(x_n)|0\rangle, \tag{36}$$

where n is even, T denotes time ordering, π denotes a sum over distinct permutations, and the higher order moments for the ground state of the quantum field may be calculated. The spectral energy density for the field is found to be the same as for electromagnetism:

$$\rho(\omega) = (\hbar/2\pi^2)\omega^3, \tag{37}$$

corresponding to an energy of $\frac{1}{2}\hbar\omega$ for each degree of freedom of the field in the ground state.

IV. STOCHASTIC QUANTIZATION

The linearized gravitational field as a quantum system may be reduced to an infinite dimensional Schrödinger equation of the form

$$\left[\sum_i\left(-\tfrac{1}{2}\hbar^2\frac{\partial^2}{\partial Q_i^{\,2}}\right) + V(Q)\right]\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{38}$$

The dimensionality of this equation can be made finite by imposing a cutoff in momentum space. Since much of the theory upon which stochastic quantization relies has been done for only finite dimensional systems, it is convenient to impose a momentum cutoff which may be taken to infinity in most expressions of interest. This is not an important limitation and it can be expected that as the theory of Markov fields advances it will eventually be eliminated. For the present, we shall assume a momentum cutoff. Let us choose the following notation

$$\psi = e^{R+iS}, \qquad \Delta_Q = \sum_i \frac{\partial^2}{\partial Q_i^{\,2}}, \tag{39}$$

where R and S are real functions.

Direct computation shows that the following equation is equivalent to Eq. (38) so long as R and S have a first time derivative and a second Q derivative, and so long as z has a nonzero real part

$$\left[-\frac{(z\hbar)^2}{2}\Delta_Q + \left(V(Q) - \frac{\hbar^2}{2}(z^2-1)\frac{\Delta_Q e^R}{e^R}\right)\right]e^{R+iS/z}$$
$$= i(z\hbar)\frac{\partial}{\partial t}e^{R+iS/z}, \tag{40}$$

where z is a constant which may be complex.

Suppose that z is purely imaginary, so that

$$z = i|z|. \tag{41}$$

Then Eq. (40) is still true if (38) is true, but (40) is only one real equation

$$\left\{\frac{(|z|\hbar)^2}{2}\Delta_Q + \left(V + \frac{\hbar^2}{2}(1+|z|^2)\frac{\Delta_Q e^R}{e^R}\right)\right\}e^{R+S/|z|}$$
$$= -|z|\hbar\frac{\partial}{\partial t}e^{R+S/|z|}. \tag{42}$$

This equation must be supplemented by another real equation since the original Schrödinger equation contains two real equations. Another equation may be generated by choosing

$$z = -i|z|, \tag{43}$$

which leads to

$$\left[\frac{(|z|\hbar)^2}{2}\Delta_Q + \left(V + \frac{\hbar^2}{2}(1+|z|^2)\frac{\Delta_Q e^R}{e^R}\right)\right]e^{R-S/|z|}$$
$$= |z|\hbar\frac{\partial}{\partial t}e^{R-S/|z|}. \tag{44}$$

Equations (42) and (44) taken together are equivalent to Schrödinger's equation by direct computation.

Consider the two real Eqs. (42) and (44). They have the mathematical form of the heat equation. Equations of this generic form are characteristic of a certain type of diffusion problem called generalized Brownian motion or Ito processes. In Ref. 4 equations of this form were derived within the context of diffusion theory. The fact that Schrödinger's equation can be rewritten in the form (42) and (44) is the basis for the stochastic interpretation of quantum mechanics.

The stochastic interpretation rests on the hypothesis that all of the experimentally verifiable predictions of quantum mechanics may be deduced from Schrödinger's equation. This hypothesis shall be assumed true in this paper.

Stochastic quantization is achieved, following Nelson,[2,3] by associating with each coordinate Q a stochastic process which is defined by the stochastic differential equation:

$$dQ_i = b_i(Q,t) + dW_i(t), \tag{45}$$

where the W's are Wiener processes which satisfy

$$E(dW_i dW_j) = 2\nu\delta_{ij}dt. \tag{46}$$

Processes defined by Eq. (45) go by the name: generalized Brownian motion, Ito processes, or multidimensional diffusion processes. The mathematical background contained in Ref. 2 is sufficient for an understanding of stochastic quantization. Other good sources for this subject are Refs. 15–22. Since the formalism of forward and backward derivatives has been developed only by Nelson,[2] a careful reading of his book is essential to an understanding of stochastic quantization. A method relying on an operator formalism, rather than the forward and backward time derivatives was presented in,[4] but it is completely equivalent to the Nelson procedure.

Many existence and uniqueness theorems for the processes defined by Eq. (45) are presented in Refs. 2, 15–22. The most important theorems show that if b satisfies a global Lipschitz condition then Eq. (45) has a solution which is a continuous Markov process and which is unique (for example, Ref. 2, p. 43).

In Ref. 2 forward and backward time derivatives D and D_* are defined by

$$Df(Q,t) = \lim_{h\to 0_+}\frac{1}{h}E\big(f(Q(t+h),t+h)$$
$$-f(Q(t),t)\,|\,Q(t)=Q\big), \tag{47}$$

$$D_*f(Q,t) = \lim_{h\to 0_+}\frac{1}{h}E\big(f(Q(t),t)$$
$$-f(Q(t-h),t-h)\,|\,Q(t)=Q\big), \tag{48}$$

and independent of any dynamical assumption, these satisfy

$$DQ_i = b_i, \tag{49}$$

$$D_* Q_i = b_{i*}, \tag{50}$$

$$b_i - b_{i*} = 2\nu \frac{\partial}{\partial Q_i} \ln[\rho(Q,t)], \tag{51}$$

$$D = \frac{\partial}{\partial t} + \sum_i b_i \frac{\partial}{\partial Q_i} + \nu \Delta_Q, \tag{52}$$

$$D_* = \frac{\partial}{\partial t} + \sum_i b_{i*} \frac{\partial}{\partial Q_i} - \nu \Delta_Q. \tag{53}$$

The dynamical assumption which leads to Schrödinger's equation is

$$\tfrac{1}{2}(DD_* + D_*D)Q_i + \frac{\beta}{8}(D - D_*)^2 Q_i = -\frac{\partial}{\partial Q_i} V, \tag{54}$$

where β is a constant and where

$$\nu = \hbar/2\sqrt{(1-\beta/2)}. \tag{55}$$

The Schrödinger wave function is written in the form

$$\psi = e^{R + izS_N}, \tag{56}$$

where

$$z = 1/(1-\beta/2)^{1/2} \tag{57}$$

and

$$b_i = 2\nu \frac{\partial}{\partial Q_i}(R + S_N). \tag{58}$$

It is straightforward to show that Eq. (54) implies that (56) satisfies (38).

If ν is to be real, then we must demand the condition

$$\beta < 2. \tag{59}$$

It is possible, because of Eq. (55), to choose any value of the diffusion parameter ν for a stochastic model of quantum mechanics. In the limit when $\nu \to 0$, the model becomes deterministic and equivalent to Bohm's hidden variable theory,[23] as was first pointed out in.[6]

In applying this formalism to linear gravity, we use the Q's in Eq. (14), but mindful of the condition (22) which relates Q's for antiparallel wave vectors. If we define

$$b_{ij} = D\bar{h}^{\text{in}}_{ij}, \tag{60}$$

$$b_{ij*} = D_*\bar{h}^{\text{in}}_{ij}, \tag{61}$$

then with the aid of the operator

$$\Delta^{ij}(\mathbf{x},t) = \frac{1}{(\sqrt{2})L^{3/2}} \sum^{\lambda,\mathbf{k}} \epsilon^{ij}(\lambda,\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \times\left(\frac{\partial}{\partial\,\mathrm{Re}Q} + i\frac{\partial}{\partial\,\mathrm{Im}Q}\right), \tag{62}$$

the forward and backward time derivatives can be expressed as

$$D = \frac{\partial}{\partial t} + \int_\Omega b_{ij}\Delta^{ij} d^3x + \nu \int_\Omega \Delta^{ij}\Delta_{ij} d^3x, \tag{63}$$

$$D_* = \frac{\partial}{\partial t} + \int_\Omega b_{ij*}\Delta^{ij} d^3x - \nu \int_\Omega \Delta^{ij}\Delta_{ij} d^3x. \tag{64}$$

If we define a random field by

$$W^{ij}(\mathbf{x},t) = \frac{1}{(\sqrt{2})L^{3/2}} \sum_{\lambda,\mathbf{k}} \epsilon^{ij}(\lambda,\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} W_{\lambda,\mathbf{k}}(t), \tag{65}$$

where the real and imaginary parts of $W_{\lambda,\mathbf{k}}$ are Wiener processes which are independent of one another and which satisfy

$$E\,[d(\mathrm{Re}W_{\lambda,\mathbf{k}})\,d(\mathrm{Re}W_{\lambda',\mathbf{k}'})] = 2\nu\delta_{\lambda\lambda'}(\delta_{\mathbf{k},\mathbf{k}'} + \delta_{\mathbf{k},-\mathbf{k}'})dt, \tag{66}$$

$$E\,[d(\mathrm{Im}W_{\lambda,\mathbf{k}})d(\mathrm{Im}W_{\lambda',\mathbf{k}'})] = 2\nu\delta_{\lambda\lambda'}(\delta_{\mathbf{k},\mathbf{k}'} - \delta_{\mathbf{k},-\mathbf{k}'})dt, \tag{67}$$

then the stochastic differential equation may be written as

$$d\bar{h}^{\text{in}}_{ij} = b_{ij}dt + dW_{ij}. \tag{68}$$

With the gravitational field expressed as in Eq. (8), and with the retarded solution still given by Eq. (9), the field equations for the free part of the field become

$$[\tfrac{1}{2}(DD_* + D_*D) + \tfrac{1}{8}\beta(D - D_*)^2 - \partial_l\partial^l]\bar{h}^{\text{in}}_{ij} = 0, \tag{69}$$

which is the stochastic version of Eq. (10) in the transverse traceless gauge.

In order to solve these equations, Schrödinger's equation must first be solved in Q space. Then the b's are calculated using Eqs. (56) and (58). Once the b's are known, Eq. (68) can in principle be solved.

In general, it is more difficult to solve the stochastic equations than to solve the usual quantum mechanical equations. Even for the free field the stochastic processes for excited states are difficult to calculate. So far, the stochastic method has proven useful in practical problems only for stationary state problems where considerable simplifications occur. See, for example, Simon[24] for a review of results in this area. Although Simon does not explicitly make the connection with Nelson's theory, many of the methods he discusses may be considered as applications of the Fenyes–Nelson model to stationary state quantum systems.

We now illustrate the theory for the ground state field. The solution to Schrödinger's equation in Q space for the ground state is

$$\psi(Q) = \prod_{\lambda,\mathbf{k}} \exp\left[-|Q(\lambda,\mathbf{k})|^2 \frac{\omega}{\hbar 128\pi G}\right], \quad \omega = |\mathbf{k}| \tag{70}$$

up to a normalization constant. The b's are found from Eqs. (56) and (58) with $S_N = 0$. One finds for the b's of Eq. (6)

$$b^{ij}(x,t) = -2\nu/(\sqrt{2})L^{3/2} \times \sum^{\lambda,\mathbf{k}} \epsilon^{ij}(\lambda,\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \frac{\omega}{\hbar 32\pi G} Q(\lambda,\mathbf{k},t). \tag{71}$$

The stochastic differential equation becomes

$$d\bar{h}^{ij} = b^{ij}dt + dW^{ij}. \tag{72}$$

Using the property of the W^{ij} in Eq. (66),

$$E\,[dW^{ij}(x,t_x)dW^{kl}(y,t_y)] = 2\nu dt\delta^{ijkl}(\mathbf{x}-\mathbf{y}), \tag{73}$$

where the delta function is that of Eq. (32), the stochastic equations may be integrated to yield

$$E\,(\bar{h}^{ij}(x)\bar{h}^{kl}(y)) = \hbar(16\pi G)\int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \times e^{-(\nu/16\pi G\hbar)|\mathbf{k}||t_x - t_y|}\delta^{ijkl}(\mathbf{k})/|\mathbf{k}|. \tag{74}$$

Since b^{ij} of Eq. (71) is linear in the Q's, the process turns out to be Gaussian with zero mean so the covariance (74) determines all higher moments from the rule (36). The covariance may also be written as a four-dimensional integral

$$E(\bar{h}^{ij}(x)\bar{h}^{kl}(y)) = (32\pi G\hbar)\int \frac{d^4k}{(2\pi)^4}\frac{\delta^{ijkl}(\mathbf{k})}{k_0^2+\mathbf{k}^2}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})} \times e^{ik_0\nu/(\hbar 16\pi G)(t_x - t_y)}. \tag{75}$$

In expressions (73)–(75) the infinite volume limit has been taken. Nelson's value for the diffusion parameter is

$$\nu_N = 16\pi G\hbar. \tag{76}$$

Comparing Eq. (75) with (35), it may be shown that the stochastic covariance in (75) may be obtained from the quantum covariance in (35) by analytically continuing

$$t_x \to -i[\nu/(\hbar 16\pi G)]t_x, \qquad t_y \to -i[\nu/(\hbar 16\pi G)]t_y. \tag{77}$$

This procedure of analytic continuation yields the Schwinger function. Using the rule (36), we obtain the general result: The moments of the stochastic theory are equal to the Schwinger functions of the quantum theory with the times scaled by the factor $\nu/(\hbar 16\pi G)$. When Nelson's value for the diffusion parameter is chosen [Eq. (76)] the stochastic covariances become equal to the Schwinger functions. This result is similar to the results obtained in scalar field theory[7,9,10] and in electromagnetism.[8,9,11]

Examining Eq. (75), we see that the covariance is not manifestly Lorentz covariant, even discounting for the fact that we have chosen a noncovariant gauge. Lorentz covariance is violated by more than just a harmless gauge transformation in Eq. (75). This is a surprising and perhaps paradoxical result which has been known for some time.[7–11] Despite this lack of manifest Lorentz covariance in the ground state, there is good reason to believe that the experimental predictions of the stochastic theory are consistent with special relativity and are in fact the same as ordinary quantum theory. The argument is as follows. Since for any real value of ν we have a stochastic model for a given solution to Schrödinger's equation, and since we believe that this equation contains all of the experimentally verifiable predictions of quantum mechanics, then there is reason to think that it is impossible to measure the diffusion parameter. If this is true, then all of the experimentally measurable predictions of the theory (scattering cross sections, line spectra, etc.) must be constants when considered as analytic functions of ν. Because of this, we may consider analytic continuations to complex ν without affecting measurable predictions of the theory. When the stochastic theory is continued to

$$\nu = i\hbar(16\pi G), \tag{78}$$

then one finds that the moments of the stochastic theory become the Green's functions of quantum field theory. This result was also found for the scalar field[10] and the electromagnetic field.[11] The stochastic theory can be expressed in a mathematical form which is identical to ordinary quantum theory when ν is continued to the value in Eq. (78) (or its complex conjugate). See, for example, Ref. 25 for a derivation of the operator formalism of quantum mechanics within this framework. The methods of[25] generalized easily to the present theory in Q space.

Stochastic quantization in the ground state for a field theory leads to moments which are analytic continuations to imaginary times of quantum moments. The stochastic interpretation suggests that a certain reality be attributed to the theory so continued. It is interesting that imaginary time continuations have played a surprisingly important role in modern physics. Euclidean field theory[26] has led to advances in constructive field theory. Complex manifold techniques[27] have led to a deeper understanding of gauge theories. Bound state problems like the Bethe–Salpeter equation are often best solved by making a "Wick rotation" to imaginary times. Perhaps the results of stochastic quantization provide an explanation for the usefulness of imaginary time continuations.

V. THE APPROXIMATION OF RANDOM CLASSICAL RADIATION

In electromagnetism it has proven interesting in several applications to approximate the ground state of the quantum field by a superposition of classical plane waves with random phases. This approximation has become known as random electrodynamics[28] and it has been found that the diffusion of charged particles in harmonic oscillator potentials and exposed to such radiation leads convincingly to Schrödinger's equation with the correct nonrelativistic Lamb shift.[29] We present this random phase approximation for the linear gravity theory in the hope that it may find similar uses and also for comparison with the stochastic theory.

We first write the metric perturbations as general solutions to the free field equations. We consider only the free field in the transverse traceless gauge

$$h^{ij}(\mathbf{x},t) = \frac{1}{(\sqrt{2})L^{3/2}}\sum_{\lambda,\mathbf{k}} \epsilon^{ij}(\lambda,\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}Q(\lambda,\mathbf{k},t), \tag{79}$$

$$Q(\lambda,\mathbf{k},t) = 32\pi G\hbar/|\mathbf{k}| \times\{\cos[\omega t + \theta_1(\lambda,\mathbf{k})] + i\cos[\omega t + \theta_2(\lambda,\mathbf{k})]\}. \tag{80}$$

Reality demands

$$\theta_1(\lambda,-\mathbf{k}) = \theta_1(\lambda,\mathbf{k}), \qquad \theta_2(\lambda,-\mathbf{k}) = \theta_2(\lambda,\mathbf{k}) + \pi. \tag{81}$$

The θ's are all independent of one another except for the conditions in Eq. (81). They are random phases which take on values from 0 to 2π. The averaging process is carried out by integrating over these phases. One finds for the covariance

$$E_\theta(h^{ij}(x)h^{kl}(y)) = 16\pi G\hbar\int\frac{d^3k}{(2\pi)^3}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}\cos(\omega(t_x-t_y))\delta^{ijkl}(\mathbf{k})/|\mathbf{k}|. \tag{82}$$

It is easy to show that this is equal to the symmetrized quantum expectation

$$E_\theta(h^{ij}(x)h^{kl}(y)) = \langle 0|\,\mathrm{Sym}(h^{ij}(x)h^{kl}(y))|0\rangle, \tag{83}$$

where Sym denotes the symmetrization operation

$$\mathrm{Sym}(\phi_1(x_1)\times\cdots\times\phi_n(x_n)) = \sum_P \frac{1}{n!}\phi_1(x_1)\times\cdots\times\phi_n(x_n), \tag{84}$$

and where P denotes a sum over all permutations of the arguments of the fields.

It can be shown that in the infinite volume limit the higher moments of the random phase average satisfy the Gaussian combinatoric rule [Eq. (36)]. This is true despite the fact that the individual Fourier coefficients in Eq. (79) are not distributed normally. The reason is the central limit theorem, as the field is a sum of an infinite number of independent random variables. Since the symmetrized quantum expectations also satisfy the Gaussian combinatoric rule, it follows that

$$\begin{aligned} &E_\theta(h^{i_1j_1}(x_1)\times\cdots\times h^{i_nj_n}(x_n)) \\ &\quad = \langle 0|\,\mathrm{Sym}(h^{i_1j_1}(x_1)\times\cdots\times h^{i_nj_n}(x_n))|0\rangle. \end{aligned} \tag{85}$$

A more detailed derivation of this result has been given for electromagnetism by Boyer[28] whose analysis can be applied with little modification to the linear gravity theory to derive Eq. (85). It is important to realize that (85) is not true for a one-dimensional oscillator or even for a finite dimensional oscillator. It can only be derived in the present case in the infinite volume limit. Thus the random phase approximation does not give a very detailed model of the quantum field accurate down to the level of a few normal modes. It is unlikely, in the author's opinion, that a consistent interpretation of the full quantum field theory could be obtained from a classical random phase approximation for field theory. This is in sharp distinction to the stochastic quantization theory of Sec. IV which allows a consistent reinterpretation of all quantum phenomena. Still, the random phase model can be useful when considering the effects of vacuum fluctuations on matter.

[1] I. Fenyes, Z. Phys. **132**, 81–106 (1952).
[2] E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton U. P., Princeton, N. J., 1967).
[3] E. Nelson, Phys. Rev. **150**, 1079 (1966).
[4] M. Davidson, Physica A **96**, 465–486 (1979).
[5] M. Davidson, Lett. Math. Phys. **3**, 271–277 (1979).
[6] D. Shucker, Lett. Math. Phys. **4**, 61–65 (1980).
[7] F. Guerra and P. Ruggiero, Phys. Rev. Lett. **31**, 1022 (1973).
[8] F. Guerra and M. I. Loffredo, Lett. Nuovo Cimento **27**, 41–45 (1980).
[9] S. M. Moore, Found. Phys. **9**, 237–259 (1979).
[10] M. Davidson, Lett. Math. Phys. **4**, 101–106 (1980).
[11] M. Davidson, "Stochastic Quantization of the Electromagnetic Field," J. Math. Phys. **22**, 2588 (1981).
[12] A. Einstein, *The Meaning of Relativity* (Princeton U. P., Princeton, N. J., 1979).
[13] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[14] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).
[15] P. Levy, *Processus Stochastique et Mouvement Brownien* (Gauthier–Villars, Paris, 1965).
[16] J. L. Doob, *Stochastic Processes* (Wiley, New York, 1953).
[17] K. Ito, *Lectures on Stochastic Processes* (Tata, Bombay, 1961).
[18] K. Ito and H. P. Mckean, Jr., *Diffusion Processes and Their Sample Paths* (Springer-Verlag, Berlin, 1965).
[19] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1971).
[20] L. Breiman, *Probability* (Addison-Wesley, Reading, Mass., 1968).
[21] E. B. Dynkin, *Markov Processes* (Springer-Verlag, Berlin, 1965).
[22] D. W. Strook and S. S. Varadhan, *Multidimensional Diffusion Processes* (Springer-Verlag, Berlin, 1979).
[23] D. Bohm, Phys. Rev. **85**, 180–193 (1952).
[24] B. Simon, *Functional Integration and Quantum Physics* (Academic, New York, 1979).
[25] M. Davidson, Lett. Math. Phys. **3**, 367–376 (1979).
[26] B. Simon, *The $P(\phi)_2$ Euclidean (Quantum) Field Theory* (Princeton U. P., Princeton, N. J., 1974).
[27] D. E. Lerner and P. D. Sommers, *Complex Manifold Techniques in Theoretical Physics* (Pitman, London, 1979).
[28] T. H. Boyer, Phys. Rev. D **11**, 790–808, 809–830 (1975).
[29] L. de la Peña and A. M. Cetto, J. Math. Phys. **20**, 469–483 (1979).

The continuum limit of a classical 3-component, one-dimensional Heisenberg model is Brownian motion on the surface of a sphere

D. Isaacson [a)]
Rutgers University, New Brunswick, New Jersey 08903 [b)]

(Received 23 October 1979; accepted for publication 18 July 1980)

We show by explicit computation that when the temperature T is chosen as a function of the lattice spacing ϵ so that the correlation length stays fixed at one as $\epsilon \rightarrow 0$, then the correlation functions of a classical one-dimensional Heisenberg model converge as $\epsilon \rightarrow 0$ to the correlation functions of Brownian motion on the surface of a sphere. We remark that this provides a "statistical mechanical" construction of Brownian motion on the surface of a sphere.

PACS numbers: 05.40. + j

I. INTRODUCTION

In this section we define the correlation functions (or moments) for Brownian motion on a sphere, and for the classical one-dimensional Heisenberg model with lattice spacing ϵ at temperature T.

In Secs. 2 and 3 we give formulas for the moments, and we show that the temperature may be chosen as a function of the lattice spacing so that as the lattice spacing goes to zero the correlation length stays fixed at one. With this normalization the correlation functions of the Heisenberg model approach the correlation functions of Brownian motion on the surface of the three sphere (S^2).

This computation was inspired by the scaling limit conjecture of Glimm and Jaffe (see Ref. 1 where the conjecture is explained and proven for ϕ_1^4 and references are given to the papers of Glimm and Jaffe). We note that the conjecture has been extended to $|\boldsymbol{\phi}|_1^4$ in Ref. 2.

The proof given in this paper is for S^2 however with slight modification it holds for S^n for when $n \geqslant 1$. When $n = 0$, S^0 is just two points and we get Brownian motion on two points (a Bernoulli process) as a limit of the spin 1/2 Ising model.[1] We comment on weak convergence of the associated measures at the end of Sec. 3. We do not prove weak convergence here. We give a direct proof of the convergence of all correlation functions using only elementary properties of the spherical harmonics.

To establish notation we briefly describe Brownian motion on S^2.[3] Let's introduce coordinate functions on S^2 by

$$\begin{aligned} \omega_1 &= \omega_1(\theta,\phi) \equiv \cos\phi \sin\theta, & 0 \leqslant \theta \leqslant \pi, \\ \omega_2 &= \omega_2(\theta,\phi) \equiv \sin\phi \sin\theta, & 0 \leqslant \phi < 2\pi, \\ \omega_3 &= \omega_3(\theta,\phi) \equiv \cos\theta. \end{aligned} \tag{1}$$

The Laplacian on S^2 is given by

$$\Delta \equiv (\sin\theta)^{-1}\partial/\partial\theta \sin\theta\, \partial/\partial\theta + (\sin\theta)^{-2}\partial^2/\partial\phi^2. \tag{2}$$

We denote the kernel of $e^{t/2\Delta}$ by $P(t;u,v)$. Thus, for $t > 0$ and $u,v \in S^2$

$$\partial/\partial t\, P(t;u,v) = (\Delta/2) P(t;u,v), \tag{3}$$

where $\delta(u,v)$ is defined by

$$\int_{S^2} \delta(u,v) h(v) d\Omega(v) \equiv h(u). \tag{4}$$

Here h is any smooth function on S^2 and

$$d\Omega(v) \equiv (1/4\pi) \sin\theta_v\, d\theta_v d\phi_v. \tag{5}$$

An interpretation of $P(t;u,v)$ is that if a particle begins Brownian motion on S^2 at u when $t = 0$, then the probability it will be in $B \subset S^2$ when $t = \tau > 0$ is

$$P(\tau;u,B) \equiv \int_B P(\tau;u,v)\, d\Omega(v). \tag{6}$$

The semigroup property $e^{(t+s)/2\Delta} = e^{t/2\Delta} e^{s/2\Delta}$ is reflected in the Markoff property

$$P(t+s;u,v) = \int_{S^2} P(t;u,z) P(s;z,v)\, d\Omega(z), \tag{7}$$

which has the interpretation that the probability $P(t+s;u,B)$ of finding the particle at time $t+s$ in the set B is the 'sum' over all regions $d\Omega(z)$ of the product of the probability for the particle to go from u at $t = 0$ into $d\Omega(z)$ at time t, with the probability that a particle starting in $d\Omega(z)$ goes into B at time s.

The correlation functions $B_n(\cdots)$ for the coordinates of the particle undergoing Brownian motion on S^2 are (for $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$)

$$B_n(t_1,t_2,\ldots,t_n;j_1,j_2,\ldots,j_n) \equiv \int_{(S^2)^n} \cdots \int w_{j_1}(u_1)\cdots w_{j_n}(u_n) P(t_2 - t_1;u_1,u_2)\cdots P(t_n - t_{n-1};u_{n-1},u_n)\, d\Omega(u_n)\cdots d\Omega(u_1). \tag{8}$$

We next define the correlation functions for the one-dimensional, 3-component, classical Heisenberg model.[4–6]

The one-dimensional Heisenberg model describes the behavior of a line of particles with spins, $\sigma^\kappa, \kappa = 0, \pm 1, \pm 2, \ldots$, where

$$\sigma^\kappa = (\sigma_1^\kappa, \sigma_2^\kappa, \sigma_3^\kappa) \quad \text{and} \quad |\sigma^\kappa| = 1. \tag{9}$$

We assume the particles are located on lattice points of spacing ϵ apart (see Fig. 1).

For any point t we make the definition

a) Research partially supported by N.S.F. Grant No. MCS-77-03568.
b) Present address: Rensselaer Polytechnic Institute, Math Dept., Troy, N.Y. 12181.

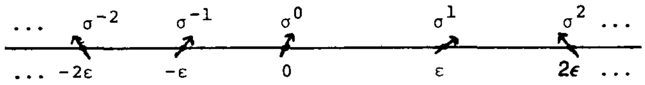


FIG. 1. Typical configuration with lattice spacing ϵ.

$$\sigma(t)\equiv\sigma^{[t/\epsilon]}. \tag{10}$$

The n-point correlation functions for the Heisenberg model at temperature $T\equiv\beta^{-1}$ are

$$\mathrm{EL4}H_n(t_1,\ldots,t_n;j_1,\ldots,j_n;\beta,\epsilon)=\langle\sigma_{j_1}(t_1)\cdots\sigma_{j_n}(t_n)\rangle_{\beta,\epsilon} \tag{11}$$

$$\equiv\lim_N z_N^{-1}\int_{|\sigma^{-N}|=1\ldots|\sigma^N|=1}\cdots\int\sigma_{j_1}(t_1)\cdots\sigma_{j_n}(t_n)\times e^{\beta\sum_{k=-N}^{N-1}\sigma^k\cdot\sigma^{k+1}}\,d\Omega(\sigma^{-N})\cdots d\Omega(\sigma^N),$$

where

$$z_N\equiv\int_{|\sigma^{-N}|=1\ldots|\sigma^N|=1}\cdots\int e^{\beta\sum_{k=-N}^{N-1}\sigma^k\cdot\sigma^{k+1}}\,d\Omega(\sigma^{-N})\cdots d\Omega(\sigma^N). \tag{12}$$

We may now state

Theorem 1: If the temperature β^{-1} is chosen so that the correlation length $\xi\equiv1$, then

$$\coth\beta-\beta^{-1}=e^{-\epsilon} \tag{13}$$

and

$$\lim_{\epsilon\to0}H_n(t_1\cdots t_n;j_1\cdots,j_n;\beta(\epsilon),\epsilon)=B_n(t_1\cdots t_n;j_1\cdots j_n). \tag{14}$$

The proof is given by direct computations in the next three sections.

II. CORRELATION FUNCTIONS FOR BROWNIAN MOTION ON S^2

Let $Y_l^m=Y_l^m(\theta,\phi)$ denote the spherical harmonics[7] normalized so that

$$\int_{S^2}\overline{Y}_l^m(v)Y_{l'}^{m'}(v)\,d\Omega(v)=\delta_{l,l'}\cdot\delta_{m,m'}. \tag{15}$$

Thus

$$Y_l^m(\theta,\phi)=N_l^mP_l^{|m|}(\cos\theta)e^{im\phi},$$
$$N_l^m=[(2l+1)(l-|m|)!/(l+|m|)!]^{1/2}. \tag{16}$$

The Y_l^m for $l=0,1,2,\ldots$ and $|m|\leqslant l$ are a complete orthonormal set of eigenfunctions [on $L_2(S^2)$] for the Laplacian, i.e.,

$$\Delta Y_l^m=-l(l+1)Y_l^m. \tag{17}$$

The spectral theorem applied to $e^{t\Delta/2}$ yields the formula

$$P(t;u,v)=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}e^{-tl(l+1)/2}Y_l^m(u)\,\overline{Y_l^m(v)}. \tag{18}$$

If we substitute (18) into (8) we get the formulas (assuming $\tau_j\equiv t_{j+1}-t_j$ for $j=1,\ldots,n-1$)

$$B_n(t_1\ldots t_n;j_1\cdots j_n) \tag{19}$$
$$=\sum_{\substack{l_1\ldots l_{n-1}\\ m_1\ldots m_{n-1}}}\exp\Big(-\sum_{j=1}^{n-1}\tau_jl_j(l_j+1)/2\Big)\langle00|w_{j_1}|l_1m_1\rangle\times\langle l_1m_1|w_{j_2}|l_2m_2\rangle\cdots\langle l_{n-1}m_{n-1}|w_{j_n}|00\rangle$$

where

$$\langle l,m|w_k|l',m'\rangle\equiv\int_{S^2}Y_l^m(u)w_k(u)\overline{Y}_{l'}^{m'}(u)\,d\Omega(u). \tag{20}$$

We remark that the series (19) reduces to a finite sum.

In the next section we give an analogous formula for the correlation functions $H_n(\cdots)$ of the Heisenberg model.

III. CORRELATION FUNCTIONS FOR THE HEISENBERG MODEL

The partition function z_N and two-point function $\langle\sigma(0)\cdot\sigma(t)\rangle_{\beta,1}$ were computed explicitly in Ref. 4, where it was pointed out that the Heisenberg model is exactly solvable. It follows from Ref. 4, and from the formulas given in the rest of this section, that

$$z_N=[\sinh\beta/\beta]^{2N}$$
$$\langle\sigma(0)\cdot\sigma(t)\rangle_{\beta,\epsilon}=e^{-\epsilon\xi^{-1}[t/\epsilon]}, \tag{21}$$

where ξ is the "correlation length" given by

$$\xi=-\epsilon[\ln(\coth\beta-\beta^{-1})]^{-1}. \tag{22}$$

We express the correlation functions in a form suitable for proving Theorem 1. Observe that if we let $k_i\equiv[t_i/\epsilon]$, then

$$\int_{|\sigma^{-N}|=1}\cdots\int_{|\sigma^N|=1}\sigma_{j_1}(t_1)\cdots\sigma_{j_n}(t_n)\times\exp\Big(\beta\sum_{k=-N}^{N-1}\sigma^k\cdot\sigma^{k+1}\Big)d\Omega(\sigma^{-N})\cdots d\Omega(\sigma^N)$$
$$=\int_{|\sigma^{k_1}|=1}\cdots\int_{|\sigma^{k_n}|=1}\sigma_{j_1}^{k_1}\cdots\sigma_{j_n}^{k_n}\exp\Big(\beta\sum_{k=k_1}^{k_n-1}\sigma^k\cdot\sigma^{k+1}\Big)$$
$$\times\Big[\int_{|\sigma^{-N}|=1}./.\int_{|\sigma^N|=1}\exp\Big(\beta\sum_{k=-N}^{k_1-1}\sigma^k\cdot\sigma^{k+1}\Big)\times\exp\Big(\beta\sum_{k=k_n}^{N}\sigma^k\cdot\sigma^{k+1}\Big)d\Omega(\sigma^{-N})./.d\Omega(\sigma^N)\Big]$$
$$\times d\Omega(\sigma^{k_1})\cdots d\Omega(\sigma^{k_n})$$

(here the / means omit the variables σ^k with $k_1\leqslant k\leqslant k_n$)

$$=[\sinh\beta/\beta]^{2N-(k_n-k_1+1)}\int\cdots\int\sigma_{j_1}^{k_1}\cdots\sigma_{j_n}^{k_n}e^{\beta\sum_{k_1}^{k_n-1}\sigma^k\cdot\sigma^{k+1}}\,d\Omega(\sigma^{k_1})\cdots d\Omega(\sigma^{k_n}). \tag{23}$$

We next use the Funk-Hecke theorem[8] to write

$$e^{(\beta\sigma^k\cdot\sigma^{k+1})}=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}c_l(\beta)Y_l^m(\sigma^k)\overline{Y}_l^m(\sigma^{k+1}), \tag{24}$$

where, if $P_l(t)$ denotes the lth Legendre polynomial

$$C_l(\beta)=\tfrac{1}{2}\int_{-1}^{+1}e^{\beta t}P_l(t)\,dt. \tag{25}$$

We substitute (24) into the integrals in (23) of the form (when a and b denote consecutive k_j).

$$\int\cdots\int\sigma^a_{j_a}\sigma^b_{j_b}\exp\beta\sum_{k=a}^{b-1}\sigma^k\cdot\sigma^{k+1}\,d\Omega(\sigma^{a+1})\cdots d\Omega(\sigma^{b-1})$$
$$=\int\cdots\int\sigma^a_{j_a}\sigma^b_{j_b}\prod_{k=a}^{b-1}\left[\sum_{l=0}^{\infty}\sum_{m=-l}^{l}c_l(\beta)Y^m_l(\sigma^k)\bar{Y}^m_l(\sigma^{k+1})\right]d\Omega(\sigma^{a+1})\cdots d\Omega(\sigma^{b-1})$$
$$=\sum_{l_a\cdots l_{b-1}}\sum_{m_a\cdots m_{b-1}}c_{l_a}\cdots c_{l_{b-1}}\int\cdots\int\sigma^a_{j_a}\sigma^b_{j_b}Y^{m_a}_{l_a}(\sigma^a)\bar{Y}^{m_a}_{l_a}(\sigma^{a+1})\cdots Y^{m_{b-1}}_{l_{b-1}}(\sigma^{b-1})\bar{Y}^{m_{b-1}}_{l_{b-1}}(\sigma^b)\,d\Omega(\sigma^{a+1})\cdots d\Omega(\sigma^{b-1})$$
$$=\sum_{l_a}\sum_{m_a}\sigma^a_{j_a}\sigma^b_{j_b}c_{l_a}(\beta)^{b-a}Y^{m_a}_{l_a}(\sigma^a)\bar{Y}^{m_a}_{l_a}(\sigma^b).\tag{26}$$

If we substitute this into $H^{(\cdots)}_n$ and use the notation $D_l(\beta)\equiv(\beta/\sinh\beta)c_l(\beta)$ then

$$H_n(t_1,\ldots,t_n;j_1,\ldots,j_n;\beta,\epsilon)$$
$$=\sum_{\substack{l_1\ldots l_{n-1}\\ m_1\ldots m_{n-1}}}D_{l_1}(\beta)^{k_2-k_1}\ldots D_{l_{n-1}}(\beta)^{k_n-k_{n-1}}\times\langle 00|\sigma_{j_1}|l_1m_1\rangle\langle l_1m_1|\sigma_{j_2}|l_2m_2\rangle$$
$$\cdots\langle l_{n-1}m_{n-1}|\sigma_{j_n}|00\rangle.\tag{27}$$

We now prove Theorem 1 by showing that if $\beta=\beta(\epsilon)$ as in (13) then

$$\lim_{\epsilon\to 0}D_l(\beta)^{k_{j+1}-k_j}=e^{-\tau_jl(l+1)/2},\tag{28}$$

where $\tau_j=t_{j+1}-t_j$.

Proof: We have that (13)

$$(e^\beta+e^{-\beta})/(e^\beta-e^{-\beta})-\beta^{-1}=e^{-\epsilon}.\tag{29}$$

We find the Laurent expansion for $\beta(\epsilon)$ by taking ϵ small on the right and β large on the left. Thus

$$1-\beta^{-1}+O(e^{-2\beta})=1-\epsilon+\epsilon^2/2+O(\epsilon^3)\tag{30}$$

which implies that

$$\beta(\epsilon)=\epsilon^{-1}+O(1).\tag{31}$$

We study $c_l(\beta)$ for large β by integrating by parts twice

$$c_l(\beta)=\tfrac{1}{2}\int_{-1}^{1}e^{\beta t}P_l(t)\,dt$$
$$=(e^\beta P_l(1)-e^{-\beta}P_l(-1))/2\beta-(e^\beta P'_l(1)-e^{-\beta}P'_l(-1))/2\beta^2+O(e^\beta/\beta^3).\tag{32}$$

The Legendre polynomials satisfy[8]

$$P_l(1)=1,\quad P_l(-1)=(-1)^l,\quad P'_l(1)=l(l+1)/2,$$
$$P'_l(-1)=l(l+1)/2(-1)^{l+1}.\tag{33}$$

Since $D_l(\beta)=c_l(\beta)(\beta/\sinh\beta)$

$$D_l(\beta)=\frac{e^\beta-e^{-\beta}(-1)^l}{e^\beta-e^{-\beta}}-\frac{l(l+1)}{2\beta}\times\left[\frac{e^\beta-e^{-\beta}(-1)^{l+1}}{e^\beta-e^{-\beta}}\right]+O(\beta^{-2}).\tag{34}$$

Now use (31) in the above to find

$$D_l(\beta(\epsilon))=1-l(l+1)\epsilon/2+O(\epsilon^2).\tag{35}$$

Thus

$$\lim_{\epsilon\to 0}D_l(\beta)^{k_{j+1}-k_j}$$
$$=\lim_{\epsilon\to 0}[1-l(l+1)\epsilon/2+O(\epsilon^2)]^{[t_{j+1}/\epsilon]-[t_j/\epsilon]}$$
$$=e^{-\tau_jl(l+1)/2}.\tag{36}$$

This concludes the proof of Theorem 1.

We remark that the techniques of the above proof could be used to prove weak convergence of the associated probability measures on the space of functions on the sphere.

We thank the referee for pointing out that another approach to proving weak convergence of the probability measures associated with the discrete time Markov semigroup $P_\epsilon(t)$, induced by the Heisenberg model (acting on functions on the surface of the unit sphere and defined for $t=0,\epsilon,2\epsilon,\cdots$), would be to show that $P_\epsilon(t)$ converges strongly to the Markov semigroup $P(t)$ of Brownian motion by showing that $(P_\epsilon(\epsilon)-I)/\epsilon\to P(0)=-(1/2)$ (Laplacian on surface of sphere) and using Chernoff's "generalized Trotter product formula"[9] which essentially says that $[P_\epsilon(t)]=[P_\epsilon(\epsilon t)]^{1/\epsilon}\to P(t)$ if $(P_\epsilon(\epsilon)-I)/\epsilon\to P'(0)$.

ACKNOWLEDGMENTS

We thank James Glimm, Dan Marchesin, and Sandy Zabell for several enjoyable discussions on this and related matters.

[1] D. Isaacson, Comm. Math. Phys. **53**, 257 (1977).
[2] D. Isaacson and D. Marchesin, Comm. Pure Appl. Math. **31**, 659 (1978).
[3] K. Yosida, *Functional Analysis* (Springer, Berlin, 1965).
[4] M. E. Fisher, Am. J. Phys. **32**, 343 (1964).
[5] G. S. Joyce, Phys. Rev. **155**, 478 (1967).
[6] H. E. Stanley, Phys. Rev. **179**, 570 (1969).
[7] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Springer, Berlin, 1966).
[8] H. Hochstadt, *The Functions of Mathematical Physics* (Wiley, New York, 1971).
[9] P. Chernoff, J. Func. Anal. **2**, 238 (1968).

A cluster expansion in field theory [a)]

James Kowall and H. M. Fried
Physics Department, Brown University, Providence, Rhode Island 02912

(Received 6 June 1980; accepted for publication 11 November 1980)

A cluster expansion is proposed for the calculation of certain Green's functions in field theory. The results presented are for $\lambda\phi^4$ theory, expanded in the number of scalar closed loops through the introduction of the composite field, $\chi = i\sqrt{\lambda/12}\,\phi^2$. As an intermediate step, Feynman path integrals are used to perform the functional integration over the field degrees of freedom. The cluster expansion that results is equivalent to perturbation theory, but at lowest order sums up all the infrared behavior of the theory. Higher orders systematically include ultraviolet effects.

PACS numbers: 11.10. – z

It is well known that the infrared behavior of abelian field theories (as in QED) shows a simple exponentiation with the soft part of every Feynman diagram summing up to an eikonal form.[1] This infrared exponentiation is easily derived by a number of approaches, but the systematic inclusion of ultraviolet corrections has not been dealt with within these schemes.[2] An exception to this is the work of Fradkin, Esposito, and Termini,[3] who have obtained modified perturbation series expressions for Green's functions in external potentials. At lowest order these solutions show infrared exponentiation, but systematically include higher order corrections to this effect.

The present paper proposes a cluster expansion, which at lowest order sums up all the infrared behavior and shows the simple exponentiation of Feynman graphs. Higher orders in the cluster expansion systematically include ultraviolet effects. This expansion has the property, that at kth order it reproduces up to kth order in perturbation theory exactly.

The calculation will be limited to $\lambda\phi^4$ theory expanded in the number of scalar closed loops through the introduction of a composite field $\chi = i\sqrt{\lambda/12}\phi^2$. These Green's functions are of interest as the $n = 0$ limit of the n-component $O(n)$ invariant scalar field theory. As an intermediate step, Feynman path integrals are introduced to enable the functional integration over the fields to be performed, and the scalar field is represented as a particle coordinate. The path integration is performed exactly to obtain the perturbation series expansion of these Green's functions. From these expressions the cluster expansion is derived.

The field theory that is considered here $\lambda\phi^4$ theory in D-dimensional Euclidean space. The Lagrangian density is

$$\mathscr{L} = -\tfrac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}M_0^2\phi^2 - \frac{\lambda_0}{4!}\phi^4,$$

with $\mathbf{x}$ a D-dimensional vector in Euclidean space. The integration notation used is

$$\int d\mathbf{x} \equiv \int d^Dx, \quad \int d\mathbf{k} \equiv \int \frac{d^Dk}{(2\pi)^D}.$$

The generating functional for Green's functions is given by the functional integral

$$\mathscr{Z}[j] = \int \mathscr{D}[\phi] \exp\left(\int \mathscr{L}(x)\,d\mathbf{x} + \int j(x)\phi(x)\,dx\right).$$

This is evaluated through the introduction of a composite field, $\chi = i\sqrt{\lambda/12}\phi^2$, by the identity

$$\exp\left(-\tfrac{1}{4!}\lambda_0\int\phi^4\right) = \int \mathscr{D}[\chi] \exp\left(-\tfrac{1}{2}\int\chi^2 + i(\lambda_0/12)^{1/2}\int\phi^2\chi\right)$$

It is convenient to consider

$$\mathscr{Z}[j,k] = \int \mathscr{D}[\chi]\int \mathscr{D}[\phi] \times \exp\left(-\tfrac{1}{2}\int\chi^2 + \int k\chi + \int j\phi - \tfrac{1}{2}\int\phi G^{-1}\phi\right),$$

with

$$G^{-1}(\mathbf{x},\mathbf{y}|\chi) = \{-\nabla^2 + m_0^2 - i(\lambda_0/3)^{1/2}\chi(\mathbf{X})\}\delta^D(\mathbf{x}-\mathbf{y}).$$

The scalar field integration is now Gaussian, and is evaluated to give

$$\mathscr{Z}[j,k] = \int D[\chi] \exp\left(-\tfrac{1}{2}\int\chi^2 + \int k\chi \times + \tfrac{1}{2}\int jG[\chi]j + L[\chi]\right),$$

with

$$L[\chi] = -\tfrac{1}{2}\operatorname{Tr}\ln[-\nabla^2 + m_0^2 - i(\lambda_0/3)^{1/2}\chi]$$

representing the sum of all one scalar closed-loop graphs interacting with the composite field $\chi(\mathbf{x})$.

