A new method of matrix transformation. I. Matrix diagonalizations via involutional transformations

Shoon K. Kim
Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122

(Received 15 February 1979; accepted for publication 27 April 1979)

It is shown that two matrices A and B of order $n\times n$ which satisfy a monic quadratic equation with two roots λ_1 and λ_2 are connected by $AT_{AB} = T_{AB}B$ where $T_{AB} = A + B - (\lambda_1 + \lambda_2)I$ with I being the $n\times n$ unit matrix (Theorem 1). The condition for T_{AB} to be involutional is that the anticommutator of $\bar{A} = A - (1/2)(\lambda_1 + \lambda_2)I$ and $\bar{B} = B - (1/2)(\lambda_1 + \lambda_2)$ is a c number (Theorem 2). A $2m\times 2m$ matrix $Q^{(2m)}$ is introduced as a typical form of a matrix which can be diagonalized by an involutional transformation. These theorems are further extended through the matrix representation of the group of the general homogeneous linear transformations, GL(n). IUH (involutional, unitary, and Hermitian) matrices are introduced and discussed. The involutional transformations are shown to play a fundamental role in the transformations of Dirac's Hamiltonian and of the field Hamiltonians which are quadratic in particle creation and annihilation operators in solid state physics.

1. INTRODUCTION

Almost ten years ago the author developed a theory of involutional matrices [1] (this work will be referred to as I), which are defined as the solutions of the simple quadratic equation,

$$x^2 = \text{const}\times I, \tag{1.1}$$

where I is the unit matrix. The two-dimensional solution of this equation is simply given by a traceless matrix; the well known examples are the Pauli spin matrices. The solutions of (1.1) in arbitrary dimensions have been obtained by the matrix representation for the group of the general homogeneous linear transformations in two dimensions, GL(2). The symmetry properties, and the eigenwert problems for the solutions have also been discussed. It has been recognized that involutional matrices have deep roots in various problems of mathematical physics. [1,2] The solutions of a more general equation $x^m = \text{const}\times I$ have also been studied by Santhanam *et al.* [3] and are given explicitly for $m = 3$. We shall, however, limit $m = 2$ in the present work.

The purpose of the present work is to show the effectiveness of involutional transformations in connecting matrices of certain types frequently encountered in physics. The convenience of an involutional transformation is obvious from the fact that the inverse transformation is the same as the original transformation. Any unitary transformation becomes an involutional transformation when the unitary matrix of the transformation is Hermitian. Such a matrix may be called an IUH (involutional, unitary, and Hermitian) matrix, where any two of the properties guarantee the third. The well known examples are Puali's spin matrices, Dirac's γ matrices, and Dirac's Hamiltonian for a free electron in the momentum representation. [4] The effectiveness of involutional transformations has also been recognized by Wigner [5] in connection with the representation theory of the Poincaré group.

Let $M(n\times n, p^{(r)}(x))$ be a set of matrices of order $n\times n$, every member of which satisfies a given polynomial equation $p^{(r)}(x) = 0$ of degree r. We can then prove a general theorem [6] that if two matrices A and B belong to the set

$$A, B\in M(n\times n, p^{(r)}(x)), \tag{1.2}$$

we can construct a polynomial T_{AB} of A and B which transforms A into B if T_{AB} is nonsingular. In the discussion of involutional transformation, however, we need to consider only the simplest special case where $r = 2$. Let $p^{(2)}(x) = x^2 - (\lambda_1 + \lambda_2)x + \lambda_1\lambda_2$. Then the transformation matrix T_{AB} which connects A and $B\in M(n\times n, p^{(2)}(x))$ is simply given by $T_{AB} = A + B - (\lambda_1 + \lambda_2)I$ (Theorem 1). Based on this theorem we shall establish the condition for T_{AB} to be involutional (Theorem 2).

In special cases where $A^2 = B^2 = I$, these theorems are closely related to the fundamental theorem of Pauli [6,7] which connects two sets of anticommuting Dirac γ matrices. Based on these theorems we can construct the spinor representation for a group of orthogonal transformations in arbitrary dimensions [6,8], of which a special case is the Lorentz group. The more direct applications of the theorems are in solving the energy eigenvalue problems in the Dirac relativistic theory of an electron; the powerful transformation due to Foldy and Wouthuysen [9,10] for a free electron and Biedenharn's transformation [11] in the Dirac–Coulomb problem are immediately obtained from Theorem 2 (Sec. 2).

Using the theorems developed in Sec. 2, we shall construct a special type of matrix $Q^{(2m)}$ of order $2m\times 2m$ which can be diagonalized by an involutional transformation. The $Q^{(2m)}$ matrix is by no means the most general form of a matrix which can be diagonalized by an involutional transformation. Nevertheless, all 2×2 matrices and the Dirac matrices are $Q^{(2m)}$ types or direct sums of $Q^{(2m)}$ type. It will be shown that when $Q^{(2m)}$ is normal, the transformation matrix which diagonalizes $Q^{(2m)}$ becomes an IUH matrix (Sec. 4, we shall extend the theorems developed in Secs. 2 and 3 to an invariant matrix [12] $S(A)$ of a matrix $A \in M(n\times m, p^2(x))$ by the matrix representation theory developed in I. This is possible since in this representation $S(A)$ is triangular when A is triangular. Moreover, if we limit A to 2×2 matrices, we can show that many symmetry properties of A are directly reflected to those of $S(A)$. As a result, if A is diagonalized by an IUH matrix, then so is $S(A)$.

In Sec. 5, we shall given some illustrative examples which may exhibit further significance to the involutional transformations. The examples considered are the Bogoliubov Hamiltonian [13,14] for phonons in a superfluid and the Hopfield Hamiltonian [15] for a polariton-photon field. These are typical field Hamiltonians in solid state physics. It is well known that diagonalizations (the reductions into the canonical forms) of these Hamiltonians require nonunitary transformations, owing to the fact that the particle creation and annihilation operators are not Hermitian. We shall show that these nonunitary transformations can be characterized by involutional transformations.

2. BASIC THEOREMS

Theorem 1: Let A and B be matrices of order $n\times n$ which satisfy a monic quadratic equation $p^{(2)}(x) = x^2 - (\lambda_1 + \lambda_2)x + \lambda_1\lambda_2 = 0$, with two roots λ_1 and λ_2. Then, A and B are connected by

$$A\,T_{AB} = T_{AB}\,B, \tag{2.1}$$

where

$$T_{AB} = A + B - (\lambda_1 + \lambda_2)I, \tag{2.1'}$$

I being the $n\times n$ unit matrix. If T_{AB} is nonsingular, then A and B are equivalent.

The proof is self evident. It is obvious that if A and B are not equivalent, then T_{AB} is singular, while the equivalence of A and B does not guarantee that T_{AB} is nonsingular. For convenience, we may call T_{AB} the characteristics transformation matrix of A and B, which may or may not be singular. Since $T_{AB} = T_{BA}$, we have

$$T_{AB}A\,T_{AB} = T^2_{AB}B = B\,T^2_{AB}, \tag{2.2}$$

that is, T^2_{AB} commutes with A and B. This property will play an important role when B is a diagonal matrix equivalent to A. The more general transformation matrix V_{AB} which connects A and B through

$$A\,V_{AB} = V_{AB}\,B \tag{2.3}$$

is given by

$$V_{AB} = F_A\,T_{AB} = T_{AB}\,F_B, \tag{2.3'}$$

where F_x is a function of x. In practical application, it is often useful to introduce a special case where $F_A = A$,- $F_B = B$. Then

$$V_{AB} = A\,T_{AB} = T_{AB}\,B = AB - \lambda_1\lambda_2 I. \tag{2.4}$$

Note that $V_{AB} \neq V_{BA}$, unless A and B commute. The transformation matrices T_{AB} and V_{AB} constitute the basis of the present work.

It is obvious that Theorem 1 is most useful when T_{AB} is nonsingular. We can easily establish such a condition for an important special case where T_{AB} is involutional. Since involutionality requires that the anticommutator of $\bar{A} = A - \frac{1}{2}(\lambda_1 + \lambda_2)I$ and $\bar{B} = B - \frac{1}{2}(\lambda_1 + \lambda_2)I$ is a c number,

$$[\bar{A}, \bar{B}]_+ = \bar{A}\bar{B} + \bar{B}\bar{A} = 2c_{AB}I, \tag{2.5}$$

we have the following theorem:

Theorem 2: Let $A, B \in M(n\times n, p^{(2)}(x))$. If the anticommuator of $\bar{A} = A - \frac{1}{2}(\lambda_1 + \lambda_2)I$ and $\bar{B} = B - \frac{1}{2}(\lambda_1 + \lambda_2)I$ is a c number ($= 2c_{AB}$), there exists an involutional transformation which brings A into B via

$$Y_{AB}AY_{AB} = B;\; Y^2_{AB} = I, \tag{2.6}$$

where

$$Y_{AB} = N_{AB}[A + B - (\lambda_1 + \lambda_2)I],$$
$$N^{-2}_{AB} = \tfrac{1}{2}(\lambda_1 - \lambda_2)^2 + 2c_{AB}, \tag{2.7}$$

provided $N^{-2}_{AB} \neq 0$. Moreover, if A and B are Hermitian, Y_{AB} is an IUH (involutional, unitary, and Hermitian) matrix.

The involutional matrix Y_{AB} acts as an exchange operator of A and B on any function $f(A,B)$ of A and B which can be expressed in a power series of A and B,

$$Y_{AB}f(A,B)\,Y_{AB} = f(B,A). \tag{2.8}$$

Thus, Y_{AB} commutes with any symmetric function of A and B, and anticommutes with any antisymmetric function of A and B.

An alternate application of Theorem 2 is to transform A into the complementary matrix B' of B defined by $B' = (\lambda_1 + \lambda_2)I - B$, since B' satisfies the same quadratic equation, $P^{(2)}(x) = 0$. The matrix of transformation $Y_{AB'}$ which brings A to B' via $Y_{AB'}AY_{AB'} = B'$ is given by

$$Y_{AB'} = N_{AB'}(A - B);\; N^{-2}_{AB'} = \tfrac{1}{2}(\lambda_1 - \lambda_2)^2 - 2c_{AB} \neq 0. \tag{2.9}$$

Note that, if $\lambda_1 \neq \lambda_2$, both of T_{AB} and $T_{AB'}$ can not be nilpotent at the same time. The transformation of B into its compliment B' gives an interesting problem since $T_{BB'} = 0$ by definition. We shall come back to this problem at the end of Sec. 3.

In a simplest special case where A and B are 2×2 equivalent matrices, the matrices $\bar{A}$, $\bar{B}$ and T_{AB} of Theorem 2 are all involutional matrices and the anticommutator of $\bar{A}$ and $\bar{B}$ is always a c number. Thus, we have the following corrollary to Theorem 2:

Corollary: The characteristic transformation matrix T_{AB} of two equivalent 2×2 matrices A and B is always involutional.

When A and B are involutional matrices satisfying $A^2 = B^2 = I$, Theorems 1 and 2 take simpler forms since $\lambda_1 + \lambda_2 = 0$, so that $T_{AB} = A + B$. In this special case Theorem 2 has very elegant applications for the Dirac theory of an electron. To see this, let $\gamma_1, \gamma_2, \cdots, \gamma_d$ be a set of d-anticommuting involutional matrices of an appropriate order,

$$\gamma_\nu\gamma_\mu + \gamma_\mu\gamma_\nu = 2\delta_{\nu\mu},\quad \nu,\mu = 1,2,\cdots,d, \tag{2.10}$$

and let u be a unit vector with constant components $(u_1, u_2, \cdots, u_d)$ in a d-dimensional space. Then a linear combination with the vector u,

$$\gamma_u = \sum_{\nu=1}^{d} u_\nu\gamma_\nu;\quad (u\cdot u) = \sum_\nu u^2_\nu = 1 \tag{2.11}$$

is also involutional. Since the anticommutator of two such linear combinations γ_u and γ_v is a c number, γ_u can be transformed into γ_v via an involutional transformation

$$Y_{uv}\,\gamma_u\,Y_{uv} = \gamma_v, \qquad (2.12)$$

where

$$Y_{uv} = (\gamma_u + \gamma_v)/(2 + 2(u\cdot v))^{1/2}; \quad (u\cdot v) = \sum u_\nu v_\nu \neq -1. \qquad (2.12')$$

This result, (2.12), may be used to construct a general theory of the spinor representation for a group of orthogonal transformations in d dimensions.[6,7,8] Alternatively, the result may be used to solve the energy eigenvalue problem for the Dirac particle. Let us assume a representation which diagonalizes γ_d. Then the involutional matrix Y_{ud} given by

$$Y_{ud} = (\gamma_u + \gamma_d)/(2 + 2u_d)^{1/2}, \quad u_d \neq -1 \qquad (2.13)$$

brings γ_u to the diagonal matrix γ_d. Since the Dirac Hamiltonian for a free electron in the momentum representation is a special case of γ_u with $d = 4$ except for a scalar factor of the energy E, the Dirac plane wave solutions[9] are completely described by the involutional matrix Y_{ud}. It can be shown that a powerful transformation due to Foldy and Wouthuysen[10] S_{FW} for a free electron and Biedenharn's[11] transformation S_B in the Dirac–Coulomb problem are simply given by $\gamma_d Y_{ud}$, which corresponds to V_{AB} of (2.4). It will be shown[6] that Y_{ud} is more effective than $\gamma_d Y_{ud}$ in these transformations on account of the fact that Y_{ud} is involutional.

3. DIAGONALIZATION OF $Q^{(2m)}$ MATRICES

We shall specialize the results obtained in Sec. 2 to the case where B equals a diagonal matrix Λ equivalent to $A \in M(n\times n, P^{(2)}(x))$ with $P^{(2)}(x)$ having two distinct roots. It can be shown that there exists at least one proper Λ which makes the characteristic transformation matrix $T = T_{A\Lambda}$ be nonsingular. We shall, however, postpone the proof of this statement to a forthcoming paper[6] where we shall develop a general theory of matrix diagonalization for a matrix $A \in M(n\times n, P^{(r)}(x))$ based on the reduced characteristic equation $P^{(r)}(x) = 0$ which has no multiple roots. In this section we simply investigate the structure and the properties of the matrix $A \in M(n\times n, P^{(2)}(x))$ which can be diagonalized by an involutional transformation.

The basic equations (2.1) and (2.1′) of Theorem 1 take the following forms, when B equals a diagonal matrix Λ equivalent to A: Let $T = T_{A\Lambda}$, then

$$AT = T\Lambda,\ TA = \Lambda T, \qquad (3.1)$$

where

$$T = A + \Lambda - (\lambda_1 + \lambda_2) I. \qquad (3.1')$$

Thus, a column (row) vector of T is an eigenvector of A provided that it is not a null vector. From (2.2) or directly from (3.1) and (3.1′), we have

$$TAT = T^2\Lambda = \Lambda T^2. \qquad (3.2)$$

Accordingly, if we assume a standard form for Λ defined by

$$\Lambda = \begin{bmatrix} \lambda_1 I_1 & 0 \\ 0 & \lambda_2 I_2 \end{bmatrix}, \qquad (3.3)$$

where I_ν is the unit $n_\nu \times n_\nu$ matrix with n_ν being the degeneracy of λ_ν, then T^2 becomes a direct sum of two submatrices of order $n_1 \times n_1$ and $n_2 \times n_2$.

In fact, if we substitute (3.1′) into one of T in T^2, we obtain

$$T^2 = \chi_1 T^{(1)} \oplus \chi_2 T^{(2)}; \quad \chi_\nu = 2\lambda_\nu - \lambda_1 - \lambda_2 \qquad (3.4)$$

where $\oplus$ denotes the direct sum and $T^{(\nu)}$ is given by a $n_\nu \times n_\nu$ submatrix of T,

$$[T^{(\nu)}]_{ij} = T_{ij}, \qquad (3.4')$$

with i,j belonging to the subspace corresponding to $\lambda_\nu I_\nu$ of Λ. This is a rather remarkable property of T, based on which one can construct a more convenient transformation matrix which becomes unitary when A is Hermitian.[6]

From (3.4), it is obvious that the condition for T to be involutional is given by

$$X_\nu T^{(\nu)} = N^{-2} I_\nu; \quad \nu = 1,2 \qquad (3.5)$$

where N is a constant independent of ν, which is equal to N_{AB} of (2.7) when $B = \Lambda$. The conditions (3.5) and (3.4) severely restrict the form of the matrix A: If we exclude the trivial case of where $n_1 \neq n_2$, we have $n_1 = n_2 = m$ and A becomes a $Q^{(2m)}$ type matrix of order $2m \times 2m$ defined by

$$Q^{(2m)} = \begin{bmatrix} a\mathbf{1} & \mathbf{b} \\ \mathbf{c} & d\mathbf{1} \end{bmatrix}; \quad \mathbf{bc} = e\mathbf{1} \qquad (3.6)$$

where a, d, e are constants, $\mathbf{1}$ is the $m \times m$ unit matrix, $\mathbf{b}$ and $\mathbf{c}$ are $m \times m$ matrices which commute with each other. The two characteristic roots are determined from $\lambda_1 + \lambda_2 = a + d$ and $\lambda_1 \lambda_2 = ad - e$, so that the condition $\lambda_1 \neq \lambda_2$ gives $(a-d)^2 + 4e \neq 0$. The trace and the determinant of $Q^{(2m)}$ are

$$trQ^{(2m)} = m(a+d), \quad \det Q^{(2m)} = (ad-e)^m. \qquad (3.6')$$

The form of $Q^{(2m)}$ matrix is rather limited, but all the 2×2 matrices are $Q^{(2m)}$ type matrices and all the Dirac matrices are $Q^{(4)}$ type or direct sums of $Q^{(2m)}$ type matrices. For convenience, we may state the above findings in a form of a theorem with the explicit form of the involutional matrix which diagonalizes a $Q^{(2m)}$ matrix.

Theorem 3: A $Q^{(2m)}$ type matrix of order $2m \times 2m$ defined by (3.6) is diagonalized by an involutional transformation through

$$Y_Q\, Q^{(2m)} Y_Q = \begin{bmatrix} \lambda_1 \mathbf{1} & 0 \\ 0 & \lambda_2 \mathbf{1} \end{bmatrix}, \quad Y_Q^2 = I, \qquad (3.7)$$

where $\lambda_1 \neq \lambda_2$ and

$$Y_Q = N_Q \begin{bmatrix} a - \lambda_2 \mathbf{1} & \mathbf{b} \\ \mathbf{c} & (d - \lambda_1)\mathbf{1} \end{bmatrix}, \qquad (3.8)$$

$$N_Q^{-2} = (\lambda_1 - \lambda_2)(a - \lambda_2) = (\lambda_2 - \lambda_1)(d - \lambda_1) \neq 0.$$

It should be noted that $N_Q^{-2} \neq 0$ can be always arranged if $\lambda_1 \neq \lambda_2$ and that Y_Q is a special case of Y_{AB} of (2.6). From (3.7), we have

$$\mathrm{tr}\, Y_Q = 0,\ \det Y_Q = (-1)^m \qquad (3.8')$$

Thus for m = odd, Y_Q describes an improper rotation (see I). It should also be noted that (2.13) is a special case of (3.8) when $Q^{(2m)}$ is involutional.

Now, according to Schur, the condition for a square matrix to be diagonalized by a unitary transformation is that the matrix be normal. A similar theorem holds for $Q^{(2m)}$.

Theorem 4: If the matrix $Q^{(2m)}$ is normal, the transformation matrix Y_Q is an IUH (involutional unitary and Hermitian) matrix.

The proof is simple. When $Q^{(2m)}$ is normal, the submatrices of $Q^{(2m)}$ satisfy the following conditions.

$$(a^* - d^*)\mathbf{b} = (a - d)\mathbf{c}^+$$
$$\mathbf{b}\mathbf{b}^+ = \mathbf{c}\mathbf{c}^+ = |e| \times \mathbf{1} \qquad (3.9)$$

where a^* and d^* are the complex conjugate of a and d respectively, and $\mathbf{b}^+$ and $\mathbf{c}^+$ are Hermitian conjugate of the submatrices **b** and **c**. From these conditions, it is a simple matter to show that Y_Q is Hermitian so that it is unitary as well.

Q.E.D.

For a 2×2 matrix, the following corollary to Theorem 4 holds.

Corollary: Any 2×2 matrix $R^{(2)}$ can be brought into a triangular form by an involutional transformation with an IUH matrix.

The proof is simple. Let $\binom{\alpha}{\beta}$ be an eigenvector of $R^{(2)}$ where α can be made real. Then the required involutional matrix Y is given by

$$Y = (\alpha^2 + |\beta|^2)^{-1/2} \begin{pmatrix} \alpha & \beta^* \\ \beta & -\alpha \end{pmatrix}, \qquad (3.10)$$

where β^* is the complex conjugate of β. Q.E.D.

Before ending this section we shall go back to the problem of transforming a matrix B into its complement $B' = (\lambda_1 + \lambda_2) I - B$. This problem is interesting since $T_{BB'} = 0$ by definition. When B and B' are equivalent and $\lambda_1 \neq \lambda_2$, we have $n_1 = n_2 = m$ so that they are matrices of order $2m \times 2m$. Let us assume further that B is a $Q^{(2m)}$ type matrix. Then there exists an involutional matrix Y_{BA} which brings B into a standard diagonal matrix Λ_B equivalent to B. Likewise there exists $Y_{B'A'}$ which brings B' into $\Lambda_{B'}$. It is obvious that an involutional matrix $J = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ brings Λ_B into $\Lambda_{B'}$. Therefore, there exists an involutional matrix $Y = Y_{BA} J Y_{B'A'}$ which brings B into B'. When B is involutional, $B' = -B$, so that Y given above satisfies $YB + BY = 0$, i.e., there exists an involutional matrix Y which anticommutes with a given involutional matrix B provided that B is a $Q^{(2m)}$ type (cf. Pauli's fundamental theorem[7] on Dirac's γ matrices).

We shall make one additional remark on the involutional matrix Y_Q defined by (3.8). It is obvious that Y_Q diverges as $\lambda_1 \to \lambda_2$ while $Y_Q^2 = 1$. It can, however, be shown that $Q^{(2m)}$ with two equal characteristics roots can be transformed into a triangular form by an involutional matrix with a mild condition that $\mathbf{bc} = e\mathbf{1} \neq 0$. In fact, the following involutional matrix,

$$Y_\epsilon = \epsilon^{-1} \begin{bmatrix} \mathbf{1} & (1-\epsilon^2)(-e)^{-1/2}\mathbf{b} \\ (-e)^{-1/2}\mathbf{c} & -\mathbf{1} \end{bmatrix} \qquad (3.11)$$

with an arbitrary parameter $\epsilon (\neq 0)$ brings $Q^{(2m)}$ with $\lambda_1 = \lambda_2$ into a triangular form (cf. Corollary to Theorem 4)

$$Y_\epsilon\, Q^{(2m)}\, Y_\epsilon = \begin{bmatrix} \frac{1}{2}(a+d)\mathbf{1} & \epsilon^2 \mathbf{1} \\ 0 & \frac{1}{2}(a+d)\mathbf{1} \end{bmatrix}. \qquad (3.12)$$

4. DIAGONALIZATION OF INVARIANT MATRICES.

We shall extend the theorems developed in Sec. 3 through a matrix representation of the group of the general homogeneous linear transformation in n dimensions, GL(n). For this purpose, it seems best to use the particular representation $S(R^{(n)})$ of a matrix $R^{(n)} \in \mathrm{GL}(n)$ introduced in I, for a variety of the symmetry properties of $R^{(n)}$ are directly reflected in $S(R^{(n)})$. This representation gives an invariant matrix[12] $S(R^{(n)})$ of $R^{(n)}$ in the sense that the elements of $S(R^{(n)})$ are polynomials in the elements of $R^{(n)}$. For definiteness, we shall describe the representation.

Let $\{f_\nu(r); \nu = 1, 2 \cdots p\}$ be a set of linearly independent functions of a vector $\mathbf{r} = (x_1, x_2 \cdots, x_n)$ in n dimensions. A representation $S(R^{(n)})$ is defined by

$$f_\nu(R^{(n)}\mathbf{r}) = \sum_{\mu=1}^{p} S(R^{(n)})_{\nu\mu} f_\mu(\mathbf{r}). \qquad (4.1)$$

It should be noted that the conventional representation due to Wigner[16] is the inverse transpose of the above representation. We take $f_\nu(\mathbf{r})$ as the qth degree monomials

$$f_\nu(\mathbf{r}) = \prod_{i=1}^{n} x_i^{\nu_i}/(\nu_i!)^{1/2}, \quad \nu_1 + \nu_2 + \cdots + \nu_n = q, \quad (4.2)$$

where ν stands for the set $\{\nu_1, \nu_2, \cdots, \nu_n;\ \sum_{i=1}^{n} \nu_i = q\}$. By definition $S(R^{(n)})$ is an invariant matrix of $R^{(n)}$ and the dimension of representation p is given by $\binom{n+q-1}{q}$. When $n = 2, p = q + 1$. To label the matrix we choose the order of ν in the decreasing order, regarding ν as a single number with n decimal points, i.e., $\nu = (\nu_1, \nu_2, \cdots, \nu_n)$. Then, when $q = 1$, the invariant matrix $S(R^{(n)})$ coincides with $R^{(n)}$.

It has been shown[1,3] that *when $R^{(n)}$ is a triangular matrix, $S(R^{(n)})$ is also triangular in shape similar to $R^{(n)}$.* Based on this theorem one can calculate the eigenvalues of $S(R^{(n)})$ simply by transforming $R^{(n)}$ into a triangular form by a suitable unitary transformation.[1] Let T be the transformation matrix which brings $R^{(n)}$ into a triangular matrix $R_t^{(n)}$ via a similarity transformation. Then we have for their representations,

$$S(T)^{-1} S(R^{(n)}) S(T) = S(R_t^{(n)}) \qquad (4.3)$$

where $S(R_t^{(n)})$ is triangular, similar in shape to $R_t^{(n)}$. According to Theorem 3, when $R_t^{(n)}$ is a $Q^{(2m)}$ type matrix with $\lambda_1 \neq \lambda_2$, it can be diagonalized by an involutional matrix Y_Q so that

$$S(Y_Q)S(Q^{(2m)})S(Y_Q) = S(\Lambda_Q), \quad n = 2m \tag{4.4}$$

where Λ_Q and $S(\Lambda_Q)$ are diagonal matrices, and $S(Y_Q)$ is involutional, $S(Y_Q)^2 = I$. This means that $S(Q^{(2m)})$ is also diagonalized by an involutional transformation. In a special case when $R = R^{(2)} \in \mathrm{GL}(2)$, $S(R) \in \mathrm{GL}(q+1)$ where q is the degree of the polynomial basis. From Theorem 2 of I, it can easily be shown that

$$S(\widetilde{R}) = \widetilde{S}(R), \quad S(R^*) = S(R)^*, \tag{4.5}$$

where $\widetilde{R}$ and R^* are the transpose and the complex conjugate of R, respectively. From (4.5) and the fact that $S(R)$ is a representation of R, we can state that if R is real, symmetric, normal, Hermitian, unitary, triangular, or diagonal, then so is $S(R)$. Combining this result with the corollary to Theorem 4, we can state that *$S(R)$ for any $R \in \mathrm{GL}(2)$ can be brought into a triangular form by a similarity transformation with an* IUH *(involutional, unitary, and Hermitian) matrix.*

The explicit form of $S(R^{(n)})$ has been given by the author for $n = 2$ in I and by Santhanam *et al.*, for $n = 3$.

5. ILLUSTRATIVE EXAMPLES FOR $Q^{(2m)}$

We have shown some important applications of the present formalism of involutional transformations on the relativistic theory of an electron in the end of Sec. 2 and on the representative theory of GL(n) in Sec. 3. In this section we shall apply the involutional transformations of $Q^{(2m)}$ type matrices for the diagonalizations of the Hamiltonians which are quadratic in the "particle" creation and annihilation operators in solid state physics.[13] It is well known that such diagonalizations require nonunitary transformations owing to the fact that the particle operators are not Hermitian. Through typical examples we shall show that such nonunitary transformations may be characterized by involutional transformations.

The examples considered are the Bogoliubov Hamiltonian which describes the phonon field in a superfluid or a magnon field[14] and the Hopfield Hamiltonian[15] which describes the polariton-photon field. In actual calculations it seems best to diagonalize the matrix which describes the equation of motion of the particle operators in the field, since it contains only commuting operators.

A. The Bogoliubov Hamiltonian[14]

$$H = \sum_k \{ s(k^2)(a_k^+ a_k + \tfrac{1}{2}) + \tfrac{1}{2} t(k^2)(a_k a_{-k} + a_{-k}^+ a_k^+) \tag{5.1}$$

where $s(k^2)$ and $t(k^2)$ are scalar functions, a_k^+ and a_k are the creation and annihilation operators which satisfy the commutation relations $[a_k, a_l^+] = \delta_{kl}$. The problem is to find a suitable linear combination α_k of a_k and a_{-k}^+ which reduces H into the canonical form

$$H = \sum_k \epsilon(k^2)(\alpha_k^+ \alpha_k + 1); \quad [\alpha_k, \alpha_k^+] = 1. \tag{5.2}$$

The equation of motion is

$$i \begin{bmatrix} \dot{a}_k \\ \dot{a}_{-k}^+ \end{bmatrix} = R^{(2)} \begin{bmatrix} a_k \\ a_{-k}^+ \end{bmatrix}; \quad R^{(2)} = \begin{bmatrix} s & t \\ -t & -s \end{bmatrix}, \tag{5.3}$$

where for simplicity the k^2 dependence of the matrix elements are suppressed and $\hbar = 1$. It is evident that $R^{(2)}$ is an involutional matrix with eigenvalues $\pm\epsilon$,

$$\epsilon = (s^2 - t^2)^{1/2} > 0. \tag{5.4}$$

Accordingly, it can be diagonalized via an involutional transformation:

$$Y_R^{(2)} R^{(2)} Y_R^{(2)} = \begin{bmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{bmatrix}; \quad Y_R^{(2)} = \begin{bmatrix} u & v \\ -v & -u \end{bmatrix}, \tag{5.5}$$

where use has been made of (3.7) and

$$u = [(s+\epsilon)/2\epsilon]^{1/2}, \quad v = [(s-\epsilon)/2\epsilon]^{1/2}; \quad u^2 - v^2 = 1. \tag{5.5'}$$

It is noted that u and v satisfy a strange normalization condition which, however, is simply due to det $Y_R^{(2)} = -1$ (cf 3.8′).

The required linear combination α_k is given by $Y_R^{(2)}$ as follows:

$$\begin{bmatrix} q_1 \alpha_k \\ q_2 \alpha_k^+ \end{bmatrix} = \begin{bmatrix} u & v \\ -v & -u \end{bmatrix} \begin{bmatrix} a_k \\ a_{-k}^+ \end{bmatrix}, \tag{5.6}$$

where q_1 and q_2 are scaling factors to be determined by the commutation relation, $[\alpha_k, \alpha_k^+] = 1$. With $q_1 = 1$, $q_2 = -1$, the above linear combination reduce to

$$\alpha_k = u\, a_k + v\, a_{-k}^+ \quad \text{or} \quad a_k = u\, \alpha_k - v\, \alpha_{-k}^+. \tag{5.7}$$

B. The polariton-photon Hamiltonian by Hopfield[14]

$$\begin{aligned} H = \sum_k \{ & f(k^2)(a_k^+ a_k + \tfrac{1}{2}) + g(k^2)(b_k^+ b_k + \tfrac{1}{2}) \\ & - ih(k^2)(a_k^+ b_k - a_k b_k^+ \\ & - a_k b_{-k} + a_k^+ b_{-k}^+)\}, \end{aligned} \tag{5.8}$$

where $f(k^2)$, $g(k^2)$, $h(k^2)$ are scalar functions of k^2; a_k, a_k^+ and b_k,
b_k^+ are two independent sets of Bose operators coupled through $h(k^2)$. The polarization dependence of these operators is suppressed for simplicity. The equation of motion is

$$i \begin{bmatrix} \dot{a}_k \\ \dot{a}_{-k}^+ \\ \dot{b}_k \\ \dot{b}_{-k}^+ \end{bmatrix} = R^{(4)} \begin{bmatrix} a_k \\ a_{-k}^+ \\ b_k \\ b_{-k}^+ \end{bmatrix}, \quad R^{(4)} = \begin{bmatrix} f & 0 & ih & ih \\ 0 & -f & ih & ih \\ -ih & ih & g & 0 \\ ih & -ih & 0 & -g \end{bmatrix}, \tag{5.9}$$

where we have suppressed the k^2 dependence of f, g, and h.

The matrix $R^{(4)}$ is traceless but it is not involutional. The characteristic equation is given by

$$[R^{(4)}]^4 - (f^2 + g^2)[R^{(4)}]^2 + (f^2 g^2 - 4fgh^2)I = 0, \tag{5.10}$$

which has four distinct roots $\pm\epsilon_1$ and $\pm\epsilon_2$. To use the theorems introduced in the foregoing sections, we regard it as the quadratic equation for $[R^{(4)}]^2$, which is a $Q^{(4)}$ type defined in Theorem 3.

$$[R^{(4)}]^2 = \begin{bmatrix} f^2\mathbf{1} & i\Gamma \\ -i\Gamma & g^2\mathbf{1} \end{bmatrix}, \Gamma = \begin{bmatrix} f+g & f-g \\ -f+g & -f-g \end{bmatrix}, \tag{5.10'}$$

where Γ is obviously a 2×2 involutional matrix. Thus, from Theorem 3, $[R^{(4)}]^2$ can be diagonalized by an involutional transformation as follows:

$$Z^{(4)}[R^{(4)}]^2 Z^{(4)} = \begin{bmatrix} \epsilon_1^2\mathbf{1} & 0 \\ 0 & \epsilon_2^2\mathbf{1} \end{bmatrix}, \tag{5.11}$$

where the explicit form of $Z^{(4)}$ can be written down immediately from (3.8). The above equation means that $Z^{(4)}R^{(4)}Z^{(4)}$ is a direct sum of two involutional matrices of order 2×2,

$$Z^{(4)}R^{(4)}Z^{(4)} = \begin{bmatrix} R_1^{(2)} & 0 \\ 0 & R_2^{(2)} \end{bmatrix}; \quad [R_\nu^{(2)}]^2 = \epsilon_\nu^2\mathbf{1}, \nu = 1,2, \tag{5.12}$$

where $R_\nu^{(2)}$ have forms similar to $R^{(2)}$ of (5.3) and hence can be diagonalized likewise; $Y_\nu^{(2)}R_\nu^{(2)}Y_\nu^{(2)} =$ diag $(\epsilon_\nu, -\epsilon_\nu)$, $\nu = 1,2$. The overall transformation W which diagonalizes $R^{(4)}$ via

$$W^{-1}R^{(4)}W = \mathrm{diag}(\epsilon_1, -\epsilon_1, \epsilon_2, -\epsilon_2) \tag{5.13}$$

is given by a product of two involutional matrices,

$$W = Y^{(4)}Z^{(4)}; \quad Y^{(4)} = \begin{bmatrix} Y_1^{(2)} & 0 \\ 0 & Y_2^{(2)} \end{bmatrix}, \tag{5.13'}$$

with $W^{-1} = Z^{(4)}Y^{(4)}$. The transformation of the particle operators is given by

$$\begin{bmatrix} \alpha_k \\ -\alpha^+_{-k} \\ \beta_k \\ -\beta^+_{-k} \end{bmatrix} = W \begin{bmatrix} a_k \\ a^+_{-k} \\ b_k \\ b^+_{-k} \end{bmatrix} \tag{5.14}$$

which diagonalizes the Hamiltonian in the form,

$$H = \sum_k \{\epsilon_1(k^2)(\alpha_k^+\alpha_k + \tfrac{1}{2}) + \epsilon_2(k^2)(\beta_k^+\beta_k + \tfrac{1}{2})\}. \tag{5.15}$$

ACKNOWLEDGMENTS

The authors wishes to thank Dr. B. K. Oh and Professor J. A. Poole at Temple University for reading the manuscript and giving helpful suggestions.

[1] Shoon K. Kim, J. Math. Phys. **10**, 1225 (1969).
[2] M.E. Fisher, Phys. Rev. **113**, 939 (1959); A. Ramakrishnan, J. Math. Anal. Appl. **20**, 9 (1967); Shoon K. Kim, J. Math. Phys. **9**, 1705 (1968).
[3] T.S. Santhanam, P.S. Chandrasekaran and N. B. Menon, J. Math. Phys. **12**, 377 (1971).
[4] P.A.M. Dirac, *Quantum Mechanics* (Oxford U.P., Oxford, England, 1958), 4th Edition.
[5] E.P. Wigner, private communication.
[6] Shoon K. Kim, to be published.
[7] W. Pauli, Ann. Inst.Henri Poincaré **6**, 109 (1936).
[8] R. Brauer and H. Weyl, Am. J. Math, **57**, 425 (1935).
[9] M.E. Rose, *Relativistic Electron Theory* (Wiley, New York, 1961).
[10] L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950); M. Cini and B. Touschek, Nuovo Cimento **7**, 422 (1958); S.K. Bose, A. Gamba, and E.C.G. Sudarshan, Phys. Rev. **113**, 166 (1959).
[11] L.C. Biedenharn, Phys. Rev. **126**, 845 (1961).
[12] D. E. Littlewood, *The Theory of Group Characters and Matrix Representations of Groups* (Clerendon, Oxford, England, 1950), 2nd edition.
[13] C. Kittel, *Quantum Theory of Solids* (*Wiley, New York*, 1963).
[14] N. Bogoliubov, J. Phys. USSR **11**, 23 (1947); also reprinted in D. Pines, *Many Body Problem* (Benjamin, New York, 1961), p. 292. See also Ref. 13, p. 26.
[15] J.J. Hopfield, Phys. Rev. **112**, 1555 (1958). See also Ref. 13, p. 44.
[16] E.P. Wigner *Group Theory* (Academic, New York, 1959).

A new method of matrix transformations. II. General theory of matrix diagonalizations via reduced characteristic equations and its application to angular momentum coupling

Shoon K. Kim
Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122

(Received 15 February 1979; accepted for publication 27 April 1979)

A general formalism is given to construct a transformation matrix which connects two matrices A and B of order $n \times n$ satisfying any given polynomial equation of degree r, $p^{(r)}(x) = 0$; $r \leq n$. The transformation matrix T_{AB} is explicitly given by a polynomial of degree $(r-1)$ in A and B based on $p^{(r)}(x)$. A special case where B is a diagonal matrix Λ equivalent to A leads to the general theory of matrix diagonalizations with the transformation matrix $T_{A\Lambda}$, which can be made nonsingular with a proper choice of Λ. In another special case where B is a constant matrix with the constant being a simple root λ_ν of $p^{(r)}(x)$, the transformation matrix T_{AB} reduces to the idempotent matrix P_ν belonging to the eigenvalue λ_ν of A. Based on the relation which exists between $T_{A\Lambda}$ and P_ν, one can construct a transformation matrix U which is more effective than $T_{A\Lambda}$ and becomes unitary when A is Hermitian. Illustrative examples of the formalism are given for the problem of angular momentum coupling.

1. INTRODUCTION

In the previous work [1] (which will be referred as I), the author has developed a theory of matrix transformation which connects two equivalent matrices of any order satisfying a quadratic equation. It has been shown that the involutional transformations play fundamental roles in transforming matrices frequently encountered in physics; for example, Dirac's Hamiltonian in the relativistic theory of an electron and the field Hamiltonian in solid state physics. In the present work we shall develop a general theory of constructing a transformation matrix T_{AB} which connects any two matrices A and B of order $n \times n$ via $AT_{AB} = BT_{AB}$, provided that A and B satisfy any given polynomial equation $p^{(r)}(x) = 0$ of degree $r \leqslant n$. The matrix T_{AB} will be given explicitly by a polynomial of $(r-1)$ degree in A and B (Theorem 1). Hereafter, we shall call T_{AB} the characteristic transformation matrix of A to B.

The present formalism contains the method of the idempotent matrix [2,3] (or the projection operator method) for constructing eigenfunctions of the matrix A as a special case where B is a constant matrix satisfying the reduced characteristic equation of A. The earliest work in this line is known as the Liverrier's method in Faddeev's modification in computational methods of linear algebra. [2]

The most important special case of the present formalism occurs when the reduced characteristic equation of A has no multiple roots. In this case, A is diagonalizable by similarity transformation. Let B be a diagonal matrix Λ equivalent to A. Then it will be shown that there exists a nonsingular characteristic transformation matrix $T_{A\Lambda}$ with a proper choice of Λ (Theorem 2). This then provides a general theory of matrix diagonalizations (Sec. 3).

Based on the fundamental theorems of matrix transformation given in Sec. 2 and 3, the mathematical properties of the characteristic transformation matrices $T \equiv T_{A\Lambda}$ and $\hat{T} \equiv T_{\Lambda A}$ will be discussed in Sec. 3. One of the most important properties is that $\hat{T}T$ is expressed by a direct sum of submatrices which are common to T and $\hat{T}$. Based on this property we shall construct a new transformation matrix U which is more convenient than T or $\hat{T}$ and becomes unitary when A is Hermitian (Sec. 4).

In Sec. 5, we shall give some very simple illustrative examples for the present formalism on the matrix diagonalizations. The examples chosen are concerned with the coupling coefficients for angular momentum wave functions. These are chosen primarily because the eigenvalues of the resultant angular momentums are known separately from the vector addition theorem. For purposes of comparison, some of the examples are taken from those introduced previously by Löwdin [3] in his work on the projection operator method applied to the angular momentum wave functions.

The present formalism of matrix transformations becomes more effective when the degree of the reduced characteristic equation is lower irrespective of the multiplicities of the characteristic roots.

2. THE FUNDAMENTAL THEOREMS

Before introducing the basic theorems we shall give some preparations. Let A be a matrix of order $n \times n$ which satisfies a monic polynomial equation of degree r $(\leqslant n)$,

$$p^{(r)}(x) = x^r + c_1 x^{r-1} + \cdots + c_r = 0, \tag{2.1}$$

with constant coefficients. From these coefficients we define a k^{th} degree polynomial by

$$x^{(k)} = x^k + c_1 x^{k-1} + \cdots + c_k, \quad k = 0, 1, \ldots, r. \tag{2.1'}$$

Then, $x^{(r)} = p^{(r)}(x)$ so that $A^{(r)} = p^{(r)}(A) = 0$. Let $M(n \times n, p^{(r)}(x))$ be a set of $n \times n$ matrices, every member of which satisfies $p^{(r)}(x) = 0$. Then, we can write for the matrix A

$$A \in M(n \times n, p^{(r)}(x)). \tag{2.2}$$

When $p^{(r)}(x) = 0$ is the equation of the least degree satisfied by the matrix A, it is called the reduced characteristic equation of A. To avoid confusion, the characteristic

 0022-2488/79/112159-09$01.00

polynomial of A given by the secular determinant is denoted by

$$D^{(n)}(x) = \det[xI - A], \tag{2.2'}$$

where I is the $n \times n$ unit matrix. The multiplicity of a root λ_ν of this equation is called the degeneracy of the eigenvalue λ_ν of A. With these preparations we state the basic theorem of matrix transformations:

Theorem 1: Let $A, B \in M(n \times n, p^{(r)}(x))$ with $p^{(r)}(x)$ of (2.1). Then A and B are connected via

$$AT_{AB} = T_{AB}B, \tag{2.3}$$

with T_{AB} given by

$$T_{AB} = \sum_{k=0}^{r-1} A^{r-1-k}B^{(k)} = \sum_{k=0}^{r-1} A^{(k)}B^{r-1-k}, \tag{2.4}$$

where $A^{(k)}$ and $B^{(k)}$ are the k^{th} degree polynomials defined by (2.1'). If T_{AB} is nonsingular, then the matrices A and B are equivalent.