The Green's functions that are considered here are the scalar propagator

$$G(\mathbf{x}-\mathbf{y}) = \frac{\delta}{\delta j(x)}\frac{\delta}{\delta j(y)}\ln\mathscr{Z}|_{j=k=0},$$

and the composite field propagator (which is simply related to the scalar four-point function),

$$D(\mathbf{x}-\mathbf{y}) = \frac{\delta}{\delta k(x)}\frac{\delta}{\delta k(y)}\ln\mathscr{Z}|_{j=k=0}.$$

These propagators will be considered in an expansion in the number of scalar closed loops via the power series

[a)]Supported in part by the U. S. Department of Energy under Contract No. DE-AC-02-76 ER03130.A005 Task A-Theoretical.

$$e^{L[X]} = \sum_{n=0}^{\infty} \frac{1}{n!}(L[\chi])^n,$$

which results in a loopwise expansion of the theory. [3a] This expansion is a simple limit of the theory, and gives the $n = 0$ behavior of the n-component scalar field theory

$$\phi = (\phi_1, \phi_2, \ldots, \phi_n),$$

with

$$\phi^2 = \sum_{\alpha=1}^{n} \phi_\alpha^2.$$

This limit is of importance to the excluded volume problem of polymer physics, and to the random potential problem in solid state physics.[4]

The expansion in the number of closed loops results in

$$G(x-y) = \sum_{n=0}^{\infty} G^{(n)}(x-y),$$

$$D(x-y) = \sum_{n=0}^{\infty} D^{(n)}(x-y),$$

with

$$D^{(0)}(x-y) = \delta^D(x-y),$$

$$D^{(1)}(x-y) = \int \mathscr{D}[\chi]\chi(x)\chi(y)L[\chi]e^{-\frac{1}{2}\int \chi^2},$$

and

$$G^{(0)}(x-y) = \int \mathscr{D}[\chi]G(x,y|\chi)e^{-\frac{1}{2}\int \chi^2}.$$

Only these Green's functions will be considered in the remainder of this paper.

In order to carry out the functional integration, proper time representations for these expressions are now introduced. These take the form

$$G(x,y|\chi) = \int_0^\infty d\xi e^{-\xi m_0^2}\rho(\xi;x,y),$$

$$L[\chi] = \frac{1}{2}\int_0^\infty \frac{d\xi}{\xi} e^{-\xi m_0^2}\int d\mathbf{x}\, \rho(\xi;x,x),$$

with the density matrix

$$\rho(\xi;x,y) = \langle \mathbf{x}| e^{-\xi[-\nabla^2 - i(\lambda_0/3)^{1/2}\chi]}|\mathbf{y}\rangle.$$

This may be represented as a Feynman path integral.[5]

$$\rho(\xi;x,y) = \int_{x(0)=y}^{x(\xi)=x} \mathscr{D}[\mathbf{p}(s)]\mathscr{D}[\mathbf{x}(s)]e^{i\int_0^\xi ds \mathbf{p}\cdot\frac{d\mathbf{x}}{ds}} \times e^{-\int_0^\xi ds[\mathbf{p}^2(s) - i\sqrt{\lambda_0/3}\,\chi(\mathbf{x}(s))]},$$

where all particle trajectories are integrated over subject to the endpoint constraints. [4a] The composite field integration is now Gaussian, and is performed to give

$$G^{(0)}(x-y) = \int_0^\infty d\xi e^{-\xi m_0^2}\mathscr{N} \times \int_{x(0)=y}^{x(\xi)=x} \mathscr{D}[\mathbf{x}(s)]e^{-S[\mathbf{x}]},$$

and

$$D^{(1)}(x-y) = -\frac{\lambda_0}{6}\int d\mathbf{k}_1\int d\mathbf{k}_2 e^{i\mathbf{k}_1\cdot\mathbf{x} - i\mathbf{k}_2\cdot\mathbf{y}} \times \int_0^\infty \frac{d\xi}{\xi} e^{-\xi m_0^2}\int_0^\xi ds_1\int_0^\xi ds_2\, \mathscr{N} \times \int_{X(0)=X(\xi)} \mathscr{D}[\mathbf{x}(s)]e^{-S[\mathbf{x}]}e^{-i\mathbf{k}_1\cdot\mathbf{x}(s_1) + i\mathbf{k}_2\cdot\mathbf{x}(s_2)},$$

where the particle action is given by

$$S[\mathbf{x}] = \frac{1}{4}\int_0^\xi ds\left(\frac{d\mathbf{x}}{ds}\right)^2 + \frac{\lambda_0}{6}\int_0^\xi ds\int_0^\xi ds'\delta^D(\mathbf{x}(s) - \mathbf{x}(s'))$$

and the normalized by

$$N = \int \mathscr{D}[\mathbf{p}(s)]e^{-\int_0^\xi ds\mathbf{p}^2(s)}.$$

Note that the particle action includes a point nonlocal selfinteraction for the scalar particle as it moves on its trajectory. The closed-loop nature of the composite field propagator is reflected in the constraint of periodic orbits.

These Green's functions are now to be expanded in perturbation series, with

$$G^{(0)}(x-y) = \sum_{n=0}^{\infty} \frac{1}{n!}(-\lambda_0/6)^n G_n^{(0)}(x-y)$$

and

$$D^{(1)}(x-y) = \sum_{n=0}^{\infty} \frac{1}{n!}(-\lambda_0/6)^{n+1} D_n^{(1)}(x-y).$$

There then follow the identifications

$$G_n^{(0)}(x-y) = \int_0^\infty d\xi e^{-\xi m_0^2}\int_0^\xi ds_1\int_0^\xi ds_1' \cdots \int_0^\xi ds_n\int_0^\xi ds_n' \times \int d\mathbf{k}_1\cdots\int d\mathbf{k}_n \mathscr{N}\int_{x(0)=y}^{x(\xi)=x} \mathscr{D}[\mathbf{x}(s)]e^{-S_n[\mathbf{x}]},$$

and

$$D_n^{(1)}(x-y) = \int d\mathbf{k}\int d\mathbf{k}' e^{i\mathbf{k}\cdot\mathbf{x} - i\mathbf{k}'\cdot\mathbf{y}}\int_0^\infty \frac{d\xi}{\xi}e^{-\xi m_0^2} \times \int_0^\xi ds\int_0^\xi ds'\int_0^\xi ds_1' \cdots \int_0^\xi ds_n \times \int_0^\xi ds_n' \cdot N\int d\mathbf{k}_1\cdots\int d\mathbf{k}_n \times \int_{\mathbf{x}(0)=\mathbf{x}(\xi)} D[\mathbf{x}(s)]e^{-S_n[\mathbf{x}] - i\mathbf{k}\cdot\mathbf{x}(s) + i\mathbf{k}'\cdot\mathbf{x}(s')}$$

where

$$S_n[\mathbf{x}] = \frac{1}{4}\int_0^\xi ds\left(\frac{d\mathbf{x}}{ds}\right)^2 - i\sum_{j=1}^{n}[\mathbf{x}(s_j) - \mathbf{x}(s_j')]\cdot\mathbf{k}_j.$$

Written in this form all the path integrals are now Gaussian, and may be evaluated. The result of this integration gives for the Fourier transforms of these Green's functions the expressions

$$\tilde{G}_n^{(0)}(\mathbf{p}) = \int_0^\infty d\xi e^{-\xi[m_0^2 + \mathbf{p}^2]}\int_0^\xi ds_1\int_0^\xi ds_1' \cdots \int_0^\xi ds_n\int_0^\xi ds_n' \times \int d\mathbf{k}_1\cdots\int d\mathbf{k}_n e^{-\sum_{i=1}^{n}|S_i - S_i'|\mathbf{k}_i^2} \times e^{-2\mathbf{p}\sum_{i=1}^{n}(S_i - S_i')\mathbf{k}_i - 2\sum_{i>j=1}^{n} m_{ij}\mathbf{k}_i\cdot\mathbf{k}_j}$$

and

$$\tilde{D}_n^{(1)}(\mathbf{k}) = \int d\mathbf{p}\int_0^\infty \frac{d\xi}{\xi}e^{-\xi[m_0^2+\mathbf{p}^2]}\int_0^\xi ds_0\int_0^\xi ds_0' e^{-|s_0s_0'|\mathbf{k}^2}$$
$$\times e^{2(S_0-S_0')\mathbf{p}\cdot\mathbf{k}}\int_0^\xi\int_0^\xi dS_1'\cdots\int_0^\xi dS_n\int_0^\xi dS_n'$$
$$\times\int d\mathbf{k}_1\cdots\int d\mathbf{k}_n \exp\Big[-\sum_{j=1}^n |s_j-s_j'|\mathbf{k}_j^2$$
$$+2\sum_{j=1}^n (s_j-s_j')\mathbf{p}\cdot\mathbf{k}_j$$
$$-2\sum_{j=1}^n m_{0j}\mathbf{k}\cdot\mathbf{k}_j - 2\sum_{i>j=1}^n m_{ij}\mathbf{k}_i\cdot\mathbf{k}_j\Big]$$

where

$$m_{ij}\equiv(s_i-s_j')\theta(s_i-s_j')-(s_i-s_j)\theta(s_i-s_j)$$
$$+(s_i'-s_j)\theta(s_i'-s_j)-(s_i'-s_j')\theta(s_i'-s_j'),$$

and

$$\theta(x)=\begin{cases}1, & x>0,\\ \frac{1}{2}, & x=0,\\ 0, & x<0.\end{cases}$$

These expressions explicitly determine the contribution of all Feynman diagrams at nth order to these Green's functions in terms of parametric integrals. It may be observed that the k_i integrals label the momentum flowing through the n vertices at nth order, while the S_i integrals label the proper times at which the vertices are arranged, and so determine the ordering of propagators in every Feynman diagram. The final integration over p in the composited field propagator is the integration over the closed-loop momentum, while s_0 and s_0' label the proper times at which the external legs are attached.

There is a natural cluster expansion that suggests itself from the form of these parametric integrals. Defining the quantity

$$f_{ij}\equiv e^{-2m_{ij}\mathbf{k}_i\cdot\mathbf{k}_j}-1,$$

the fully interacting propagators may be written as

$$\tilde{G}^{(0)}(\mathbf{p})=\int_0^\infty d\xi e^{-\xi(m_0^2+\mathbf{p}^2)}\sum_{n=0}^\infty\frac{1}{n!}\Big(-\frac{\lambda_0}{6}\Big)^n$$
$$\times\prod_{j=1}^n\int_0^\xi ds_j\int_0^\xi ds_j'\int d\mathbf{k}_j\cdot e^{-|s_j-s_j'|\mathbf{k}_j^2+2(s_j-s_j')\mathbf{p}\cdot\mathbf{k}_j}$$
$$\times\prod_{i>j=1}^n(1+f_{ij}),$$

and

$$\tilde{D}^{(1)}(\mathbf{k})=-\frac{\lambda_0}{6}\int d\mathbf{p}\int\frac{d\xi}{\xi}e^{-\xi[m_0^2+\mathbf{p}^2]}\int_0^\xi ds_0\int_0^\xi ds_0'$$
$$\times e^{-|s_0-s_0'|\mathbf{k}^2+2(s_0-s_0')\mathbf{p}\cdot\mathbf{k}}\sum_{n=0}^\infty\frac{1}{n!}\Big(-\frac{\lambda_0}{6}\Big)^n$$
$$\times\prod_{j=1}^n\int_0^\xi ds_j\int_0^\xi ds_j'\int d\mathbf{k}_j e^{-|s_j-s_j'|\mathbf{k}_j^2+2\{(s_j-s_j')\mathbf{p}-m_{0j}\mathbf{k}\}\cdot\mathbf{k}_j}$$
$$\times\prod_{i>j=1}^n(1+f_{ij}).$$

Notice that the m_{ij} are bounded

$$|m_{ij}|\leqslant\xi,$$

while on dimensional grounds in the above integrals it is expected that

$$O(\xi)\sim 1/\mathbf{p}^2.$$

This then indicates that the f_{ij} are small for small momentum transfers; that is,

$$|f_{ij}|\ll 1 \quad \text{for all} \quad |\mathbf{k}_i|\ll|\mathbf{p}|,$$

and that an expansion in the f_{ij} will sum up the effect of soft exchanges at lowest order.[5a)]

This expansion is now performed. Following the usual treatment of statistical mechanics,[6] a cluster expansion is defined, which at lowest order sums up all infrared effects and in higher orders systematically includes ultraviolet effects. The result for the scalar propagator is

$$\tilde{G}^{(0)}(\mathbf{p})=\int_0^\infty d\xi e^{-\xi[m_0^2\exp[\mathbf{p}^2]]}\exp\Big[\sum_{l=1}^\infty\Big(-\frac{\lambda_0}{6}\Big)^l b_l(\xi,\mathbf{p})\Big],$$

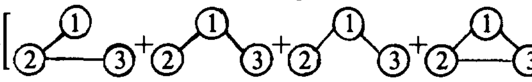
where the b_l are cluster integrals,)

$$b_l=\frac{1}{l!}[\text{sum of all possible } l\text{-clusters}],$$

and the l-clusters are l- particle graphs with each particle attached to at least one line and directly or indirectly to all other $l-1$ patricles. Some examples are

$$b_1=\text{①}=\int_0^\xi ds_1\int_0^\xi ds_1'\int d\mathbf{k}_1 e^{-|s_1-s_1'|\mathbf{k}_1^2+2(s_1-s_1')\mathbf{p}\cdot\mathbf{k}_1}$$

$$b_2=\frac{1}{2!}[\text{①}-\text{②}]=\frac{1}{2!}\int_0^\xi ds_1\int_0^\xi ds_1'\int_0^\xi ds_2\int_0^\xi ds_2'\int d\mathbf{k}_1\int d\mathbf{k}_2$$
$$\times f_{12}\exp[-|s_1-s_1'|\mathbf{k}_1^2-|s_2-s_2'|\mathbf{k}_2^2+2$$
$$\times(s_1-s_1')\mathbf{p}\cdot\mathbf{k}_1+2(s_2-s_2')\mathbf{p}\cdot\mathbf{k}_2],$$

$b_3=\frac{1}{3!}\Big[$  $\Big]$,

etc.

This result clearly shows the exponentiation of Feynman diagrams that is characteristic of summing the soft part of every graph. However, this result goes far beyond the simple infrared exponentiation; it is exact. The expansion to k th order in the cluster expansion will give up to the k th order in perturbation theory, exactly. The rules for writing down the cluster integrals are also very simple.

The result for the composite field propagator is similar

$$\tilde{D}^{(1)}=-\frac{\lambda_0}{6}\int d\mathbf{p}\int_0^\infty\frac{d\xi}{\xi}e^{-\xi[m_0^2+\mathbf{p}^2]}\int_0^\xi ds_0\int_0^\xi ds_0'$$
$$\times\exp\{-|s_0-s_0'|\mathbf{k}^2+2(s_0-s_0')\mathbf{p}\cdot\mathbf{k}\}$$
$$\times\exp\Big\{\sum_{l=1}^\infty\Big(-\frac{\lambda_0}{6}\Big)^l C_l\Big\},$$

where these cluster integrals are given as

$$C_1=\int_0^\xi ds_1\int_0^\xi ds_1'\int d\mathbf{k}_1 e^{-|s_1-s_1'|\mathbf{k}_1^2+2\mathbf{k}_1\cdot[(s_1-s_1')\mathbf{p}-m_{01}\mathbf{k}]},$$

$$C_2=\frac{1}{2!}\prod_{i=1,2}\Big[\int_0^\xi ds_i\int_0^\xi ds_i'\int d\mathbf{k}_i e^{-|s_i-s_i'|\mathbf{k}_i^2}$$
$$\times e^{2\mathbf{k}_i\cdot[(s_i-s_i')\mathbf{p}-m_{0i}\mathbf{k}]}\Big]f_{12}, \text{ etc.}$$

Similar expressions may be written for all the n-point Green's functions of the theory.

The techniques presented here depends upon the possibility of replacing functional integrals over field degrees of freedom by Feynman path integration.[5] The expansions generated are systematic ones, and should be applicable to a variety of problems with more ease than those of Fradkin, *et al.*[3]

[1]F. Block and A. Nordsieck, Phys. Rev. **52**, 54 (1937); D. R. Yennie, S. Frautschi, and H. Suura, Ann. Phys. **13**, 379 (1961); H. M. Fried, *Functional Methods and Models in Quantum Field Theory* (MIT Press, Cambridge, Mass., 1972).

[2]H. M. Fried, J. Phys. Lett. **40**, 89 (1979); Nucl. Phys. B (to be published) Brown Univ. HET preprint.

[3]E. S. Fradkin, V. Esposito, and S. Termini, Rev. Nuovo Cimento, Ser. I, Vol. 2, 498 (1970). See also E. S. Fradkin, Nucl. Phys. **76**, 588 (1966).

[3a]This loopwise expansion should not be confused with the ordinary loop expansion in field theory. The number of closed loops here refers to the propagation of the ϕ field in the presence of the composite χ field, with the effective interaction of the form $\phi^2\chi$. Note that the χ propagator is a δ function.

[4]V. J. Emery, Phys. Rev. B **11**, 239 (1975).

[4a]Here the integration measure follows Feynman's original formulation; see Feynman and Hibbs, *Quantum Mechanics and Path Integrals*, (McGraw–Hill, New York, 1965).

[5]For a different application of the loopwise expansion and use of Feyman path integrals see J. Kowall, "A Semiclassical Calculation of the Photon Propagator in Two-Dimensional Scalar QED," Brown Univ. HET Preprint 406.

[5a]Note that one difference here from the usual statistical mechanics case is that the f_{ij} are unbounded.

[6]K. Huang, *Statistical Mechanics* (Wiley, New York, 1963).

Similarity solutions of nonlinear Dirac equations and conserved currents

W. -H. Steeb and W. Erig
Universität Paderborn, Theoretische Physik, D-479 Paderborn, West Germany
W. Strampp
Gesamthochschule Kassel, Fachbereich Mathematik, D-3500 Kassel, West Germany

(Received 31 March 1981; accepted for publication 30 July 1981)

Nonlinear Dirac equations in one space dimension and three space dimensions are studied. Using the continuous symmetries of the nonlinear Dirac equations we reduce the system of nonlinear partial differential equations to a system of nonlinear ordinary differential equations, applying group theoretical methods, and give solutions of these equations. Moreover, we determine conserved currents using the continuous symmetries of the nonlinear Dirac equation.

PACS numbers: 11.30. – j, 02.20. – b, 02.40. + m

I. INTRODUCTION

Recently, several authors[1–4] have investigated nonlinear Dirac equations with fourth-order self-coupling. In the case of one space and one time dimension exact localized solutions have been described.[1,2] Solutions have also been given in four-dimensional space–time.[3,4] Takahashi[3,4] has shown that stringlike and ball-like soliton solutions exist in four-dimensional space–time.

The purpose of the present paper is twofold. First of all we demonstrate how with the knowledge of continuous symmetries, solutions of nonlinear Dirac equations can be obtained. We compare the solutions with those given by the authors cited above. For investigation of the continuous symmetries we adopt the modern approach due to Cartan. This means we cast the field equation (i.e., a system of partial differential equations) into an equivalent system of differential forms[5–8] and calculate the Lie derivative of these differential forms with respect to the infinitesimal generators (i.e., symmetry generator). As symmetry generators we consider space and time translations and an infinitesimal generator which is associated with a gauge transformation. The second purpose of the present paper is to derive conserved currents and conservation laws. Here we adopt the Hamilton–Cartan formalism[9–11] (jet bundle formalism) for first-order Lagrangians. In this approach the field equations, i.e., the nonlinear Dirac equation, is derived from a two-form when we study one space and one time dimension and from a four-form when we study three space and one time dimension. Noether's theorem can easily be formulated within this approach.[11] With the help of an example we show that conserved currents and conservation laws can also be obtained without the knowledge of a Lagrangian and a symmetry generator. Only the differential forms which are equivalent to the field equations are taken into account. The approach is similar to that given by Estabrook and Wahlquist.[12] Moreover, we describe a third possibility which to our knowledge is not known so far for obtaining conserved current, namely, taking into account the symmetry generator and the differential forms which are equivalent to the field equation.

The case with one space and one time dimension is described in detail, while the case with three space and one time dimension is only briefly studied since the approach is the same.

Throughout, the type of nonlinearity is given by the scalar interaction. However, the extension to other types of interactions like vector, tensor, axial vector, or pseudoscalar interaction is straightforward.

II. NONLINEAR DIRAC EQUATION

The Dirac equation with rest mass m_0 can be written as

$$\sum_{k=1}^{3} \hbar \frac{\partial}{\partial x_k}(\gamma_k \psi) - i\hbar \frac{\partial}{\partial x_4}(\gamma_4 \psi) + m_0 c\psi = 0, \qquad (2.1)$$

where $x_4 = ct$ and $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)^T$ (T means transpose). $\gamma_1, \gamma_2, \gamma_3$, and γ_4 are the following 4×4 matrices

$$\gamma_1 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$\gamma_3 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix},$$

$$\gamma_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \qquad (2.2)$$

Throughout we assume that $m_0 > 0$.
Introducing the quantity

$$\lambda = \hbar / m_0 c \qquad (2.3)$$

which has the dimension of a length, we obtain

$$\lambda \sum_{k=1}^{3} \frac{\partial}{\partial x_k}(\gamma_k \psi) - \lambda i \frac{\partial}{\partial x_4}(\gamma_4 \psi) + \psi = 0. \qquad (2.4)$$

In the following we study the nonlinear Dirac equation of the form

$$\lambda \sum_{k=1}^{3} \frac{\partial}{\partial x_k}(\gamma_k \psi) - \lambda i \frac{\partial}{\partial x_4}(\gamma_4 \psi) + \psi + \lambda^3 \epsilon \psi(\bar{\psi}\psi) = 0, \qquad (2.5)$$

 0022-2488/82/010145-09$02.50

where ϵ is a real parameter (coupling constant) and $\bar{\psi}\equiv(\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*)$. This means that we investigate scalar Fermi interaction. We mention that the extension of the following approach to other Fermi interactions such as vector interaction, pseudoscalar interaction and so on is straightforward. Now we put $\psi_j(x)\equiv u_j(x) + iv_j(x)$, where $j = 1,...,4$ and $x\equiv(x_1, x_2, x_3, x_4)$. $u_j(x)$ and $v_j(x)$ are real fields. If we insert $\psi_j(x)\equiv u_j(x) + iv_j(x)$ into the partial differential equations (2.5) we obtain a real system of partial differential equations. In the following we mainly investigate the case with one space dimension $[x\equiv(x_1, x_4)]$. In this case we have

$$\lambda \frac{\partial}{\partial x_1}(\gamma_1\psi) - \lambda i \frac{\partial}{\partial x_4}(\gamma_4\psi) + \psi(1 + \lambda\epsilon\bar{\psi}\psi) = 0. \tag{2.6}$$

Then we obtain the following nonlinear coupled system of eight partial differential equations;

$$\begin{aligned}
&\lambda \frac{\partial v_4(x)}{\partial x_1} + \lambda \frac{\partial v_1(x)}{\partial x_4} + u_1(x)[1 + \lambda\epsilon K] = 0,\\
&\lambda \frac{\partial v_3(x)}{\partial x_1} + \lambda \frac{\partial v_2(x)}{\partial x_4} + u_2(x)[1 + \lambda\epsilon K] = 0,\\
&-\lambda \frac{\partial v_2(x)}{\partial x_1} - \lambda \frac{\partial v_3(x)}{\partial x_4} + u_3(x)[1 + \lambda\epsilon K] = 0,\\
&-\lambda \frac{\partial v_1(x)}{\partial x_1} - \lambda \frac{\partial v_4(x)}{\partial x_4} + u_4(x)[1 + \lambda\epsilon K] = 0,\\
&-\lambda \frac{\partial u_4(x)}{\partial x_1} - \lambda \frac{\partial u_1(x)}{\partial x_4} + v_1(x)[1 + \lambda\epsilon K] = 0,\\
&-\lambda \frac{\partial u_3(x)}{\partial x_1} - \lambda \frac{\partial u_2(x)}{\partial x_4} + v_2(x)[1 + \lambda\epsilon K] = 0,\\
&\lambda \frac{\partial u_2(x)}{\partial x_1} + \lambda \frac{\partial u_3(x)}{\partial x_4} + v_3(x)[1 + \lambda\epsilon K] = 0,\\
&\lambda \frac{\partial u_1(x)}{\partial x_1} + \lambda \frac{\partial u_4(x)}{\partial x_4} + v_4(x)[1 + \lambda\epsilon K] = 0,
\end{aligned} \tag{2.7}$$

where

$$K(u(x),v(x))\equiv \sum_{j=1}^{2}(u_j^2(x) + v_j^2(x)) - \sum_{j=3}^{4}(u_j^2(x) + v_j^2(x)) \tag{2.8}$$

with $u\equiv(u_1, u_2, u_3, u_4)$ and $v\equiv(v_1, v_2, v_3, v_4)$.

III. NONLINEAR DIRAC EQUATION AND SYMMETRIES

Now we use group theoretical methods for reducing the nonlinear system of partial differential equations (2.6) into a nonlinear system of ordinary differential equations. For this purpose we cast the system of partial differential equations into an equivalent set of differential forms.[5–8] We put $\partial u_i(x)/\partial x_j \rightarrow p_{ij}$ and $\partial v_i(x)/\partial x_j \rightarrow q_{ij}$ $(i,j = 1,2,3,4)$. Then we find

$$\begin{aligned}
&F_1(u_1,...,v_4,p_{11},...,q_{44})\\
&\quad\equiv\lambda(q_{41} + q_{14}) + u_1(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_2(....................)\\
&\quad\equiv\lambda(q_{31} + q_{24}) + u_2(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_3(....................)\\
&\quad\equiv\lambda(-q_{21} - q_{34}) + u_3(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_4(....................)\\
&\quad\equiv\lambda(-q_{11} - q_{44}) + u_4(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_5(....................)\\
&\quad\equiv\lambda(-p_{41} - p_{14}) + v_1(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_6(....................)\\
&\quad\equiv\lambda(-p_{31} - p_{24}) + v_2(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_7(....................)\\
&\quad\equiv\lambda(p_{21} + p_{34}) + v_3(1 + \lambda\epsilon K(u,v)) = 0,\\
&F_8(....................)\\
&\quad\equiv\lambda(p_{11} + p_{44}) + v_4(1 + \lambda\epsilon K(u,v)) = 0,
\end{aligned} \tag{3.1}$$

and

$$\begin{aligned}
&\alpha_i\equiv du_i - p_{i1}dx_1 - p_{i4}dx_4,\\
&\beta_i\equiv dv_i - q_{i1}dx_1 - q_{i4}dx_4,
\end{aligned} \tag{3.2}$$

where $i = 1,2,3,4$. Instead of considering the nonlinear Dirac equation (2.5) we consider the differential system given by Eqs. (3.1) and (3.2) for investigating the symmetries.

A comment about the approach is in order. We consider the nonlinear system of differential equations within the jet bundle formalism[9–11] (compare also Sec. 5). The quantities α_i and β_i are called contact forms.

In a previous paper[5] the authors have shown that the differential system (3.1) and (3.2) is invariant under the infinitesimal generator

$$\bar{Z} = \sum_{k=1}^{4}\left(u_k \frac{\partial}{\partial v_k} - v_k \frac{\partial}{\partial u_k}\right) + \sum_{k=1}^{4}\sum_{i=1}^{4}\left(p_{ki}\frac{\partial}{\partial q_{ki}} - q_{ki}\frac{\partial}{\partial p_{ki}}\right). \tag{3.3}$$

It follows that the nonlinear Dirac equation is invariant under the transformation group generated by the infinitesimal generator $Z = \Sigma_k(u_k\partial/\partial v_k - v_k\partial/\partial u_k)$. $\bar{Z}$ is the once-extended vector field of Z.

Moreover, the nonlinear Dirac equation is invariant under space and time translation, since the partial differential equations do not depend explicitly on time and space coordinates. The infinitesimal generators taken the form

$$X = \frac{\partial}{\partial x_1}, \quad T = \frac{\partial}{\partial x_4}. \tag{3.4}$$

The once-extended infinitesimal generators are of the same form. Instead of considering the infinitesimal generators X and T for deriving conserved currents and similarity solutions we can take the infinitesimal generators

$$U = -\sum_{i=1}^{4}\left(p_{i1}\frac{\partial}{\partial u_i} + q_{i1}\frac{\partial}{\partial v_i}\right) \tag{3.5a}$$

and

$$V = -\sum_{i=1}^{4}\left(p_{i2}\frac{\partial}{\partial u_i} + q_{i2}\frac{\partial}{\partial v_i}\right). \tag{3.5b}$$

We notice that $X\lrcorner\alpha_i = U\lrcorner\alpha_i$, $X\lrcorner\beta_i = U\lrcorner\beta_i$, and $T\lrcorner\alpha_i = V\lrcorner\alpha_i$, $T\lrcorner\beta_i = V\lrcorner\beta_i$.

IV. SIMILARITY SOLUTIONS

In this section we derive similarity solutions to the nonlinear Dirac equation (2.6). For this purpose we consider a linear combination of the infinitesimal generators X, T, and Z, i.e.,

$$Y = a_1 \frac{\partial}{\partial x_1} + a_4 \frac{\partial}{\partial x_4} + a_5\Big(\sum_{k=1}^{4} u_k \frac{\partial}{\partial v_k} - v_k \frac{\partial}{\partial u_k}\Big), \tag{4.1}$$

where $a_1, a_4, a_5 \in \mathbb{R}$. The contraction of the one-forms $\alpha_1, ..., \alpha_4, \beta_1, ..., \beta_4$ by the vector field Y yields the zero-forms

$$Y \rfloor \alpha_i = -a_1 p_{i1} - a_4 p_{i4} - a_5 v_i,$$
$$Y \rfloor \beta_i = -a_1 q_{i1} - a_4 q_{i4} + a_5 u_i, \tag{4.2}$$

where $i = 1, ..., 4$. In order to obtain a similarity solution to the nonlinear Dirac equation we consider the linear partial differential equations which are given by $s^*(Y \rfloor \alpha_i) = 0$ and $s^*(Y \rfloor \beta_i) = 0$ where $i = 1, ..., 4$. s is the mapping $s(x) = (x, u(x), v(x), p(x), q(x))$ and s^* is the pull back mapping induced by s. We obtain a linear system of partial differential equations of first order, namely

$$a_1 \frac{\partial u_i(x)}{\partial x_1} + a_4 \frac{\partial u_i(x)}{\partial x_4} + a_5 v_i(x) = 0,$$
$$a_1 \frac{\partial v_i(x)}{\partial x_1} + a_4 \frac{\partial v_i(x)}{\partial x_4} - a_5 u_i(x) = 0 \tag{4.3}$$

for $i = 1, ..., 4$.

Let $a_1 \neq 0$. Then we find as a solution to Eq. (4.3),

$$u_i(x_1,x_4) = f_i(\eta)\cos(a_5 x_1/a_1) - g_i(\eta)\sin(a_5 x_1/a_1),$$
$$v_i(x_1,x_4) = f_i(\eta)\sin(a_5 x_1/a_1) + g_i(\eta)\cos(a_5 x_1/a_1), \tag{4.4}$$

where $\eta \equiv a_1 x_4 - a_4 x_1$. $f_i(\eta)$ and $g_i(\eta)$ are smooth functions. Now let $a_4 \neq 0$. Then we find as a solution to Eq. (2.10),

$$u_i(x_1,x_4) = f_i(\eta)\cos(a_5 x_4/a_4) - g_i(\eta)\sin(a_5 x_4/a_4),$$
$$v_i(x_1,x_4) = f_i(\eta)\sin(a_5 x_4/a_4) + g_i(\eta)\cos(a_5 x_4/a_4). \tag{4.5}$$

Now the theory tells us that when we insert the functions u_i and v_i into the nonlinear Dirac equation (2.6) we obtain a system of ordinary differential equations, where the independent variable is η. Consequently, taking into account symmetry generators we have reduced a system with two independent variables to a system with one independent variable. The quantity η is called the similarity variable.

Inserting Eq. (4.4) into Eq. (2.6) and after some algebraic manipulation we obtain the following coupled nonlinear system of ordinary differential equations:

$$\lambda\Big(-a_1 \frac{df_1}{d\eta} + a_4 \frac{df_4}{d\eta} + \frac{a_5}{a_1} g_4\Big) + g_1(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_1 \frac{dg_1}{d\eta} - a_4 \frac{dg_4}{d\eta} + \frac{a_5}{a_1} f_4\Big) + f_1(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_4 \frac{df_1}{d\eta} - a_1 \frac{df_4}{d\eta} + \frac{a_5}{a_1} g_1\Big) - g_4(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_4 \frac{dg_1}{d\eta} + a_1 \frac{dg_4}{d\eta} + \frac{a_5}{a_1} f_1\Big) - f_4(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_1 \frac{df_2}{d\eta} + a_4 \frac{df_3}{d\eta} + \frac{a_5}{a_1} g_3\Big) + g_2(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_1 \frac{dg_2}{d\eta} - a_4 \frac{dg_3}{d\eta} + \frac{a_5}{a_1} f_3\Big) + f_2(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_4 \frac{df_2}{d\eta} - a_1 \frac{df_3}{d\eta} + \frac{a_5}{a_1} g_2\Big) - g_3(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_4 \frac{dg_2}{d\eta} + a_1 \frac{dg_3}{d\eta} + \frac{a_5}{a_1} f_2\Big) - f_3(1 + \lambda\epsilon K(f,g)) = 0, \tag{4.6}$$

where

$$K(f,g) = f_1^2 + g_1^2 + f_2^2 + g_2^2 - f_3^2 - g_3^2 - f_4^2 - g_4^2.$$

Inserting Eq. (4.5) into Eq. (2.6) and after some algebraic manipulation we obtain the following coupled nonlinear system of ordinary differential equations:

$$\lambda\Big(a_1 \frac{df_1}{d\eta} - a_4 \frac{df_4}{d\eta} - \frac{a_5}{a_4} g_1\Big) - g_1(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_1 \frac{dg_1}{d\eta} - a_4 \frac{dg_4}{d\eta} + \frac{a_5}{a_4} f_1\Big) + f_1(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_4 \frac{df_1}{d\eta} + a_5 \frac{df_4}{d\eta} - \frac{a_5}{a_4} g_4\Big) + g_4(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_4 \frac{dg_1}{d\eta} + a_1 \frac{dg_4}{d\eta} + \frac{a_5}{a_4} f_4\Big) - f_4(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_2 \frac{df_2}{d\eta} - a_1 \frac{df_3}{d\eta} + \frac{a_5}{a_2} g_3\Big) - g_3(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(a_2 \frac{dg_2}{d\eta} - a_1 \frac{dg_3}{d\eta} - \frac{a_5}{a_2} f_3\Big) + f_3(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_1 \frac{df_2}{d\eta} + a_2 \frac{df_3}{d\eta} + \frac{a_5}{a_2} g_2\Big) + g_2(1 + \lambda\epsilon K(f,g)) = 0,$$
$$\lambda\Big(-a_1 \frac{dg_2}{d\eta} + a_2 \frac{dg_3}{d\eta} - \frac{a_5}{a_2} f_2\Big) - f_2(1 + \lambda\epsilon K(f,g)) = 0. \tag{4.7}$$

In the following we are only interested in localized (confined) solutions. Hence we require that for $x_1 \to \pm\infty$ the functions u_i and v_i ($i = 1,...,4$) vanish. We consider solutions of the type given by Eq. (4.5). This means we study solutions with oscillation in time. Thus we have to investigate the Eqs. (4.7). It is obvious that the general solution of this coupled system of nonlinear ordinary differential equations cannot be given explicitly. However, for particular cases we can find solutions which can be given explicitly. For example, let $f_1(\eta) = f_3(\eta) = g_2(\eta) = g_3(\eta) = 0$ and $f_4(\eta) = g_1(\eta) = 0$. Moreover, we put $a_1 = 0$. We obtain the following system of ordinary differential equations:

$$\frac{df_1}{d\eta} = -\frac{a_5}{a_4^2} g_4 + \frac{1}{a_4\lambda} g_4(1 + \lambda\epsilon(f_1^2 - g_4^2)),$$
$$\frac{dg_4}{d\eta} = \frac{a_5}{a_4^2} f_1 + \frac{1}{a_4\lambda} f_1(1 + \lambda\epsilon(f_1^2 - g_4^2)). \tag{4.8}$$

Since $a_1 = 0$, we obtain $\eta = -a_4 x_1$ and therefore

$$\frac{df_1}{dx_1} = \Big(\frac{a_5}{a_4} - \frac{1}{\lambda}\Big) g_4 - g_4\epsilon(f_1^2 - g_4^2),$$
$$\frac{dg_4}{dx_1} = \Big(-\frac{a_5}{a_4} - \frac{1}{\lambda}\Big) f_1 - f_1\epsilon(f_1^2 - g_4^2), \tag{4.9}$$

where $1/k = -a_5/a_4$. k has the dimension of a length and is a positive quantity. In the following we put $\epsilon < 0$ (attractive force) and $1/\lambda > 1/k$ since we consider confined solutions. We find as a solution to Eq. (4.9),

$$f_1(x_1) = \Big(\frac{2(1/\lambda - 1/k)}{-\epsilon}\Big)^{1/2} \frac{1}{\cosh C_2 x_1 (1 - C_1^2 \tanh^2 C_2 x_1)},$$

$$g_4(x_1) = \left(\frac{2(1/\lambda - 1/k)}{-\epsilon}\right)^{1/2} \frac{C_1 \tanh C_2 x_1}{\cosh C_2 x_1 (1 - C_1^2 \tanh^2 C_2 x_1)}, \tag{4.10}$$

where

$$C_1 = ((k - \lambda)/(k + \lambda))^{1/2}$$

and

$$C_2 = (1/\lambda^2 - 1/k^2)^{1/2}. \tag{4.11}$$

As a consequence we obtain

$$\psi^\dagger \psi = f_1^2 + g_4^2 = \frac{2(1/\lambda - 1/k)}{-\epsilon} \times \frac{1 + C_1^2 \tanh^2 C_2 x_1}{\cosh^2 C_2 x_1 (1 - C_1^2 \tanh^2 C_2 x_1)^2} \tag{4.12}$$

and

$$\bar{\psi}\psi = f_1^2 - g_4^2 = \frac{2(1/\lambda - 1/k)}{-\epsilon} \times \frac{1}{\cosh^2 C_2 x_1 (1 - C_1^2 \tanh^2 C_2 x_1)}. \tag{4.13}$$

The quantity k is determined by the equation

$$\int_{+\infty}^{-\infty} \psi^\dagger \psi \, dx_1 = 1. \tag{4.14}$$

We see that both $\psi^\dagger\psi$ and $\bar{\psi}\psi$ vanish rapidly as $x_1 \to \pm \infty$.

We have found the solutions given by Lee *et al.*,[1] where the solutions given by these authors have been written in a somewhat circumstantial manner. In order to find further solutions of Eqs. (4.7) we must solve these equations numerically.

V. CONSERVED CURRENTS

We now calculate the local conservation laws which are associated with the symmetries described by the vector fields (3.3) and (3.4). Our approach to the calculus is based on the theory of jets. The theory of jets is applied to the calculus of variations.[9–11]

Let us briefly describe the approach. Let M be an oriented manifold of dimension m, with local coordinates x_i and volume m form Ω given in these coordinates by $\Omega = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_m$. Let N be an n-dimensional manifold with local coordinates z_j and let (E, π, M) be a fiber bundle with fiber N. The k jet bundle of local sections of (E, π, M) is denoted by $J^k(E)$.

In the case with one space dimension we have $M = \mathbb{R}^2$, $N = \mathbb{R}^8$, and $(E, \pi, M) \equiv (M \times N, pr_1, M)$. Let (x_i, z_j) be a coordinate system on E and (x_i, z_j, z_{ji}) the corresponding coordinates on $J^1(E)$. The Cartan fundamental form (an n form) defined on $J^1(E)$ is given by

$$\Theta = \left(L - \sum_{j=1}^{8} \sum_{i=1}^{2} \frac{\partial L}{\partial z_{ji}} z_{ji}\right)\Omega + \sum_{j=1}^{8} \sum_{i=1}^{2} \frac{\partial L}{\partial z_{ji}} dz_j \wedge \left(\frac{\partial}{\partial x_i} \lrcorner \Omega\right), \tag{5.1}$$

where $\Omega = dx_1 \wedge dx_2$ and $L: J^1(E) \to \mathbb{R}$. In physics L is called the Lagrangian density.

In order to adopt the notation of the Secs. 2 and 3 we set $x = (x_1, x_2) = (x_1, x_4)$, $(z_1, \cdots, z_4) = (u_1, \cdots, u_4) = u$, $(z_5, \cdots, z_8) = (v_1, \cdots, v_4) = v$, $(z_{11}, \cdots, z_{44}) = (p_{11}, \cdots, p_{44}) = p$, and $(z_{51}, \cdots, z_{84}) = (q_{11}, \cdots, q_{44}) = q$.

For the nonlinear Dirac equation L takes the form

$$\begin{aligned} L = \lambda (& - u_4 q_{11} + v_4 p_{11} - u_3 q_{21} + v_3 p_{21} \\ & - u_2 q_{31} + v_2 p_{31} - u_1 q_{41} + v_1 p_{41} \\ & - u_1 q_{14} + v_1 p_{14} - u_2 q_{24} + v_2 p_{24} \\ & - u_3 q_{34} + v_3 p_{34} - u_4 q_{44} + v_4 p_{44}) - K(1 + \lambda \epsilon K), \end{aligned} \tag{5.2}$$

where

$$K(u,v) \equiv \sum_{j=1}^{2} (u_j^2 + v_j^2) - \sum_{j=3}^{4} (u_j^2 + v_j^2).$$

We mention that the exterior differential systems generated by $\{ V \lrcorner d\Theta \}$, where V denotes the vector fields on $J^1(E)$ which are vertical over M, is equivalent to the Euler–Lagrange equations for L.

A straightforward calculation yields

$$\left(L - \sum_{j=1}^{4} \left(\frac{\partial L}{\partial p_{j1}} p_{j1} + \frac{\partial L}{\partial p_{j4}} p_{j4} + \frac{\partial L}{\partial q_{j1}} q_{j1} + \frac{\partial L}{\partial q_{j4}} q_{j4}\right)\right) = -K(1 + \lambda \epsilon K), \tag{5.3}$$

and

$$\begin{aligned} & \sum_{j=1}^{4} \left(\frac{\partial L}{\partial p_{j1}} du_j \wedge dx_4 - \frac{\partial L}{\partial p_{j4}} du_j \wedge dx_1 + \frac{\partial L}{\partial q_{j1}} dv_j \wedge dx_4 - \frac{\partial L}{\partial q_{j4}} dv_j \wedge dx_1\right) \\ & = \lambda (v_4 du_1 \wedge dx_4 + v_3 du_2 \wedge dx_4 + v_2 du_3 \wedge dx_4 + v_1 du_4 \wedge dx_4 - v_1 du_1 \wedge dx_1 - v_2 du_2 \wedge dx_1 - v_3 du_3 \wedge dx_1 - v_4 du_4 \wedge dx_1 \\ & \quad - u_4 dv_1 \wedge dx_4 - u_3 dv_2 \wedge dx_4 - u_2 dv_3 \wedge dx_4 - u_1 dv_4 \wedge dx_4 + u_1 dv_1 \wedge dx_1 + u_2 dv_2 \wedge dx_1 + u_3 dv_3 \wedge dx_1 + u_4 dv_4 \wedge dx_1). \end{aligned} \tag{5.4}$$

In order to determine the conserved current we have to calculate the Lie derivative of the two-form

$$\begin{aligned} \Theta = & -K(1 + \lambda \epsilon K) dx_1 \wedge dx_4 + \lambda (v_4 du_1 - u_4 dv_1 + v_3 du_2 - u_3 dv_2 + v_2 du_3 - u_2 dv_3 + v_1 du_4 - u_1 dv_4) \wedge dx_4 \\ & + \lambda (u_1 dv_1 - v_1 du_1 + u_2 dv_2 - v_2 du_2 + u_3 dv_3 - v_3 du_3 + u_4 dv_4 - v_4 du_4) \wedge dx_1 \end{aligned} \tag{5.5}$$

with respect to the vector fields X, T, and Z given by Eqs. (3.3) and (3.4). We find

$$L_X \Theta = 0, \quad L_T \Theta = 0 \tag{5.6}$$

and

$$L_Z \Theta = 0. \tag{5.7}$$

In order to obtain the conserved currents we have to calculate the contraction $X \lrcorner \Theta$, $T \lrcorner \Theta$, and $Z \lrcorner \Theta$. We obtain

$$X\lrcorner\Theta = -K(1+\lambda\epsilon K)\,dx_4 - \lambda\sum_{i=1}^{4}(u_i\,dv_i - v_i\,du_i), \tag{5.8}$$

$$T\lrcorner\Theta = K(1+\lambda\epsilon K)\,dx_1 - \lambda\sum_{i=1}^{4}(v_{5-i}\,du_i - u_{5-i}\,dv_i), \tag{5.9}$$

$$Z\lrcorner\Theta = -2\lambda\,(u_1u_4 + u_2u_3 + v_1v_4 + v_2v_3)\,dx_4 + \lambda\,(\sum_{i=1}^{4}(u_i^2 + v_1^2))\,dx_1. \tag{5.10}$$

Consequently, the conserved currents are given by

$$-s^*(X\lrcorner\Theta) = \left[\lambda\sum_{i=1}^{4}\left(u_i(x)\frac{\partial v_i(x)}{\partial x_4} - v_i(x)\frac{\partial u_i(x)}{\partial x_4}\right)\right]dx_1$$
$$+\left[\lambda\sum_{i=1}^{4}\left(u_i(x)\frac{\partial v_i(x)}{\partial x_4} - v_i(x)\frac{\partial u_i(x)}{\partial x_4}\right) + K(u(x),v(x))(1+\lambda\epsilon K(u(x),v(x)))\right]dx_4,$$

$$-s^*(T\lrcorner\Theta) = \left[\lambda\sum_{i=1}^{4}\left(v_{5-i}(x)\frac{\partial u_i(x)}{\partial x_1} - u_{5-i}(x)\frac{\partial v_i(x)}{\partial x_1}\right) - K(u(x),v(x))(1+\lambda\epsilon K(u(x),v(x)))\right]dx_1$$
$$+\left[\lambda\sum_{i=1}^{4}\left(v_{5-i}(x)\frac{\partial u_i(x)}{\partial x_4} - u_{5-i}(x)\frac{\partial v_i(x)}{\partial x_4}\right)\right]dx_4,$$

and

$$-s^*(Z\lrcorner\Theta) = -\lambda\left[\sum_{i=1}^{4}(u_i^2(x) + v_i^2(x))\right]dx_1 + 2\lambda\,[u_1(x)u_4(x) + u_2(x)u_3(x) + v_1(x)v_4(x) + v_2(x)v_3(x)]\,dx_4,$$

where s is the following mapping: $s(x) = (x,u(x),v(x),\partial u(x)/\partial x,\partial v(x)/\partial x)$; s^* is the pull back mapping induced by s. Notice that $s^*d(\cdot)\equiv ds^*(\cdot)$. We are interested in solutions $u_i(x),v_i(x)(i = 1,2,3,4)$ which vanish at infinity, i.e., we are looking for localized solutions. Then we have

$$\frac{d}{dx_4}\int_{-\infty}^{+\infty}\sum_{i=1}^{4}\left(u_i(x)\frac{\partial v_i(x)}{\partial x_4} - v_i(x)\frac{\partial u_i(x)}{\partial x_4}\right)dx_1 = 0, \tag{5.11}$$

$$\frac{d}{dx_4}\int_{-\infty}^{+\infty}\left[\lambda\sum_{i=1}^{4}\left(v_{5-i}(x)\frac{\partial u_i(x)}{\partial x_1} - u_{5-i}(x)\frac{\partial v_i(x)}{\partial x_1}\right)\right.$$
$$\left. - K(u(x),v(x))(1+\lambda\epsilon K(u(x),v(x)))\right]dx_1 = 0, \tag{5.12}$$

$$\frac{d}{dx_4}\int_{-\infty}^{+\infty}\left[\sum_{i=1}^{4}(u_i^2(x) + v_i^2(x))\right]dx_1 = 0. \tag{5.13}$$

Instead of considering the vector fields X and T for obtaining the conserved currents we can also find the conserved currents with the help of the vector fields U and V which are given by Eq. (3.5). By a straightforward calculation we find

$$L_U\Theta = \sum_{i=1}^{4}p_{i1}\left[\frac{\partial}{\partial u_i}(K(1+\lambda\epsilon K))\right]dx_1\wedge dx_4$$
$$+\sum_{i=1}^{4}q_{i1}\left[\frac{\partial}{\partial v_i}(K(1+\lambda\epsilon K))\right]dx_1\wedge dx_4$$
$$+\lambda\sum_{i=1}^{4}(v_i\,dp_{i1}\wedge dx_1 - p_{i1}\,dv_i\wedge dx_1)$$
$$+\lambda\sum_{i=1}^{4}(p_{i1}dv_{5-i}\wedge dx_4 - v_{5-i}dp_{i1}\wedge dx_4)$$
$$+\lambda\sum_{i=1}^{4}(-u_i dq_{i1}\wedge dx_1 + q_{i1}du_i\wedge dx_1)$$
$$+\lambda\sum_{i=1}^{4}(-q_{i1}du_{5-i}\wedge dx_4 + u_{5-i}dq_{i1}\wedge dx_4).$$

Owing to the identities

$$\sum_{i=1}^{4}(\frac{\partial}{\partial u_i}(K(1+\lambda\epsilon K)))\alpha_i\wedge dx_4 \equiv d(K(1+\lambda\epsilon K)dx_4)$$
$$-\sum_{i=1}^{4}p_{i1}\left(\frac{\partial}{\partial u_i}K(1+\lambda\epsilon K)\right)dx_1\wedge dx_4,$$
$$p_{i1}dx_1 \equiv -\alpha_i + du_i - p_{i4}\,dx_4,$$
$$dp_{i1}\wedge dx_1 \equiv -d\alpha - dp_{i4}\wedge dx_4,$$
$$du_i\wedge dx_1 \equiv \alpha_i\wedge dx_1 - p_{i4}dx_1\wedge dx_4,$$

and so on, where α_i is given by Eq. (3.2), we obtain

$$L_U\Theta = d\chi \text{ modulo (contact forms } \alpha_i,\beta_i),$$

where

$$\chi = K(1+\lambda\epsilon K)\,dx_4$$
$$+\sum_{i=1}^{4}(-u_iq_{i1}\,dx_1 + v_ip_{i1}\,dx_1$$
$$+u_{5-i}q_{i1}\,dx_4 - v_{5-i}p_{i1}\,dx_4)$$
$$+\lambda\sum_{i=1}^{4}(u_i\,dv_i - v_i\,du_i).$$

The conserved current is given by $\chi - U\lrcorner\Theta$ and we find the result given by Eq. (5.8).

In the following we show that conserved currents can also be obtained without the knowledge of the two-form Θ and therefore without the knowledge of the Lagrangian density L. The approach is as follows (compare also Estabrook and Wahlquist[12]): We make the ansatz

$$\omega = f_1(u,v)\,dx_1 + f_2(u,v)\,dx_4, \tag{5.14}$$

where f_1 and f_2 are smooth functions. Let

$$J \equiv \langle F_1,\ldots,F_8,\ dF_1,\ldots,dF_8,\ \alpha_1,\ldots,\beta_4,\ d\alpha_1,\ldots,d\beta_4\rangle \tag{5.15}$$

denote the ideal generated by $F_1,\ldots,d\beta_4$. $F_1,\ldots,d\beta_4$ are given by Eq. (2.7). The condition for obtaining the conservation laws is as follows: If

$$d\omega\in\langle F_1,\ldots,F_8,\ dF_1,\ldots,dF_8,\ \alpha_1,\ldots,\beta_4,\ d\alpha_1,\ldots,d\beta_4\rangle, \tag{5.16}$$

then

$$s^*d\omega = 0, \tag{5.17}$$

where s is the mapping $s(x) = (x, u(x), v(x), p(x), q(x))$. We do not describe a general treatment, but we show that the one-form given by Eq. (5.10) can be derived. For this purpose let us construct a convenient two-form which is an element of the ideal generated by $F_1, \ldots, d\beta_4$.

Let

$$\begin{aligned}
\sigma_1 &= \beta_1 \wedge dx_1 - \beta_4 \wedge dx_4,\\
\sigma_2 &= \beta_2 \wedge dx_1 - \beta_3 \wedge dx_4,\\
\sigma_3 &= -\beta_3 \wedge dx_1 + \beta_2 \wedge dx_4,\\
\sigma_4 &= -\beta_4 \wedge dx_1 + \beta_1 \wedge dx_4,\\
\sigma_5 &= -\alpha_1 \wedge dx_1 + \alpha_4 \wedge dx_4,\\
\sigma_6 &= -\alpha_2 \wedge dx_1 + \alpha_3 \wedge dx_4,\\
\sigma_7 &= \alpha_3 \wedge dx_1 - \alpha_2 \wedge dx_4,\\
\sigma_8 &= \alpha_4 \wedge dx_1 - \alpha_1 \wedge dx_4.
\end{aligned} \tag{5.18}$$

It is obvious that the two-forms $\sigma_i (i = 1, \ldots, 8)$ are elements of the ideal. Since both f_1 and f_2 do not depend on p_{ji} and q_{ji} we have to eliminate the terms which contain p_{ji} and q_{ji}. Therefore, we consider the two-forms $\tau_j \equiv \lambda\sigma_j - F_j\, dx_1 \wedge dx_4$ which are elements of the ideal. We find

$$\begin{aligned}
\tau_1 &= \lambda\,(dv_1 \wedge dx_1 - dv_4 \wedge dx_4)\\
&\quad - u_1(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_2 &= \lambda\,(dv_2 \wedge dx_1 - dv_3 \wedge dx_4)\\
&\quad - u_2(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_3 &= \lambda\,(-dv_3 \wedge dx_1 + dv_2 \wedge dx_4)\\
&\quad - u_3(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_4 &= \lambda\,(-dv_4 \wedge dx_1 + dv_1 \wedge dx_4)\\
&\quad - u_4(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_5 &= \lambda\,(-du_1 \wedge dx_1 + du_4 \wedge dx_4)\\
&\quad - v_1(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_6 &= \lambda\,(-du_2 \wedge dx_1 + du_3 \wedge dx_4)\\
&\quad - v_2(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_7 &= \lambda\,(du_3 \wedge dx_1 - du_2 \wedge dx_4)\\
&\quad - v_3(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4,\\
\tau_8 &= \lambda\,(du_4 \wedge dx_1 - du_1 \wedge dx_4)\\
&\quad - v_4(1 + \lambda\epsilon K(u,v))\, dx_1 \wedge dx_4.
\end{aligned} \tag{5.19}$$

We notice that the conditions $s^*\tau_j = 0\ (j = 1, \ldots, 8)$ lead to the Eqs. (2.6). It is worthwhile to mention that in the approach described by Harrison and Estabrook[6] the two-forms given by Eq. (5.19) are the starting point for investigating the symmetries.

We consider now the two-form

$$\begin{aligned}
\tau = \lambda\,(&v_1\tau_1 + v_2\tau_2 - v_3\tau_3 - v_4\tau_4 - u_1\tau_5 - u_2\tau_6\\
&+ u_3\tau_7 + u_4\tau_8)
\end{aligned} \tag{5.20}$$

which is again an element of the ideal. It follows that

$$\begin{aligned}
\tau &= 2\lambda\left(\sum_{i=1}^{4} (v_i\, dv_i + u_i du_i) \wedge dx_1\right) \qquad (5.21)\\
&\quad + 2\lambda\,(-v_1\, dv_4 - v_2\, dv_3 - v_3\, dv_2 - v_4\, dv_1\\
&\quad - u_1\, du_4 - u_2\, du_3 - u_3\, du_2 - u_4\, du_4) \wedge dx_4.
\end{aligned}$$

Now the two-form τ can be represented as the exterior derivative of the one-form ω, i.e., $\tau = d\omega$, where

$$f_1(u,v) = \lambda \sum_{i=1}^{4} (u_i^2 + v_i^2), \tag{5.22}$$

$$f_2(u,v) = -2\lambda\,(u_1u_4 + u_2u_3 + v_1v_4 + v_2v_3). \tag{5.23}$$

Finally, we mention that the two-form Θ can be expressed with the help of the two-forms $\tau_i (i = 1, \ldots, 8)$, namely

$$\begin{aligned}
\Theta = &-u_1\tau_1 - u_2\tau_2 + u_3\tau_3 + u_4\tau_4 - v_1\tau_5\\
&- v_2\tau_6 + v_3\tau_7 + v_4\tau_8.
\end{aligned} \tag{5.24}$$

It follows that $\Theta \in \langle F_1, \ldots, d\beta_4 \rangle$, but $d\Theta \neq 0$ and therefore Θ cannot be obtained as the exterior derivative of a one-form. For field equations which can be derived from a Lagrangian density it is obvious that $\Theta \in \langle F_1, \ldots, d\beta_4 \rangle$.

Thus far we have derived the conserved current given by Eq. (5.10) applying two approaches. In the first one we have taken into account the Cartan form Θ which contains the Lagrangian density and the symmetry generator Z. In the second approach we have only considered the differential forms which are equivalent to the nonlinear Dirac equation. Now we describe a third approach for obtaining the conserved current given by Eq. (5.10), where we take into account the symmetry generator Z and the differential forms which are equivalent to the nonlinear Dirac equation. Hence we consider the differential forms given by Eq. (5.19).

First of all let us consider the two-form

$$\chi = \sum_{j=1}^{8} h_j(u,v)\tau_j, \tag{5.25}$$

where the two-forms τ_j are given by Eq. (5.25). h_j are smooth functions. Let Z be the symmetry generator given by Eq. (3.4) and let $\bar{J}$ be the ideal given by Eq. (5.19).

If $Z \rfloor d\chi \in \bar{J}$, then taking into account

$$L_Z(\bar{J}) \subset \bar{J} \tag{5.26}$$

and the identity

$$L_Z\chi \equiv Z \rfloor d\chi + d\,(Z \rfloor \chi) \tag{5.27}$$

it follows that the one-form $Z \rfloor \chi$ is conserved current since $d\,(Z \rfloor \chi) \in \bar{J}$. Now we have to determine the unknown functions $h_j(u,v)$ with the help of the equation $Z \rfloor d\chi \in \bar{J}$. This condition gives the solution

$$\begin{aligned}
h_1(u,v) &= u_1, & h_2(u,v) &= u_2,\\
h_3(u,v) &= -u_3, & h_4(u,v) &= -u_4,\\
h_5(u,v) &= v_1, & h_6(u,v) &= v_2,\\
h_7(u,v) &= -v_3, & h_8(u,v) &= -v_4.
\end{aligned} \tag{5.28}$$

Consequently, we find $\chi = \Theta$, where Θ is given by Eq. (5.24).

A comment about the different approaches is in order. For obtaining conservation laws due to Noether's generalized theorem[11] we use the whole contact ideal, including all differential consequences of the given field equation. If the field equation admits a Lagrangian formulation there exists a well-defined correspondence between symmetries and conservation laws.[11] Therefore Noether's generalized theorem is superior to other approaches for finding conservation laws of field equation in Lagrangian form. The approach due to Estabrook and Wahlquist[12] only uses a subideal $\bar{J}$ of the contact ideal J. Therefore this method requires fewer variables.

The reason for this is that this approach tends to determine the conservation $\omega = f_1\, dx_1 + f_2\, dx_4$ from the condition $d\omega \in \bar{J}$. This condition leads to a differential equation for the quantities f_1 and f_2. This differential equation which is highly overdetermined looks similar to the determining equations for the symmetries. The same holds true for our third approach which is also based upon the condition $d\omega \in \bar{J}$. However, we take into account known symmetries of the field equation under consideration.

The merit of the second and third approach is that they are applicable to the case where the field equation under consideration admits no Lagrangian formulation. To see this let us consider an example. For the diffusion equation

$$u_t = u_{xx} + u$$

there is no Lagrangian density. Now we determine conservation laws. First of all we put $u_1 = u$, $u_2 = u_x$, $x_1 = x$, and $x_2 = t$. Let us express the diffusion equation by two-forms

$$\tau_1 = du_1 \wedge dx_2 - u_2 dx_1 \wedge dx_2,$$
$$\tau_2 = du_1 \wedge dx_1 + du_2 \wedge dx_2 + u_1 dx_1 \wedge dx_2.$$

We make the ansatz

$$\omega = f_1(x_1,x_2,u_1,u_2)\, dx_1 + f_2(x_1,x_2,u_1,u_2)\, dx_2$$

which is an extension of the ansatz given by Eq. (5.14). We obtain with the technique described above that

$$\omega = (-\alpha \sin x_1 + \beta \cos x_1)u_1\, dx_1 + ((\alpha \cos x_1 + \beta \sin x_1)u_1 + (-\alpha \sin x_1 + \beta \cos x_1)u_2)\, dx_2$$

is a conservation law of $u_t = u_{xx} + u$, where α and β are two arbitrary constants. Taking into account the invariance of the diffusion equation under the infinitesimal generator

$$-2x_2 \frac{\partial}{\partial x_1} + x_1 u_1 \frac{\partial}{\partial u_1},$$

we obtain by the third approach that

$$\omega = x_1 \exp(-x_2)u_1\, dx_1 + (x_1 \exp(-x_2) - 1)u_2\, dx_2$$

is a conservation law.

VI. FOUR-DIMENSIONAL SPACE–TIME

In this section we show in four-dimensional space–time how solutions can be obtained using continuous symmetry. We consider a particular case, demonstrating how cylindrically symmetric solutions can be derived. Such solutions have been considered by Takahashi.[4] Now we study the nonlinear Dirac equation with scalar coupling, namely

$$\sum_{k=1}^{3} \lambda \frac{\partial}{\partial x_k}(\gamma_k \psi) - i\lambda \frac{\partial}{\partial x_4}(\gamma_4 \psi) + \psi(1 + \epsilon\lambda^3 \bar{\psi}\psi) = 0. \tag{6.1}$$

Again we put $\psi_j(x) = u_j(x) + iv_j(x)$, where $u_j(x)$ and $v_j(x)$ are real fields. In the same manner as in Sec. 2, we cast the system of partial differential equations into an equivalent system of differential forms. To investigate the continuous symmetries we study the Lie derivative of the differential forms with respect to infinitesimal generators.

Let us consider now cylindrically symmetric solutions whose axis lies along the x_3 direction. We recall that the rotation group on the plane (x_1,x_2) has the form $-x_1\partial/\partial x_2 + x_2\partial/\partial x_1$. Now we consider the infinitesimal generator

$$R = -x_1 \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1} + u_1 \frac{\partial}{\partial v_1} - v_1 \frac{\partial}{\partial u_1}. \tag{6.2}$$

We calculate the once-extended infinitesimal generator and require that the zero-forms $F_1,\dots,F_8$ are invariant under this once-extended infinitesimal generator. This requirement leads to the condition that $u_3(x) = v_3(x) = 0$. Consequently, the derivatives of these fields with respect to $x_1,\dots,x_4$ also vanish. Then the zero-forms $F_1,\dots,F_8$ take the form

$$\begin{aligned}
F_1 &\equiv \lambda(-p_{41} - q_{42} - p_{14}) + v_1(1 + \lambda^3\epsilon K) = 0,\\
F_2 &\equiv \lambda(p_{43} - p_{24}) + v_2(1 + \lambda^3\epsilon K) = 0,\\
F_3 &\equiv \lambda(p_{13} - p_{21} + q_{22}) = 0,\\
F_4 &\equiv \lambda(p_{11} - q_{12} - p_{23} + p_{44}) + v_4(1 + \lambda^3\epsilon K) = 0,\\
F_5 &\equiv \lambda(q_{14} - q_{41} + p_{42}) - u_1(1 + \lambda^3\epsilon K) = 0,\\
F_6 &\equiv \lambda(q_{43} - q_{24}) - u_2(1 + \lambda^3\epsilon K) = 0,\\
F_7 &\equiv \lambda(q_{21} - p_{22} + q_{13}) = 0,\\
F_8 &\equiv \lambda(q_{11} + p_{12} - q_{23} + q_{44}) - u_4(1 + \lambda^3\epsilon K) = 0,
\end{aligned} \tag{6.3}$$

where

$$K = u_1^2 + v_1^2 + u_2^2 + v_2^2 - u_4^2 - v_4^2. \tag{6.4}$$

Let $\bar{R}$ denote the once-extended infinitesimal generator of R. We find

$$L_{\bar{R}}F_1 = -F_5, \quad L_{\bar{R}}F_2 = 0, \quad L_{\bar{R}}F_3 = -F_7,$$
$$L_{\bar{R}}F_4 = 0, \tag{6.5}$$
$$L_{\bar{R}}F_5 = F_1, \quad L_{\bar{R}}F_6 = 0,$$
$$L_{\bar{R}}F_7 = F_3, L_{\bar{R}}F_8 = 0. \tag{6.6}$$

Since we study cylindrically symmetric solutions whose axis lies along the x_3 direction the fields do not depend on x_3. Thus in Eq. (6.3) the terms p_{i3} and q_{i3} can be omitted. Moreover, we assume that $u_2(x) = 0$ and $v_2(x) = 0$. This means we require that the differential forms are invariant under $\partial/\partial u_2$. Then the zero-forms given by Eq. (6.3) take the form

$$\begin{aligned}
F_1 &\equiv \lambda(-p_{41} - q_{42} - p_{14}) + v_1(1 + \lambda^3\epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2))\\
&= 0,\\
F_4 &\equiv \lambda(p_{11} - q_{12} + p_{44}) + v_4(1 + \lambda^3\epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2))\\
&= 0,\\
F_5 &\equiv \lambda(-q_{14} - q_{41} + p_{42}) - u_1(1 + \lambda^3\epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2))\\
&= 0,\\
F_8 &\equiv \lambda(q_{11} + p_{12} + q_{44}) - u_4(1 + \lambda^3\epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2))\\
&= 0.
\end{aligned} \tag{6.7}$$

The conditions $(j = 1,4)$

$$s^*(R \lrcorner \alpha_j) = 0,$$
$$s^*(R \lrcorner \beta_j) = 0, \tag{6.8}$$

lead to the following linear partial differential equations $[x = (x_1,x_2,x_4)]$:

$$x_1 \frac{\partial u_1(x)}{\partial x_2} - x_2 \frac{\partial u_1(x)}{\partial x_1} - v_1(x) = 0,$$
$$x_1 \frac{\partial v_1(x)}{\partial x_2} - x_2 \frac{\partial v_1(x)}{\partial x_1} + u_1(x) = 0,$$

$$x_1 \frac{\partial u_4(x)}{\partial x_2} - x_2 \frac{\partial u_4(x)}{\partial x_1} = 0,$$
$$x_1 \frac{\partial v_4(x)}{\partial x_2} - x_2 \frac{\partial v_4(x)}{\partial x_1} = 0. \tag{6.9}$$

The solution to the linear partial differential equations is given by

$$\begin{aligned} u_1(x_1,x_2,x_4) &= \bar{u}_1(\rho,x_4)x_1 + \bar{v}_1(\rho,x_4)x_2, \\ v_1(x_1,x_2,x_4) &= \bar{v}_1(\rho,x_4)x_1 - \bar{u}_1(\rho,x_4)x_2, \\ u_4(x_1,x_2,x_4) &= \bar{u}_4(\rho,x_4), \\ v_4(x_1,x_2,x_4) &= \bar{v}_4(\rho,x_4), \end{aligned} \tag{6.10}$$

where $\rho^2 \equiv x_1^2 + x_2^2$ is the similarity variable. When we insert the Eq. (6.10) into the nonlinear Dirac equation we obtain a nonlinear system of differential equations with two independent variables, namely ρ and x_4. The dependent variables are $\bar{u}_1(\rho,x_4)$, $\bar{v}_1(\rho,x_4)$, and so on.

Let us now further reduce the nonlinear Dirac equation. We use the fact that the Dirac equation is invariant under the transformation group which is associated with the infinitesimal generator

$$Z = a_4 \frac{\partial}{\partial x_4} - u_1 \frac{\partial}{\partial v_1} + v_1 \frac{\partial}{\partial u_1} - u_4 \frac{\partial}{\partial v_4} + v_4 \frac{\partial}{\partial u_4}, \tag{6.11}$$

where a_4 is a constant which has the dimension of a length ($a_4 \neq 0$). The conditions $s^*(Z \rfloor \alpha_j) = 0$ and $s^*(Z \rfloor \beta_j) = 0$ ($j \in \{1,4\}$) lead to the following linear system partial differential equations:

$$\begin{aligned} a_4 \frac{\partial u_1(x)}{\partial x_4} - v_1(x) &= 0, \\ a_4 \frac{\partial v_1(x)}{\partial x_4} + u_1(x) &= 0, \\ a_4 \frac{\partial u_4(x)}{\partial x_4} - v_4(x) &= 0, \\ a_4 \frac{\partial v_4(x)}{\partial x_4} + u_4(x) &= 0. \end{aligned} \tag{6.12}$$

The solution to the partial differential equations is given by

$$\begin{aligned} u_1(x_1,x_2,x_4) &= \bar{u}_1(x_1,x_2)\cos(x_4/a_4) + \bar{v}_1(x_1,x_2)\sin(x_4/a_4), \\ v_1(x_1,x_2,x_4) &= \bar{v}_1(x_1,x_2)\cos(x_4/a_4) - \bar{u}_1(x_1,x_2)\sin(x_4/a_4), \\ u_4(x_1,x_2,x_4) &= \bar{u}_4(x_1,x_2)\cos(x_4/a_4) + \bar{v}_4(x_1,x_2)\sin(x_4/a_4), \\ v_4(x_1,x_2,x_4) &= \bar{v}_4(x_1,x_2)\cos(x_4/a_4) - \bar{u}_4(x_1,x_2)\sin(x_4/a_4). \end{aligned} \tag{6.13}$$

Now Eq. (6.10) and Eq. (6.13) can be combined. We find that

$$\begin{aligned} u_1(x_1,x_2,x_4) &= \bar{u}_1(\rho)(x_1 \cos(x_4/a_4) - x_2 \sin(x_4/a_4) \\ &\quad + \bar{v}_1(\rho)(x_2 \cos(x_4/a_4) + x_1 \sin(x_4/a_4), \\ v_1(x_1,x_2,x_4) &= \bar{v}_1(\rho)(x_1 \cos(x_4/a_4) - x_2 \sin(x_4/a_4) \\ &\quad - \bar{u}_1(\rho)(x_2 \cos(x_4/a_4) + x_1 \sin(x_4/a_4), \\ u_4(x_1,x_2,x_4) &= \bar{u}_4(\rho) \cos(x_4/a_4) + \bar{v}_4(\rho) \sin(x_4/a_4), \\ v_4(x_1,x_2,x_4) &= \bar{v}_4(\rho) \cos(x_4/a_4) - \bar{u}_4(\rho) \sin(x_4/a_4). \end{aligned} \tag{6.14}$$

Thus, the similarity form of the solution can be written as a sum of products, where in each term the time coordinate is separated from the space variables. Thus with group theoretical methods we have derived the separation ansatz (6.14). Inserting Eq. (6.14) into the nonlinear Dirac equation [compare Eq. (6.3)] we obtain a nonlinear system of ordinary differential equations where the independent variable is the quantity ρ. By a straightforward calculation we find

$$\begin{aligned} &\lambda \left(\rho \frac{d\bar{u}_1}{d\rho} + 2\bar{u}_1 + \frac{\bar{v}_4}{a_4} \right) + \bar{v}_4 [1 + \lambda^3 \epsilon(\bar{u}_1^2 + \bar{v}_1^2 - \bar{u}_4^2 - \bar{v}_4^2)] \\ &\quad = 0, \\ &\lambda \left(\frac{d\bar{v}_4}{d\rho} - \frac{\rho}{a_4} \bar{u}_1 \right) + \rho \bar{u}_1 [1 + \lambda^3 \epsilon(\bar{u}_1^2 + \bar{v}_1^2 - \bar{u}_4^2 - \bar{v}_4^2)] \\ &\quad = 0, \\ &\lambda \left(\rho \frac{d\bar{v}_1}{d\rho} + 2\bar{v}_1 - \frac{\bar{u}_4}{a_4} \right) - \bar{u}_4 [1 + \lambda^3 \epsilon(\bar{u}_1^2 + \bar{v}_1^2 - \bar{u}_4^2 - \bar{v}_4^2)] \\ &\quad = 0, \\ &\lambda \left(\frac{d\bar{u}_4}{d\rho} + \frac{\rho}{a_4} \bar{v}_1 \right) - \rho \bar{v}_1 [1 + \lambda^3 \epsilon(\bar{u}_1^2 + \bar{v}_1^2 - \bar{u}_4^2 - \bar{v}_4^2)] \\ &\quad = 0. \end{aligned} \tag{6.15}$$

If we put $\bar{u}_1(\rho) = \bar{v}_4(\rho) = 0$ or $\bar{v}_1(\rho) = \bar{u}_4(\rho) = 0$, then we obtain a system of two coupled differential equations (compare also Takahashi[4]).

Finally we mention that we can derive conserved currents with the help of the symmetry generators given by Eqs. (6.2) and (6.11). In the present case we have

$$\begin{aligned} L = \lambda (&- u_4 q_{11} + v_4 p_{11} - u_1 q_{41} + v_1 p_{41} \\ &- u_4 p_{12} - v_4 q_{12} + u_1 p_{42} + v_1 q_{42} \\ &- u_1 q_{14} + v_1 p_{14} - u_4 q_{44} + v_4 p_{44}) \\ &- (u_1^2 + v_1^2 - u_4^2 - v_4^2)(1 + \lambda^3 \epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2). \end{aligned} \tag{6.16}$$

Consequently, the Hamilton–Cartan form is given by

$$\begin{aligned} \Theta = &-(u_1^2 + v_1^2 - u_4^2 - v_4^2)(1 + \lambda^3 \epsilon(u_1^2 + v_1^2 - u_4^2 - v_4^2)) \\ &\times dx_1 \wedge dx_2 \wedge dx_4 \\ &+ \lambda (v_4 \, du_1 + v_1 \, du_4 - u_4 \, dv_1 - u_1 \, dv_4) \wedge dx_2 \wedge dx_4 \\ &+ \lambda (u_4 \, du_1 - u_1 \, du_4 + v_4 \, dv_1 - v_1 \, dv_4) \wedge dx_1 \wedge dx_4 \\ &+ \lambda (v_1 \, du_1 + v_4 \, du_4 - u_1 \, dv_1 - u_4 \, dv_4) \wedge dx_1 \wedge dx_2. \end{aligned} \tag{6.17}$$

We find

$$L_R \Theta = 0$$

and

$$L_Z \Theta = 0. \tag{6.18}$$

We notice that the infinitesimal generators R and Z form a basis of an abelian Lie algebra.

VII. CONCLUSION

We have demonstrated for a nonlinear Dirac equation how the knowledge of infinitesimal symmetry generators can be used for deriving similarity solutions and conserved currents. With the help of infinitesimal symmetry generators we are able to reduce the system of nonlinear partial differential equations to a nonlinear system of ordinary differential equations. For particular cases we are able to find solutions which can be given explicitly. Moreover, we have described three different approaches for obtaining conserved currents.

In the first one the infinitesimal symmetry generator and the Hamilton–Cartan form has been taken into account. In the second, the starting point was only those differential forms which are equivalent to the nonlinear Dirac equation. In this approach the Hamilton–Cartan form (which in the present modern formulation plays the role of the Lagrangian density) and the infinitesimal symmetry generators are not taken into account. Finally, in the third approach, we have used the infinitesimal symmetry generators and the differential forms which are equivalent to the nonlinear Dirac equation. The Hamilton–Cartan form is not used in this approach either.

[1] S. Y. Lee, T. K. Kuo, and A. Gavrielides, Phys. Rev. D **12**, 2249 (1975).
[2] S. J. Chang, S. D. Ellis, and B. W. Lee, Phys. Rev. D **11**, 3572 (1975).
[3] K. Takahashi, Prog. Theor. Phys. **61**, 1251 (1979).
[4] K. Takahashi, J. Math. Phys. **20**, 1232 (1979).
[5] J. Dieudonné, *Treatise on Analysis* (Academic, New York, 1974), Chap. 18.
[6] B. K. Harrison and F. B. Estabrook, J. Math. Phys. **12**, 653 (1971).
[7] W.-H. Steeb, J. Math. Phys. **21**, 1656 (1980).
[8] W.-H. Steeb, W. Erig, and W. Strampp, J. Math. Phys. **22**, 970 (1981).
[9] H. Goldschmidt and S. Sternberg, Ann. Inst. Fourier (Grenoble) **23**, 203 (1973).
[10] D. Krupka, Czechoslovak Math. J. **27**, 114 (1977).
[11] W. F. Shadwick. Lett. Math. Phys. **4**, 241 (1980).
[12] F. B. Estabrook and H. Wahlquist, J. Math. Phys. **16**, 1 (1975).

T matrix for Coulomb-nuclear admixtures

B. Talukdar, N. Mallick, M. Kundu, and R. N. Panigrahi
Department of Physics, Visva-Bharati University, Santiniketan-731235, West Bengal, India

(Received 6 August 1980; accepted for publication 24 October 1980)

A model is proposed for calculating the T matrix for Coulomb-distorted separable nuclear potentials without the use of the Gell-Mann–Goldberger two-potential theorem. Analytical expressions for the Jost function and T matrix are presented for the Yamaguchi plus screened Coulomb potential treated in the Ecker–Weizel approximation.

PACS numbers: 21.30. + y

The Gell-Mann–Goldberger (GG) two-potential theorem[1] has been extensively used to construct expressions for Coulomb-distorted nuclear T matrices. Such studies are often tailored to be appropriate for their possible usuage in nuclear few-body problems, as, for example, in the work of van Haeringen and van Wageningen[2] and references therein. Applicability of the two-potential formula is directly related to the existence and/or completeness of wave operators for the scattering system.[3] The presence of long-range forces in addition to the nuclear potential tends to pose serious problems with respect to this.[4] With this in mind, we present a coordinate-space approach to the Coulomb-nuclear problem, which does not make explicit use of the GG theorem. In general, the method proposed will work for Coulomb plus separable potentials of arbitrary rank. However, for a model calculation we choose to work with the following interaction.

(i) The Yamaguchi potential[5] is used to represent the basic nuclear interaction.

(ii) The pure Coulomb field is replaced by a screened potential, which we treat in the Ecker–Weizel approximation.[6]

Consider the s-wave Schrödinger equation

$$\left[\frac{d^2}{dr^2}+k^2+V_0\frac{e^{-\mu r}}{r}\mu r\right]f(k,r)=d(k)e^{-\alpha r}, \tag{1}$$

where

$$d(k)=\lambda\int_0^\infty e^{-\alpha s}f(k,s)\,ds. \tag{2}$$

Here V_0 and μ are related to the coupling constant and screening parameter for the Coulomb field, and λ and α to the strength and range of the nuclear interaction. To solve Eq. (1) with Jost boundary conditions[7] we employ the standard substitutions

$$f(k,r)=e^{ikr}\vartheta(r)\quad\text{and}\quad \mu r=-\ln y \tag{3}$$

and arrive at a nonhomogeneous hypergeometric differential equation

$$y(1-y)\frac{d^2\vartheta}{dy^2}+[c-(a+b+1)y]\frac{d\vartheta}{dy}-ab\vartheta=\frac{d(k)}{\mu^2}[y^{\sigma-1}-y^\sigma] \tag{4}$$

studied by Babister.[8] The parameters a, b, c, and σ are given by

$$\begin{aligned}a&=-(i/\mu)[k-(k^2-\gamma\mu^2)^{1/2}],\\ b&=-(i/\mu)[k+(k^2-\gamma\mu^2)^{1/2}],\\ c&=1-2ik/\mu\quad\text{and}\quad\sigma=(\alpha+ik)/\mu.\end{aligned} \tag{5}$$

In writing Eqs. (4) and (5) we have treated

$$-\gamma=V_0(1-y)/\ln y \tag{6}$$

as a constant in the spirit of the Ecker–Weizel approximation, certain features of which have been discussed by Lam and Varshni.[9]

The Jost solution associated with the complementary function of Eq. (4) has been obtained by one of us.[10,11] Following a suggestion given elsewhere[12,13] we regard $d(k)$ as a constant and obtain a particular integral in the form[8]

$$\vartheta^P(y)=\frac{d(k)}{\alpha^2+k^2}[y^\sigma-\gamma f_{\sigma+1}(a,b;c;y)]. \tag{7}$$

The relation between f_σ and generalized hypergeometric function is given in Eq. (6.18) of Ref. 8. The complete primitive of Eq. (1) is

$$\begin{aligned}f(k,y)&=y^{-ik/\mu}[1-\gamma f_1(a,b;c;y)]\\&\quad+\frac{d(k)y^{-ik/\mu}}{\alpha^2+k^2}[y^\sigma-\gamma f_{\sigma+1}(a,b;c,y)],\end{aligned} \tag{8}$$

where the first term stands for the screened Coulomb Jost solution.[10,11] Interestingly $f(k,y)$ in Eq. (8) satisfies the Jost boundary condition since $f_\sigma(a,b;c;y)\to 0$ as $y\to 0$ $(r\to\infty)$ for $\mathrm{Re}\,\sigma>0$,[8] which is true in our case.

Our immediate concern is to determine the value of $d(k)$. To that end we combine Eqs. (2) and (8). This gives

$$d(k)=\frac{\lambda}{D(k)}\left[\frac{1}{\alpha-ik}-\frac{\gamma}{c(\alpha-ik+\mu)}\times{}_4F_3\left(\begin{matrix}1 & 1+a & 1+b & 1+\sigma^*\\ & 2 & 1+c & 2+\sigma^*\end{matrix}\middle|1\right)\right], \tag{9}$$

with

$$D(k)=1-\frac{\lambda}{2\alpha(\alpha^2+k^2)}\,\frac{\gamma\lambda\mu^2}{(2\alpha+\mu)(\alpha+ik+\mu)(\alpha+ik+c\mu)(\alpha^2+k^2)}\cdot{}_4F_3\left(\begin{matrix}1 & 1+a+\sigma & 1+b+\sigma & 1+\sigma+\sigma^*\\ & 2+\sigma & 1+c+\sigma & 2+\sigma+\sigma^*\end{matrix}\middle|1\right). \tag{10}$$

The results in Eqs. (9) and (10) have been obtained by employing an integral used in Ref. 11. The ${}_4F_3$ hypergeometric function occuring here is absolutely convergent.[14]

The Jost function[7] obtained from Eq. (8) is given by

$$f(k) = [1 - \gamma f_1(a,b;c;1)] + \frac{d(k)}{\alpha^2 + k^2}[1 - \gamma f_{\sigma+1}(a,b;c;1)]. \tag{11}$$

The function f in Eq. (11) can be written in terms of gamma functions employing Eq. (6.185) of Ref. 8 and a genereliza-tion of the Dixon's theorem of Ref. 14. A useful check on the fairly complicated formula (11) is that it yields the well-known Yamaguchi–Jost function when the Coulomb field is turned off. In possession of the Jost function $f(k)$ the on-shell T matrix can be determined from

$$T(k) = (f(k) - f(-k))/i\pi f(k). \tag{12}$$

We conclude by noting that Fuda and Whiting[15] have shown how to write the half-off-shell T matrix in terms of the off-shell Jost function. The off-shell Jost solutions for the screened Coulomb field are now available.[11] Thus in principle the half-off-shell T matrix can also be obtained by the method presented here. However, the fully off-shell case will need some further consideration. The treatment of the unscreened Coulomb field within the framework of this approach is an interesting problem with which to deal.

[1]M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91**, 398 (1953).
[2]H. van Haeringen and R. van Wageningen, J. Math. Phys. **16**, 1441 (1975).
[3]Z. Bajzer, Il Nuovo Cimento A **22**, 300 (1974).
[4]E. Prugovecki and J. Zorbas, J. Math. Phys. **14**, 1398 (1973); W. W. Zachary, J. Math. Phys. **17**, 1056 (1976); **18**, 536 (1977); A. W. Saenz and W. W. Zachary, J. Math. Phys. **19**, 1688 (1978).
[5]Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[6]G. Ecker and W. Weizel, Ann. Phys. (Leipzig) **17**, 126 (1956).
[7]R. Jost, Helv. Phys. Acta **20**, 256 (1947).
[8]A. W. Babister, *Transcendental Functions Satisfying Nonhomogeneous Linear Differential Equations* (Macmillan, New York, 1967), Chap. 6.
[9]C. S. Lam and Y. P. Varshni, Phys. Lett. A **59**, 363 (1975).
[10]B. Talukdar, R. N. Chaudhuri, U. Das, and P. Banerjee, J. Math. Phys. **19**, 1654 (1978).
[11]U. Das, S. Chakravarty, and B. Talukdar, J. Math. Phys. **20**, 887 (1979).
[12]B. Talukdar, U. Das, and S. Chakravarty, Phys. Rev. C **19**, 322 (1979).
[13]B. Talukdar and U. Das, Pramana **13**, 525 (1979).
[14]L. J. Slater, *Generalized Hypergeometric Functions* (Cambridge U. P., Cambridge, 1966), p. 45.
[15]M. G. Fuda and J. S. Whiting, Phys. Rev. C **8**, 1255 (1973).

Four Euclidean conformal group approach to the multiphoton processes in the H-atom

Jean-Pierre Gazeau [a)]

Centre de Recherche de Mathématiques Appliquées, Université de Montréal, Québec, Canada

(Received 9 October 1980; accepted for publication 26 November 1980)

A compact analytical form, suitable for any analytic continuation, is obtained for the following bound–bound N-photon transition matrix element,

$$I_{nlm\to n'l'm'} \equiv \left\langle n'l'm' \left| \prod_{i=1}^{N} \begin{Bmatrix} \vec{p}\cdot\vec{\epsilon}_i \\ \vec{r}\cdot\vec{\epsilon}_i \end{Bmatrix} e^{\vec{k}_i\cdot\vec{r}} G(E_i) \right| nlm \right\rangle,$$

where $G(E)$ is the Coulomb Green's function. We show that $I_{nlm\to n'l'm'}$ is a "linear superposition" of matrix elements $\mathscr{T}^1(g)_{n'l'm'\to nlm}$ of some irreducible representation $\mathscr{T}^1$ of a semigroup $G_{\leqslant}^{-1}$ contained in the four Euclidean conformal group $G = \mathrm{SU}^*(4) \approx \overline{\mathrm{SO}_0(1,5)}$. This "linear superposition" is understood in the general framework of the theory of the distributions on a Lie group. The final result is a linear combination of special functions known as "generalized Euler functions."

PACS numbers: 32.80.Kf, 31.15. + q, 03.65.Fd

INTRODUCTION

It has become possible over the past two decades to observe, with sufficiently large intensity of the light source, interaction processes with atoms in which each electronic transition involves the net absorption, emission, or scattering of more than one photon.[1]

But, the evaluation of transition probabilities for multiphoton processes requires tedious calculations of matrix elements issued from the time-dependent perturbation theory. In the specific case of interaction of light with the H-atom, the latter arise in the general form (for a N-photon process)

$$I_{nlm\to n'l'm'} = \langle n'l'm' | \mathscr{O} | nlm \rangle = (\psi_{n'l'm'}, \mathscr{O}\psi_{nlm})_{L^2_{\mathbb{C}}(R^3)}. \tag{1}$$

$\psi_{nlm}, \psi_{n'l'm'}$ are H-bound state wave functions, and $\mathscr{O}$ is the operator

$$\prod_{i=1}^{N} A_i e^{i\vec{k}_i\cdot\vec{r}} G(E_i),$$

where $G(E)$ is the Coulomb resolvent $(\vec{p}^2/2m - E - g/r)^{-1}$ and $A_i = \vec{p}\cdot\vec{\epsilon}_i$ or $\vec{r}\cdot\vec{\epsilon}_i$, with $\vec{\epsilon}_i\cdot\vec{k}_i = 0$.

It should be noted that the last factor $G(E_N)$ in $\mathscr{O}$ can be removed from (1) by using the evident equality

$$G(E_N)\psi_{nlm} = (E_n - E_N)^{-1}\psi_{nlm}. \tag{2}$$

However, we prefer to keep it for writing convenience.

Exact analytic expressions for the matrix element (1) have been known for a long time in the two-photon case, in the dipole approximation or with retardation effects, between bound or continuum states.[1-8] In more recent works, Coulomb Green's function[4] or Sturmian techniques[5] have been used to obtain analytical expressions for transition amplitudes of multiphoton processes for $N \geqslant 3$ in the dipole approximation.

This work generalizes to an arbitrary N a group theoretical technique already used to obtain a compact form for the two-photon matrix element, with retardation effect.[6] We show that $I_{nlm\to n'l'm'}$ is a "linear superposition" of matrix elements $\mathscr{T}^1(g)_{n'l'm',nlm}$ of some irreducible representation $\mathscr{T}^1$ of a semigroup, denoted by $G_{\leqslant}^{-1}$, contained in the four Euclidean conformal group $G = \mathrm{SU}^*(4) \approx \overline{\mathrm{SO}_0(1,5)}$

$$I_{nlm\to n'l'm'} = \int_{G_{\leqslant}^{-1}} dS(g)\, \mathscr{T}^1(g)_{n'l'm',nlm},$$

where $\int_{G_{\leqslant}^{-1}} dS(g)$ symbolizes a product of differential/integral operators involving the g variable of the special function $\mathscr{T}^1(g)_{n'l'm',nlm}$. It should be underlined that by its use of the SU*(4) group and the "distribution theory on Lie groups" outfit, our method differs from the SO(4,2) techniques used in the past by Fronsdal,[7] Huff,[8] or Barut[9] for treating the one- or two-photon problems. Rather, the claim of the present paper is to illustrate some of the practical aspects of a very general and meaningful group theoretical structure of the quantum mechanics which has been displayed in a recent work.[10] Also, we should like to point out that somewhat esoteric mathematics can sometimes play a nonnegligible part in some practical computations of Physics.

In the first section, we define the Fock transformation which permits to introduce the Hilbert space $L^2_{\mathbb{C}}(\mathrm{SU}(2))$, and we describe it with the aid of quaternionic conformal actions.

In Sec. II, we present the group $G = \mathrm{SU}^*(4)$, the subgroup $G_< = \mathrm{Sp}(1,1) \approx \overline{\mathrm{SO}_0(1,4)}$, and a semigroup $G_{\leqslant} \supset G_<$. Then, we define the representation $\mathscr{T}^1$ of $G_{\leqslant}^{-1}$.

In Sec. III, the tempered vector distribution on G, with support in the semigroup $G_{\leqslant}^{-1}$, are introduced. Useful and concrete examples of such distributions are given in Sec. IV, as also in Sec. V their connection with physical operators of interest.

Finally, in Sec. VI, the matrix element $I_{nlm\to n'l'm'}$ is explicated in terms of "generalized Euler integrals," extending the familiar hypergeometric functions occurring in the two-photon case.

[a)] Permanent address: Laboratoire de Chimie Physique, 11 rue Pierre et Marie Curie, 75005 Paris, France.