The proof is elementary. The second equality of (2.4) is an identity which follows from the definition of $x^{(k)}$ and the rearrangement of the summations. Equation (2.3) follows from $A^{(r)} = B^{(r)}$ and the recursion formula,

$$x^{(k)} = xx^{(k-1)} + c_kI; \quad k = 1,2,\cdots,r-1,$$
$$x^{(0)} = I,$$

where x stands for A and B. Q.E.D.

It should be noted that the above theorem can be generalized by replacing the assumption $A^{(r)} = B^{(r)} = 0$ by $A^{(r)} = B^{(r)}$, since only the latter condition is needed for the proof. This generalization may be useful when r is small. However, we shall not consider this generalization any further in the present work.

It is obvious that if A and B are not equivalent then T_{AB} is singular and also that the equivalence of A and B does not guarantee that T_{AB} is nonsingular. We may call T_{AB} the characteristic transformation matrix of A to B which may or may not be singular. It should be noted that $T_{AB} \neq T_{BA}$ in general except when A and B commute or $r = 2$. For the latter case

$$T_{AB} = T_{BA} = A + B - (\lambda_1 + \lambda_2)I, \tag{2.5}$$

where λ_1 and λ_2 are the characteristic roots of A (B). This simple special case has many important applications which have been extensively discussed in the previous work.[1]

From (2.3) we can show that the product $T_{AB}T_{BA}$ commutes with A since

$$T_{AB}BT_{BA} = T_{AB}T_{BA}A = AT_{AB}T_{BA}. \tag{2.6}$$

If A and B are Hermitian, we have $T_{AB} = T_{BA}^+$ where T_{BA}^+ is the Hermitian conjugate of T_{BA} so that $T_{AB}T_{BA}$ is positive semidefinite.

The more general transformation matrix V_{AB} which connects A to B through

$$AV_{AB} = V_{AB}B \tag{2.7}$$

is given by

$$V_{AB} = F_AT_{AB} = T_{AB}F_B, \tag{2.7'}$$

where F_A (F_B) is a function of A (B). In practical applications it is often useful to introduce the special case where $F_A = A$, $F_B = B$. Then,

$$V_{AB} = AT_{AB} = T_{AB}B = -c^r + \sum_{k=1}^{r-1} AA_{k-1}B^{(r-k)}$$
$$= -c^r + \sum_{k=1}^{r-1} A^{(r-k)}B_{k-1}B, \tag{2.8}$$

which becomes,[1] for $r = 2$,

$$V_{AB} = AB - \lambda_1\lambda_2 I. \tag{2.8'}$$

It is most desirable to give some simple criteria for the existence of a nonsingular characteristic transformation matrix T_{AB}. We have established such a criterion in a simple special case where A, B, and T_B are involutional in the previous work I. We shall also establish it for the most important special case of a matrix diagonalization in Sec. 3. However, even if T_{AB} is singular, Theorem 1 can provide very significant consequences in some cases. A typical example is the case where T_{AB} becomes an indempotent matrix with B being a constant matrix. We shall now discuss this special case.

A. Idempotent matrix

Let $A, B \in M(n \times n, p^{(r)}(x))$ and $\lambda_1, \lambda_2, \cdots, \lambda_r$ be the roots of the polynomial $p^{(r)}(x)$, of which some of the roots could be equal. Then, if B is a constant matrix equal to $\lambda_\nu I$ where λ_ν is one of the characteristic roots of A, it can be shown that the characteristic transformation matrix T_{AB} equals the quotient of $p^{(r)}(x)$ and $(A - \lambda_\nu I)$, namely,

$$T_{AB} = \prod_{\sigma \neq \nu}^{r} (A - \lambda_\sigma I), \quad B = \lambda_\nu I. \tag{2.9}$$

If we asume further that $p^{(r)}(x)$ is the reduced characteristic polynomial of A, and λ_ν is its simple root, then T_{AB} becomes proportional to the idempotent matrix P_ν belonging to the eigenvalue λ_ν of A defined by

$$P_\nu = \chi_\nu^{-1} \prod_{\sigma \neq \nu}^{r} (A - \lambda_\sigma I), \quad \chi_\nu = \prod_{\sigma \neq \nu}^{r} (\lambda_\nu - \lambda_\sigma) \neq 0. \tag{2.10}$$

Let the degeneracy of the eigenvalue λ_ν be n_ν. Then, there exist n_ν and only n_ν linearly independent columns (rows) in P_ν which provide all the eigenvectors of A belonging to the eigenvalue λ_ν, since the rank of P_ν is n_ν, which can be seen easily from Jordan's canonical from of P_ν (see also 3.8). The calculation of these eigenvectors is facilitated by the following formula,

$$P_\nu = \chi_\nu^{-1} \prod_{\sigma \neq \nu}^{r} (A - \lambda_\sigma I), \quad \chi_\nu = \prod_{\sigma \neq \nu}^{r} (\lambda_\nu - \lambda_\sigma) \neq 0. \tag{2.10}$$

where use has been made of (2.4), (2.9) and (2.10). If all the roots of the reduced characteristic polynomial $p^{(r)}(x)$ are distinct, then we have the orthogonality and the closure relations

$$P_\nu P_\mu = \delta_{\nu\mu}P_\nu, \quad \sum_{\nu=1}^{r} P_\nu = I, \quad \nu,\mu = 1,2,\cdots,r, \tag{2.11}$$

where $\delta_{\nu\mu}$ is Kronecker's delata. In particular, when $n = r$ and $p^{(r)}(x)$ becomes the characteristic polynomial $D^{(n)}(x)$ of A, and we have a set of convenient recursion formulas,

$$A^{(k)} = AA^{(k-1)} + c_k I, \quad c_k = -k^{-1}\mathrm{tr}AA^{(k-1)}, \quad k = 1,2,\cdots,n, \tag{2.12}$$

where tr ⋯ denotes the trace. The set of Eq. (2.10′) and (2.12) constitute Liverrier's method in Faddeev's modification [2] for calculating the eigenvectors of A.

3. THEORY OF MATRIX DIAGONALIZATIONS

We shall now develop a general theory of matrix diagonalizations based on Theorem 1. The condition for a matrix A to be diagonalized by a similarity transformation is that the reduced characteristic equation of A has no multiple roots. This condition may be expressed as follows:

$$A \in M\,(n \times n, p^{[r]}(x))$$

where $p^{[r]}(x)$ denotes a polynomial of degree r with all distinct roots. With this preparation we state the basic theorem for matrix diagonalizations.

Theorem 2: Let $A \in M\,(n \times n, p^{(r)}(x))$ and Λ be a diagonal matrix equivalent to A. Then, there exists at least one diagonal matrix Λ which makes the characteristic transformation matrices $T \equiv T_{A\Lambda}$ and/or $\hat{T} \equiv T_{\Lambda A}$ nonsingular, so that

$$T^{-1}AT = \Lambda, \quad \text{and/or} \quad \hat{T}A\hat{T}^{-1} = \Lambda. \tag{3.1}$$

The proof of this theorem is somewhat involved. Before proceeding with the proof we shall give some preparations. From (2.4), the explicit forms of T and $\hat{T}$ are give by

$$T = T_{A\Lambda} = \sum_{k=0}^{r-1} A^{r-1-k}\Lambda^{(k)} = \sum_{k=0}^{r-1} A^{(k)}\Lambda^{r-1-k}$$
$$\hat{T} = T_{\Lambda A} = \sum_{k=0}^{r-1} \Lambda^{r-1-k}A^{(k)} = \sum_{k=0}^{r-1} \Lambda^{(k)}A^{r-1-k} \tag{3.2}$$

In a special case where $r = 2$, we have

$$T = \hat{T} = A + \Lambda - (\lambda_1 + \lambda_2)I, \tag{3.3}$$

which has been extensively studied in the previous work I. Hereafter, we shall call T and $\hat{T}$ the characteristic transformation matrices of A even if they are singular. It is obvious that the column (row) vectors are eigenvectors of A even if $T(\hat{T})$ is singular, unless they are null vectors.

Now a diagonal matrix Λ equivalent to A is characterized by a sequence of the whole set of n characteristic roots of A. Let the distinct characteristic roots of A be $\lambda_1, \lambda_2, \cdots, \lambda_r$ and their respective degeneracies be $n_1, n_2, \cdots n_r$. Then $n_1 + n_2 + \cdots + n_r = n$. The matrix elements of Λ may be expressed by

$$\Lambda_{st} = \lambda_{\nu = g(s)}\delta_{st} \quad \nu = 1,2,\cdots,r \quad s,t = 1,2,\cdots,n \tag{3.4}$$

where δ_{st} is the Kronecker delta and $\nu = g(s)$ defines the sequence of the characteristic roots on the principal diagonal of Λ. The inverse $g^{-1}(\nu)$ is a multivalued function of ν which gives the n_ν indices of the columns or the rows belonging to λ_ν. It is obvious that the intersection of two sets of indices belonging to two different roots λ_ν and $\lambda_{\nu'}$ are null,

$$g^{-1}(\nu) \cap g^{-1}(\nu') = \emptyset, \quad \text{for } \nu \neq \nu'. \tag{3.5}$$

This simple and obvious property plays an important role in the proof. If a sequence $\nu = g(s)$ is such that the equal roots are placed on the consecutive positions, then the corresponding Λ is called a standard form of Λ.

Now, let us denote the j^{th} column (row) vectors of any matrix M by $M_{\cdot j}$ ($M_{j\cdot}$). Then we can state that $T_{\cdot j}$ ($\hat{T}_{j\cdot}$) belongs to the eigenvalue λ_ν of A if $\nu = g(j)$. Moreover, from (2.10′) and (3.2) we have

$$T_{\cdot j} = \chi_\nu [P_\nu]_{\cdot j}, \quad T_{j\cdot} = \chi_\nu [P_\nu]_{j\cdot}; \quad \nu = g(j) \quad j = 1,2,\cdots n, \tag{3.6}$$

which are the crucial relations for the proof of the existence of a nonsingular $T(\hat{T})$. One has to show that all the n columns (rows) of $T_{\cdot j}$ ($T_{j\cdot}$) can be made linearly independent with a proper choice of the sequence $\nu = g(j)$.

By assumption, there exists a nonsingular matrix G which diagonalizes A. Then G also diagonalizes P_ν,

$$G^{-1}P_\nu G = D_\nu \quad \nu = 1,2,\cdots,r, \tag{3.7}$$

where D_ν is a diagonal matrix which has n_ν unit elements,

$$[D_\nu]_{ij} = \delta_{\nu,h(i)}\delta_{ij}, \tag{3.8}$$

with $\nu = h(i)$ which defines the sequence of λ_ν in the diagonal matrix $D = G^{-1}AG$. Substitution of $P_\nu = GD_\nu G^{-1}$ into the first of (3.6) yields

$$T_{\cdot j} = \chi_\nu \sum_{k \in h^{-1}(\nu)} G_{\cdot k}[G^{-1}]_{kj}; \quad j \in g^{-1}(\nu), \tag{3.9}$$

where $G_{\cdot k}$ is the k^{th} column vector of G.

Since G is nonsingular, all the column vectors $G_{\cdot k}$ are linearly independent. Hence, if the $n_\nu \times n_\nu$ coefficient matrix $\Delta^{(\nu)}$ of (3.9) with elements.

$$\Delta^{(\nu)}_{kj} = [G^{-1}]_{kj}; \quad k \in h^{-1}(\nu), \quad j \in g^{-1}(\nu), \tag{3.10}$$

is nonsingular then the n_ν column vector $T_{\cdot j}, j \in g^{-1}(\nu)$, are linearly independent. Now, det $[\Delta^{(\nu)}]$ is nothing but a minor of det$[G^{-1}]$. Thus, if we apply the Laplace theorem on expansion of a determinant to det $[G^{-1}]$ with respect to n_ν columns given by $h^{-1}(\nu)$ changing ν successively from $\nu = 1$ to $\nu = r$, we must have at least one set of nonvanishing minors such that

$$\prod_{\nu=1}^{r} \det[\Delta^{(\nu)}] \neq 0, \tag{3.11}$$

with conditions

$$h^{-1}(\nu) \cap h^{-1}(\nu') = \emptyset, \quad g^{-1}(\nu) \cap g^{-1}(\nu') = \emptyset, \quad \text{for } \nu \neq \nu'. \tag{3.12}$$

This means that there exists at least one sequence $\nu = g(j)$ of λ_ν defining Λ which makes T nonsingular for a given sequence $\nu = h(k)$.

An analogous proof holds for $\hat{T}$. In fact, the row vector $\hat{T}_{i\cdot}$ is given by

$$\hat{T}_{i\cdot} = \chi_\nu \sum_{k \in h^{-1}(\nu)} G_{ik}[G^{-1}]_{k\cdot}, \quad i \in g^{-1}(\nu), \tag{3.13}$$

proceeding as before we can find a set of nonvanishing minors of det$[G]$ which determines a sequence $\nu = g'(i)$ of λ_ν defining Λ which makes $\hat{T}$ nonsingular. Q.E.D.

It is noted that the proper sequence $\nu = g(i)$ and $\nu = g'(i)$ for T and $\hat{T}$ respectively need not be the same except for $r = 2$ unless there exists a certain symmetry in A. In most cases, however, one can find such a sequence for which both

T and $\hat{T}$ are nonsingular. The number of ways defining $\nu = g(i)$ which makes T nonsingular is in anywhere between 1 and $n!/(n_1!n_2!\cdots, n_r!)$. The minimum number occurs in the exceptional case where A is triangular. In this case the transformation matrix T and $\hat{T}$ are also triangular similar in shape to that of A, and the diagonal elements of T and $\hat{T}$ are given by

$$T_{ss} = \hat{T}_{ss} = \sum_{\sigma \neq g(s)}^{r} (A_{ss} - \lambda_\sigma); \quad s = 1,2,\cdots,r, \tag{3.14}$$

where $g(s)$ describes the assumed sequence of the characteristic roots in A (see 4.2). Thus, to obtain nonsingular $T(\hat{T})$ one must take Λ equal to the diagonal part of A, i.e.,

$$A_{ss} = \lambda_{g(s)}.$$

It is noted here that an alternate form of $T(\hat{T})$ is of interest which has the form analogous to the indempotent matrix P_ν. Let θ be a cyclic permutation of $\lambda_1,\lambda_2,\cdots,\lambda_r$ on the principal diagonal of Λ;

$$\theta = (\lambda_1,\lambda_2,\cdots,\lambda_r). \tag{3.15}$$

We apply θ on Λ repeatedly to generate a set of diagonal matrices Λ_i by

$$\Lambda = \Lambda_1, \Lambda_2, = \theta\Lambda_1, \cdots, \Lambda_r = \theta\Lambda_{r-1}. \tag{3.16}$$

Then using the relation between the coefficients and the roots of $p^{[r]}(x) = 0$ we can show that

$$p^{[r]}(A) = O_{A\Lambda} \prod_{i=1}^{h} (A - \Lambda_i) = 0_{\Lambda A} \prod_{i=1}^{r} (A - \Lambda_i) = 0, \tag{3.17}$$

where $O_{A\Lambda}$ and $O_{\Lambda A}$ are operators which order the products of A and Λ in the orders written in the subscripts; for example,

$$O_{A\Lambda}(A - \Lambda_1)(A - \Lambda_2) = A^2 - A(\Lambda_1 + \Lambda_2) + \Lambda_1\Lambda_2, \tag{3.18}$$

$$O_{\Lambda A}(A - \Lambda_1)(A - \Lambda_2) = A^2 - (\Lambda_1 + \Lambda_2)A + \Lambda_1\Lambda_2,$$

Then, (3.17) gives

$$T = O_{A\Lambda} \prod_{i=2}^{r} (A - \Lambda_i), \quad \hat{T} = O_{\Lambda A} \prod_{i=2}^{r} (A - \Lambda_i), \tag{3.19}$$

which satisfy $AT = T\Lambda$ and $\hat{T}A = \Lambda\hat{T}$, respectively. We can easily show that these satisfy the expansion formula of T and $\hat{T}$ given by (3.2).

When we have found a proper sequence of λ_ν in Λ which makes $T(\hat{T})$ nonsingular it is always possible to reduce Λ into a standard form where the equal roots are placed in the consecutive positions by unitary transformation of A and Λ which is merely a simultaneous renumbering of the columns and the rows of A and Λ. Hereafter, we shall assume such a renumbering has been performed so that the proper Λ takes a standard form unless otherwise specified. As one can see in the practical applications, a standard sequence is frequently a proper sequence for a given matrix A without any simultaneous renumbering.

4. THE BASIC PROPERTIES OF T AND $\hat{T}$

We shall discuss the basic properties of the characteristic transformation matrices T and $\hat{T}$ which stem from their relations with the idempotent matrices given by (3.6). Then, these properites will be used in Sec. 5 to construct a new transformation matrix U which is more effective then $T(\hat{T})$ in the sense that its inverse is written down immediately and it becomes unitary when A is Hermitian. Let $\nu = g(j)$ be any sequence of λ_ν in Λ equivalent to the matrix $A \in M$ ($n \times n$, $p^{[r]}(x)$). Then from (3.6) we have

$$[T\hat{T}]_{ij} = \begin{cases} 0, & \text{for } g(i) \neq g(j), \\ \chi_\nu T_{ij}^{(\nu)}, & \text{for } \nu = g(i) = g(j)' \end{cases} \tag{4.1}$$

where $T^{(\nu)}$ is a $n_\nu \times n_\nu$ submatrix common to T and $\hat{T}$

$$T_{ij}^{(\nu)} = T_{ij} = \hat{T}_{ij}; \quad i,j \in g^{-1}(\nu). \tag{4.2}$$

These are remarkable properties of T and $\hat{T}$. If we assume a standard form for Λ, we can express it in the form of a direct sum:

$$\Lambda = \lambda_1 I_1 \oplus \lambda_2 I_2 \oplus \cdots \oplus \lambda_r I_r, \tag{4.3}$$

where I_ν is the $n_\nu \times n_\nu$ unit matrix. Then $\hat{T}T$ and $\hat{T}AT$ also take the forms of direct sums:

$$\hat{T}T = \chi_1 T^{(1)} \oplus \chi_1 T^{(2)} \oplus \cdots \chi_r T^{(r)}, \tag{4.4}$$

$$\hat{T}AT = \lambda_1 \chi_1 T^{(1)} \oplus \lambda_2 \chi_2 T^{(2)} \oplus \cdots \oplus \lambda_r \chi_r T^{(r)}. \tag{4.5}$$

These equations suggest an alternate method of diagonalizing the matrix A when both T and $\hat{T}$ are nonsingular. We may regard $\chi_\nu T_{ij}^{(\nu)}$ as the scalar products of two sets of vectors $\{\hat{T}_{i\cdot}\}$ and $\{T_{\cdot j}\}$ belonging to λ_ν of A. Then the mutual orthogonalization of these two vector sets leads to the desired diagonalization of A (see Sec. 5).

From (4.4) we have for the determinant of $\hat{T}T$

$$\det \hat{T}T = \prod_{\nu=1}^{r} \chi_\nu^{n_\nu} \det T^{(\nu)}. \tag{4.6}$$

Thus, if both T and $\hat{T}$ are nonsingular, then all $T^{(\nu)}$, $\nu = 1,2,\cdots r$, are nonsingular and vice versa. If one of $T^{(\nu)}$ is singular, at least one of T and $\hat{T}$ must be singular. In a special case where one of $T^{(\nu)}$ is a null matrix and $\hat{T}$ is nonsingular, then all the column vectors of T belonging to the eigenvalue λ_ν are also null, because

$$\sum_{k=0}^{n} \hat{T}_{\cdot k} T_{kj} = 0, \quad \text{for all } j \in g^{-1}(\nu). \tag{4.7}$$

Hereafter, we shall assume that all $T^{(\nu)}$ are nonsingular unless otherwise specified.

If the matrix A is Hermitian, then $\hat{T} = T^+$ and $\hat{T}T$ becomes positive definite. Hence, the following quadratic forms with respect to a set of complex variables $\{x_i \in g^{-1}(\nu)\}$ for each ν becomes positive definite,

$$\chi_\nu \sum_{i,j \in g^{-1}(\nu)} x_i^* T_{ij}^{(\nu)} x_j = \sum_{k=1}^{n} \Big| \sum_i T_{ki} x_i \Big|^2 > 0, \tag{4.8}$$

where x_i^* is the complex conjugate of x_i and the number of the variables x_i is arbitrary as long as the indices i belong to $g^{-1}(\nu)$. Consequently, *all the principal minors of $\chi_\nu T^{(\nu)}$ are positive*; for example, the signs of all diagonal elements of $T^{(\nu)}$ are given by the sign of χ_ν:

$$\chi_\nu T_{ii}^{(\nu)} = \sum_k |T_{ik}|^2 > 0, \quad \nu = 1,2,\cdots,r. \tag{4.9}$$

If any one of the diagonal elements, T_{ii} is zero, then T becomes singular and the ith column and row become null. These properties, which come from (4.8), will be used to diagonalize $T^{(\nu)}$ by the Gaussian elimination procedure discussed in Sec. 5.

5. CONSTRUCTION OF A TRANSFORMATION MATRIX U

It has been shown by (4.4) and (4.5) in Sec. 4 that the complete diagonalization of the matrix A can be achieved by reducing each submatrix $\chi_\nu T^{(\nu)}$ of order $n_\nu \times n_\nu$ into the unit $n_\nu \times n_\nu$ matrix. Obviously, the present formalism can be used for this. However, it is more effective to use the method of successive elimination since $\chi_\nu T_{ij}^{(\nu)}$ can be regarded as the scalar products of two vector sets $\{\hat{T}_{i\cdot}\}$ and $\{T_{\cdot j}\}$ according to (4.4). The method is based on the well known Gaussian procedure, which transforms a square matrix $T^{(\nu)}$ into a triangular form by another triangular matrix of the opposite shape. It serves as an effective algorithm for the Schmidt orthogonalization process. Previously, Löwdin[3] used the elimination method in his well known work of the projection operator method in constructing an orthogonal set of angular momentum eigenfunctions, where he has based his argument on the idempotent nature of the projection operator and the "turn over rule" of the Hermitian character of the operator. We shall see that the present method is free from all these assumptions.

Let us introduce two nonsingular matrices $C^{(\nu)}$ and $\bar{C}^{(\nu)}$ of order $n_\nu \times n_\nu$ and diagonalize $T^{(\nu)}$ via $C^{(\nu)}T^{(\nu)}\bar{C}^{(\nu)}$. To this end, we put

$$S^{(\nu)} = T^{(\nu)}\bar{C}^{(\nu)}, \quad \bar{S}^{(\nu)} = C^{(\nu)}T^{(\nu)}. \tag{5.1}$$

Then, we require $S^{(\nu)}$ to be, say, a lower triangular form assuming an upper triangular form with unit diagonal elements for $\bar{C}^{(\nu)}$. Likewise, we require $\bar{S}^{(\nu)}$ to be upper triangular assuming a lower triangular form with unit diagonal elements for $C^{(\nu)}$. It is well known that there exists a unique solution for each pair of $\{S^{(\nu)}, \bar{C}^{(\nu)}\}$ and $\{\bar{S}^{(\nu)}, C^{(\nu)}\}$, if $T^{(\nu)}$ satisfies the condition that a certain number of the principal minors of $T^{(\nu)}$ are nonvanishing. This condition will be discussed later by (5.12). From (5.1) we have

$$C^{(\nu)}T^{(\nu)}\bar{C}^{(\nu)} = C^{(\nu)}S^{(\nu)} = \bar{S}^{(\nu)}\bar{C}^{(\nu)}, \tag{5.2}$$

which is diagonal, since the second equality means that the product of upper triangular matrices is equal to the product of lower triangular matrices;

$$[C^{(\nu)}T^{(\nu)}\bar{C}^{(\nu)}]_{ij} = \delta_{ij}\, S_{ii}^{(\nu)} = \delta_{ij}\, \bar{S}_{ii}^{(\nu)}; \quad i,j \in g^{-1}(\nu). \tag{5.3}$$

It is obvious that the above argument holds if we interchange all the upper and lower triangular forms.

When A is Hermitian, $T^{(\nu)}$ is also Hermitian so that it is only necessary to construct one of $S^{(\nu)}$ and $\bar{S}^{(\nu)}$ since

$$\bar{S}^{(\nu)} = [S^{(\nu)}]^+, \quad \bar{C}^{(\nu)} = [C^{(\nu)}]^+.$$

Now we define a pair of transformation matrices of order $n \times n$ by

$$S = T\bar{C}, \quad \bar{S} = C\hat{T}, \tag{5.4}$$

where $\bar{C}$ and C are direct sums of $\bar{C}^{(\nu)}$ and $C^{(\nu)}$,

$$\bar{C} = \sum_{\nu=1}^{r} \oplus \bar{C}^{(\nu)}, \quad C = \sum_{\nu=1}^{r} \oplus C^{(\nu)}. \tag{5.5}$$

Then S and $\bar{S}$ coincide with $S^{(\nu)}$ and $\bar{S}^{(\nu)}$ respectively in each degenerate subspace of λ_ν, and

$$\sum_k \bar{S}_{ik}S_{kj} = \delta_{ij}\,\chi_{g(i)}S_{ii}; \quad i,j = 1,2,\cdots,n, \tag{5.6}$$

where use has been made of (4.4) and (5.3)~(5.5), and $\chi_{g(j)}$ means χ_ν with $\nu = g(j)$.

Finally, the required transformation matrix U which diagonalizes A via $U^{-1}AU = \Lambda$ is given by

$$U_{ij} = S_{ij}N_j, \quad \{U^{-1}\}_{ij} = N_i\bar{S}_{ij}, \tag{5.7}$$

with

$$N_j^{-2} = \chi_{g(j)}S_{jj} \neq 0, \quad [U^{-1}]_{ii} = U_{ii} = 1/(N_i\chi_{g(i)}) \qquad i,j = 1,2,\cdots,n. \tag{5.8}$$

Here it is essential for the existence of U that $S_{jj} \neq 0$ for all j. As will be shown in (5.12), this condition is equivalent to the condition that the Gaussian elimination procedure assumed in (5.1) is valid. It is also noted that the argument of each normalization constant N_j has to be assigned appropriately, once for each index. If necessary, one may avoid this arbitrariness using S^{-1} given by

$$[S^{-1}]_{ij} = N_i^2\bar{S}_{ij} \tag{5.9}$$

instead of U^{-1}. In a special case when there exist no degeneracies in the eigenvalues of A, we have $S = T$ and $\bar{S} = \hat{T}$ so that

$$U_{ij} = T_{ij}N_j, \quad [U^{-1}]_{ij} = N_i\hat{T}_{ij}; \quad N_j^{-2} = \chi_{g(j)}T_{jj}; \qquad i,j = 1,2,\cdots,n. \tag{5.10}$$

Finally, when A is Hermitian, U becomes unitary and

$$N_j^{-2} = \chi_{g(j)}S_{jj} > 0, \tag{5.11}$$

which follows from (5.6) since $\bar{S} = S^+$.

Now we shall discuss the condition that the Gaussian elimination procedure assumed by (5.1) is valid. To this end, we first note that the diagonal elements $S_{ii}^{(\nu)}$ of (5.6) are related to the determinants of the leading submatrices of $T^{(\nu)}$ by

$$\prod_{t=1}^{q} = S_{a+t,a+t}^{(\nu)} = \det[T_{ij}^{(\nu)}; \quad i,j = a+1, a+2,\cdots,a+q] \qquad q = 1,2,\cdots,n_\nu, \tag{5.12}$$

where we have assumed for definiteness that $S^{(\nu)}$ is lower triangular and a is the minimum index of the columns and rows of T for $T^{(\nu)}$. It is well known that the Gaussian elimination procedure assumed by (5.1) is possible on condition[2] that all the determinants on the right hand side of (5.12) for $q = 1,2,\cdots, n_\nu$ are nonvanishing. We have already shown by (4.8) that this condition holds when the matrix A is Hermitian. In the general case, we have shown by (4.6) only that $\det T^{(\nu)} \neq 0$ for all ν with the assumption that both T and $\hat{T}$ are nonsingular. Accordingly, in order to satisfy the above condition it may be necessary to renumber the columns and rows of A and T in each degenerate subspace where the numbering has so far been arbitrary. If the renumbering does not work we simply take

$$C^{(\nu)} = [\chi_\nu T^{(\nu)}]^{-1}, \quad \bar{C}^{(\nu)} = I_\nu \tag{5.13}$$

in (5.1). Then $S = T$ and $\bar{S} = T^{-1}$ from (4.4), (5.4) and (5.5).

Frequently, we encounter the problem of simultaneous transformation of a matrix H which commutes with the matrix $A \in M\,(n \times n, p^{[r]}(x))$. Then, we have the following generalizationof (4.5),

$$[THT]_{ij} = \begin{cases} 0, & \text{for } g(i)\neq g(j), \\ \chi_\nu [HT]_{ij} = \chi_\nu [\hat{T}H]_{ij}, & \text{for } \nu = g(i) = g(j). \end{cases} \tag{5.14}$$

In terms of the transformation matrix U, these equations give

$$[U^{-1}HU]_{ij} = 0, \quad \text{for } g(i)\neq g(j),$$

$$[U^{-1}HU]_{ij} = (1/U_{ii}) \sum_{k<i} \sum_{k'>j} C^{(\nu)}_{ik} H_{kk'} U_{k'j} \tag{5.15}$$

$$= (1/U_{jj}) \sum_{k>i} \sum_{k'<j} [U^{-1}]_{jk} H_{kk'} \bar{C}^{(\nu)}_{k'j}, \quad \text{for } i,j\in q^{-1}(\nu),$$

where for definiteness it is assumed that in the subspace of λ_ν, $C^{(\nu)}$ and U are lower triangular and $\bar{C}^{(\nu)}$ and U^{-1} are upper triangular. From (5.15) one can seen that the off-diagonal elements of RHS of (5.15) do not contain any diagonal elements of H. This is important since it prevents the off-diagonal elements from becoming large. It was Löwdin [3] who first recognized a result analogous to this in his work on the projection operator method. The present treatment is more general and explicit.

6. ILLUSTRATIVE EXAMPLES

In order to give some illustrations of the present formalism we shall discuss some very simple examples. In the actual calculation of the transformation matrix T or $\hat{T}$, one may simplify the calculation by using the fact that any similarity transformation which diagonalizes a matrix $A \in M\,(n \times n, p^{[r]}(x))$ also diagonalizes its linear transform,

$$A \to aA + bI, \tag{6.1}$$

where a and b are constants and $a\neq 0$. It can be easily shown that the matrix U given by Eq. (5.7) is invariant for this linear mapping while the matrix $T\,(\hat{T})$ is mapped into $a^{r-1}T\,(\hat{T})$ where r is the degree of the reduced characteristic equation. Hereafter, we shall freely use such a linear mapping.

The examples chosen here are concerned with the coupling coefficients of the total angular momentum wave functions, for which the eigenvalues are known from the vector addition theorem. For purposes of comparison, some of the examples are taken from those previously introduced by Löwdin [3] in his well known work on the projection operator method.

We shall use a standard set of notations: $\mathbf{L}$ and $\mathbf{S}$ are orbital and spin angular momentum, respectively, and $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Their ladder operators are defined by $\mathbf{M}_\pm = \mathbf{M}_x \pm i\mathbf{M}_y$ where $\mathbf{M}$ stands for $\mathbf{J}$, $\mathbf{L}$, and $\mathbf{S}$.

A. A spin system [3]

Consider the total spin of a four electron system with $S_z = 1$. Let α and β be the elementary spinors. Then the basis of representation is given by

$$\phi_1 = \alpha\alpha\alpha\beta \quad \phi_2 = \alpha\alpha\beta\alpha \quad \phi_3 = \alpha\alpha\beta\alpha \quad \phi_4 = \beta\alpha\alpha\alpha, \tag{6.2}$$

where the particle coordinates in order are suppressed. The allowed total spin quantum numbers are $S = 1$ and 2 with degeneracies 3 and 1, respectively. To construct the eigenfunctions of $\mathbf{S}^2$, we shall take the matrix representation of $\|\mathbf{S}_-\ \mathbf{S}_+\|$ given by the basis of (6.2) as the matrix A:

$$A = \|\mathbf{S}_-\mathbf{S}_+\| = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \tag{6.3}$$

The relations between the total spin quantum number S, λ_ν, the eigenvalues of A, n_ν, the degeneracies, and χ_ν of (2.10) are given by

$$\begin{array}{lccccc} & S & \lambda_\nu & n_\nu & \chi_\nu & \\ \nu=1 & 1 & 0 & 3 & -4. \\ \nu=1 & 2 & 4 & 1 & 4 \end{array} \tag{6.4}$$

The reduced characteristic equation of A is

$$A^2 - 4A = 0. \tag{6.5}$$

Choosing the following standard form for the diagonal matrix,

$$\Lambda = \mathrm{diag}[0,0,0,4], \tag{6.6}$$

we can immediately write down the transformation matrices T and $\hat{T}$ using (3.3),

$$T = \hat{T} = A + \Lambda - 4I = \left[\begin{array}{ccc:c} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ \hdashline 1 & 1 & 1 & 1 \end{array}\right]. \tag{6.7}$$

One can easily see from this result that T is nonsingular for any sequence of the characteristic roots in Λ. The submatrices $T^{(1)}$ and $T^{(2)}$ corresponding to λ_1 and λ_2 are shown by the dotted lines in (6.7). They are symmetric since A is symmetric. Note that the signs of the diagonal elements of T coincide with those χ_1 and χ_2 satisfying (4.9). By successive elimination of the elements above the principal diagonal of T in the first three columns leaving the first column intact, we obtain the matrix S defined by (5.4) for the present case. Then normalizing each column of S using (5.7), we obtain the unitary transformation matrix U,

$$U = \begin{bmatrix} -\frac{3}{\sqrt{12}} & 0 & 0 & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & -\frac{2}{\sqrt{6}} & 0 & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix}. \tag{6.8}$$

From (6.4) and (6.6), the first three columns of U give the coupling coefficients for the eigenfunctions belonging to $\lambda_1 = 0$ or $S = 1$. The last column gives those for $\lambda_2 = 4$ or $S = 2$. These are in complete agreement with Löwdin's result based on the projection operator method. The present method is definitely simpler than his method for this example, since the latter requires construction of two idempotent matrices.

B. The atomic wave functions[3]

The problem is to construct the atomic state wave functions for a given electronic configuration. The present formalism may give a very general scheme to handle this problem. However, we shall postpone the general treatment and consider a typical example of constructing $\mathbf{L}^2$ eigenfunctions for a given configuration d^3. For simplicity, set $L_z = 2$, then the basis of representation is given by a set of the six Slater determinants:

$$\phi_1 = (21|\bar{1}),\quad \phi_2 = (20|0),\quad \phi_3 = (2\bar{1}|1),\quad \phi_4 = (10|1),\quad \phi_5 = (2\bar{2}|2),\quad \phi_6 = (1\bar{1}|2). \tag{6.9}$$

Here, the notation $(21|\bar{1})$, for example, is an abbreviation of the Slater determinant,

$$(21|\bar{1}) = (nd\,2\alpha,\ nd\,1\alpha,\ nd\bar{1}\beta), \tag{6.10}$$

where each one-electron function is characterized by four quantum numbers, $(n, l, m_l\ m_s)$. From the vector addition theorem we expect the total orbital angular momentum quantum number L to be 5, 4, 3, and 2 with degeneracies 1, 1, 2, and 2 respectively. We take

$$A = \tfrac{1}{2}\|\mathbf{L}_-\mathbf{L}_+\| - 5I = \begin{bmatrix} -2 & 3 & 0 & 0 & 0 & 0 \\ 3 & 1 & 3 & \sqrt{6} & 0 & 0 \\ 0 & 3 & 0 & \sqrt{6} & 2 & 2 \\ 0 & \sqrt{6} & \sqrt{6} & -1 & 0 & \sqrt{6} \\ 0 & 0 & 2 & 0 & -3 & 2 \\ 0 & 0 & 2 & \sqrt{6} & 2 & 0 \end{bmatrix}. \tag{6.11}$$

Then, the relations between L, λ_ν, n_ν, and χ_ν are given by

$$\begin{array}{lrrrr} & L & \lambda_\nu & n_\nu & \chi_\nu \\ \nu=1 & 5 & 7 & 1 & 540, \\ \nu=2 & 4 & 2 & 1 & -140, \\ \nu=3 & 3 & -2 & 2 & 108, \\ \nu=4 & 2 & -5 & 2 & -252, \end{array} \tag{6.12}$$

The reduced characteristic equation of A is given by

$$A^4 - 2A^3 - 39A^2 + 8A + 140\ I = 0. \tag{6.13}$$

Assuming a standard form for A,

$$A = \text{diag}\,[7, 2, -2, -2, -5, -5], \tag{6.14}$$

we obtain, from (3.2),

$$T = \begin{bmatrix} 18 & -36 & -27 & -9\sqrt{6} & 18 & 36 \\ 54 & -48 & 0 & 0 & -18 & -36 \\ 54 & 12 & 45 & -9\sqrt{6} & 78 & -12 \\ 18\sqrt{6} & 4\sqrt{6} & -9\sqrt{6} & 54 & -30\sqrt{6} & 24\sqrt{6} \\ 18 & 24 & 18 & -18\sqrt{6} & -138 & 60 \\ 36 & 48 & -36 & 0 & 60 & -48 \end{bmatrix}, \tag{6.15}$$

where the 2×2 submatrices surrounded by the dotted lines are $T^{(3)}$ and $T^{(4)}$ corresponding to $\lambda_3 = -2$ and $\lambda_4 = -5$, respectively. We orthogonalize the third and the fourth columns of T by bringing $T^{(3)}$ to an upper triangular form and likewise for the fifth and sixth of T and obtain the transformation matrix S of (5.4). By normalizing each column of S using (5.7), one obtains the unitary transformation matrix U,

$$U = \begin{bmatrix} \sqrt{\frac{1}{30}} & -\sqrt{\frac{27}{140}} & -\sqrt{\frac{1}{3}} & -\sqrt{\frac{1}{12}} & \frac{1}{2} & \sqrt{\frac{3}{28}} \\ \sqrt{\frac{3}{10}} & -\sqrt{\frac{12}{35}} & 0 & 0 & -\frac{1}{2} & -\sqrt{\frac{3}{28}} \\ \sqrt{\frac{3}{10}} & \sqrt{\frac{3}{140}} & \sqrt{\frac{1}{3}} & -\sqrt{\frac{1}{12}} & \frac{1}{2} & -\sqrt{\frac{1}{84}} \\ \sqrt{\frac{1}{5}} & \sqrt{\frac{1}{70}} & 0 & \sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{2}{7}} \\ \sqrt{\frac{1}{30}} & \sqrt{\frac{3}{35}} & 0 & -\sqrt{\frac{1}{3}} & -\frac{1}{2} & \sqrt{\frac{25}{84}} \\ \sqrt{\frac{2}{15}} & \sqrt{\frac{12}{35}} & -\sqrt{\frac{1}{3}} & 0 & 0 & -\sqrt{\frac{4}{21}} \end{bmatrix}. \tag{6.16}$$

From (6.14) and the conversion relation (6.12), the coupling coefficients for the eigenfunctions of $\mathbf{L}^2$ are given by the columns of U as follows:

$$U_{\cdot 1}\in L=5,\quad U_{\cdot 2}\in L=4,\quad (U_{\cdot 3},U_{\cdot 4})\in L=3,\quad (U_{\cdot 5},U_{\cdot 6})\in L=2. \tag{6.17}$$

The last two columns corresponding to D^2 are in complete agreement with Löwdin's result [3] based on the projection operator method. He did not calculate the rest, which requires further calculation of three more idempotent matrices.

C. The vector coupling coefficients

The vector coupling coefficients for two angular momenta or simply the Wigner coefficients [5] have completely been worked out by Wigner based on the theory of group representations. One of the reasons for choosing this well known problem is to incorporate the sign convention of the coupling coefficients due to Condon–Shortley [6] and Wigner [5] into the present formalism. For the sake of simplicity of notation, we may state that the problem is to construct the eigenfunctions of $\mathbf{J}^2$ from the direct product basis functions $|lsm_l m_s\rangle$ of the two angular momenta $\mathbf{L}$ and $\mathbf{S}$, with $l \geqslant s$ and a given $m = m_l + m_s$. Then, the order of the matrix $\|\mathbf{J}^2\|$ is $(2s+1)\times(2s+1)$ in a subspace of a given set of l, s and m. The present formalism can give a very general treatement for this problem. However, we shall discuss here very simple special cases of $s=\frac{1}{2}$ and 1. In actual calculation it is simplest to take $A=\|L\cdot S\|$. There is no degeneracy in this problem. Accordingly, we may take the proper sequence $\nu = g(i)$ in A such that $\nu = i$.