1. FOCK TRANSFORMATION[10–12]

Let p_0 be a real momentumlike parameter. In the momentum representation, a one-particle wavefunction $\hat{\psi}$, the Fourier transform of ψ, will be viewed as a function of the 4-vector $x = (p_0,\vec{p})$.

We denote by $\underline{F}_{p_0,+1}$ the space of the complex functions square integrable with respect to the scalar product

$$(\psi(p_0),\psi'(p_0))_{p_0,+1} \equiv \int d\vec{p}\, \hat{\psi}^*(p_0,\vec{p})\,[(p_0^2+\vec{p}^2)/2p_0^2\,]\hat{\psi}(p_0,\vec{p}). \tag{3}$$

$\underline{F}_{p_0,+1}$ is a Hilbert space. There exists an isomorphism $\mathscr{F}_{p_0}$, called Fock transformation, mapping $\underline{F}_{p_0,+1}$ onto $\underline{E} \equiv L^2_{\mathbb{C}}(S^3 \approx \mathrm{SU}(2))$, provided with the scalar product:

$$(\phi,\phi')_{\underline{E}} \equiv \int_{\mathrm{SU}(2)} d\mu(\xi)\,\phi^*(\xi)\phi'(\xi),$$

where $d\mu(\xi)$ is the O(4)-invariant measure on S^3 or equivalently the Haar measure on SU(2).

$\mathscr{F}_{p_0}$ is defined in the following way:

$$\mathscr{F}_{p_0}(\psi(p_0))(\xi) = 4p_0^{3/2}|\underline{1}+\xi|^{-4}\hat{\psi}(s(p_0)\cdot\xi) \tag{4}$$

where $\underline{1} = (1,\vec{0})$, $|x|$ is the Euclidean norm in R^4, and $(p_0,\vec{p}) = s(p_0)\cdot\xi$ is the stereographic projection of $\xi \in S^3$. This "Fock stereographic projection" is conveniently described by identifying the four Euclidean space with the quaternion field $\mathbb{H}$. Any quaternion x will be written $x = (x_0,\vec{x})$ where x_0 is the scalar part and $\vec{x}$ the vector part, or yet:

$$x = x_0\underline{1} + x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3,$$

with $\vec{e}_i\vec{e}_j = \vec{e}_k$, (i,j,k) being an even permutation of (1,2,3).

The conjugate $\bar{x}$ is defined by $\bar{x} = (x_0, -\vec{x})$. Any scalar λ will be confounded with the quaternion $\underline{\lambda} = (\lambda,\vec{0})$.

Let $g = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$ be a 2×2 quaternionic matrix. Its conformal action on $\mathbb{H}$ is given by

$$g\cdot x = (ax+b)(cx+d)^{-1}. \tag{5}$$

The projection $s(p_0)$ is the special singular conformal action

$$s(p_0) = \frac{1}{\sqrt{2p_0}}\begin{pmatrix} 2p_0 & \underline{0} \\ \underline{1} & \underline{1}\end{pmatrix}. \tag{6}$$

It establishes a one-to-one correspondence between $S^3 \approx \mathrm{SU}(2)$ and the compactified hyperplane of the quaternions having the same scalar part p_0.

Our purpose is to compute (1) by using the transformation $\mathscr{F}_{p_0}$ with different values of p_0 and a certain representation of the conformal action (5) which will now be explicated.

2. A SEMIGROUP REPRESENTATION

$G \equiv \mathrm{SU}^*(4)$ is the simple, simply connected, Lie group of the 2×2 quaternionic matrices which are elements of SL(4,$\mathbb{C}$),[13] precisely:

$$G = \left\{ g = \begin{pmatrix} a & b \\ c & d\end{pmatrix}, a,b,c,d \in \mathbb{H}, |c||d|\,|ac^{-1} - bd^{-1}| = 1\right\}. \tag{7}$$

SU*(4) is also the universal covering of $\mathrm{SO}_0(1,5)$. A semigroup, denoted by $G_<$, is defined as the maximal subset of G preserving the unit ball in $\mathbb{H}$ under its conformal action:

$$G_< = \{g \in G,\ |x| \leqslant 1 \Rightarrow |g\cdot x| \leqslant 1\}. \tag{8a}$$

It can be shown that

$$G_< = \left\{ g = \begin{pmatrix} a & b \\ c & d\end{pmatrix} \in G,\ |b\bar{d} - a\bar{c}| \leqslant |d|^2 - |c|^2 - 1\right\}. \tag{8b}$$

The simple, simply connected, subgroup $G_< \equiv \mathrm{Sp}(1,1) \approx \overline{\mathrm{SO}_0(1,4)}$[14] is contained in $G_<$:

$$G_< = \left\{ h = \begin{pmatrix} a & b \\ c & d\end{pmatrix} \in G,\ Jh^+J = h^{-1}\right\}, \tag{9a}$$

where

$$h^+ = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d}\end{pmatrix},\quad J = \begin{pmatrix} \underline{1} & \underline{0} \\ \underline{0} & -\underline{1}\end{pmatrix};$$

$G_<$ can be defined equivalently by

$$G_< = \{h \in G_<, |x| = 1 \Rightarrow |h\cdot x| = 1\}. \tag{9b}$$

Thus $G_<$ preserve the interior of the unit ball and separately its frontier S^3 under its conformal action.

The above characterizations can be understood along the following partition of G in three sheets:

$$G = \mathscr{N}^< \cup \mathscr{N}^> \cup \mathscr{N}^=$$

with $\mathscr{N}^{<,>,=} = \{g = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right); |c| <,>,= |d|\}$. For all $g \in \mathscr{N}^<$, we have the factorization:

$$g = \mathrm{tn}(K)\mathrm{dn}(t)h \tag{10}$$

h is an element of $G_<$, dn(t) is the conformal dilatation,

$$\mathrm{dn}(t) = \begin{pmatrix} e^{t/2} & \underline{0} \\ \underline{0} & e^{-t/2}\end{pmatrix}, \tag{11}$$

and tn(K) acts conformally on $\mathbb{H}$ as a simple translation:

$$\mathrm{tn}(K) = \begin{pmatrix} \underline{1} & K \\ \underline{0} & \underline{1}\end{pmatrix},\quad K = (b\bar{d} - a\bar{c})e^t. \tag{12}$$

It is apparent that $G_< \subset G_< \subset \mathscr{N}^<$. It should be noted that the stereographic projection $s(p_0)$ is an element of $\mathscr{N}^=$.

An irreducible bounded representation $\mathscr{T}^\tau, \tau \in \mathbb{C}$, of $G_<$ is given by

$$\phi \in \underline{E} = L^2_{\mathbb{C}}(\mathrm{SU}(2)),\quad h \in G_<,$$

$$\mathscr{T}^\tau(h)\phi(\xi) = [\alpha(\xi,h)]^{-2\tau}\phi(h^{-1}\cdot\xi), \tag{13}$$

where $\alpha(x,g)$, $x \in \mathbb{H}$, $g \in G$, is a multiplier defined by

$$\alpha(x,g) = |cx+d| \quad \text{if } g^{-1} = \begin{pmatrix} a & b \\ c & d\end{pmatrix} \tag{14a}$$

and verifying

$$\alpha(x,g_1g_2) = \alpha(x,g_1)\alpha(g_1^{-1}\cdot x,g_2). \tag{14b}$$

We extend $\mathscr{T}^\tau$ to $G_<^{-1}$ by means of the "harmonic extension" of ϕ

$$\phi \in \underline{E},\quad g \in G_<^{-1},\quad \mathscr{T}^\tau(g)\phi(\xi) = [\alpha(\xi,g)]^{-2\tau}\Phi(g^{-1}\cdot\xi), \tag{15}$$

where

$$\Phi(x) = \frac{1}{2\pi^2}\int_{\mathrm{SU}(2)} d\mu(\xi')\,\frac{1-|x|^2}{|\xi'-x|^4}\,\phi(\xi'),\quad |x| < 1, \tag{16}$$

and

$$\Phi(\rho\xi) \xrightarrow[\rho\to 1_-]{} \phi(\xi) \quad \text{almost everywhere.}$$

$(1/2\pi^2)(1-|x|^2)/|\xi'-x|^4$ is the Poisson kernel in $\mathbb{R}^4$.[15] We recall that its integral on SU(2) is equal to 1. In the following, we are interested exclusively in the $\tau=1$ or 2 case.

3. TEMPERED VECTOR DISTRIBUTION ON G

A. Schwartz space $\mathscr{S}(U,\underline{E})$

The Lie algebra SU*(4) of G is the set of 2×2 quaternionic matrices X with $(\mathrm{Tr}X)_0=0$. It will be denoted by $\mathfrak{g}$. The Killing form on $\mathfrak{g}$ is given by

$$(X,Y)\in\mathfrak{g}\times\mathfrak{g}\rightarrow B(X,Y)=8[\mathrm{Tr}(XY)]_0.$$

The Cartan involution is defined by

$$X\overset{\theta}{\rightarrow}-X^{+}.$$

We put $\|X\|^2=-B(X,\theta X)$.

Let $g=k(g)\exp X(g)$, the Cartan decomposition of G; $k(g)\in\mathrm{Spin}(5)$ and $\exp X(g)=[g^{+}g]^{1/2}$.

We introduce with Harish-Chandra,[16]

$$\sigma(g)\equiv\|X(g)\|. \tag{17}$$

The "seminorm" σ takes into account these elements which do not belong to any compact subgroup of G: $\sigma(g)=0$ for $g\in\mathrm{Spin}(5)$.

For instance $\sigma(\mathrm{dn}(t))=2|t|$, where it is whorthwhile to note that $\mathrm{dn}(t)$, given by (11), is an element of the Cartan subgroup A of G.[13] We have also $\sigma(l(t))=2|t|$, where

$$l(t)\equiv\begin{pmatrix}\underline{\cosh t/2} & \underline{\sinh t/2}\\ \underline{\sinh t/2} & \underline{\cosh t/2}\end{pmatrix}\in A_{<} \tag{18}$$

$A_{<}$ is the Cartan subgroup of $G_{<}$.[14] The following properties should also be noted:

$$\sigma(g)=\sigma(g^{+})=\sigma(g^{-1}),$$
$$\sigma(g_1g_2)\leqslant\sigma(g_1)+\sigma(g_2), \tag{19}$$

for all $g_1,g_2\in G$.

Now, any element X of g is identified with the left invariant (resp. right invariant) vector field $\overleftarrow{X}$ (resp. $\overrightarrow{X}$) on G[17]

$$f\in C^{\infty}(G),\quad g\in G,$$
$$(\overleftarrow{X}f)(g)=(df(g(\exp tX))/dt)|_{t=0} \tag{20}$$

[resp. $(\overrightarrow{X}f)(g)=(df((\exp tX)g)/dt)|_{t=0}$].

We extend in the usual manner (Birkhoff–Witt) the isomorphism $X\rightarrow\overleftarrow{X}$ (resp. anti-isomorphism $X\rightarrow\overrightarrow{X}$) to an homomorphism $Z\rightarrow\overleftarrow{Z}$ (resp. antihomomorphism $Z\rightarrow\overrightarrow{Z}$) of the enveloping algebra $\mathfrak{G}$ of $\mathfrak{g}$ into the algebra of the differential operators on G. That extension is unique and it is clear that $\overleftarrow{Z}$ and $\overrightarrow{Z}'$ commute for all Z,Z' in $\mathfrak{G}$.

Let U be an open subset of G. The "Schwartz space" $\mathscr{S}(U,\mathbb{C})\subset C^{\infty}(U,\mathbb{C})$ is the Hausdorff complete locally convex space of the function $f: U\rightarrow\mathbb{C}$, provided with the seminorms ($r\geqslant 0$):

$$\nu_{Z,Z',r}(f)=\sup_{U}|(\overleftarrow{Z}\overrightarrow{Z}'f)(g)|[1+\sigma(g)]^{r}. \tag{21}$$

In the same way, we define $\mathscr{S}(U,\underline{E})$, the space of the functions f: $U\rightarrow E$, satisfying $(\phi,f)_{\underline{E}}\in\mathscr{S}(U,\mathbb{C})$ for all $\phi\in\underline{E}$, provided with a "weak" topology induced by that of $\mathscr{S}(U,\mathbb{C})$:

$$f_i\xrightarrow[i\rightarrow+\infty]{}0\quad\text{in }\mathscr{S}(U,\underline{E})\Leftrightarrow(\phi,f_i)_{\underline{E}}\xrightarrow[i\rightarrow+\infty]{}0\ \text{in}\ \mathscr{S}(\mathrm{U},\mathbb{C}),$$

for all $\phi\in\underline{E}$.

B. Tempered $\underline{E}$-distribution on G

By "tempered E-distribution T on G" we shall understand an element of $\mathscr{L}(\mathscr{S}(G,\underline{E}),\underline{E})$, i.e., a continuous linear map from $\mathscr{S}(G,\underline{E})$ to $\underline{E}$, where the continuity is defined in a weak sense:

$$f\in\mathscr{S}(G,\underline{E})\overset{T}{\longrightarrow}\langle T,f\rangle\equiv\int_G dT(g)f(g),$$

and T is such that the map $f\rightarrow(\phi\ \langle T,f\rangle)_{\underline{E}}$ is continuous for all ϕ in $\underline{E}$.

We denote the convolution product by

$$\langle S*T,f\rangle\equiv\int_G dS(g_1)\int_G dT(g_2)f(g_1g_2).$$

We designate by $\mathscr{S}'$ the space of the tempered $\underline{E}$-distribution. [Also T will eventually denote an element of $\mathscr{L}(\mathscr{S}(G,\mathbb{C}),\mathbb{C})$, i.e., a scalar distribution. Thus, we shall not distinguish between $(\phi,\langle T,f\rangle)_{\underline{E}}$ and $\langle T,(\phi,f)_{\underline{E}}\rangle$.]

It is important for the following to note that there exists a natural isomorphism[17]: $Z\in\mathfrak{G}\rightarrow T_Z\in\mathscr{S}'_{\{e\}}$ between $\mathfrak{G}$ and the algebra of the scalar distributions with support $\{e\}$, where e is the unit element of G, such that

$$T_{ZZ'}=T_Z*T_{Z'},\quad T_{[Z,Z']}=T_Z*T_{Z'}-T_{Z'}*T_Z$$
$$\equiv[T_Z,T_{Z'}],\quad\langle T_Z,f\rangle=(\overrightarrow{Z}f)(e)=(\overleftarrow{Z}f)(e). \tag{20'}$$

Now, $\mathscr{S}'_{\leqslant}(*)$ designates the space of the tempered $\underline{E}$-distribution on G, with support in $G_{\leqslant}^{-1}$, provided with the convolution $*$ (not defined everywhere),

Let us define the subspace $\underline{E}_1\subset\underline{E}$:

$$\underline{E}_1=\Big\{\phi;g\overset{\mathscr{T}^{(\phi',\phi)}}{\longrightarrow}(\phi',\mathscr{T}^{1}(g)\phi)_{\underline{E}}:\mathscr{T}^{(\phi',\phi)}\in\mathscr{S}(U,\mathbb{C})$$
$$\times\text{ for all open }U\subset G_{\leqslant}^{-1}\text{ and for all }\phi'\in\underline{E}\Big\}.$$

Lemma: $\underline{E}_1$ is dense in $\underline{E}$.

Proof: Indeed, the S^3 harmonics $Ynlm$ (Appendix A) are elements of $\underline{E}_1$. This fact can be proved by considering some properties of these basis elements in the Hilbert space $\underline{E}$.

—First, it is easy to check that for all Z,Z' in $\mathfrak{G}$, there exists a constant $C_{Z,Z'}$ such that

$$|(\overleftarrow{Z}\overrightarrow{Z}'A)(\xi,g)|\leqslant C_{Z,Z'}A(\xi,g) \tag{21'}$$

for all $\xi\in\mathrm{SU}(2)$ and $g\in G$, with $A(\xi,g)\equiv[\alpha(\xi,g)]^{-2}$.

—There exists $Z_{nlm}\in\mathfrak{G}$ such that

$$Y_{nlm}(\xi)=\langle T_{Z_{nlm}},A(\xi)\rangle, \tag{22}$$

where $A(\xi)(g)\equiv A(\xi,g)$.

—Afterwards, we have the equalities

$$\mathscr{T}^{1}(g)Y_{nlm}(\xi)=\mathscr{T}^{1}(g)\langle T_{Z_{nlm}},A(\xi)\rangle$$
$$=\langle T_{Z_{nlm}},(\mathscr{T}^{1}(g)A)(\xi)\rangle$$
$$=(\overleftarrow{Z}_{nlm}A(\xi))(g). \tag{23}$$

—Consider now $\mathscr{T}^{1}(g)Y_{nlm}(\xi)$ as a function of g. Then, from (21) and (23)

$$(\overleftarrow{Z}\overrightarrow{Z}'\mathscr{T}^{1}Y_{nlm}(\xi))(g)=|(\overleftarrow{Z}''_{nlm}\overrightarrow{Z}'A(\xi))(g)|\leqslant C_{Z''_{nlm}Z'}A(\xi,g), \tag{24}$$

where $Z''_{nlm} = ZZ_{nlm}$.

—For all $\phi \in \underline{E}$, the Schwartz inequality entails

$$|(\phi,(\overleftrightarrow{Z}\overrightarrow{Z}'\mathscr{T}^1 Y_{nlm})(g))_{\underline{E}}| \leqslant C_{Z''_{nlm}Z'} \|A(g)\|_{\underline{E}} \|\phi\|_{\underline{E}}, \tag{25}$$

where $A(g)(\xi) = A(\xi,g)$.

—$\|A(g)\|_{\underline{E}}$ can be easily evaluated from properties of the Poisson kernel; for $g^{-1} = \binom{a\ b}{c\ d} \in G_<$,

$$\begin{aligned}\|A(g)\|^2_{\underline{E}} &= \int_{\mathrm{SU}(2)} d\mu(\xi)|c\xi + d|^{-4} \\ &= \frac{|d|^{-2}2\pi^2}{|d|^2 - |c|^2}\frac{1}{2\pi^2}\int_{\mathrm{SU}(2)} d\mu(\xi)\frac{1-|d^{-1}c|^2}{|\xi - (-d^{-1}c)|^4} \\ &= \frac{|d|^{-2}2\pi^2}{|d|^2 - |c|^2} \leqslant 2\pi^2|d|^{-2}.\end{aligned} \tag{26}$$

The last inequality results from the Definition (8b) of the semigroup $G_\leqslant$.

—Finally, if we define $d(g^{-1})$ by $g^{-1} = \binom{a\ b}{c\ d(g^{-1})}$,

$$\nu_{Z,Z',r}((\phi,\mathscr{T}^1 Y_{nlm})_E) = \sup_{g\in U}|(\phi,(\overleftrightarrow{Z}\overrightarrow{Z}'\mathscr{T}^1 Y_{nlm})(g))_{\underline{E}}|[1+\sigma(g)]^r$$

$$\leqslant C_{Z''_{nlm}Z'}\|\phi\|_{\underline{E}}(2\pi^2)^{1/2}\sup_{g\in U}[[1+\sigma(g)]^r|d(g^{-1})|^{-1}].$$

Now, it can be shown easily that $[1+\sigma(g)]^r|d(g^{-1})|^{-1}$ is bounded for any g in $G_\leqslant^{-1}$; it is sufficient (and trivial) to check that boundedness for $g = \mathrm{dn}(t)$, $t<0$, and $g = l(t)$ for all t. Q.E.D.

The above lemma is important since it permits one to associate to any tempered $\underline{E}$-distribution T on G with support in $G_\leqslant^{-1}$, a linear operator denoted by $\mathscr{T}^1(T)$, constructed with the aid of a suitable smoothing. Let g be an element of $G_\leqslant^{-1}$. From Eq. (10), $g^{-1} = \mathrm{tn}(K)\mathrm{dn}(t)h$, $h\in\mathrm{Sp}(1,1)$, $t\leqslant 0$. We define:

$$s(g) = \begin{cases} e(\exp - (1-t^2)^{-1}), & 0\leqslant t<1 \\ 0, & t\geqslant 1, \\ 1, & t<0.\end{cases} \tag{27}$$

Then, for all $\phi\in\underline{E}_1$ and $T\in\mathscr{S}'_\leqslant(*)$,

$$\begin{aligned}\mathscr{T}^1(T)\phi &\equiv \langle T,\mathscr{T}^1\phi\rangle \\ &= \int_G dT(g)\, s(g)\mathscr{T}^1(g)\phi \\ &\equiv \int_G dT(g)\,\mathscr{T}^1(g)\phi.\end{aligned} \tag{28}$$

We have the fundamental property

$$\mathscr{T}^1(S*T) = \mathscr{T}^1(S)\mathscr{T}^1(T). \tag{29}$$

4. MINIMAL GLOSSARY FOR PRACTICAL COMPUTATIONS

(a) Particularly important are the following three elements of $\mathfrak{g}$, the Lie algebra of G:

$$D = \frac{1}{2}\begin{pmatrix}\underline{1} & \underline{0}\\ \underline{0} & \underline{1}\end{pmatrix},\quad L = \frac{1}{2}\begin{pmatrix}\underline{0} & \underline{1}\\ \underline{1} & \underline{0}\end{pmatrix},\quad \Omega = \frac{1}{2}\begin{pmatrix}\underline{1} & \underline{1}\\ -\underline{1} & -\underline{1}\end{pmatrix}. \tag{30}$$

Note that $D^2 = L^2 = \frac{1}{4}I$, $\Omega^2 = 0$.

These three matrices generate a Lie algebra isomorphic to $\mathrm{sl}(2,R)$:

$$[D,\Omega] = L,\quad [L,\Omega] = \Omega,\quad [D,L] = \Omega - D. \tag{31}$$

By exponentiating we obtain

$$\exp tD = \mathrm{dn}(t),\quad \exp tL = l(t),$$

$$\exp t\Omega \equiv \omega(t) = \begin{pmatrix}\underline{1+t/2} & \underline{t/2}\\ \underline{-t/2} & \underline{1-t/2}\end{pmatrix}. \tag{32}$$

Now, we introduce $\Omega_i \equiv \Omega\vec{e}_i$ for defining $\omega(\vec{x})$:

$$\exp\vec{x}\cdot\vec{\Omega} \equiv \omega(\vec{x}) = \begin{pmatrix}(1,\vec{x}/2) & (0,\vec{x}/2)\\ (0,-\vec{x}/2) & (1,-\vec{x}/2)\end{pmatrix}. \tag{33}$$

The matrices $l(t)$ and $\omega(\vec{x})$ occur in the Iwasawa–Cartan decomposition[14] of $G_< = \mathrm{Sp}(1,1) = K_< A_< N_<$:

$$h\in G_<,\quad h = kl(t)\omega(\vec{x}),\quad k\in\mathrm{Spin}(4) = K_<. \tag{34}$$

(b) *Distribution T_D and Coulomb–Sturmian operator $\mathfrak{R}_0^{-1}$*: It should be noted that $\mathscr{T}^1(T_D) = \partial_t\mathscr{T}^1(\mathrm{dn}(t))|_{t=0}$, in our framework, since $\mathrm{dn}(t)\in G_\leqslant^{-1}$ for $t\geqslant 0$. The generator D has a special importance with regard to its connection with the restriction of the four-dimensional Laplace operator to S^3. We define the operator $\mathfrak{R}_0^{-1}$ by

$$\mathfrak{R}_0^{-1}\phi(\xi) = \frac{1}{2\pi^2}\int_{\mathrm{SU}(2)} d\mu(\xi')|\xi-\xi'|^{-2}\phi(\xi'). \tag{35}$$

$\mathfrak{R}_0^{-1}$ is a compact self-adjoint operator on $L^2_{\mathbb{C}}(\mathrm{SU}(2))$, with eigenvalues $1/n$, $n = 1,2,\cdots$. Its link with the Coulomb–Sturmian operator[6,10] is well known since Fock.[18–20]

Then $\mathscr{T}^1(T_D) = -\mathfrak{R}_0$ (on $\underline{E}_1$). The $*$-inverse of T_D is defined by

$$\langle T_D^{-1},f\rangle = -\int_0^{+\infty} dt\, f(\mathrm{dn}(t)).$$

Then $\mathscr{T}^1(T_D^{-1}) = -\mathfrak{R}_0^{-1}$ (on $\underline{E}$). Explicitly,

$$\begin{aligned}\mathfrak{R}_0^{-1}\phi(\xi) &= -\int_0^{+\infty} dt\,\mathscr{T}^1(\mathrm{dn}(t))\phi(\xi) \\ &= -\int_0^{+\infty} dt\, e^{-t}\Phi(e^{-t}\xi).\end{aligned} \tag{36}$$

The operator $\mathfrak{R}_0$ is interwinning for the representation $\mathscr{T}^1$ and its "contragradient", namely,

$$\mathfrak{R}_0\mathscr{T}^1(g) = \mathscr{T}^1(Jg^+J)\mathfrak{R}_0. \tag{37a}$$

We have also,

$$\mathfrak{R}_0\mathscr{T}^1(g) = (|d|^2 - |c|^2)\mathscr{T}^2(g)\mathfrak{R}_0 \tag{37b}$$

for all $g^{-1} = \binom{a\ b}{c\ d}\in G_\leqslant$. On the other hand, it can be shown that

$$\mathscr{T}^{1+}(h^{-1}) = \mathscr{T}^2(h) \tag{37c}$$

for all $h\in G_<$.

Thus, $\mathscr{T}^1(T_D*\delta(h)*T_D^{-1}) = \mathscr{T}^2(\delta(h))$, (37d)

where $\delta(g)$ is the Dirac distribution on G.

From the above it should be noted that the representations $\mathscr{T}^1$ and $\mathscr{T}^2$ of G are equivalent to one complementary series unitary irreducible representation.[21]

(c) *Distribution T_Ω and T_L*: Ω is also very useful because we have

$$\mathscr{T}^1(T_\Omega * T_D^{-1})\phi(\xi) = (1+\xi_0)\phi(\xi) = \tfrac{1}{2}|\underline{1}+\xi|^2\phi(\xi). \tag{38}$$

More generally,

$$\mathscr{T}^1([T_D,T_{\Omega_i}]*T_D^{-1})\phi(\xi)=\xi_i\phi(\xi). \tag{39a}$$

It may be possible to simplify the above relation by using Eq. (31);

$$[T_D,T_{\Omega_i}]=T_{[D,\Omega_i]}=T_{L_i},$$

where $L_i \equiv L\vec{e}_i$.

But the matrix

$$\exp tL_i=\begin{pmatrix}\underline{\cos(t/2)} & \sin(t/2)\vec{e}_i\\ \sin(t/2)\vec{e}_i & \underline{\cos(t/2)}\end{pmatrix}=l(t)$$

is not an element of $G_{\leqslant}^{-1}$ for all t.

Nevertheless,

$$\mathscr{T}^1(T_{L_i}*T_D^{-1})\phi(\xi)=-\partial_{t'=0}\int_0^\infty dt\mathscr{T}^1(l_i(t')\mathrm{dn}(t))\phi(\xi)$$

makes sense, because at $t>0$ fixed, $l_i(t')\mathrm{dn}(t)$ is in $G_{\leqslant}^{-1}$ for t' satisfying $(1+\sin t')/\cos t' \leqslant e^t$.

Thus

$$\mathscr{T}^1(T_{L_i}*T_D^{-1})=\xi_i \quad (\text{on } \underline{E}_1). \tag{39b}$$

(d) *Operator* $\mathscr{T}^1(T_X)$, $X\in\mathfrak{g}$: More generally, let $t\to g(t)=\exp tX$ be a one parameter subgroup of G such that $g(t)\in G_{\leqslant}^{-1}$ for $0\leqslant t<\epsilon$. We put

$$g^{-1}(t)=\exp -tX=\begin{pmatrix}a(t) & b(t)\\ c(t) & d(t)\end{pmatrix}$$

or

$$-X=\begin{pmatrix}\dot a(0) & \dot b(0)\\ \dot c(0) & \dot d(0)\end{pmatrix},$$

with $b(0)=c(0)=\underline{0}$, $a(0)=d(0)=\underline{1}$. Then

$$\begin{aligned}\mathscr{T}^1(T_X)\phi(\xi)&=\partial_{t=0}\cdot\mathscr{T}^1(\exp tX)\phi(\xi)\\ &=[-2((\dot d(0))_0+\bar{\dot c}(0)\cdot x)(1+x\cdot\partial_x)\\ &\quad+(\dot a(0)x+x\bar{\dot d}(0)+\dot b(0)\\ &\quad+|x|^2\bar{\dot c}(0))\cdot\partial_x]\Phi(x)|_{x=\xi},\end{aligned} \tag{40}$$

where $x\cdot x'=(x\bar x')_0=x_0x_0'+\vec x\cdot\vec x'$, and $\partial_x=(\partial_{x_0},\vec\nabla_{\vec x})$.

Let u be a quaternion. We define the operators $C(u)$, $C'(u)$, $\mathscr{L}(u)$, $\tilde{\mathscr{L}}(u)$, acting on $L^2_{\mathbb{C}}(\mathrm{SU}(2))$, by

$$C(u)\phi(\xi)\equiv(u\cdot\xi)\phi(\xi), \tag{41a}$$

$$C'(u)\phi(\xi)\equiv(u\cdot\partial_x)\Phi(x)|_{x=\xi}, \tag{41b}$$

$$\mathscr{L}(u)\phi(\xi)\equiv(xu\cdot\partial_x)\Phi(x)|_{x=\xi}, \tag{41c}$$

$$\tilde{\mathscr{L}}(u)\phi(\xi)\equiv(ux\cdot\partial_x)\Phi(x)|_{x=\xi}. \tag{41d}$$

Then,

$$\begin{aligned}\mathscr{T}^1(T_X)=&-2((\dot d(0))_0I+C(\bar{\dot c}(0)))(I+\mathscr{L}(\underline{1}))\\ &+\tilde{\mathscr{L}}(\dot a(0))+\mathscr{L}(\bar{\dot d}(0))+C'(\dot b(0)+\bar{\dot c}(0)).\end{aligned} \tag{42}$$

It should be noted that $I+\mathscr{L}(\underline{1})=\mathfrak{N}_0$ (on $\underline{E}_1$).

(e) *Basis in* $L^2_{\mathbb{C}}(\mathrm{SU}(2))$ *and matrix elements*: We choose as basis in $L^2_{\mathbb{C}}(\mathrm{SU}(2))$ the set of the S^3 harmonics Y_{nlm}, eigenvectors of the operator $\mathfrak{N}_0^{-1}$:

$$\mathfrak{N}_0^{-1}Y_{nlm}=(1/n)Y_{nlm}. \tag{43}$$

Recall that these functions are given in Appendix A.

In Appendix B we explicit the matrix elements of the operator $\mathscr{T}^1(g)$, $g\in G_{\leqslant}^{-1}$.

$$\mathscr{T}^1(g)_{nlm,n'l'm'}=\int_{\mathrm{SU}(2)}d\mu(\xi)Y^*_{nlm}(\xi)|c\xi+d|^{-2}\mathscr{Y}_{n'l'm'}(g^{-1}\cdot\xi), \tag{44}$$

where $\mathscr{Y}_{nlm}$ is the harmonic polynomial ("solid harmonic") deduced from the surface harmonic by the homogeneity formula

$$\mathscr{Y}_{nlm}(x)=|x|^{n-1}Y_{nlm}(x/|x|).$$

Likewise, we give in Appendix C the matrix elements of the operators $C(u)$, $C'(u)$, $\mathscr{L}_{0i}$, $\mathscr{L}_{ij}$. The latter are defined by

$$\begin{aligned}\mathscr{L}_{0i}&=\tfrac12[\tilde{\mathscr{L}}(\vec e_i)+\mathscr{L}(\vec e_i)]\\ &=x_0\partial_i-x_i\partial_0|_{x=\xi},\end{aligned} \tag{45a}$$

$$\begin{aligned}\tilde{\mathscr{L}}_{jk}&=\tfrac12[\tilde{\mathscr{L}}(\vec e_i)-\mathscr{L}(\vec e_i)]\\ &=x_j\partial_k-x_k\partial_j|_{x=\xi},\end{aligned} \tag{45b}$$

where (i,j,k) is an even permutation of $(1,2,3)$.

5. LITTLE LEXICON FOR PHYSICAL OPERATORS

In this section we examine the diverse dynamical variables A occurring in physical calculations, their respective $\mathscr{F}_{p_0}$-transform $\tilde A_{p_0}=\mathscr{F}_{p_0}A\mathscr{F}_{p_0}$, and the distribution $T(A_{p_0})$ defined formally by

$$\tilde A_{p_0}=\mathscr{T}'(T(A_{p_0})) \quad (\text{on } \underline{E}_1). \tag{46}$$

The operator identity (46) displays a sort of completeness property: any "physically reasonable" dynamical variable A has its Fock transform $\tilde A_{p_0}$ in the linear span generated by the representation operators $\mathscr{T}^1(g)$ for g varying in the semigroup $G_{\leqslant}^{-1}$.[10]

A. Free resolvent

We start with the free particle Hamiltonian resolvent $(H_0-E)^{-1}=2m(p_0^2+\vec p^2)^{-1}$ for $E=-p_0^2/2m<0$.

Its Fock transform is simply

$$2m\mathscr{F}_{p_0}(p_0^2+\vec p^2)^{-1}\mathscr{F}_{p_0}^{-1}=2m|\underline{1}+\xi|^2/4p_0^2, \tag{47}$$

and by using the Eq. (38),

$$2m\mathscr{F}_{p_0}(p_0^2+\vec p^2)^{-1}\mathscr{F}_{p_0}^{-1}=(m/p_0^2)\mathscr{T}^1(T_\Omega*T_D^{-1}). \tag{48}$$

B. Scalar product

As a direct consequence, it should be noted that the usual $L^2_{\mathbb{C}}(\mathbb{R}^3)$ scalar product is $\mathscr{F}_{p_0}$-transformed in the following way:

$$\begin{aligned}(\psi,\psi')_{L^2_{\mathbb{C}}(\mathbb{R}^3)}&=(\phi,(1+\xi_0)\phi')_{\underline{E}}\\ &=(\phi,\mathscr{T}^1(T_\Omega*T_D^{-1})\phi')_{\underline{E}},\end{aligned} \tag{49}$$

with $\phi=\mathscr{F}_{p_0}\psi$, $\phi'=\mathscr{F}_{p_0}\psi'$.

C. Yukawa–Coulomb potential

The Fock transform of the Yukawa potential (and its Coulomb limit $\mu=0$) is

$$\begin{aligned}&\mathscr{F}_{p_0}(e^{-\mu r}/r)\mathscr{F}_{p_0}^{-1}\\ &=-(p_0/\hbar)\mathscr{T}^1(T_D*T_\Omega^{-1}*\delta(\omega(\mu\hbar/p_0)*T_D^{-1})).\end{aligned} \tag{50}$$

The Coulomb–Sturmian operator is $(H_0-E)^{-1}(1/r)$. This operator is compact although not symmetrical.

Its Fock transform is compact self-adjoint on $\underline{E}$. From (49) and (50),

$$2m\mathscr{F}_{p_0}(p_0^2+\vec{p}^2)^{-1}(1/r)\mathscr{F}_{p_0}^{-1} = -(m/p_0\hbar)\mathscr{T}^1(T_D^{-1}) = (m/p_0\hbar)\mathfrak{N}_0^{-1}. \qquad (51)$$

D. Coulomb–Green function

The Coulomb resolvent is defined by

$$G(E) = [\vec{p}^2/2m - E - g/r]^{-1} = [I - 2mg(p_0^2+\vec{p}^2)^{-1}(1/r)]^{-1}2m(p_0^2+\vec{p}^2)^{-1},$$

$$p_0 = (-2mE)^{1/2}.$$

We have for $E<0$:

$$\mathscr{F}_{p_0}G(E)\mathscr{F}_{p_0}^{-1} = (m/p_0^2)\mathscr{T}^1((\delta(e)+\nu T_D^{-1})^{-1}*T_\Omega*T_D^{-1}),$$

where $\nu = mg/p_0\hbar$.

Now, the $*$-inverse of $\delta(e)+\nu T_D^{-1}$ exists in $\mathscr{S}'_{\leqslant}(*)$. It is given by[19]

$$[\delta(e)+\nu T_D^{-1}]^{-1} = \delta(e)+\nu T_{0\nu},$$

with

$$\langle T_{0\nu},f\rangle \equiv \int_0^{+\infty} dt e^{\nu t} f(\mathrm{dn}(t)). \qquad (52a)$$

It is trivially verified that

$$T_D^{-1}*(\delta(e)+\nu T_{0\nu}) = -T_{0\nu}.$$

Hence,

$$\mathscr{F}_{p_0}G(E)\mathscr{F}_{p_0}^{-1} = -(m/p_0^2)\mathscr{T}^1(T_D*T_{0\nu}*T_\Omega*T_D^{-1}). \qquad (52b)$$

E. Galilean boost $e^{i\vec{k}\cdot\vec{r}}$

For practical purposes, it is possible to extend formally the representation $\mathscr{T}^\tau$ so that the singular conformal transformations $s(p_0)$ and $s^{-1}(p_0)$ and the Fock transformation are included in our formalism;

$$\hat{\psi} = \mathscr{F}_{p_0}^{-1}\phi = \sqrt{p_0}\,\mathscr{T}^2(s(p_0))\phi,$$

$$\phi = \mathscr{F}_{p_0}\hat{\psi} = \frac{1}{\sqrt{p_0}}\mathscr{T}^2(s^{-1}(p_0))\hat{\psi}. \qquad (53)$$

The action of the operator $e^{i\vec{k}\cdot\vec{r}}$ on a momentum wavefunction $\hat{\psi}(\vec{p})$ can be written

$$e^{i\vec{k}\cdot\vec{r}}\hat{\psi}(\vec{p}) = \hat{\psi}(\vec{p}-\hbar\vec{k}) = (\mathscr{T}^2(\mathrm{tn}(0,\hbar\vec{k}))\hat{\psi})(\vec{p}). \qquad (54)$$

tn is defined in Eq. (12). From Eqs. (53), (54), and (37d) we obtain

$$\mathscr{F}_{p_0}e^{i\vec{k}\cdot\vec{r}}\mathscr{F}_{p_0}^{-1} = \mathscr{T}^2(\omega(\hbar\vec{k}/p_0)) = \mathscr{T}^1(T_D*\delta(\omega(\hbar\vec{k}/p_0))*T_D^{-1}) \qquad (55)$$

where

$$\omega(\hbar\vec{k}/p_0) = s^{-1}(p_0)\mathrm{tn}((0,\hbar\vec{k}))s(p_0)$$

is defined by Eq. (33).

F. Scalar boost or "tilt"[22]

The expression $\mathscr{F}_{p_0}\mathscr{F}_{p_0'}^{-1}$ does not mean anything. On the other hand, $\mathscr{F}_{p_0}\mathscr{T}^2(\mathrm{tn}\underline{\Delta_{00'}})\mathscr{F}_{p_0'}^{-1}$ makes sense, where $\Delta_{00'} = p_0 - p_0'$ and

$$(\mathscr{T}^2(\mathrm{tn}(\underline{\Delta_{00'}}))\hat{\psi})(p_0,\vec{p}) = \hat{\psi}(p_0',\vec{p}).$$

By using (53) and (37d), we obtain

$$\mathscr{F}_{p_0}\mathscr{T}^2(\mathrm{tn}(\underline{\Delta_{00'}}))\mathscr{F}_{p_0'}^{-1} = (p_0'/p_0)^{1/2}\mathscr{T}^2(l(\lambda_{00'})) = (p_0'/p_0)^{1/2}\mathscr{T}^1(\mathrm{T_D}*\delta(l(\lambda_{00'}))*T_D^{-1}), \qquad (56)$$

where

$$l(\lambda_{00'}) = s^{-1}(p_0)\mathrm{tn}(\underline{\Delta_{00'}})s(p_0') = \frac{1}{2(p_0p_0')^{1/2}}\begin{pmatrix} p_0+p_0' & p_0-p_0' \\ p_0-p_0' & p_0+p_0' \end{pmatrix} = \exp\lambda_{00'}L, \quad \lambda_{00'} = \mathrm{Log}(\underline{p_0/p_0'}). \qquad (57)$$

G. Operator $\vec{r}$

From $\vec{r} = -i\nabla_{\vec{k}}e^{i\vec{k}\cdot\vec{r}}|_{\vec{k}=\vec{0}}$ and from (55), we obtain immediately,

$$\mathscr{F}_{p_0}\vec{r}\mathscr{F}_{p_0}^{-1} = -i(\hbar/p_0)\mathscr{T}^1(T_D*T_{\vec{\Omega}}*T_D^{-1}). \qquad (58)$$

H. Operator $\vec{p}$

From the relation $\vec{\xi} = 2p_0\vec{p}/(p_0^2+\vec{p}^2)$ if $\xi = (\xi_0,\vec{\xi}) = s^{-1}(p_0)\cdot(p_0,\vec{p})$ and from Eq. (40), we can write

$$\mathscr{F}_{p_0}\vec{p}\mathscr{F}_{p_0}^{-1} = p_0\mathscr{T}^1(T_D*T_\Omega^{-1}*T_{\vec{L}}*T_D^{-1}). \qquad (59)$$

I. Coulomb bound states

Finally, we recall the well-known relation between the Coulomb bound states ψ_{nlm} and the S^3 harmonics

$$\mathscr{F}_{p_{0n}}\psi_{nlm} = Y_{nlm}, \qquad (60)$$

with $p_{0n} = \sqrt{-2mE_n} = mg/\hbar n$.

E_n is the nth level energy.

6. N-PHOTON TRANSITION MATRIX ELEMENT

We return now to our initial motivation: the calculation of the matrix element (1). By using the scalar product transformation (49) and the Eq. (60) that expression becomes

$$(\psi_{n'l'm'},\mathscr{O}\psi_{nlm})_{L^2_{\mathbb{C}}(\mathbf{R}^3)} = (Y_{n'l'm'},\mathscr{T}^1(T_\Omega*T_D^{-1})\mathscr{F}_{p_{0n'}}\mathscr{O}\mathscr{F}^{-1}_{p_{0n}}Y_{nlm})_{\underline{E}}, \qquad (61)$$

where

$$\mathscr{F}_{p_{0n'}}\mathscr{O}\mathscr{F}_{p_{0n}}^{-1} = \mathscr{F}_{p_{0n'}}\mathscr{T}^2(\mathrm{tn}(\underline{\Delta_{n'1}}))\mathscr{F}_{p_{01}}^{-1} \times\Big[\prod_{i=1}^{N}\mathscr{F}_{p_{0i}}A_i\mathscr{F}_{p_{0i}}^{-1}\mathscr{F}_{p_{0i}}e^{i\vec{k}_i\cdot\vec{r}}\mathscr{F}_{p_{0i}}^{-1}\mathscr{F}_{p_{0i}}G(E_i) \times\mathscr{F}^{-1}_{p_{0i}}\mathscr{F}_{p_{0i}}\mathscr{T}^2(\mathrm{tn}(\underline{\Delta_{ii+1}}))\mathscr{F}_{p_{0i}}^{-1}\mathscr{F}_{p_{0i+1}}^{-1}\Big], \qquad (62)$$

with

$$\Delta_{n'1} = p_{0n'} - p_{01}, \quad \Delta_{ii+1} = p_{0i} - p_{0i+1}$$

$$p_{0N+1} \equiv p_{0n}, \quad p_{0i} = (-2mE_i)^{1/2}.$$

The operators $\mathscr{T}^2(\mathrm{tn}(\underline{\Delta_{ii+1}}))$ ensure the connection between the ith and $(i+1)$th processes.

We introduce

$$T(A_i) = \begin{cases} p_{0i} T_{\Omega}^{-1} * T_{\vec{\epsilon}_i \cdot \vec{L}} & \text{for } A_i = \vec{\epsilon}_i \cdot \vec{p}, \qquad (63a) \\ -(i\hbar/p_{0i}) T_{\vec{\epsilon}_i \cdot \vec{\Omega}} & \text{for } A_i = \vec{\epsilon}_i \cdot \vec{r}, \qquad (63b) \end{cases}$$

then, from (52), (55), and (56), we obtain

$$I_{nlm \to n'l'm'} = \mathscr{T}^1(S)_{n'l'm',nlm} = \int_{G_<} dS(g) (\mathscr{T}^1(g))_{n'l'm',nlm}. \qquad (64a)$$

S is the distribution

$$S = \left[(-m)^N \Big/ \left(\prod_{i=1}^{N} p_{0i}^2\right)\right] (p_{0n}/p_{0n'})^{1/2} \{T_\Omega * \delta(l(\lambda_{n'1})) * \left[\prod_{i=1}^{N} (*) T_{A_i} * \delta(\omega(\hbar \vec{k}_i/p_{0i})) * T_{0\nu_i}\right] * T_\Omega * \delta(l(\lambda_{ii+1})) * T_D^{-1}\}, \qquad (64b)$$

with

$$\lambda_{n'1} = \log(p_{0n'}/p_{01}), \quad \lambda_{ii+1} = \log(p_{0i}/p_{0i+1}),$$
$$\nu_i = mg/p_{0i}.$$

Commutation rules and simplifications

Each distribution with punctual support T_X, $X \in \mathfrak{g}$ involved in Eq. (64b) can be brought to the left (or to the right) of the expression of S by using the following property:

$$\begin{aligned} \mathscr{T}^1(\delta(g_1) * T_X * \delta(g_2)) &= \partial_{t=0} \cdot \mathscr{T}^1(g_1(\exp tX) g_2) \\ &= \partial_{t=0} \cdot \mathscr{T}^1((\exp t g_1 X g_1^{-1}) g_1 g_2) \\ &= \mathscr{T}^1(T_{g_1 X g_1^{-1}} * \delta(g_1 g_2)). \end{aligned} \qquad (65)$$

On the other hand, we have a specific and very useful commutation rule

$$T_\Omega * \delta(l(\lambda)) * T_\Omega^{-1} = e^\lambda \delta(l(\lambda)). \qquad (66)$$

This rule permits one to eliminate systematically the cumbersome presence of the factor T_Ω^{-1} in expression (63).

We have also

$$\begin{aligned} l(\lambda_{00'}) \omega(\hbar \vec{k}/p_0') l(\lambda_{0'0''}) &= l(\lambda_{00''}) \omega(\hbar \vec{k}/p_0'') \\ &= \omega(\hbar \vec{k}/p_0) l(\lambda_{00''}). \end{aligned} \qquad (67)$$

Finally, from $[\vec{\epsilon} \cdot \vec{p}, e^{i\vec{k} \cdot \vec{r}}] = 0$ when $\vec{\epsilon} \cdot \vec{k} = 0$, we deduce

$$[T_\Omega^{-1} * T_{\vec{\epsilon} \cdot \vec{L}}, \delta(\omega(\hbar \vec{k}/p_0))] = 0. \qquad (68)$$

Let us consider now the (usual) situations where we have $A_i = \vec{\epsilon}_i \cdot \vec{p}$ for all i ("situation P) or $A_i = \vec{\epsilon}_i \cdot \vec{r}$ for all i ("situation R").

$$C = \left[(-m)^N \Big/ \prod_{i=1}^{N} p_{0i}\right] (p_{0n'}/p_{0n})^{1/2}, \qquad (P)$$

$$C = \left[(i\hbar m)^N \Big/ \prod_{i=1}^{N} p_{0i}^3\right] (p_{0n'}/p_{0n})^{1/2}. \qquad (R)$$

Then, by using (52a), (65), and (66), we obtain

$$S = C \int_0^{+\infty} dt_1 \cdots \int_0^{+\infty} dt_N \left[\prod_{i=1}^{N} (*) e^{\nu_i t_i} S_i\right] * \delta(g_n) * T_\Omega * T_D^{-1}, \qquad (69)$$

where

$$S_i = \begin{cases} T_{X_i}, & X_i = g_i \vec{\epsilon}_i \cdot \vec{L} g_i^{-1}, \quad (P) \\ T_{g_i \Omega g_i^{-1}} * T_{X_i}, & X_i = g_i \vec{\epsilon}_i \cdot \vec{\Omega} g_i^{-1}, \quad (R) \end{cases}$$

with

$$g_1 = l(\lambda_{n'1}),$$
$$g_{i+1} = g_i \omega(\hbar \vec{k}_i/p_{0i}) \mathrm{dn}(t_i) l(\lambda_{ii+1}),$$
$$g_n = g_{N+1}.$$

The explicit form of the matrix element $\mathscr{T}^1(S)_{n'l'm',nlm}$ is evidently intricate

$$\begin{aligned} \mathscr{T}^1(S)_{n'l'm',nlm} &= -\frac{C}{n} \int_0^{+\infty} dt_1 \cdots \int_0^{+\infty} dt_N \exp\left(\sum_{i=1}^{N} \nu_i t_i\right) \\ &\times \sum_{n_N, l_N, m_N} \mathscr{T}^1\left(\prod_{i=1}^{N} (*) S_i\right)_{n'l'm', n_N l_N m_N} \\ &\times \mathscr{T}^1(\delta(g_n) * T_\Omega)_{n_N l_N m_N, nlm}. \end{aligned} \qquad (70)$$

Let us remark that the matrix element

$$\begin{aligned} &\mathscr{T}^1\left(\prod_{i=1}^{N} (*) S_i\right)_{n'l'm', n_N l_N m_N} \\ &= \sum_{\substack{n_i, l_i, m_i \\ 1 \leqslant i \leqslant N-1}} \prod_{i=1}^{N} \mathscr{T}^1(S_i)_{n_{i-1} l_{i-1} m_{i-1}, n_i l_i m_i} \\ &(n_0 l_0 m_0 \equiv n'l'm') \end{aligned}$$

is equal to 0 except if $|n_N - n'|$, $|l_N - l'|$, $|m_N - m'| \leqslant \{^{N}_{2N}$. This fact is easily understood from Eq. (42) and the expressions of the matrix elements given in Appendix C.

On the other hand, the last integral in t_N can be carried out since the relation (2) implies

$$\begin{aligned} &-(m/p_{0N}^2) \mathscr{T}^1(T_{0\nu_N} * T_\Omega * \delta(l(\lambda_{Nn}))) Y_{nlm} \\ &= -(m/p_{0N} p_{0n}) \mathscr{T}^1(T_{0\nu_N} * \delta(l(\lambda_{Nn})) * T_\Omega) Y_{nlm} \\ &= (E_n - E_N)^{-1} \mathscr{T}^1(l(\lambda_{Nn})) Y_{nlm}. \end{aligned} \qquad (71)$$

Finally, we make the change of variables

$$z_i = e^{-t_i}.$$

Then, from the general structure of the matrix element $\mathscr{T}^1(g)_{nlm,n'l'm'}$ given in Appendix B, it is apparent that the final expression of $I_{nlm \to n'l'm'}$ is a *finite* linear combination of "generalized Euler integrals"[23]:

$$\mathscr{G}_s(p_i, q_i; r; g) = \int_0^1 dz_1 \cdots \int_0^1 dz_s \left[\prod_{i=1}^{s} z_i^{p_i - 1} (1 - z_i)^{q_i - 1}\right] |d(g(z_1, \cdots, z_2)^{-1})|^{-2r}, \qquad (72)$$

where $g^{-1} = \left(\begin{smallmatrix} a & b \\ c & d(g^{-1}) \end{smallmatrix}\right) \in G_<$, $r \geqslant 1$. In our specific case, $s = N - 1$, $r \leqslant n + n' + N$, and

$$g = \left[\prod_{i=1}^{N-1} l(\lambda_{i-1i}) \omega\left(\hbar \frac{\vec{k}_i}{p_{0i}}\right) \mathrm{dn}(t_i)\right] l(\lambda_{N-1n}) \omega\left(\hbar \frac{\vec{k}_N}{p_{0n}}\right), \qquad (73)$$

with $\lambda_{01} = \lambda_{n'1}$.

In the two photon case we recover well-known formulas with hypergeometric or Appell functions.[1–6] It should be noted that all these expressions can be analytically continued to positive energies $E_i = p_{0i}^2 > 0$ or equivalently to purely imaginary values of ν_i.

CONCLUSION

As a final comment, let us emphasize that the above result can be thought of as illustrating particularly a phenomenon which we conjecture to be a general feature of the interacting one- (and surely multi-) particle quantum mechanics. As we have already asserted[10], the completeness of the operatores $\mathscr{T}^1(g)$ in the sense that $\mathscr{F}_{p_0}A\mathscr{F}_{p_0}^{-1} = \int_{G_{<}^{-1}} dT^A(g)\mathscr{T}^1(g)$ for any physically "reasonable" operator A, makes it strongly tempting to claim that any matrix element of a dynamical variable A between two arbitrary square integrable physical state ψ_i,ψ_f always has the form

$$(\psi_f,A\psi_i)_{L^2_c(\mathbb{R}^3)} = \int_{G_{<}^{-1}} \mathrm{d}S^A_{if}(g)|\mathrm{d}(g^{-1})|^{-2}.$$

ACKNOWLEDGMENTS

The author is grateful to J. Patera and P. Winternitz for their hospitality at the Centre de recherche de mathématiques appliquées de l'Université de Montréal.

APPENDIX A

—Coulomb bound states in configuration space.

$$\psi_{nlm}(\vec{r}) = N_{nl}\left(\frac{2p_{0n}}{\hbar}\right)^{l+3/2} r^l e^{-p_{0n}r/\hbar} L^{2l+1}_{n-l-1}\left(\frac{2p_{0n}}{\hbar}r\right)Y_{lm}(\hat{r}), \tag{A1}$$

with

$$N_{nl} = (-1)^{n-l-1}((n-l-1)!/2n(n+l)!)^{1/2},$$

$$p_{0n} = \sqrt{-2mE_n}.$$

—Coulomb bound states in momentum space.

$$\begin{aligned}\hat{\psi}_{nlm}(\vec{p}) &= \frac{1}{(2p\hbar)^{3/2}}\int d\vec{r} e^{-(i/\hbar)\vec{p}\cdot\vec{r}}\psi_{nlm}(\vec{r}) \\ &= \{4p_{0n}^{5/2}/(p_{0n}^2+\vec{p}^2)^2\}Y_{nlm}(\xi_n),\end{aligned} \tag{A2}$$

$$\begin{aligned}\xi_n &= s^{-1}(p_{0n})\cdot(p_{0n},\vec{p}) \\ &= \left(\frac{p_{0n}^2-\vec{p}^2}{p_{0n}^2+\vec{p}^2}, \frac{2p_{0n}\vec{p}}{p_{0n}^2+\vec{p}^2}\right).\end{aligned} \tag{A3}$$

—S^3-Harmonics.

$$Y_{nlm}(\xi) = M_{nl}C^{l+1}_{n-l-1}(\cos\alpha)(\sin\alpha)^l Y_{lm}(\theta,\varphi),$$

$$\xi = (\alpha,\theta,\varphi),$$

$$M_{nl} = (-i)^l l!2^{l+1}\left(\frac{(n-l-1)!n}{(n+l)!2\pi}\right)^{1/2}.$$

Our present definition differs from that given in Ref. 6: we have included the phase factor $(-i)^l$. The Y_{lm}'s are the normalized S^2-harmonics defined in Edmonds.[24] The C^α_k are the Gegenbauer polynomials defined in Magnus *et al.*[25]

APPENDIX B

Representation matrix elements $\mathscr{T}^1(g)_{nlm,n'l'm'}$ are defined for $g^{-1} = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)\in\mathscr{N}^<$, i.e., for $|c| < |d|$:

$$\begin{aligned}\mathscr{T}^1(g)_{nlm,n'l'm'} &= \int_{\mathrm{SU}(2)} d\mu(\xi)Y^*_{nlm}(\xi)|c\xi+d|^{-2}\mathscr{Y}_{n'l'm'}(g^{-1}\cdot\xi) \\ &= \left[(2l+1)(2l+1)\frac{n'}{n}\right]^{1/2}\sum_{\substack{m_1,m_2 \\ m_1',m_2'}}(-1)^{j+j'-m_2-m_2'-1}\begin{pmatrix} j & j & l \\ m_1 & -m_2 & m\end{pmatrix}\begin{pmatrix} j' & j' & l' \\ m_1' & -m_2' & m'\end{pmatrix}\tilde{\mathscr{T}}^1(g)_{jm_1m_2,j'm_1'm_2'},\end{aligned} \tag{B1}$$

where

$$j = \frac{n-1}{2},\quad j' = \frac{n'-1}{2},\quad \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3\end{pmatrix}$$

is a $3j$-coefficient.[24]

$$\begin{aligned}\tilde{\mathscr{T}}^1(g)&_{jm_1,m_2,j'm_1'm_2'} \\ &= (\sigma^j_{m_2}\sigma^{j'}_{m_2'}/\sigma^j_{m_1}\sigma^{j'}_{m_1'})\sum_{\substack{j_i,m_{i1},m_{i2} \\ 1\leqslant i\leqslant 4}}\delta_{j',j_1+j_2}\delta_{j_4,j'+j_3} \\ &\times\delta_{j,j_1+j_3}\delta_{m_1',m_{11}+m_{21}} \\ &\delta_{m_{42},m_2'+m_{31}}\delta_{m_{41},m_2+m_{22}} \\ &\times\delta_{m_1,m_{12}+m_{32}}(-1)^{2j_3}\prod_{i=1}^4\sigma^{j_i}_{m_{i1}m_{i2}} \\ &\times\mathscr{D}^{j_i}(g_i)_{m_{i1}m_{i2}}/|d|^{4j_4+2}(\sigma^{j_4}_{m_{41}m_{42}})^2,\end{aligned} \tag{B2}$$

with

$$g_1 = a,\quad g_2 = b,\quad g_3 = c,\quad g_4 = \bar{d},$$

$$\sigma^j_m = [(j-m)!(j+m)!]^{-1/2},$$

$$\sigma^j_{m_1m_2} = \sigma^j_{m_1}\sigma^j_{m_2};$$

the $\mathscr{D}^j(x)_{m_1m_2}$ are the homogeneous harmonic polynomials on $\mathbb{H}$ extending the usual matrix elements of the unitary irreducible representations of SU(2):

$$\begin{aligned}\mathscr{D}^j(x)&_{m_1m_2} \\ &= \frac{(-1)^{m_1-m_2}}{\sigma^j_{m_1m_2}}\sum_t\frac{(x_0+ix_1)^{j-m_2-t}}{(j-m_2-t)!}\frac{(x_0-ix_1)^{j+m_1-t}}{(j+m_1-t)!} \\ &\times\frac{(x_3+ix_2)^t}{t!}\frac{(-x_3+ix_2)^{(t+m_2-m_1)}}{(t+m_2-m_1)!}.\end{aligned} \tag{B3}$$

$$x\equiv\begin{pmatrix} x_0+ix_1 & -x_3+ix_2 \\ x_3+ix_2 & x_0-ix_1\end{pmatrix}\in\mathbb{R}^+\times\mathrm{SU}(2)\approx\mathbb{H}.$$

For $g\in\mathrm{SL}(2,\mathbb{R})\cap\mathscr{N}^<$, the matrix element $\mathscr{T}^1(g)_{nlm,n'l'm'}$ is reduced to a hypergeometric polynomial.

$$\begin{aligned}\mathscr{T}^1(g)_{nlm,n'l'm'} &= \delta_{ll'}\delta_{mm'}(n'/n)^{1/2}[(n_>-l-1)!(n_>+l)! \\ &/(n_<-l-1)!(n_<+l)]^{1/2}d^{-(n_>+l+1)} \\ &a^{n_<-l-1}\{[\gamma(b,c)]^{n_>-n_<} \\ &\times/(n_>-n_<)!\}{}_2F_1(l+1-n_<,n_> \\ &+l+1,n_>-n_<+1;bc/ad),\end{aligned}$$

$$n_{\gtrless} = {}^{\sup}_{\inf}(n,n'),\quad \gamma(b,c) = \begin{cases} b & \text{if } n_> = n' \\ -c & \text{if } n_> = n\end{cases}$$

with $g^{-1} = \begin{pmatrix} a & b \\ c & d\end{pmatrix}$.

APPENDIX C

—Operator $C(u)$.

For $u \in \mathbb{H}$ we write

$$C(u) = u_0 C(\underline{1}) + \sum_i u_i C(\vec{e}_i),$$

$$u_1 = u_z,\ u_2 = u_x,\ u_3 = u_y,$$

$$C(\underline{1})_{nlm,n'l'm'} = \delta_{ll'}\delta_{mm'}[\delta_{n,n'-1}\alpha(n',l) + \delta_{n,n'+1}\alpha(n,l)], \tag{C1}$$

with

$$\alpha(n,l) = \tfrac{1}{2}[(n+l)(n-l-1)/n(n-1)]^{1/2},$$

$$\begin{aligned} C(\vec{e}_1)_{nlm,n'l'm'} &= i\delta_{mm'}\{-\delta_{n,n'-1}[\delta_{l,l'-1}a(n',l')b(l',m') \\ &+ \delta_{l,l'+1}a(n',-l'-1)b(l'+1,m')] \\ &+ \delta_{n,n'+1}[\delta_{l,l'-1}a(n'+1,-l')b(l',m') \\ &+ \delta_{l,l'+1}a(n'+1,l'+1)b(l'+1,m')]\}, \end{aligned} \tag{C2}$$

with

$$a(n,l) = \tfrac{1}{2}[(n+l)(n+l-1)/n(n-1)]^{1/2},$$

$$b(l,m) = [(l+m)(l-m)/(2l+1)(2l-1)]^{1/2},$$

$$C(\vec{e}_2)_{nlm,n'l'm'} = \tfrac{1}{2}i(\delta_{m,m'-1}V^{m'}_{nl,n'l'} - \delta_{m,m'+1}V^{-m'}_{nl,n'l'}), \tag{C3a}$$

$$C(\vec{e}_3)_{nlm,n'l'm'} = -\tfrac{1}{2}(\delta_{m,m'-1}V^{m'}_{nl,n'l'} + \delta_{m,m'+1}V^{-m'}_{nl,n'l'}). \tag{C3b}$$

We have introduced here,

$$\begin{aligned} V^{m'}_{nl,n'l'} &= \delta_{n,n'-1}[\delta_{l,l'-1}a(n',l')c(l',m') \\ &- \delta_{l,l'+1}a(n',-l'-1)c(l'+1,1-m')] \\ &- \delta_{n,n'+1}[\delta_{l,l'-1}a(n'+1,-l')c(l',m') \\ &- \delta_{l,l'+1}a(n'+1,l'+1)c(l'+1,1-m')], \end{aligned} \tag{C3c}$$

with

$$c(l,m) = [(l+m)(l+m-1)/(2l+1)(2l-1)]^{1/2}.$$

—Operator $C'(u)$.

We have the relation

$$C'(u)_{nlm,n'l'm'} = 2n'\delta_{n,n'-1}C(u)_{nlm,n'l'm'}. \tag{C4}$$

—Operators $\mathscr{L}_{0i}$.

$$\begin{aligned} (\mathscr{L}_{01})_{nlm,n'l'm'} &= -i\delta_{nn'}\delta_{mm'}[\delta_{l,l'-1}b(l',m')d(n',l') \\ &+ \delta_{l,l'+1}b(l'+1,m')d(n',l'+1)], \end{aligned} \tag{C5}$$

where $d(n,l) = [(n+l)(n-l)]^{1/2}$.

$$(\mathscr{L}_{02})_{nlm,n'l',m'} = -\tfrac{1}{2}i\delta_{nn'}(\delta_{m,m'-1}A^{m'}_{nl,n'l'} - \delta_{m,m'+1}A^{-m'}_{nl,n'l'}), \tag{C6a}$$

$$(\mathscr{L}_{03})_{nlm,n'l'm'} = \tfrac{1}{2}\delta_{nn'}(\delta_{m,m'-1}A^{m'}_{nl,n'l'} + \delta_{m,m'+1}A^{-m'}_{nl,n'l'}). \tag{C6b}$$

We have introduced here,

$$\begin{aligned} A^{m'}_{nl,n'l'} &= -\delta_{l,l'-1}c(l',m')d(n',l') \\ &+ \delta_{l,l'+1}c(l'+1,1-m')d(n',l'+1). \end{aligned} \tag{C6c}$$

—Operators $\mathscr{L}_{ij}$.

$$(\mathscr{L}_{23})_{nlm,n'l'm'} = im'\delta_{nn'}\delta_{ll'}\delta_{mm'}, \tag{C7}$$

$$\begin{aligned} (\mathscr{L}_{31})_{nlm,n'l'm'} &= \tfrac{1}{2}i\delta_{nn'}\delta_{ll'}\{\delta_{m,m'-1}[(l'+m')(l'-m'+1)]^{1/2} \\ &+ \delta_{m,m'+1}[l'+m'+1)(l'-m')]^{1/2}\}, \end{aligned} \tag{C8}$$

$$\begin{aligned} (\mathscr{L}_{12})_{nlm,n'l'm'} &= -\tfrac{1}{2}\delta_{nn'}\delta_{ll'}\{\delta_{m,m'-1}[(l'+m')(l'-m'+1)]^{1/2} \\ &- \delta_{m,m'+1}[(l'+m'+1)(l'-m')]^{1/2}\}. \end{aligned} \tag{C9}$$

[1] P. Lambropoulos, "Topics on Multiphoton Processes in Atoms," in *Advances in Atomic and Molecular Physics*, edited by D. R. Bates and B. Bederson (Academic, New York, 1976), Vol. 12, p. 87 and references therein.

[2] M. Gavrila, Phys. Rev. **163**, 147 (1967).

[3] S. Klarsfeld, Lett. Nuovo Cimento, Ser. I, **1**, 682 (1969); **2**, 548 (1969); **3**, 395 (1970); Ya. I. Granovskii, Zh. Eksp. Teor. Fiz. **56**, 605 (1969) [Sov. Phys. JETP **29**, 333 (1969)].

[4] E. Karule, J. Phys. B **4**, L67 (1971); "Atomic Processes, Report of the Latvian Academy of Sciences," Paper No. YΔ K539.188, pp. 5–24.

[5] A. Maquet, Phys. Rev. A **15**, 1088 (1977).

[6] J. P. Gazeau, J. Math. Phys. **19**, 1048 (1978).

[7] C. Fronsdal, Phys. Rev. **179**, 1513 (1969).

[8] R. W. Huff, Phys. Rev. **186**, 1367 (1969).

[9] A. O. Barut and H. Kleinert, Phys. Rev. **156**, 1541 (1967); **157**, 1180 (1967); A. O. Barut and R. Wilson, Phys. Rev. A **13**, 918 (1976).

[10] J. P. Gazeau, Lett. Math. Phys. **3**, 285 (1979); unpublished thesis (Université de Paris VI, 1978).

[11] J. P. Gazeau and A. Maquet, Phys. Rev. A **20**, 797 (1979).

[12] M. C. Dumont-Lepage, N. Gani, J. P. Gazeau, A. Ronveaux, J. Phys. A **13**, 1243 (1980).

[13] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).

[14] R. Takahashi, Bull. Soc. Math. France, **91**, 289 (1963); S. Ström, Ann. Inst. Henri Poincaré **13**, 77 (1970).

[15] R. R. Coifman and G. Weiss, *Analyse harmonique non commutative sur certains espaces homogènes*, Lecture notes in Mathematics, Vol. 242 (Springer-Verlag, Berlin, 1971).

[16] Harish Chandra, Acta Math. **116**, 1 (1966).

[17] G. Warner, *Harmonic Analysis on Semi-simple Lie Groups* (Springer-Verlag, Berlin, 1972).

[18] W. Fock, Z. Phys. **98**, 145 (1935).

[19] J. Schwinger, J. Math. Phys. **5**, 1606 (1964).

[20] M. Bander and C. Itzyckson, Rev. Mod. Phys. **38**, 330 (1966).

[21] R. L. Lipsman, *Group Representations*, Lecture Notes in Mathematics, Vol. 388 (Springer-Verlag, Berlin, 1974).

[22] A. O. Barut, in *Lie Algebras, Applications and Computational Methods*, edited by B. Kolman (Soc. Ind. Appl. Math., Philadelphia, 1973), p. 79.

[23] H. Exton, *Multiple Hypergeometric Functions and Applications* (Ellis Harwood, Chichester, 1976).

[24] A. R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P., Princeton, N.J., 1957).

[25] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer-Verlag, Berlin, 1966).

Exact invariants for a class of time-dependent nonlinear Hamiltonian systems

H. Ralph Lewis
Los Alamos National Laboratory, Los Alamos, New Mexico 87545
P. G. L. Leach
Department of Applied Mathematics, La Trobe University, Bundoora, 3083, Australia

(Received 15 January 1980; accepted for publication 11 May 1981)

A method of generalizing a class of invariants for a time-dependent linear oscillator is developed for the motion of a mass point in one dimension with a general time-dependent nonlinear potential. Formulas are derived for the allowable time-dependent potentials and for the corresponding invariants. The method by which these conclusions are reached is interesting theoretically and is explained in detail.

PACS numbers: 46.10. + z

I. INTRODUCTION

The question of the existence of invariants (constants of the motion or first integrals) is one of central importance in the study of any dynamical system, be it classical or quantum. If a sufficient number of invariants be known, the motion may be describable without actually integrating the equations of motion. For many systems the Hamiltonian provides a first integral; but there are systems of practical importance for which the Hamiltonian is time-dependent and, therefore, is not an invariant. Such a Hamiltonian occurs in the description of the motion of a charged particle in a time-dependent electromagnetic field.[1]

Various methods have been used to obtain approximate solutions for such time-dependent problems. The usual methods are the adiabatic approximation, the sudden approximation and time-dependent perturbation techniques. A simple time-dependent problem of interest has the Hamiltonian

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\Omega^2(t)\,q^2. \tag{1.1}$$

An adiabatic invariant for (1.1) was given at the first Solvay Congress in 1911 when (1.1) was used as an approximate Hamiltonian for the slowly lengthening pendulum.[2] An exact invariant was used by Courant and Snyder in discussing particle accelerators.[1] That invariant was obtained independently by Lewis[3] by applying Kruskal's asymptotic method[4] to (1.1) in closed form. By a systematic application of Kruskal's method, Sarlet[5] generalized the work of Lewis by elaborating classes of Hamiltonians which are susceptible to a particular closed-form treatment. Leach[6] showed that time-dependent linear canonical transformations could be applied fruitfully to (1.1) and, indeed, to the whole class of time-dependent quadratic Hamiltonians. Recently, the formal development of nonlinear time-dependent canonical transformations in series form has been undertaken,[7] but it appears that they will not be of practical use.[8] Two other methods, the method of the Lie theory of extended groups and Noether's theorem, provide an indirect approach *via* the determination of the generators of symmetry transformations. To each generator there corresponds a constant of the motion. These methods have been applied to one-dimensional linear systems,[9] to n-dimensional linear systems[10] and to some nonlinear systems.[11] The systematic development of the study of Ermakov systems[12,13] is also providing useful results.

It appears that the series method and Kruskal's method may have reached the limits of their ability to provide exact solutions for time-dependent problems. The Lie and Noether approaches have been criticized on the grounds that they are indirect methods and involve considerable calculation.[14] Although Ermakov systems, which are coupled second-order equations, provide a direct method, it is necessary to guess what is often called the auxiliary equation.[15] Because there are important practical problems for which exact solutions would be desirable, it is appropriate to seek a new method. In this article we return to the starting point of the discussion of canonical transformations. In particular, we consider the Hamiltonian for the motion of a mass point in an arbitrary, time-dependent, one-dimensional potential and examine canonical transformations which are a generalization of that found by Lewis[3] for the quadratic Hamiltonian (1.1). Our basic result is that this generalization extends considerably the class of Hamiltonians for which an exact invariant can be found. The Ermakov systems treated by Ray and Reid[12] are included in the class. Apart from any practical value of our result, the method of obtaining it is of interest for the theory of canonical transformations. It may also be possible to find invariants for an even wider class of Hamiltonians by generalizing the method. Our result has been used in plasma physics to derive a class of exact, nonlinear, time-dependent solutions of the Vlasov–Poisson equations.[16]

We consider the problem of transforming the Hamiltonian

$$H = \tfrac{1}{2}p^2 + V(q,t) \tag{1.2}$$

to a new Hamiltonian

$$K = K(P,\rho), \tag{1.3}$$

by means of a canonical transformation of the form

$$\begin{aligned} Q &= Q(q,p,\rho,\dot\rho),\\ P &= P(q,p,\rho,\dot\rho), \end{aligned} \tag{1.4}$$

 0022-2488/82/010165-11$02.50

where $\rho = \rho(t)$ is a function of t which is to be determined. A dot over a symbol denotes differentiation with respect to t. The functions Q and P are not to depend on t except implicitly through dependence on ρ and $\dot\rho$. Equations (1.3) and (1.4) are of the form found by Lewis[3] for the case in which $V(q,t)$ is quadratic in q: $V(q,t) = (1/2)\Omega^2(t)q^2$. In that case, $\rho(t)$ is *any* particular solution of a certain differential equation and each such $\rho(t)$ gives a specific function P which is an invariant.

It might seen that the choice (1.2) is unnecessarily restrictive in that, for example, it excludes the so-called Kanai Hamiltonian[17] for the damped harmonic oscillator. Consider, however, the class of Hamiltonians represented by

$$H(q,p,t) = \tfrac{1}{2}f(t)p^2 + g(q,t), \tag{1.5}$$

which includes the Kanai Hamiltonian. Hamilton's equations are

$$\frac{dq}{dt} = f(t)p,$$
$$\frac{dp}{dt} = -\frac{\partial g(q,t)}{\partial q}. \tag{1.6}$$

The change of time variable from t to a variable τ,

$$\tau = \int_{t_0}^{t} f(t')dt', \tag{1.7}$$

renders (1.6) as

$$\frac{dq}{d\tau} = p,$$
$$\frac{dp}{d\tau} = -\frac{\partial V(q,t)}{\partial q}, \tag{1.8}$$

where

$$V(q,t) = g(q,t)/f(t). \tag{1.9}$$

The system (1.8) is of the form (1.2).

It turns out that a crucial feature of our analysis is solving a partial differential equation which can be obtained by requiring that the time rate of change of Q be consistent with H, K, and the transformation (1.4). From (1.3) and (1.4) we have

$$\frac{dQ}{dt} = \frac{\partial K}{\partial P}$$

and (1.10)

$$\frac{dQ}{dt} = \dot q\frac{\partial Q}{\partial q} + \dot p\frac{\partial Q}{\partial p} + \dot\rho\frac{\partial Q}{\partial \rho} + \ddot\rho\frac{\partial Q}{\partial \dot\rho}.$$

requiring the equality of these two expressions for dQ/dt and using Hamilton's equations for $\dot q$ and $\dot p$, we have

$$p\frac{\partial Q}{\partial q} - \frac{\partial V}{\partial q}\frac{\partial Q}{\partial p} + \dot\rho\frac{\partial Q}{\partial \rho} + \ddot\rho\frac{\partial Q}{\partial \dot\rho} = \frac{\partial K}{\partial P}. \tag{1.11}$$

The solution of our problem is intimately connected with the integration of the linear first-order partial differential equation (1.11).

In Sec. II, we discuss canonical transformations for which the generating function is only a function of the old coordinate, the old momentum and time and in which the new canonical variables are given from the outset as functions of the old canonical variables and time. Results of the discussion are applied in Sec. III to transformations whose explicit time dependence is of the form (1.4). General conditions, which are necessary in order that such transformations be canonical, are derived and satisfied. In Sec. IV, we examine the determination of $Q(q,p,\rho,\dot\rho)$ and $P(q,p,\rho,\dot\rho)$ in more detail and find the conditions which govern the class of allowable potentials. Although the class of allowable potentials does not consist of all functions of the independent variables q and t, it is described by an arbitrary function of a single argument and includes time-dependent potentials which can be arbitrarily nonlinear in q. In Sec. V we obtain formulas for $Q(q,p,\rho,\dot\rho)$ and $P(q,p,\rho,\dot\rho)$ and in Sec. VI we discuss our results with some illustrations.

II. CANONICAL TRANSFORMATIONS WITH GENERATING FUNCTIONS *F(q, p, t)*

A canonical transformation between two sets of canonical variables is usually discussed[18] in terms of a generating function which is a function of a mixture of the old and new variables. This is not always satisfactory. For example, when it is desired to obtain one set of variables as explicit functions of the other set, there can be difficulty in inverting the functions.[19] Here we discuss the transformation in terms of a generating function which is a function only of the original variables.[20]

Consider a canonical transformation

$$Q = Q(q,p,t), \quad P = P(q,p,t). \tag{2.1}$$

In order that the transformation be canonical, the Poisson bracket between Q and P must be unity

$$[Q,P]_{q,p} \equiv \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = 1. \tag{2.2}$$

The Lagrangian of the system expressed in terms of (q,p,t) can differ from the Lagrangian expressed in terms of (Q,P,t) by at most a total time derivative calculated along a phase trajectory.[21] Therefore,

$$p\frac{dq}{dt} - H = P\frac{dQ}{dt} - K + \frac{dF}{dt}, \tag{2.3}$$

where K is the transformed Hamiltonian and $F(q,p,t)$ is the

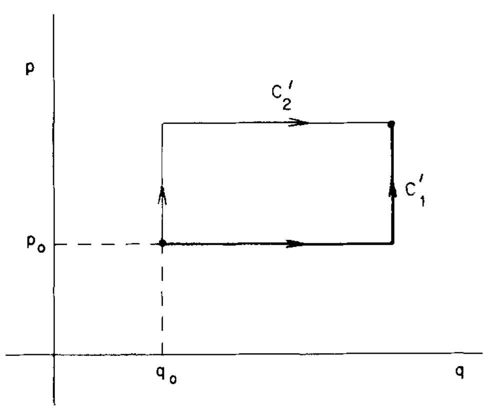


Fig. 1. Paths C_1' and C_2' in the (q,p) plane which define the paths C_1 and C_2 in the (Q,P) plane. C_1 and C_2 are the images of C_1' and C_2', respectively.

generating function of the transformation. Notice that Q, P, and F are functions only of the old variables and time. By writing this equation explicitly in terms of (q, p, t) and setting to zero the coefficients of dq/dt, dp/dt and the part not involving either, we obtain the conditions which must be satisfied in order that Hamilton's principle give the correct equations of motion in terms of the new canonical variables:

$$p - P\frac{\partial Q}{\partial q} - \frac{\partial F}{\partial q} = 0, \tag{2.4}$$

$$P\frac{\partial Q}{\partial p} + \frac{\partial F}{\partial p} = 0, \tag{2.5}$$

$$H + P\frac{\partial Q}{\partial t} - K + \frac{\partial F}{\partial t} = 0. \tag{2.6}$$

Provided Q and F are sufficiently well-behaved to permit the interchange of the order of differentiation (an assumption which is maintained for all functions in this work), the canonical requirement (2.2) follows directly from $(\partial/\partial q)(2.5) + (\partial/\partial p)(2.4)$.

Using (2.4) and (2.5) we can obtain an expression for $F(q, p, t)$ in terms of Q and P. From (2.4),

$$F(q, p, t) = p(q - q_0) - \int_{q_0}^{q} P\frac{\partial Q}{\partial q'}dq' + \psi_1(p, t), \tag{2.7}$$

where $\psi_1(p, t)$ is an arbitrary function of p and t, and $q = q_0(t)$ is an arbitrary function of t. Substituting (2.7) into (2.5) and making use of the Poisson bracket requirement on Q and P, we see that $\psi_1(p, t)$ satisfies

$$\frac{\partial \psi_1}{\partial p} = -P\frac{\partial Q}{\partial p}\bigg|_{q = q_0}. \tag{2.8}$$

We integrate (2.8) and substitute the result into (2.7) to obtain

$$F(q, p, t) = p(q - q_0) - \int_{q_0}^{q} P\frac{\partial Q}{\partial q'}dq' - \int_{p_0}^{p} P\frac{\partial Q}{\partial p'}\bigg|_{q = q_0} dp', \tag{2.9}$$

in which $p_0 = p_0(t)$ is an arbitrary function of t. Alternatively, we could integrate (2.5) first, and substitute the result into (2.4). In this way we find the following different, but equally valid, expression for the generating function,

$$F(q, p, t) = p_0(q - q_0) - \int_{q_0}^{q} P\frac{\partial Q}{\partial q'}\bigg|_{p = p_0} dq' - \int_{p_0}^{p} P\frac{\partial Q}{\partial p'}dp', \tag{2.10}$$

The equivalence of these two expressions for $F(q, p, t)$ can be verified by manipulating (2.9) to obtain (2.10), or vice versa. The basic reason why the two expressions for $F(q, p, t)$ given by (2.9) and (2.10) are equivalent is that the area in the phase plane bounded by a closed curve is invariant under a canonical transformation. This can be seen as follows. Equations (2.9) and (2.10) may be written as

$$F(q, p, t) = p(q - q_0) + \int_{C_2} P\,dQ \tag{2.11}$$

and

$$F(q, p, t) = p_0(q - q_0) - \int_{C_1} P\,dQ, \tag{2.12}$$

where the integrals are line integrals along paths C_2 and C_1 in the (Q, P) plane, and where C_1 and C_2 are the images in the (Q, P) plane of the paths C_1' and C_2' in the (q, p) plane which are shown in Fig. 1. Subtracting (2.12) from (2.11), we have

$$(p - p_0)(q - q_0) = -\int_{C_1 + C_2} P\,dQ. \tag{2.13}$$

In order to show that this equation is an identity, we transform the right-hand side. Let $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ be unit vectors in the directions of increasing q and p, respectively. Define $\hat{\mathbf{z}}$ by $\hat{\mathbf{z}} = \hat{\mathbf{q}} \times \hat{\mathbf{p}}$ and a gradient operator by $\nabla = \hat{\mathbf{q}}(\partial/\partial q) + \hat{\mathbf{p}}(\partial/\partial p)$. Using Stokes' theorem, the line integral in (2.13) can be transformed to a surface integral

$$\begin{aligned}\int_{C_1 + C_2} P\,dQ &= \int_{C_1' + C_2'} d\mathbf{l}\cdot\left(\hat{\mathbf{q}}P\frac{\partial Q}{\partial q} + \hat{\mathbf{p}}P\frac{\partial Q}{\partial p}\right)\\ &= \int_{C_1' + C_2'} d\mathbf{l}\cdot P\nabla Q\\ &= \int_{R'} dq\,dp[\nabla\times(P\nabla Q)]\cdot\hat{\mathbf{z}}\\ &= \int_{R'} dq\,dp[\nabla P\times\nabla Q]\cdot\hat{\mathbf{z}}\\ &= \int_{R'} dq\,dp\left(\frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} - \frac{\partial P}{\partial p}\frac{\partial Q}{\partial q}\right)\\ &= -\int_{R'} dq\,dp = -(p - p_0)(q - q_0),\end{aligned} \tag{2.14}$$

where $d\mathbf{l}$ is a differential line element along the path $C_1' + C_2'$. Thus, (2.13) is an identity based on the area-preservation property of canonical transformations.

III. APPLICATION TO THE TRANSFORMATION $Q = Q(q, p, \rho, \dot{\rho})$, $P = P(q, p, \rho, \dot{\rho})$

We now consider the particular problem which is of interest to us. We take the old Hamiltonian to be

$$H = \tfrac{1}{2}p^2 + V(q, t), \tag{3.1}$$

with $V(q, t)$ arbitrary, and we consider canonical transformations of the form $Q = Q(q, p, \rho, \dot{\rho})$, $P = P(q, p, \rho, \dot{\rho})$ which produce a transformed Hamiltonian

$$K = K(P, \rho). \tag{3.2}$$

The explicit time dependence of Q, P, and K is given in terms of the as yet unknown function $\rho(t)$ and its first derivative $\dot{\rho}(t)$.

Equations (2.4) and (2.5) have been satisfied by solving for $F(q, p, t)$. The remaining equation to be solved is (2.6). We can postpone dealing with the term $\partial F/\partial t$ in (2.6) by using (2.4)–(2.6) to derive two conditions which do not involve F. One condition is obtained by differentiating (2.6) with respect to q and using the time derivative of (2.4) to eliminate $\partial^2 F/\partial q\partial t$; the other is obtained by combining the derivative of (2.6) with respect to p with the time derivative of (2.5). The two conditions are

$$\frac{\partial K}{\partial P}\frac{\partial P}{\partial q} = \frac{\partial V}{\partial q} + \dot{\rho}\left[\frac{\partial P}{\partial q}\frac{\partial Q}{\partial \rho} - \frac{\partial P}{\partial \rho}\frac{\partial Q}{\partial q}\right] + \ddot{\rho}\left[\frac{\partial P}{\partial q}\frac{\partial Q}{\partial \dot{\rho}} - \frac{\partial P}{\partial \dot{\rho}}\frac{\partial Q}{\partial q}\right], \tag{3.3}$$

$$\frac{\partial K}{\partial P}\frac{\partial P}{\partial p} = p + \dot{\rho}\left[\frac{\partial P}{\partial p}\frac{\partial Q}{\partial \rho} - \frac{\partial P}{\partial \rho}\frac{\partial Q}{\partial p}\right] + \ddot{\rho}\left[\frac{\partial P}{\partial p}\frac{\partial Q}{\partial \dot{\rho}} - \frac{\partial P}{\partial \dot{\rho}}\frac{\partial Q}{\partial p}\right]. \tag{3.4}$$

We do not consider the case $\rho(t)$ equal to a linear function of t; that would be equivalent to considering the most general time-dependent canonical transformation. Furthermore, in the domain of t over which the equations of motion are to be solved, we assume that $\dot{\rho}$ cannot be expressed as a function of ρ, and that ρ and $\dot{\rho}$ cannot be inverted to give t as a function of ρ and $\dot{\rho}$. All quantities in (3.4) except $\ddot{\rho}$ manifestly involve t only through ρ and $\dot{\rho}$. Therefore, either the coefficient of $\ddot{\rho}$ must vanish, or $\ddot{\rho}$ must be expressible completely in terms of ρ and $\dot{\rho}$, or both. If $\ddot{\rho}$ were expressed completely in terms of ρ and $\dot{\rho}$, then (3.3) would require that the time dependence of $\partial V/\partial q$ be expressed completely in terms of ρ and $\dot{\rho}$. This possibility will be treated in a subsequent publication. In the present paper, we consider that at least part of the time dependence of $\partial V/\partial q$ is not expressed in terms of ρ and $\dot{\rho}$, but is given explicitly in terms of t. Furthermore, we shall assume that the time dependence of $\partial V/\partial q$ does *not* involve $\dot{\rho}$. Because explicit dependence of $\partial V/\partial q$ on t is allowed, the coefficient of $\ddot{\rho}$ in (3.4) must vanish,

$$\frac{\partial P}{\partial p}\frac{\partial Q}{\partial \dot{\rho}} - \frac{\partial P}{\partial \dot{\rho}}\frac{\partial Q}{\partial p} = 0. \tag{3.5}$$

In terms of the gradient operator ∇ in $(p, \dot{\rho})$ space, (3.5) is

$$(\nabla P)\times(\nabla Q) = 0. \tag{3.6}$$

The general solution of (3.6) is

$$P = \Gamma(Q, q, \rho), \tag{3.7}$$

where Γ is an arbitrary differentiable function.

In (3.3), the first and third terms involve t only through q, p, ρ and $\dot{\rho}$. For this to be true, we must have

$$\frac{\partial V}{\partial q} + \ddot{\rho}\left[\frac{\partial P}{\partial q}\frac{\partial Q}{\partial \dot{\rho}} - \frac{\partial P}{\partial \dot{\rho}}\frac{\partial Q}{\partial q}\right] = f(q, p, \rho, \dot{\rho}), \tag{3.8}$$

where f is arbitrary. Substituting for P from (3.7) into (3.8), we obtain

$$f = \frac{\partial V}{\partial q} + \ddot{\rho}\frac{\partial \Gamma}{\partial q}\frac{\partial Q}{\partial \dot{\rho}}; \tag{3.9}$$

that is, ρ satisfies the second order differential equation

$$\ddot{\rho} = \left(f - \frac{\partial V}{\partial q}\right)\Big/\left(\frac{\partial \Gamma}{\partial q}\frac{\partial Q}{\partial \dot{\rho}}\right). \tag{3.10}$$

Clearly, the right-hand side of (3.10) must be independent of q and p. This is a condition of f, Γ, and Q, the implications of which we now develop.

Substituting (3.7) into (3.3) and (3.4), we have

$$\frac{\partial K}{\partial P}\left(\frac{\partial \Gamma}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial \Gamma}{\partial q}\right) = f + \dot{\rho}\left(\frac{\partial \Gamma}{\partial q}\frac{\partial Q}{\partial \rho} - \frac{\partial \Gamma}{\partial \rho}\frac{\partial Q}{\partial q}\right) \tag{3.11}$$

and

$$\frac{\partial K}{\partial P}\frac{\partial \Gamma}{\partial Q}\frac{\partial Q}{\partial p} = p - \dot{\rho}\frac{\partial \Gamma}{\partial \rho}\frac{\partial Q}{\partial p}. \tag{3.12}$$

Substituting (3.7) into the Poisson bracket requirement for the canonical variables Q and P, we obtain

$$\frac{\partial Q}{\partial p}\frac{\partial \Gamma}{\partial q} = -1. \tag{3.13}$$

Using (3.13) in (3.12) we have

$$\frac{\partial K}{\partial P}\frac{\partial \Gamma}{\partial Q} = -p\frac{\partial \Gamma}{\partial q} - \dot{\rho}\frac{\partial \Gamma}{\partial p}. \tag{3.14}$$

Since the left-hand side of (3.14) is manifestly a function of Q, q, and ρ only, the right-hand side must also be a function of Q, q, and ρ only. Therefore, the $(p, \dot{\rho})$ dependence of the right-hand side is only the result of the dependence of Q on p and $\dot{\rho}$. This is equivalent to

$$(\nabla Q)\times\left\{\nabla\left(p\frac{\partial \Gamma}{\partial q} + \dot{\rho}\frac{\partial \Gamma}{\partial \rho}\right)\right\} = 0,$$

where again ∇ is the gradient operator in $(p, \dot{\rho})$ space. This simplifies to

$$\frac{\partial Q}{\partial \dot{\rho}}\frac{\partial \Gamma}{\partial q} - \frac{\partial Q}{\partial p}\frac{\partial \Gamma}{\partial \rho} = 0. \tag{3.15}$$

Equations (3.14) and (3.15) are a pair of equations linear in $\partial \Gamma/\partial q$ and $\partial \Gamma/\partial \rho$ whose solution is

$$\frac{\partial \Gamma}{\partial q} = -\left[\frac{\partial K}{\partial P}\frac{\partial Q}{\partial p}\Big/\left(p\frac{\partial Q}{\partial p} + \dot{\rho}\frac{\partial Q}{\partial \dot{\rho}}\right)\right]\frac{\partial \Gamma}{\partial Q}, \tag{3.16}$$

$$\frac{\partial \Gamma}{\partial p} = -\left[\frac{\partial K}{\partial P}\frac{\partial Q}{\partial \dot{\rho}}\Big/\left(p\frac{\partial Q}{\partial p} + \dot{\rho}\frac{\partial Q}{\partial \dot{\rho}}\right)\right]\frac{\partial \Gamma}{\partial Q}. \tag{3.17}$$

Combining (3.13) with (3.16) and (3.17) we have

$$\frac{\partial \Gamma}{\partial q} = -1\Big/\frac{\partial Q}{\partial p}, \tag{3.18}$$

$$\frac{\partial \Gamma}{\partial Q} = \left(p\frac{\partial Q}{\partial p} + \dot{\rho}\frac{\partial Q}{\partial \dot{\rho}}\right)\Big/\left[\frac{\partial K}{\partial P}\left(\frac{\partial Q}{\partial p}\right)^2\right], \tag{3.19}$$

$$\frac{\partial \Gamma}{\partial \rho} = -\frac{\partial Q}{\partial \dot{\rho}}\Big/\left(\frac{\partial Q}{\partial p}\right)^2. \tag{3.20}$$

The left-hand sides of (3.18)–(3.20) are functions of Q, q, and ρ only. Hence their right-hand sides depend on p and $\dot{\rho}$ only through dependence on Q. Applying this to (3.18) yields

$$\frac{\partial Q}{\partial \dot{\rho}}\frac{\partial^2 Q}{\partial p^2} = \frac{\partial Q}{\partial p}\frac{\partial^2 Q}{\partial \dot{\rho}\partial p},$$

which is equivalent to

$$\frac{\partial}{\partial p}\left(\frac{\partial Q}{\partial \dot{\rho}}\Big/\frac{\partial Q}{\partial p}\right) = -\frac{\partial}{\partial p}\left(\frac{\partial Q}{\partial \dot{\rho}}\frac{\partial \Gamma}{\partial q}\right) = 0. \tag{3.21}$$

From the second relation in (3.21) it follows that the denominator of (3.10) is independent of p and hence f is independent of p. Requiring that the right-hand side of (3.20) be a function only of Q, q, and ρ and using (3.21), we obtain

$$\frac{\partial}{\partial \dot\rho}\left(\frac{\partial Q}{\partial \dot\rho}\Big/\frac{\partial Q}{\partial p}\right) = -\frac{\partial}{\partial \dot\rho}\left(\frac{\partial Q}{\partial \dot\rho}\frac{\partial \Gamma}{\partial q}\right) = 0. \quad (3.22)$$

Thus the denominator of (3.10) is also independent of $\dot\rho$. Manipulation of (3.19) does not produce any additional information; the requirement that the right-hand side of (3.19) be only a function of Q, q, and p can be combined with (3.21) and (3.22) to give (3.21) again.