1. S = ½ case

The matrix $A=\|L\cdot S\|$ is of order 2×2 described by the basis of representation.

$$\phi_1 = |l,\tfrac{1}{2},m-\tfrac{1}{2},\tfrac{1}{2}\rangle,\quad \phi_2 = |l,\tfrac{1}{2},m+\tfrac{1}{2},-\tfrac{1}{2}\rangle. \tag{6.18}$$

In this special case of $r=2$, we can write down the T matrix without further calculation using (3.3) once the matrix elements of A are known. The eigenvalues of A are $\lambda_1 = \frac{1}{2}l$, $\lambda_2 = -\frac{1}{2}(l+1)$ corresponding to $j=l+\frac{1}{2}, j=l-\frac{1}{2}$ respectively. Taking

$$A = \operatorname{diag}\left[\tfrac{1}{2}l,\ -\tfrac{1}{2}(l+1)\right] \tag{6.19}$$

we have, from (3.3) and (5.10) with the positive square roots for the normalization constants,

$$Y=\begin{bmatrix}\left(\dfrac{(l+m+1/2)}{2l+1}\right)^{1/2} & \left(\dfrac{l-m+1/2}{2l+1}\right)^{1/2}\\ \left(\dfrac{l-m+1/2}{2l+1}\right)^{1/2} & -\left(\dfrac{l+m+1/2}{2l+1}\right)^{1/2}\end{bmatrix}, \tag{6.20}$$

where we have written Y instead of U following the notation used in the previous work I. This matrix Y has the IUH symmetry (involutional, unitary, and Hermitian) and is real and symmetric. Despite these symmetries, the columns of Y do not satisfy the sign convention of the Condon-Shortley and Wigner. In fact, the Wigner coefficients are given by the columns of $\mathbf{U}=\mathbf{Y}\sigma$ wl ere $\sigma = \operatorname{diag}[1,\ -1]$. Hereafter, in the calculation of the vector coupling coefficients we shall take the signs of the normalization constants in (5.10) such that all the row elements of U corresponding to the minimum m_s $(=-s)$ become positive. Then we can show that the matrix U given by (5.10) satisfies the sign convention of Condon–Shortley and Wigner [5] (see the next example).

2. S = 1 case

In terms of the basis of representation with a given set of $l, s=1, m_j=m$,

$$\phi_1 = |l,1,m-1,1\rangle,\quad \phi_2 = |l,1,m,0\rangle,$$
$$\phi_3 = |l,1,m+1,-1\rangle, \tag{6.21}$$

we can easily write down the matrix elements of $A=\|\mathbf{L}\cdot\mathbf{S}\|$, which satisfies the characteristic equation

$$A^3 + 2A^2 - (l^2+l-1)A - l(l+1)I = 0. \tag{6.22}$$

The relations between j, λ_ν and χ_ν are

$$\begin{array}{llll} & j & \lambda_\nu & \chi_\nu \\ \nu=1, & l+1 & l & (l+1)(2l+1), \\ \nu=2, & l & -1 & -l(l+1), \\ \nu=3, & l-1 & -(l+1) & l(2l+1). \end{array} \tag{6.23}$$

Assuming $\nu = i$, namely,

$$A = \operatorname{diag}[l,\ -1,\ -(l+1)], \tag{6.24}$$

we can calculate the matrix T from

$$T = A^2 + AA^{(1)} + A^{(2)}. \tag{6.25}$$

Then, from (5.10), we have

$$U=\begin{bmatrix}\left(\dfrac{(l+m)(l+m+1)}{2(l+1)(2l+1)}\right)^{1/2} & -\left(\dfrac{(l+m)(l-m+1)}{2l(l+1)}\right)^{1/2} & \left(\dfrac{(l-m)(l-m+1)}{2l(2l+1)}\right)^{1/2}\\ \left(\dfrac{(l-m+1)(l+m+1)}{(l+1)(2l+1)}\right)^{1/2} & \dfrac{m}{[l(l+1)]^{1/2}} & -\left(\dfrac{(l+m)(l-m)}{l(2l+1)}\right)^{1/2}\\ \left(\dfrac{(l-m)(l-m+1)}{2(l+1)(2l+1)}\right)^{1/2} & \left(\dfrac{(l-m)(l+m+1)}{2l(l+1)}\right)^{1/2} & \left(\dfrac{(l+m)(l+m+1)}{2l(2l+1)}\right)^{1/2}\end{bmatrix}, \tag{6.26}$$

where the signs of all the last row elements corresponding to the minimum m_s $(=-1)$ are made positive to satisfy the convention of Condon-Shortley. It is noted that the column vectors $U_{\cdot 1}$, $U_{\cdot 2}$ and $U_{\cdot 3}$ give the coupling coefficients corresponding to $j=l+1$, l, and $l-1$, respectively.

ACKNOWLEDGMENTS

The author wishes to thank Dr. B.K. Oh and Professor J.A. Poole at Temple University for reading the manuscript and giving helpful suggestions.

[1]Shoon K.Kim, J. Math. Phys., see preceding paper in this issue.
[2]V.N. Faddeeva, *Computational Methods of Linear Algebra* (Dover, New York, 1959), p. 177 and references therein.
[3]P.O. Löwdin, Rev. Mod. Phys. **36**, 966 (1964), and reference therein.
[4]D.E. Littlewood, *The Theory of Group Characters* (Oxford, Clarendon, 1950), second edition.
[5]E.P. Wigner, *Group Theory* (Academic, New York, 1959), p. 190.
[6]E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra* (Cambridge U.P., Cambridge, 1935), p. 76.

Representation groups for semiunitary projective representations of finite groups [a),b)]

José F. Cariñena

Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, Spain

Mariano Santander

Departamento de Física Teórica, Facultad de Ciencias, Universidad de Valladolid, Spain

(Received 29 May 1978)

Given a finite group G and a subgroup K of index 1 or 2, a method is developed for finding all the finite groups $\bar{G}$ (up to equivalence) such that any semiunitary projective representation of G can be lifted to a semiunitary representation of $\bar{G}$. The method is used in simple but interesting groups in physics.

1. INTRODUCTION

The aim of this paper is to examine in a systematic way a new method for obtaining all the semiunitary (unitary/antiunitary) projective representations of a finite group G from the semiunitary representations of an auxiliary finite group $\bar{G}$.[1]

It must be pointed out that the case of a connected Lie group has been considered in an earlier paper[2] and that of a nonconnected Lie group will be treated in a forthcoming paper.

This paper is motivated by the fact that the semiunitary projective representations arise in quantum physics as a consequence of Wigner's theorem[3] which implies that any symmetry group of a quantum system must be realized by means of a semiunitary projective representation in the space of the states of the system.

The method that will be developed here reduces the study of all the (semiunitary) projective representations of a finite group G, to the problem of the (semiunitary) representations of its auxiliary group $\bar{G}$ once that this last group is known. This method is simpler than the standard methods that are usually applied when one considers unitary representations (via, e.g., twisted algebra[4,5]).

It is worthwhile to point out that this method is different from the well known procedure of obtaining those semiunitary projective representations of a group G corresponding to a given (cohomology class of) factor system, from the semiunitary representations of the middle group of the associated extension of G by the circle group.[6] Notice that in our approach only one group is used while in Ref. 6 a family of auxiliary groups is employed.

The study of unitary projective representations of a finite group was done almost exhaustively by Schur[7] in a series of three papers. Schur showed that the unitary projective representations of G can be obtained from the unitary representations of a new finite group $\bar{G}$, called the representation group (darstellungsgruppe) for G. However this group $\bar{G}$ is not univocally determined by G.

Janssen[8] proved recently that a similar result holds for the more general case of semiunitary projective representations of a finite group G. Then the following question arises; Why reconsider an old problem already solved? The reason is twofold: Firstly, Janssen's method of construction of the representation group cannot be extended to topological groups while our method permits this generalization. Secondly, we give a new general method (without any reference to the defining relations) of finding all the representation groups, and further they are classified. Consequently our method allows us to choose in each case the simplest representation group: This results in a simplification of the calculations needed in some particular problems (see, e.g., the worked example in Sec. 9 for the Vierergruppe).

It is interesting to point out that this new simple method may be useful in several fields, like theoretical physics, solid state, and perhaps quantum chemistry. Taking into account that semiunitary projective representations of groups appear in fundamental problems in such fields whereas cohomology theory is not usually in the mathematical background of their practitioners, we confine ourselves to a relatively elementary level of this theory, which suffices for our purpose, and we have avoided the use of more advanced techniques—as, e.g., the spectral sequences—with whose help some of the results contained in this paper could be easily obtained.

In Sec. 2. we establish some basic definitions and notations, due to the fact that, unfortunately, the terms involved have different meanings for different authors. The most important properties concerning these definitions are included in the Appendix. In Sec. 3. the concept of splitting group is defined and characterized, and we classify the representation groups in Sec. 4. Section 5 deals with representation groups. In Sec. 6 we develop a new general method for obtaining all the representation groups for any finite group G; we discuss in Sec. 7 the characterization of the different classes of representation groups and we give a theorem concerning the unicity of (splitting classes of) representation

[a)]This work has been supported by Instituto de Estudios Nucleares.

[b)]Partly from the Ph.D. Thesis quoted in Ref. 1.

 0022-2488/79/112168-10$01.00

groups. In Sec. 8 the connection between semiunitary projective and linear representations, in particular the irreducible case, is considered. Finally, Sec. 9 applies the present theory to some examples of great interest in Q.M. We do not enter here in discussing the possible applications of semiunitary projective representations.[9]

2. NOTATIONS AND BASIC DEFINITIONS

In the following H will denote a separable complex Hilbert space, and $\overline{H}$ the projective space associated with H. The group of unitary or antiunitary operators in H (the so-called semiunitary group) will be denoted $\Gamma\,\mathrm{U}(H)$ and $\mathrm{U}(H)$ will be the unitary subgroup. By T we mean, as usual, the circle group $\mathrm{U}(1)$, identified with the center of $\mathrm{U}(H)$ (unitary scalar operators $e^{i\alpha}I$); the factor groups $\Gamma\,\mathrm{U}(H)/T$ and $\mathrm{U}(H)/T$ will be denoted $P\Gamma\,\mathrm{U}(H)$ and $P\,\mathrm{U}(H)$, respectively. We call π the canonical epimorphism $\pi:\Gamma\,\mathrm{U}(H)\rightarrow P\Gamma\,\mathrm{U}(H)$.

Let G be a finite group (noted multiplicatively) and K a subgroup of index 1 or 2 of G. An action of G on T denoted by $*K$, will be defined by

$$(*K)(g)=\begin{cases}\lambda\rightarrow\lambda, & g\in K,\\ \lambda\rightarrow\lambda^*, & g\in G-K,\end{cases}$$

where λ^* stands for the complex conjugate of $\lambda\in T$. If no risk of confusion arises, we use $\lambda\rightarrow\lambda^g$ as a shorthand notation for the action $*K$. The corresponding Eilenberg–MacLane (normalized) cochain complex as well as the corresponding subgroups and quotients [10] are denoted $C^{\bullet}_{*K}(G,T)$, etc. Whenever K is clear from the context, we shall write $*$ instead of of $*K$, and when $K=G$ we will also omit $*$.

Next, let us introduce some definitions that will be used in the following sections.

Definition 1: A semiunitary representation (hereafter SUR) of (G,K) is an homomorphism $D:G\rightarrow\Gamma\,\mathrm{U}(H)$ such that $D(K)\subset\mathrm{U}(H)$ and $D(G-K)\subset\Gamma\,\mathrm{U}(H)-\mathrm{U}(H)$.

Definition 2: A semiunitary projective representation (SUPR) of (G,K) is an homomorphism $P:G\rightarrow P\Gamma\,\mathrm{U}(H)$ such that $P(K)\subset P\,\mathrm{U}(H)$ and $P(G-K)\subset P\Gamma\,\mathrm{U}(H)-$ $-P\,\mathrm{U}(H)$.

Definition 3: A semiunitary multiplier representation (SUMR) of (G,K) is a map $R:G\rightarrow\Gamma\,\mathrm{U}(H)$ such that $\pi\circ R$ is a SUPR of (G,K) and $R(1)=1$.

Definition 4: A SUMR, R of (G,K) is said to be a (multiplier) lifting of the SUPR P of (G,K) if $\pi\circ R=P$. In this case one also says that P is the SUPR of (G,K) associated with R.

Definition 5: Let $\overline{G}$ be a finite group and $p:\overline{G}\rightarrow G$ an epimorphism. If P is a SUPR of (G,K), a SUR, R of $(\overline{G},\overline{K})$ [with $\overline{K}=p^{-1}(K)$] is said to be a (linear) lifting of P [with respect to the "covering" $(\overline{G},p)$] if $P\circ p=\pi\circ R$.

Other basic concepts, as equivalence, irreducibility, etc., are included in the Appendix.

3. THE SPLITTING GROUPS

Almost everyone knows that, in general, not every SUPR of a finite group (G,K) admits a (linear) lifting to a SUR of (G,K) (with the identity epimorphism). Here we consider the possibility of finding a group $\overline{G}$ (and an epimorphism $p:\overline{G}\rightarrow G$) such that *any* SUPR of (G,K) can be lifted to a SUR of $(\overline{G},\overline{K})$.

Definition 6: Let $\overline{G}$ be a finite group and $p:\overline{G}\rightarrow G$ be an epimorphism. Then $(\overline{G},p)$ is said to be a splitting group for (G,K) if any SUPR of (G,K) can be lifted to a SUR of $(\overline{G},\overline{K})(\overline{K}=p^{-1}(K))$.

The term "splitting group"has been used by Moore[11] in another slightly different context, and also by Cattaneo.[12]

From now on, and when K is clear from the context, a SUPR of G will be one of (G,K), and so on.

Next we show that to every "general" splitting group (if it exists) we can associate another splitting group of a more restricted kind, namely extensions of G by Abelian finite kernels A with a definite action $-K$ of G on A given by

$$(-K)(g)=\begin{cases}a\rightarrow a, & g\in K,\\ a\rightarrow a^{-1}, & g\in G-K.\end{cases}$$

This task is carried out in two steps. First let us show there is no real loss of generality (in the above sense) if we demand the kernel of the epimorphism $p:\overline{G}\rightarrow G$ to be Abelian. Let G_1 be a "general" splitting group with respect to $p_1:G_1\rightarrow G$. If D is a SUR of G_1 which lifts some SUPR P of G, it is clear that $D(\ker p_1)\subset T$, and therefore $\chi_D\equiv D|_{\ker p_1}$ is a unidimensional UR of $\ker p_1$. Because of the Abelian nature of T we have $\chi_D((\ker p_1)')=1$. Then $(\ker p_1)'$ is contained in the kernel of any SUR of G_1 which lifts any SUPR of G. But $(\ker p_1)'$ being a characteristic subgroup of the normal subgroup $\ker p_1$, $(\ker p_1)'$ is also normal in G_1 and we can define $G_2=G_1/(\ker p_1)'$ and $p_2:G_2\rightarrow G$ as $p_2[g_1(\ker p_1)']=p_1(g_1)$. On the other hand, D gives rise naturally to a SUR of G_2 which is also a lifting of P. The new kernel, $\ker p_2$, is Abelian because it is the Abelianized of $\ker p_1$. Then G_2 is a splitting group for G with an Abelian kernel $\ker p_2$.

Henceforth we restrict our search for splitting groups to extensions of G by Abelian kernels, A, $1\rightarrow A\rightarrow\overline{G}\rightarrow G\rightarrow 1$. Any (equivalence of class of) extension of G by A is characterized by an action Θ of G on A, $\Theta:G\rightarrow\mathrm{Aut}\,A$ and a (cohomology class $\bar{w}$ of) Θ-factor system $w:G\times G\rightarrow A$ If $w\in\bar{w}$, the group G is isomorphic (equivalent as extension) to the group called $G_{\bar{w}}$ and obtained giving in the set of pairs, $(a,g)\in A\times G$ the following composition law,

$$(a,g)(b,h)=(ab^{\Theta(g)}w(g,h),gh).$$

The epimorphism $p:\overline{G}\rightarrow G$ is defined in terms of $G_{\bar{w}}$ by $p(a,g)=g$. The explicit reference to this p will be omitted in the following. We turn now to the second reduction step, in which we show that one can choose, without any real loss of generality the action $-K$ instead of the most general Θ. Let G_1 be a splitting group for G, with Abelian kernel A_1 and "general" action $\Theta:G\rightarrow\mathrm{Aut}\,A_1$. If D is a SUR of G_1 which lifts some SUPR P of G, and $\chi_D\equiv D|_{\ker A_1}\in\hat{A}_1$, the relation $(1,g)(a,1)=(a^{\Theta(g)},g)$ which holds in $G_{\bar{w}}$, implies $[\chi_D(a)]^g=\chi_D(a^{\Theta(g)})$, or $\chi_D((a^{-1})^{(-K)(g)}a^{\Theta(g)})=1$, that is, $(a^{-1})^{(-K)(g)}a^{\Theta(g)}\in\ker\chi_D$, these relations being accomplished for all $\chi_D\in\hat{A}_1$ arising from those SUR's of $G_{\bar{w}}$ which are lifts of the SUPR's of G. Let S be the subgroup of $\hat{A}_1$ whose elements are all such χ_D, and define $B=\cap_{\chi\in S}\ker\chi$. It is clear that B is a normal subgroup of A_1, and again it is also normal in G_1, so that defining $G_2=G_1/B$ and $p_2(g_1B)=p_1(g_1)$ we have a new splitting group G_2 because B is contained in the kernel of every "relevant"SUR of G_1. Furthermore, the induced action of G on the new kernel $A_2=A_1/B$ is just, as one easily sees, the action $-K$ of G on A_2.

Then we have obtained:

Proposition 1: In the search of splitting groups for (G,K) we can restrict our study to the ("right-hand part" of) extensions of G by Abelian kernels A with action $-K$ of G on A.

Now we fix our attention on the conditions which must satisfy A and $\bar{w}$ in order for $G_{\bar{w}}$ to be a splitting group. With this purpose let us introduce some auxiliary homomorphisms. Every $\chi\in\hat{A}$ defines a homomorphism $\dot{\chi}:Z^2_-(G,A)\to Z^2_*(G,T)$, $\dot{\chi}(w)=\chi\circ w$, compatible with the equivalence modulo the respective B^2 and therefore induces another homomorphism $\dot{\bar{\chi}}:H^2_-(G,A)\to H^2_*(G,T)$, $\dot{\bar{\chi}}(\bar{w})=\dot{\chi}(w)$ with any $w\in\bar{w}$. Similarly, every $w\in Z^2_-(G,A)$ defines a homomorphism $\dot{w}:\hat{A}\to Z^2_*(G,T)$ by $\dot{w}(\chi)=\dot{\chi}(w)$. Moreover, if w_1 and w_2 are cohomologous, so are $\dot{w}_1(\chi)$ and $\dot{w}_2(\chi)$ for any $\chi\in\hat{A}$, and one can define an homomorphism $\dot{\bar{w}}:\hat{A}\to H^2_*(G,T)$ by $\dot{\bar{w}}(\chi)=\dot{\bar{\chi}}(\bar{w})$ In terms of such homomorphisms, we have the following results:

Proposition 2: A SUPR P of (G,K) can be lifted to a SUR of $(G_{\bar{w}}, K_{\bar{w}})$ iff there exists a $\chi\in\hat{A}$ such that $\dot{\bar{\chi}}(\bar{w})$ is the cohomology class of P.

Proof: If D is a SUR of $G_{\bar{w}}$ lifting P, the cohomology class of P can be easily found to be $\dot{\bar{\chi}}_D(\bar{w})$. [Here $\chi_D\in\hat{A}$ is determined as before by $D|_A(a)=\chi_D(a)\mathbb{1}$.] Conversely, let us assume that P is a SUPR of (G,K) with cohomology class $\bar{\zeta}$. If there is a $\chi\in\hat{A}$ such that $\bar{\zeta}=\dot{\bar{\chi}}(\bar{w})$, take a section $r:G\to G_{\bar{w}}$ and let $w_r\in\bar{w}$ be the corresponding lifting of $\bar{w}$, and define $\zeta_r=\dot{\chi}(w_r)$. Now ζ_r lies in $\bar{\zeta}$ and Proposition 1 (Appendix) shows that there exists a multiplier lifting R_{ζ_r} of P, with factor system $\zeta_r\in\bar{\zeta}$. The application $D:G_{\bar{w}}\to\Gamma\,\mathrm{U}(H)$ defined by $R(a,g)_r=\chi(a)R_{\zeta_r}(g)$ is a SUR of $G_{\bar{w}}$ which lifts P.

Theorem 1: If $\bar{\zeta}\in H^2_*(G,T)$, there is at least a SUPR of (G,K) with $\bar{\zeta}$ as its cohomology class.

Proof: It can be done by a straightforward generalization of the Bargmann's argument [13]. Take an arbitrary $\zeta\in\bar{\zeta}$, and, on the vector space of all complex functions on G, define $[R_\zeta(g)F](h)=\zeta(g,g^{-1}h)F^g(g^{-1}h)$. Then R_ζ is a SUMR of G and $\pi\circ R_\zeta=P$ is a SUPR of G with cohomology class $\bar{\zeta}$. In the case $K=G$ this theorem is a particular case of a result of Mackey.[14]

From Proposition 2 and Theorem 1 it follows:

Theorem 2: $G_{\bar{w}}$ is a splitting group iff $\dot{\bar{w}}$ is an epimorphism.

Proof: Every element of $H^2_*(G,T)$ must be obtained as $\dot{\bar{\chi}}(\bar{w})\equiv\dot{\bar{w}}(\chi)$ with some $\chi\in\hat{A}$.

As $\dot{\bar{w}}$ is an epimorphism, we have that Card$(\hat{A})$ (the order of $\hat{A}$) is some multiple of Card $H^2_*(G,T)$ and consequently, the order of any splitting group $G_{\bar{w}}$ is some multiple of Card $H^2_*(G,T)\times$ Card G. The lowest order for $G_{\bar{w}}$ is Card $H^2_{*G}(G,T)\times$ Card G.

Definition 7: The group $G_{\bar{w}}$ is said to be a representation group for (G,K) if it is a splitting group of minimal order.

For representation groups we have

Theorem 3: $G_{\bar{w}}$ is a representation group iff $\dot{\bar{w}}$ is an isomorphism.

Proof: The order of $G_{\bar{w}}$ must be Card $H^2_{*H}(G,T)\times$ Card G and therefore $\hat{A}$ must have the same order as $H^2_*(G,T)$. This, and the fact that $\dot{\bar{w}}$ must be an epimorphism implies that $\dot{\bar{w}}$ is an isomorphism. The converse is trivial.

A useful characterization of the kernel of $\dot{\bar{w}}$ is furnished by the next proposition.

Proposition 3: The kernel of $\dot{\bar{w}}$ consists of those elements of $\hat{A}$ which are restrictions to A of the crossed homomorphisms $\gamma:G_{\bar{w}}\to T$.

Proof: Let r be a section $r:G\to G_{\bar{w}}$ and w_r its factor system. If $\chi\in\ker\dot{\bar{w}}$, as $\dot{\chi}(w_r)\in B^2_*(G,T)$, there exists $\Lambda\in C^1_*(G,T)$ such that $\chi\circ w_r=\delta\Lambda$. Define $\gamma:G_{\bar{w}}\to T$ by $\gamma(a,g)_r=\chi(a)\Lambda(g)$. It is clear that $\gamma\in Z^1_*(G_{\bar{w}},T)$, and $\gamma|_A\equiv\chi$. Conversely, if $\chi\in\hat{A}$ is the restriction to A of some $\gamma\in Z^1_*(G,T)$, $\chi\equiv\gamma|_A$, then the relations

$$\gamma[(a,g)_r(b,h)_r]=\gamma(a,g)_r\gamma^g(b,h)_r,$$

$$\gamma(a,g)_r=\chi(a)\gamma(1,g)_r,$$

hold and with a simple computation we obtain

$$\chi(w_r(g,h))=\frac{\gamma(1,g)_r\gamma^g(1,h)_r}{\gamma(1,gh)_r},$$

i.e., $\dot{\bar{w}}(\chi)=1$.

4. EQUIVALENCE OF SPLITTING GROUPS

We begin this section with a definition about the equivalence of splitting groups. According to our Definition 6, a natural concept of equivalence ought to refer only to the epimorphisms $p:\bar{G}\to G$.

Definition 8: Two splitting groups for G, $(\bar{G},p)$, and $(\bar{G}',p')$ will be called equivalent if there exists an isomorphism $R:\bar{G}\to\bar{G}'$ such that $p=p'\circ R$

This relation is obviously an equivalence relation.

As we have seen in Propostion 1, splitting groups arise essentially only as the ("right-hand part" of) extensions of G by Abelian groups A. Thus, if one knows $\bar{G}$ as some extension of G by A (which demands to know $p:\bar{G}\to G$ and $i:A\to\bar{G}$) one can also consider the equivalence of extensions (in the usual sense) as a criterium of equivalence of splitting groups. Of course there is a close relation between these two equivalences which we are going to analyze, in order to show that, from our point of view, this difference is, in some sense, irrelevant.

In fact it is well known [15] that if $(\bar{G},p)$ is a splitting group for G with a kernel isomorphic to some A, and we denote by i the canonical injection i:Ker$p \to \overline{G}$, any isomorphism α:Ker$p \to A$ defines a new extension (unique up to equivalence) $1 \to A \to \overline{G}' \to G \to 1$ and an isomorphism R (because of the short five lemma) such that the diagram

$$\begin{tikzcd} 1 \arrow[r] & \mathrm{Ker}\,p \arrow[r, "i"] \arrow[d, "\alpha"'] & \overline{G} \arrow[r, "p"] \arrow[d, "R"] & G \arrow[r] & 1 \\ 1 \arrow[r] & A \arrow[r, "i"] & \overline{G}' \arrow[r, "p'"] & G \arrow[r] & 1 \end{tikzcd}$$

is commutative. It is clear that $(\overline{G}',p')$ is also a splitting group equivalent (cf. Definition 8) to $(\overline{G},p)$. Conversely, if we have two equivalent splitting groups $(\overline{G},p)$ and $(\overline{G}',p')$, let i,i' be the canonical injections i:Ker$p \to \overline{G}$, i':Ker$p' \to \overline{G}'$; it is clear that the restriction of the R that realizes the equivalence to Ker$p$ is an isomorphism σ:Ker$p \to$Kerp'. Now if one identifies Ker p and Kerp' with A, through the isomorphisms $\alpha\circ\sigma$ and α, respectively (α arbitrary), and consider the extensions of G by A obtained from $1 \to \mathrm{Ker}p \xrightarrow{i} \overline{G} \xrightarrow{p} G \to 1$ through $\alpha\circ\sigma$ and from $1 \to \mathrm{Ker}p' \xrightarrow{i'} \overline{G}' \xrightarrow{p'} G \to 1$, through α, it is evident that such extensions of G by A are equivalent in the usual sense.

Now let $\phi\in$Aut A, and use the obvious generalization of the same notations previously employed, $\dot{\phi}$ for the homomorphism $Z^2_-(G,A) \to Z^2_-(G,A)$, $\dot{\phi}(w) = \phi\circ w$, and $\ddot{\phi}$ for the homomorphism $H^2_-(G,A) \to H^2_-(G,A)$, $\ddot{\phi}(\bar{w}) = \overline{\dot{\phi}(w)}$ with any $w\in\bar{w}$. It is easy to show:

Proposition 4: Let (G,p) and (G',p') be two splitting groups, which are extensions of G by A with action $-K$, and let $\bar{w}$ and $\bar{w}'$ be the corresponding cohomology classes. Then $\overline{G}$ and $\overline{G}'$ are equivalent iff there exists $\phi\in$Aut$A$ such that $\bar{w}' = \ddot{\phi}(\bar{w})$.

The extension equivalence classes of splitting groups for (G,K) with kernel A and action $-K$ are in one-to-one correspondence with the subset of $H^2_-(G,A)$ whose elements $\bar{w}$ verify that $\dot{\bar{w}}$ is an isomorphism. On the other hand, the equivalence classes (cf. Definition 8) are in one to one correspondence with the orbits of Aut$A$ [with action $\phi \to (\bar{w} \to \ddot{\phi}(\bar{w}))$] in such a subset of $H^2_-(G,A)$. It is a trivial task to establish the corresponding results for isomorphism of splitting groups.

5. REPRESENTATION GROUPS

In the folllowing we shall restrict ourselves to the representation groups. If some extension of G by A has a middle group $G_{\bar{w}}$ which is a representation group, one derives easily from Theorem 3 that $\hat{A}$ must be isomorphic to $H^2_{\cdot}(G,T)$, and, of course, A itself must be isomorphic to $[H^2_{\cdot}(G,T)]^\wedge$. Then, for finite groups, the kernel of any extension of G leading to a representation group is essentially unique ($\approx [H^2_{\cdot}(G,T)]^\wedge$). Because of the autoduality of finite Abelian groups, one could also take $A \approx H^2_{\cdot}(G,T)$ as Janssen does, but as we are looking for a procedure that can be generalized to Lie groups [in which case $H^2_{\cdot}(G,T)$ is no longer finite,], then we shall always choose $A \approx [H^2_{\cdot}(G,T)]^\wedge$. As a direct consequence of Theorem 3 and Proposition 3 we have:

Proposition 5: Let $G_{\bar{w}}$ be the middle group of the extension of G by $[H^2_{\cdot}(G,T)^\wedge]$ that is labelled by $\bar{w}\in H^2_-(G,[H^2_{\cdot}(G,T)]^\wedge)$. Then $G_{\bar{w}}$ is a representation group [for (G,K)] iff the restriction to $[H^2_{\cdot}(G,T)]^\wedge$ of every crossed homomorphism $\gamma\in Z^1_{\cdot}(G_{\bar{w}},T)$ is trivial.

In order to translate this condition into a more easy form we draw our attention to the groups $Z^1_{\cdot}(G,T)$ and $H^1_{\cdot}(G,T)$. Let $\bar{w}$ be an arbitrary element of $H^2_-(G,A)$ (with an arbitrary A), and let $1 \to A \to G_{\bar{w}} \to G \to 1$ be the corresponding extension. The surjective homomorphism p:$G_{\bar{w}} \to G$ induces an injective homorphism ψ:$Z^1_{\cdot}(G,T) \to Z^1_{\cdot}(G_{\bar{w}},T)$ defined by $\psi(\delta) = \delta\circ p$.

The elements belonging to the image of ψ are trivial on A. Moreover $\psi(B^1_{\cdot}(G,T)) \subset B^1_{\cdot}(G_{\bar{w}},T)$ and ψ induces a new injective homomorphism, $\bar{\psi}$:$H^1_{\cdot}(G,T) \to H^1_{\cdot}(G_{\bar{w}},T)$.

If we take $A = [H^2_{\cdot}(G,T)]^\wedge$, Proposition 5 show us that $G_{\bar{w}}$ is a representation group iff ψ is also surjective, because then every $\gamma\in Z^1_{\cdot}(G_{\bar{w}},T)$ is trivial on A. In such a case ψ and $\bar{\psi}$ are isomorphisms. We have obtained:

Theorem 4: $G_{\bar{w}}$ is a representation group iff the canonical mapping $\bar{\psi}$ is an isomorphism.

The elements of $Z^1_{\cdot}(G,T)$ and $H^1_{\cdot}(G,T)$ are closely related to the semiunitary unidimensional representations of (G,K). In fact, let $\gamma\in Z^1_{\cdot}(G,T)$ and $\mathbb{K}$ be some (fixed) antiunitary operator in a unidimensional Hilbert space H, and define an application R_γ:$G \to \Gamma\,\mathrm{U}(H)$ by

$$R_\gamma: \begin{cases} g \to \gamma(g), & g\in K, \\ g \to \gamma(g)\mathbb{K}, & g\in G-K. \end{cases}$$

Then, it is clear that R_γ is a unidimensional SUR of (G,K). Conversely we can associate some $\gamma\in Z^1_{\cdot}(G,T)$ to every unidimensional SUR of (G,K), by "removing" $\mathbb{K}$.

The study of $Z^1_{\cdot}(G,T)$ is therefore equivalent to the one of unidimensional SUR's of (G,K), which can be done by means of the traditional method of induced representations (with antiunitary operators[16]). Take some n-dimensional UR of K,Δ, and induce from Δ a SUR of (G,K); we obtain a SUR of (G,K) whose dimension is n, $2n$, $2n$, depending on the "Wigner type" of Δ being I, II, III. Then, unidimensional SUR's of (G,K) are just induced by type I unidimensional UR's of K, which form a subgroup of $\hat{K}$ (unidimensional UR's of K) and will be denoted $\hat{K}^{(\mathrm{I})}$. Then we have,[16]

$$\hat{K}^{(\mathrm{I})} = \{\sigma\in\hat{K} \mid \sigma(h) = \sigma^*(g_0^{-1}hg_0) \quad \text{and} \quad \sigma(g_0^2) = 1, \ \forall\ h\in K,\ g_0\in G-K\}.$$

There is a useful test, due to Dimmock[17] which gives the Wigner type of any UR of K in terms of his character. In our case, the UR's of K and their characters coincide, and we have

$$\hat{K}^{(\mathrm{I})} = \Big\{\sigma\in\hat{K} \mid \sum_{h\in K}\sigma((g_0h)^2) = \mathrm{Card}\,K\Big\}.$$

Proposition 6: There is a (noncanonical) isomorphism $\widehat{K}^{(1)} \otimes T \to Z^{1}_{*}(G,T)$.

Proof: Select an arbitrary but fixed $g_0 \in G - K$ and define $\varphi_{g_0}: \widehat{K}^{(1)} \otimes T \to Z^{1}_{*}(G,T)$ as follows: If $(\sigma,\xi) \in \widehat{K}^{(1)} \otimes T$, then

$$[\varphi_{g_0}(\sigma,\xi)](g) = \begin{cases} \sigma(g), & g \in K, \\ \xi\sigma^{*}(g_0^{-1}g), & g \in G - K \end{cases}.$$

The remaining part is trivial.

The image of T by φ_{g_0} is just $B^{1}_{*}(G,T)$ and hence we can state:

Proposition 7: There is a (canonical) isomorphism $\widehat{K}^{(1)} \to H^{1}_{*}(G,T)$.

If $K = G$, the crossed homomorphisms are homomorphisms and (as there are not "Wigner types") Proposition 6 reduces to the well-known statement that $H^{1}(G,T)$ is the group $\mathrm{Hom}\,(G,T) = \widehat{G}$.

By making use of these results, we obtain from Theorem 4:

Corollary 1: $G_{\bar{w}}$ is a representation group [for (G,K)] iff $\widehat{K}^{(1)}_{\bar{w}}$ and $\widehat{K}^{(1)}$ are isomorphic.

The explicit form of Dimmock's test can be modified for $\widehat{K}^{(1)}_{\bar{w}}$ in the following way: If $w \in \bar{w}$, and $g_0 \in G - K$, we have $\widehat{K}^{(1)}_{\bar{w}} = \{\sigma \in \widehat{K}_{\bar{w}} \mid \Sigma_{h \in K}\, \sigma(w(g_0h, g_0h), (g_0h)^2) = \mathrm{Card}\, K\}$.

In the particular case $K = G$, from Corollary 1 we have: $G_{\bar{w}}$ is a representation group iff $\widehat{G}_{\bar{w}}$ and $\widehat{G}$ are isomorphic; this condition is equivalent to saying that $A = [H^{2}_{*}(G,T)]\hat{}$ is contained in the derived group $(G_{\bar{w}})'$ of $G_{\bar{w}}$. Hence we meet Schur's well-known result[7,5]:

Corollary 2: The middle group $G_{\bar{w}}$ of some central extension of G by (a group isomorphic to) $[H^{2}(G,T)]\hat{}$ is a representation group for (G,G) iff $[H^{2}(G,T)]\hat{} \subset (G_{\bar{w}})'$.

An alternative, more general view of the results of Secs. 3 and 5 is provided by considering the inflation-restriction sequence for $1 \to A \to G_{\bar{w}} \to G \to 1$. In fact, Propositions 3 and 5 and Theorem 4 are trivial consequences of the exactness of that sequence,[12]

$$0 \to H^{1}_{*}(G,T) \xrightarrow{\bar{\Psi}} H^{1}_{*}(G_{\bar{w}},T) \to \hat{A} \xrightarrow{\dot{\bar{w}}} H^{2}_{*}(G,T) \to H^{2}_{*}(G_{\bar{w}},T).$$

Our maps $\dot{\bar{w}}$ and $\bar{\Psi}$ are, respectively, the corresponding transgression and inflation maps. In this context, Theorems 2 and 3 can also be formulated as follows: $G_{\bar{w}}$ is a splitting group iff the inflation map $H^{2}_{*}(G,T) \to H^{2}_{*}(G_{\bar{w}},T)$ is trivial and $G_{\bar{w}}$ is a representation group if this last map is trivial and the inflation map $\bar{\Psi}$ is bijective; both theorems are but particular cases of results by Cattaneo (comments preceding Remark 1 of Ref. 12). It is also clear that a sufficient condition for $G_{\bar{w}}$ being a splitting group is the triviality of $H^{2}_{*}(G_{\bar{w}},T)$.

6. A METHOD FOR FINDING THE REPRESENTATION GROUPS

In this section we shall give an explicit method of construction of representation groups. The shorthands H^{2}_{*}, Z^{2}_{*}, B^{2}_{*} will be used for $H^{2}_{*}(G,T)$, etc.

Let $G_{\bar{w}}$ be a representation group and $w \in \bar{w}$ a lifting of $\bar{w}$. In the following diagram

$$\begin{tikzcd}
Z^{2}_{-K}(G,[H^{2}_{*}(G,T)]\hat{}) \arrow[rr, "\pi"] \arrow[dd, "\dot{\chi}"'] & & H^{2}_{-K}(G,[H^{2}_{*}(G,T)]\hat{}) \arrow[dd, "\dot{\bar{\chi}}"] \\
& {[H^{2}_{*K}(G,T)]\hat{}} \arrow[dl, "\dot{w}"'] \arrow[dr, "\dot{\bar{w}}"] & \\
Z^{2}_{*K}(G,T) \arrow[rr, "\pi"] & & H^{2}_{*K}(G,T)
\end{tikzcd}$$

the map $\dot{\bar{w}}$ is an isomorphism, and as $\dot{\bar{w}} = \pi\dot{w}$, $\dot{w}$ must be an injective homomorphism. The isormorphism $\dot{\bar{w}}: H^{2}_{*}\hat{\hat{}} \to H^{2}_{*}$ can be factorized in only one way, $\dot{\bar{w}} = \theta^{-1} \circ \hat{\hat{}}^{-1}$, with $\theta \in \mathrm{Aut}\, H^{2}_{*}$ and $\hat{\hat{}}$ the canonical isomorphism $H^{2}_{*} \to (H^{2}_{*})\hat{\hat{}}$; $(\hat{\hat{}}\bar{\xi})(\chi) = \chi(\bar{\xi})$ for any $\chi \in (H^{2}_{*})\hat{}$ and $\bar{\xi} \in H^{2}_{*}$. Let $s: H^{2}_{*} \to Z^{2}_{*}$ be defined by $s = \dot{w} \circ \dot{\bar{w}}^{-1}$; it is clear that s is a section which is also an (injective) homomorphism. In some sense, we can consider the isomorphism $\dot{\bar{w}}$ as labelled by θ, and $\dot{w}$ by θ and s. If $\bar{\xi} \in H^{2}_{*}$, the image $s(\bar{\xi})$ is given by $s(\bar{\sigma}) = \{(\hat{\hat{}}[\theta(\bar{\xi})])\}\ (w)$, or, equivalently, by $(s(\bar{\xi}))(g,h) = w(g,h)\,(\theta(\bar{\xi}))$. Then, in terms of (s,θ), w can be written

$$w(g,h): \bar{\xi} \to (s[\theta^{-1}(\bar{\xi})])(g,h).$$

Conversely, if s is a homomorphic section $s: H^{2}_{*} \to Z^{2}_{*}$ and $\theta \in \mathrm{Aut}\, \mathrm{H}^{2}_{*}$, the map $w_{\theta s}: G \times G \to (H^{2})\hat{}$ defined by

$$w_{\theta s}(g,h): \bar{\xi} \to (s[\theta^{-1}(\bar{\xi})])(g,h)$$

is a cocycle, $w_{\theta s} \in Z^{2}_{-}(G, [H^{2}_{*}(G,T)]\hat{})$ and $(\hat{\hat{}}[\theta(\bar{\xi})])\bar{w}_{\theta s} = s(\bar{\xi})$, that is, $\dot{w}^{\circ}_{\theta s}\hat{\hat{}} \circ \theta = s$ and therefore, $\dot{w}_{\theta s}$ is an injective homomorphism. Furthermore, the $\dot{\bar{w}}_{\theta s}$ is easily computed to be $\dot{\bar{w}}_{\theta s} = \theta^{-1} \cdot \hat{\hat{}}^{-1}$ and, as it is an isomorphism, the corresponding $G_{\bar{w}_{\theta s}}$ is a representation group. Then:

Theorem 5: Let s be a section $s: H^{2}_{*}(G,T) \to Z^{2}_{*}(G,T)$ which is also a homomorphism and $\theta \in \mathrm{Aut}\, H^{2}_{*}(G,T)$. If $w_{\theta s}$ is the mapping $w_{\theta s}: G \times G \to [H^{2}_{*}(G,T)]\hat{}$ defined as

$$w_{\theta s}(g,h): \bar{\xi} \to (s[\theta^{-1}(\bar{\xi})])(g,h),$$

then $w_{\theta s} \in Z^{2}_{-K}(G, [H^{2}_{*}(G,T)]\hat{})$ and $G_{\bar{w}_{\theta s}}$ is a representation group. Conversely, if $G_{\bar{w}}$ is a representation group for (G,K), there is a homomorphic section $s: H^{2}_{*} \to Z^{2}_{*}$ and an automorphism $\theta \in \mathrm{Aut}\, H^{2}_{*}(G,T)$ such that $\bar{w}_{\theta s} = \bar{w}$.

By making use of this Theorem one can get the following result about the existence of representation groups:

Corollary (Theorem 4 of Ref. 8): There is (at least) a representation group for (G,K).

Proof. The Abelian group $B^{2}_{*}(G,T)$ is divisible, and therefore an injective Z module. Any Abelian extension with injective kernel is inessential[15] and hence there is a section $H^{2}_{*} \to Z^{2}_{*}$ which is a homomorphism.

7. EQUIVALENCE AND UNICITY OF REPRESENTATION GROUPS

From Theorem 4, one sees that the subset of $Z^2_-(G,[H^2_*(G,T)]^\wedge)$ whose elements w verify that $\tilde{w}$ is an isomorphism can be "parametrized" by the set $S\times\Omega$ where S is the set of all homomorphic sections $H^2_*\to Z^2_*$ and Ω the set Aut H^2_*. Furthermore, this "parameter space" is very useful in the actual construction of representation groups. Then we are naturally lead to the question of studying, in terms of $S\times\Omega$, equivalence of the representation groups. If two pairs (s,θ) and (s',θ') have associated (through Theorem 4) two equivalent representation groups, we shall say that (s,θ) and (s',θ') are equivalent.

At the first place, and before discussing such a problem, we remark that S can be put in one-to-one correspondence with Hom $(H^2_*(G,T), B^2_*(G,T))$ in the following way: If we pick out an arbitrary but fixed homomorphic section $s_0:H^2_*\to Z^2_*$, every homomorphic section s can be "factorized" as a product of s_0 by some $\lambda\in$Hom (H^2_*, B^2_*), $s(\bar\xi)=s_0(\bar\xi)\,\lambda(\bar\xi)$, or $s=s_0\lambda$ and conversely; then, for the chosen section s_0, the correspondence $S\to$ Hom (H^2_*, B^2_*) is $s_0\lambda\to\lambda$.

Let us discuss firstly when (s,θ) and (s',θ') lead to two representation groups which are equivalent as extensions of G by $(H^2_*)^\wedge$. Because of the identity $\tilde{w}_{\theta s}={}^\circ\theta^{-1}\hat{}^{-1}$, an obviously necessary condition is $\theta=\theta'$. Suppose now that (s,θ) and (s',θ) are extension equivalent; then $w_{s\theta}\,[w_{s'\theta}]^{-1}$ is trivial [as an element of $Z^2_-(G,H^2_*)$]. Calling $\Xi_{s,s',\theta}\equiv w_{s\theta}\,[w_{s'\theta}]^{-1}$, we have

$$\Xi_{s,s',\theta}(g,h):\bar\xi\to\{s\,[\theta^{-1}(\bar\xi)]\,(g,h)\}\,\{s'\,[\theta^{-1}(\bar\xi)]\,(g,h)\}^{-1}.$$

A slight play shows that the value of $\Xi_{s,s',\theta}(g,h)$ in $\bar\xi$ is $\{s\,[\theta^{-1}(\bar\xi)]\cdot s'\,[\theta^{-1}(\bar\xi^{-1})]\}\,(g,h)$. If we put $s=s_0\sigma$, $s'=s_0\sigma'$ with $\sigma,\sigma'\in$Hom $(H^2_*B^2_*)$, $\lambda=\sigma(\sigma')^{-1}$, and $\Xi_{\lambda,\theta}=\Xi_{s,s',\theta}$ we obtain that $(s_0\sigma,\theta)$ and $(s_0\sigma\lambda,\theta)$ are extension equivalent iff $\Xi_{\lambda,\theta}$, defined by

$$\Xi_{\lambda,\theta}(g,h):\bar\xi\to(\lambda\,[\theta^{-1}(\bar\xi)])\,(g,h)$$

is trivial [in $Z^2_-(G,(H^2_*)^\wedge)$]

The triviality of $\Xi_{\lambda,\theta}$ depends only on λ; in fact, if for some λ, $\Xi_{\lambda,1}$ is trivial, there is a map $\Lambda_\lambda:G\to(H^2_*)^\wedge$, such that $\Xi_{\lambda,1}=\delta\Lambda_\lambda$. In this case, $\Xi_{\lambda,\theta}=\delta(\hat\theta^{-1}\Lambda_\lambda)$ (where $\hat\theta^{-1}$ is the adjoint automorphism of θ^{-1}). On the other hand, it is clear that the subset of Hom(H^2_*, B^2_*) whose elements λ have a trivial $\Xi_{\lambda\theta}$ is a subgroup of Hom (H^2_*, B^2_*), which will be denoted Hom$_t(H^2_*, B^2_*)$. Then, (s,θ) and (s',θ') lead to representation groups equivalent as extensions iff $\theta=\theta'$ and $s'=s\lambda$ for some $\lambda\in$Hom$_t(H^2_*, B^2_*)$. According to Proposition 4, the representation groups obtained from (s,θ) and (s',θ') are equivalent (in the sense of Definition 8) iff $W_{s'\theta'}$ and $\hat\phi\,(W_{s\theta})$ are cohomologous for some $\phi\in$Aut $(H^2_*)^\wedge$. First we show that, for any $\theta\in$Aut H^2_* and $\phi\in$Aut$(H^2_*)^\wedge$ we have $\hat\phi\,(w_{s,\theta})=w_{s,\hat\phi^{-1}\theta}$, where $\hat\phi$ is the adjoint map of ϕ, H^2_* and $H^{2\,\hat{}\hat{}}_*$ being identified as usual. By means of a simple computation it follows:

$$[(\hat\phi\,(w_{s\theta}))\,(g,h)]\,(\bar\xi)$$
$$=[\phi\,(w_{s\theta}\,(g,h))]\,(\bar\xi)=w_{s,\theta}\,(g,h)\,(\hat\phi(\bar\xi))$$
$$=w_{s,\hat\phi^{-1}\theta}\,(g,h)\,(\bar\xi).$$

Then (s,θ) and (s',θ') are equivalent iff there exists $\phi\in$Aut $(H^2_*)^\wedge$, such that (s,θ) and $(s',\hat\phi^{-1}\theta')$ have associated the same extension class, i.e., iff $s'=s\lambda$ [$\lambda\in$Hom$_t(H^2_*, B^2_*)$] and $\theta=\hat\phi^{-1}\theta'$. This last condition being always automatically satisfied for an appropriate ϕ, we obtain:

Proposition 8: Two pairs (s,θ), (s',θ') are equivalent iff $s'=s\lambda$ for some $\lambda\in$Hom$_t(H^2_*, B^2_*)$. The "parameter space" for the equivalence classes of representation groups is

$$\frac{\text{Hom}\,(H^2_*, B^2_*)}{\text{Hom}_t\,(H^2_*, B^2_*)}.$$

We exhibit out last results for practical convenience:

Theorem 6: Let s_0 be a fixed homomorphic section $s_0:H^2_*(G,T)\to Z^2_*(G,T)$. For each class

$$[\lambda]\in\frac{\text{Hom}\,(H^2_*, B^2_*)}{\text{Hom}_t\,(H^2_*, B^2_*)}.$$

select an arbitrary element $\lambda\in[\lambda]$. Then the mapping $w_\lambda:G\times G\to(H^2_*)^\wedge$, defined by

$$w_\lambda\,(g,h):\bar\xi\to(s_0(\bar\xi))\,(g,h)\cdot(\lambda\,(\bar\xi)\,(g,h)),$$

verifies $w_\lambda\in Z^2_-(G,(H^2_*)^\wedge)$ and the associated group $G_{\tilde w_\lambda}$ is a representation group. Conversely, if $G_{\tilde w}$ is a representation group for (G,K), there is a $[\lambda]$ such that, for each $\lambda\in[\lambda]$, $G_{\tilde w_\lambda}$, and $G_{\tilde w}$ are equivalent.

In the following theorem we derive a characterization of the elements of Hom$_t(H^2_*, B^2_*)$. In order to do this, remember the exact sequence of Abelian groups,

$$1\to Z^1_{*K}(G,T)\to C^1_{*K}(G,T)\xrightarrow{\delta}B^2_{*K}(G,T)\to 1.$$

The group $C^1_{*K}(G,T)$ is isomorphic to $T^{(\mathrm{Card}.G-1)}$ and $Z^1_{*K}(G,T)$ is isomorphic to $\hat K^{(1)}\otimes T$, where $\hat K^{(1)}$ is a finite group. We have:

Theorem 7: Let $\lambda\in$Hom(H^2_*, B^2_*). Then $\lambda\in$Hom$_t(H^2_*, B^2_*)$ iff there is a section $\sigma:B^2_*\to C^1_*$ whose restriction to the subgroup $\lambda\,(H^2_*)$ is a homomorphism.

Proof: Suppose $\lambda\in$Hom$_t\,(H^2_*, B^2_*)$. The triviality of Ξ_λ permits us to choose $\Xi\in C^1_{*K}(G,(H^2_*)^\wedge)$ such that $\Xi_\lambda=\delta\Xi$. For each $\bar\xi\in H^2_*$, define a function $\Lambda_{\bar\xi}:G\to T$ as $\Lambda_{\bar\xi}(g)=(\Xi(g))\,(\bar\xi)$. By construction $\Lambda_{\bar\xi}$ verifies $\Lambda_{\bar\xi}\Lambda_{\bar\xi'}=\Lambda_{\bar\xi\bar\xi'}$ and furthermore $\delta\,(\Lambda_{\bar\xi})=\lambda\,(\bar\xi)$. For the elements of B^2_* which are in $\lambda(H^2_*)$ one can define $\sigma(\lambda\,(\bar\xi))=\Lambda_{\bar\xi}$, and complete σ over $B^2_*-\lambda\,(H^2_*)$ in order to build a section. Then σ has the required properties.

For the "if" part, suppose that for some $\lambda\in$Hom (H^2_*, B^2_*), we have a section $\sigma:B^2_*\to C^1_*$ which is an homomorphism if one is restricted to $\lambda\,(H^2_*)$. For each $\bar\xi\in H^2_*$, $\lambda\,(\bar\xi)\in B^2_*$, and therefore, $\lambda\,(\bar\xi)=\delta(\sigma(\lambda\,(\bar\xi)))$, and, as a consequence of σ being an homomorphism on $\lambda\,(H^2_*)$, for every $g\in G$ the mapping $\bar\xi\to\sigma(\lambda\,(\bar\xi))\,(g)$ is in $(H^2_*)^\wedge$. Next we define $\Xi:G\to(H^2_*)^\wedge$ as $\Xi(g)\,(\bar\xi)=\sigma(\lambda\,(\bar\xi))\,(g)$. It is now clear that $\Xi_\lambda=\delta\Xi$.

This result enables us to consider the following unicity criterium for representation groups (up to equivalence):

Theorem 8: The representation group for (G,K) is unique (up to equivalence) if $\widehat{K}^{(1)}$ is trivial.[18]

Proof: The isomorphism $Z^1_{*K}(G,T) \approx \widehat{K}^{(1)} \otimes T$, shows that $Z^1_{*K}(G,T)$ is divisible iff $\widehat{K}^{(1)}$ is trivial. In this case, by the argument invoked in the demonstration of Theorem 5, there exists a global homomorphic section $B^2 \to C^1$ and then by Theorem 7, $\mathrm{Hom}_t(H^2_*, B^2_*) = \mathrm{Hom}(H^2_*, B^2_*)$. Proposition 8 shows the required result.

Theorem 8 sharpens and generalizes Schur's result[19] about the unicity (up to isomorphisms) of representation groups for the unitary case which we have obtained here as a corollary:

Corollary 1: The representation group for (G,G) is unique (up to equivalence and also up to isomorphism) if $\mathrm{Hom}(G,T)$ is trivial.

Proof: For the unitary case $\hat{K}^{(1)} = \mathrm{Hom}(G,T)$, and of course if two representation groups are equivalent, they are also isomorphic.

A more general result than Theorem 8 in the sense that it applies also to some topological groups, and of which our Theorem 8 is a particular case is given by Cattaneo.[18]

8. SEMIUNITARY PROJECTIVE REPRESENTATIONS OF (G,K)

Once a representation (or splitting) group $(\overline{G},p)$ for (G,K) is known, the SUPR's of (G,K) can be derived from the SUR's of $(\overline{G},\overline{K})$. As already remarked in Sec. 3, only the SUR's of $(\overline{G},\overline{K})$ such that $D(\ker p) \subset T$ can appear as liftings of the SUPR's of (G,K); it is easy to show that, in fact, all these representations appear as liftings of the SUPR's of (G,K). So:

Proposition 9: Let $(\bar{G},p)$ be a representation (or splitting) group for (G,K), and let R be a SUR of $(\overline{G},\overline{K})$. Then there exists a SUPR, P of (G,K) such that $P \circ p = \pi \circ R$ iff $R(\ker p) \subset T$.

In this case, let $r: G \to \overline{G}$ be an arbitrary normalized section; the mapping $g \to R(r(g))$ is a multiplier lifiting of P.

Furthermore, (projective) equivalence of SUPR's reflects itself into pseudoequivalence of the corresponding liftings (see the Appendix for the relevant definitions). Then, given a representation (or splitting) group $(\overline{G},p)$ for (G,K), all the (projective equivalence classes of) SUPR's of (G,K) can be obtained in the following way:

(i) Select, among the SUR's of $(\overline{G},\overline{K})$, those satisfying $R(\ker p) \subset T$:

(ii) Classify these SUR's into classes of pseudoequivalence (after Proposition 2 in the Appendix). Each such a class corresponds to a class of SUPR's of (G,K) in a one-to-one way.

There is an important case in which step (i) can be overlooked: If we are interested in the *irreducible* SUPR's of (G,K), it is clear that the corresponding liftings are also irreducible, and automatically, in virtue of the generalized Schur lemma[20] verify $R(\ker p) \subset T$. Thus, for ISUPR's we have:

Theorem 9: If $(\overline{G},\overline{K})$ is a representation (or splitting) group for (G,K), there is a one-to-one correspondence between the classes of equivalence of ISUPR's of (G,K) and the classes of pseudoequivalence of ISUR's of $(\overline{G},\overline{K})$.

To be complete, one must also indicate how the ISUR's of (G,K) can be calculated. This point is adequately treated in the literature[16] so we refrain from doing so here.

9. SOME EXAMPLES

Finally, we show how the preceding techniques can be used in the search of representation groups (and of SUPR's) of simple but very interesting examples.

To use the theory for some group (G,K) we have to know:

(a) $H^2_{*K}(G,T)$ as abstract group;

(b) One arbitrary homomorphic section $s_0: H^2_{*K} \to Z^2_{*K}$ (for each cohomology class, a cocycle of this class satisfying the appropriate product relations);

(c) $B^2_*(G,T)$, both as abstract group and as group of applications $G \times G \to T$.

The most difficult point is, of course (a), because (b) is very simple in each case and (c) is mechanical. We do not enter here in the problem of the determination of $H^2_{*K}(G,T)$ which is adequately discussed in the literature.[7,8,21]

In order to maintain some uniformity in our examples we adopt the following notations- $Z^2_{*K}(G,T)$ will be denoted as some subgroup of T^n, and every $\zeta \in Z^2_{*K}(G,T)$ is a n-tuple of elements of T, $(m,n,\cdots,p,q,\cdots)$ whose value in (g,h) will be given in a table. For $B^2_{*K}(G,T)$ we use a similar notation, and for $H^2_{*K}(G,T)$ the conventional one as equivalence classes of cocycles in the form $[m,n,\cdots] = \overline{(m,n,\ldots,p,q,\cdots)}$. (Remember that if $\zeta \in Z^2_*(G,T)$, then $\bar{\zeta}$ is the cohomology class of ζ.)

(1) The group (C_2, C_1)

The corresponding problem arises in quantum mechanics for the time reversal. In this case $G = C_2$ generated by α, with $\alpha^2 = 1$, and $K = C_1$ is the trivial subgroup. It is easy to show that:

$$B^2_{*C_1}(C_2, T) = 1,$$

$$Z^2_{*C_1}(C_2, T) = {}_2T, \quad (m) \equiv \begin{array}{c|c} & \alpha \\ \hline \alpha & m \end{array}, \quad m = \pm 1,$$

$$H^2_{*C_1}(C_2, T) = C_2, \quad [m] = \overline{(m)}, \quad m = \pm 1.$$

There is only one homomorphic section $s_0: H^2_* \to Z^2_*$, given by $s_0[m] = (m)$. Because $B^2_* = 1$, we have $\mathrm{Hom}(H^2_*, B^2_*) = 1$ and then there is only one (class of equivalence of) representation group. Now $(H^2_*)\hat{}$ is the cyclic group C_2 generated by the ϵ defined as $\epsilon[m] = m$. The w_1 of Theorem 6 is given by $w_1(\alpha,\alpha) = \epsilon$, and the corresponding representation group is the extension of $C_2 = \langle\alpha\rangle$ by $C_2 = \langle\epsilon\rangle$ with this factor system; the middle group is $C_4 = \langle\delta\rangle$ with projection $p(\delta) = \alpha$, and the subgroup $K_{\bar{\omega}_1}$ is $K_{\bar{\omega}_1} \approx C_2 = \langle\delta^2\rangle$.

This subgroup has only two irreducible unitary representations $\Delta_{\pm}(\delta^2) = \pm 1$. Only Δ_+ is of type I. The induced irreducible SUR's of C_4 are given by $D_+(\delta) = \mathbb{K}$ and $D_-(\delta) = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}\mathbb{K}$. Because of the different dimension, these representations cannot be pseudoequivalent, and we finally obtain the well known results (which we have rederived here only to show our method in the simplest example):

There are two classes of ISUPR's of (C_2, C_1), and a lifting of each class if given by

$$R_+(\alpha) = \mathbb{K}, \quad R_-(\alpha) = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}\mathbb{K}.$$

(2) The group (V,C_2)

The Klein's Vierergruppe V arises in physics, among other ways, as the group of connected components of classical kinematical groups.[22] We discuss here the SUPR's of (V,C_2), with $G = V = \langle\alpha,\beta\rangle, \alpha^2 = \beta^2 = 1, \alpha\beta = \beta\alpha$, and $K = C_2 = \langle\alpha\rangle$. Now:

$B^2_{*C_2}(V,T) = T \otimes T,$

		α	β	$\alpha\beta$
$(p,q) \equiv$	α	p	q^{-1}	pq
	β	$p^{-1}q^{-1}$	1	$p^{-1}q^{-1}$
	$\alpha\beta$	q	q	1

$Z^2_{*C_2}(V,T) = {}_2T \otimes {}_2T \otimes T \otimes T,$

		α	β	$\alpha\beta$
$(m,n,p,q) \equiv$	α	p	nq^{-1}	npq
	β	$mp^{-1}q^{-1}$	n	$mnp^{-1}q^{-1}$
	$\alpha\beta$	mq	q	m

$H^2_{*C_2}(V,T) = C_2 \otimes C_2, \quad [m,n] = \overline{(m,n,p,q)}.$

We choose the section $s_0: H^2_* \to Z^2_*$ given by $s_0[m,n] = (m,n,1,1)$. Furthermore $\mathrm{Hom}\,(H^2_*, B^2_*) = \mathrm{Hom}\,(H^2_*, T\otimes T) = (H^2_*)^\wedge \otimes (H^2_*)^\wedge$.

Let ϵ,η be the generators of $(H^2_*)^\wedge$ given by

$$\epsilon[m,n] = m, \quad \eta[m,n] = n.$$

Then each $\lambda \in \mathrm{Hom}\,(H^2_*, B^2_*)$ can be parametrized by a pair of elements of $(H^2_*)^\wedge$, in the form $\lambda \equiv (\epsilon^i\eta^j, \epsilon^k\eta^l)$, where $i,j,k,l, = 0,1$.

If (a, b, c) is the element of $C^1_{*C_2}(V,T)$ given by $\alpha \to a$, $\beta \to b$, $\alpha\beta \to c$, it is a trivial task to verify $\delta(a,b,c) = (a^2, c/ba)$. If $\epsilon^i\eta^j \neq 1$, there is no section $B^2_* \to C^1_*$ which restricted to $\lambda(H^2_*)$ can be a homomorphism, while such a section exists if $\epsilon^i\eta^j = 1$. Thus $\mathrm{Hom}_t(H^2_*, B^2_*) = \{(1, \epsilon^k\eta^l)\}$. The four classes of representation groups correspond, e.g., to the factor systems w_λ with $\lambda \equiv (\epsilon^i\eta^j, 1), i,j, = 0,1$. We obtain

$w_\lambda(g,h): [m,n] \to (m,n,1,1)(g,h) \times (1,1,m^i n^j, 1)(g,h)$

and, in an explicit form

$w_{\lambda_{ij}}(g,h)$	α	β	$\alpha\beta$
α	$\epsilon^i\eta^j$	η	$\epsilon^i\eta^{j+1}$
β	$\epsilon^{1-i}\eta^j$	η	$\epsilon^{1-i}\eta^{1-j}$
$\alpha\beta$	ϵ	1	ϵ.

The four corresponding extensions of $V \approx \langle\alpha,\beta\rangle$ by $(H^2_*)^\wedge = \langle\epsilon,\eta\rangle$ lead to representations of each class. The cases $(i,j) = (0,0), (0,1), (1,0)$ furnish non-Abelian groups whereas $(i,j) = (1,1)$, whose factor system is

$w_{\lambda_{11}}(g,h)$	α	β	$\alpha\beta$
α	$\epsilon\eta$	η	ϵ
β	η	η	1
$\alpha\beta$	ϵ	1	ϵ ,

gives rise to an Abelian group isomorphic to $C_4 \otimes C_4 = \langle\delta,\gamma\rangle$ [the isomorphism $V_{\bar{\omega}_{\lambda_{11}}} \approx C_4 \otimes C_4$ is given by $(1,\alpha) \leftrightarrow \delta$, $(1,\beta) \leftrightarrow \delta$], and with projection $p(\delta) = \alpha$, $p(\gamma) = \beta$ and the subgroup $K_{\bar{\omega}_{\lambda_{11}}} \approx C_4 \otimes C_2 = \langle\delta,\gamma^2\rangle$. We shall not go into details of the structure of the other representation groups, because this, being Abelian, is easier to use. Next we derive the ISUPR's of (V, C_2) by means of this representation group $V_{\bar{\omega}_{\lambda_{11}}}$ which we denote simply $\bar{V}$. The IUR's of $\bar{C}_2$ are labelled $\Delta_{i^n,\pm}$ and given by $\Delta_{i^n,\pm}(\delta,\gamma^2) = \pm i^n$. If we choose $g_0 = \delta$, Dimmock's test gives

Δ	$\Delta_{1,+}$	$\Delta_{-1,+}$	$\Delta_{i,+}$	$\Delta_{-i,+}$	$\Delta_{1,-}$	$\Delta_{-1,-}$	$\Delta_{i,-}$	$\Delta_{-i,-}$
$\sum\chi_\Delta$	8	8	0	0	-8	-8	0	0

The ISUR's of $(\bar{V}, \bar{C}_2)$ induced from $\Delta_{1,+}$ and $\Delta_{-1,+}$ are pseudoequivalent, as well as those induced from $\Delta_{1,-}$ and $\Delta_{-1,-}$. Then we obtain four classes of pseudoequivalence of ISUR's of $(\bar{V}, \bar{C}_2)$, and therefore four classes of ISUPR's of (V,C_2): a lifting of each one is given by

R	Induced from	$R(\alpha)$	$R(\beta)$	$R(\alpha\beta)$
R_+	$\Delta_{1,+}$	$\mathbb{1}$	$\mathbb{K}$	$\mathbb{K}$
R_{i+}	$\Delta_{i,+}, \Delta_{-i,+}$	$i\mathbb{1}_{1,1}$	$\mathbb{A}\,\mathbb{K}$	$i\mathbb{S}\mathbb{K}$
R_-	$\Delta_{1,-}$	$\mathbb{1}$	$\mathbb{S}\mathbb{K}$	$\mathbb{S}\mathbb{K}$
R_{i-}	$\Delta_{i,-}, \Delta_{-i,-}$	$i\mathbb{1}_{1,1}$	$-\mathbb{S}\mathbb{K}$	$-i\mathbb{A}\mathbb{K}$,

where $\mathbb{K}$ is the complex conjugation in some fixed basis and

$$\mathbb{1} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad \mathbb{1}_{1,1} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad \mathbb{S} = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}.$$

These four ISUPR's of (V,C_2) are related with the "types" of the elementary particles. We remark that, as a consequence of the complete character of our method of construction of representation groups, we can choose a "simpler" representation group (in this case, an Abelian one), whereas Janssen's method leads to one of the non-Abelian ones. [8]

(3) The group (Q,C_4)

The quaternionic group Q is not a crystallographic group, and its ISUPR's have not been studied, at least to the best of our knowledge. The UPR's of (Q,Q) can be found from the UR of (Q,Q) because $H^2_{*Q}(Q,T) = 1$.[23] Now, we are going to find the SUPR's of Q with respect to a subgroup of index 2. Here $G \equiv Q = \langle \alpha,\beta \rangle$, with $\alpha^4 = 1, \beta^2 = \alpha^2, \alpha\beta = \beta\alpha^3$, and $K = C_4 = \langle \alpha \rangle$. Now

$$B^2_{*C_4}(Q,T) = \bigotimes_{i=1}^{6} T\ (pqrstu),\ Z^2_{*C_4}(Q,T) = {}_2T \otimes (\bigotimes_{i=1}^{6} T)\ (m,pqrstu),\ H^2_{*C_4}(Q,T) = {}_2T\ \ [m] = \overline{(m,pqrstu)}.$$

(p,q,r,s,t,u)	α	α^2	α^3	β	$\alpha\beta$	$\alpha^2\beta$	$\alpha^3\beta$
α	$pq^{-2}r^{-1}$	$q^{-1}r^{-2}$	$pq^{-1}r^{-1}$	u	$pq^{-2}r^{-1}su^{-1}$	$s^{-1}tu$	$pq^{-2}r^{-3}t^{-1}u^{-1}$
α^2	$q^{-1}r^{-2}$	r^{-2}	q^{-1}	s	t	$r^{-2}s^{-1}$	r^2t^{-1}
α^3	$pq^{-1}r^{-1}$	q^{-1}	pr	$q\,r^2tu$	$pq^{-1}r^{-1}u^{-1}$	$qs^{-1}u$	$pq^{-1}r^{-1}st^{-1}u^{-1}$
β	$p^{-1}q^2r^3tu$	r^2s	$p^{-1}qru$	r	$p^{-1}q^2r^2u$	rs	$p^{-1}qr^3tu$
$\alpha\beta$	u^{-1}	r^2t	$q^{-1}su^{-1}$	$q^{-1}r^{-1}u^{-1}$	r	rsu^{-1}	rt
$\alpha^2\beta$	$p^{-1}q^2rs^{-1}u$	s^{-1}	$p^{-1}q\,rs^{-1}tu$	$r^{-1}s^{-1}$	$p^{-1}qs^{-1}u$	r	$p^{-1}q^2rs^{-1}tu$
$\alpha^3\beta$	$st^{-1}u^{-1}$	t^{-1}	$q^{-1}r^{-2}t^{-1}u^{-1}$	$r^{-1}t^{-1}u^{-1}$	$r^{-1}t^{-1}$	$q^{-1}r^{-1}st^{-1}u^{-1}$	r

$(m,1,1,1,1,1,1)$	α	α^2	α^3	β	$\alpha\beta$	$\alpha^2\beta$	$\alpha^3\beta$
α	1	1	1	1	1	1	1
α^2	1	1	1	1	1	1	1
α^3	1	1	1	1	1	1	1
β	m	1	m	1	m	1	m
$\alpha\beta$	m	1	m	1	m	1	m
$\alpha^2\beta$	m	1	m	1	m	1	m
$\alpha^3\beta$	m	1	m	1	m	1	m

w_{λ_i}	α	α^2	α^3	β	$\alpha\beta$	$\alpha^2\beta$	$\alpha^3\beta$
α	ϵ^i	1	ϵ^i	1	ϵ^i	1	ϵ^i
α^2	1	1	1	1	1	1	1
α^3	ϵ^i	1	ϵ^i	1	ϵ^i	1	ϵ^i
β	ϵ^{1+i}	1	ϵ^{1+i}	1	ϵ^{1+i}	1	ϵ^{1+i}
$\alpha\beta$	ϵ	1	ϵ	1	ϵ	1	ϵ
$\alpha^2\beta$	ϵ^{1+i}	1	ϵ^{1+i}	1	ϵ^{1+i}	1	ϵ^{1+i}
$\alpha^3\beta$	ϵ	1	ϵ	1	ϵ	1	ϵ

We choose s_0 as $s_0[m] = (m,1,1,1,1,1,1)$. Let ϵ be the generator of $(H^2_*)\hat{}$, $\epsilon[m] = m$, and then $\mathrm{Hom}(H^2_*, B^2_*) = \otimes^6_{i=1}(H^2_*)\hat{}$. By means of Theorem 7, it is easy to show that a $\lambda \in \mathrm{Hom}(H^2_*, B^2_*)$, parametrized by a 6-tuple of elements of $(H^2_*)\hat{}$ lies in $\mathrm{Hom}_t(H^2_*, B^2_*)$ iff the first element of the 6-tuple is trivial. Then, a representation group in each of the two splitting classes is obtained with $\lambda = (\epsilon^i,1,1,1,1,1)$, $i = 0,1$. The corresponding w_{λ_i} is

The representation groups so obtained are two non-Abelian groups, which, although in different classes, are isomorphic; the one corresponding to $\lambda = (1,1,1,1,1,1)$ is the group generated by δ,γ with relation $\delta^4 = 1, \gamma^4 = 1$, $\delta\gamma = \gamma^{-1}\delta$, and projection $p(\delta) = \alpha, p(\gamma) = \beta\alpha$. We do not enter into the actual construction of the ISUPR's of (Q, C_4) from the ISUR's of this representation group, and we only quote the result obtained: There are four classes of ISUPR's of (Q,C_4).

APPENDIX

If R is a SUMR of (G,K), $\pi(R(g)R(h)) = \pi(R(gh))$ for any pair $(g,h)\in G\times G$ and therefore there is a map $\omega: G\times G\to, T$ (the so-called multiplier, or factor system of G) such that $R(g)R(h) = \omega(g,h)R(gh)$. Associativity implies $\omega(g,h)\omega(gh,k) = \omega(g,hk)\omega^{g}(h,k)$. So, any SUMR of (G,K) determines an element of $Z^2_{\cdot}(G,T)$ (cocycle). Such a SUMR R will be written R_ω.

To any SUMR R_ω is associated the SUPR $P=\pi\circ R_\omega$. If P is a SUPR of (G,K) and $\sigma:(P\Gamma\,\mathrm{U}(H)\to\Gamma\,\mathrm{U}(H)$ is a normalized section ($\pi\circ\sigma = \mathrm{id}_{P\Gamma\,\mathrm{U}}$, $\sigma(1)=1$), $P_\sigma = \sigma\circ P$ is a SUMR of (G,K) which is also said to be a lifting of P. When two SUMR's R_ω and $Q_{\omega'}$ have associated the same SUPR of (G,K), there exists a map $\Lambda\in C^1_{\cdot}(G,T)$ such that

$$\omega'(g,h) = \omega(g,h)\frac{\Lambda(g)\Lambda^{g}(h)}{\Lambda(gh)} = (\omega\delta\Lambda)(g,h),$$

that is to say, ω' and ω are cohomologous. So, to any SUPR P of (G,K) corresponds an element of $H^2_{\cdot}(G,T) = Z^2_{\cdot}(G,T)/B^2_{\cdot}(G,T)$. However, different SUPR's with the same cohomology class $\omega\in H^2_{\cdot}(G,T)$ can exist.

Proposition 1: Let P be a SUPR of (G,K) with cohomology class $\bar\omega$. If $\omega\in Z^2_{\cdot}(G,T)$ is an arbitrary lifting of $\bar\omega$ there exists a SUMR R_ω which is a lifting of $P_{\bar\omega}$.

Proof: Let σ: $P\Gamma\,\mathrm{U}(H)\to\Gamma\,\mathrm{U}(H)$ be an arbitrary normalized section. Then $Q=\sigma\circ P_{\bar\omega}$ is a SUMR with a multiplier to be denoted ω^{σ}. As ω^{σ} and ω are cohomologous, $\omega^{\sigma} = \omega\delta\Lambda$ with $\Lambda\in C^1_{\cdot}(G,T)$. If $R:G\to\Gamma\,\mathrm{U}(H)$ is defined by $R(g)=\Lambda(g)Q(g)$, one can verify easily that R is a SUMR of (G,K) with factor system ω.

The definitions of equivalence we adopt are:

Definition 1: Two SUR's of (G,K), D, and D' (with the same support space) are said to be equivalent if there is $A\in\Gamma\,\mathrm{U}(H)$ such that $AD(g) = D'(g)A$ for any $g\in G$. If $A\in\mathrm{U}(H)$ one calls it unitary equivalence and if $A\in\Gamma\,\mathrm{U}(H) - \mathrm{U}(H)$ one calls it antiunitary equivalence.

Definition 2: Two SUPR's of (G,K), P and P' are said to be (projectively) equivalent if there exists $S\in P\Gamma\,\mathrm{U}(H)$ such that $SP(g) = P'(g)S$ for any $g\in G$. If $S\in P\,\mathrm{U}(H)$ we call it unitary equivalence and if $S\in P\Gamma\,\mathrm{U}(H) - P\,\mathrm{U}(H)$ we call it antiunitary equivalence.

With respect to semiunitary multiplier representations:

Definition 3: Let R_ω and R'_ω be two SUMR's of (G,K):

(i) R_ω and $R'_{\omega'}$ are said to be equivalent if there is $A\in\Gamma\,\mathrm{U}(H)$ such that $AR_\omega(g) = R'_{\omega'}(g)(A)$;

(ii) R_ω and $R'_{\omega'}$ are said to be pseudoequivalent if $\pi\circ R$ and $\pi\circ R'_{\omega'}$ are equivalent,

(iii) R_ω and $R'_{\omega'}$ are similar, if $\pi\circ R_\omega$ and $\pi\circ R'_{\omega'}$ are identical.

Definition 4: Let R and R' be two SUR's of $(\bar G,\bar K)$ such that $R(\ker p)\subset T\, R'(\ker p)\subset T$; we will say that R and R' are pseudoequivalent if the associated SUPR's of (G,K) are equivalent.

It is evident that an intrinsic criterion of pseudoequivalence, without making any reference to the associated SUPR is the following:

Proposition 2: Two SUR, R and R' of $(\bar G,\bar K)$ such that $R(\ker p)\subset T$, $R'(\ker p)\subset T$ are pseudoequivalent iff there exists a $V\in\Gamma\,\mathrm{U}(H)$ and a mapping $\lambda:\bar G\to T$ such that

$$R'(a,g) = \lambda(a,g)VR(a,g)V^{-1}.$$

The definitions of reducibility of the representations are as always:

Definition 5: A SUR of (G,K) is irreducible if the only invariant subspaces are $\{0\}$ and H.

Definition 6: A SUPR of (G,K) is irreducible if the only invariant (projective) subspace is H.

Definition 7: A SUMR R of (G,K) is irreducible if the only invariant subspaces under the set $\{R(g)|\, g\in G\}$ are $\{0\}$ and H.

[1] M. Santander, "Linealización de representaciones proyectivas semiunitarias de grupos," Universidad de Valladolid, 1977.
[2] J.F. Cariñena and M. Santander, J. Math. Phys. **16**, 1416 (1975).
[3] E.P. Wigner, Ann. Math., **40**, 149 (1939); also U. Ulhorn, Ark. Fysik **23**, 307 (1962); V. Bargmann, J. Math. Phys. **5**, 862 (1964). V.S. Varadarajan, *Geometry of Quantum Theory* (Van Nostrand, New York, 1968, 1970), Vols. I and II.
[4] W.G. Harter, J. Math. Phys. **10**, 739 (1969).
[5] I.M. Isaacs, *Character Theory of Finite Groups* (Academic, New York, 1976), pp. 176–7.
[6] For the more general case of Polish G and semiunitary representations, see, e.g., U. Cattaneo, Rep. Math. Phys. **9**, 31 (1976) and references therein for the previous work; for finite G the results of this paper, of course, hold with some evident simplifications.
[7] J. Schur, J. Reine Angew. Math. **127**, 20 (1904); **132**, 85 (1907); **139**, 155 (1911).
[8] T. Janssen, J. Math. Phys. **13**, 342 (1972).
[9] See, e.g., N.B. Backhouse, Physica **70**, 505 (1970) and references therein; also J.M. Levy-Leblond, "Galilei Group and Galilean invariance," in *Group Theory and Its Applications*, edited by E. Loebl (Academic, New York, 1971), Vol. II, and references therein.
[10] See, e.g., U. Cattaneo and A. Janner, J. Math. Phys. **15**, 1155 (1974), and references therein. This paper deals with a more general case than finite groups.
[11] C.C. Moore, Trans. Am. Math. Soc. **113**, 40, 64 (1964).
[12] U. Cattaneo, J. Math. Phys. **19**, 452 (1978).
[13] V. Bargmann, Ann. Math. **59**, 1 (1954).
[14] G.W. Mackey, Acta Math. **99**, 265 (1968), Theorem 4.2.
[15] S. MacLane, *Homology* (Springer, Berlin, 1967).
[16] E.P. Wigner, *Group Theory and Its Applications* (Academic, New York, 1971), Sec. 26. Also L. Janssen and M. Boon, *Theory of Finite Groups. Applications in Physics* (North-Holland, Amsterdam, 1967); R. Shaw and J. Lever, Commun. Math. Phys. **38**, 257 (1974).
[17] J.O. Dimmock, J. Math. Phys. **4**, 1307 (1963).
[18] U. Cattaneo, J. Math. Phys. **19**, 767 (1978), Appendix B.
[19] J. Schur, J. Reine Angew. Math. **127**, 20, (1904), Satz IV.
[20] See, e.g., Ref. 7, Lemma 3.
[21] L. Michel, "Invariance in Q.M. and group extensions," in *Istanbul Lectures*, 1962, edited by F. Gursey (Gordon & Breach, New York, 1964); C.J. Bradley and D.E. Wallis, Q. J. Math. (Oxford) **25**, 85 (1974); P. Rudra, J. Math. Phys. **17**, 509, 512 (1976); P.M. Van den Broek, J. Phys. A. Math. Gen. **10**, 649 (1977) and references therein.
[22] H. Bacry and J.M. Levy-Leblond, J. Math. Phys. **9**, 1605 (1968).
[23] L. Michel, Ref. 21, p. 169.

Micu-type invariants of simple Lie algebras

Vishnu K. Agrawala

Department of Natural Sciences and Technology, Point Park College, Pittsburgh, Pennsylvania 15222

(Received 27 September 1978)

An algorithm is presented for computing the eigenvalues of Micu-type invariants for Lie algebras A_r, B_r, C_r, and D_r. Inasmuch as the definitions of these invariants is recursive and not in terms of a closed formula, only a recursion relation is obtained for computation of their eigenvalues, which is easily programmable on a computer. Some of the formulas developed here are valid in arbitrary complex semisimple Lie algebras and may be useful in other contexts.

1. INTRODUCTION

An algorithm was described some time ago[1] for computation of various types of invariants for the Lie algebras A_r corresponding to the unitary unimodular groups SU(n), $n = r + 1$. If one is interested only in Racah-type invariants (denoted by type C in Ref. 1), this algorithm can be applied in its basic form itself to other simple Lie algebras as well. Explicit computations of Racah-type invariants by this procedure have been published recently for Lie algebras A_r, B_r, C_r, and D_r[2] and for Lie groups U(n), O(n), and Sp(n).[3]

In order to apply this algorithm for other invariants, one possible approach is to derive recursion relations for the quantities involved by using commutation relations and other algebraic manipulations. A few examples of non-Racah-type invariants for SU(n) were treated this way in Ref. 1. Here we obtain recursion formulas that are helpful in the computation of Micu-type[4] invariants for the Lie algebras A_r, B_r, C_r and D_r. In as much as these invariants are defined recursively, it is not surprising that closed formulas are not obtained. However, the recursion formulas are amply suited for machine computations. Computer programs for the algorithm developed here are being prepared and will be available shortly to the interested readers.

Concerning general aspects of notation, throughout this paper, G will stand for a compact connected Lie group and L its Lie algebra, complexified if necessary. The complex field will be denoted by $\mathbb{C}$. Also we assume G is semisimple and finit-dimensional so that L also has these properties. The universal enveloping algebra of L will be denoted by U.

In Sec. 5 and onward, a further assumption of simplicity will be imposed on G and L. Further specializations will be pointed out when they occur.

The product in the group G will be denoted by juxtaposition gh for $g,h\in G$ and the product in the Lie algebra L is denoted by $[X,Y]$ for all $X,Y\in L$. Juxtaposition is also used for products in U and for compositions of mappings, a map α followed by a map β being written as $\beta\alpha$. A dot as in $g\cdot m$ or $X\cdot m$ will be used to denote the action of the group or the algebra on a module. Also, we shall identify the canonical image of L in U with L itself although the Lie product $[X,Y]$ is then replaced by the commutator $XY - YX$.

Greek indices run from 1 through dim$L$, whereas the range of Latin indices will depend on the context. Summation over repeated indices is implied unless stated otherwise.

2. LIE ALGEBRAS AND THEIR MODULES

We summarize here some useful facts about Lie groups and Lie algebras and their representations. This is for reference purposes, mainly to establish notation, to clarify the sense in which some common terms have been used in this paper and to have easy access to the basic results we need. More details and proofs may be found in standard texts.[5-9]

Quite often it is necessary to translate statements about G to statements involving its Lie algebra L and vice versa. This is achieved by noting the fact that for every $X\in L$, there is a unique one-parameter subgroup of G whose elements are denoted by $\exp(tX)$, with parameter t, and satisfy

$$\exp(0) = 1, \tag{2.1}$$

$$\frac{d}{dt}(\exp(tX))_{t=0} = X. \tag{2.2}$$

An important application of (2.1) and (2.2) for us is that any inner automorphism of G induces an automorphism of L. Let Adg denote the conjugation by $g\in G$: $(\mathrm{Ad}g)(h) = ghg^{-1}$ for each $h\in G$. Then for each $X\in L$, we define

$$(\mathrm{Ad}g)(X) = \left(\frac{d}{dt}(g\exp(tX)g^{-1})\right)_{t=0}$$

$$= g\left(\frac{d}{dt}\exp(tX)\right)_{t=0} g^{-1} = gXg^{-1}. \tag{2.3}$$

Adg is an automorphism of L and the set of all such automorphisms is a group called the *adjoint group* AdG of G.

By a *module* M over G we shall mean a finite-dimensional vector space M over complex numbers such that for every $g\in G$ and every $m\in M$, a group action $g\cdot m$ is defined and belongs to M and satisfies the following properties:

(a) $e\cdot m = m$, where e is the identity element of G,

(b) for $g,h\in G$ and $m\in M$, $g\cdot(h\cdot m) = (gh)\cdot m$.

(c) $g\cdot m$ is a $\mathbb{C}$-linear function of m, (2.4)

(d) $g\cdot m$ is a continuous function of g and m.

 0022-2488/79/112178-09$01.00

Elements of G are then representable by elements of $\mathrm{Aut}_{\mathbb{C}}M$, or even by matrices of order equal to the dimension of M.

We can then regard M as a module over L with $X\in L$ acting on $m\in M$ through

$$X\cdot m = \left(\frac{d}{dt}\exp(tX)\cdot m\right)_{t=0}. \tag{2.5}$$

It is easy to verify the following properties for all $X,Y\in L$ and $m\in M$:

$$\text{(a) } [X,Y]\cdot m = X\cdot(Y\cdot m) - Y\cdot(X\cdot m), \tag{2.6}$$

(b) $X\cdot m$ is linear in X and m.

We shall write $X(M)$ for the linear operator in M representing the element X of L.

Lastly, M can be regarded as a module over the associative algebra U with the additional condition $1_U m = m$.

A submodule $S\subset M$ is an invariant subspace of M, i.e., $g\cdot S\subset S$ for all $g\in G$, or $X\cdot S\subset S$ for all $X\in L$. M is called *irreducible* if its only submodules are the trivial submodule $\{0\}$ and M itself.

In particular, L itself is an L module giving the *adjoint representation* in which each $X\in L$ is represented by the linear operator $\mathrm{ad}X: L\to L$ such that $(\mathrm{ad}X)(Y) = [X,Y]$ for all $Y\in L$. the trace operation in the adjoint representation defines the so-called *Killing form*

$$(X,Y) = \mathrm{Tr}_L(\mathrm{ad}X,\mathrm{ad}Y), \tag{2.7}$$

which is a nondegenerate bilinear form on $L\times L$ if L is semisimple.