Equations (3.21) and (3.22) are equivalent to the first-order partial differential equation

$$\frac{\partial Q}{\partial \dot\rho} - h(q,\rho)\frac{\partial Q}{\partial p} = 0, \quad (3.23)$$

where h is an arbitrary function. The general solution is

$$Q(q, p, \rho, \dot\rho) = R(\xi, q, \rho), \quad (3.24)$$

where ξ is defined by

$$\xi = p + h(q, \rho)\dot\rho. \quad (3.25)$$

By using (3.24), we can rewrite (3.18)–(3.20) as

$$\frac{\partial \Gamma}{\partial q} = -1\Big/\frac{\partial R}{\partial \xi}, \quad (3.26)$$

$$\frac{\partial \Gamma}{\partial R} = \xi\Big/\left(\frac{\partial K}{\partial P}\frac{\partial R}{\partial \xi}\right), \quad (3.27)$$

$$\frac{\partial \Gamma}{\partial \rho} = -h(q,\rho)\Big/\frac{\partial R}{\partial \xi}. \quad (3.28)$$

Because Γ is only a function of R, q, and ρ, the right-hand sides of (3.26)–(3.28) must be expressible in terms of ξ, q, and ρ only.

We now write the condition (1.11) in terms of the independent variables $(q, \xi, \rho, \dot\rho)$ instead of $(q, p, \rho, \dot\rho)$.

$$\dot\rho(\xi - h\dot\rho)\frac{\partial R}{\partial \xi}\frac{\partial h}{\partial q} + (\xi - h\dot\rho)\frac{\partial R}{\partial q} - \frac{\partial V}{\partial q}\frac{\partial R}{\partial \xi} + \dot\rho^2\frac{\partial R}{\partial \xi}\frac{\partial h}{\partial \rho} + \dot\rho\frac{\partial R}{\partial \rho} + g(\rho,\dot\rho,t)\frac{\partial R}{\partial \xi}h = \frac{\partial K}{\partial P}, \quad (3.29)$$

where we have taken ρ to satisfy the equation

$$\ddot\rho = g(\rho, \dot\rho, t), \quad (3.30)$$

in which, at present, g is an arbitrary function. We now assume that $h(q,\rho)\,\partial R/\partial \xi$ does not vanish identically. In this case, because we have assumed that $\partial V/\partial q$ does not involve $\dot\rho$, we see immediately from (3.29) that $g(\rho, \dot\rho, t)$ can be written as

$$g(\rho, \dot\rho, t) = g_0(\rho, t) + g_1(\rho, t)\dot\rho + g_2(\rho, t)\dot\rho^2. \quad (3.31)$$

Now (3.29) is a quadratic function of $\dot\rho$ and the coefficient of each power of $\dot\rho$ must vanish separately. These three conditions are

$$h\frac{\partial h}{\partial q} - \frac{\partial h}{\partial \rho} = g_2 h, \quad (3.32)$$

$$\left(\xi\frac{\partial h}{\partial q} + g_1 h\right)\frac{\partial R}{\partial \xi} - h\frac{\partial R}{\partial q} + \frac{\partial R}{\partial \rho} = 0, \quad (3.33)$$

$$\xi\frac{\partial R}{\partial q} + \left(g_0 h - \frac{\partial V}{\partial q}\right)\frac{\partial R}{\partial \xi} = \frac{\partial K}{\partial P}. \quad (3.34)$$

From (3.32) and (3.33) it is apparent that g_2 and g_1 must be functions of ρ only; they cannot depend on t.

IV. THE CLASS OF ADMISSIBLE POTENTIALS

Equation (3.32) may be solved by the method of characteristics. The solution is given by

$$F_1(q_0, h_0) = 0, \quad (4.1)$$

where F_1 is an arbitrary function of its arguments,

$$q_0 = q + ah_0, \quad h_0 = h\Big/\frac{da}{d\rho}, \quad (4.2)$$

and $a = a(\rho)$ is such that

$$g_2(\rho) = -\frac{d^2a}{d\rho^2}\Big/\frac{da}{d\rho}. \quad (4.3)$$

It is more convenient to work in terms of $a(\rho)$ rather than ρ. From (3.30) and (3.31), a is seen to satisfy the second-order ordinary differential equation

$$\ddot a = f_0(a, t) + \dot a f_1(a) \quad (4.4)$$

where $\cdot$ denotes differentiation with respect to t and

$$f_0(a, t) = g_0(\rho(a), t)\frac{da}{d\rho}, \quad f_1(a) = g_1(\rho(a)). \quad (4.5)$$

It is implied in (4.3) that $da/d\rho \neq 0$; therefore, about any given value of ρ there exists a neighborhood in which $a(\rho)$ may be uniquely inverted to give $\rho(a)$. In terms of a and h_0 instead of ρ and h, Eqs. (3.32), (3.33), and (3.34) become

$$h_0\frac{\partial h_0}{\partial q} - \frac{\partial h_0}{\partial a} = 0, \quad (4.6)$$

$$\left(\xi\frac{\partial h_0}{\partial q} + f_1 h_0\right)\frac{\partial R_0}{\partial \xi} - h_0\frac{\partial R_0}{\partial q} + \frac{\partial R_0}{\partial a} = 0, \quad (4.7)$$

$$\xi\frac{\partial R_0}{\partial q} + \left(f_0 h_0 - \frac{\partial V}{\partial q}\right)\frac{\partial R_0}{\partial \xi} = \frac{\partial K}{\partial P}, \quad (4.8)$$

where

$$R_0(\xi, q, a) = R(\xi, q, \rho(a)) \quad (4.9)$$

and now h_0 is taken as expressed in terms of q and a instead of q and ρ.

In light of the characteristics of (3.32) given by setting q_0 and h_0 equal to constants in (4.2), we rewrite (4.7) in terms of the variables ξ, u and a where

$$u = F_2(q_0, h_0), \quad (4.10)$$

F_2 being any arbitrary function of q_0 and h_0 which is functionally independent of F_1, i.e.,

$$F_2(q_0, h_0) \neq G\{F_1(q_0, h_0)\}. \quad (4.11)$$

In the discussion which follows it is assumed that Eqs. (4.1) and (4.10) are at least locally invertible so that q_0 and h_0 may be expressed as functions of u. The condition for this is that

$$\frac{\partial F_1}{\partial h_0}\frac{\partial F_2}{\partial q_0} - \frac{\partial F_1}{\partial q_0}\frac{\partial F_2}{\partial h_0} \neq 0. \quad (4.12)$$

Writing

$$R_0(\xi, q, a) = \bar R_0(\xi, u, a), \quad (4.13)$$

(4.7) becomes

$$\left(\xi\frac{\partial h_0}{\partial q} + f_1 h_0\right)\frac{\partial \bar R_0}{\partial \xi} + \frac{\partial \bar R_0}{\partial a} = 0, \quad (4.14)$$

in which it is understood that h_0 and $\partial h_0/\partial q$ are written in terms of u and a from inversion of (4.1) and (4.10) in some appropriate domain of the variables. From (4.14) it is evident that $u=$ const defines one family of characteristics. The equation for the other family of characteristics is

$$\frac{d\xi}{da}-\xi\frac{\partial h_0}{\partial q}=f_1h_0. \tag{4.15}$$

The solution may be written in terms of the variable v given by

$$v=\xi\exp\left\{-\int^a da'\,\frac{\partial h_0}{\partial q}\right\}-h_0\int^a da' f_1(a')\exp\left\{-\int^{a'} da''\,\frac{\partial h_0}{\partial q}\right\}, \tag{4.16}$$

in which h_0 is to be expressed as a function of u alone. That is, the general solution of (4.15) is obtained by holding v constant in (4.16). Hence the solution of (3.33) is

$$R(\xi,q,p)=R_1(u,v), \tag{4.17}$$

where R_1 is an arbitrary function of u and v, which are defined by (4.10) and (4.16), respectively. Defining

$$X(u,a)=\exp\left\{-\int^a da'\,\frac{\partial h_0}{\partial q}\right\}, \tag{4.18}$$

$$Y(u,a)=h_0\int^a da' f_1(a')X(u,a'), \tag{4.19}$$

in which again h_0 is to be expressed as a function of u alone, we have

$$v=\xi X-Y \leftrightarrow \xi=(v+Y)/X. \tag{4.20}$$

Rewriting (4.8) in terms of $R_1(u,v)$ and the variables u and v, we have

$$\frac{v+Y}{X}\frac{\partial u}{\partial q}\frac{\partial R_1}{\partial u}+\left\{\frac{v+Y}{X}\frac{\partial v}{\partial q}+\left(f_0h_0-\frac{\partial V}{\partial q}\right)\frac{\partial v}{\partial \xi}\right\}\frac{\partial R_1}{\partial v}=\frac{\partial K}{\partial P}. \tag{4.21}$$

From (3.27),

$$\frac{\partial K}{\partial P}=\xi\Big/\left(\frac{\partial \Gamma}{\partial R}\frac{\partial R}{\partial \xi}\right). \tag{4.22}$$

Combining (3.26) and (3.28), we find

$$h\frac{\partial \Gamma}{\partial q}-\frac{\partial \Gamma}{\partial p}=0 \tag{4.23}$$

$$\leftrightarrow\quad h_0\frac{\partial \Gamma_0}{\partial q}-\frac{\partial \Gamma_0}{\partial a}=0, \tag{4.24}$$

where

$$\Gamma_0(R,q,a)=\Gamma(R,q,p). \tag{4.25}$$

The solution is

$$\Gamma(R,q,p)=\Gamma_1(R_1,u). \tag{4.26}$$

Using (4.20) and (4.26), (4.22) becomes

$$\frac{\partial K}{\partial P}=\frac{v+Y}{X^2}f(u,v), \tag{4.27}$$

where

$$f(u,v)=\left(\frac{\partial \Gamma_1}{\partial R_1}\frac{\partial R_1}{\partial v}\right)^{-1}. \tag{4.28}$$

Thus (4.21) may be rewritten as

$$\frac{v+Y}{X^2}X\frac{\partial u}{\partial q}\frac{\partial R_1}{\partial u}+\left\{\frac{v+Y}{X^2}\left(\frac{v+Y}{X}\frac{\partial X}{\partial u}-\frac{\partial Y}{\partial u}\right)X\frac{\partial u}{\partial q}+\left(f_0h_0-\frac{\partial V}{\partial q}\right)X\right\}\frac{\partial R_1}{\partial v}=\frac{v+Y}{X^2}f(u,v). \tag{4.29}$$

For (4.29) to be self-consistent, it must be equivalent to an equation expressed solely in terms of u and v; i.e., it may contain the function a explicitly only through a common multiplicative factor. Hence we may write

$$X\frac{\partial u}{\partial q}=M(u,v) \tag{4.30}$$

and

$$\frac{v+Y}{X^2}\left(\frac{v+Y}{X}\frac{\partial X}{\partial u}-\frac{\partial Y}{\partial u}\right)X\frac{\partial u}{\partial q}+\left(f_0h_0-\frac{\partial V}{\partial q}\right)X=\frac{v+Y}{X^2}E(u,v), \tag{4.31}$$

where $M(u,v)$ and $E(u,v)$ are arbitrary functions of their arguments.

In terms of the definition of u in (4.10), (4.30) may be written as

$$X\left\{\frac{\partial F_2}{\partial q_0}\left(1+a\frac{\partial h_0}{\partial q}\right)+\frac{\partial F_2}{\partial h_0}\frac{\partial h_0}{\partial q}\right\}=M(u,v). \tag{4.32}$$

From the definition of $X(u,a)$ in (4.18) we find

$$\frac{\partial X(u,a)}{\partial a}=-X\frac{\partial h_0}{\partial q}, \tag{4.33}$$

which shows that $\partial h_0/\partial q$ is expressible completely in terms of u and a. The functions $\partial F_2/\partial q_0$ and $\partial F_2/\partial h_0$ can be expressed as functions of u alone because q_0 and h_0 can be expressed as functions of u alone. Therefore, the left-hand side of (4.32) is manifestly independent of v, which means that $M(u,v)$ is only a function of u: $M(u,v)\to M(u)$. (4.32) can now be written as

$$\frac{\partial F_2}{\partial q_0}\left(X-a\frac{\partial X}{\partial a}\right)-\frac{\partial F_2}{\partial h_0}\frac{\partial X}{\partial a}=M(u). \tag{4.34}$$

Since the right-hand side of (4.34) depends upon u only and both $\partial F_2/\partial q_0$ and $\partial F_2/\partial h_0$ can be expressed in terms of u alone, the coefficients of $\partial F_2/\partial q_0$ and $\partial F_2/\partial h_0$ must be functions of u alone, i.e.,

$$X-a\frac{\partial X}{\partial a}=A(u),\qquad \frac{\partial X}{\partial a}=B(u), \tag{4.35}$$

where $A(u)$ and $B(u)$ are arbitrary functions of u, not both of which are zero. Thus $X(u,a)$ can be expressed as

$$X(u,a)=A(u)+aB(u). \tag{4.36}$$

Substituting (4.36) into (4.33), we find

$$\frac{\partial h_0}{\partial q}=-\frac{\partial}{\partial a}\ln[A(u)+aB(u)], \tag{4.37}$$

which checks with (4.18). By using (4.36) in (4.19) we can write $Y(u,a)$ as

$$Y(u,a) = Ah_0 \int^a f_1(a')da' + Bh_0 \int^a a' f_1(a')\, da'. \quad (4.38)$$

The functions $A(u)$ and $B(u)$ can be expressed in terms of the function $F_1(q_0, h_0)$ that was introduced in (4.1). By differentiating (4.1) with respect to q and using (4.2) we obtain

$$\frac{\partial h_0}{\partial q} = -1 \bigg/ \left(\frac{\partial F_1}{\partial h_0} \bigg/ \frac{\partial F_1}{\partial q_0} + a \right). \quad (4.39)$$

By comparing this with (4.37) we then find

$$A(u) = C(u) \frac{\partial F_1}{\partial h_0}, \quad (4.40)$$

$$B(u) = C(u) \frac{\partial F_1}{\partial q_0}, \quad (4.41)$$

where C is an arbitrary nonzero function. These equations express the derivatives of $F_1(q_0, h_0)$ in terms of $u = F_2(q_0, h_0)$ through the functions A and B. The consistency condition for the existence of a solution of (4.40) and (4.41) is

$$\left[\frac{A(u)}{C(u)} \right]' \frac{\partial F_2}{\partial q_0} - \left[\frac{B(u)}{C(u)} \right]' \frac{\partial F_2}{\partial h_0} = 0. \quad (4.42)$$

By using (4.35) we can write (4.34) as

$$A(u) \frac{\partial F_2}{\partial q_0} - B(u) \frac{\partial F_2}{\partial h_0} = M(u). \quad (4.43)$$

Before continuing with the analysis of (4.42) and (4.43), we begin to examine condition (4.31). It will turn out that information derived from examining (4.31) will be important in our treatment of (4.42) and (4.43). We write (4.31) in the form

$$\left[\frac{Y}{X} \frac{\partial X}{\partial u} - \frac{\partial Y}{\partial u} \right] M(u) + v \frac{1}{X} \frac{\partial X}{\partial u} M(u) + \frac{X^3}{v+Y} \left[f_0 h_0 - \frac{\partial V}{\partial q} \right] = E(u, v). \quad (4.44)$$

Each coefficient of an expansion of the left-hand side in powers of v must be independent of a and t. This implies that the quantities Y, Z_1, and Z_2 must be independent of a, where

$$Z_1 = \frac{1}{X} \frac{\partial X}{\partial u}, \quad (4.45)$$

$$Z_2 = \frac{Y}{X} \frac{\partial X}{\partial u} - \frac{\partial Y}{\partial u}. \quad (4.46)$$

For Z_1 to be independent of a, we must have

$$0 = \frac{\partial Z_1}{\partial a} = \frac{A(u) B'(u) - B(u) A'(u)}{[A(u) + aB(u)]^2}, \quad (4.47)$$

the general solution of which is

$$c_1 A(u) + c_2 B(u) = 0, \quad (4.48)$$

where c_1 and c_2 are constants, not both of which are zero. Using (4.48) in (4.38) we find

$$Y(u,a) = \begin{cases} Ah_0 \int^a da' f_1(a') \left[1 - \frac{c_1}{c_2} a' \right], & \text{if } c_2 \neq 0 \\ Bh_0 \int^a da' f_1(a') \left[a' - \frac{c_2}{c_1} \right], & \text{if } c_1 \neq 0. \end{cases} \quad (4.49)$$

From this, $\partial Y / \partial a = 0$ implies

$$f_1(a) = 0, \quad (4.50)$$

which implies

$$Y(u,a) = 0. \quad (4.51)$$

Because $Y(u,a) = 0$, we also have $Z_2 = 0$, which means that Z_2 is independent of a as required.

The linear dependence of $A(u)$ and $B(u)$ restricts the form of $F_1(q_0, h_0)$. Combining (4.48) with (4.40) and (4.41) we find

$$c_2 \frac{\partial F_1}{\partial q_0} + c_1 \frac{\partial F_1}{\partial h_0} = 0, \quad (4.52)$$

the general solution of which is

$$F_1(q_0, h_0) = F_3(\eta), \quad (4.53)$$

where

$$\eta = c_1 q_0 - c_2 h_0 \quad (4.54)$$

and F_3 is an arbitrary function. The function $h_0(q, a)$ is defined by (4.1), which now reduces to

$$F_3(\eta) = 0. \quad (4.55)$$

The function F_3 cannot be identically zero; if it were, (4.55) would not define $h_0(q, a)$. Therefore, the solution of (4.55) must be

$$\eta = c_0 = \text{const}, \quad (4.56)$$

which implies

$$h_0(q_0, a) = (c_1 q - c_0)/(c_2 - c_1 a). \quad (4.57)$$

Since $\partial h_0 / \partial q$ depends only on a, we can calculate $X(u,a)$ directly from (4.18),

$$X(u,a) = c_2 - c_1 a, \quad (4.58)$$

where we have made a particular choice of the irrelevant arbitrary constant in the definition of X. Comparing with (4.36) we find

$$A(u) = c_2, \quad (4.59)$$

$$B(u) = -c_1. \quad (4.60)$$

We now make a particular choice of $F_2(q_0, h_0)$ which is functionally independent of η, and therefore of $F_1(q_0, h_0)$, as required,

$$u = F_2(q_0, h_0) = c_2 q_0 + c_1 h_0. \quad (4.61)$$

We then have the following relations for u, q_0 and h_0:

$$u = [(c_1^2 + c_2^2) q - c_0(c_1 + c_2 a)]/(c_2 - c_1 a), \quad (4.62)$$

$$q_0 = (c_2 u + c_0 c_1)/(c_1^2 + c_2^2), \quad (4.63)$$

$$h_0 = (c_1 u - c_0 c_2)/(c_1^2 + c_2^2). \quad (4.64)$$

Substitution into (4.42) shows that $C(u)$ must be a constant, which we choose to be minus one,

$$C(u) = -1. \quad (4.65)$$

Consistent with (4.40), (4.41), and (4.53), we take

$$F_1(q_0, h_0) = F_3(\eta) = \eta - c_0. \quad (4.66)$$

Finally, from (4.43) we find

$$M(u) = c_1^2 + c_2^2. \quad (4.67)$$

Notice that $X(u, a)$, $A(u)$, $B(u)$, $C(u)$ and $M(u)$ are all constant functions of u.

We now conclude our examination of (4.44) to determine the class of potentials which can be treated by the procedure developed in this paper. Equation (4.44) can now be written as

$$X^3\left[f_0h_0 - \frac{\partial V}{\partial q}\right] = vE(u, v). \tag{4.68}$$

Because the left-hand side of this equation is independent of v, it is appropriate to introduce an arbitrary function of u, $W(u)$, by

$$E(u, v) = -\frac{W'(u)}{v}. \tag{4.69}$$

Then the admissible functions $\partial V/\partial q$ are given by

$$\frac{\partial V}{\partial q} = f_0(a, t)\left(\frac{c_1u - c_0c_2}{c_1^2 + c_2^2}\right) + \frac{W'(u)}{(c_2 - c_1a)^3}; \tag{4.70}$$

and the admissible potentials are given by

$$V(q, t) = \frac{f_0(a, t)}{(c_2 - c_1a)}\left[\frac{1}{2}c_1q^2 - c_0q\right] + \frac{W(u)}{(c_1^2 + c_2^2)(c_2 - c_1a)^2}, \tag{4.71}$$

where we have chosen the irrelevant additive function of t in such a way as to simplify the expression. Reiterating, we point out that c_0, c_1, and c_2 are arbitrary constants such that c_1 and c_2 are not both zero, $f_0(a, t)$ and $W(u)$ are arbitrary functions, and $a(t)$ is any function satisfying

$$\ddot{a} = f_0(a, t). \tag{4.72}$$

The expression for the admissible potentials may be viewed in two ways. The first is constructive; i.e., given $f_0(a, t)$, (4.72) may be solved for a and the potentials compatible with that choice of f_0 and a deduced. The second is eliminative; i.e., one can ask whether a given potential is in the class of admissible potentials.

V. THE TRANSFORMED HAMILTONIAN AND THE INVARIANT

We may now proceed to obtain the transformed Hamiltonian, the canonical transformation, and the invariant associated with the admissible potentials. From the results of the previous section, (3.34) can be written as

$$(c_1^2 + c_2^2)v\frac{\partial R_1}{\partial u} - W'(u)\frac{\partial R_1}{\partial v} = vf(u, v). \tag{5.1}$$

Since c_1 and c_2 were introduced as the coefficients in a relation expressing linear dependence, we may, and do, normalize them according to

$$c_1^2 + c_2^2 = 1. \tag{5.2}$$

Then (5.1) takes the simple form

$$v\frac{\partial R_1}{\partial u} - W'(u)\frac{\partial R_1}{\partial v} = vf(u, v). \tag{5.3}$$

The right-hand side of (5.3) is related to the transformed Hamiltonian by

$$\frac{\partial K}{\partial P} = \frac{vf(u, v)}{(c_2 - c_1a)^2}. \tag{5.4}$$

We now choose

$$K(P, \rho) = \frac{P}{(c_2 - c_1a)^2}, \tag{5.5}$$

so that

$$vf(u, v) = 1. \tag{5.6}$$

There is no loss of generality with this choice because we shall be able to find the canonical transformation corresponding to it. The equation for $R_1(u, v)$ is now

$$v\frac{\partial R_1}{\partial u} - W'(u)\frac{\partial R_1}{\partial v} = 1, \tag{5.7}$$

whose solution is

$$R_1(u, v) = \frac{v}{|v|}\int^u \frac{du'}{\{2[\alpha(u, v) - W(u')]\}^{1/2}} + T[\alpha(u, v)], \tag{5.8}$$

where

$$\alpha(u, v) = \tfrac{1}{2}v^2 + W(u) \tag{5.9}$$

and T is an arbitrary function.

The invariant may be calculated using (3.26) and (3.27) which, written in terms of $\Gamma_1(R_1, u)$ and $R_1(u, v)$, are

$$\frac{\partial \Gamma_1}{\partial u} = -1\bigg/\frac{\partial R_1}{\partial v}, \tag{5.10}$$

$$\frac{\partial \Gamma_1}{\partial R_1} = v\bigg/\frac{\partial R_1}{\partial v}, \tag{5.11}$$

where v and $\partial R_1/\partial v$ are to be expressed in terms of u and R_1 by solving (5.8) for $v = v(u, R_1)$ in terms of u and R_1. Let

$$J(u, v) = \Gamma_1[R_1(u, v), u]. \tag{5.12}$$

From (5.11),

$$\frac{\partial J}{\partial v} = \frac{\partial \Gamma_1}{\partial R_1}\frac{\partial R_1}{\partial v} = v, \tag{5.13}$$

which implies

$$J(u, v) = \tfrac{1}{2}v^2 + S(u). \tag{5.14}$$

Combining (5.10) and (5.11) we find

$$v\frac{\partial \Gamma_1}{\partial u} + \frac{\partial \Gamma_1}{\partial R_1} = 0, \tag{5.15}$$

which can be solved as follows. We first derive expressions for the derivatives of $v(u, R_1)$ by considering the identity

$$R_1 = R_1[u, v(u, R_1)]. \tag{5.16}$$

Differentiating the identity with respect to R_1 gives

$$\frac{\partial v}{\partial R_1} = 1\bigg/\frac{\partial R_1}{\partial v}. \tag{5.17}$$

Differentiating the identity with respect to u and using (5.7) to eliminate $\partial R_1/\partial u$ gives

$$v\frac{\partial v}{\partial u} = -1\bigg/\frac{\partial R_1}{\partial v} - W'(u). \tag{5.18}$$

Substituting (5.14) into (5.15), considering $v = v(u, R_1)$, then gives

$$-vW'(u) + vS'(u) = 0, \tag{5.19}$$

which we can solve by taking

$$S(u) = W(u). \tag{5.20}$$

Therefore, we have

$$J(u, v) = \Gamma_1(R_1, u) = \frac{1}{2} v^2 + W(u). \tag{5.21}$$

We now summarize our results for the canonical transformation in terms of the variables q, p, a, and $\dot a$. A Hamiltonian

$$H = \tfrac{1}{2}p^2 + \frac{f_0(a, t)}{(c_2 - c_1 a)} [\tfrac{1}{2}c_1 q^2 - c_0 q] + \frac{W(u)}{(c_2 - c_1 a)^2}, \tag{5.22}$$

where

$$u = \frac{q - c_0(c_1 + c_2 a)}{c_2 - c_1 a}, \tag{5.23}$$

c_0, c_1, and c_2 are arbitrary constants such that $c_1^2 + c_2^2 = 1$, $f_0(a, t)$ and $W(u)$ are arbitrary functions, and $a(t)$ is any function satisfying

$$\ddot a = f_0(a, t), \tag{5.24}$$

is transformed to a new Hamiltonian

$$K(P, a) = P/(c_2 - c_1 a)^2 \tag{5.25}$$

by the canonical transformation

$$Q(q, p, a, \dot a) = \frac{v}{|v|} \int^u \frac{du'}{\{2[\alpha(u, v) - W(u')]\}^{1/2}} + T[\alpha(u, v)], \tag{5.26}$$

$$P(q, p, a, \dot a) = \tfrac{1}{2}[(c_2 - c_1 a)p + \dot a(c_1 q - c_0)]^2 + W(u), \tag{5.27}$$

where

$$\alpha(u, v) = \tfrac{1}{2}v^2 + W(u), \tag{5.28}$$

$$v = (c_2 - c_1 a)p + \dot a(c_1 q - c_0), \tag{5.29}$$

and T is an arbitrary function. The equations of motion for Q and P are

$$\frac{dQ}{dt} = \frac{\partial K}{\partial P} = \frac{1}{(c_2 - c_1 a)^2}, \tag{5.30}$$

$$\frac{dP}{dt} = -\frac{\partial K}{\partial Q} = 0. \tag{5.31}$$

Therefore, $P(q, p, a, \dot a)$ is an exact, in general explicitly time-dependent, invariant of the motion induced by H.

VI. DISCUSSION

In the preceding sections we have investigated Hamiltonians of the type

$$H = \tfrac{1}{2}p^2 + V(q, t) \tag{6.1}$$

and determined a class for which we could find an exact invariant. To determine this class we postulated the existence of a canonical transformation

$$Q = Q(q, p, \rho, \dot\rho), \quad P = P(q, p, \rho, \dot\rho), \tag{6.2}$$

in which time dependence occurs only through the function $\rho(t)$ and its derivative. It was also assumed that $\ddot\rho$ was not identicaly zero. The transformed Hamiltonian was to take the form

$$K = K(P, \rho), \tag{6.3}$$

so that the invariant would be the transformed momentum, P. Apart from the postulates above, we worked within the framework of the theory of canonical transformations in Hamiltonian mechanics. Although the particular approach taken in the use of canonical transformations was unconventional in that the generating function was a function of the original canonical variables and time only, the possibility of such an approach is inherent in at least one published work.[20] However, in that work there is no suggestion of applying such a generating function in a practical context.

Toward the end of our calculations it became apparent that there was no loss of generality in taking the transformed Hamiltonian as

$$K = P/[c_2 - c_1 a(t)]^2, \tag{6.4}$$

where $a(t)$ is a function of t only and c_1 and c_2 are constants. Thus the transformation is essentially one to action-angle variables. Indeed the transformation may be put exactly into the context of a transformation to action-angle variables if a generalized canonical transformation (cf. Ref. 22) is used. If in addition to (6.2) we introduce a new time variable

$$\tau = \tau(t) = \int^t [c_2 - c_1 a(t')]^{-2} dt', \tag{6.5}$$

so that now

$$K = K(P) = P, \tag{6.6}$$

then the invariant P is the action.

At this stage it is instructive to see how the results obtained here apply to some simple problems. For the first problem we consider that well-known paradigm, the time-dependent linear oscillator with Hamiltonian

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2(t)q^2. \tag{6.7}$$

Comparing this with (5.22) it is obvious that $W(u)$ is at most quadratic in u and we write it as

$$W(u) = \tfrac{1}{2}\alpha u^2 + \beta u + \gamma, \tag{6.8}$$

α, β, and γ being constants. Equating coefficients of like powers of q in (6.7) and (5.22) [with W as given in (6.8)], it is apparent that β, γ, and c_0 are zero. This leaves

$$-c_1 f_0(a, t) = g_0(\rho, t) = -\omega^2(t)\rho + \alpha/\rho^3, \tag{6.9}$$

where we have related ρ and a by

$$\rho = c_2 - c_1 a. \tag{6.10}$$

Thus $\rho(t)$ is a solution of the second-order equation

$$\ddot\rho + \omega^2(t)\rho = \alpha/\rho^3. \tag{6.11}$$

The constant α is open to choice. In view of the form of (6.11) we set it as one or zero to obtain

$$\ddot\rho + \omega^2(t)\rho = 1/\rho^3 \tag{6.12}$$

or

$$\ddot\rho + \omega^2(t)\rho = 0. \tag{6.13}$$

The solutions to these two equations are related to one another. The solutions of (6.12) can be expressed in terms of two linearly independent solutions of (6.13) [see the second work listed under Ref. 3]; however, for our purposes it is more convenient to take $\rho(t)$ as a particular solution of (6.12).

Then (6.13) has the solution set $\{\rho \cos\tau, \rho \sin\tau\}$ where τ is given by (6.5). We may use ρ, $\rho \cos\tau$ and $\rho \sin\tau$ to obtain three first integrals from (5.27). They are

$$I_1 = \tfrac{1}{2}[(\rho p - \dot\rho q)^2 + q^2/\rho^2], \qquad (6.14)$$

$$I_2 = \tfrac{1}{2}[\rho p \cos\tau - (\dot\rho \cos\tau - \rho^{-1} \sin\tau) q]^2, \qquad (6.15)$$

$$I_3 = \tfrac{1}{2}[\rho p \sin\tau - (\dot\rho \sin\tau + \rho^{-1} \cos\tau) q]^2. \qquad (6.16)$$

The first of these is the invariant as given directly by (5.27) with $\alpha = 1$ and can be found in Ref. 3. The second and third are obtained by substituting $\rho \cos\tau$ and $\rho \sin\tau$ for ρ in (5.27) with $\alpha = 0$. It will be observed that

$$I_1 = I_2 + I_3, \qquad (6.17)$$

i.e., these three integrals are not linearly independent. The existence of three linearly independent quadratic first integrals is known from the work of Lutsky,[9] which uses Noether's theorem, and of Leach (see the first work listed under Ref. 9), which uses the method of the Lie theory of extended groups. We may reconstruct their result as follows. From the expressions for I_2 and I_3 it is apparent that

$$I_4 = (q/\rho) \cos\tau - (\rho p - \dot\rho q) \sin\tau \qquad (6.18)$$

and

$$I_5 = (q/\rho) \sin\tau + (\rho p - \dot\rho q) \cos\tau \qquad (6.19)$$

are also first integrals of the motion. Since

$$[I_4, I_5] = 1, \qquad (6.20)$$

I_4 and I_5 are functionally independent and may be taken as the two functionally independent integrals of the system. It is evident that the following are three linearly independent quadratic first integrals:

$$J_1 = I_4^2 + I_5^2 = I_1, \qquad (6.21)$$

$$J_2 = I_4 I_5 = [I_1, I_2] = [I_3, I_1], \qquad (6.22)$$

$$J_3 = I_4^2 - I_5^2 = [J_1, J_2] = I_3 - I_2. \qquad (6.23)$$

The integrals J_1, J_2, and J_3 are in the form given in the work cited.

This type of result is usual for linear systems, but not for nonlinear systems. It is interesting to observe that the three linearly independent quadratic first integrals given in (6.14)–(6.16) are, as it were, on an equal footing in our treatment, just as they are in the Lie and Noether treatments. Any one of (6.14)–(6.16) could be taken as the transformed momentum and could be obtained by canonical transformation from the original Hamiltonian (6.7).

As our prime motivation for this work was to find nonlinear time-dependent systems for which an exact first integral can be determined, we consider as another example the system with Hamiltonian

$$H = \tfrac{1}{2} p^2 + \tfrac{1}{2} q^2 + \tfrac{1}{3} B(t) q^3. \qquad (6.24)$$

This has recently been treated using the Lie method applied to the corresponding Newtonian equation of motion.[23] Comparing (6.24) with (5.22), $W(u)$ must be a cubic polynomial,

$$W(u) = \tfrac{1}{3}\alpha u^3 + \tfrac{1}{2}\beta u^2 + \gamma u + \delta, \qquad (6.25)$$

α, β, γ and δ being constants. Clearly, c_0, γ and δ are zero. Then

$$\ddot\rho + \rho = \beta / \rho^3, \qquad (6.26)$$

$$B(t) = \alpha / \rho^5. \qquad (6.27)$$

If we set $\beta = 0$, $B(t)$ takes the form

$$B(t) = K_1 \cos^{-5}(t + \epsilon_1), \qquad (6.28)$$

and, for $\beta = 1$,

$$B(t) = K_2[C + (C^2 - 1)^{1/2} \cos(2\tau + \epsilon_2)]^{-5/2}, \qquad (6.29)$$

where K_1, K_2, $C\,(>1)$, ϵ_1, and ϵ_2 are constants. In both cases, for a given $B(t)$ of the form (6.28) or (6.29), $\rho(t)$ has been found and can only be used to determine one invariant. The results given here are in accordance with the results obtained using the Lie method.

In the two examples considered above we have started with a given Hamiltonian with a time-dependent potential and determined whether it fits in with the permissible form given by (5.22). In the case of the time-dependent linear oscillator, no restriction was placed on $\omega^2(t)$. For the anharmonic oscillator, an invariant was found only if $B(t)$ is of a particular form, viz., that given by (6.28) or (6.29). As we have already remarked (see the final comments in Sec. IV), the results obtained in this paper may be viewed in two ways. The first, as in the two examples, is to test a given potential to determine whether it is of the permitted form and, if so, then to construct the invariant. The second viewpoint is to deduce classes of potentials for which a first integral can be found from these results. This in effect reduces to a choice of the function $a(t)$ and the constants c_0, c_1, and c_2. To take a simple example of this, suppose

$$c_2 - c_1 a(t) = -(1 + t^2), \quad c_0 = 0. \qquad (6.30)$$

The invariant is then

$$I(q, p, t) = \tfrac{1}{2}[(1 + t^2) p - 2t q]^2 + W\left(\frac{q}{1 + t^2}\right) \qquad (6.31)$$

and the class of Hamiltonians which admit such an invariant is given by

$$H(q, p, t) = \tfrac{1}{2} p^2 + \frac{q^2}{1 + t^2} + \frac{1}{(1 + t^2)^2} W\left(\frac{q}{1 + t^2}\right). \qquad (6.32)$$

Given the formulas presented in this paper, it is a straightforward exercise to apply them to a particular problem. This in itself represents an advantage over the Lie and Noether methods, for both of which it is necessary to solve a set of partial differential equations to determine the generators of symmetry transformations for each particular problem. We have, in effect, been able to include most of the computationl work within the general theory. Admittedly we do not obtain the generators of symmetry transformations and so have no knowledge of any group-theoretic properties of the problem under consideration. However, as is well known, one-dimensional linear systems all exhibit SL(3, R) symmetry and nonlinear one-dimensional systems possess at most only one generator of a symmetry transformation. Thus nothing has really been lost. Indeed the practice of searching for first integrals of the motion *via* symmetry groups, when only the former are of interest, has received recently some adverse criticism as being unnecessarily circuitous.[14]

Another method for the investigation of invariants, which has been revived and developed by Ray and Reid,[12] is found in the study of Ermakov systems.[13] The results which they have obtained are contained in the results given in this paper. Their work is continuing,[24] and it will be of some interest to see how it develops.

The application of our results to the corresponding quantum mechanical problem is immediate. An intermediate stage in the progress from the original canonical coordinates q, p to the transformed coordinates Q, P may be written as a time-dependent linear point transformation to coordinates q', p' given by

$$q' = \frac{q - c_0(c_1 + c_2 a)}{c_2 - c_1 a}, \tag{6.33}$$

$$p' = (c_2 - c_1 a) p + \dot{a}(c_1 q - c_0). \tag{6.34}$$

The Hamiltonian H given by (5.22) is transformed to

$$H' = \frac{\frac{1}{2}p'^2 + W(q')}{(c_2 - c_1 a)^2} \tag{6.35}$$

and, under the change of time scale given by (6.5), (6.35) is equivalent to

$$\tilde{H} = \tfrac{1}{2}p'^2 + W(q'). \tag{6.36}$$

To the extent that a wavefunction for the Schrödinger equation for (6.36) can be obtained, a wavefunction for the original problem may be obtained in the same way as for the time-dependent oscillator (cf. Ref. 25).

It is appropriate to ask whether it is possible to find time-dependent potentials of a more general form than those presented here for which a first integral can also be found. To obtain the results given here we assumed a transformation of the form

$$Q = Q(q, p, \rho, \dot{\rho}), \quad P = P(q, p, \rho, \dot{\rho}) \tag{6.37}$$

and a transformed Hamiltonian of the form

$$K = K(P, \rho). \tag{6.38}$$

One generalization would be to allow more time-dependent parameters, assuming

$$Q = Q(q, p, \rho_1, \dot{\rho}_1, \rho_2, \dot{\rho}_2, \ldots),$$

$$P = P(q, p, \rho_1, \dot{\rho}_1, \rho_2, \dot{\rho}_2, \ldots). \tag{6.39}$$

We hope to report on this possibility in the future.

ACKNOWLEDGMENTS

One of us (P.G.L.L.) wishes to thank the Los Alamos National Laboratory for its hospitality during a visit when part of this work was done and the Department of Mathematics, La Trobe University, for a travel grant which made the visit possible.

[1] Cf. E. D. Courant and H. D. Snyder, Ann. Phys. **3**, 1 (1958); E. W. Seymour, R. B. Leipnik, and A. F. Nicholson, Austral. J. Phys. **18**, 883 (1965).
[2] This is a problem of more than academic interest; see E. A. Poe, "The Pit and the Pendulum," Broadway J., **1** (1865).
[3] H. R. Lewis, Jr., Phys. Rev. Lett. **18**, 510 (1967) and Erratum, Phys. Rev. Lett. **18**, 636 (1967); H. R. Lewis, Jr., J. Math. Phys. **9**, 1976 (1968); H. R. Lewis, Jr. and W. B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969); and H. R. Lewis, Jr., Phys. Rev. **172**, 1313 (1968).
[4] M. Kruskal, J. Math. Phys. **3**, 806 (1962).
[5] W. Sarlet, Ann. Phys. **92**, 232 (1975); *ibid.* **92**, 248 (1975); W. Sarlet, *Proceedings of* ICNO'*75* (*International Conference on Non-Linear Oscillators, 1975*), edited by G. Schmidt (Akademie, Berlin, 1977).
[6] P. G. L. Leach, J. Austral. Math. Soc. Ser. B **20**, 97 (1977); P. G. L. Leach, J. Math. Phys. **18**, 1608 (1977).
[7] P. G. L. Leach, J. Math. Phys. **20**, 86 (1979).
[8] This is the essential result of P. G. L. Leach, J. Math. Phys. **20**, 96 (1979).
[9] P. G. L. Leach, J. Math. Phys. **21**, 300 (1980); P. G. L. Leach, J. Phys. A **13**, 1991 (1980); P. G. L. Leach, J. Austral. Math. Soc. Ser. B **22**, 12 (1980); M. Lutsky, J. Phys. A **11**, 249 (1978); J. R. Ray and J. L. Reid, J. Math. Phys. **20**, 2054 (1979).
[10] G. E. Prince and C. J. Eliezer, J. Phys. A **13**, 815 (1980).
[11] G. E. Prince and C. J. Eliezer, "On the Lie Symmetries of the Classical Kepler Problem" (Preprint AM-79:06, Department of Applied Mathematics, La Trobe University, Bundoora, 3083, Australia).
[12] J. L. Reid and J. R. Ray, J. Math. Phys. **21**, 1583 (1980); J. R. Ray, Phys. Lett. A **78**, 4 (1980); J. R. Ray, Lett. Nuovo Cimento **27**, 424 (1980) and the references cited therein.
[13] V. P. Ermakov, Univ. Izv. Kiev **20**, 1 (1880).
[14] W. Sarlet and L. Y. Bahar, "Quadratic Integrals for Linear Nonconservative Systems and their Connection with the Inverse Problem of Lagrangian Mechanics" (Preprint RUGsam 80-16, Septemeber 1980, Instituut voor Theoretische Mechanica, Rijksuniversiteit Gent. B-9000 Gent, Belgium).
[15] C. J. Eliezer and A. Gray, SIAM J. Appl. Math. **30**, 463 (1976).
[16] H. R. Lewis and K. R. Symon (in preparation).
[17] E. Kanai, Prog. Theoret. Phys. **5**, 440 (1968); see also P. Caldriola, Nuovo Cimento **8**, 393 (1941).
[18] Cf. H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 1950); L. A. Pars, *A Treatise on Analytical Dynamics* (Heinemann, London, 1965); E. J. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed. (Cambridge U. P., Cambridge, 1937).
[19] Cf. Leach, Ref. 7 above.
[20] See also E. C. G. Sudarshan and N. Mukunda, *Classical Mechanics: A Modern Perspective* (Wiley, New York, 1974), pp. 34-36.
[21] Cf. J. W. Leech, *Classical Mechanics* (Metheun, London, 1965), p. 70, Eq. (7.6).
[22] Cf. A Munier, J. R. Burgan, M. Feix, and E. Fijalkow, "Schrödinger Equation with Time-Dependent Boundary Conditions" (preprint, Commissariat a l'Energie Atomique, Centre d'Etudes de LIMEIL, Villeneuve St. Georges, France, September, 1980).
[23] P. G. L. Leach, J. Math. Phys. **22**, 465 (1981).
[24] J. R. Ray (private communication), September 1980.
[25] P. G. L. Leach, J. Math. Phys. **18**, 1902 (1977).

Current responses of first and second order in a collisionless plasma. I. Stationary plasma

Jonas Larsson
Department of Plasma Physics, Umea University, S-901 87 Umea, Sweden

(Received 23 December 1980; accepted for publication 21 August 1981)

The current responses of first and second order, due to an electromagnetic perturbation of a stationary but otherwise arbitrary solution of the relativistic Vlasov–Maxwell equations, are studied. In particular the symmetries leading to the (approximative, due to wave-particle interaction) conservation of wave energy for an *inhomogeneous* plasma are considered. Thereby we clarify the mathematical structure of certain previously derived formulas for the response operators and also make these more readily accessible for applications.