Given G modules M and N, we can form the tensor product $M\otimes N$ of the two vector spaces and make it into a G module by defining the action of $g\in G$ by

$$g\cdot(m\otimes n) + (g\cdot m)\otimes(g\cdot n) \tag{2.8}$$

for all $m\in M$, $n\in N$, and extended to all of $M\otimes N$ by linearity. In terms of Lie algebra L this equation translates as

$$X\cdot(m\otimes n) = (X\cdot m)\otimes n + m\otimes(X\cdot n) \tag{2.9}$$

for all $X\in L$.

On the other side, we shall write $\mathrm{lin}(M,N)$ for the vector space of all $\mathbb{C}$-linear maps from M to N. Then $\mathrm{lin}(M,N)$ can be made into a G (or L) module by defining the action of $g\in G$ on $\alpha\in\mathrm{lin}(M,N)$ by

$$(g\cdot\alpha)(m) = g\cdot(\alpha(g^{-1}\cdot m)) \tag{2.10}$$

for all $m\in M$. Elements $X\in L$ act according to

$$(X\cdot\alpha)(m) = X\cdot(\alpha m) - \alpha(X\cdot m). \tag{2.11}$$

The complex numbers $\mathbb{C}$ form a G module (resp. L module) under the trivial operation $g\cdot c = c(X\cdot c = 0)$ for all $g\in G$ ($X\in L$) and $c\in\mathbb{C}$. Therefore, the action of G (resp. L) on the dual module $M^* = \mathrm{lin}(M,\mathbb{C})$ is given by

$$(g\cdot\mu)(m) = \mu(g^{-1}\cdot m), \quad (X\cdot\mu)(m) = -\mu(X\cdot m), \tag{2.12}$$

for each $g\in G$ ($X\in L$), $\mu\in M^*$, and $m\in M$. With these definitions, $\mathrm{lin}(M,N)$ and $N\otimes M^*$ are isomorphic as modules, that is, there is a vector space-isomorphism between them commuting with elements of G (or L). This isomorphism allows us to regard a typical element $\Sigma_{i\in I} n_i\otimes\mu_i$ of $N\otimes M^*$ as a linear operator from M to N where

$$\sum_{i\in I}(n_i\otimes\mu_i)(m) = \sum_{i\in I}\mu_i(m)n_i \tag{2.13}$$

for each $m\in M$.

A submodule of $\mathrm{lin}(M,N)$ is termed a *tensor operator submodule*. These may be irreducible or reducible. The irreducible ones have been called *irreducible tensorial sets* by Fano and Racah[10] and their individual members have been called *irreducible tensor operators*. In particular, if $\mathrm{lin}(M,N)$ has a submodule V which is isomorphic to L (adjoint representation), then V is called a *vector operator submodule*, its individual members being called *vector operators*.

Of particular interest in the invariant theory is the *scalar operator submodule* S of $\mathrm{lin}(M,N)$ consisting of all G-invariant elements, that is, elements $\alpha\in\mathrm{lin}(M,N)$ such that

$$g\cdot\alpha = \alpha, \quad X\cdot\alpha = 0 \tag{2.14}$$

for all $g\in G$ and all $X\in L$. In view of (2.10), this says

$$g\cdot(\alpha m) = \alpha(g\cdot m), \quad X\cdot(\alpha m) = \alpha(X\cdot m) \tag{2.15}$$

for all $m\in M$. Thus elements of S are module homomorphisms and so S is often appropriately denoted by $\mathrm{Hom}_G(M,N)$ or $\mathrm{Hom}_L(M,N)$.

If H denotes a Cartan subalgebra of L (semisimple), then roots and weights are linear forms over H, that is, belong to H^*. The Killing form induces a nondegenerate bilinear form in H^* which we shall again denote by $(\ ,\ )$. If $\alpha_1,\ldots,\alpha_r$ (r = rank of L) denote the simple roots, then for any weight μ, the numbers $m_i = 2(\mu,\alpha_i)/(\alpha_i,\alpha_i)$, $i = 1,\ldots,r$, are nonnegative integers; they will be called the *Dynkin indices* of μ and we shall write $\mu = (m_1,\ldots,m_r)$. Every irreducible module of L is uniquely specified by the Dynkin indices of its highest weight.

The modules having highest weights $\omega_1 = (1,0,\ldots,0)$, $\omega_2 = (0,1,0,\ldots,0),\ldots,\omega_r = (0,0,\ldots,1)$ are called the *basic modules* and the one with highest weight ω_1 will be termed the *fundamental module*. The Dynkin indices of a weight are simply its components relative to the basis $\{\omega_1,\ldots,\omega_r\}$. The Dynkin indices of simple roots are nothing but columns of the Cartan matrix of L.[11]

From now on, M will denote an irreducible module with highest weight $\mu = (m_1,\ldots,m_r)$. The eigenvalue $\gamma(M)$ of the second order Casimir invariant is given by Weyl's formula

$$\gamma(M) = (\mu,\mu + 2\delta), \tag{2.16}$$

where δ is half the sum of positive roots. The Dynkin indices of δ are $(1,1,\ldots,1)$. Weyl also gave a formula for the dimension of M,

$$\dim M = \prod_{\alpha>0}(\alpha,\mu+\delta)\Big/\prod_{\alpha>0}(\alpha,\delta), \tag{2.17}$$

where the product is over all positive roots α.

Practical computations involving these formulas can be performed easily with the help of the matrix of the bilinear form (,) in H^* relative to the basis $\{\omega_1,...,\omega_r\}$.[11] If one prefers, the use of this matrix can be avoided by using the Young indices $[f_1,...,f_r]$ to specify an irreducible module of L; but then the formulas are different for each Lie algebra and may be found in Refs. 2 and 3.

3. CONSTRUCTION OF INVARIANTS

We define an invariant of G or L as any element I of the center of the universal enveloping algebra U of L. Since U is generated by L, I belongs to the center of U if and only if I commutes with every element of L. This again is equivalent to the condition[12] $\exp(tX)I\exp(-tX) = I$ for all $X\in L$, that is, I is invariant under the action of the adjoint group. In any module M over G or L, I is represented by a scalar operator $I(M)\in \mathrm{Hom}_G(M,M)$, that is,

$$gI(M) = I(M)g \tag{3.1}$$

for every $g\in G$. If M is an irreducible module, then by Schur's Lemma, $I(M)$ is a multiple $i(M)$ of the unit operator 1_M in M. We are interested in computing the number $i(M)$ known as the *eigenvalue of I in the irreducible module M*. For a semisimple Lie group of rank r there are exactly r algebraically independent invariants[13] and their eigenvalues can be used instead of Dynkin indices to specify irreducible modules of L.[14]

For explicit construction of invariants, let $\{G_\alpha | \alpha = 1,...,\dim L\}$ denote any basis of L. (Throughout this paper, greek indices run from 1 through $\dim L$). We define a metric tensor $g_{\alpha\beta}$ using the Killing form,

$$g_{\alpha\beta} = (G_\alpha, G_\beta) = \mathrm{Tr}_L(\mathrm{ad}G_\alpha, \mathrm{ad}G_\beta). \tag{3.2}$$

Thus $(g_{\alpha\beta})$ is just the matrix of the Killing form relative to the above basis. Since L is semisimple, there exists a reciprocal metric tensor $g^{\alpha\beta}$, with the help of which we can raise indices to define the correlate basis for L, namely

$$G^\alpha = g^{\alpha\beta}G_\beta. \tag{3.3}$$

(Note that summation is implied over repeated indices.)

The simplest invariant is that of Casimir which is of order 2 and is defined by

$$\Gamma = G_\alpha G^\alpha, \tag{3.4}$$

where G_α and G^α are considered as members of the universal enveloping algebra U. This definition is independent of the basis chosen[12] and normalizes Γ to have the eigenvalue 1 in the adjoint representation.

To define other invariants, let F denote any faithful module over L and let $F_\alpha = G_\alpha(F)$ denote the linear operators representing the basis $\{G_\alpha\}$ in this module. Racah-type invariants are defined by

$$C_k = (\mathrm{Tr}_F(F_{\alpha_1}\cdots F_{\alpha_k}))G^{\alpha_1}\cdots G^{\alpha_k}, \tag{3.5}$$

where $k = 2,3,\cdots$. If L is simple, then any L module other than $\{0\}$ and $\mathbb{C}$ may be used for F. Gruber and O'Raifeartaigh,[15] in discussing the algebraic independence of invariants, chose F to be the fundamental module for almost all the classical simple groups. With F as the fundamental module, the eigenvalues of the Racah-type invariants have been explicitly calculated in Ref. 1 for SU(n), in Ref. 3 for the Lie groups U(n), O(n), and Sp(n), and in Ref. 2 for the Lie algebras A_r, B_r, C_r, D_r, and G_2.

Another sequence of invariants has been defined by Micu[4] as follows. Let F denote any faithful module over L. Then by semisimplicity of L, the trace form $\mathrm{Tr}_F(X(F), Y(F))$, $X, Y\in L$, is nondegenerate, i.e., if $\mathrm{Tr}_F(X(F), Y(F)) = 0$ for every $Y\in L$, then $X = 0$. If L is simple, then F could be any nontrivial module. First one defines the collection of numbers

$$u_{\alpha\beta\gamma\delta} = \mathrm{Tr}_F(F_\alpha F_\beta F_\gamma F_\delta). \tag{3.6}$$

Using these coefficients, one defines a sequence of "vector" elements in U recursively,

$$\begin{aligned} v_\delta^{(1)} &= G_\delta, \\ v_\delta^{(3)} &= u_{\alpha\beta\gamma\delta}G^\alpha G^\beta G^\gamma, \\ v_\delta^{(2k-1)} &= u_{\alpha\beta\gamma\delta}v^\alpha_{(2k-3)}G^\beta G^\gamma, \quad k = 1,2,\cdots. \end{aligned} \tag{3.7}$$

It is shown in Appendix A that for each $k = 1,2,\cdots$, the elements $v_\delta^{(2k-1)}$, $k = 1,2,\cdots$ are linearly independent and satisfy the same commutation relations as the basis elements G_δ of L.

Micu-type invariants are then defined as

$$\begin{aligned} U_2 &= v_\delta^{(1)}G^\delta = G_\delta G^\delta, \\ U_4 &= v_\delta^{(3)}G^\delta = u_{\alpha\beta\gamma\delta}G^\alpha G^\beta G^\gamma G^\delta, \\ &\cdots\cdots\cdots\cdots \\ U_{2k} &= v_\delta^{(2k-1)}G^\delta = u_{\alpha\beta\gamma\delta}v^\alpha_{(2k-3)}G^\beta G^\gamma G^\delta, \end{aligned} \tag{3.8}$$

$k = 1,2,\cdots$. Their invariant character is established in Appendix A.

It should be pointed out that Micu constructed his invariants using the symmetrized coefficients

$$s_{\alpha\beta\gamma\delta} = \mathrm{Tr}_F((F_\alpha F_\beta F_\gamma)_{\mathrm{symm}}F_\delta)$$

instead of the unsymmetrized $u_{\alpha\beta\gamma\delta}$. The unsymmetrized invariants are, of course, simpler and more likely to occur in applications. Part of the algorithm described here involves permuting the factors in $F_\alpha F_\beta F_\gamma$, so that it can be easily adapted to cover the symmetrized invariants, if necessary.

4. SOME LIE ALGEBRA IDENTITIES

We derive here some useful results valid in an arbitrary semisimple Lie algebra. The notation is as before except $G_\alpha(M)$ denotes the representation of G_α in a module M. The most fundamental identity, of course, is the commutation relations

$$G_\alpha(M)G_\beta(M) - G_\beta(M)G_\alpha(M) = f_{\alpha\beta}^{\ \ \gamma}G_\gamma(M), \tag{4.1}$$

where the $f_{\alpha\beta}^{\ \ \gamma}$ are the structure constants of L, obviously skew symmetric in α and β. Defining $f_{\alpha\beta\gamma} = g_{\gamma\delta}f_{\alpha\beta}^{\ \ \delta}$, it is easy to show that the $f_{\alpha\beta\gamma}$ are unchanged under a cyclic permutation of indices and are skew symmetric under interchange of any two indices.

Let M and N be irreducible L modules. If L is an irreducible tensor operator submodule of lin(M,N), then the relation (2.11) written more explicitly reads

$$[G_\alpha(T)](t) = G_\alpha(N)t - tG_\alpha(M) \quad (4.2)$$

for each $t \in T$. Applying $G^\alpha(T)$ to this equation we get

$$\Gamma(T)(t) = [G^\alpha(N)t - tG^\alpha(M)][G_\alpha(N)t - tG_\alpha(M)]$$

which, after simplification, gives

$$G^\alpha(N)tG_\alpha(M) = \tfrac{1}{2}[\gamma(M) + \gamma(N) - \gamma(T)]t. \quad (4.3)$$

Now suppose T is isomorphic to the adjoint representation module L with an isomorphism $\sigma: L \to T$ which maps the basis $\{G_\alpha\}$ of L into the basis $\{t_\alpha = \sigma(G_\alpha)\}$ of T. then

$$\begin{aligned} G_\alpha(T)(t_\beta) &= \sigma^{-1}\sigma G_\alpha(T)\sigma^{-1}\sigma t_\beta \\ &= \sigma^{-1}(\mathrm{ad}G_\alpha)G_\beta \\ &= \sigma^{-1}[G_\alpha, G_\beta] \\ &= \sigma^{-1}(f_{\alpha\beta}{}^{\gamma} G_\gamma) \\ &= f_{\alpha\beta}{}^{\gamma} t_\gamma. \end{aligned} \quad (4.4)$$

Therefore, from (2.11),

$$G_\alpha(N)t_\beta - t_\beta G_\alpha(M) = f_{\alpha\beta}{}^{\gamma} t_\gamma. \quad (4.5)$$

Contracting this with G^β and using (4.3), we get

$$f_{\alpha\beta\gamma}t^\beta G^\gamma(M) = \tfrac{1}{2}[\gamma(M) - \gamma(N) - 1]t_\alpha. \quad (4.6)$$

In particular, with $M = N$, this simplifies to

$$f_{\alpha\beta\gamma}t^\beta G^\gamma = -\tfrac{1}{2}t_\alpha. \quad (4.7)$$

Specializing even more by setting $t^\beta = G^\beta(M)$, we obtain

$$f_{\alpha\beta\gamma}G^\beta(M)G^\gamma(M) = -\tfrac{1}{2}G_\alpha(M). \quad (4.8)$$

We now derive Cutkosky's identity that is at the root of algorithms described here and in Refs. 1, 2, and 3. Let N and M be L modules and let

$$N \otimes M = \sum_{\lambda \in \Lambda} M(\lambda) \quad (4.9)$$

be a direct sum decomposition (Clebsch–Gordan series) of their tensor product in terms of irreducible submodules $M(\lambda)$. Let $\pi_\lambda : N \otimes M \to M(\lambda)$ be the canonical projections and let $i_\lambda : M(\lambda) \to N \otimes M$ be the natural injections. The maps π_λ and i_λ are scalar operators satisfying

$$\sum_{\lambda \in \Lambda} i_\lambda \pi_\lambda = 1_{N \otimes M}, \quad (4.10)$$

and $i_\lambda \pi_\lambda$ is idempotent for each $\lambda \in \Lambda$. Since π_λ is a scalar operator, (2.15) gives

$$G_\alpha(\lambda)\pi_\lambda = \pi_\lambda(G_\alpha(N) \otimes 1_M + 1_N \otimes G_\alpha(M)), \quad (4.11)$$

where the 1's denote the unit operators and we have written $G_\alpha(\lambda)$ for $G_\alpha(M(\lambda))$. A similar equation holds for $G^\alpha(\lambda)\pi_\lambda$ which, on contracting with $G_\alpha(\lambda)$ and using (4.11) gives us

$$\begin{aligned} \gamma(\lambda)\pi_\lambda = \pi_\lambda(&\gamma(N)1_N \otimes 1_M + 2G^\alpha(N) \otimes G_\alpha(M) \\ &+ \gamma(M)1_N \otimes 1_M), \end{aligned} \quad (4.12)$$

where we have written $\gamma(\lambda)$ for $\gamma(M(\lambda))$. Multiplying (4.12) on the left by i_λ and summing over λ we obtain Cutkosky's identity

$$G^\alpha(N) \otimes G_\alpha(M) = \sum_{\lambda \in \Lambda} \tfrac{1}{2}[\gamma(\lambda) - \gamma(M) - \gamma(N)]i_\lambda \pi_\lambda. \quad (4.13)$$

5. COMPUTATIONS IN SIMPLE LIE ALGEBRAS

From now on we shall restrict L to be a simple Lie algebra and F will denote the fundamental module. In that case, every nontrivial module is faithful and the trace form in any module is proportional to the Killing form.[17] More explicitly,

$$\mathrm{Tr}_N(G_\alpha(N)G_\beta(N)) = \frac{\dim N \gamma(N)}{\dim L} g_{\alpha\beta}. \quad (5.1)$$

Equation (5.1) can be used to write the eigenvalues of invariants as the trace of a scalar operator in $F \otimes M$ which is defined below. In an irreducible module M,

$$U_{2k}(M) = u_{2k}(M)1_M, \quad (5.2)$$

where $u_{2k}(M)$ is the eigenvalue of $U_{2k}(M)$ in M. Therefore, writing $V_\alpha^{(2k-1)} = v_\alpha^{(2k-1)}(M)$ and $M_\alpha = G_\alpha(M)$,

$$\begin{aligned} u_{2k}(M) &= \frac{1}{\dim M}\mathrm{Tr}_M U_{2k}(M) \\ &= \frac{1}{\dim M}\mathrm{Tr}_M(V_\delta^{(2k-1)}M^\delta) \\ &= \frac{1}{\dim M}\mathrm{Tr}_M(g_{\alpha\beta}V^\alpha_{(2k-1)}M^\beta). \end{aligned} \quad (5.3)$$

Using (5.1), this becomes

$$u_{2k}(M) = \frac{\dim L}{\dim M \dim F\, \gamma(F)}\mathrm{Tr}_{F \otimes M}(Q_{2k-1}Q_1), \quad (5.4)$$

where $Q_{2k-1} = F_\alpha \otimes V^\alpha_{(2k-1)}$. It is shown in Appendix A that the Q_{2k-1} are scalar linear operators belonging to $\mathrm{Hom}_L(F \otimes M, F \otimes M)$.

To calculate the trace appearing in (5.4), first note that because of (4.10) we can write

$$\mathrm{Tr}_{F \otimes M}(Q_s Q_t) = \mathrm{Tr}_{F \otimes M}\sum_{\lambda \in \Lambda}\sum_{\lambda' \in \Lambda} Q_s i_\lambda \pi_\lambda Q_t i_{\lambda'} \pi_{\lambda'}. \quad (5.5)$$

Here $\pi_\lambda Q_t i_{\lambda'} \in \mathrm{lin}(M(\lambda'), M(\lambda))$ is a scalar operator between irreducible modules $M(\lambda)$ and $M(\lambda')$. Therefore, by Schur's lemma, it is equal to 0 if $\lambda \neq \lambda'$ [note that if F is the fundamental module, then all $M(\lambda)$, $\lambda \in \Lambda$ are nonisomorphic] and a multiple $q_t(\lambda)$ of the unit operator 1_λ when $\lambda = \lambda'$. Similar remarks apply to Q_s and since the trace of the idempotent operator $i_\lambda \pi_\lambda$ equals $\dim(\lambda)$, we get

$$\mathrm{Tr}_{F \otimes M}(Q_s Q_t) = \sum_{\lambda \in \Lambda}\dim(\lambda)q_s(\lambda)q_t(\lambda), \quad (5.6)$$

where $\dim(\lambda) = \dim(M(\lambda))$ and $q_s(\lambda) = q_s(M(\lambda))$.

Clearly the same technique can be applied to an arbitrary number of factors giving

TABLE I

Quantity	$A_r = \mathrm{su}(n,\mathbb{C})$	B_r or $D_r = \mathrm{so}(n,\mathbb{C})$	$C_r = \mathrm{sp}(n,\mathbb{C})$
Highest weight of fundamental module	ω_1	ω_1	ω_1
$\dim F = n$	$r+1$	$2r+1$ (B_r) $2r$ (D_r)	$n = 2r$
$\dim L$	n^2-1	$\frac{n(n-1)}{2}$	$\frac{n(n+1)}{2}$
Highest weight of adjoint rep. module	$\omega_1+\omega_2$	ω_2	$2\omega_1$
$\gamma(F)$	$\frac{n^2-1}{2n}$	$\frac{n-1}{2(n-2)}$	$\frac{n+1}{2(n+1)}$
$\gamma(\omega_2)$	$\frac{(n-2)(n+1)}{n^2}$	1	$\frac{n}{n+2}$
$\gamma(2\omega_1)$	$\frac{(n+2)(n-1)}{n^2}$	$\frac{n}{n-2}$	1
$\mathrm{Tr}_F(F_\alpha F_\beta)$	$\frac{1}{2n} g_{\alpha\beta}$	$\frac{1}{n-2} g_{\alpha\beta}$	$\frac{1}{n+2} g_{\alpha\beta}$
$(F_{ab})_{ij}$	$\delta_{ai}\delta_{bj} - \frac{1}{n}\delta_{ab}\delta_{ij}$	$\delta_{ai}\delta_{bj} - \delta_{aj}\delta_{bi}$	$\delta_{ai}E_{bj} + \delta_{bi}E_{aj}$
$g_{(ab)(cd)}$	$2n\left(\delta_{ad}\delta_{bc} - \frac{1}{n}\delta_{ab}\delta_{cd}\right)$	$2(n-2)(\delta_{ad}\delta_{bc} - \delta_{ac}\delta_{bd})$	$-2(n+2)(E_{ac}E_{bd} + E_{bc}E_{ad})$
$g^{(ab)(cd)}$	$\frac{1}{2n}\left(\delta_{ad}\delta_{bc} - \frac{1}{n}\delta_{ab}\delta_{cd}\right)$	$\frac{1}{8(n-2)}(\delta_{ad}\delta_{bc} - \delta_{ac}\delta_{bd})$	$\frac{1}{8(n+2)}(E_{ac}E_{bd} + E_{bc}E_{ad})$
$(F^{ab})_{ij}$	$\frac{1}{2n}(\delta_{ai}\delta_{bj} - \frac{1}{n}\delta_{ab}\delta_{ij})$	$\frac{1}{4(n-2)}(\delta_{ai}\delta_{bj} - \delta_{aj}\delta_{bi})$	$\frac{1}{4(n+2)}(E_{ai}\delta_{bj} + E_{bi}\delta_{aj})$
$(F_\alpha)_{ij}(F^\alpha)_{kl}$ $= (F_{ab})_{ij}(F^{ba})_{kl}$	$\frac{1}{2n}\left(\delta_{il}\delta_{kj} - \frac{1}{n}\delta_{ij}\delta_{kl}\right)$	$\frac{1}{2(n-2)}(\delta_{il}\delta_{jk} - \delta_{ik}\delta_{jl})$	$\frac{1}{2(n+2)}(\delta_{il}\delta_{jk} + E_{ik}E_{jl})$
$u_{\alpha\beta\gamma\delta}F^\delta$	$\frac{1}{2n}\Big(F_\alpha F_\beta F_\gamma$ $- \frac{1}{n}\mathrm{Tr}_F(F_\alpha F_\beta F_\gamma)1_F\Big)$	$\frac{1}{2(n-2)}(F_\alpha F_\beta F_\gamma + F_\gamma F_\beta F_\alpha)$	$\frac{1}{2(n+2)}(F_\alpha F_\beta F_\gamma + F_\gamma F_\beta F_\alpha)$

$$\mathrm{Tr}_{F\otimes M}(Q_{s_1}^{n_1}\cdots Q_{s_p}^{n_p}) = \sum_{\lambda\in\Lambda} \dim(\lambda) q_{s_1}^{n_1}(\lambda)\cdots q_{s_p}^{n_p}(\lambda). \quad (5.7)$$

Applying (5.6) to (5.4), we get

$$u_{2k}(M) = \frac{\dim L}{\dim F\gamma(F)\dim M}\sum_{\lambda\in\Lambda}\dim(\lambda) q_{2k-1}(\lambda) q_1(\lambda). \quad (5.8)$$

The rest of the paper is devoted to developing a recursion formula for $q_{2k-1}(\lambda)$. Note that since $Q_1 = F_\alpha \otimes M^\alpha$, the value of $q_1(\lambda)$ is already available from (4.13) as

$$q_1(\lambda) = \tfrac{1}{2}[\gamma(\lambda) - \gamma(M) - \gamma(F)]. \quad (5.9)$$

It will be necessary for us to reverse the order of the first three factors in a product of the kind $F_\gamma F_\beta F_\alpha \otimes V^\alpha M^\beta M^\gamma$, where V^α is a vector operator in M. Such reversal is also necessary if one is interested in computing the invariants defined by means of the symmetrized coefficients $s_{\alpha\beta\gamma\delta}$ as given by Micu.[4] The main technique is repeated use of the commutation relations (4.1). To begin with, (4.1) and (4.8) give

$$F_\beta F_\alpha \otimes V^\alpha M^\beta = F_\alpha F_\beta \otimes V^\alpha M^\beta + \tfrac{1}{2}F_\alpha \otimes V^\alpha. \quad (5.10)$$

Using (5.10), (4.1), and then (5.10) again, we get

$$F_\gamma F_\beta F_\alpha \otimes V^\alpha M^\beta M^\gamma = F_\gamma F_\alpha F_\beta \otimes V^\alpha M^\beta M^\gamma + \tfrac{1}{2}F_\alpha F_\beta \otimes V^\alpha M^\beta + \tfrac{1}{4}F_\alpha \otimes V^\alpha. \quad (5.11)$$

The first term on the right-hand side can be simplified in the same manner and the result subsituted in (5.11) gives

$$F_\gamma F_\beta F_\alpha \otimes V^\alpha M^\beta M^\gamma = F_\alpha F_\beta F_\gamma \otimes V^\alpha M^\beta M^\gamma + [\gamma(F) + \tfrac{1}{2}]F_\alpha F_\beta \otimes V^\alpha M^\beta + \tfrac{1}{4}F_\alpha \otimes V^\alpha - F_\sigma F_\alpha F_\beta F^\sigma \otimes V^\alpha M^\beta. \quad (5.12)$$

Reduction of the last term is somewhat indirect as follows.

Let $F\otimes F^* = \Sigma_{i\in I} T(i)$ be a decomposition as a direct sum of irreducible submodules $T(i)$. The linear operator $F_\alpha F_\beta$ belongs to $\mathrm{lin}(F,F) \cong F\otimes F^*$ and hence can be written uniquely as a sum of irreducible tensor operators,

$$F_\alpha F_\beta = \sum_{i\in I} T_{\alpha\beta}(i), \quad (5.13)$$

where $T_{\alpha\beta}(i)$ belongs to $T(i)$ for each $i\in I$. Applying (4.3) we get

$$F_\sigma F_\alpha F_\beta F^\sigma = \sum_{i\in I} F_\sigma T_{\alpha\beta}(i) F^\sigma$$

$$= \sum_{i\in I} [\gamma(F) - \tfrac{1}{2}\gamma(T(i))] T_{\alpha\beta}(i)$$

$$= \gamma(F) F_\alpha F_\beta - \tfrac{1}{2} \sum_{i\in I} \gamma(T(i)) T_{\alpha\beta}(i). \tag{5.14}$$

Substitution of this into Eq. (5.12) gives

$$\begin{aligned} F_\gamma F_\beta F_\alpha \otimes V^\alpha M^\beta M^\gamma &= F_\alpha F_\beta F_\gamma \otimes V^\alpha M^\beta M^\gamma + \tfrac{1}{2} F_\alpha F_\beta \otimes V^\alpha M^\beta \\ &\quad + \tfrac{1}{4} F_\alpha \otimes V^\alpha + \tfrac{1}{2} \sum_{i\in I} \gamma(T(i)) T_{\alpha\beta}(i) \otimes V^\alpha G^\beta. \end{aligned} \tag{5.15}$$

For the rest of the computations, we have to work in a specific Lie algebra employing a special basis for the fundamental representation. The elements F_α are labeled as F_{ab} by the double index $\alpha \sim (a,b)$, where all latin indices run from 1 through $n = \dim F$. Explicit matrix elements, the metrics $g_{(ab)(cd)}$ and $g^{(ab)(cd)}$ together with other pertinent information are collected in Table I for the Lie algebras A_r, B_r, C_r, and D_r. These data are used to compute expressions for $(F_{ab})_{ij}(F^{ba})_{kl}$ and $u_{\alpha\beta\gamma\delta}F^\delta$ which are also given there. These will give the needed recursion formula because

$$Q_{2k-1} = F^\delta \otimes V^{(2k-1)}_\delta = u_{\alpha\beta\gamma\delta} F^\delta \otimes V^\alpha_{(2k-3)} M^\beta M^\gamma, \tag{5.16}$$

and the (k,l) element of $u_{\alpha\beta\gamma\delta}F^\delta$ is $(F_\alpha F_\beta F_\gamma)_{ji}(F_\delta)_{ij}(F^\delta)_{kl}$.

The numbering of the simple roots used by us is the same as in Ref. 5.

6. LIE ALGEBRA A_r

This is the Lie algebra of the special (det = 1) unitary group SU(n) and will also be denoted by su(n,ℂ), where $n = r + 1$. The Clebsch–Gordan series for $F \otimes M$ can be obtained easily from the Young diagram techniques.[18,19] Here we present the result in terms of Dynkin indices:

$$\begin{aligned} (1,0,\ldots,0) &\otimes (m_1,m_2,\ldots,m_r) \\ &= \omega_1 \otimes \mu \\ &= (\mu + \omega_1) \oplus (\mu - \omega_1 + \omega_2) \oplus \cdots \otimes (\mu - \omega_{r-1} + \omega_r) \\ &\quad \oplus (\mu - \omega_r). \end{aligned} \tag{6.1}$$

Now Dynkin indices of highest weights must be nonnegative numbers; hence any term on the right-hand side of (6.1) that contains a negative Dynkin index must be omitted.

The fundamental representation of $L = \mathrm{su}(n,\mathbb{C})$ consists of all $n \times n$ complex matrices of trace zero and so any $\dim L = n^2 - 1$ traceless matrices which are linearly independent will constitute a basis of L in the fundamental representation. One such choice is given in Table I. Substituting $u_{\alpha\beta\gamma\delta}F^\delta$ and other information from Table I into Eq. (5.15) we obtain the following recursion relation,

$$Q_{2k-1} = \frac{1}{2n}\left(Q_{2k-3}Q_1^2 - \frac{1}{n\dim M}(\mathrm{Tr}_{F\otimes M} Q_{2k-3}Q_1^2) 1_F \otimes 1_M\right). \tag{6.2}$$

Projecting this onto the submodule $M(\lambda) \subset F \otimes M$ and substituting for the trace from (5.6), we get

$$q_{2k-1}(\lambda) = \frac{1}{2n}\left(q_{2k-3}(\lambda) q_1^2(\lambda) - \frac{1}{n\dim M}\sum_{\lambda\in\Lambda} \dim(\lambda) \times q_{2k-3}(\lambda) q_1^2(\lambda) 1_{F\otimes M}\right), \tag{6.3}$$

where the summation is over all submodules $M(\lambda)$, $\lambda \in \Lambda$, in the Clebsch–Gordan series (4.9) for $F \otimes M$.

The value of $q_1(\lambda)$ is given by (5.9) so that (6.3) allows us to calculate successively the values $q_3(\lambda)$, $q_5(\lambda)$, etc. With these and the values $\dim L = n^2 - 1$, $\dim F = n$, $\gamma(F) = (n^2 - 1)/(2n)$, (5.8) reduces finally to

$$u_{2k}(M) = \frac{n-2}{2n} \frac{1}{\dim M} \sum_{\lambda\in\Lambda} \dim(\lambda) q_{2k-3}(\lambda) q_1^3(\lambda). \tag{6.4}$$

7. LIE ALGEBRAS B_r AND D_r

These are the simple Lie algebras of the special orthogonal groups SO(n), where $n = 2r + 1$ (B_r) or $n = 2r$ (D_r). For most part, these algebras can be treated together and so we shall denote them collectively by so(n,ℂ).

The Clebsch–Gordan series for $F \otimes M$ was calculated by Murnaghan[20] for the orthogonal groups O(n). For the group SO(n), his formulas have to be slightly modified as described by Okubo.[2] Their formulas translated in terms of Dynkin indices give for B_r,

$$\begin{aligned} \omega_1 \otimes \mu &= (1,0,\ldots,0) \otimes (m_1,\ldots,m_r) \\ &= (\mu + \omega_1) \oplus (\mu - \omega_1 + \omega_2) \oplus \cdots \oplus (\mu - \omega_{r-2} + \omega_{r-1}) \\ &\quad \oplus (\mu - \omega_{r-1} + 2\omega_r) \oplus (\mu - \omega_1) \oplus (\mu + \omega_1 - \omega_2) \oplus \cdots \\ &\quad \oplus (\mu + \omega_{r-2} - \omega_{r-1}) \oplus (\mu + \omega_{r-1} - 2\omega_r) \oplus (m_1,\ldots,m_r \\ &\quad - \delta_{0,m_r}). \end{aligned} \tag{7.1}$$

Any term on the right-hand side in which a negative Dynkin index occurs, must be omitted. Moreover, the last term is just μ if $m_r \neq 0$, but is omitted if $m_r = 0$.

For D_r we have

$$\begin{aligned} \omega_1 \otimes \mu &= (1,0,\ldots,0) \otimes (m_1,\ldots,m_r) \\ &= (\mu + \omega_1) \oplus (\mu - \omega_1 + \omega_2) \oplus \cdots \oplus (\mu - \omega_{r-2} + \omega_{r-1} + \omega_r) \\ &\quad \oplus (\mu - \omega_{r-1} + \omega_r) \oplus (\mu - \omega_1) \oplus (\mu + \omega_1 - \omega_2) \oplus \cdots \end{aligned}$$

$$\oplus(\mu+\omega_{r-2}-\omega_{r-1}-\omega_r)\oplus(\mu+\omega_{r-1}-\omega_r)$$

$$\oplus(m_1,\ldots,m_r-2m_r\delta_{m_r m_{r-1}}). \tag{7.2}$$

Again, all terms with any negative Dynkin index have to be excluded and the last term is simply μ if $m_r \neq m_{r-1}$ but must be omitted if $m_r = m_{r-1}$.

The fundamental representation of so(n,$\mathbb{C}$) consists of all $n\times n$ skew matrices. A particular basis choice is shown in Table I in which we have also gathered some other relevant information. Using the reduction of $u_{\alpha\beta\gamma\delta}F^\delta$ given there, we find

$$Q_{2k-1}=\frac{1}{2(n-2)}(F_\alpha F_\beta F_\gamma+F_\gamma F_\beta F_\alpha)\otimes V^\alpha_{(2k-3)}G^\beta G^\gamma. \tag{7.3}$$

To reverse the order of factors in $F_\gamma F_\beta F_\alpha$, we use (5.12) which needs a simplication of $F_\alpha F_\beta F_\gamma F^\alpha$. This is achieved by applying formula (5.14) in the special case of so(n,$\mathbb{C}$). For both B_r and D_r, $F\otimes F=\omega_1\otimes\omega_1$ decomposes into irreducible modules with highest weights 0, ω_2, and $2\omega_1$. This corresponds to the decomposition

$$F_\beta F_\gamma=T_{\beta\gamma}(0)+T_{\beta\gamma}(\omega_2)+T_{\beta\gamma}(2\omega_1), \tag{7.4}$$

where the

$$T_{\beta\gamma}(0)=\frac{1}{n(n-2)}g_{\beta\gamma}1_F \tag{7.5}$$

are multiples of the unit operator,

$$T_{\beta\gamma}(\omega_2)=\tfrac{1}{2}(F_\beta F_\gamma-F_\gamma F_\beta) \tag{7.6}$$

are skew linear operators, and

$$T_{\beta\gamma}(2\omega_1)=\tfrac{1}{2}(F_\beta F_\gamma+F_\gamma F_\beta)-\frac{1}{n(n-2)}g_{\beta\gamma}1_F \tag{7.7}$$

are traceless symmetric linear operators. Using these explicit tensors and supplying the values of the Casimir operators, we get

$$F_\alpha F_\beta F_\gamma F^\alpha=\frac{1}{2(n-2)}\times\left(-F_\beta F_\gamma+f_{\beta\gamma\sigma}F^\sigma+\frac{1}{(n-2)}g_{\beta\gamma}1_F\right). \tag{7.8}$$

Substitution of this in (5.12) followed by some algebra gives

$$F_\gamma F_\beta F_\alpha\otimes V^\alpha M^\beta M^\gamma$$
$$=F_\alpha F_\beta F_\gamma\otimes V^\alpha M^\beta M^\gamma+\frac{n-1}{n-2}F_\alpha F_\beta\otimes V^\alpha M^\beta$$
$$+\frac{n-1}{4(n-2)}F_\alpha\otimes V^\alpha-\frac{1}{2(n-2)^2}1_F\otimes V^\alpha M_\alpha.$$

Finally, this used in (7.3) gives

$$q_{2k-1}(\lambda)=\frac{1}{2(n-2)}\left(2q_{2k-3}(\lambda)q_1^2(\lambda)+\frac{n-1}{n-2}q_{2k-3}(\lambda)q_1(\lambda)+\frac{n-1}{4(n-2)}q_{2k-3}(\lambda)-\frac{u_{2k-2}(M)}{2(n-2)^2}1_\lambda\right).$$

With $q_1(\lambda)$ given by (5.9), this allows successive computation of $q_3(\lambda)$, $q_5(\lambda)$, etc., which then can be used in (6.4) to determine $u_4(M)$, $u_6(M)$, etc.

8. LIE ALGEBRA C_r

The simple Lie algebra $L=C_r$ is the (complex) Lie algebra of the symplectic group Sp(n), $n=2r$ and is often denoted by sp(n,$\mathbb{C}$). The Clebsch–Gordan series for $F\otimes M$ is given in Ref. 2 which, written in terms of the Dynkin indices is

$$\omega_1\otimes\mu=(\mu+\omega_1)\oplus(\mu-\omega_1+\omega_2)\oplus\cdots\oplus(\mu-\omega_{r-1}+\omega_r)$$
$$\oplus(\mu-\omega_1)\oplus(\mu+\omega_1-\omega_2)\oplus\cdots$$
$$\oplus(\mu+\omega_{r-1}-\omega_r). \tag{8.1}$$

As usual, any term in which a negative Dynkin index occurs must be omitted from (8.1).

The fundamental module F is of dimension $n=2r$ and carries a nondegenerate alternating bilinear form. The matrix E of this bilinear form can be chosen to be

$$E=\left(\begin{array}{c|c}0 & 1\\ \hline -1 & 0\end{array}\right), \tag{8.2}$$

where each block is a matrix of order r. The fundamental representation of L consists of all complex $n\times n$ matrices X satisfying

$$EXE=X^T. \tag{8.3}$$

Any $[(n(n+1)/2]$ such matrices will form a basis of L in this representation. One such choice is shown in Table I. Substituting $u_{\alpha\beta\gamma\delta}F^\delta$ from Table I into Eq. (5.15), we get

$$Q_{2k-1}=\frac{1}{2(n+1)}(F_\alpha F_\beta F_\gamma+F_\beta F_\gamma F_\alpha)\otimes V^\alpha_{(2k-3)}G^\beta G^\gamma. \tag{8.4}$$

Since $F\otimes F$ decomposes into irreducible modules with highest weights 0, ω_2, $2\omega_1$, we can write

$$F_\alpha F_\beta=T_{\alpha\beta}(0)+T_{\alpha\beta}(\omega_2)+T(2\omega_1). \tag{8.5}$$

The tensor operators

$$T_{\alpha\beta}(0)=\frac{1}{n(n+2)}g_{\alpha\beta}1_F \tag{8.6}$$

are multiples of the unit operator,

$$T_{\alpha\beta}(\omega_2)=\tfrac{1}{2}(F_\alpha F_\beta+F_\beta F_\alpha)-\frac{1}{n(n+2)}g_{\alpha\beta}1 \tag{8.7}$$

are traceless symplectic-symmetric operators, and

$$T_{\alpha\beta}(2\omega_1)=\tfrac{1}{2}(F_\alpha F_\beta-F_\beta F_\alpha) \tag{8.8}$$

are skew-symplectic operators. Using these explicit tensors in (5.14) and simplifying, we get

$$F_\sigma F_\alpha F_\beta F^\sigma = \frac{1}{2(n+2)}\left(F_\alpha F_\beta - f_{\alpha\beta\gamma}F^\gamma + \frac{1}{n+2}g_{\alpha\beta}1_F\right). \tag{8.9}$$

This, when substituted in (5.12) gives

$$F_\gamma F_\beta F_\alpha \otimes V^\alpha M^\beta M^\gamma$$
$$= F_\alpha F_\beta F_\gamma \otimes V^\alpha M^\beta M^\gamma + \frac{n+1}{n+2}F_\alpha F_\beta \otimes V^\alpha M^\beta$$
$$+ \frac{n+1}{4(n+2)}F_\alpha \otimes V^\alpha - \frac{1}{2(n+2)^2}1_F \otimes V^\alpha M_\alpha. \tag{8.10}$$

Finally, this used in (8.4) gives the recursion formula

$$q_{2k-1}(\lambda)$$
$$= \frac{1}{2(n+2)}\left(2q_{2k-3}(\lambda)q_1^2(\lambda) + \frac{n+1}{n+2}q_{2k-3}(\lambda)q_1(\lambda)\right.$$
$$\left.+ \frac{n+1}{4(n+2)}q_{2k-3}(\lambda) - \frac{u_{2k}(M)}{2(n+2)^2}1_\lambda\right). \tag{8.11}$$

ACKNOWLEDGMENT

The author wishes to thank Professor C.Y. Chao for helpful discussions and a critical reading of the manuscript.

APPENDIX A

Here we show that the elements $v_\delta^{(2k-1)}$ $(k = 1,2,\cdots)$ of U defined in (3.7) are linearly independent and satisfy the same commutation relations as the basis elements G_α of L. As a corollary, it follows at once that the elements U_{2k} $(k = 1,2,\cdots)$ of U defined in (3.8) are indeed invariants and also that the operators $Q_{2k-1} = F_\delta \otimes V^\delta_{(2k-1)}$, where $V^\delta_{(2k-1)} = v^\delta_{(2k-1)}(M)$, are scalar operators acting in $F \otimes M$. the proofs are broken down in the following lemmas. The notation is the same as in the rest of the paper, that is, F is a faithful module, M an arbitrary module, G_α $(\alpha = 1,\ldots,\dim L)$ constitute a basis of L which are represented as linear operators $G_\alpha(M)$ in M. Greek indices run from 1 through $\dim L$ and summation over repeated indices is implied. We also abbreviate $F_\alpha = G_\alpha(F)$ and $M_\alpha = G_\alpha(M)$.