PACS numbers: 52.35.Fp, 52.40.Fd

I. INTRODUCTION

In a homogeneous plasma, mathematically described by the Vlasov–Maxwell equations, we may calculate the linear and quadratic conductivity tensors[1,2] in the form $(\kappa = (\omega, k))$,

$$\sigma_{ij}(\kappa) = i\int a_{ij}(\kappa,\mathbf{v})d\mathbf{v}, \tag{1.1}$$

$$\sigma_{ijk}(\kappa_1,\kappa_2) = i\int a_{ijk}(\kappa_1,\kappa_2,\mathbf{v})d\mathbf{v}, \tag{1.2}$$

where we have the symmetries

$$a_{ij}(\kappa_1,\mathbf{v}) = a_{ji}(\kappa_2,\mathbf{v}) \quad \text{for} \quad \kappa_1+\kappa_2=0 \tag{1.3}$$

and

$$\begin{aligned} a_{ijk}(\kappa_1,\kappa_2,\mathbf{v}) &= a_{kij}(\kappa_3,\kappa_1,\mathbf{v}) = a_{jki}(\kappa_2,\kappa_3,\mathbf{v}) = a_{ikj}(\kappa_2,\kappa_1,\mathbf{v}) \\ &= a_{jik}(\kappa_3,\kappa_2,\mathbf{v}) = a_{kji}(\kappa_1,\kappa_3,\mathbf{v}) \\ &\quad \text{for} \quad \kappa_1+\kappa_2+\kappa_3=0. \end{aligned} \tag{1.4}$$

Poles of the functions a_{ij} and a_{ijk} prevent the symmetries (1.3), (1.4) to be exactly inherited by σ_{ij} and σ_{ijk}. However, in many cases of interest, we have relations like (1.3), (1.4) approximately valid for the conductivity tensors and this means a considerable simplification as wave energy and momentum then are approximately conserved.

It is demonstrated in this paper how (1.1)–(1.4) may be generalized to the case of inhomogeneous plasmas. This will be achieved by means of explicit formulas. Our unperturbed plasma is stationary and may have two directions of homogeneity (for example, a homogeneous plasma in the half-space $z>0$), one direction of homogeneity (for example, a cylindrical plasma) or no such directions (e.g., a tokamak plasma). In order to discuss a typical symmetry result of this paper let us consider the last case.

A perturbation $\mathbf{A}(\mathbf{r})e^{i\omega t}$ of the electromagnetic vector potential (we here choose the gauge with vanishing scalar potential, although, in the following sections only gauge invariant expressions will appear) induces the linear current response $\delta\mathbf{J}_\omega[\mathbf{A}](\mathbf{r})e^{i\omega t}$. Defining $\langle\mathbf{A}_1,\mathbf{A}_2\rangle = \int\mathbf{A}_1^*(\mathbf{r})\cdot\mathbf{A}_2(\mathbf{r})d\mathbf{r}$ we make a suitable space of square integrable functions into a Hilbert space. Corresponding to (1.1) and (1.3) we obtain

$$\int\mathbf{A}_1(\mathbf{r})\cdot\delta\mathbf{J}_\omega[\mathbf{A}_2](\mathbf{r})d\mathbf{r} = \int a[\omega,\mathbf{A}_1,\mathbf{A}_2](\mathbf{r},\mathbf{v})d\mathbf{r}\,d\mathbf{v}, \tag{1.5}$$

$$a[\omega_1,\mathbf{A}_1,\mathbf{A}_2] = a[\omega_2,\mathbf{A}_2,\mathbf{A}_1], \quad \text{for} \quad \omega_1+\omega_2=0. \tag{1.6}$$

Thus if we may neglect pole contributions in (1.5)

$$\begin{aligned} \langle\mathbf{A}_1,\delta\mathbf{J}_\omega[\mathbf{A}_2]\rangle &= \int a[\omega,\mathbf{A}_1^*,\mathbf{A}_2]\,d\mathbf{r}\,d\mathbf{v} \\ &= \int a[-\omega,\mathbf{A}_2,\mathbf{A}_1^*]\,d\mathbf{r}\,d\mathbf{v} \\ &= \langle\delta\mathbf{J}_\omega[\mathbf{A}_1],\mathbf{A}_2\rangle \end{aligned} \tag{1.7}$$

and then $\delta\mathbf{J}_\omega$ is Hermitian.

The relations (1.2) and (1.4) may be generalized in a similar way, now involving the second-order current response $\delta\mathbf{J}_{\omega_1\omega_2}$,

$$\begin{aligned} &\int\mathbf{A}_3(\mathbf{r})\cdot\delta\mathbf{J}_{\omega_1\omega_2}[\mathbf{A}_1,\mathbf{A}_2](\mathbf{r})\,d\mathbf{r} \\ &\quad = \int a[\omega_1,\omega_2,\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3](\mathbf{r},\mathbf{v})\,d\mathbf{r}\,d\mathbf{v}, \end{aligned} \tag{1.8}$$

$$a[\omega_1,\omega_2,\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3] = a[\omega_\alpha,\omega_\beta,\mathbf{A}_\alpha,\mathbf{A}_\beta,\mathbf{A}_\gamma], \tag{1.9}$$

where $\{\alpha,\beta,\gamma\} = \{1,2,3\}$ and $\omega_1+\omega_2+\omega_3=0$.

In order to obtain (1.6) and (1.9) we have the following sufficient condition on the unperturbed Vlasov operator:

$$\begin{aligned} &\lim_{\eta\to 0+}\left((-i\omega+\eta)+\mathbf{v}\cdot\frac{\partial}{\partial\mathbf{r}}+q(\mathbf{E}_0(\mathbf{r})+\mathbf{v}\times\mathbf{B}_0(\mathbf{r}))\cdot\frac{\partial}{\partial\mathbf{p}}\right)^{-1} \\ &= \lim_{\eta\to 0+}\left((-i\omega-\eta)+\mathbf{v}\cdot\frac{\partial}{\partial\mathbf{r}}+q(\mathbf{E}_0(\mathbf{r})+\mathbf{v}\times\mathbf{B}_0(\mathbf{r}))\cdot\frac{\partial}{\partial\mathbf{p}}\right)^{-1}. \end{aligned} \tag{1.10}$$

It is easily demonstrated that (1.10) follows if all unperturbed particle orbits are periodic. This result has some interest in the fluid limit but for kinetic applications we certainly need to treat more general cases.

The two operators in (1.10) may be calculated in terms of the unperturbed orbits and we notice that the condition (1.10) means that such an orbit is in some sense symmetric in time around any of its points. The Poincaré recurrence theorem and the ergodic theorems of Birkhoff give some information in this direction.[3] In this paper, however, we have no ambition to derive (1.10) from results of general dynamics since this probably, if possible, would be a most difficult task. Instead we consider the case when the unperturbed orbits belong to a most natural and mathematically interesting class of functions, namely, the almost periodic functions

 0022-2488/82/010176-07$02.50

(discovered by H. Bohr[4]). The almost periodic functions have several attractive features, for example, the definition of almost periodicity is physically appealing, almost periodicity is connected with Liapunov stability,[3] almost periodic orbits have the necessary recurrence properties and the orbits of nearly integrable systems being conditionally periodic[5] are also almost periodic. By this last point it will follow that the symmetry result in Ref. 6 is contained in this paper.

It is instructive to consider what the poles of $a[\omega,\mathbf{A}_1,\mathbf{A}_2](\mathbf{r},\mathbf{v})$ in (1.5) look like [the poles in (1.8) may, of course, be investigated in a similar way]. The unperturbed orbit through $(\mathbf{r},\mathbf{v})$ being almost periodic determines a countable set of frequencies $\{\Omega_n(\mathbf{r},\mathbf{v})|n=1,2...\}$, we assume this set is closed under subtraction (otherwise we just include the necessary frequencies, in spite of the corresponding Fourier coefficients of the unperturbed orbit are zero). We then obtain the poles in (1.5) from factors $(\omega-\Omega_n(\mathbf{r},\mathbf{v}))^{-1}$ and this seems physically natural.

The notation and coordinate free formalism of Refs. 7 and 8 will be used without repeating the definitions. Thus all formulas will be covariant. Another advantage is that we may write formulas covering at the same time the three different cases with two, one, or zero spatial directions of homogeneity. In Sec. 2 general formulas for the first- and second-order current responses are given and in Sec. 3 these are specialized to the stationary case with an unspecified dimension on the space of homogeneity. In Sec. 4 the case of almost periodic orbits is considered.

2. GENERAL UNPERTURBED STATE

From (3.1)–(3.8) in Ref. 7 and (3.14) in Ref. 8 we obtain formulas for the linear and bilinear four-current responses $\delta J^{(1)}[\phi]$ and $\delta J^{(2)}[\phi,\phi]$ due to an electromagnetic perturbation with four-potential ϕ. The unperturbed state may be a quite arbitrary (space-time dependent) solution of the (relativistic, multicomponent) Vlasov–Maxwell equations in the presence of an external four-current. In the notations of Refs. 7 and 8 we obtain

Result 1: Take $\phi_0\in L^0(E,V)$ and $\phi_1,\phi_2\in L_0(E,V)$. Then

$$\int\phi_0\cdot\delta J^{(1)}[\phi_1]\,dP$$
$$=2^{-1}qc\int_{E\times S}f_0(P,u)[\delta x(1)\cdot\nabla_E(u\cdot\phi_0)+\delta u(1)\cdot\phi_0$$
$$+\delta x(0)\cdot\nabla_E(u\cdot\phi_1)+\delta u(0)\cdot\phi_1]\,dP\,du, \tag{2.1}$$

$$\int_E\phi_0\cdot\delta J^{(2)}[\phi_1,\phi_2]dP=\sum_{\substack{\alpha,\beta,\gamma=0\\ \alpha\neq\beta\neq\gamma\neq\alpha}}^{2}2^{-1}qc\int_{E\times S}f_0(P,u),$$

$$[6^{-1}\delta x(0)\otimes\delta x(1)\otimes\delta x(2)\vdots\nabla_E\otimes\nabla_E\otimes\nabla_E(u\cdot\Phi_0)$$
$$+2^{-1}\delta x(\alpha)\otimes\delta x(\beta)\otimes\delta u(\gamma)\vdots,$$
$$\nabla_E\otimes\nabla_E\otimes\Phi_0+2^{-1}\delta x(\alpha)\otimes\delta x(\beta):\nabla_E\otimes\nabla_E(u\cdot\phi_\gamma)$$
$$+\delta x(\alpha)\otimes\delta u(\beta):\nabla_E\otimes\phi_\gamma]dP\,du. \tag{2.2}$$

Here $\delta x(j)$ and $\delta u(j)$ are determined by the equations

$$D_0\delta x(j)=\delta u(j), \tag{2.3}$$

$$D_0\delta u(j)-qm_0^{-1}c^{-2}\nabla_E\wedge\Phi_0\cdot\delta u(j)-qm_0^{-1}c^{-2}$$
$$\times\delta x(j)\cdot\nabla_E(\nabla_E\wedge\Phi_0\cdot u)$$
$$=qm_0^{-1}c^{-2}\nabla_E\wedge\phi_j\cdot u \tag{2.4}$$

with the boundary conditions

$$\delta x(j),\ \delta u(j)\to 0\quad\text{towards the past if } j=1\text{ or }2, \tag{2.5}$$

$$\delta x(0),\ \delta u(0)\to 0\quad\text{towards the future.} \tag{2.6}$$

Remark 1: The results (2.1), (2.2) look perfectly symmetric with respect to perturbations of the indecies 0, 1 and 0, 1, 2, respectively. Note, however, that index 0 plays a particular role, this is seen from (2.5), (2.6) and $\phi_0\in L^0(E,V)$ while $\phi_1,\phi_2\in L_0(E,V)$.

Result 2: The function $\delta x(j)$ defined by (2.3)–(2.6) may be expressed in terms of D_0^{-1} (the inverse of the unperturbed Vlasov-operator D_0) as

$$\delta x(j)=-qm_0^{-1}c^{-2}\{\nabla_SD_0^{-1}(u\cdot\phi_j)$$
$$+D_0^{-1}(u\cdot D_0\nabla_SD_0^{-1}(u\cdot\phi_j))u\}, \tag{2.7}$$

where D_0^{-1} is subject to boundary conditions in accordance with (2.5) and (2.6).

Proof: We introduce the notations $\alpha=qm_0^{-1}c^{-2}$ and

$$\delta\tilde{x}=-\alpha\nabla_SD_0^{-1}(\phi\cdot u),\quad\delta\tilde{u}=-\alpha D_0\nabla_SD_0^{-1}(\phi\cdot u). \tag{2.8}$$

Let us in this proof assume that (2.7) and (2.3) defines δx and δu and prove that the relation (2.4) then is satisfied. From (2.8), (2.7), and (2.3) we obtain

$$\delta x=\delta\tilde{x}+D_0^{-1}(u\cdot\delta\tilde{u})u, \tag{2.9}$$

$$\delta u=\delta\tilde{u}+D_0(D_0^{-1}(u\cdot\delta\tilde{u})u). \tag{2.10}$$

It is straightforward to demonstrate by means of (2.9), (2.10), and $D_0=u\cdot\nabla_E+\alpha(\nabla_E\wedge\Phi\cdot u)\cdot\nabla_S$ that

$$D_0\delta u\ -\alpha\nabla_E\wedge\Phi_0\cdot\delta u-\alpha\delta x\cdot\nabla_E(\nabla_E\wedge\Phi_0\cdot u)$$
$$=D_0(\delta\tilde{u}+u\cdot\delta\tilde{u}u)-\alpha\nabla_E\wedge\Phi\cdot\delta\tilde{u}$$
$$-\alpha\delta\tilde{x}\cdot\nabla_E(\nabla_E\wedge\Phi_0\cdot u) \tag{2.11}$$

and thus the left-hand side in (2.4) has been expressed in terms of $\delta\tilde{x}$ and $\delta\tilde{u}$. The next step is to express the right-hand side of (2.11) in terms of Φ_0 and ϕ. From (2.8) we obtain

$$\delta\tilde{u}=-\alpha\nabla_S(\phi\cdot u)+\alpha(\nabla_E+u\ \ u\cdot\nabla_E)D_0^{-1}(\phi\cdot u)$$
$$-\alpha^2\nabla_E\wedge\Phi_0\cdot\nabla_SD_0^{-1}(\phi\cdot u) \tag{2.12}$$

and from (2.12)

$$\delta\tilde{u}+u\cdot\delta\tilde{u}\ \ u=-\alpha\phi+\alpha(\nabla_E-\alpha\nabla_E\wedge\Phi_0\cdot\nabla_S)D_0^{-1}(\phi\cdot u). \tag{2.13}$$

Substitution of (2.13) and (2.8) in the right-hand side of (2.11) yields after a substantial amount of algebra the result (2.4). Three identities that have been used in the derivations above are

$$\nabla_Su=I+u\otimes u,\quad D_0u=\alpha\nabla_E\wedge\Phi_0\cdot u, \tag{2.14}$$

$$D_0\nabla_S=\nabla_SD_0-(\nabla_E+uu\cdot\nabla_E)+\alpha\nabla_E\wedge\Phi_0\cdot\nabla_S. \tag{2.15}$$

Lemma 1: For $v\in V$ denote by T_v the translation operator defined on any function ϕ from E to some vector space by $(T_v\phi)(P)=\phi(P-v)$. Now if $T_vf_0=f_0$ [i.e., $f_0(P-v,u)=f_0(P,u)$] and $T_v\nabla_E\wedge\Phi_0=\nabla_E\wedge\Phi_0$ and if A is an operator (linear or nonlinear) determined only by the unperturbed state and defined on a space of functions $\{\phi\,|\phi:E\to V\}$, then $T_v(A(\phi))=A(T_v\phi)$.

Remark 2: Lemma 1 is somewhat trivial but it will still be useful to have it explicitly stated. The content of the lemma is perhaps most clearly seen if we represent the functions ϕ, $T_v\phi$, $A(\phi)$, $T_vA(\phi)$ by their graphs in $E \times V$.

3. STATIONARY UNPERTURBED STATE

We now consider the situation when the unperturbed state has one or several directions of homogeneity with at least one of these directions timelike.

An arbitrary event $0\in E$ is choosen and then to each event $P\in E$ there corresponds a vector $x\in V$ such that $P = 0 + x$. Often in notation we identify x and $P = 0 + x$ and write $x\in E$, etc.

Definition 1: The vector spaces V_h and V_i are defined by $V_h = \{$all homogenity directions of the unperturbed state$\} = \{a\in V \mid f_0(P + \lambda a, u)$ and $\nabla_E \wedge \Phi_0(P + \lambda a)$ are independent of the real parameter $\lambda\}$ and $V_i = V_h^{\perp} = \{a\in V \mid a\cdot b = 0$ for all $b\in V_h\}$.

Lemma 2: $V_h + V_i = V$ and each $a\in V$ has a unique decomposition $a = a_h + a_i$, where $a_h\in V_h$ and $a_i\in V_i$.

Proof: Choose a Lorentz coordinate system (e_0, e_1, e_2, e_3) such that $(e_0,\ldots,e_{m-1})$ span $V_h (m = 1, 2, 3, 4)$. This is possible to do since V_h contains a timelike vector (just elementary linear algebra). Then $(e_m,\ldots,e_3)$ span V_i. The lemma follows since each element in V, V_h, or V_i may be expressed as a unique linear combination from (e_0, e_1, e_2, e_3), $(e_0,\ldots,e_{m-1})$ or $(e_m,\ldots,e_3)$, respectively.

Definition 2: (a) Subscripts i or h on a vector in V is defined by $a_h + a_i = a$ for $a\in V$ and $a_h\in V_h$ and $a_i\in V_i$ (Lemma 2),

(b) $\nabla_E = \nabla_i + \nabla_h$ defines ∇_i and ∇_h in analogy to (a) above.

(c) For $\kappa\in V_h^+$ (the plus sign stands for complexification, cf. Ref. 8) we define $\nabla_\kappa = i\kappa + \nabla_i$.

(d) For $\kappa\in V_h^+$ we define

$$D_\kappa = u\cdot\nabla_\kappa + qm_0^{-1}c^{-2}(\nabla_E \wedge \Phi_0\cdot u)\cdot\nabla_S.$$

Remark 3: If $\mathrm{Im}\kappa$ is timelike (Im = imaginary part of) then we have a natural definition, including boundary conditions, for the inverse D_κ^{-1} of D_κ. This is easily seen by integrating along unperturbed orbits using that $D_\kappa A(x_i,u) = e^{-i\kappa\cdot x_h}D_0(e^{i\kappa\cdot x_h}A(x_i,u))$ and assuming that $A(x_i,u)$ and $B(x_i,u)$ is bounded on each particle orbit while solving the equation $D_\kappa A = B$ for A. For real κ, however, we must indicate if an infinitesimal imaginary part of κ is directed towards the past or the future. The notation $D_{\kappa-}^{-1}$ is used in the former and $D_{\kappa+}^{-1}$ in the latter case. The symmetry results we want to demonstrate are valid when $D_{\kappa-}^{-1} = D_{\kappa+}^{-1}$, this follows from explicit expressions below in this section. In the next section it will be investigated what sort of conditions on the unperturbed state is needed in order to imply $D_{\kappa-}^{-1} = D_{\kappa+}^{-1}$.

Definition 3: (a) $P(V_i,V^+)$ is a set of functions $\psi: V_i \rightarrow V^+$ such that $\int_{V_i}|a\cdot\psi(x_i)|^2dx_i < \infty$ for all $a\in V$. We sometimes denote this set of functions with P only.

(b) Let L be a Lorentz coordinate system (e_0, e_1, e_2, e_3) then we define $P_L = P_L(V_i,V^+)$ by $P_L = \{\psi\in P \mid e_0\cdot\psi = 0\}$.

(c) The bilinear form $\langle,\rangle$ on P is defined by $\langle\psi_1,\psi_2\rangle = \int_{V_i}\psi_1^*\cdot\psi_2\,dx_i$.

Remark 4: The bilinear form $\langle,\rangle$ is positive definite on P_L but not on P. We consider P_L to be a Hilbert space with scalar product $\langle,\rangle$. When P_L instead of P is considered this only means that we have choosen a particular electromagnetic gauge, namely, the radiation gauge in the frame L.

Definition 4: (a) For κ, κ_1, $\kappa_2\in V_h^+$ with imaginary parts directed towards the future, the linear operator $\delta J_\kappa: P(V_i,V^+) \rightarrow P(V_i,V^+)$ and the bilinear operator $\delta J_{\kappa_1\kappa_2}: P\times P \rightarrow P$ are defined by

$$\delta J_\kappa[\psi](x_i) = \delta J^{(1)}[\phi](x)\exp(-i\kappa\cdot x), \tag{3.1}$$

$$\delta J_{\kappa_1\kappa_2}[\psi_1,\psi_2](x_i) = \delta J^{(2)}[\phi_1,\phi_2](x) \times\exp(-i(\kappa_1+\kappa_2)\cdot x), \tag{3.2}$$

where

$$\phi(x) = \psi(x_i)\exp(i\kappa\cdot x), \tag{3.3}$$

$$\phi_j(x) = \psi_j(x_i)\exp(i\kappa_j\cdot x), \quad j = 1 \text{ or } 2. \tag{3.4}$$

(b) For $\kappa\in V_h^+$ with $\mathrm{Im}\kappa$ timelike the linear operators δx_κ and δu_κ on P are defined by

$$D_\kappa\delta x_\kappa[\psi] = \delta u_\kappa[\psi], \tag{3.5}$$

$$\begin{aligned} D_\kappa\delta u_\kappa[\psi] &- qm_0^{-1}c^{-2}\nabla_E\wedge\Phi_0\cdot\delta u_\kappa[\psi] \\ &- qm_0^{-1}c^{-2}\delta x_\kappa[\psi]\cdot\nabla_E(\nabla_E\wedge\Phi_0\cdot u) \\ &= qm_0^{-1}c^{-2}\nabla_\kappa\wedge\psi\cdot u. \end{aligned} \tag{3.6}$$

Remark 5: We have to show that the right-hand sides in (3.1) and (3.2) are independent of x_h, i.e., $v\in V_h \Rightarrow T_v$ acts as the identity operator on the right-hand sides of (3.1) and (3.2). Let us check (3.1). For $v\in V_h$ we have from Lemma 1 $T_v\delta J^{(1)}[\phi] = \delta J^{(1)}[T_v\phi]$. Thus

$$\begin{aligned} T_v(\delta J^{(1)}[\phi](x)\exp(-i\kappa\cdot x)) &= T_v(\delta J^{(1)}[\phi](x)), \\ T_v(\exp(-i\kappa\cdot x)) &= \delta J^{(1)}[T_v\phi](x)\exp(i\kappa\cdot v)\exp(-i\kappa\cdot x) \\ &= \delta J^{(1)}[\phi\exp(-i\kappa\cdot v)](x)\exp(i\kappa\cdot v)\exp(-i\kappa\cdot x) \\ &= \delta J^{(1)}[\phi](x)\exp(-i\kappa\cdot x). \end{aligned}$$

In the same way we may check (3.2) then we use

$$T_v\delta J^{(2)}[\phi,\phi](x) = \delta J^{(2)}[T_v\phi,T_v\phi](x). \qquad \text{(Lemma 1).}$$

Remark 6: Since $\mathrm{Im}\kappa$, $\mathrm{Im}\kappa_1$, and $\mathrm{Im}\kappa_2$ are directed towards the future in (a) above it follows that ϕ, ϕ_1, $\phi_2\in L_0(E,V)$ and they may thus be used as arguments of $\delta J^{(1)}$ and $\delta J^{(2)}$ as in (3.1) and (3.2).

Remark 7: The operators δx_κ and δu_κ defined by (3.5), (3.6) may alternatively be defined by

$$\delta x_\kappa[\psi](x_i) = \delta x^{(1)}[\phi](x)\exp(-i\kappa\cdot x), \tag{3.7}$$

$$\delta u_\kappa[\psi](x_i) = \delta u^{(1)}[\phi](x)\exp(-i\kappa\cdot x), \tag{3.8}$$

where $\phi(x) = \psi(x_i)\exp(i\kappa\cdot x)$ and $\kappa\in V_h^+$ with $\mathrm{Im}\kappa$ directed towards the future. For $\mathrm{Im}\kappa$ directed towards the past we just replace the operators $\delta x^{(1)}$ and $\delta u^{(1)}$ in (3.7) and (3.8) by $\delta x^{(1-)}$ and $\delta u^{(1-)}$ defined on $L^0(E,V)$ by $\delta x^{(1-)}[\phi_0] = \delta x(0)$ and $\delta u^{(1-)}[\phi_0] = \delta u(0)$ with $\delta x(0)$ and $\delta u(0)$ as in (2.3), (2.4), and (2.6). The right-hand side of (3.7) and (3.8) is independent of x_h and this follows as in remark 5. Now (3.5) and (3.6) is easily obtained from (2.3), (2.4), and (3.7), (3.8).

Remark 8: The operator δJ_κ may in a natural way be defined also for $\mathrm{Im}\kappa$ directed towards the past. Define the operator $\delta J^{(1-)}:L^0(E,V)\to L^0(E,V)$ so that $\delta J^{(1-)}[\phi]$ is the linear part of the "current response" due to ϕ and is subject to the boundary condition that it vanishes towards the future.

From (2.1) it easily follows that

$$\int_E \phi_0\cdot\delta J^{(1)}[\phi_1]dP=\int_E \phi_1\cdot\delta J^{(1-)}[\phi_0]dP, \tag{3.9}$$

for $\phi_0\in L^0(E,V)$ and $\phi_1\in L_0(E,V)$. Alternatively (3.9) may be used as the definition of $\delta J^{(1-)}$. Now define

$$\delta J_\kappa[\psi]=\delta J^{(1-)}[\phi]\exp(-i\kappa\cdot x) \tag{3.10}$$

for $\phi(x)=\psi(x_i)\exp(i\kappa\cdot x)$ with $\kappa\in V_h^+$ and $\mathrm{Im}\kappa$ directed towards the past.

It follows from (3.9) that $\langle\psi_0,\delta J_\kappa[\psi_1]\rangle = \langle\delta J_{\kappa^*}[\psi_0],\psi_1\rangle$ and thus $\Pi_L\circ(\delta J_\kappa+\delta J_{\kappa^*})$ is an Hermitian operator on P_L. Here $L=(e_0,e_1,e_2,e_3)$ denotes a Lorentz frame and the operator Π_L takes the spatial part of a vector with respect to this frame, i.e., $\Pi_L(v)=v+e_0\cdot ve_0$. Thus

$$\Pi_L\circ\delta J_\kappa=2^{-1}\Pi_L\circ(\delta J_\kappa+\delta J_{\kappa^*})+2^{-1}\Pi_L\circ(\delta J_\kappa-\delta J_{\kappa^*}) \tag{3.11}$$

expresses $\Pi_L\circ\delta J_\kappa$ as the sum of one Hermitian and one anti-Hermitian operator on P_L.

Result 3: (a) Take $\psi_0,\psi_1\in P(V_i,V^+)$ and $\kappa_0,\kappa_1\in V_h^+$ such that $\kappa_0+\kappa_1=0$ and $\mathrm{Im}\kappa_1$ directed towards the future. Then

$$\int_{V_i}\psi_0\cdot\delta J_{\kappa_1}[\psi_1]dx_i=2^{-1}qc\int_{V_i\times S}f_0(x_i,u)[\delta\tilde x(1)\cdot\nabla_{\kappa_0}(u\cdot\psi_0)+\delta\tilde x(0)\cdot\nabla_{\kappa_1}(u\cdot\psi_1)+\delta\tilde u(1)\cdot\psi_0+\delta\tilde u(0)\cdot\psi_1]\,dx_i\,du, \tag{3.12}$$

where

$$\delta\tilde x(j)=\delta x_{\kappa_j}[\psi_j],\quad \delta\tilde u(j)=\delta x_{\kappa_j}[\psi_j]. \tag{3.13}$$

(b) Take $\psi_0,\psi_1,\psi_2\in P(V_i,V^+)$ and $\kappa_0,\kappa_1,\kappa_2\in V_h^+$ such that $\kappa_0+\kappa_1+\kappa_2=0$ and $\mathrm{Im}\kappa_1$, $\mathrm{Im}\kappa_2$ are directed towards the future. Then

$$\begin{aligned}\int_{V_i}\psi_0\cdot\delta J_{\kappa_1,\kappa_2}[\psi_1,\psi_2]dx_i=2^{-1}qc\sum_{\substack{\alpha,\beta,\gamma=0\\ \alpha\neq\beta\neq\gamma\neq\alpha}}^{2}\int_{V_i\times S}f_0(P,u),\\
\{6^{-1}\delta\tilde x(0)\otimes\delta\tilde x(1)\otimes\delta\tilde x(2):\nabla_E\otimes\nabla_E\otimes\nabla_E(u\cdot\Phi_0)\\
+2^{-1}\delta\tilde x(\alpha)\otimes\delta\tilde x(\beta)\otimes\delta\tilde u(\gamma):\nabla_E\otimes\nabla_E\otimes\Phi_0\\
+2^{-1}\delta\tilde x(\alpha)\otimes\delta\tilde x(\beta):\\
\nabla_{\kappa_\gamma}\otimes\nabla_{\kappa_\gamma}(u\cdot\psi_\gamma)+\delta\tilde x(\alpha)\otimes\delta\tilde u(\beta):\nabla_{\kappa_\gamma}\otimes\psi_\gamma\}dx_i\,du,\end{aligned} \tag{3.14}$$

where we also use definition (3.13).

(c) Take $\kappa\in V_h^+$ with $\mathrm{Im}\kappa$ timelike and $\psi\in P$, then

$$\begin{aligned}\delta x_\kappa[\psi]=&-qm_0^{-1}c^{-2}\{\nabla_S D_\kappa^{-1}(u\cdot\psi)\\&+D_\kappa^{-1}(u\cdot D_\kappa\nabla_S D_\kappa^{-1}(u\cdot\psi))u\}.\end{aligned} \tag{3.15}$$

Definition 5: For $\kappa\in V_h$ (note $\mathrm{Im}\kappa=0$) define

$$D_{\kappa+}^{-1}=\lim_{\lambda\to0+}D_{\kappa+i\lambda e}^{-1}, \tag{3.16}$$

$$D_{\kappa-}^{-1}=\lim_{\lambda\to0+}D_{\kappa-i\lambda e}^{-1}, \tag{3.17}$$

where $e\in S$.

Corollary 1: If the limits in (3.16) and (3.17) exist and are equal we define for $\kappa\in V_h$

$$D_\kappa^{-1}=D_{\kappa\pm}^{-1} \tag{3.18}$$

and then (3.15) together with (3.5) defines δx_κ and δu_κ also for κ with $\mathrm{Im}\kappa=0$. Then, for $\kappa,\kappa_1,\kappa_2\in V_h$, the integrands of (3.12) and (3.14) are perfectly symmetric with respect to perturbation of (0,1) and (0,1,2), respectively.

Proof of Corollary 1: Due to (3.18) the only source of asymmetry is removed.

Corollary 2: If $D_{\kappa+}^{-1}=D_{\kappa-}^{-1}$ and if we may neglect pole contributions in (3.12) (with $\kappa=\kappa_1=-\kappa_0$, $\kappa\in V_h$) then $\Pi_L\circ\delta J_\kappa$ is an Hermitian operator on $P_L(V_i,V^+)$, (cf. remark 8).

Proof: By the assumptions

$$\int_{V_i}\psi_0\cdot\delta J_\kappa[\psi_1]dx_i=\int_{V_i}\psi_1\cdot\delta J_{-\kappa}[\psi_0]\,dx_i \tag{3.19}$$

thus

$$\begin{aligned}\langle\psi_0,\delta J_\kappa^{(1)}[\psi_1]\rangle&=\int_{V_i}\psi_0{}^*\cdot\delta J_\kappa^{(1)}[\psi_1]dx_i\\
&=\int_{V_i}\delta J^{(1)}_{-\kappa}[\psi_0{}^*]\cdot\psi_1dx_i=\int_{V_i}\delta J_\kappa^{(1)}[\psi_0]^*\cdot\psi_1dx_i\\
&=\langle\delta J_\kappa^{(1)}[\psi_0],\psi_1\rangle.\end{aligned} \tag{3.20}$$

Remark 9: The Fourier transform with respect to x_h of a function $G(x)$ is

$$G_\kappa(x_i)=\int_{V_h}G(x)\exp(-i\kappa\cdot x)dx_h \tag{3.21}$$

and the inverse transform

$$G(x)=(2\pi)^{-m}\int_{V_h}G_\kappa(x_i)\exp(i\kappa\cdot x)d\kappa, \tag{3.22}$$

where m is the dimension of V_h. In (3.21) we take $\kappa\in V_h$ or sometimes $\kappa\in V_h^+$ with timelike imaginary part and correspondingly we may have to choose some path of integration in V_h^+ in (3.22) [see (2.1), (2.2), and Remark 1–3 in Ref. 8].

Lemma 3: We have the following Fourier transforms

$$\int_{V_h}\delta J^{(1)}[\phi](x)\exp(-i\kappa\cdot x)dx_h=\delta J_\kappa[\phi_\kappa](x_i), \tag{3.23}$$

$$\begin{aligned}\int_{V_h}&\delta J^{(2)}[\phi,\phi](x)\exp(-i\kappa\cdot x)dx_h\\
&=(2\pi)^{-m}\int_{V_h}\delta J_{\kappa',\kappa-\kappa'}[\phi_{\kappa'},\phi_{\kappa-\kappa'}]\,d\kappa',\end{aligned} \tag{3.24}$$

$$\int_{V_h}\delta x^{(1)}[\phi](x)\exp(-i\kappa\cdot x)dx_h=\delta x_\kappa[\phi_\kappa], \tag{3.25}$$

$$\int_{V_h}\delta u^{(1)}[\phi](x)\exp(-i\kappa\cdot x)dx_h=\delta u_\kappa[\phi_\kappa], \tag{3.26}$$

where $\phi_\kappa(x_i)$ is the Fourier transform of $\phi(x)$ with respect to x_h.

Proof: Calculate the inverse transforms of (3.23)–(3.26). The results follows directly from (3.1), (3.2) and (3.7), (3.8).

Proof of result 3: Consider case (a). Take κ_0 and κ_1 as in this result. Take $\phi_1(x)=\psi_1(x_i)\exp(i\kappa_1\cdot x)$ and $\phi_0\in L^0(E,V)\cap L_0(E,V)$. Substitution in (2.1) then yields (3.12) by means of (3.23), (3.25), (3.26), (3.7), and (3.8). In a similar way one obtains (3.14) from (2.2).

4. ALMOST PERIODIC MOTION

The purpose of this section is to give, in terms of the unperturbed particle orbits, a sufficient condition for (3.18) in Corollary 1. Consider the unperturbed plasma in Sec. 3 and choose a vector $e_0 \in S \cap V_h$. Let $(x(t), u(t)) \in E \times S$ be the motion of a plasma particle with parameter t as e_0 time, i.e., ordinary time for an observer with e_0 as time axis (i.e., $e_0 \cdot x(t) = -ct$). Denote the projection of this orbit on $V_i \times S$ by $\Gamma_i = \{(x_i(t), u(t)) | t \in R\}$.

Result 4: Let A be a continuous function on $V_i \times S$ and κ a vector in V_h. If $x_i(t)$ and $u(t)$ are almost periodic functions (in the sence of H. Bohr[4]) and $x_h(t)$ is the sum of a linear and an almost periodic function, and

$$\overline{\lim_{\eta \to 0+}} |(D^{-1}_{\kappa + i\eta e_0} A)(x_i, u)| \quad \text{is bounded on} \quad \Gamma_i \tag{4.1}$$

then

$$(D^{-1}_{\kappa +} A)(x_i, u) = (D^{-1}_{k -} A)(x_i, u) \quad \text{on} \quad \Gamma_i. \tag{4.2}$$

Remark 10: Denote the module of the almost periodic part of $(x(t), u(t))$ by $\{\Omega_n | n \text{ interger}\}$. (Each almost periodic function determines a countable spectrum of Fourier exponents and the module is then the smallest set which contains this spectrum and is closed under subtraction.) According to the assumptions in Result 4 we may write

$$x_h(t) = cte_0 + t\mathbf{v}_d + \mathbf{r}_a(t), \tag{4.3}$$

where $e_0 \cdot \mathbf{v}_d = e_0 \cdot \mathbf{r}_a(t) = 0$ and $\mathbf{r}_a(t)$ is almost periodic. Define ω and $\mathbf{k}$ by

$$\kappa = c^{-1} \omega e_0 + \mathbf{k}. \tag{4.4}$$

It follows from the proof of result 4 below that

$$(D^{-1}_{\kappa \pm})(x_i(t), u(t)) \sim i \sum (\omega - \mathbf{k} \cdot \mathbf{v}_d - \Omega_n)^{-1} a_n \times \exp(i \Omega_n t), \tag{4.5}$$

where a_n is determined from

$$a(t) \sim \sum a_n \exp i(\Omega_n + \mathbf{k} \cdot \mathbf{v}_d - \omega) t \tag{4.6}$$

and $a(t)$ is given by (4.36). The assumption (4.1) is in general needed to prove (4.5). The reason for this is the small denominators in (4.5). The module $\{\Omega_n\}$ is indeed dense on R except when the almost periodic part of the motion is exactly periodic.

Remark 11: Each point (x, u) in the plasma determines a set $\Gamma(x, u) \subset E \times S$ consisting of all points of the particle orbit through (x, u). Together with $e_0 \in S$ the point (x, u) determines a module $\{\Omega_n(x, u)\}$ and a drift velocity $\mathbf{v}_d(x, u)$ [with $e_0 \cdot \mathbf{v}_d(x, u) = 0$] if the particle orbit parametrized in e_0 time have the properties assumed in Result 4. From the homogeneity in the V_h directions it follows that $\mathbf{v}_d(x, u) = \mathbf{v}_d(x_i, u)$ and $\{\Omega_n(x, u)\} = \{\Omega_n(x_i, u)\}$. Thus from (4.5) and (3.15) we observe that the poles in (3.12) and (3.14) are due to factors $[\omega - \mathbf{k} \cdot \mathbf{v}_d(x_i, u) - \Omega_n(x_i, u)]^{-1}$.

Remark 12: If a particle orbit have the properties assumed in result 4 when parametrized in e_0 time then this is true also in e-time where e is any element in $S \cap V_h$. This will now be demonstrated. Let $s = h(t)$ be the relation between e-time s and e_0-time t with respect to the particle orbit $x(t)$ so that $x(h^{-1}(s))$ is the particle orbit in e time, that is

$$h(t) = -c^{-1} e \cdot x(t). \tag{4.7}$$

We observe directly that $h(t)$ is the sum of a linear and an almost periodic function and that $h(t)$ is strictly increasing $[t_2 > t_1 \Rightarrow x(t_2) - x(t_1)$ future oriented $\Rightarrow e \cdot (x(t_2) - x(t_1)) < 0 \Rightarrow h(t_2) > h(t_1)]$. From Lemmas 5 and 6 below it is easy to see that it is now sufficient for us to show $h^{-1}(s)$ is uniformly continuous. We use the following simple result: There exist a $k > 0$ such that

$$|v \cdot e| > k\, |v \cdot e_0| \quad \text{for all timelike } v. \tag{4.8}$$

[Choose $k > 0$ such that $e - ke_0$ is future oriented. For $u \in S$ we then have $u \cdot (e - ke_0) < 0 \Rightarrow -ku \cdot e_0 < -u \cdot e \Rightarrow |u \cdot e| > k\, |u \cdot e_0| \Rightarrow (4.8)$].

Given $\epsilon > 0$, take $\delta = k\epsilon$; then

$$|h(t_2) - h(t_1)| < \delta \Rightarrow c^{-1} |e \cdot (x(t_2) - x(t_1))| < \delta \Rightarrow k\, |t_2 - t_1| < \delta \Rightarrow |t_2 - t_1| < \epsilon \tag{4.9}$$

and thus $h^{-1}(s)$ is uniformly continuous.

Lemma 4: Let $g(t)$, $t \in R^n$, be a continuous function and $h(s)$ an almost periodic function with values in R^n. Then $g(h(s))$ is almost periodic and the module of $g(h(s))$ is contained in the module of $h(s)$.

Proof: It is sufficient to prove that given an $\epsilon > 0$ there exists a trigonometric polynomial $a(s)$

$$a(s) = \sum a_n \exp(i \lambda_n s) \tag{4.10}$$

such that λ_n is an element in the module of $h(s)$ and

$$|g(h(s)) - a(s)| < \epsilon, \quad \text{for } s \in R. \tag{4.11}$$

First define $K \subset R^n$ by (choose any norm on R^n)

$$K = \{t \in R^n \,|\, |t| \leqslant \sup_{s \in R} |h(s)| + 1\} \tag{4.12}$$

then K is a compact set since $h(s)$ is bounded. Being continuous $g(t)$ is also uniformly continuous on K and thus we may choose a δ with $0 < \delta < 1$ such that for t_1, $t_2 \in K$ and $|t_1 - t_2| < \delta$,

$$|g(t_1) - g(t_2)| < \epsilon/2. \tag{4.13}$$

By Weierstrass approximation theorem there exists a polynomial $p(t)$ in n variables $t \in R^n$ such that

$$|g(t) - p(t)| < \epsilon/2, \quad t \in K. \tag{4.14}$$

Since $h(s)$ is almost periodic there exists a trigonometric polynomial $b(s) = \Sigma b_n \exp(i \lambda_n s)$ with λ_n from the module of $h(s)$ and

$$|h(s) - b(s)| < \delta < 1, \quad s \in R. \tag{4.15}$$

Now it easily follows from (4.13) and (4.15) that $a(s) = p(b(s))$ have the required properties.

Lemma 5: Let $s = h(t)$ be a strictly increasing function from R to R which is the sum of a linear and an almost periodic function. Then the inverse $t = h^{-1}(s)$ is also the sum of a linear and an almost periodic function if and only if $h^{-1}(s)$ is uniformly continuous.

Lemma 6: Let the real valued functions $g(s)$ be the sum of a linear and an almost periodic function and let $h(s)$ be an almost periodic function. Then $h(g(s))$ is almost periodic.

Proof of Lemmas 5 *and* 6: See, Ref. 9

Lemma 7: Let $a(s)$ be an almost periodic function and define

$$a_\eta(s)=e^{-\eta s}\int_{-\infty}^{s} e^{\eta s'}a(s')ds',\quad n>0, \tag{4.16}$$

$$a^\eta(s)=-e^{\eta s}\int_{s}^{\infty} e^{-\eta s'}a(s')ds',\quad \eta>0. \tag{4.17}$$

Then $a_\eta(s)$ and $a^\eta(s)$ are almost periodic with the same Fourier exponents as $a(s)$. Furthermore, the functions $a_\eta(s)$ and $a^\eta(s)$ are majorized[4] by $\eta^{-1}a(s)$.

Proof: Let τ be a translation number of $\eta^{-1}a(s)$ corresponding to $\epsilon>0$. Then, by definition

$$|\eta^{-1}a(s+\tau)-\eta^{-1}a(s)|<\epsilon,\ \text{ for } s\in R. \tag{4.18}$$

We easily obtain

$$a_\eta(s+\tau)-a_\eta(s)$$
$$=e^{-\eta s}\int_{-\infty}^{s} e^{\eta s'}(a(s'+\tau)-a(s'))ds' \tag{4.19}$$

and then by means of (4.18)

$$|a_\eta(s+\tau)-a_\eta(s)|<\epsilon \tag{4.20}$$

and thus $a_\eta(s)$ is almost periodic and majorized by $\eta^{-1}a(s)$. For arbitrary but fixed $\eta>0$ we may expand $a_\eta(s)$ and $a(s)$ in Fourier series

$$a(s)\sim\sum_n A_n\exp(i\lambda_n s), \tag{4.21}$$

$$a_\eta(s)\sim\sum_n B_n\exp(i\lambda_n S), \tag{4.22}$$

where $\{\lambda_n\}$ is the set of all Fourier exponents in $a(s)$ and $a_\eta(s)$. From (4.16)

$$\frac{d}{ds}(a_\eta(s))=-\eta a_\eta(s)+a(s) \tag{4.23}$$

and thus

$$a_\eta(s)\sim\sum(i\lambda_n+\eta)^{-1}A_n\exp(i\lambda_n s) \tag{4.24}$$

and so $a_\eta(s)$ have the same Fourier exponents as $a(s)$. All statements about $a_\eta(s)$ is now proved and $a^\eta(s)$ may be treated in a similar way. The result corresponding to (4.24) is

$$a^\eta(s)\sim\sum(i\lambda_n-\eta)^{-1}A_n\exp(i\lambda_n s). \tag{4.25}$$

Lemma 8: Let $a(s)$ be an almost periodic function. Then (a), (b) and (c) below are equivalent

(a) $\int_0^s a(s')ds'$ is a bounded function of $s\in R$.

(b) $\overline{\lim_{\eta\to0+}}|a_\eta(s)|$ is a bounded function of $s\in R$.

(c) $\lim_{\eta\to0+}a_\eta(s)=\lim_{\eta\to0+}a^\eta(s)$ is a bounded function of $s\in R$.