Lemma 1: If $u_{\alpha\beta\gamma\delta} = \mathrm{Tr}_F(F_\alpha F_\beta F_\gamma F_\delta)$, then for all complex numbers c^δ satisfying $c^\delta u_{\alpha\beta\gamma\delta} = 0$, we must have $c^\delta = 0$ for each δ.

Proof: Since L is semisimple and F is faithful, the trace form in the representation F is nondegenerate. If for all α, β, γ, we have $0 = c^\delta u_{\alpha\beta\gamma\delta} = \mathrm{Tr}_F(F_\alpha F_\beta F_\gamma c^\delta F_\delta)$, then by the nondegenerate nature of the trace form, it follows that $c^\delta F_\delta = 0$, which implies each $c^\delta = 0$ by linear independence of F_δ.

Lemma 2: For each $k = 1,2,\ldots$, the elements $v_\delta^{(2k-1)}$ defined in (3.7) are linearly independent.

Proof: $V_\delta^{(1)} = G_\alpha$ are linearly independent because they form a basis of L. Now let $c^1,\ldots,c^{\dim L}$ be complex numbers such that

$$c^\delta v_\delta^{(3)} = c^\delta u_{\alpha\beta\gamma\delta}G^\alpha G^\beta G^\gamma = 0.$$

By the Poincaré–Birkhoff–Witt theorem,[21] the set of $G^\alpha G^\beta G^\gamma$ is linearly independent as α, β, γ vary from 1 through $\dim L$. Hence $c^\delta u_{\alpha\beta\gamma\delta} = 0$ for all α, β, and γ which, by Lemma 1, implies $c^\delta = 0$ for each δ. Similarly $c^\delta v_\delta^{(5)} = 0$ implies $c^\delta(u_{\alpha\beta\gamma\delta}v^\alpha_{(3)}G^\beta G^\gamma) = 0$; that is, $c^\delta u_{\alpha\beta\gamma\delta}u_{\mu\nu\rho\alpha} \times G^\mu G^\nu G^f G^\beta G^\gamma = 0$ which gives $c^\delta u_{\alpha\nu\gamma\delta}u_{\mu\nu\rho\alpha} = 0$ and hence, by Lemma 1, $c^\delta u_{\alpha\beta\gamma\delta} = 0$ so that again by Lemma 1, $c^\delta = 0$ for each δ. Although the notation becomes cumbersome, it is clear that this procedure can be continued for higher order v_δ's and so the result follows by induction.

Lemma 3: For each $k = 1,2,\cdots$ the $v_\delta^{(2k-1)}$ satisfy the commutation relations

$$G_\alpha v_\beta^{(2k-1)} - v_\beta^{(2k-1)}G_\alpha = f_{\alpha\beta}{}^{\gamma}_{\cdot}v_\gamma^{(2k-1)}.$$

Proof: We use induction on k. For $k = 1$, the result is obvious because $v_\delta^{(1)} = G_\delta$. Now assume that the result is true for all positive integers less than k.

Then

$$G_\alpha v_\beta^{(2k-1)} - v_\beta^{(2k-1)}G_\alpha$$
$$= u_{\mu\nu\rho\beta}(G_\alpha v^\mu_{(2k-3)}G^\nu G^\rho - v^\mu_{(2k-3)}G^\nu G^\rho G_\alpha).$$

Repeated use of commutation relations converts the right-hand side to

$$\mathrm{Tr}_F(F_\mu F_\nu F_\rho F_\beta[f^\rho_{\cdot\sigma\alpha}v^\mu_{(2k-3)}G^\nu G^\sigma + f^\nu_{\cdot\sigma\alpha}v^\mu_{(2k-3)}G^\sigma G^\rho$$
$$+ f^\mu_{\cdot\sigma\alpha}v^\sigma G^\nu G^\rho].$$

To simplify this, we associate the f coefficients with the F operators and use the commutation relations for the F operators backwards. Further repeated use of commutation relations (among the F's) to rearrange the order of factors reduces this to $f_{\alpha\beta}{}^{\gamma}_{\cdot}v_\mu^{(2k-1)}$.

Lemma 4: For each $k = 1,2,\cdots$, the element U_{2k} defined in (3.8) commutes with every element of L.

Proof: It is enough to show that U_{2k} commutes with every basis element of L. Using the commutation relations twice we get

$$G_\alpha U_{2k} = G_\alpha v_\beta G^\beta = v_\beta G_\alpha G^\beta - f_{\beta\alpha}{}^{\sigma}_{\cdot}v_\sigma G^\beta$$
$$= v_\beta G^\beta G_\alpha - f^\beta_{\cdot\alpha\sigma}v_\beta G^\sigma - f_{\sigma\alpha}{}^{\beta}_{\cdot}v_\beta G^\sigma$$
$$= U_{2k}G_\alpha.$$

Lemma 5: For each $k = 1,2,\cdots$, $Q_{2k-1} = F_\beta \otimes V^\beta_{(2k-1)}$ is a scalar operator in $F \otimes M$, where we have written $V^\beta_{(2k-1)} = v^\beta_{(2k-1)}(M)$.

Proof: Again, it suffices to prove that Q_{2k-1} commutes with every basis element of L represented in $F \otimes M$, that is, with $F_\alpha \otimes 1_M + 1_F \otimes M_\alpha$. Now

$$(F_\alpha \otimes 1_M + 1_F \otimes M_\alpha)$$
$$\times (F_\beta \otimes V^\beta_{(2k-1)}) - F_\beta \otimes V^\beta_{(2k-1)}(F_\alpha \otimes 1_M + 1_F \otimes M_\alpha)$$
$$= (F_\alpha F_\beta - F_\beta F_\alpha) \otimes V^\beta_{(2k-1)} + F_\beta \otimes (M_\alpha V^\beta_{(2k-1)}$$
$$- V^\beta_{(2k-1)}M_\alpha)$$

$$
\begin{aligned}
&= f_{\alpha\beta\sigma} F^{\sigma} \otimes V^{\beta}_{(2k-1)} + f_{\alpha\sigma\beta} F^{\sigma} \otimes V^{\beta}_{(2k-1)} \\
&= 0.
\end{aligned}
$$

[1]V.K. Agrawala and J.G. Belinfante, BIT **11**, 1 (1971).

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

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

[4]M. Micu, Nucl. Phys. **60**, 353 (1964).

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

[6]H. Freudenthal and H. de Vries, *Linear Lie Groups* (Academic, New York, 1969).

[7]J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, New York, 1972).

[8]J.F. Adams, *Lectures on Lie Groups* (Benjamin, New York, 1962).

[9]L.S. Pontryagin, *Topological Groups* (Gordan and Breach, New York, 1966).

[10]U. Fano and G. Racah, *Irreducible Tensorial Sets* (Acadamic, New York, 1959).

[11]V.K. Agrawala and J.G. Belinfante, BIT **9**, 301 (1969).

[12]See Ref. 6, p. 211.

[13]C. Chevalley, Amer. J. Math. **77**, 778 (1955).

[14]Harish-Chandra, Ann. Math. (N.Y.) (2) **50**, 900 (1949).

[15]B. Gruber and L. O'Raifeartaigh, J. Math. Phys. **5**, 1796 (1964).

[16]I learned of this identity from Prof. R.E. Cutkosky who pointed out its similarity with the well-known expression for **L·S** in the angular momentum theory. However, a rigorous proof seems to have appeared for the first time in Ref. 1.

[17]Ref. 7, p. 118.

[18]C. Itzykson and M. Nauenberg, Rev. Mod. Phys. **38**, 95 (1966).

[19]A. Pais, Rev. Mod. Phys. **38**, 215 (1966).

[20]F.D. Murnaghan, *The Theory of Group Representations* (Dover, New York, 1963), p. 278.

[21]See, for example, Ref. 5, p. 159.

Quantum dynamics of Hamiltonians perturbed by pulses

Mark J. Christensen and John Piepenbrink
School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 27 February 1979; accepted for publication 15 June 1979)

There has been a recent controversy between Gzyl [J. Math Phys. **18**, 1327 (1977)] and Blume [J. Math. Phys. **19**, 2004 (1978)] concerning the correct way to integrate the equation $i\,\partial\psi/\partial t = H_0\psi + \delta(t)V\psi$ where δ is the Dirac delta function. In this paper we: (1) Suggest some physical reasons for rejecting Gzyl's procedure and (2) Extend the limiting procedure suggested the Blume to unbounded Hamiltonians.

1. INTRODUCTION

Various authors have recently considered quantum mechanical systems having Hamiltonians of the form:

$$H = H_0 + \delta V, \tag{1.1}$$

where H_0 is the time-independent background Hamiltonian δ is the Dirac delta function, and V is a potential function. Such Hamiltonians can be thought of as representing a system which evolves according to the equation,

$$i\frac{dx}{dt} = H_0 x \tag{1.2}$$

for $t\neq 0$. At $t = 0$, however, a potential pulse of spatial distribution V arrives. Clauser and Blume,[1] as well as Gzyl,[2] have considered Hamiltonians of the slightly more general form,

$$H = H_0 + \sum_i \delta(t - T_i)V_i\,, \tag{1.3}$$

in which T_i, the time of arrival of the ith pulse described by V_i, is a random variable.

Such Hamiltonians were used by Clauser and Blume to describe the line shape of Mössbauer spectra. In the course of their treatment they claim that the correct solution to the equation

$$i\frac{dx}{dt} = H_0 x + \delta(t - T)Vx \tag{1.4}$$

is

$$x(t) = \begin{cases} \exp(-iH_0 t)x(0) & \text{if } 0\leqslant t < T, \\ \exp(-iH_0(t-T))\exp(-iV)x(T-) & \text{if } t > T, \end{cases} \tag{1.5}$$

where

$$x(T-) = \operatorname*{limit}_{t\nearrow T} x(t) \quad\text{and}\quad x(T+) = \operatorname*{limit}_{t\searrow T} x(t)\,.$$

In particular it follows that,

$$x(T+) = \exp(-iV)x(T-)\,. \tag{1.6}$$

Gzyl, however, writes (1.4) in differential form,

$$dx = -iH_0 x\,dt - i\delta(t - T)Vx\,dt\,. \tag{1.7}$$

Integrating this expression and treating $\delta(t - T)$ as a point mass at T he obtains,

$$x(T+) = (I - iV)x(T-)\,. \tag{1.8}$$

Obviously, (1.6) and (1.8) are very different. In order to determine which is in fact physically appropriate we make two remarks. First, imagine that the two shocks are applied to the system. Suppose that at time T_1 the potential pulse V_1 arrives and at time T_2, V_2 arrives, with $T_2 > T_1$. Then the formula of Clauser and Blume would predict that,

$$x_{cb}(T_2+) = \exp(-iV_2)\exp[-iH_0(T_2 - T_1)] \times \exp(-iV_1)x(T_1-)\,, \tag{1.9}$$

while that of Gzyl would lead one to expect that,

$$x_g(T_2) = (I - iV_2)\exp[iH_0(T_2 - T_1)](I - iV_1)x(T_1-)\,. \tag{1.10}$$

If the two shocks were flowed together so that $T_2 \to T_1$, then we see from (1.9) that

$$x_{cb}(T_1+) = \exp(-iV_2)\exp(-iV_1)x(T_1-)\,, \tag{1.11}$$

while from (1.10) we would get

$$\begin{aligned} x_g(T_1+) &= (I - iV_2)(I - iV_1)x(T_1-) \\ &= [I - V_2V_1 - i(V_2 + V_1)]x(T_1-)\,. \end{aligned} \tag{1.12}$$

Now suppose that V_2 and V_1 commute. Then (1.11) becomes,

$$x_{cb}(T_1+) = \exp[-i(V_2 + V_1)]x(T_1-)\,, \tag{1.13}$$

which is exactly what we would expect from two shocks arriving as one. This is obviously not the case in (1.12) since then we would expect,

$$x_g(T_1+) = [I - i(V_2 + V_1)]x(T_1-)\,.$$

A second objection to Gzyl's interpretation of (1.4) is that in practice it is impossible for a potential to be turned on instantaneously. Indeed, if this were to happen it would contradict relativity, since the potential could be a carrier of information. Instead, (1.4) and (1.7) should be thought of as representing a system in which the switching off and on simply occurs very rapidly.

This remark is especially important since it suggests the sense of Eq. (1.4) and how its solution should be obtained. Since the pulse must actually arrive during a very brief interval of time, Eq. (1.4) should be replaced by a sequence of approximating problems;

$$\begin{aligned} i\frac{dx_\epsilon}{dt} &= H_0 x_\epsilon + f_\epsilon(t - T)Vx_\epsilon\,, \\ x_\epsilon(0) &= x(0)\,, \end{aligned} \tag{1.14}$$

where each f_ϵ is nonnegative, continuously differentiable with compact support satisfying

$$\int_{-\infty}^{+\infty} f_\epsilon(s)\,ds = 1 = \underset{n\to 0}{\text{limit}} \int_{T-\delta}^{T+\delta} f_\eta(t-T)\,dt$$

for every $\epsilon > 0$ and $\delta > 0$. Finding the solution to (1.4) in this interpretation is now reduced to solving the problems (1.14) and, in the appropriate topology, taking the limit of the $x_\epsilon(t)$ as ϵ tends to zero.

This is the approach taken by Blume[3] in his discussion of Gzyl's formula (1.8). However in his justification of (1.6) Blume assumes that H_0 can be ignored during the interaction, i.e., when $f_\epsilon \neq 0$. This is appropriate for the Hamiltonians considered by Blume, Clauser, and Gzyl. In general, however, this is a dangerous assumption since H_0 and V will often be unbounded operators, thereby admitting the possibility of a nontrivial interaction between H_0 and V during the limiting process. If, however, H_0 is bounded then it is easy to see that,

$$\lim_{\epsilon\to 0} x_\epsilon(T+) = \exp(-iV)x(T-) \tag{1.15}$$

in the norm topology of the Hilbert space in which x exists. Indeed, Clauser showed this for $H_0 = 0$ in his example.

The purpose of this paper is to show that, in fact, for general classes of H_0 and V, bounded or not, (1.15) holds.

In Sec. 2 we state the precise conditions under which we show (1.15) holds. In Sec. 3 the necessary methematical apparatus is assembled and the result is proved.

2. THE MAIN RESULT

Let $\mathscr{H}$ denote a complex Hilbert space that is separable, i.e., $\mathscr{H}$ has a complete orthonormal sequence. $\langle\phi,\psi\rangle$ is the inner product of vectors ϕ and ψ in $\mathscr{H}$ with $\|\phi\| = \langle\phi,\phi\rangle^{1/2}$ the norm.

Each operator A on $\mathscr{H}$ will be linear and have a dense domain $\mathscr{D}(A)$ in $\mathscr{H}$. A is *closed* if whenever $\phi_n \to \phi$ and $A\phi_n \to \psi$ for a sequence $\{\phi_n\}$ in $\mathscr{D}(A)$, it follows that $\phi\in\mathscr{D}(A)$ and $A\phi = \psi$. A is *closable* if there is a closed operator A_1 with $\mathscr{D}(A_1)\supseteq\mathscr{D}(A)$ and $A_1\phi = A\phi$ for all ϕ in $\mathscr{D}(A)$. It turns out that if A is closable there is a minimal closed operator $\bar{A}$ extending A. $\bar{A}$ is called the *closure* of A.

An operator A is *bounded* if the norm of A

$$\|A\| = \sup_{\|\phi\|=1} \|A\phi\|$$

is finite. A bounded operator always has a unique extension to a bounded operator with domain equal to all of $\mathscr{H}$. The extension has the same norm.

A is *symmetric* if the equation

$$\langle A\phi,\psi\rangle = \langle\phi,A\psi\rangle$$

holds for each pair of vectors ϕ and ψ in $\mathscr{D}(A)$. The adjoint of A, denoted by A^*, is defined for each vector ψ for which there is a vector η satisfying

$$\langle A\phi,\psi\rangle = \langle\phi,\eta\rangle$$

for all $\phi\in\mathscr{D}(A)$. We write $\eta = A^*\psi$. Note that if A is symmetric, A^* is an extension of A. If $A^* = A$ we say that A is *self-adjoint*.

It is well known that if H_0 is a self-adjoint operator, the Schrödinger equation

$$i\frac{d\phi}{dt} = H_0\phi \tag{2.1}$$

generates a unitary group $\exp[-iH_0t]$ of operators so that any solution $\phi(t)$ of (2.1) can be written as

$$\phi(t) = \exp[-iH_0t]\phi(0). \tag{2.2}$$

We must consider two self-adjoint operators H_0 and V and the associated Schrödinger equation

$$i\frac{d\phi}{dt} = [H_0 + f_\epsilon(t)V]\phi\,. \tag{2.3}$$

Here f_ϵ is a smooth approximation of the δ function $\delta(t-1)$. By translation of the time scale we here set T, the shock time, equal to 1. If $f(t)$ is any continuously differentiable function which is non-negative on $(-\infty,\infty)$ and satisfies the conditions

$$f(t) = 0 \ \text{ for } \ t\notin(0,2)$$

and

$$\int_0^2 f(t)\,dt = 1\,,$$

then we set

$$f_\epsilon(t) = (1/\epsilon)f[1 + (t-1)/\epsilon]$$

for $0 < \epsilon \leqslant 1$.

Now by letting $1 + (t-1)/\epsilon$ be the new time variable, which we denote by t, (2.3) becomes

$$i\frac{d\phi}{dt} = [\epsilon H_0 + f(t)V]\phi\,. \tag{2.4}$$

If $\epsilon = 0$, it is easy to show that

$$\phi(t) = \exp\left[-i\int_0^t f(s)\,ds\,V\right]\phi(0) \tag{2.5}$$

is the general solution and hence that the equation

$$\phi(2) = \exp[-iV]\phi(0)$$

gives the result of Blume.[3] The significant problem is to determine whether the solutions of (2.4) for positive ϵ tend to (2.5) as $\epsilon \to 0$. Its significance derives from the fact that, in general, H_0 and V are unbounded.

Since there are two possibly unbounded operators in the present case care must be used in defining the operator $\epsilon H_0 + f(t)V$ on the right side of (2.4). We are basing our result on the recent work of Dollard and Friedman[4] on evolution equations in Banach spaces. We find that the following hypotheses are convenient for our purpose. (For examples of Hamiltonian and potentials satisfying these conditions see Simon.[5])

(i) $\mathscr{D}(H_0)\cap\mathscr{D}(V)$ is dense in $\mathscr{H}$. For each $\xi \geqslant 0$, $H_0 + \xi V$, defined on $\mathscr{D}(H_0)\cap\mathscr{D}(V)$, has a self-adjoint closure, denoted by $H(\xi)$.

(ii) $H(\xi)$ has a bounded inverse $H(\xi)^{-1}$ for each $\xi \geqslant 0$. Furthermore, as a function ξ $H(\xi)^{-1}$ is *weakly bounded* on each finite interval $[0,L]$. This means that for any vectors ϕ and ψ the functions $\langle\phi,H(\xi)^{-1}\psi\rangle$ is bounded on each finite interval $[0,L]$.

(iii) The operator $V[H_0 + \xi V]^{-1}$ is bounded on

$(H_0 + \xi V)[\mathscr{D}(H_0) \cap \mathscr{D}(V)]$. Since (i) and (ii) will imply that this space is dense in $\mathscr{H}$, we denote by $C(\xi)$ the unique extension of $V[H_0 + \xi V]^{-1}$ to all of $\mathscr{H}$. We require $C(\xi)$ be *weakly continuous* as a function of ξ. This means that for each pair of vectors ϕ and ψ, the function $\langle \phi, C(\xi)\psi\rangle$ is continuous for $\epsilon \geqslant 0$.

We can now state our results.

Theorem 2.1: *Suppose* (i)–(iii) *hold. Then for each vector* $\phi_0 \in \mathscr{D}(H_0)$, *there is precisely one solution* $\phi_\epsilon(t)$ *of* (2.4) *with* $\phi_\epsilon(0) = \phi_0$. *Furthermore we have the formula*

$$\phi_\epsilon(t) = \exp\left[-i\epsilon H\left(\frac{1}{\epsilon}\int_0^t f(s)\,ds\right)\right]\phi_0, \tag{2.6}$$

where $H(\xi)$ *is the closure of* $H_0 + \xi V$.

We will discuss the proof of this theorem in Sec. 3, as well as the proof of its important:

Corollary 2.1: $\lim_{\epsilon \to 0}\phi_\epsilon(t) = \exp\left(-i\int_0^t f(s)\,ds V\right)\phi_0$.

3. LEMMAS AND THE PROOF OF THEOREM 2.1 AND COROLLARY 2.1

In addition to the results of Dollard and Friedman[4] we will employ the following two lemmas that can be found in Kato.[6]

Lemma 3.1: Let T_n and T be self-adjoint operators. Suppose $\mathscr{D}$ is a dense subspace contained in the domains of T_n and T, and T equals the closure of its restriction to $\mathscr{D}$. If $T_n u \to Tu$ as $n \to +\infty$ for each u in $\mathscr{D}$, then for any nonreal number ζ,

$$(T_n - \zeta I)^{-1} \to (T - \zeta I)^{-1} \text{ as } n \to +\infty. \tag{3.1}$$

The next lemma is a special case of Theorem 2.6 on page 502 of Ref. 6.

Lemma 3.2: Suppose T_n and T are self-adjoint and (3.1) holds for some nonreal number ζ. Then it follows that for each t

$$e^{-itT_n} \to e^{-itT} \text{ as } n \to +\infty. \tag{3.2}$$

The convergence in (3.1) and (3.2) is strong point-wise convergence.

We proceed with the proof of Theorem 2.1. For simplicity we establish Eq. (2.6) in the case $\epsilon = 1$. The hypotheses are such that the proof is valid for any positive ϵ.

Following Dollard and Friedman we set $A(t) = H[f(t)]$, the closure of $H_0 + f(t)V$ on $\mathscr{D}(H_0) \cap \mathscr{D}(V)$. Then (2.4) becomes

$$i\frac{d\phi}{dt} = A(t)\phi. \tag{3.3}$$

To employ the results in Ref. 4 we must verify the following conditions:

(A) $A^{-1}(t)$ has a weakly continuous weak derivative $(d/dt)A^{-1}(t)$.

(B) $A(t)(d/dt)A^{-1}(t)$ is weakly continuous in t and bounded for each t.

To show that (A) is true fix t. Then on the space $A(t)[\mathscr{D}(H_0) \cap \mathscr{D}(V)]$ we have

$$A^{-1}(s) - A^{-1}(t) = A^{-1}(s)[A(t) - A(s)]A^{-1}(t).$$

But on $\mathscr{D}(H_0) \cap \mathscr{D}(V)$, $A(t) - A(s) = [f(t) - f(s)]V$. Hence by (iii) we may write

$$A^{-1}(s) - A^{-1}(t) = [f(t) - f(s)]A^{-1}(s)C(f(t)). \tag{3.4}$$

Since (3.4) holds on the dense subspace $A(t)[\mathscr{D}(H_0) \cap \mathscr{D}(V)]$ of $\mathscr{H}$ and both sides are bounded (3.4) is true on all of $\mathscr{H}$. Now since $A^{-1}(s)$ is weakly bounded by (ii), (3.4) shows that $A^{-1}(t)$ is weakly continuous. Furthermore,

$$\begin{aligned}\lim_{s \to t}\left\langle\frac{A^{-1}(s) - A^{-1}(t)}{s - t}\phi, \psi\right\rangle &= -\lim_{s \to t}\frac{f(s) - f(t)}{s - t}\langle A^{-1}(s)C(f(t))\phi, \psi\rangle \\ &= -f'(t)\langle A^{-1}(t)C(f(t))\phi, \psi\rangle,\end{aligned}$$

since f is differentiable and $A^{-1}(s)$ is now weakly continuous. Thus

$$\frac{d}{dt}A^{-1}(t) = -f'(t)A^{-1}(t)C(f(t)), \tag{3.5}$$

and (A) holds since by (iii) $C(f(t))$ is weakly continuous. Finally we see that (B) is also true since

$$A(t)\frac{d}{dt}A^{-1}(t) = -f'(t)C(f(t)).$$

Now we will prove that the solution $\phi(t)$ of (3.3) satisfying the initial condition $\phi(0) = \phi_0$ is given by the formula

$$\phi(t) = \exp\left[-iH\left(\int_0^t f(s)\,ds\right)\right]\phi_0. \tag{3.6}$$

Again we note that (2.6) for any $\epsilon > 0$ can be similarly derived.

For any positive number λ set

$$A_\lambda(t) = \lambda A(t)[\lambda + iA(t)]^{-1}$$

and consider the initial value problem

$$\begin{aligned} i\frac{d\phi_\lambda}{dt} &= A_\lambda(t)\phi_\lambda, \\ \phi_\lambda(0) &= \phi_0. \end{aligned} \tag{3.7}$$

Since (A) and (B) hold, we may collect some results from Ref. 4 in the form of a lemma.

Lemma 3.3: The following statements are true:

(a) For each t, $A_\lambda(t)$ is bounded and is norm continuous as a function of t.

(b) For each $\psi \in \mathscr{D}(A(t))$,

$$\|A_\lambda(t)\psi\| \leqslant \|A(t)\psi\| \quad\text{and}\quad \lim_{\lambda \to +\infty} A_\lambda(t)\psi = A(t)\psi.$$

(c) For each t, $\lim_{\lambda \to +\infty}\phi_\lambda(t) = \phi(t)$, where $\phi(t)$ is the solution of (3.3) with $\phi(0) = \phi_0$.

Term-by-term differentiation of the exponential series will now show that

$$\phi_\lambda(t) = \exp\left(-i\int_0^t A_\lambda(s)\,ds\right)\phi_0.$$

Thus (3.6) now follows if we can show that

$$\exp\left(-i\int_0^t A_\lambda(s)\,ds\right)\phi_0 \to \exp\left[-iH\left(\int_0^t f(s)\,ds\right)\right]\phi_0 \tag{3.8}$$

as $\lambda \to +\infty$.

Let $\lambda = \lambda_n$ where $\{\lambda_n\}$ is an arbitrary sequence tending to $+\infty$, and let ψ be an arbitrary vector in $\mathscr{D}(H_0) \cap \mathscr{D}(V)$. Then for each s, $0 \leqslant s \leqslant t$,

$$A(s)\psi = H_0\psi + f(s)V\psi,$$

$$\|A(s)\psi - A_{\lambda_n}(s)\psi\| \to 0 \quad \text{as } n \to +\infty,$$

and

$$\|A_{\lambda_n}(s)\psi\| \leqslant \|A(s)\psi\|.$$

So by Lebsegue's dominated covergence theorem,

$$\lim_{n \to +\infty} \int_0^t A_{\lambda_n}(s)\psi\, ds = \int_0^t A(s)\psi\, ds = H\left(\int_0^t f(s)\, ds\right)\psi.$$

So if

$$T_n \equiv \int_0^t A_{\lambda_n}(s)\psi\, ds \quad \text{and} \quad T \equiv H\left(\int_0^t f(s)\, ds\right)\psi,$$

Lemmas 3.1 and 3.2 show that

$$e^{-iT_n}\phi_0 \to e^{-iT}\phi_0 \quad \text{as } n \to +\infty.$$

This proves (3.8), and completes the proof of Theorem 2.1.

Corollary 2.1 follows now by another application of Lemmas 3.1 and 3.2 to Eq. (2.6).

[1] M.J. Clauser and M. Blume, Phys. Rev. B **3**, 583–91 (1971).
[2] H. Gzyl, J. Math. Phys. **18**, 1327 (1977).
[3] M. Blume, J. Math. Phys. **19**, 2004 (1978).
[4] J.D. Dollard and C.N. Friedman, J. Func. Anal. **28**, 309–54 (1978).
[5] B. Simon, *Quantum Mechanics for Hamiltonian Defined as Quadratic Forms* (Princeton, N.J., 1971).
[6] T. Kato, *Perturbation Theory for Linear Operator* (Springer-Verlag, New York, 1966).

On the asymptotic behavior of the derivatives of Airy functions [a]

Philip A. Maurone and Alain J. Phares
Department of Physics, Villanova University, Villanova, Pennsylvania 19085

(Received 6 March 1979; accepted for publication 15 May 1979)

We present a noniterative functional solution to the three-term recursion relation satisfied by the higher order derivatives of the Airy function. This solution allows one to obtain the asymptotic behaviors of these derivatives for large argument and also for large order.

In two recent publications[1,2] the linear potential as a quark confinement mechanism has been discussed for the case of nonzero angular momentum. It is well known that Airy's equation and its convergent solution, $Ai(z)$, are useful in describing zero angular momentum states. In Refs. 1 and 2 it is seen that not only the Airy function but its derivatives are also important when discussing nonzero angular momentum states. It is the purpose of this note to present some new useful properties of the derivatives of the Airy function.

It is well known that $W^{(m)}(z)$, the mth derivative of the Airy function, can be expressed in the form

$$W^{(m)}(z) = P_m(z)W(z) + Q_m(z)W'(z), \tag{1}$$

where $P_m(z)$ and $Q_m(z)$ are polynomials. Indeed this result can be easily seen from the recurrence relations

$$P_{m+1}(z) = P'_m(z) + zQ_m(z),$$
$$Q_{m+1}(z) = P_m(z) + Q'_m(z) \tag{2}$$

obtained by use of Airy's equation[3]

$$W''(z) - zW(z) = 0, \tag{3}$$

and the starting values

$$P_0(z) = 1, \quad Q_0(z) = 0.$$

On the other hand, no explicit expressions for the polynomials $P_m(z)$ and $Q_m(z)$ are known. The recurrence relations (2) are coupled and so far no one has been able to solve them. Instead, we replace them by one single linear finite-difference equation, which, in turn, is solved using the recent combinatorics function technique.[4] Taking repeated derivatives of Eq. (3), one obtains the three-term recurrence relation among the higher order derivatives of the Airy function, namely[5]

$$W^{(m)} = (m-2)W^{(m-3)} + zW^{(m-2)}. \tag{4}$$

Systematic application of the technique of Ref. 4 to Eq. (4) leads to the results (the square brackets refer to integer divisions)

$$W^{(3p)}(z) = 3^p \sum_{j=0}^{1} z^j W^{(j)}(z) \sum_{n=0}^{[(p-j)/2]} (3z^3)^n \times \sum_{q=0}^{[3n+j]} \frac{(-1)^q}{q!(3n+j-q)!} \times \frac{\Gamma(p+n+(j+1-q)/3)}{\Gamma((j+1-q)/3)}, \tag{5a}$$

$$W^{(3p+1)}(z) = 3^{p+1} z^2 \sum_{j=0}^{1} (3z)^{-j} W^{(j)}(z) \sum_{n=0}^{[(p+j-1)/2]} (3z^3)^n \times \sum_{q=0}^{[3n+2-j]} \frac{(-1)^q}{q!(3n+2-j-q)!} \times \frac{\Gamma(p+n+1+(j+1-q)/3)}{\Gamma((j+1-q)/3)}, \tag{5b}$$

$$W^{(3p+2)}(z) = 3^{p+1} z \sum_{j=0}^{1} z^j W^{(j)}(z) \sum_{n=0}^{[(p-j)/2]} (3z^3)^n \times \sum_{q=0}^{[3n+j+1]} \frac{(-1)^q}{q!(3n+j+1-q)!} \times \frac{\Gamma(p+n+1+(j+1-q)/3)}{\Gamma((j+1-q)/3)}, \tag{5c}$$

Comparing Eqs. (1) and (5), $P_m(z)$ and $Q_m(z)$ are then easily identified. The resulting expressions for P_m and Q_m can then be verified by substituting them in recurrence relations (2), the same way expressions (5) can also be shown to satisfy Eq. (4).

Equations (5a)–(5c) provide easy asymptotic expresions for large values of z and for a given order p. This asymptotic behavior is based on the well known[3] asymptotic expansions for $W(z)$ and $W'(z)$. Obviously, all higher order derivatives of $W(z)$ will fall off at infinity if $W(z)$ is the well behaved $Ai(z)$, and will blow up if $W(z)$ is any linear combination of $Ai(z)$ and $Bi(z)$.

Asymptotic behaviors of $W^{(m)}(z)$ for large order m can also be computed using Eqs. (5a)–(5c). This is perhaps less trivial to obtain. By setting $m = 3p + i$ ($i = 0,1,2$), we see in Eqs. (5a)–(5c) the dependence on p appear at only two places: the upper limit on the n-summation and the argument of one Γ function.

The details of the calculations leading to Eqs. (5a)–(5c) and other formulas (not available in the literature) can be found in Ref. 6.

[a] Work supported in part by a grant from the Office of Research and Patent Affairs, Villanova University.

[1] A.F. Antippa and A.J. Phares, J. Math. Phys. **19**, 385 (1978).
[2] A.J. Phares, J. Math. Phys. **19**, 2239 (1978).
[3] See, for example, *Handbook of Mathematical Functions*, edited by M. Abramowitz and I.A. Stegun (National Bureau of Standards, Washington, D.C., 1964).
[4] This is a technique that allows one to solve any one-dimensional multiterm linear recursion relation. This technique is formally developed in A.F. Antippa and A.J. Phares, J. Math. Phys. **18**, 173 (1977); A.J. Phares, J. Math. Phys. **18**, 1838 (1977); and A.F. Antippa, J. Math. Phys. **18**, 2214 (1977).
[5] This property is not new but presented in P.L. Ferreira and J.A. Castilho Alcarás, Lett. Nuovo Cimento **15**, 500 (1975).
[6] P.A. Maurone and A.J. Phares, Villanova preprint #TH23-03 (1978).

Schwinger equations

Jerzy Hańćkowiak

Institute of Theoretical Physics, University of Wrocław, Cybulskiego 36, 50-205 Wrocław, Poland
(Received 24 March 1978)

A compact description of the Schwinger equation is given and its different transformations and projections are considered.

In the case of systems with finite degrees of freedom the process of solving a given system of equations is in principle finished if we reach an integral form of solutions. For systems with infinite degrees of freedom like quantum field theory (qft) or statistical physics (st), a completely different situation occurs because of troubles with the effective approximation of the resulting integrals (functional integrals). It turns out that in such cases it is useful sometimes to come back to the starting equations, and by means of those one can try to find new approximations to considered quantities. In the work presented we concentrate on qft formulated by means of the functional integral.

$$\mathscr{V}[\alpha] = \mathscr{N}^{-1} \int d\mu[\phi]\, e^{\mathscr{W}[+i\varphi]} e^{i(\varphi,\alpha)} \tag{1.1}$$

with the measure

$$\int d\mu[\varphi]\, e^{i(\varphi,\alpha)} = e^{(\alpha, K^{-1}\alpha)/2} \tag{1.2}$$

and an r-component classical field $\alpha(x)$.

We consider here the Schwinger equation (5.2) fulfilled by the functional $\mathscr{V}$. We would like to show in this paper that the Schwinger equation can be used as a good starting point for different approximations of the functional $\mathscr{V}$. To do this, the equation has to be described in the standard form of one linear, inhomogeneous equation, Sec. 5, and transformed in different ways.

The introduction of a general notion of projection operators was found to be very fruitful for investigation of the Schwinger equation. By this we mean an operation P with only idempotent property, $\mathrm{P} = \mathrm{P}^2$. It turns out that there are a lot of operations in nature and theory connected with the above property. For example, stirring sweet tea or substracting infinities from Feynman's graphs are projection operators. There are different possibilities of applying the projector operators to the Schwinger equation. First of all, we can use them to separate a chosen process from others. Secondly, the projection operators can also be used to obtain the nonoscillator perturbation approach to qft. Thirdly, by means of the projectors (projection operators) we can improve mathematical correctness of the formulas obtained.

I. THE CONTENT OF THE WORK

In Secs. 2 and 3 we give some remarks and consider general functional equations. In Sec. 4 we discuss the projection operators. In Sec. 6 we consider projection properties of the Schwinger equation. It turns out that the free part of the Schwinger operator is the lowering operator ($P_i \amalg\!\amalg_0 P_j \neq 0$ only for $j < i$) but the interacting part is the nonlowering operator. The latter property is responsible for linking of equations for n-point functions in infinite chain. The possibility of decoupling of such a chain of equations is considered in Sec. 7. The most spacious is Sec. 8, where we demonstrate different applications of the projected Schwinger equation. In particular we derive there the Feshbach equations and we propose the modified Tamm–Dancoff method for calculating kernels of those integral equations. The Dyson equations, as a particular case of the Feshbach equations, are derived there. It turns out now that the kernels of these equations are the inverses of projections of the Schwinger operator. We also discuss in this section the application of the projected Schwinger equation to renormalization of a theory. This is possible since the usual subtraction operations are a projection operators.

Section 9 contains variational formulation of the Schwinger equations and some remarks about the improvement of calculations when we do not know the exact properties of the considered operators. In particular, we consider there the method of effective operators.

Section 10 contains similarity transformations of the Schwinger equation. We also give there a definition of general symmetrical transformations of the Schwinger equation.

II. GENERAL REMARKS

A. Operator and vector equations

Let us consider a linear inhomogeneous equation

$$\mathrm{AX} = \mathrm{Y}, \tag{2.1}$$

where all occurring quantities are operators. This is an operator equation. The vector equation is obtained if operators on both sides of (2.1) act on the vector $\mathscr{V}_0 \in D(\mathrm{X})$ and $D(\mathrm{Y})$. We then get

$$\mathrm{A}\mathscr{X} = \mathscr{Y} \tag{2.2}$$

with

$$\mathscr{X} = \mathrm{X}\mathscr{V}_0, \quad \mathscr{Y} = \mathrm{Y}\mathscr{V}_0 \tag{2.3}$$

In spite of the fact that in physics vector equations are more original, we also consider the operator equations because they are more general and may lead to calculational consequences (Sec. 9C).

B. Solution of (2.1)

To solve (2.1), we usually transform that equation

 0022-2488/79/112192-12$01.00

$$\mathrm{BAX} = \mathrm{BY} \tag{2.4}$$

to obtain a simpler one. For B, which has an inverse operator, (2.4) is equivalent to (2.1). If A^{-1} exists we may express the solution of (2.1) or (2.4) in the form of

$$\mathrm{X} = (\mathrm{BA})^{-1}\mathrm{BY}. \tag{2.5}$$

The solution does not depend on a freely chosen operator B and is indeed equal to

$$\mathrm{X} = \mathrm{A}^{-1}\mathrm{Y}. \tag{2.6}$$

However, in the expression (2.5) we can make use of the fact that the inverse BA can be easier to calculate than the inverse A.

The standard choice of the operator B is the following one[1]:

$$\mathrm{B} = (\mathrm{A} + \mathrm{A}_1)^{-1}, \tag{2.7}$$

where A_1 is any operator. Then from (2.1) we have

$$\mathrm{X} = \mathrm{BA}_1\mathrm{X} + \mathrm{BY}. \tag{2.8}$$

The operator B is successfully chosen if (2.7) is calculable and the spectral radius ρ of the operator BA is less than 1. We then apply to (2.8) the method of successive approximation:

$$\mathrm{X}_n = \mathrm{BA}_1\mathrm{X}_{n-1} + \mathrm{BY}. \tag{2.9}$$

In a vector version equation (2.9) becomes

$$\mathscr{X}_n = \mathrm{BA}_1\mathscr{X}_{n-1} + \mathrm{B}\mathscr{Y}. \tag{2.10}$$

For $\rho(\mathrm{BA}_1) < 1$ and any ϵ, $0 < \epsilon < 1 - \rho(\mathrm{BA}_1)$, the sequence $\mathscr{X}_n$ tends to the unique solution $\mathscr{X}^*$ and the following estimation holds true

$$\|\mathscr{X}_n - \mathscr{X}^*\| \leqslant c(\epsilon)[\rho(\mathrm{BA}_1) + \epsilon]^n \|\mathscr{X}_0 - \mathrm{BA}_1\mathscr{X}_0 - f\| \tag{2.11}$$

(see Ref. 1, Chap. I, Sec. 2).

Often, it is hard to find the spectral radius of a considered operator. For such cases special procedure has been developed in Sec. 9.

C. Projected equations

Particularly useful to analyze the linear equation (2.1) or (2.2) are projection operators (projectors) which by definition have to be only idempotent,[2]

$$\mathrm{P} = \mathrm{P}^2. \tag{2.12}$$

Any projector P divides the initial space E into two complementary linear manifolds

$$\mathrm{E} = \mathrm{M} \oplus \mathrm{N}, \tag{2.13}$$

where

$$\mathrm{M} = \mathrm{PE}, \quad \mathrm{N} = (\mathrm{I} - \mathrm{P})\mathrm{E}. \tag{2.14}$$

The opposite statement is also true. Equation (2.13) corresponds to a pair of complementary projectors

$$\mathrm{P} + \mathrm{Q} = \mathrm{I} \tag{2.15}$$

such that

$$\mathrm{PE} = \mathrm{M}, \quad \mathrm{QE} = \mathrm{N}.$$

Using a pair of complementary projectors (2.15) we may rewrite Eq. (2.1) in an equivalent, projected form as

$$\mathrm{PAPX} + \mathrm{PAQX} = \mathrm{PY}, \tag{2.16a}$$

$$\mathrm{QAQX} + \mathrm{QAPX} = \mathrm{QY}. \tag{2.16b}$$

We may use only one equation, e.g., (2.16b), but then the projection PX has to be treated as a given quantity. The choice of PX cannot lead beyond the class of possible solutions (2.1). It means that for Eq. (2.1) whose general solution is parametrized by a finite number of f parameters, P can project on an f-dimensional manifold at the most. In addition, if the kernel of a projected operator QAQ is trivial, Eq. (2.16b) is equivalent to (2.11).

III. FUNCTIONAL EQUATIONS

In statistic and quantum physics the following functional equations appear:

$$A[\alpha,\delta/\delta\alpha]\mathscr{X}[\alpha] = \mathscr{Y}[\alpha], \tag{3.1}$$

where A and $\mathscr{Y}$ are given quantities and the functional differential operator A has a general, normal order form of

$$A[\alpha,\delta/\delta\alpha] = \sum_{n,m}\frac{1}{m!}\frac{1}{n!}\int a_{mn}(x_{(m)},y_{(n)}) \times \alpha(x_1)\cdots\alpha(x_m)\frac{\delta}{\delta\alpha(y_1)}\cdots\frac{\delta}{\delta\alpha(y_n)}. \tag{3.2}$$

The useful representation of the operator A, defined by (3.2), expresses in an explicit way the dependence on the functional derivatives. That is,

$$A[\alpha,\delta/\delta\alpha] \leftrightarrow \mathrm{A}[\alpha,\beta] = e^{\alpha\beta}A[\alpha,\beta], \tag{3.3}$$

where β are now functions and $\alpha\beta = (\alpha,\beta)$. Following Rzewuski,[3] we call this representation the matrix representation of the operator A. Using the Volterra series produced by the formula

$$\mathscr{Y}[\alpha] = e^{\alpha\delta/\xi}\mathscr{Y}[\xi] = |_{\xi=0}, \tag{3.4}$$

we express, by means of that representation Eq. (3.1) as

$$\mathrm{A}[\alpha,\delta/\delta\xi]\mathscr{X}[\xi]_{\xi=0} = \mathscr{Y}[\alpha]. \tag{3.5a}$$

We may consider the lhs of equality (3.5) as the *definition* of action of the operator A represented by the double, functional power series, on the vector $\mathscr{X}$, represented by the single functional power series. In other words, we can rewrite (3.5) in an equivalent way as

$$\mathrm{A}\mathscr{X} = \mathscr{Y}. \tag{3.5b}$$

This is the *vector equation*. The *operator equation* is obtained if the vectors $\mathscr{X}$ and $\mathscr{Y}$ are substituted for the operators X and Y which, we assume, admit also representations in the form of a double functional series. We have then

$$\mathrm{A}[\alpha,\delta/\delta\xi]\mathrm{X}[\xi,\beta]_{\xi=0} = \mathrm{Y}[\alpha,\beta] \tag{3.6a}$$

with

$$\mathscr{X}[\alpha] = \mathrm{X}[\alpha,0], \quad \mathscr{Y}[\alpha] = \mathrm{Y}[\alpha,0]. \tag{3.7}$$

In an abbreviated notation we can write (3.6a) as follows,

$$\mathrm{AX} = \mathrm{Y}. \tag{3.6b}$$

IV. UNIT AND PROJECTION OPERATORS

Using the lhs of the formula (3.6a) as a definition of the operator multiplication, we can chek at once that the matrix representation of the unit operator is given by

$$\mathrm{I}[\alpha,\beta] = e^{\alpha\beta} = \sum \frac{1}{n!}(\alpha\beta)^n. \tag{4.1}$$

According to the same *multiplication law*, the operators

$$\mathrm{P}_i[\alpha,\beta] = \frac{1}{i!}(\alpha\beta)^i \tag{4.2}$$

are projectors and

$$\mathrm{P}_i[\alpha,\delta/\delta\xi]\mathrm{P}_j|[\xi,\beta]|_{\xi=0} = \delta_{ij}\mathrm{P}_j[\alpha,\beta] \tag{4.3a}$$

or

$$\mathrm{P}_i\mathrm{P}_j = \delta_{ij}\mathrm{P}_j. \tag{4.3b}$$

We have also

$$\mathrm{I} = \sum_i \mathrm{P}_i. \tag{4.4}$$

Any matrix representation of the operator can be represented as follows:

$$\mathrm{A}[\alpha,\beta] = \sum_{i,j} \mathrm{P}_i[\alpha,\delta/\delta\xi]A[\xi,\delta/\delta\gamma]\mathrm{P}_j[\gamma,\beta]|_{\xi=\gamma=0} \tag{4.5a}$$

or in a shortened form

$$\mathrm{A} = \sum_{i,j} \mathrm{P}_i\mathrm{A}\mathrm{P}_j. \tag{4.5b}$$

The projected operator $\mathrm{P}_i\mathrm{A}\mathrm{P}_j$ is represented by the functional monomial of order i in α and j in β:

$$\mathrm{P}_i\mathrm{A}\mathrm{P}_j)[\alpha,\beta] = \int dx_{(i)}dy_{(j)}\alpha(x_1)\cdots\alpha(x_i)\mathrm{A}_{ij}(x_{(i)},y_{(j)})\beta(y_1)\cdots\beta(y_j), \tag{4.6}$$

where, e.g., $x_{(i)} = (x_1,\ldots,x_i)$. Thus we emphasize explicitly the dependence of the functional A on two kinds of function α and β. This provides useful information for investigating the decoupled properties of the Schwinger equation.

V. SCHWINGER EQUATION AND DIFFERENT COUPLING CONSTANT EXPANSIONS

The functional $\mathscr{V}$ defined by (1.1) and (1.2) and expressed by

$$\mathscr{V}[\alpha] = \mathscr{N}^{-1}\exp(\mathscr{W}[\delta/\delta\alpha])\exp(\tfrac{1}{2}\alpha K^{-1}\alpha), \tag{5.1}$$

fulfills the Schwinger equation

$$\delta/\delta\alpha(x)\mathscr{V} - [(K^{-1}\alpha)(x) + (K^{-1}\mathscr{W}'[\delta/\delta\alpha])(x)]\mathscr{V} = 0, \quad x\in R^d \tag{5.2a}$$

where, e.g.,

$$(K^{-1}\alpha)(x) = \int K^{-1}(x,y)\alpha(y)\,dy.$$

In qft K^{-1} is porportional to the free propagators and $\mathscr{W}$ is connected with an interacting part of the action integral. It can be seen that the Schwinger equation (5.2) has a particular form of the functional equation (3.1) with

$$A[x;\alpha,\delta/\delta\alpha] = \frac{\delta}{\delta\alpha(x)} - (K^{-1}\alpha)(x) - (K^{-1}\mathscr{W}'[\delta/\delta\alpha]). \tag{5.3}$$

The matrix representation of the operator (5.3) is

$$\mathrm{A}[x;\alpha,\beta] = [\beta(x) - (K^{-1}\alpha)(x)] - (K^{-1}\mathscr{W}'[\beta])(x)]e^{\alpha\beta}, \tag{5.4}$$

and the Schwinger equation (5.2) can be described as

$$\mathrm{A}[x;\alpha,\delta/\delta\xi]\mathscr{V}[\xi]|_{\xi=0} = 0, \quad x\in R^d. \tag{5.2b}$$

The functional derivative in the matrix representation is given by the functional $\mathrm{D}[x;\alpha,\beta] = \beta(x)e^{\alpha\beta}$. It is easy to check that an inverse operation for the functional derivative, defined by the relation $\int dx\, \mathrm{D}^{-1}[x;\alpha,\delta/\delta\xi]\mathrm{D}[x;\xi,\beta|]_{\xi=0} = e^{\alpha\beta} - 1$, is given by the formula $\mathrm{D}^{-1}[x;\alpha,\beta] = \alpha(x)\int_0^1 d\lambda e^{\lambda\alpha\beta} = \alpha(x)(\alpha\beta)^{-1} \times (e^{\alpha\beta} - 1)$.[4] Applying the above formula to the operator A we get

$$\int dx D^{-1}[x;\alpha,\delta/\delta\xi]\mathrm{A}[x;\xi,\beta]|_{\xi=0} = e^{\alpha\beta} - 1 - \alpha K^{-1}\alpha g(\alpha\beta) - \alpha K^{-1}\mathscr{W}'[\beta]f(\alpha\beta) \tag{5.5}$$

with

$$f(t) = \int_0^1 d\lambda e^{\lambda t}, \quad g(t) = f^1(t).$$

Introducing the Schwinger operator

$$\mathrm{I}[\alpha,\beta] - \mathrm{Ш}[\alpha,\beta] = \mathrm{I}[\alpha,\beta] - \mathrm{Ш}_0[\alpha,\beta] - \mathrm{Ш}_1[\alpha,\beta] \tag{5.6}$$

with

$$\mathrm{Ш}_0 = \alpha K^{-1}\alpha\cdot g(\alpha\beta) \quad \text{(free term)}, \tag{5.7a}$$

$$\mathrm{Ш}_1 = \alpha K^{-1}\mathscr{W}'[\beta]f(\alpha\beta) \quad \text{(interacting term)}, \tag{5.7b}$$

we describe the Schwinger equation in the form of

$$\mathscr{V}[\alpha] - \mathrm{Ш}[\alpha,\delta/\delta\xi]\mathscr{V}[\xi]|_{\xi=0} = 1, \tag{5.8a}$$

where we have used the normalized condition $\mathscr{V}[0] = 1$. To abbrevate this,

$$(\mathrm{I} - \mathrm{Ш})\mathscr{V} = \mathbb{1}, \tag{5.8b}$$

where $\mathbb{1}$ is the vector with only one component different from zero, $\mathrm{P}_0\mathbb{1} = 1$.

In this way we have arrived at the Schwinger equation described in the standard form (Sec. 2) of *one* linear, inhomogeneous equation. Further, apart from the vector Schwinger Eq. (5.8) we consider the operator Schwinger equation

$$\mathrm{X}[\alpha,\beta] - \mathrm{Ш}[\alpha,\delta/\delta\xi]\mathrm{X}[\xi,\beta]|_{\xi=0} = \mathrm{Y}[\alpha,\beta] \tag{5.9a}$$

with

$$\mathrm{Y}[\alpha,0] = \mathbb{1} \quad \text{and} \quad \mathscr{V}[\alpha] = \mathrm{X}[\alpha,0], \tag{5.10}$$

which we also describe in an abbreviated way as

$$(\mathrm{I} - \mathrm{Ш})\mathrm{X} = \mathrm{Y}. \tag{5.9b}$$

We may choose the unit operator for Y, because it is seen from (4.1) that (5.10) is fulfilled. The Schwinger equation then can be described in the form of an equation for the right *inverse operator*

$$(\mathrm{I} - \text{Ш})\mathrm{X} = \mathrm{I}. \tag{5.11}$$

If we assume that the right and the left inverse operators are the same we get a new equation for X,

$$\mathrm{X}(\mathrm{I} - \text{Ш}) = \mathrm{I}. \tag{5.12}$$

A. Perturbation series

The usual perturbation series for the solution of Schwinger equation is obtained after rewriting (5.11) as

$$(\mathrm{I} - (\mathrm{I} - \text{Ш}_0)^{-1}\text{Ш}_1)\mathrm{X} = (\mathrm{I} - \text{Ш}_0)^{-1} \tag{5.13}$$

and representing X in the form of the Neumann series

$$\mathrm{X} = \sum_{j=0}^{\infty} [(\mathrm{I} - \text{Ш}_0)^{-1}\text{Ш}_1]^j(\mathrm{I} - \text{Ш}_0)^{-1}. \tag{5.14}$$

For

$$\mathscr{Y} = \mathrm{XP}_0, \tag{5.15}$$

(5.14) produces the same formula as (5.1).

B. Inverse coupling constant series

The series, useful for a large coupling constant is obtained after transforming (5.11) into

$$(\mathrm{I} - (\mathrm{I} - \text{Ш}_1)^{-1}\text{Ш}_0)\mathrm{X} = (\mathrm{I} - \text{Ш}_1)^{-1}. \tag{5.16}$$

The Neumann series for

$$\mathrm{X} = \sum_{j=0}^{\infty} [(\mathrm{I} - \text{Ш}_1)^{-1}\text{Ш}_0]^j(\mathrm{I} - \text{Ш}_1)^{-1}. \tag{5.17}$$

Introducing explicit dependence on the coupling constant

$$\text{Ш}_1 = g\,\text{Ш}_1 \tag{5.18}$$

and using the so-called the first Neumann series for the resolvent,[2]

$$(\mathrm{I} - g\text{Ш}_1)^{-1} = g^{-1}(g^{-1}\mathrm{I} - \text{Ш}_1)^{-1}$$

$$= -\sum_{n=0}^{\infty} g^{-n-1}\text{Ш}_1^{-n-1}, \tag{5.19}$$

we get

$$\mathrm{X} = \sum_{j=0}^{\infty}(-1)^{j+1}\left[\sum_{n=0}^{\infty} g^{-n-1}\widetilde{\text{Ш}_1^{-n-1}}\text{Ш}_0\right]^j \times \sum_{m=0}^{\infty} g^{-m-1}\text{Ш}_1^{-m-1}. \tag{5.20}$$

The first terms are

$$\mathrm{X} = g^{-1}\widetilde{\text{Ш}_1^{-1}} + g^{-2}[\widetilde{\text{Ш}_1^{-1}}\text{Ш}_0\widetilde{\text{Ш}_1^{-1}} - \text{Ш}_1^{-2}] + o(g^{-3}). \tag{5.21}$$

We would like to note here that applying appropriate inverse operators or applying Neumann series cannot be, in general, justified operations, which is seen most directly when particular terms of the given expansion are the divergent quantities. The projected Schwinger equations appear to be partially remedial for these difficulties.

VI. PROJECTION PROPERTIES OF THE SCHWINGER OPERATORS

The functions f and g appearing in the definition (5.6), (5.7) of the Schwinger operators are

$$f(t) = t^{-1}(e^t - 1) = \sum_0^{\infty}\frac{t^n}{n!(n+1)} \tag{6.1a}$$

and

$$g(t) = \sum_0^{\infty}\frac{t^n}{n!(n+2)}. \tag{6.1b}$$

Hence we get for $i > 1$

$$\mathrm{P}_i\text{Ш}_0 = \mathrm{P}_i\text{Ш}_0\mathrm{P}_{i-2} \tag{6.2}$$

and $\mathrm{P}_i\text{Ш}_0 = 0$ for $i = 0,1$, where projectors P_i are defined in Sec. 4. It means that the free part of the Schwinger operator admits the following expansion

$$\text{Ш}_0 = \sum_2^{\infty}\mathrm{P}_i\text{Ш}_0\mathrm{P}_{i-2}. \tag{6.3}$$

This is a subdiagonal operator

$$\text{Ш}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \mathrm{P}_2\text{Ш}_0\mathrm{P}_0 & 0 & 0 & 0 & \cdots \\ 0 & \mathrm{P}_3\text{Ш}_0\mathrm{P}_1 & 0 & 0 & \cdots \\ 0 & 0 & \mathrm{P}_4\text{Ш}_0\mathrm{P}_2 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdots \end{pmatrix}.$$

In general, a more complicated case is the interacting part of the Schwinger operator. For

$$\mathscr{W}[\beta] = \sum_0^{\infty}\frac{1}{n!}\int w_n(x_1,\dots,x_n)\beta(x_1)\cdots\beta(x_n)\,dx_{(n)} \tag{6.4}$$

we get from (6.1a) and (5.7b) the following formula,

$$\mathrm{P}_i\text{Ш}_1 = \sum_{j=i-1}^{\infty}\mathrm{P}_i\text{Ш}_1\mathrm{P}_j. \tag{6.5}$$

For the usual interactions the sum (6.4) begins with $n = 3$, and (6.5) is substituted for

$$\mathrm{P}_i\text{Ш}_1 = \sum_{j=i+1}^{\infty}\mathrm{P}_i\text{Ш}_1 P_j. \tag{6.6}$$

In this case

$$\text{Ш}_1 = \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\mathrm{P}_i\text{Ш}_1\mathrm{P}_j. \tag{6.7}$$

In the matrix representation of Ш_1, all elments are placed over the principle diagonal.

In both cases (6.5) and (6.6) we have for the projections of the whole Schwinger operator

$$\mathrm{P}_i(\mathrm{I} - \text{Ш}) = \mathrm{P}_i - \sum_{j=i-2}\mathrm{P}_i\text{Ш}\mathrm{P}_j \tag{6.8}$$

and the expression

$$\text{Ш} = \sum_{i=1}^{\infty}\sum_{j=i-2}^{\infty}\mathrm{P}_i\text{Ш}\mathrm{P}_j. \tag{6.9}$$

From the expansions (6.3), (6.7), and (6.9) we get for the right projections:

$$\text{Ш}_0\mathrm{P}_j = \mathrm{P}_{j+2}\text{Ш}_0\mathrm{P}_j, \quad j = 0,1,\cdots, \tag{6.10}$$

$$\text{Ш}_1 P_j = \sum_{i=1}^{j-1} P_i \text{Ш}_1 P_j, \; j = 1,2,\cdots, \tag{6.11}$$

$$\text{Ш} P_j = \sum_{i=1}^{j+2} P_i \text{Ш}_1 P_j. \tag{6.12}$$

Definition: The polynomial interactions are characterized by the finite terms in the expression (6.4). We have, e.g.,

$$\mathscr{W} = \sum_0^m \frac{1}{n!} \int w_n(x_{(n)})\, \beta(x_1)\cdots\beta(x_n)\, dx_{(n)}. \tag{6.13}$$

The terms with $n = 0,1,2$ ususally disappear in QFT. The term with $n = 0$ gives a trivial multiplicative change of $\mathscr{V}$ [see def. (5.1)]. The term with $n = 1$ disappears from vacuum stability. The term with $n = 2$ usually enters the free part of Ш_0. In this case in the formulas (6.5)–(6.7) sums over j are finite.

A. Inverse operators

Inverse operators $(I - \text{Ш}_0)^{-1}$ and $(I - \text{Ш}_1)^{-1}$ which appear in the transformed Schwinger equation, Sec. 5, have interesting properties. *We assume* that in the space considered they can be expressd in the form of the Neumann series

$$(I - \text{Ш}_0)^{-1} = \sum_0^\infty \text{Ш}_0^l, \tag{6.14}$$

$$(I - \text{Ш}_1)^{-1} = \sum_0^\infty \text{Ш}_1^l. \tag{6.15}$$

Then for the left projection, we have from (6.2) and (6.14) the following expansion,

$$P_m (I - \text{Ш}_0)^{-1}$$

$$= \begin{cases} \sum_{l=0}^{m/2} P_m \text{Ш}_0 P_{m-2} \cdots P_{m-2l+2} \text{Ш}_0 P_{m-2l}, & \text{for even } m, \\ \sum_{l=0}^{(m-1)/2} P_m \text{Ш}_0 P_{m-2} \cdots P_{m-2l+2} \text{Ш}_0 P_{m-2l}, & \text{for odd } m, \end{cases} \tag{6.16}$$

and only the finite terms of the expansion (6.14) contribute to the projection (6.16).

The same situation appears for the right projection of the operator $(I - \text{Ш}_1)^{-1}$ in the case of (6.7). We have then from (6.15) and (6.11) the following equality,

$$(I - \text{Ш}_1)^{-1} P_m = P_m + \sum_n \sum_{i_1,\ldots,i_n} P_{i_1} \text{Ш}_1 P_{i_2} \cdots P_{i_n} \text{Ш}_1 P_m. \tag{6.17}$$

But because of $i_n \leqslant m-1, i_{n-1} \leqslant m-2$, and so on, we obtain $i_1 < 1$ for appropriately large n. It means that in the case of interaction (6.7), the following is true.

Theorem: The right projection of (6.15) contains only finite terms of the corresonding infinite sum.

B. Monomial interactions

Definition: For the monomial interactions the functional $\mathscr{W}$ appearing in the Schwinger operator (5.6) has the form

$$\mathscr{W}[\beta] = \int dx_{(n)}\, w_n(x_{(n)}) \beta(x_1)\cdots\beta(x_n), \tag{6.18}$$

where $n > 2$.

For $n = 3,4$ the interaction is normalizable, for $n > 4$ unrenormalizable. Now the operators Ш and Ш_1 have very simple projective properties. For general n we have

$$P_i \text{Ш} = P_i \text{Ш}_1 P_{i+n-2}, \tag{6.19}$$

where $i > 0$, $n > 2$;

$$\text{Ш}_1 P_i = P_{i-n+2} \text{Ш}_1 P_i, \tag{6.20}$$

where $i \geqslant n-2 > 0$. In other cases projections are equal to zero. For the whole Ш

$$P_i \text{Ш} = P_i(\text{Ш}_0 P_{i-2} + \text{Ш}_1 P_{i+n-2})$$

$$= P_i \text{Ш}(P_{i-2} + P_{i+n-2}) \tag{6.21}$$

and

$$\text{Ш} P_i = (P_{i+2} + P_{i-n+2}) \text{Ш} P_i. \tag{6.22}$$

The projection

$$(I - \text{Ш}_1)^{-1} P_i = P_i + \sum_l P_{i-l(n-2)} \text{Ш}_1 \cdots P_{i-n+2} \text{Ш}_1 P_i. \tag{6.23}$$

Taking into account that $n - 2 > 0$ and that $P_j \text{Ш}_1 = 0$ for $j < 1$ we see that all terms of the expansion (6.15) with appropriate large l disappear and we have only finite terms in (6.23). See the above theorem.

VII. DECOUPLED EQUATIONS

The usual Schwinger equation described in the form (5.11)

$$(I - \text{Ш})X = I$$

is the infinite chain of coupled equations for projections $P_i X$. For example, in the case of monomial interactions (6.18) [see Eq. (6.21)] we have

$$(P_i - P_i \text{Ш}(P_{i-2} + P_{i+n-2}))X = P_i, \quad i = 2,3,\cdots. \tag{7.1}$$

Because $n - 2 > 0$, we see that the Schwinger equation links the $P_i X$ projection with lower and higher projections and leads, therefore, to an infinite chain of equations.

Definition: The projection P decouples the equation

$$AX = Y \tag{7.2}$$

if

$$PA = PAP. \tag{7.3}$$

In a general case of the projector P, for which the inverse operator of the projected operator QAQ exists, we may transform (7.2) in such a manner that

$$BAX = BY \tag{7.4}$$

is decoupled for that projector. We have to take

$$B = B_Q = I - A(QAQ)^{-1}, \tag{7.5}$$

where Q is a conjugated projector to P,

$$Q + P = I. \tag{7.6}$$

It is easy to prove that

$$PB_Q A = PB_Q AP. \tag{7.7}$$

Indeed,

$$
\begin{aligned}
PB_Q A &= PA - PA(QAQ)^{-1}A \\
&= PA - PA(QAQ)^{-1}A(A+P) \\
&= PA - PAQ - PA(QAQ)^{-1}AP \\
&= PAP - PA(QAQ)^{-1}AP \\
&= PB_Q AP.
\end{aligned}
$$

Similarly

$$
\begin{aligned}
B_Q AP &= AP - A(QAQ)^{-1}AP \\
&= AP - QAP - PA(QAQ)^{-1}AP \\
&= PAP - PA(QAQ)^{-1}AP \\
&= PB_Q AP.
\end{aligned}
$$

Hence we see that P commutes with $B_Q A$,

$$[P, B_Q A] = 0, \tag{7.8}$$

and as a consequence we have (7.7). It is, however, a problem to calculate the *decoupled transformation B_Q* which contains the inverse operator $(QAQ)^{-1}$.

We'll show now that in particular cases of the small and large coupling constant, decoupling is practically realized without using (7.5). Special equations which admit decoupling for particular projectors are equations with operators A for which

$$P_i A P_j = 0 \quad \text{for } j > i. \tag{7.9}$$

For projectors

$$P^i \equiv \sum_{l=o}^{i} P_l \tag{7.10}$$

we get

$$P^i A = P^i A P^i,$$

where we have used the following properties of the projectors considered

$$P_l P^i = P^i P_l = P_l \tag{7.11}$$

A. Small coupling constant

Now we can transform the Schwinger equation (5.11) in such a manner that decoupling is realized approximately in the case of the small coupling constant. We shall make use of the fact that projections of Ш_0 are lowering, Eq. (6.2). Multiplying (5.11) by $(I - \text{Ш}_0)^{-1}$ and using (6.14) we get

$$\Big(I - \sum_0^{\infty} \text{Ш}_0^l \text{Ш}_1\Big)X = (I - \text{Ш}_0)^{-1}. \tag{7.12}$$

The projections of Ш_1 are not lowering. Using the lowering property of Ш_0 we shall always find for the polynomial interactions (6.13) an appropriate power of Ш_0 so that $\text{Ш}_0^l \text{Ш}_1$ will fulfill (7.9). To be more precise we assume that

$$\text{Ш}_1 = \sum_{m=3}^{r} \text{Ш}_{1m}, \tag{7.13}$$

where Ш_{1m} describes the monomial interaction of the power m. It means that

$$P_i \text{Ш}_{1m} = P_i \text{Ш}_{1m} P_{i+m-2}. \tag{7.14}$$

Hence we have

$$
\begin{aligned}
P_i \text{Ш}_0^l \text{Ш}_{1m} &= P_i \text{Ш}_0^l P_{i-2l} \text{Ш}_{1m} \\
&= P_i \text{Ш}_0^l \text{Ш}_{1m} P_{i-2l+m-2}.
\end{aligned} \tag{7.15}
$$

The minimal power l for which the product $\text{Ш}_0^l \text{Ш}_{1m}$ fulfills (7.9) is

$$l = \left[\frac{m-2}{2}\right], \tag{7.16}$$

where

$$\left[\frac{m-2}{2}\right] = \begin{cases} \dfrac{m-2}{2} & \text{for even } m, \\[2mm] \dfrac{m-1}{2} & \text{for odd } m \end{cases}.$$

Then the terms in the lhs of (7.12) that do not fulfill (7.9), are as follows,

$$B_m = \sum_{l=0}^{[(m-2)/2]-1} \text{Ш}_0^l \text{Ш}_{1m}, \quad m = 3,\ldots,r. \tag{7.17}$$

Assuming that the inverse operator

$$B = \Big(I - \sum_{m=3}^{r} B_m\Big)^{-1} \tag{7.18}$$

exists we exclude all terms (7.17) from the Schwinger equation (7.12) and we get

$$\Big(I - B \sum_{m=3}^{r} \sum_{l=[(m-2)/2]}^{\infty} \text{Ш}_0^l \text{Ш}_{1m}\Big)X = B(I - \text{Ш}_0)^{-1}. \tag{7.19}$$

For a small coupling constant $g(\text{Ш}_1 = g\widetilde{\text{Ш}}_1)$ we may substitute the unit operator for B. We take into account only the first order terms in parenthesis. In this way, we arrive at the approximated Schwinger equation

$$\Big(I - \sum_{m=3}^{r} \sum_{l=[(m-2)/2]}^{\infty} \text{Ш}_0^l \text{Ш}_{1m}\Big)X \approx B(I - \text{Ш}_0)^{-1} \tag{7.20}$$

which admits decoupling for any operator P^i defined by (7.10). It is worth noting that to derive (7.20) we did not assume analytic properties of X or even B with respect to the coupling constant. We only assumed (6.14) and that for a small coupling constant the operator B in parenthesis is approximately equal to the unit operator. It means that Eqs. (7.20) can be especially useful in the case of nonanalytic dependence of X on a coupling constant. As a particular case we consider the monomial interaction with $m = r = 4$. From (7.16), $l = [(m-2)/2] = 1$ and from (7.17) we have,

$$B_m = B_4 = \text{Ш}_{14} = \text{Ш}_1. \tag{7.21}$$

From (7.21) and (7.18) we have

$$B = (I - \text{Ш}_1)^{-1}. \tag{7.22}$$

The equation (7.19) is now

$$
\begin{aligned}
&(I - (I - \text{Ш}_1)^{-1}[(I - \text{Ш}_0)^{-1} - I]\text{Ш}_1)X \\
&\quad = (I - \text{Ш}_1)^{-1}(I - \text{Ш}_0)^{-1}
\end{aligned} \tag{7.23}
$$

and by analogy to (7.20),

$$(\mathrm{I}-[(\mathrm{I}-Ш_0)^{-1}-\mathrm{I}]Ш_1)X \approx (\mathrm{I}-Ш_1)^{-1}(\mathrm{I}-Ш_0)^{-1}. \tag{7.24}$$

The generalization of Eq. (7.23) is as follows,

$$\left(\mathrm{I}-\left(\mathrm{I}-\sum_{i=0}^{l}Ш_0^iШ_1\right)^{-1}\left[(\mathrm{I}-Ш_0)^{-1}-\sum_{i=0}^{l}Ш_0^i\right]Ш_1\right)X = \left(\mathrm{I}-\sum_{i=0}^{l}Ш_0^iШ_1\right)^{-1}(\mathrm{I}-Ш_0)^{-1}, \tag{7.25}$$

which can be used to derive approximated decoupled equations for the polynomial interactions. In the case of, e.g., $r=m=4$ and $l>0$, the decoupling is of a higher order than one. It means that we may retain the higher order terms with respect to the coupling constant in a Neumann expansion of the operator $(\mathrm{I}-\Sigma_{i=0}^{l}Ш_0^iШ_1)^{-1}$ occurring in the lhs of (7.25) and have a further decoupled equation with respect to the projectors P^n.

B. Large coupling constant

Definition: We mean that then the term $Ш_1$ is dominated with respect to $Ш_0$ and that we may consider, instead of (5.11), the equation

$$(\mathrm{I}-Ш_1)X=\mathrm{I}. \tag{7.26}$$

We now transform Eq. (7.26) as follows,

$$(\mathrm{I}-Ш_1-Ш_1^+ + Ш_1^+Ш_1)X=(\mathrm{I}-Ш_1^+).$$

Omitting the linear terms and unit operator,

$$Ш_1^+Ш_1X=\mathrm{I}-Ш_1^+. \tag{7.27}$$

The vector version of that equation for

$$\mathscr{V}=X\mathrm{P}_0\mathbb{1}$$

is even simpler because $Ш_1^+\mathrm{P}_0=0$. For the monomial interaction (6.18) due to (6.19),

$$\mathrm{P}_iШ_1^+=\mathrm{P}_iШ_1^+\mathrm{P}_{i+n+2}. \tag{7.28}$$

Hence

$$\mathrm{P}_{i+n-2}Ш_1^+\mathrm{P}_iШ_1\mathrm{P}_{i+n-2}\mathscr{V}=\mathrm{P}_{i+n-2}P_0=0 \tag{7.29}$$

for $i=0,1,\cdots$. This is a decoupled equation for projections $\mathrm{P}_{i+n-2}\mathscr{V}$ for a large coupling constant.

VIII. THE FESHBACH EQUATIONS AND OTHER PROJECTIONS OF THE SCHWINGER EQUATION

A. Feshbach equations

The Feshbach equations which are widely known in atomic and nuclear physics, are equations for distinguished components of the stationary wavefunction[5] (see also Ref. 3, Chap. 4). They are used in the case of multidimensional projectors for each one of the projected equations (2.16a) or (2.16b) is not equivalent to the original equation (2.1) (see Sec. 2C). Then, if we want to consider only, e.g., PX projection, we have to eliminate the component QX from the second equation to get the Feshbach equation. For that purpose we may also use the operator (7.5).

We now derive the Feshbach equations for the Schwinger equation. For the sake of simplicity we use the vector version (5.8) of the Schwinger equation which is multiplied by the operator (7.5). In the considered case

$$\mathrm{B}=\mathrm{B}_\mathrm{Q}=\mathrm{I}-(\mathrm{I}-Ш)[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]^{-1} \tag{8.1}$$

and

$$[\mathrm{P},\mathrm{B}_\mathrm{Q}(\mathrm{I}-Ш)]=0. \tag{8.2}$$

From (8.1) and (8.2),

$$\begin{aligned}\mathrm{PB}_\mathrm{Q}(\mathrm{I}-Ш)&=\mathrm{PB}_\mathrm{Q}(\mathrm{I}-Ш)\mathrm{P}\\&=\mathrm{P}\{\mathrm{I}-(\mathrm{I}-Ш)[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]^{-1}\}\\&\quad\times(\mathrm{I}-Ш)\mathrm{P}\\&=\mathrm{P}(\mathrm{I}-Ш)\mathrm{P}-\mathrm{P}Ш\\&\quad\times[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]^{-1}Ш\mathrm{P}.\end{aligned} \tag{8.3}$$

Also taking into account that the vacuum

$$\mathbb{1}=\mathrm{P}_0\mathbb{1}, \tag{8.4}$$

and choosing the projector Q so that

$$\mathrm{QP}_0=0, \tag{8.5}$$

we get from (5.8) the following equation,

$$\mathrm{P}(\mathrm{I}-Ш)\mathrm{P}\mathscr{V}-\mathrm{P}Ш[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]^{-1}Ш\mathrm{P}\mathscr{V}=\mathrm{PP}_0\mathbb{1}. \tag{8.6}$$

This is the Feshbach equation for the $\mathrm{P}\mathscr{V}$ component of the functional $\mathscr{V}[\alpha]$ fulfilling the Schwinger equation. Contrary to the latter one they are not overlapping equations. For the projection $\mathrm{P}=\mathrm{P}_n$, we get the integral equation only for the n-point function, generated by the component

$$(\mathrm{P}_n\mathscr{V})[\alpha]=(1/n!)\int dx_{(n)}\,(\delta\mathscr{V}[\xi]/\delta\xi(x_1)\cdots\delta\xi(x_n))\alpha(x_1)\cdots\alpha(x_n)_{\xi=0}. \tag{8.7}$$

The serious problem is to find the inverse operator appearing in the Feshbach equation

$$\mathrm{X}_\mathrm{Q}=[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]^{-1}, \tag{8.8}$$

which by definition fulfills the equality

$$[\mathrm{Q}(\mathrm{I}-Ш)\mathrm{Q}]\mathrm{X}_\mathrm{Q}=\mathrm{Q}. \tag{8.9}$$

The fact that for the polynomial interaction only particular components of the operator X_Q occur in (8.6) is favorable condition. We therefore, may repeat our trick with the operator B, constructed now for Eq. (8.9), and derive the Feshbach equation for that equation. Assuming that

$$\mathrm{P}Ш=\mathrm{P}Ш(\mathrm{P}+\Delta\mathrm{P}), \tag{8.10a}$$

$$Ш\mathrm{P}=(\Delta\mathrm{P}'+\mathrm{P})Ш\mathrm{P}, \tag{8.10b}$$

and taking into account that for the projectors $\Delta\mathrm{P}$ and $\Delta\mathrm{P}'$

$$\Delta\mathrm{P}\cdot\mathrm{Q}=\mathrm{Q}\Delta\mathrm{P}=\Delta\mathrm{P}, \tag{8.11a}$$

$$\Delta\mathrm{P}'\mathrm{Q}=\mathrm{Q}\Delta\mathrm{P}'=\Delta\mathrm{P}', \tag{8.11b}$$

introducing, moreover, the complementary projector $\Delta\mathrm{Q}$ in the space QE,

$$\Delta\mathrm{Q}+\Delta\mathrm{P}=\mathrm{Q}, \tag{8.12}$$

for which

$$\Delta Q\cdot Q = Q\Delta Q = \Delta Q, \quad \Delta P\Delta Q = 0, \tag{8.13}$$

we get the following Feshbach equation for the projected Schwinger equation (8.9)

$$\begin{aligned}&\Delta\, P(I - \text{Ш})\Delta\, PX_Q\Delta\, P' \\ &\quad + \Delta\, P\text{Ш}[\Delta\, Q(I - \text{Ш})\Delta\, Q]^{-1}\text{Ш}\Delta\, PX_Q\Delta\, P' \\ &\quad = \Delta\, P\Delta\, P' + \Delta\, P\text{Ш}[\Delta\, Q(I - \text{Ш})\Delta\, Q]^{-1}\Delta\, P'. \end{aligned}\tag{8.14}$$

This is the equation for a part of the kernel occurring in the second term of the Feshbach equation (8.6) [see (8.8) and (8.10)]. It is important to note that the inverse operator appearing in (8.14) is defined in a smaller space than the analogous inverse operator appearing in (8.6).

B. Modified Tamm–Dancoff method

Although in the case of the Schwinger equation there are not exact theorems like the Fadeev theorems for three-body scattering, the following approximation method can be suggested to calculate the inverse $\Delta Q(I - \text{Ш})\Delta Q$. For

$$X_{\Delta Q} = [\Delta Q(I - \text{Ш})\Delta Q]^{-1} \tag{8.15}$$

we have the equation

$$[\Delta Q(I - \text{Ш})\Delta Q]X_{\Delta Q} = \Delta Q. \tag{8.16}$$

Because $\Delta Q\text{Ш}\Delta Q$ is already defined on a subspace, we divide it into two parts using a new pair of the complementary projectors

$$R_{\Delta Q} + T_{\Delta Q} = \Delta Q. \tag{8.17}$$

We get

$$(I - R_{\Delta Q}\text{Ш}_{\Delta Q} - T_{\Delta Q}\text{Ш}_{\Delta Q})X_{\Delta Q} = \Delta Q. \tag{8.18}$$

Hence, after the transformation we arrive at the equation

$$\begin{aligned}&(I - (\Delta Q - R_{\Delta Q}\text{Ш}_{\Delta Q})^{-1}T_{\Delta Q}\text{Ш}_{\Delta Q})X_{\Delta Q} \\ &\quad = (\Delta Q - R_{\Delta Q}\text{Ш}_{\Delta Q})^{-1}\Delta Q\end{aligned}\tag{8.19}$$

which we can try to solve by means of the perturbation method using the operator $T_{\Delta Q}\text{Ш}_{\Delta Q}$ as a perturbation. However, we have to find the inverse operator occurring in (8.19). There is a very useful formula that reduces a calculation of that inverse operator to the subspace RE of the initial space E on which the Schwinger equation is considered. It is

$$\begin{aligned}&(\Delta Q - R_{\Delta Q}\text{Ш}_{\Delta Q})^{-1} \\ &\quad = [(R_{\Delta Q} - R_{\Delta Q}\text{Ш}_{\Delta Q}R_{\Delta Q})^{-1} \\ &\qquad \times(R_{\Delta Q}\text{Ш}_{\Delta Q}T_{\Delta Q} + R_{\Delta Q} + T_{\Delta Q})]\end{aligned}\tag{8.20}$$

[see Eq. (8.17)]. We check (8.20), multiplying this equality from the right by

$$\begin{aligned}\Delta Q - R_{\Delta Q}\text{Ш}_{\Delta Q} &= R_{\Delta Q} + T_{\Delta Q} - R_{\Delta Q}\text{Ш}_{\Delta Q}R_{\Delta Q} \\ &\quad - R_{\Delta Q}\text{Ш}_{\Delta Q}T_{\Delta Q}.\end{aligned}\tag{8.21}$$

One can get a similar formula for the right division of the operator

$$\text{Ш}_{\Delta Q} = \text{Ш}_{\Delta Q}R_{\Delta Q} + \text{Ш}_{\Delta Q}T_{\Delta Q}. \tag{8.22}$$

From the above formulas it can be seen that at successive stages of calculation, the Schwinger operator is projected on fewer and fewer spaces, which can be decisive in obtaining successful results.

C. The Dyson equations

We consider further the interaction

$$\mathscr{W}[\beta] = \int d^{\,n}x\, dt e^{-t^2K^{-1}(0)/2}\tilde{\mathscr{L}}_{\text{int}}(t)e^{it\beta(x)}, \tag{8.23}$$

which takes into account the normal ordering. $\tilde{\mathscr{L}}_{\text{int}}$, is any one variable function, which for the polynomial interaction is a sum of derivatives of Dirac's function. We consider the interaction for which

$$\tilde{\mathscr{L}}_{\text{int}}(t) = g\delta^{(4)}(t), \quad (\phi^4). \tag{8.24}$$

Then

$$\mathscr{W}[\beta] = g\int d^{\,n}x(\beta^4(x) - 6K^{-1}(0)\beta^2(x)). \tag{8.25}$$

We have, therefore,

$$P_i\text{Ш}_1 = \begin{cases}P_i\text{Ш}_1(P_{i+2} + P_i) & \text{for } i>0, \\ 0 \quad \text{for } i=0\end{cases}, \tag{8.26a}$$

and

$$\text{Ш}_1P_i = \begin{cases}(P_{i-2} + P_i)\text{Ш}_1P_i & \text{for } i>2, \\ P_1\text{Ш}P_1 & \text{for } i=1, \\ 0 & \text{for } i=0\end{cases} \tag{8.26b}$$

We introduce the projectors

$$P = P^2 = P_0 + P_1 + P_2 \quad \text{and} \quad Q = Q^2 \tag{8.27}$$

for which $P^2 + Q^2 = I$. Eq. (8.5) is now fulfilled. For the interaction (8.25)

$$P_{2n+1}\mathscr{V} = 0 \quad \text{for } n = 0,1,2,\cdots \tag{8.28}$$

Hence from (8.6) we get

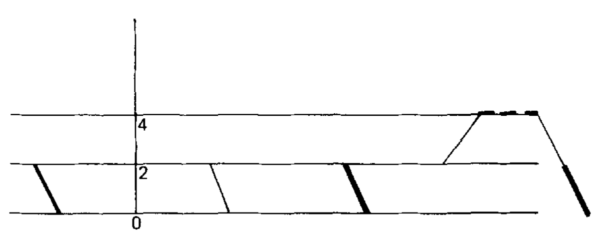
$$\begin{aligned}&P_2\mathscr{V} - P_2\text{Ш}_0P_0 - P_2\text{Ш}_1P_2\mathscr{V} \\ &\quad - P_2\text{Ш}_1P_4X_{Q'}P_4\text{Ш}_0P_2\mathscr{V} = 0\end{aligned}\tag{8.29}$$

This is the equation for the two-point function represented here by (8.7) for $n = 2$. This equation is similar to the Dyson equation but now the third and forth terms connected with the mass operator are given in a more manifest way, which facilitates applying different approximation methods like the one proposed in Sec. 8B.

For the zeroth order approximation

$$P_2\mathscr{V}^{(0)} = P_2XP_0 = P_2\text{Ш}_0P_0. \tag{8.30}$$

That justifies the following graphical representation for Eq. (8.29)



(8.29′)

where, e.g., the thick line corresponds to $P_2\mathscr{Y}$, the thin line with the same slope and level corresponds to its zeroth order approximation, and the disconnected line corresponds to the projected operator $P_4X_{Q^\cdot}P_4$. It is possible to derive a similar equation for the four-point function.