Proof: We prove (c)$\Rightarrow$(b)$\Rightarrow$(a)$\Rightarrow$(c). Here (c)$\Rightarrow$(b) is trivial and (b)$\Rightarrow$(a) easily follows from

$$\left|\int_0^s a(s')ds'\right|\leqslant|a_\eta(s)|+|a_\eta(0)|+\eta\int_0^s|a_\eta(s')|\,ds', \tag{4.26}$$

which we obtain by partial integration. From (a) and (4.21) it follows that

$$d(s)\sim\sum_n(i\lambda_n)^{-1}A_n\exp(i\lambda_n s) \tag{4.27}$$

is an almost periodic primitive function of $a(s)$. We will show that

$$\{d(s)-a_\eta(s)|\eta>0\} \tag{4.28}$$

is a majorizable set of functions and that

$$\lim_{T\to\infty}T^{-1}\int_0^T|d(s)-a_\eta(s)|^2ds\to0\quad\text{when } \eta\to0+. \tag{4.29}$$

Then it follows (Ref. 4, Sec. 73) that

$$a_\eta(s)\to d(s)\quad\text{when } \eta\to0+ \tag{4.30}$$

and in a similar way,

$$a^\eta(s)\to d(s)\quad\text{when } \eta\to0+ \tag{4.31}$$

and thus (c) is obtained from (a). By the identity

$$d(s)-a_\eta(s)=\eta d_\eta(s) \tag{4.32}$$

and Lemma 7 applied to $d(s)$ we obtain that the set (4.28) is majorized by $d(s)$. From (4.24), (4.21), and Parseval's equation for almost periodic functions we get

$$\lim_{T\to\infty}T^{-1}\int_0^T|d(s)-a_\eta(s)|^2ds$$
$$=\sum_n|A_n|^2\lambda_n^{-2}\eta^2(\eta^2+\lambda_n^2)^{-1}. \tag{4.33}$$

Given $\epsilon>0$ choose N so large that

$$\sum_{|n|>N}|A_n|^2\lambda_n^{-2}<\epsilon \tag{4.34}$$

and this is possible to do since the series in (4.34) is convergent [since (4.27) determines an almost periodic function]. Now the right-hand side in (4.33) is smaller than

$$\sum_{|n|\leqslant N}|A_n|^2\lambda_n^{-2}\eta^2(\eta^2+\lambda_n^2)^{-1}+\epsilon \tag{4.35}$$

and (4.29) easily follows.

Proof of Result 4: Define

$$a(t)=[|x'(t)\cdot x'(t)|]^{1/2}A(x_i(t),u(t))\exp[i\kappa\cdot x(t)]. \tag{4.36}$$

Then $a(t)$ is almost periodic with Fourier exponents in the set $\{\Omega_n-\omega+\mathbf{k}\cdot\mathbf{v}_d|\Omega_n$ is an element in the module of the almost periodic part of $(x(t),u(t))\}$ [see (4.4) and Lemma 4]. By integration along the unperturbed orbit $(x(t),u(t))$ we obtain (see Remark 13 below)

$$(D^{-1}_{\kappa+i\eta e_0}A)(x_i(t),u(t))=a_{\tilde\eta}(t)\exp[-i\kappa\cdot x(t)],\quad \tilde\eta=c\eta, \tag{4.37}$$

$$(D^{-1}_{\kappa-i\eta e_0}A)(x_i(t),u(t))=a^{\tilde\eta}(t)\exp[-i\kappa\cdot x(t)]. \tag{4.38}$$

The condition (4.1) implies by means of (4.37) that (b) in Lemma 8 is satisfied by $a(t)$ in (4.36). From this lemma and (4.37) and (4.38) we obtain (4.2).

Remark 13: Let $(x(t),u(t))$ be an unperturbed orbit, where t is e_0 time and let $(\tilde x(s),\tilde u(s))$ be the same orbit parametrized by arclength so that s/c is proper time for the particle. The relation between s and t is

$$\frac{ds}{dt}=[|x'(t)\cdot x'(t)|]^{1/2}. \tag{4.39}$$

The equation for the orbit is simplest in proper time

$$\tilde{x}'(s) = \tilde{u}(s), \tag{4.40}$$

$$\tilde{u}'(s) = qn_0^{-1}c^{-2}\nabla_E \wedge \Phi_0(\tilde{x}(s))\cdot\tilde{u}(s). \tag{4.41}$$

Let $B(x,u)$ be a function on $E \times S$ then

$$(D_0 B)(\tilde{x}(s),\tilde{u}(s)) = \frac{d}{ds} B(\tilde{x}(s),\tilde{u}(s)) \tag{4.42}$$

and so we may calculate D_0^{-1} by integration along unperturbed orbits.

5. SUMMARY

The first- and second-order conductivity tensors of a homogeneous Vlasov-plasma possess (approximately) certain well-known symmetry relations, which lead to the (approximative) conservation of wave energy. The purpose of this paper is to consider what happens when the plasma is *inhomogeneous*. Our starting point is the previously derived formulas (2.1)–(2.6). It is quite straightforward that (2.1)–(2.6) leads to the expected symmetries [i.e., (1.1)–(1.4)] when the plasma is homogeneous. However, in order to find the symmetries implied by (2.1)–(2.6) for an inhomogeneous plasma, essential use is made of result 2, which is new and give the solution of (2.3)–(2.6)—i.e., the Lagrangian pertubations $\delta x(j)$ and $\delta u(j)$—in terms of the inverse unperturbed Vlasov operator. This inverse operator may be calculated by integration along unperturbed particle orbits which thus now explicitly enters in the response operator formulas. Assuming a stationary unperturbed state and by Fourier transformation in time and eventual spatial homogeneity directions we get in Sec. 3 from (2.1)–(2.7) the condition $D_{\kappa+}^{-1} = D_{\kappa-}^{-1}$ for the wanted symmetries to be valid. In particular the fulfillment of this condition leads to an Hermitian (if we neglect pole contributions) dispersion operator (this follows from corollary 2) and, as will be seen in part II of this paper,[10] also to the Manley–Rowe relations in three wave interaction. It is now natural to search for conditions on the unperturbed particle orbits that leads to $D_{\kappa+}^{-1} = D_{\kappa-}^{-1}$. In Sec. 4 a class of almost periodic motions is considered and in result 4 we obtain a result of the wanted form.

Two reasons for the choice of studying almost periodic motions are (a) it is a sufficiently large class of motions to be physically interesting and (b) there is a good theory for generalized Fourier series of almost periodic functions. It is (b) that make the analysis in Sec. 4 possible.

[1]S. Ichimaru, *Basic Principles of Plasma Physics* (Benjamin, London, 1973).

[2]J. Larsson, J. Plasma Phys. **21**, 519 (1979).

[3]G. R. Sell, *Topological Dynamics and Ordinary Differential Equations* (Van Nostrand, London, 1971).

[4]H. Bohr, *Almost Periodic Functions* (Chelsea, New York, 1951).

[5]V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1978).

[6]S. Johnston and A. N. Kaufman, J. Plasma Phys. **22**, 105 (1979).

[7]J. Larsson, J. Math. Phys. **20**, 1321 (1979).

[8]J. Larsson, J. Math. Phys. **20**, 1331 (1979).

[9]H. Bohr, Collected Mathematical Works II, Danish Mathematical Society (Copenhagen 1952), paper C30 "über fastperiodische Bewegungen auf einem Kreis."—with B. Jessen.

[10]J. Larsson, J. Math. Phys. **23**, 183 (1982), companion paper.

Current responses of first and second order in a collisionless plasma. II. Three-wave interaction

Jonas Larsson
Department of Plasma Physics, Umea University, S-901 87 Umea, Sweden

(Received 12 March 1981; accepted for publication 21 August 1981)

The three-wave interaction in a possibly strongly inhomogeneous plasma is considered. Coupled mode equations are derived with coefficients expressed in terms of the response operators treated in part I of this paper.

PACS numbers: 52.35.Fp, 52.40.Fd

I. INTRODUCTION

The purpose of this part II is to derive the coupled mode equations for waves in an inhomogeneous plasma. The coefficients will be expressed in terms of the first-and second-order conductivity operators and thereby illustrate how the formulas and results in part I[1] may be used.

There are in general great qualitative differences between resonant wave interaction in a strongly inhomogeneous and a homogeneous plasma.[2] We will only consider the situation where a pure resonant three-wave interaction is possible. The requirements for this is now much stronger than in the homogeneous case. Still the equations derived will be valid in situations of considerable physical interest. For example, the resonant interaction of three surface modes may be considered.[3] The symmetries that were discussed in I are of great importance for the properties of the coupled mode equations. Actually we have to assume from the outset that the normal modes, to first order, are determined by Hermitian operators.

The class of three-wave interaction processes we are going to consider is the simplest possible where still the inhomogeneity of the plasma may be essential in the calculation of the coupling coefficients. Note that for a weakly inhomogeneous plasma we may use WKB analysis and obtain the mode coupling equations,[4] but the coefficients are then calculated for a homogeneous plasma. We take the unperturbed state of the plasma as in Sec. 3 (I). As in I we decompose the four-vector space V as $V = V_h + V_i$, where V_h contains all directions of homogeneity of the plasma and $V_i = V_h^{\perp}$. Correspondingly we uniquely write a four-vector x as $x = x_h + x_i$, where $x_h \in V_h$ and $x_i \in V_i$. Let the plasma have normal modes of the form

$$\psi(x_i)\exp(i\kappa\cdot x_h) + \text{complex conjugate}, \tag{1.1}$$

for $\kappa = \kappa_1, \kappa_2$, and κ_3, where

$$\kappa_1 + \kappa_2 + \kappa_3 = 0 \tag{1.2}$$

and $\kappa_1, \kappa_2, \kappa_3 \in V_h$. [In a Lorentz frame (e_0, e_1, e_2, e_3) we may write (1.2) as $\omega_1 + \omega_2 + \omega_3 = 0$ and $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ if we define $\kappa_j = c^{-1}\omega_j e_0 + \mathbf{k}_j$.] We may then consider the interaction between three wave packets of the form

$$\psi_j(x_h, x_i)\exp(i\kappa_j\cdot x) + \text{c.c.}, \quad j = 1, 2, 3. \tag{1.3}$$

Each wave packet in (1.3) is a superposition of normal modes with κ close to κ_j. The amplitude $\psi_j(x_h, x_i)$ has a slow variation in x_h in comparison with $\exp i\kappa_j\cdot x$ ($= \exp i\kappa_j\cdot x_h$ since $\kappa_j \in V_h$). In principle one could write down a self-consistent set of equations describing the development of $\psi_j(x_h, x_i), j = 1, 2, 3$ including the resonant wave interaction. However, already the linear evolution of the wave packets may be very complicated. In general we have little control of the change in the x_i dependence with time and additional assumptions are thus needed in order to get resonably simple equations. Two important mechanisms behind the complicated behavior in the general case may be distinguished. First, there may be several branches of normal modes represented in each wave packet. [The Hermitian operator $\Pi_L \circ H_\kappa$ defined in Sec. 2, determines to first order the branches of normal modes by dispersion relations $\lambda(\kappa) = 0$, where $\lambda(\kappa)$ is some eigenvalue of $\Pi_L \circ H_\kappa$ and continuous in κ.] Normal modes with κ close to κ_j and belonging to one single branch have an approximately common x_i dependence while there are in general no simple relation between normal modes from different branches. This means that we may have a fast and complicated linear development of the x_i dependence of a wave packet if several branches of normal modes are present. We now come to the second mechanism. Let us consider the development of the κ_j-wave packet if only normal modes from one single branch are present. Let $(\psi_j(x_i)\exp i\kappa_j\cdot x + \text{c.c.})$ be a normal mode of this branch. Then we have

$$\psi_j(x_h, x_i) = A_j(x_h, x_i)\,\psi_j(x_i), \tag{1.4}$$

where the complex valued function $A_j(x_h, x_i)$ varies slowly in x_h and x_i in comparison with $\exp(i\kappa_j\cdot x)$ and $\psi_j(x_i)$, respectively. When the inhomogeneity also is weak we have the situation in Ref. 4 and WKB analysis may be used. In the present paper we will consider a complementary case, involving a possibly strong inhomogeneity, where the typical wavelengths of the normal modes are of the same order as the size of the plasma. Then the slow variation of $A_j(x_h, x_i)$ in x_i effectively means that it is constant in x_i. Another situation when this may happen is when the normal modes are localized, take for example surface waves. Then the thickness of the layer close to the surface, where the wave fields are essentially different from zero, enters instead of the plasma size. In order to obtain the coupled mode equations given in the result and in Corollary 2 of the next section we thus need to assume that (A) only one branch of normal modes exists for κ close to κ_j for $j = 1, 2$, and 3, (b) $A_j(x_h, x_i)$ in (1.4) is independent of x_i.

The resonant three-wave interaction is then described by equations of the same form as in a homogeneous plasma. The inhomogeneity enters in the calculation of the coefficients in these equations.

2. THE COUPLED MODE EQUATIONS

We use the notation from part I and Ref. 5. The electromagnetic waves are governed by the wave equation for the perturbation ϕ of the 4-potential

$$\nabla_E\cdot(\nabla_E \wedge \phi) = -(\mu_0/\epsilon_0)^{1/2}(\delta J^{(1)}[\phi] + \delta J^{(2)}[\phi,\phi]). \quad (2.1)$$

Here ∇_E is the four-dimensional gradient operator on the event space. The second-order current response is included in (2.1) since we are considering the three-wave interaction. We Fourier transform (2.1) with respect to V_h and obtain

$$\nabla_\kappa\cdot(\nabla_\kappa \wedge \phi_\kappa) + (\mu_0/\epsilon_0)^{1/2}\delta J_\kappa[\phi_\kappa]$$
$$= -(\mu_0/\epsilon_0)^{1/2}(2\pi)^{-m}\int_{V_h}\delta J_{\kappa',\kappa-\kappa'}[\phi_{\kappa'},\phi_{\kappa-\kappa'}]d\kappa'. \quad (2.2)$$

As in I we have $\nabla_\kappa = i\kappa + \nabla_i, \phi_\kappa(x_i)$ is the Fourier transform of $\phi(x)$ and m is the dimension of V_h. The linear terms have been collected on the left-hand side in (2.2) and we rewrite these as

$$(H_\kappa + ih_k)\phi_\kappa. \quad (2.3)$$

The operators H_κ and h_κ are defined on [see definition 3 (I)] $P(V_i, V^+)$. Take $\kappa \in V_h$ and let $\kappa+ (\kappa-)$ denote that a future (past) oriented infinitesimal imaginary part is present in κ. We define [c.f. remark 8 (I) and definition 5 (I)]

$$H_\kappa\psi = \nabla_\kappa\cdot(\nabla_\kappa \wedge \psi) + (\mu_0/\epsilon_0)^{1/2}2^{-1}(\delta J_{\kappa+} + \delta J_{\kappa-}), \quad (2.4)$$

$$h_\kappa\psi = -(\mu_0/\epsilon_0)^{1/2}i2^{-1}(\delta J_{\kappa+} - \delta J_{\kappa-}). \quad (2.5)$$

The resonant particle contributions is contained in h_κ.

Convention: The " + " sign but never the " − " sign will be omitted after "κ" in most places below. We observe that "$\kappa+$" indicate that we consider the causal (i.e., the physical) response of the perturbation.

Lemma 1: Take $\kappa \in V_h$ and a Lorentz frame L. Then $\Pi_L \circ H_\kappa$ and $\Pi_L \circ h_\kappa$ are Hermitian operators on $P_L(V_i, V^+)$.

Proof: The operator Π_L is defined in remark 8(I) and takes the spatial part of a 4-vector with respect to L. From remark 8(I) it follows that only the Hermitian property of $\Pi_L(\nabla_\kappa\cdot(\nabla_\kappa \wedge \psi))$ remains to be demonstrated. We make use of the identity

$$\psi_1^*\cdot(\nabla_\kappa\cdot(\nabla_\kappa \wedge \psi_2)) + \nabla_i\cdot(\psi_1^*\cdot\nabla_\kappa \wedge \psi_2)$$
$$= \psi_2\cdot(\nabla_\kappa\cdot(\nabla_\kappa \wedge \psi_1))^* + \nabla_i\cdot(\psi_2\cdot(\nabla_\kappa \wedge \psi_1)^*), \quad (2.6)$$

where $\psi_1, \psi_2 \in P_L(V_i, V^+)$. By integrating (2.6) over V_i we obtain

$$\langle\psi_1, \nabla_\kappa\cdot(\nabla_\kappa \wedge \psi_2)\rangle = \langle(\nabla_\kappa\cdot(\nabla_\kappa \wedge \psi_1), \psi_2\rangle \quad (2.7)$$

and now the Hermitian property we wanted to show follows easily and Lemma 1 is proved.

Given $\kappa \in V_h$ we associate with each $\psi \in P(V_i, V^+)$ the 4-potential

$$\phi(x) = \psi(x_i)\exp(i\kappa\cdot x) + \text{c.c.} \quad (2.8)$$

and the corresponding electromagnetic field tensor $-\nabla_E \wedge \phi$. The necessary and sufficient condition for ψ_1 and ψ_2 in $P(V_i, V^+)$ to represent the same electromagnetic field is

$$\nabla_\kappa \wedge \psi_1 = \nabla_\kappa \wedge \psi_2. \quad (2.9)$$

Accordingly we make the following definition.

Definition: A function $f(\psi)$ of $\psi \in P(V_i, V^+)$ is gauge invariant with respect to $\kappa \in V_h$ if

$$\nabla_\kappa \wedge \psi_1 = \nabla_\kappa \wedge \psi_2 \Rightarrow f(\psi_1) = f(\psi_2). \quad (2.10)$$

Lemma 2: Take $\kappa \in V_h$ and a Lorentz frame $L = (e_0, e_1, e_2, e_3)$ such that $e_0\cdot\kappa \neq 0$. For each $\psi \in P(V_i, V^+)$ there exist a unique $\chi \in P_L(V_i, V^+)$ such that $\nabla_\kappa \wedge \psi = \nabla_\kappa \wedge \chi$. In terms of the electric field $(\mathbf{E}(x_i)\exp i\kappa\cdot x + \text{c.c.})$ in the Lorentz frame L due to ψ and κ we get $\chi(x_i) = -i(c/\omega)\mathbf{E}(x_i)$ with $\omega = -ce_0\cdot\kappa$.

Proof: The existence of χ follows by checking that $-(\nabla_\kappa \wedge \chi(x_i)\exp i\kappa\cdot x + \text{c.c.})$ with $\chi(x_i) = -i(c/\omega)\mathbf{E}(x_i)$ is equal to the electromagnetic field tensor due to κ and ψ. This is straightforward to do if we use Faraday's law.

The uniqueness of χ follows from $e_0\cdot(\nabla_\kappa \wedge \psi) = e_0\cdot(\nabla_\kappa \wedge \chi) = ie_0\cdot\kappa\chi$ and $e_0\cdot\kappa \neq 0$.

Remark 1: In Lemma 1 we consider operators defined on $P_L(V_i, V^+) \subset P(V_i, V^+)$. This is necessary in order to obtain Hermitian operators [see Remark 4 (I)]. From Lemma 2 we may observe that a particular gauge which uniquely determines the 4-potential as a very simple function of the electric field then has been chosen.

Result: Let ψ_j and κ_j be given for $j = 1, 2$, and 3 such that $\psi_j \in P(V_i, V^+)$ and $\kappa_j \in V_h$ and

$$H_{\kappa_j}\psi_j = 0 \text{ and } \kappa_1 + \kappa_2 + \kappa_3 = 0. \quad (2.11)$$

Assume that:

(i) To first order we may neglect h_{κ_j} in comparison with H_{κ_j}. Thus

$$\psi_j(x_i)\exp(i_{\kappa_j}\cdot x) + \text{c.c.} \quad (2.12)$$

are normal modes.

(ii) Any wave packet of normal modes with κ close to κ_j will, if an appropriate gauge is chosen, be proportional to $\psi_j(x_i)$ (see Remark 3 below). Then

$$\phi(x) = \sum_{j=1}^{3} A_j(x_h)\psi_j(x_i)\exp(i\kappa_j\cdot x) + \text{c.c.} \quad (2.13)$$

is approximately a solution of (2.1) provided the complex amplitudes $A_j(x_h)$ satisfy the mode coupling equation

$$[\nabla_h\langle\psi_3, H_\kappa\psi_3\rangle]_{\kappa=\kappa_3}\cdot\nabla_h A_3^* - \langle\psi_3, h_{\kappa_3}\psi_3\rangle A_3^*$$
$$= 2i(\mu_0/\epsilon_0)^{1/2}\langle\psi_3^*, \delta J_{\kappa_1+,\kappa_2+}[\psi_1, \psi_2]\rangle A_1A_2 \quad (2.14)$$

and the two equations obtained by permutating 1, 2, and 3. The first ∇_h in (2.14) acts on $\kappa \in V_h$ and the second on $x_h \in V_h$. We note some properties of the coefficients. The 4-vector $[\nabla_h\langle\psi_j, H_\kappa\psi_j\rangle]_{\kappa=\kappa_j}$ is gauge invariant with respect to κ_j. It is proportional to the group four velocity $u_{gj} \in V_h \cap S$ of wave j. In a Lorentz frame $L = (e_0, e_1, e_2, e_3)$ we have the relation

$$e_0\cdot[\nabla_h\langle\psi_j, H_\kappa\psi_j\rangle]_{\kappa=\kappa_j} = (\epsilon_0\omega_j)^{-1}W_j, \quad (2.15)$$

where $\omega_j = -ce_0\cdot\kappa_j$ is the frequency of wave j and $|A_j(x_h)|^2W_j$ is the over V_i integrated energy density of wave j. The quantity $\langle\psi_j, h_{\kappa_j}\psi_j\rangle$ is real and gauge invariant with respect to κ_j. It is proportional to the linear damping. Finally $\langle\psi_3^*, \delta J_{\kappa_1,\kappa_2}[\psi_1, \psi_2]\rangle$ is gauge invariant with respect to κ_1, κ_2, or κ_3 when considered as a function of ψ_1, ψ_2, or ψ_3, respectively.

Proof: See Secs. 3 and 4.

Corollary 1: If $D_{\kappa+}^{-1} = D_{\kappa-}^{-1}$ for $\kappa = \kappa_1$, κ_2 and κ_3 and wave–particle interactions are neglected then [see corollary 1–2 (I)] $h_{\kappa_j} = 0$ and $\langle\psi_3^*, \delta J_{\kappa_1,\kappa_2}[\psi_1, \psi_2]\rangle$ is independent of permutations of 1, 2, and 3. The Manley–Rowe relations are then satisfied, i.e.,

$$[\nabla_h \langle \psi_j, H_\kappa \psi_j \rangle]_{\kappa=\kappa_j} \cdot \nabla_h |A_j(x_h)|^2 \tag{2.16}$$

is independent of $j, j \in \{1,2,3\}$.

Proof: The Manley–Rowe relations follows from (2.14) and the two other equations obtained by permutations of 1, 2, and 3 if we make use of $\langle \psi_j, h_{\kappa_j} \psi_j \rangle = 0$ and the symmetry of the coupling coefficient.

Corollary 2: With the assumptions in the result we may write (2.14) in a Lorentz frame L (e_0, e_1, e_2, e_3) as

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_{g3} \cdot \frac{\partial}{\partial \mathbf{r}} + \gamma_3 \right) A_3^*(t,\mathbf{r}) = \frac{i\omega_3}{W_3} V(3,1,2) A_1(t,\mathbf{r}) A_2(t,\mathbf{r}). \tag{2.17}$$

We define $cte_0 + \mathbf{r} = x_h$, $(\omega/c)e_0 + \mathbf{k} = \kappa$ for $j = 1,2$, and 3

$$W_j = \epsilon_0 \omega_j \left(\frac{\partial}{\partial \omega} \langle \psi_j, H_\kappa \psi_j \rangle \right)_{\kappa=\kappa_j}, \tag{2.18}$$

$$v_{gj} = -\left(\frac{\partial}{\partial \mathbf{k}} \langle \psi_j, H_\kappa \psi_j \rangle \Big/ \frac{\partial}{\partial \omega} \langle \psi_j, H_\kappa \psi_j \rangle \right)_{\kappa=\kappa_j}, \tag{2.19}$$

$$\gamma_j = \epsilon_0 \omega_j \langle \psi_j, h_{\kappa_j} \psi_j \rangle W_j^{-1}, \tag{2.20}$$

$$V(3,1,2) = -2c^{-1} \langle \psi_3^*, \delta J_{\kappa_1,\kappa_2}[\psi_1, \psi_2] \rangle. \tag{2.21}$$

Proof: Corollary 2 follows in a straightfoward way from the result.

Remark 2: By assumption (i) in the result the first-order linear wave spectrum is determined by an essentially Hermitian operator. This is of course an important simplification due to the nice properties of such operators. The assumption is essential for the derivation of (2.14).

Remark 3: Assumption (ii) in the result is easily seen to be equivalent to (a) and (b) in the end of the Introduction. Without (a) a pure three-wave interaction process is no longer possible since also modes on the other branches would be excited. In the case of many such modes this may be a fast randomization mechanism.[2]

3. A COROLLARY OF RESULT 1 IN REFERENCE 5

In order to prove the gauge invariance of the coefficients in (2.14) we need (for $m = 1$ and 2) the following corollary of Result 1 in Ref. 4.

Corollary 3: (a) The quantity

$$\int_E \phi_0(x) \cdot \delta J^{(m)}[\phi_1, \ldots, \phi_m](x) dx, \tag{3.1}$$

appearing in Ref. 4 in (3.1) and (3.2), is gauge invariant in ϕ_j for $j = 0, 1, \ldots, m$.

(b) $\delta J^{(m)}[\phi_1, \phi_2, \ldots, \phi_m]$ is gauge invariant in ϕ_j for $j = 1, \ldots, m$.

(c) $\nabla_E \cdot \delta J^{(m)}[\phi_1, \ldots \phi_m] = 0$ (3.2)

Proof: In (a) we need to prove that if $\nabla_E \wedge \phi_j = \nabla_E \wedge \phi_j'$ for $j = 0, \ldots, m$ then a substitution of ϕ_j' instead of ϕ_j in (3.1) does not change the value of the quantity (3.1). This follows easily by the use of (4.22), (4.23), and (4.25) (in Ref. 5) in (3.1)–(3.3) (Ref. 5). We directly obtain (b) from (a) so now only (c) remains. From (a) we have

$$\nabla_E \wedge \phi_0 = 0 \Longrightarrow \int_E \phi_0(x) \cdot \delta J^{(m)}[\phi_1, \ldots, \phi_m](x) dx = 0. \tag{3.3}$$

Since $\nabla_E \wedge (\nabla_E B) = 0$ for any scalar function $B(x)$ we obtain from (3.3) that

$$\int_E \nabla_E B(x) \cdot \delta J^{(m)}[\phi_1, \ldots, \phi_m](x) dx = 0. \tag{3.4}$$

By partial integration of (3.4) we easily see that (3.2) follows.

4. DERIVATION OF THE RESULT

The derivation of the result is divided in the following steps:

(a) The gauge invariance of the coefficients in (2.14) is proved.

(b) We prove that the imaginary parts of the coefficients on the left hand side vanishes.

(c) Equation (2.14) is then derived. Due to (a) above it is now sufficient to consider the case when $\psi_1, \psi_2, \psi_3 \in P_L(V_i, V^+)$, where L is some Lorentz frame.

(d) We show that $[\nabla_h \langle \psi_j, H_\kappa \psi_j \rangle]_{\kappa=\kappa_j}$ is proportional to the group four velocity

(e) We demonstrate that the wave energy density integrated over V_i and with respect to a Lorentz frame $L = (e_0, e_1, e_2, e_3)$ is equal to $|A_j(x_h)|^2 W_j$ where W_j is given by (2.18). From this (2.15) follows.

Proof of (a): First we demonstrate the gauge invariance of $\langle \psi, H_\kappa \psi \rangle$, i.e., for $\psi, \psi' \in P(V_i, V^+)$ we want to show the implication

$$\nabla_\kappa \wedge \psi = \nabla_\kappa \wedge \psi' \Longrightarrow \langle \psi, H_\kappa \psi \rangle = \langle \psi', H_\kappa \psi' \rangle. \tag{4.1}$$

From remark 8(I) and (2.7) (which is valid for ψ_1, $\psi_2 \in P(V_i, V^+)$) we easily obtain

$$\langle \psi_1, H_\kappa \psi_2 \rangle = \langle H_\kappa \psi_1, \psi_2 \rangle. \tag{4.2}$$

It follows from Corollary 3(b) that $H_\kappa \psi$ is gauge invariant with respect to κ. If $\nabla_\kappa \wedge \psi = \nabla_\kappa \wedge \psi'$ we now obtain

$$\begin{aligned} \langle \psi, H_\kappa \psi \rangle &= \langle \psi, H_\kappa \psi' \rangle = \langle H_\kappa \psi, \psi' \rangle = \langle H_\kappa \psi', \psi' \rangle \\ &= \langle \psi', H_\kappa \psi' \rangle. \end{aligned} \tag{4.3}$$

We have now derived (4.1). The formulas (4.1)–(4.3) remains valid if we replace H with h and now the gauge invariance of the coefficients on the left-hand side in (2.14) follows. In the right-hand side we must prove that $\langle \psi_3^*, \delta J_{\kappa_1,\kappa_2}[\psi_1, \psi_2] \rangle$ is a gauge invariant function of ψ_1, ψ_2, and ψ_3 with respect to κ_1, κ_2, and κ_3, respectively. The statement conserning ψ_1 and ψ_2 follows easily from Corollary 3 (b). In order to prove it for ψ_3 it is sufficient to demonstrate that the implication ($\psi \in P(V_i, V^+)$)

$$\nabla_{\kappa_3} \wedge \psi = 0 \Longrightarrow \langle \psi^*, \delta J_{\kappa_1,\kappa_2}[\psi_1, \psi_2] \rangle = 0. \tag{4.4}$$

However, $\nabla_{\kappa_3} \wedge \psi = 0$ implies the existence of a unique complexvalued function $A(x_i)$ such that

$$\nabla_{\kappa_3} A(x_i) = \psi(x_i). \tag{4.5}$$

(*Proof*: Write $\psi = \psi_h + \psi_i$ in accordance with definition 2(a) (I). Observe that $\nabla_{\kappa_3} \wedge \psi = 0$ is equivalent the three relations (1): $i\kappa_3 \wedge \psi_h = 0$, (2): $\nabla_i \wedge \psi_i = 0$ and (3): $i\kappa_3 \wedge \psi_i + \nabla_i \wedge \psi_h = 0$. By means of (1) we have κ_3 parallel to ψ_h and we define a complex valued function $A(x_i)$ by $\psi_h = iA\kappa_3$. By inserting this expression in (3) we obtain $i\kappa_3 \wedge (\psi_i - \nabla_i A) = 0$ which implies $\psi_i = \nabla_i A$. Then also (2) is satisfied and the existence and uniqueness of A have been proved.) By substitution of (4.5) in (4.4), by partial integration and the use of $\kappa_3 = -(\kappa_1 + \kappa_2)$ we obtain

$$\langle (\nabla_{\kappa_3} A)^*, \delta J_{\kappa_1,\kappa_2}[\psi_1,\psi_2]\rangle = -\int_{V_i} A \nabla_{\kappa_1+\kappa_2}\cdot \delta J_{\kappa_1,\kappa_2}[\psi_1,\psi_2]dx_i. \tag{4.6}$$

From corollary 3(c) we easily obtain

$$\nabla_{\kappa_1+\kappa_2}\cdot \delta J_{\kappa_1+\kappa_2}[\psi_1,\psi_2] = 0. \tag{4.7}$$

Now (4.4) follows from (4.6) and (4.7).

Proof of (b): Take $\kappa \in V_h$ and $\psi \in P(V_i, V^+)$. According to Lemma 2 there exist a $\chi \in P_L(V_i, V^+)$ such that $\nabla_\kappa \wedge \chi = \nabla_\kappa \wedge \psi$. Due to (a) above we have

$$\langle \psi, H_\kappa \psi\rangle = \langle \chi, H_\kappa \chi\rangle = \langle \chi, \Pi_L \circ H_\kappa \chi\rangle. \tag{4.8}$$

The last term in (4.8) is real since $\Pi_L \circ H_\kappa$ is Hermitian. The reality of $\langle \psi, h_\kappa \psi\rangle$ is proved in the same way and (b) follows easily.

Proof of (c): Take ψ_1, ψ_2, and ψ_3 in $P_L(V_i, V^+)$. We use the symbol $O(\epsilon^n)$ to denote any quantity of order ϵ^n or smaller, where ϵ is a small parameter. We assume that

$$A_j = O(\epsilon), \quad \psi_j = O(1), \quad \Delta\kappa = O(\epsilon), \quad H_{\kappa_j} = O(1),$$
$$h_{\kappa_j} = O(\epsilon), (j = 1,2,3), \tag{4.9}$$

where $\Delta\kappa$ is the width in κ space of the wave packets.

Fourier transformation of (2.13) with respect to V_h yields

$$\phi_\kappa = \sum_{j=1}^{3} (A_{j\kappa-\kappa_j}\psi_j + (A_j^*)_{\kappa+\kappa_j}\psi_j^*) + O(\epsilon^2), \tag{4.10}$$

where $A_{j\kappa-\kappa_j} \neq 0$ only for $\kappa - \kappa_j = O(\epsilon)$. For $\kappa - \kappa_3 = O(\epsilon)$ we have

$$\langle \psi_3, (H_\kappa + ih_\kappa)\phi_\kappa\rangle = \langle \psi_3, [H_{\kappa_3} + (\kappa - \kappa_3)\cdot(\nabla_h H_\kappa)_{\kappa=\kappa_3} + ih_{\kappa_3} + O(\epsilon^2)]\phi_\kappa\rangle. \tag{4.11}$$

Since $\psi_3 \in P_L(V_i, V^+)$ and $\Pi_L \circ H_{\kappa_3}$ is Hermitian and (2.11) we obtain $\langle \psi_3, H_{\kappa_3}(\ldots)\rangle = 0$. Thus

$$\langle \psi_3, (H_\kappa + ih_\kappa)\phi_\kappa\rangle = \langle \psi_3, [(\kappa - \kappa_3)\cdot(\nabla_h H_\kappa)_{\kappa=\kappa_3} + ih_{\kappa_3} + O(\epsilon^2)]\phi_\kappa\rangle. \tag{4.12}$$

We now substitute (4.10) in (4.12)

$$\langle \psi_3, (H_\kappa + ih_\kappa)\phi_\kappa\rangle = [(\kappa - \kappa_3)\cdot(\nabla_h \langle \psi_3, H_\kappa \psi_3\rangle)_{\kappa=\kappa_3} + i\langle \psi_3, h_{\kappa_3}\psi_3\rangle]A_{3\kappa-\kappa_3} + O(\epsilon^3) \tag{4.13}$$

From (4.13) we then obtain by the inverse Fourier transform [see (3.22) (I)]

$$\begin{aligned}(2\pi)^{-m}\int_{\kappa=\kappa_3+O(\epsilon)} &\langle \psi_3, (H_\kappa + ih_\kappa)\phi_\kappa\rangle \exp i(\kappa - \kappa_3)\cdot x d\kappa \\ &= -i(\nabla_h \langle \psi_3, H_\kappa \psi_3\rangle)_{\kappa=\kappa_3}\cdot \nabla_h A_3(x_h) \\ &\quad + i\langle \psi_3, h_{\kappa_3}\psi_3\rangle A_3(x_h) + O(\epsilon^3).\end{aligned} \tag{4.14}$$

Now substitute in (4.14) the right-hand side of (2.2) in place of $(H_\kappa + ih_\kappa)\phi_\kappa$. We get

$$\begin{aligned}(2\pi)^{-m}\int_{\kappa=\kappa_3+O(\epsilon)} &\langle \psi_3, -(\mu_0/\epsilon_0)^{1/2}(2\pi)^{-m} \\ &\times \int_{V_h} \delta J_{\kappa',\kappa-\kappa'}[\phi_{\kappa'}, \phi_{\kappa-\kappa'}]d\kappa'\rangle \\ &\times \exp i(\kappa - \kappa_3)\cdot x d\kappa = -2(\mu_0/\epsilon_0)^{1/2} \\ &\times \langle \psi_3, \delta J_{-\kappa_1,-\kappa_2}[\psi_1^*, \psi_2^*]\rangle A_1^* A_2^*.\end{aligned} \tag{4.15}$$

Now (2.14) follows from (2.2), (2.3), (4.14), and (4.15).

Proof of (d): Let us consider wave 3. We deal below with κ satisfying $\kappa = \kappa_3 + O(\epsilon)$. According to (a) above we may assume that $\psi_3 \in P_L(V_i, V^+)$. Let $\lambda_L(\kappa)$ be continuous in κ and an eigenvalue of the Hermitian operator $\Pi_L \circ H_\kappa$ and satisfying $\lambda_L(\kappa_3) = 0$. We denote the corresponding eigenvectors $\chi(\kappa)$. Then according to assumption (ii) in the result the plasma have only one branch of normal modes for κ close to κ_3 (see Remark 3). Thus $\lambda_L(\kappa)$ is uniquely determined by the requirements above. The eigenvector $\chi(\kappa_3)$ is proportional to ψ_3 and we normalize so that $\chi(\kappa_3) = \psi_3$. The dispersion relation of mode 3 is $\lambda_L(\kappa) = 0$. The ordinary group velocity in frame L is $-[[\partial\lambda_L(\kappa)/\partial\mathbf{k}]/[\partial\lambda_L(\kappa)/\partial\omega]]_{\kappa=\kappa_3}$, where $\kappa = (\omega/c)e_0 + \mathbf{k}$. Thus the group 4-velocity $u_{g_3} \in S \cap V_h$ is proportional to $(\nabla_h \lambda_L(\kappa))_{\kappa=\kappa_3}$. Now, if

$$\langle \psi_3, H_\kappa \psi_3\rangle = \lambda_L(\kappa)\langle \psi_3, \psi_3\rangle + O(\epsilon^2) \tag{4.16}$$

then it follows that u_{g_3} is proportional to $(\nabla_h \langle \psi_3, H_\kappa \psi_3\rangle)_{\kappa=\kappa_3}$ as we want to show. In order to prove (4.16) observe that

$$\langle \psi_3, H_\kappa O(1)\rangle = O(\epsilon), \tag{4.17}$$

$$\langle O(1), \Pi_L \circ H_\kappa(\psi_3)\rangle = O(\epsilon). \tag{4.18}$$

The relations (4.17) and (4.18) follows from $\psi_3 \in P_L(V_i, V^+)$, the Hermitian property of $\Pi_L \circ H_\kappa$ and $H_{\kappa_3}\psi_3 = 0$. From (4.17), (4.18), and $\chi(\kappa) = \psi_3 + O(\epsilon)$ we obtain

$$\langle \chi(\kappa), H_\kappa \chi(\kappa)\rangle = \langle \psi_3, H_\kappa \psi_3\rangle + O(\epsilon^2). \tag{4.19}$$

We also have

$$\begin{aligned}\langle \chi(\kappa), H_\kappa \chi(\kappa)\rangle &= \lambda_L(\kappa)\langle \chi(\kappa), \chi(\kappa)\rangle \\ &= \lambda_L(\kappa)\langle \psi_3, \psi_3\rangle + O(\epsilon^2),\end{aligned} \tag{4.20}$$

where $\lambda_L(\kappa) = O(\epsilon)$ is used. The relations (4.19) and (4.20) gives (4.16).

Proof of (e): The wave energy in a lossless medium is equal to the work required from external sources to generate the wave. In an almost lossless plasma, like the one we consider [assumption (i) in the result], it is reasonable to define the wave energy as the work required to generate the wave when dissipation is neglected. We want to calculate the over V_i integrated wave energy density W_3 of the normal mode

$$\phi_3(x) = \psi_3 \exp(i\kappa_3\cdot x) + \text{c.c.} \tag{4.21}$$

The wave energy is defined with respect to some Lorentz frame L. In accordance with the discussion above we calculate W_3 as follows. Take $\phi(x)$ so that $\phi(x)$ approaches $\phi_3(x)$ towards the future and vanishes towards the past. Let $J_{\text{ext}}(x)$ be the 4-current that linearily produces $\phi(x)$ if dissipation is neglected, i.e., [see (2.1) and Remark 8 (I)]

$$\begin{aligned}\nabla_E\cdot(\nabla_E \wedge \phi) + (\mu_0/\epsilon_0)^{1/2}2^{-1}(\delta J^{(1)}[\phi] + \delta J^{(1-)}[\phi]) \\ = -(\mu_0/\epsilon_0)^{1/2}J_{\text{ext}}.\end{aligned} \tag{4.22}$$

Let E be the electric field due to ϕ and $\mathbf{J}_{\text{ext}}$ the external current in the frame $L = (e_0, e_1, e_2, e_3)$. Take W as the energy density which is related to ϕ as W_3 is to ϕ_3. Then we have

$$W'(t) = -\overline{\int_{V_i} \mathbf{E}(x)\cdot \mathbf{J}_{\text{ext}}(x)dx_i}, \tag{4.23}$$

where $t = -c^{-1}e_0\cdot x$ and the line above the integral denotes that variations in V_h on the scale $\exp i\kappa_3\cdot x$ have been averaged away. We obtain

$$W_3 = W(\infty) = -\int_{-\infty}^{\infty} dt \overline{\int_{V_i} \mathbf{E}(x)\cdot\mathbf{J}_{\mathrm{ext}}(x) dx_i}. \quad (4.24)$$

Let $a(t)$ be a slowly growing function with $a(-\infty) = 0$ and $a(\infty) = 1$. We then take

$$\phi(x) = a(t)\phi_3(x). \quad (4.25)$$

The 4-current J_{ext} is determined by (4.22). Let $a(\omega)$ denote the Fourier transform of $a(t)$. We assume that $a(\omega)\neq 0$ only for a sufficiently small ω. We have

$$a(t)\phi(x) = (2\pi)^{-1}\int [a(\omega-\omega_3)\psi(x_i)\exp(i\mathbf{k}_3\cdot\mathbf{r}) + a(\omega+\omega_3)\psi^*(x_i)\exp-(i\mathbf{k}_3\cdot\mathbf{r})]e^{-i\omega t}d\omega. \quad (4.26)$$

From (4.22), (4.26), (2.4), and Remark 8 (I) we obtain (for typographical reasons we write $H_\kappa[\psi] = H[\kappa,\psi]$)

$$(2\pi)^{-1}\int \{a(\omega-\omega_3)H[(\omega/c)e_0+\mathbf{k}_3,\psi_3](x_i)\exp i\mathbf{k}_3\cdot x + a(\omega+\omega_3)H[(\omega/c)e_0-\mathbf{k}_3,\psi_3^*](x_i)\exp i(-\mathbf{k}_3\cdot x)\}e^{i\omega t}d\omega$$
$$= -(\mu_0/\epsilon_0)^{1/2}J_{\mathrm{ext}}. \quad (4.27)$$

We may choose $\psi_3\in P_L(V_i, V^+)$ and then by Lemma 2

$$\mathbf{E}(x) = a(t)i(\omega_3/c)(\psi_3(x_i)\exp(i\kappa_3\cdot x) - \mathrm{c.c.}). \quad (4.28)$$

We now substitute (4.27) and (4.28) in (4.24)

$$W_3 = i(\omega_3/c)(\epsilon_0/\mu_0)^{1/2}(2\pi)^{-1}\int_{-\infty}^{\infty} dt\int d\omega\, a(t)\{-\langle\psi_3^*,H[(\omega/c)e_0-\mathbf{k}_3,\psi_3^*]\rangle$$
$$\times a(\omega+\omega_3)\exp i(-\omega_3-\omega)t + \langle\psi_3, H[(\omega/c)e_0+\mathbf{k}_3,\psi_3]\rangle a(\omega-\omega_3)\exp i(\omega_3-\omega)t\}. \quad (4.29)$$

In the first term of the integrand in (4.29) only ω close to ω_3 will contribute (for only then $a(\omega+\omega_3)\neq 0$) to the integral and in the second term only ω close to ω_3 will contribute.

We accordingly may approximate the integrand by linear terms in $(\omega+\omega_3)$ and $(\omega-\omega_3)$. We also use

$$\langle\psi_3^*, H[(\omega/c)e_0-\mathbf{k}_3,\psi_3^*]\rangle = \langle\psi_3, H[-(\omega/c)e_0+\mathbf{k}_3,\psi_3]\rangle \quad (4.30)$$

and then we obtain from (4.29)

$$W_3 = \epsilon_0\omega_3\left[\frac{\partial}{\partial\omega}\langle\psi_3,H[(\omega/c)e_0+\mathbf{k}_3,\psi_3]\rangle\right]_{\omega=\omega_3}\int_{-\infty}^{\infty} dt\, a(t)(2\pi)^{-1}\int d\omega$$
$$\cdot[-i(\omega+\omega_3)a(\omega+\omega_3)\exp(-i(\omega+\omega_3)t) - i(\omega-\omega_3)a(\omega-\omega_3)\exp(-i(\omega-\omega_3)t)]. \quad (4.31)$$

The ω integral in (4.31) is equal to $2a'(t)$ and for the t-integral we then obtain

$$\int_{-\infty}^{\infty} dt\, a(t)2a'(t)dt = a(\infty)^2 - a(-\infty)^2 = 1. \quad (4.32)$$

5. SUMMARY

The main result of this paper are the coupled mode equations in the form (2.14) or in "standards" form (2.17). Explicit expressions for the coefficients may be found in part I of this paper. In corollary 1 we see how the symmetries considered in part I leads to the Manley–Rowe relations. As have been stressed in the Introduction and in the Result there are, for an inhomogeneous plasma, additional conditions for these coupled mode equations to be valid (see also remark 2). There are, however no specific restrictions on the geometry of the plasma and the formulas may still be used for a great variety of situations. We finally remark that this part II has been written in a model independent way only requiring knowledge of the response operators and the results may thus be useful also for fluid and/or collisional plasmas.

Note added in proof: The Lorentz frames $L = (e_0,e_1,e_2,e_3)$ of interest in this paper have the property that $e_0,\dots,e_{m-1}\in V_h$ while $e_m,\dots,e_3\in V_i$, where m is the dimension of V_H. This property of L has implicitly been assumed in some derivations above.

[1] J. Larsson, J. Math. Phys. 22, 176 (1982), companion paper.
[2] K. Rypdal, "On the coupled-mode theory for ducted waves," Physica Scripta 23, 277 (1981).
[3] Nonlinear interaction between waves in plasma with sharp boundary have recently received increasing interest. See, for example, J.-M. Larsen and F. W. Crawford, Int. J. Electron. 46, 577 (1979); 47, 317 (1979); A. G. Sitenko and V. N. Pavlenko, Sov. Phys. JETP 47, 65 (1979).
[4] E. S. Weibel, Plasma Phys. 16, 921 (1974).
[5] J. Larsson, J. Math. Phys. 20, 1321 (1979).