For that purpose we have to choose the projector

$$P^4 = P_0 + P_1 + P_2 + P_3 + P_4 \tag{8.31a}$$

and

$$Q^4 = I - P^4. \tag{8.31b}$$

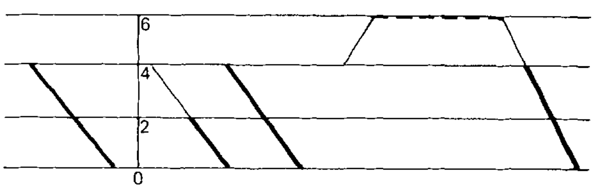
Then we rewrite (8.6) as

$$P^4\mathscr{Y} - P^4\text{Ш}P^4\mathscr{Y} - P^4\text{Ш}X_{Q^4}\text{Ш}P^4\mathscr{Y} = P_0. \tag{8.32}$$

Hence, after projection with P_4 we get

$$P_4\mathscr{Y} - P_4\text{Ш}_0P_2\mathscr{Y} - P_4\text{Ш}_1P_4\mathscr{Y} - P_4\text{Ш}P_6X_{Q^4}P_6\text{Ш}_0P_4\mathscr{Y} = 0. \tag{8.33}$$

In a graphic representation we have

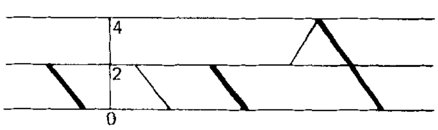


(8.33′)

To compare the Schwinger equation with Eq. (8.29) and Eq. (8.33) we multiply the Schwinger equation (5.8) by P_2 and we get

$$P_2\mathscr{Y} - P_2\text{Ш}_0P_0 - P_2\text{Ш}_1(P_2 + P_4)\mathscr{Y} = 0. \tag{8.34}$$

The graphic form for (8.34) is



(8.34′)

The system of equations (8.29′), (8.33′), and (8.34′) is not closed because no kernels presented by the disconnected lines are known in analytic form. Other relations can be obtained if we use the following projected Schwinger equation,

$$Q(I - \text{Ш})Q\mathscr{Y} + Q(I - \text{Ш})P\mathscr{Y} = 0. \tag{8.35}$$

This equation can be used to obtain the formula expressing the component $Q\mathscr{Y}$ by the component $P\mathscr{Y}$,

$$Q\mathscr{Y} = X_Q\text{Ш}P\mathscr{Y} \tag{8.36}$$

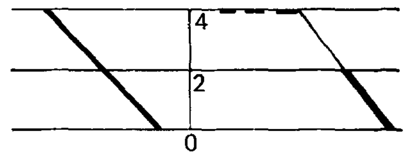
[see (8.8)]. Obviously, this equation cannot be considered alone because, in general we do not know the projection $P\mathscr{Y}$. However, it may establish a complement to the Schwinger equation (8.34). For the two-point function, using $Q = Q^2$, $P = P^2$, we get in the case of φ^4 interaction the following equation,

$$Q^2\mathscr{Y} = X_{Q^\cdot}P_4\text{Ш}_0P_2\mathscr{Y}. \tag{8.37}$$

Hence we have

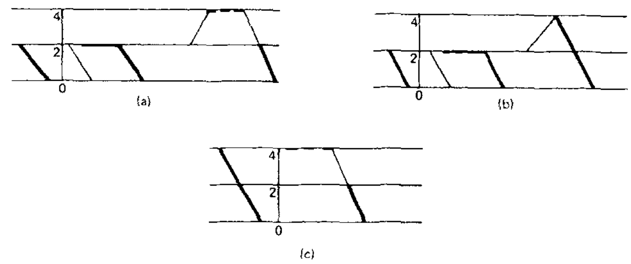
$$P_4\mathscr{Y} = P_4X_{Q^\cdot}P_4\text{Ш}_0P_2\mathscr{Y}. \tag{8.38}$$

In a graphic representation this becomes

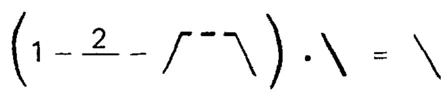


(8.38′)

Similar to (8.29), an unknown in an analytic form $P_4X_{Q^\cdot}P_4$ occurs. Collecting (8.29′) and (8.34′) we get three equations:

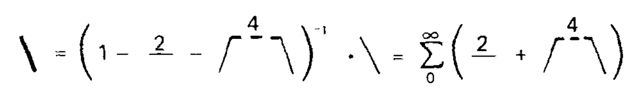


We see that they are not independent equations. We get (a) from (b) and (c). Introducing the corresponding graphic multiplication we can rewrite (a) as



(8.39)

Hence the solution is



The quantity --- is here represented by

$$P_4X_{Q^\cdot}P_4 = P_4[Q^2(I - \text{Ш})Q^2]^{-1}P_4. \tag{8.40}$$

This is an eight-point function which can be calculated by means of the same approximation technique proposed before.

D. Renormalization of the theory

Projected Schwinger equations in the vector version,

$$P(I - \text{Ш})P\mathscr{Y} - P\text{Ш}Q\mathscr{Y} = P\mathbb{1}, \tag{8.41}$$

$$Q(I - \text{Ш})Q\mathscr{Y} - Q\text{Ш}P\mathscr{Y} = Q\mathbb{1}, \tag{8.42}$$

are analogous to Eq. (2.16). The existence of the only one, perhaps formal, perturbation solution is the specific feature of the Schwinger equation. It means that only one of the equations (8.41) and (8.42), in the case of nontrivial projector, cannot be equivalent to the original Schwinger equation (5.8). In other words, for any nontrivial projector P, $P\mathscr{Y}$ cannot be freely fixed. On the other hand, to develop the divergentless perturbation theory we have to use the nontrivial projector P to eliminate some divergences. It turns out that $P\mathscr{Y}$ chosen in agreement with certain physical demands is inconsistent with the original Schwinger equation. This allows us to renormalize bare parameters of the theory. We now give some examples of projectors which can be used to remove divergences from the perturbation expansion or to embed additional phenomenological parameters in the Schwinger equation. We consider more general projectors than (4.2),

$$\mathrm{P}_n[\alpha,\beta] = \frac{1}{n!}\int dx_{(n)}\, dy_{(n)}\, \alpha(x_1)\cdots\alpha(x_n)$$

$$\times \mathrm{P}_n(x_{(n)},y_{(n)})\beta(y_1)\cdots\beta(y_n), \tag{8.43}$$

with

$$\int dz_{(n)}\, \mathrm{P}_n(x_{(n)},z_{(n)})\mathrm{P}_n(z_{(n)},y_{(n)}) = \mathrm{P}_n(x_{(n)},y_{(n)}). \tag{8.44}$$

Particular examples are projectors with

Example 1:

$$\mathrm{P}_n(x_{(n)},y_{(n)}) = \delta(\alpha_{(n)} - y_{(n)}), \tag{8.45}$$

where $a_{(n)} = (a_1,\ldots,a_n)$ are fixed. Of course (8.44) is fulfilled and

$$\mathrm{P}_n[\alpha,\beta] = \frac{1}{n!}\left[\int dx\alpha(x)\right]^n \beta(a_1)\cdots\beta(a_n). \tag{8.46}$$

We have for $\mathscr{V}$ [see (8.7)]

$$(\mathrm{P}_n\mathscr{V})[\alpha] = \frac{1}{n!}\frac{\delta^n\mathscr{V}[\zeta]}{\delta\zeta(a_1)\cdots\delta(a_n)}\bigg|_{\zeta=0}\left(\int dx\alpha(x)\right)^n. \tag{8.47}$$

Example 2: We construct here the projectors which exclude from n-point functions their first m terms in a Taylor expansion. We take

$$\mathrm{P}_n(x_{(n)},y_{(n)}) = \sum_{\substack{l=0\\ l_1+\cdots+l_n=l}}^{m} \frac{(x_1-a_1)^{l_1}}{l_1!}\cdots\frac{(x_n-a_n)^{l_n}}{l_n!}$$

$$\times \mathscr{D}^l\delta(a_{(n)} - y_{(n)}), \tag{8.48}$$

where

$$\mathscr{D}^l = \left(\frac{d}{dy_1}\right)^{l_1}\cdots\left(\frac{d}{dy_n}\right)^{l_n}.$$

Because of the coefficients $l_1!\cdots l_n!$, (8.44) is fulfilled. We have

$$\mathrm{P}_n[\alpha_1,\beta] = \frac{1}{n!}\sum_0^m(-1)^l\int dx_{(n)}\, dy_{(n)}$$

$$\times\frac{(x_1-a_1)^{l_1}}{l_1!}\alpha(x_1)\cdots\frac{(x_n-a_n)^{l_n}}{l_n!}\alpha(x_n)$$

$$\times\left(\frac{d}{da_1}\right)^{l_1}\beta(a_1)\cdots\left(\frac{d}{da_n}\right)^{l_n}\beta(a_n) \tag{8.49}$$

and

$$(\mathrm{P}_n\mathscr{V})[x] = \frac{1}{n!}\int dx_{(n)}\alpha(x_1)\cdots\alpha(x_n)$$

$$\times\sum_l^m\frac{(x_1-a_1)^{l_1}}{l_1!}\cdots\frac{(x_n-a_n)^{l_n}}{l_n!}\mathscr{D}^l\omega_n(a_{(n)}). \tag{8.50}$$

Similar formulas in momentum space are true. We have to change only all the quantities by their Fourier transforms and $x\to p$.

IX. IMPROVEMENT OF CALCULATIONS

In our paper we have used the Neumann series [see (5.14) and (5.17)] to express solutions of the considered equations. Unfortunately, we cannot resolve whether there are norms for which the spectral radius of the transformed or projected Schwinger equation is less than one. In such a situation we can only try a different method of calculation. Some variational principles have particular advantages in such cases.

A. Variational principle

The variational principle in the case of the linear equation

$$\mathrm{A}\mathscr{X} = \mathscr{Y}, \tag{9.1}$$

defined in the final space F to which $\mathrm{A}\mathscr{X}$ and $\mathscr{Y}$ belong, can be chosen as follows,

$$\mathrm{F}[\mathscr{X}] = \|\mathrm{A}\mathscr{X} - \mathscr{Y}\|. \tag{9.2}$$

It is easy to derive estimations [see (9.4)] which justify the functional (9.2) for the operator A with a bounded inverse. Having some approximated solution, which we label as $\mathscr{X}^*$, we define

$$\mathscr{Y}^* = \mathrm{A}\mathscr{X}^*. \tag{9.3}$$

Hence for the solution $\mathscr{X}$ we have

$$\mathscr{X} - \mathscr{X}^* = \mathrm{A}^{-1}(\mathscr{X} - \mathscr{X}^*)$$

and

$$\|\mathscr{X} - \mathscr{X}^*\| \leqslant \|\mathrm{A}^{-1}\|\mathrm{F}[\mathscr{X}^*]. \tag{9.4}$$

B. Modified equations

Sometimes we have to deal with Eq. (9.1) in which the left- and right-hand side are modified,

$$(\mathrm{A}+\mathrm{B})\mathscr{X}^* = \mathscr{Y} + \mathscr{Z}, \tag{9.5}$$

e.g., $\mathrm{B} = -\mathrm{QA}$ and $\mathscr{Z} = -\mathrm{Q}\mathscr{Y}$, where Q is a projector. We are interested now in the relation of the solution (9.1) to the solution (9.5). For the defined E, and bounded operators A and B, we have the following estimation,[6]

$$\frac{\|\mathscr{X} - \mathscr{X}^*\|}{\|\mathscr{X}\|} \leqslant \frac{\mu(\mathrm{A})}{1-\mu(\mathrm{A})\|\mathrm{A}\|/\|\mathrm{B}\|}\left(\frac{\|\mathrm{B}\|}{\|\mathrm{A}\|} + \frac{\|\mathscr{Z}\|}{\|\mathscr{Y}\|}\right), \tag{9.6}$$

where

$$\mu(\mathrm{A}) = \|\mathrm{A}\|\,\|\mathrm{A}^{-1}\|, \quad \|\mathrm{A}^{-1}\|\,\|\mathrm{B}\| < 1. \tag{9.7}$$

The inequality (8.17) enables us to approximate the occurring relative error if we are going from Eq. (9.1) to (9.5). For a small μ (e.g., $\|\mathrm{A}^{-1}$ large) even in the case of relatively large $\|\mathrm{B}\|$ and $\mathscr{Z}$ such that

$$\frac{\|\mathrm{B}\|}{\|\mathrm{A}\|} \approx 1, \quad \frac{\|\mathscr{Z}\|}{\|\mathscr{Y}\|} \approx 1, \tag{9.8}$$

the relative error is a small quantity. It means then, that in the case of (9.7) and (9.8), we can significantly change the form of the initial equation (9.1). For a large μ, (9.6) does not give any restriction for the relative error because the rhs can take a negative value.

Now let us observe that the rhs of the estimation (9.6) can also be used as a variational principle (see Sec. 9C). Taking

$$\mathrm{A} = \mathrm{A}_{\mathrm{eff}}, \quad \mathrm{B} = \mathrm{A} - \mathrm{A}_{\mathrm{eff}},$$

and

$$\mathscr{D} = 0, \tag{9.9}$$

we have to choose the parameters of A_{eff} or X_{app} [see (9.12)] so that

$$F[X_{app}] = \frac{\mu(A_{eff})}{1 - \mu(A_{eff})\|A - A_{eff}\|/\|A_{eff}\|} \cdot \frac{\|A - A_{eff}\|}{\|A_{eff}\|}, \tag{9.10}$$

at the condition $\|A_{eff}^{-1}\|\,\|A - A_{eff}\| < 1$, has the minimal value.

C. Effective operators

We consider here the *operator* equation

$$AX = Y,$$

with only one solution. If we know the solution X we may find the operator A describing that equation

$$A = YX^{-1}. \tag{9.11}$$

This circumstance enables us, in the case of any approximated solution of the operator equation, to find the corresponding effective operator

$$A_{eff} = YX_{app}^{-1}. \tag{9.12}$$

Now we can decompose the operator A in the original equation as follows,

$$(A_{eff} + A_1)X = Y, \tag{9.13}$$

where

$$A_1 = A - A_{eff}. \tag{9.14}$$

The standard form for (9.13) is obtained (Sec. 2) if we multiply (9.13) by A_{eff}^{-1}. We get then [see (9.12)]

$$(I + X_{app}Y^{-1}A_1)X = X_{app}. \tag{9.15}$$

X. SIMILARITY TRANSFORMATIONS OF THE SCHWINGER EQUATION

By means of any nonsingular operator U we may transform the Schwinger equation

$$(I - \text{Ш})\mathscr{V} = \mathbb{1},$$

as follows

$$(I - \text{Ш}')\mathscr{V}' = U\mathbb{1}, \tag{10.1}$$

where the transformed Schwinger operator

$$\text{Ш}' = U\text{Ш}U^{-1} \tag{10.2}$$

and

$$\mathscr{V}' = U\mathscr{V}. \tag{10.3}$$

A. Symmetrical transformations

In the case of the Schrödinger equation, symmetric transformations are defined by the unitary U which does not change the shape of the equation. Such a definition can be realized here if

$$\text{Ш}' = \text{Ш} \tag{10.4}$$

and

$$U\mathbb{1} = \mathbb{1}. \tag{10.5}$$

The last condition may be interpreted as the invariance of the vacuum. When there exists only one solution of the Schwinger equation (10.4) and (10.5), this means that the generating functional is a scalar

$$\mathscr{V}' = \mathscr{V}. \tag{10.6}$$

For the simplest transformations U given in the matrix form

$$U[\alpha,\beta] = e^{\alpha\Lambda\beta}, \tag{10.7}$$

(10.6) means

$$\mathscr{V}'[\alpha] = U[\alpha,\delta/\delta\xi]\mathscr{V}[\xi]|_{\xi=0} = \mathscr{V}[\Lambda\alpha] = \mathscr{V}[\alpha], \tag{10.8}$$

where Λ describes the space–time and internal symmetricity of the theory. Now $U\mathbb{1} = \mathbb{1}$ and

$$U^{-1}[\alpha,\beta] = e^{\alpha\Lambda^{-1}\beta}. \tag{10.9}$$

B. Amputated transformations

With the functional $\mathscr{V}$ is correlated the amputated functional

$$\mathscr{S} = U\mathscr{V}, \tag{10.10}$$

with

$$U[\alpha,\beta] = e^{\alpha K\beta}, \tag{10.11}$$

where K is defined by the differential operators occurring in the free part of the field equations.

The S matrix is constructed from the functional $\mathscr{S}$ as follows,[3]

$$S[\alpha,\beta] = \mathscr{S}[K\gamma]|_{\gamma = q_0[x;\tilde{\alpha},\tilde{\beta}]},$$

where $q_0[x;\widetilde{\alpha,\beta}]$ are general solutions of the classical field equations. Denoting in the case of (10.11),

$$\text{Ш}' = \text{Ш}^a, \tag{10.12}$$

we get for the functional $\mathscr{S}$

$$(I - \text{Ш}^a)\mathscr{S} = \mathbb{1}. \tag{10.13}$$

It is important that the projection properties of the operator Ш^a are the same as Ш because of the diagonal form of the transformation (10.11) and U^{-1}. We have, e.g.,

$$U = \sum_i P_i U P_i. \tag{10.14}$$

C. The general symmetric transformations

By the general symmetric transformations, we mean here any similarity transformations which do not change the physical content of the Schwinger equation. If we confine ourselves to the mass shell, the transformations which do not change the S matrix elements are symmetric transformations. Such transformations have been considered by Rzewuski.[7] More general transformations are obtained if the invariance concerns only absolute values of the transition amplitudes or if it concerns only a finite amount of the amplitudes.

[1]M.A. Krasnosielski *et al.*, *Priblizhennye rasheniya operatornych uravneniy* (Moscow, 1969), Chap. II, Sec. 7.
[2]T. Kato, *Perturbation Theory for Linear Operators* (Berlin, 1966), p. 20.
[3]J. Rzewuski, *Field Theory* (Name of publisher, London, 1969), Vol. II.
[4]W. Garczyński and J. Hańćkowiak, 1978, Preprint.
[5]H. Feshbach, Ann. Phys. (N.Y.) **19**, 287 (1962).
[6]M.K. Gavurin, *Lekcii po metodam vychisleniy* (Moskva, 1971).
[7]J. Rzewuski, Acta Phys. Polon. **31**, 19 (1967).

An application of the maximum principle to the study of essential self-adjointness of Dirac operators. I

J. J. Landgren

Department of Mathematics, University of Tennessee, Chattanooga, Tennessee 37402

P. A. Rejto[a)]

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

(Received 25 October 1978)

We formulate a theorem saying that the Dirac operator corresponding to a one-electron atomic ion is essentially self-adjoint on the usual domain, provided that the nuclear charge Z is less than 118. Furthermore, for such nuclear charges the domains of the closure of the free particle and total Dirac operators are equal. In the present part I of this paper we prove this theorem for the part of the operator over each of the usual reducing subspaces.

1. INTRODUCTION

Rellich showed,[1,2] that the Dirac operator corresponding to a one-electron atomic ion is essentially self-adjoint provided that the nuclear charge Z is less than 118. Independently of Rellich, using the Rellich–Kato perturbation theorem, Kato showed [3c] that this is the case for Z less than 67. At the same time he emphasized that the domain of the closure of this operator equals the domain of the closure of the free particle Dirac operator. The aim of this paper is to show domain preserving essential self-adjointness for the entire range of values considered by Rellich,[1,2] According to the recent result of Parthasaraty[4] for Z outside the Kato range the bound of the potential with respect to the free particle operator is no longer less than one and so the Rellich–Kato perturbation theorem does not apply to them. Nevertheless, we show that the free particle operator can be replaced by an approximate operator which was introduced by Rellich and Weidmann in connection with their study of the separated Dirac operator.

In Sec. 2, in the main Theorem 2.1 we describe this domain preserving essential self-adjointness in specific terms. Then in Corollary 2.1 we show that domain preserving essential self-adjointness also holds if we add a perturbation whose bound with respect to the free particle Dirac operator is zero. An interesting special case of Corollary 2.1 is the case of potential perturbations. Therefore, in Corollary 2.2 with the aid of the Stummel bound[5,6] we formulate conditions on the potential which ensure that the corresponding operator has bound zero with respect to the free particle Dirac operator. Corollary 2.2 is interesting inasmuch as it gives a common generalization of the results of Rellich,[1] Kato,[3c] and Evans,[7,8] Corollary 2.1 also allows us to extend a result of Gross.[9] This extension is formulated in Corollary 2.3. Note that this extension is based on the Gross criterion for a potential to have bound zero with respect to the free particle Dirac operator. For various overlapping classes of such potentials we refer to the works of Prosser,[10] Thompson,[11] Eckhardt,[12] Guillot-Schmitt,[13] and Weder.[14]

[a)]Supported by NSF grant MCS78-02199.

In the short abstract Sec. 3 we show that Theorem 2.1 implies Corollary 2.1.

In Sec. 4 we prepare for the proof of Theorem 2.1. We start this preparation by emphasizing that the Dirac operator admits a complete family of reducing subspaces, on each of which it acts like a system of ordinary differential operators.[15,16,17,18] We study the Dirac operator by studying its restrictions to these reducing subspaces. For brevity we call these restrictions the parts of the operator. In Theorem 4.1 we show that Theorem 2.1 holds for each of these parts of the Dirac operator. Then we show that Theorem 4.1 implies the essential self-adjointness conclusion of the main Theorem 2.1. To show the domain preservation conclusion we need our estimates for the parts of the operator to be uniform in the reducing subspace parameter. This is implied by Theorem 4.2.

In Sec. 5 we prove Theorem 4.1. This proof makes essential use of the approximate potential of definition (5.4) which was introduced by Rellich,[1,2] and Weidmann.[19] The difficult part of the proof of Theorem 4.1 is isolated in Theorem 5.1. It says that zero is in the resolvent set of the approximate operator and that the part of the Coulomb operator is bounded with respect to this approximate operator. We start the proof of Theorem 5.1 by proving that the formal part of the Weyl–Weidmann[19] construction can be carried out for this approximate operator. This is the statement of Lemma 5.1. Then in Lemma 5.2, using a result of Schur–Holmgren–Carleman[20] we show that the kernel of Lemma 5.1 defines a bounded operator. Hence, the formal construction of Lemma 5.1 gives the resolvent kernel. Finally, in Lemma 5.3 we show that the part of the Coulomb potential times the resolvent of the approximate operator is also bounded. Lemma 5.3 is the key part of the proof of Theorem 5.1. We prove it again by using a result of Schur–Holmgren–Carleman.[20]

Theorem 4.2 will be proved with the aid of the maximum principle in the second part of this paper.

Finally we note that our Corollary 2.1 overlaps with an elegant result of Schmincke.[21] For related work on Dirac

 0022-2488/79/112204-08$01.00

operators we refer to the excellent paper of Kalf–Schmincke–Walter–Wüst[22] and to the recent paper of Nenciu.[23]

2. FORMULATION OF THE RESULT

Let $g_{\mu\nu}$ be the components of the Lorentz matrix, that is,

$$g_{00} = 1, \quad g^{11} = g^{22} = g^{33} = -1,$$

and

$$g^{\mu\nu} = 0 \quad \text{for } \mu \neq \nu, \quad \mu,\nu = 0,1,2,3.$$

Next let $\mathscr{C}_4$ be the four-dimensional complex Euclidean space. The Dirac matrices $\{\gamma^\mu\}$ are then defined as a set of unitary matrices acting on $\mathscr{C}_4$, which satisfy the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\mu\gamma^\nu = 2g^{\mu\nu}I.$$

As usual we define $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$ to be the space of $\mathscr{C}_4$-valued infinitely differentiable complex-valued functions with compact support on $\mathscr{R}_3$, the three-dimensional real Euclidean space. We denote partial derivatives by

$$D_k f(x) = \frac{\partial f}{\partial x_k}(x), \quad f \in \mathfrak{C}_0^\infty(\mathscr{R}_3), \quad k = 1,2,3.$$

Similar to Kato[3c] and Reed-Simon,[8] we define the free particle Dirac operator to be the closure of

$$H(0) = \sum_{k=1}^{3} \gamma^0\gamma^k \frac{1}{i} D_k + \gamma^0 \quad \text{on } \mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{2.1}$$

Next set

$$r(x) = (x_1^2 + x_2^2 + x_3^2)^{1/2} \tag{2.2}$$

and

$$M\left(\frac{1}{r}\right)f(x) = \frac{1}{r(x)} f(x), \quad f \in \mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{2.3}$$

Then in a one-electron atomic ion, the interaction between the electron and the nucleus is given by

$$V_0(e) = -eM(1/r) \quad \textit{on } \mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{2.4}$$

Here e is the adjusted nuclear charge defined by $e = Z\alpha$, where Z is the nuclear charge and α is the fine structure constant. This scaling was introduced by Kato.[3c]

With the aid of these notations, the one-electron Dirac operator corresponding to an atomic ion with adjusted nuclear charge e is defined to be

$$H(e) = H(0) + V(e) \quad \text{on } \mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{2.5}$$

It is an elementary fact that the matrices $\{\gamma^0\gamma^k\}$ and γ^0 are symmetric. This shows that the operators $H(0)$ and $H(e)$ are also symmetric. The aim of this paper is to show that *they are essentially self-adjoint and that the domains of the closures of $H(0)$ and $H(e)$ are equal* provided that e satisfies the Rellich[1,2] condition. This is the statement of our main theorem which follows. In it for a given operator T acting in an abstract Hilbert space $\mathfrak{a}$ we denote by $\rho(T)$ its resolvent set. Following Dunford–Schwartz,[24] for a given complex number we set

$$R(\mu,T) = (\mu I - T)^{-1}, \quad \mu \in \rho(T).$$

We shall also make use of the well known fact that a given symmetric operator A is self-adjoint, if and only if and only if $\pm i$ are in $\rho(A)$.[24c,3d] For brevity we denote an operator and its closure by the same symbol.

Theorem 2.1: *Let the operators $H(0)$ and $H(e)$ be defined by relations* (2.1) *and* (2.5). *Suppose that the real number e satisfies*

$$e \in (-\sqrt{3}/2, \sqrt{3}/2). \tag{2.6}$$

Then the operator $H(e)$ is essentially self-adjoint on $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$; equivalently its closure is such that

$$\pm i \in \rho(H(e)). \tag{2.7}$$

Furthermore,

$$H(0)R(\pm i, H(e)) \in \mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_3,\mathscr{C}_4)), \tag{2.8}$$

and hence

$$\mathfrak{D}(H(e)) = \mathfrak{D}(H(0)). \tag{2.9}$$

Next we formulate a corollary of the main Theorem 2.1. In it we assume that the operator V_1 has $H(0)$ bound zero. It is not difficult to see that this is equivalent to

$$\lim_{\mu \to \pm\infty} \|V_1 R(i\mu, H(0))\| = 0. \tag{2.10}$$

Corollary 2.1: *Suppose that the real number e satisfies assumption* (2.6) *and that the operator V_1 is symmetric and satisfies assumption* (2.10). *Then the operator $H(e) + V_1$ is essentially self-adjoint on $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$; equivalently its closure is such that*

$$\pm i \in \rho(H(e) + V_1). \tag{2.11}$$

Furthermore,

$$\mathfrak{D}(H(e) + V_1) = \mathfrak{D}(H(e)). \tag{2.12}$$

An interesting special case of Corollary 2.1 is the case of a potential. Accordingly, we assume that V_1 is of the form

$$V_1 f(x) = p(x)f(x), \quad f \in \mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4), \tag{2.13}$$

where p is a real valued function in $\mathfrak{L}_{2,\mathrm{loc}}(\mathscr{R}_3)$. To describe conditions which ensure that this operator has $H(0)$ bound zero we need the Stummel bound of p.[5,6] For a given pair of positive numbers, say (α,k), this is defined by

$$S(\alpha,k)(p) = \sup_{x \in \mathscr{R}_3} \mathrm{SM}(\alpha,k)(p)(x), \tag{2.14}$$

where

$$\mathrm{SM}(\alpha,k)(p)(x) = \int_{|x-y|<1} \frac{|p(y)|^k}{|x-y|^\alpha} dy. \tag{2.15}$$

In the following corollary we formulate an essential self-adjointness criterion for such potential perturbations with the aid of the Stummel bound.

Corollary 2.2. *Suppose that the real number e satisfies assumption* (2.6) *and that the operator V_1 is symmetric and of the form* (2.13). *Suppose, further, that to the corresponding potential p there is a number α such that*

$$\alpha > 1 \quad \textit{and} \quad \mathrm{S}(\alpha,2)(p) < \infty. \tag{2.16}$$

Then the operator $H(e) + V_1$ is essentially self-adjoint on $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$; equivalently conclusion (2.11) *holds for its closure. Furthermore, conclusion* (2.12) *holds.*

Corollary 2.2 is interesting inasmuch as it gives a common generalization of results of Rellich [1] Kato [3c] and Evans. [7] For the case of $p = 0$ it extends the Rellich result, inasmuch as Rellich did not have conclusion (2.12). It extends the Kato result inasmuch as instead of assumption (2.6), Kato had the more restrictive assumption

$$e\in(-\tfrac{1}{2},\tfrac{1}{2}). \tag{2.17}$$

For the case of $p\neq 0$ it extends the Evans result, inasmuch as Evans[7] used the Kato assumption (2.17) instead of our assumption (2.6). At the same time it extends the Evans assumption on the potential p. Specifically, he assumed that for each positive number ρ,

$$\sup_{|x|\leqslant\rho} \mathrm{SM}(\alpha,2)(p) < \infty, \quad \alpha > 1 \tag{2.18}$$

and that

$$\limsup_{|x|\to\infty} |p(x)| < \infty. \tag{2.19}$$

It is not difficult to show that the Evans assumptions (2.18) and (2.19) imply our assumption (2.16).

Evans also showed that his assumptions imply that the operator V_1 of definition (2.13) has $H(0)$ bound zero. A straightforward adaptation of his proof shows that in this implication one can replace his assumptions by our assumption (2.16). Therefore, Corollary 2.2 is a consequence of Corollary 2.1.

Corollary 2.1 also allows us to extend a result of Gross.[9] For this extension we need one of his results.[9] This says that

$$p\in\mathfrak{L}_3(\mathscr{R}_3) \tag{2.20}$$

implies that the corresponding operator of definition (2.13) has $H(0)$ bound zero. Combining this implication with Corollary 2.1 yields the following.

Corollary 2.3: *Suppose that the real number e satisfies assumption* (2.6) *and that the operator* V_1 *is symmetric and of the form* (2.13). *Suppose, further, that the corresponding potential p satisfies assumption* (2.20). *Then the operator* $H(e) + V_1$ *is essentially self-adjoint on* $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$; *equivalently conclusion* (2.11) *holds for its closure. Furthermore, conclusion* (2.12) *holds.*

3. THEOREM 2.1 IMPLIES COROLLARY 2.1

In this short section we show that the main Theorem 2.1 implies Corollary 2.1. To show this, we need the fact that for each self-adjoint operator A and for each real μ, [3a]

$$\|R(i\mu, A)\| \leqslant 1/|\mu|. \tag{3.1}$$

Clearly,

$$(iI - A)R(i\mu, A) = I + i(1-\mu)R(i\mu, A).$$

Inserting relation (3.1) in this relation shows that

$$\limsup_{\mu\to\pm\infty} \|(iI - A)R(i\mu, A)\| < \infty. \tag{3.2}$$

Conclusion (2.7) allows us to apply relation (3.2) to the operator $H(e)$. This yields

$$\limsup_{\mu\to\pm\infty} \|(iI - H(e))R(i\mu,H(e))\| < \infty. \tag{3.3}$$

Combining relation (3.3) with conclusion (2.8) we find that

$$\limsup_{\mu\to\pm\infty} \|H(0)R(i\mu,H(e))\| < \infty. \tag{3.4}$$

Relation (3.4) and another application of the abstract relation (3.1) together show that

$$\limsup_{\mu\to\pm\infty} \|(i\mu I - H(0))R(i\mu,H(e))\| < \infty. \tag{3.5}$$

Inserting relation (3.5) in assumption (2.10) we arrive at

$$\lim_{\mu\to\pm\infty} \|V_1 R(i\mu,H(e))\| = 0. \tag{3.6}$$

Relation (3.6) allows us to apply the Rellich–Kato theorem [3b] to the operators $H(e)$, and $H(e) + V_1$ and conclude the validity of Corollary 2.1.

4. PREPARATIONS FOR THE PROOF OF THEOREM 2.1

Our proof of Theorem 2.1 will make essential use of the fact that the operator of definition (2.5) admits a complete family of reducing subspaces on each of which it acts like a system of ordinary differential operators. The proof of this fact is similar to the proof of the well known fact that the corresponding eigenvalue equation can be solved by separation of variables.[15,16,18] This proof is implicit in the work of Rellich[1] and for completeness it was also given elsewhere.[17]

To describe these systems of ordinary differential operators we introduce some notations. First we let D denote differentiation;

$$Df(\rho) = f'(\rho), \quad f\in\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2), \tag{4.1}$$

and we set

$$J = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix}. \tag{4.2}$$

Secondly for a given matrix valued function P in $\mathfrak{L}_{2,\mathrm{loc}}(\mathscr{R}_+,\mathscr{C}_2)$ we let $M(P)$ denote the corresponding multiplication operator;

$$M(P)f(\rho) = P(\rho)f(\rho), \quad f\in\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2) \tag{4.3}$$

Then we define the operator $L(P)$ to be the closure of

$$L(P) = JD + M(P) \quad \text{on } \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2) \tag{4.4}$$

Thirdly, for each pair of real numbers (e,κ) we set

$$C_0(e,\kappa) = \begin{pmatrix} -e & \kappa \\ \kappa & -e \end{pmatrix}, \tag{4.5}$$

and

$$C_\infty = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.6}$$

Then define the matrix valued function $P(e,\kappa)$ by

$$P(e,\kappa)(\rho) = C_\infty + \frac{1}{\rho} C_0(e,\kappa), \quad \rho\in\mathscr{R}_+. \tag{4.7}$$

These notations allow us to describe the parts of the operator $H_0(e)$. The statement which follows is implied by a result formulated elsewhere.[17] *There is a family of mutually orthogonal orthoprojectors,* $\{O(\kappa)\}$, *such that*

$$\sum_{\kappa=-\infty}^{\kappa=+\infty} O(\kappa) = I \quad \text{on } \mathfrak{L}_2(\mathscr{R}_3,\mathscr{C}_4) \tag{4.8}$$

and for each $\kappa = \pm 1, \pm 2, \cdots$

$$O(\kappa)H(e) = H(e)O(\kappa) \quad \text{on } O(\kappa)\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{4.9}$$

Furthermore, there is a family of partial isometries $\{U(\kappa)\}$ *mapping* $\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4)$ *onto* $\mathscr{C}_{2|\kappa|}\otimes\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2)$ *such that*

$$U^*(\kappa)U(\kappa)=O(\kappa), \tag{4.10}$$

and

$$U(\kappa)H(e)=(I\otimes L(P(e,\kappa)))U(\kappa) \quad \text{on} \quad O(\kappa)\mathfrak{C}_0^\infty(\mathscr{R}_3,\mathscr{C}_4). \tag{4.11}$$

Here, I is the identity operator on $\mathscr{C}_{2|\kappa|}$. Roughly speaking these relations say that the part of the operator $H(e)$ over the reducing subspace $O(\kappa)\mathfrak{L}_2(\mathscr{R}_3,\mathscr{C}_4)$ is unitarily equivalent to $2|\kappa|$ copies of the operator $L(P(e,\kappa))$.

As a first step in the proof of the main Theorem 2.1 we show that it holds for each part of the operator $H(e)$. The validity of conclusion (2.7) for these parts is a result of Rellich[1] and Weidman.[19] We state it for future reference,

$$\pm i\in\rho(L(P(e,\kappa))), \quad \kappa=\pm 1,\pm 2,\cdots. \tag{4.12}$$

The validity of the other two conclusions for these parts is the statement of the theorem which follows.

Theorem 4.1: *Let the operators* $L(P(0,\kappa))$ *and* $L(P(e,\kappa))$ *be defined by relations* (4.4) *and* (4.7). *Suppose that the real number e satisfies assumption* (2.6), *Then for each* $\kappa=\pm 1,\pm 2,\cdots$,

$$L(P(0,\kappa))R(\pm i,L(P(e,\kappa)))\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)), \tag{4.13}$$

and hence

$$\mathfrak{D}(L(P(e,\kappa)))=\mathfrak{D}(L(P(0,\kappa))). \tag{4.14}$$

As a second step in the proof of the main Theorem 2.1 we show that Theorem 4.1 implies conclusion (2.7). To do this we note the elementary fact that the inverse of an orthogonal sum equals the orthogonal sum of the inverses. As is well known[25] the norm of an orthogonal sum equals the supermum of the norms. These two facts and relations (4.11) and (4.12) together show that each of the two operators $(\pm iI-H(e))$ admits a possibly unbounded inverse for which,

$$\|(\pm iI-H(e))^{-1}\|=\sup_\kappa\|R(\pm i,L(P(e,\kappa)))\|. \tag{4.15}$$

Since the operator $L(P(e,\kappa))$ is symmetric relation (4.12) allows us to apply the abstract relation (3.1) to it. This yields

$$\|R(\pm i,L(P(e,\kappa)))\|\leqslant 1.$$

Inserting this estimate in relation (4.15) proves conclusion (2.7).

As a third step in the proof of the main Theorem 2.1 we formulate a theorem which implies conclusion (2.8). For this purpose we note that according to relation (4.9) the reducing subspaces are independent of the parameter e. Hence, similarly to relation (4.15),

$$\|H(0)R(\pm i,H(e))\|=\sup_\kappa\|L(P(0,\kappa))\times R(\pm i,L(P(e,\kappa)))\|. \tag{4.16}$$

To estimate the supermum on the right recall definitions (4.3), (4.4), and (4.7). They show that if for $P(\rho)=\rho$ we set $M(P)=M$ then

$$L(P(e,\kappa))=L(P(0,\kappa))-eM^{-1}. \tag{4.17}$$

In the following theorem we show that the $L(P(0,\kappa))$ bound of M^{-1} is small for large $|\kappa|$.

Theorem 4.2: *Under the assumptions of Theorem* 2.2,

$$\lim_{|\kappa|\to\infty}\|M^{-1}R(\pm i,L(P(0,\kappa)))\|=0. \tag{4.18}$$

We shall prove Theorem 4.2 in the forthcoming part II of this paper. At present we show that Theorems 4.1 and 4.2 together imply conclusion (2.8) of the main Theorem 2.1.

Theorem 4.2 and relation (4.17) together with the Rellich–Kato[3b] theorem show that for large enough $|\kappa|$,

$$R(\pm i,L(P(e,\kappa))=R(\pm i,L(P(0,\kappa)))[I+eM^{-1}\times R(\pm i,L(P(0,\kappa)))]^{-1}. \tag{4.19}$$

At the same time Theorem 4.2 together with the usual estimate for the norm of a convergent Neumann series[5] yields

$$\limsup_{|\kappa|\to\infty}\|[I+eM^{-1}R(\pm i,L(P(0,\kappa)))]^{-1}\|<\infty. \tag{4.20}$$

Combining relations (4.19) and (4.20) we obtain

$$\limsup_{|\kappa|\to\infty}\|L(P(0,\kappa))R(\pm i,L(P(e,\kappa)))\|<\infty. \tag{4.21}$$

Inserting estimate (4.21) and conclusion (4.13) of Theorem 4.1 in relation (4.16) we arrive at the validity of conclusion (2.8) of Theorem 2.1.

Since we defined the domains of our operators by closure conclusion (2.9) is an immediate consequence of conclusion (2.8).

5. THE PROOF OF THEOREM 4.1

In this section we prove Theorem 4.1. We start this proof with a reformulation of conclusion (4.13). Specifically, we claim that it is equivalent to

$$M^{-1}R(\pm i,L(P(e,\kappa)))\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)). \tag{5.1}$$

To see this we note that relation (4.17) implies the following version of the second resolvent equation[26]

$$\begin{aligned} R(\pm i,L(P(e,\kappa))) &= R(\pm i,L(P(0,\kappa)))-R(\pm i,L(P(0,\kappa)))\\ &\quad\times eM^{-1}R(\pm i,L(P(e,\kappa))) \end{aligned} \tag{5.2}$$

on $(\pm iI-L(P(e,\kappa)))\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2)$.

As is well known relation (4.12) implies that

$$(\pm iI-L(P(e,\kappa)))\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2) \quad \textit{is dense in} \quad \mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2). \tag{5.3}$$

Relations (5.2) and (5.3) together show the equivalence of relation (5.1) and conclusion (4.13).

To prove relation (5.1) we employ an approximate potential which was introduced by Rellich[1,2] and Weidmann.[19] Specifically, we set

$$Q(e,\kappa)(\rho)=\begin{cases}(1/\rho)C_0(e,\kappa) & \rho\in(0,1)\\ C_\infty & \rho\in(1,\infty).\end{cases} \tag{5.4}$$

Weidmann used this approximate potential in his proof of the self-adjointness of the original operator $L(P(e,\kappa))$. His proof also shows the self-adjointness of the approximate operator $L(Q(e,\kappa))$. Hence each nonreal complex number is in $\rho(L(Q(e,\kappa)))$. In the theorem which follows we show that zero is also in this resolvent set and that M^{-1} is bounded with respect to $L(Q(e,\kappa))$. At the same time this theorem isolates the technical part of the proof of relation (5.1).

Theorem 5.1: *Suppose that the real number e satisfies assumption* (2.6). *Let the Rellich–Weidmann approximate potential $Q(e,\kappa)$ be defined by relation* (5.4). *Then the closure of the corresponding operator of definition* (4.4) *is such that*

$$O\in\rho(L(Q(e,\kappa))), \tag{5.5}$$

and

$$M^{-1}R(O,L(Q(e,\kappa)))\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)). \tag{5.6}$$

We start the proof of conclusion (5.5) by showing that for this operator the formal part of the Weyl–Weidmann construction[19] can be carried out. This is the statement of the lemma which follows.

Lemma 5.1: *Suppose that e satisfies assumption* (2.6). *Then for each $\kappa=\pm1,\pm2,\cdots$, the equation,*

$$Jk'(e,\kappa)+Q(e,\kappa)k(e,\kappa)=0 \tag{5.7}$$

admits two linearly independent solutions, $k^l(e,\kappa)$ and $k^r(e,\kappa)$ such that

$$k^l(e,\kappa)\in\mathfrak{L}_2(0,1) \tag{5.8a}$$

and

$$k^r(e,\kappa)\in\mathfrak{L}_2(1,\infty). \tag{5.8b}$$

In the proof of this and the following lemmas we omit the parameter κ. To prove conclusions (5.8) note that according to definition (5.4) for any two points ρ_1 and ρ_2 which are both in the interval $(0,1)$ or $(1,\infty)$,

$$JQ(e)(\rho_1)JQ(e)(\rho_2)=JQ(e)(\rho_2)JQ(e)(\rho_1). \tag{5.9}$$

This commutation property allows us to give a fundamental matrix of Eq. (5.7) in each of these two intervals. In fact, elementary algegra yields,

$$Y^l(e)(\rho)=\exp[-\ln\rho JC_0(e)],\quad \rho\in(0,1), \tag{5.10a}$$

and

$$Y^r(e)(\rho)=\exp[-(\rho-1)JC_\infty],\quad \rho\in(1,\infty). \tag{5.10b}$$

Definitions (4.3), (4.5), and (4.6) show that

$$JC_0(e,\kappa)=\begin{pmatrix}-\kappa & e\\ -e & \kappa\end{pmatrix} \tag{5.11}$$

and

$$JC_\infty=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}. \tag{5.12}$$

These formulas, in turn, show that the spectra of these matrices are given by

$$\sigma(JC_0(e))=\{-(\kappa^2-e^2)^{1/2},+(\kappa^2-e^2)^{1/2}\} \tag{5.13}$$

and

$$\sigma(JC_\infty)=\{-1,+1\}. \tag{5.14}$$

Hence there are unit vectors $a^\pm$, $b^\pm$ in $\mathscr{C}_2$ such that

$$-JC_0(e)a^\pm=\pm(\kappa^2-e^2)^{1/2}a^\pm \tag{5.15$\pm$}$$

and

$$-JC_\infty b^\pm=\pm b^\pm. \tag{5.16$\pm$}$$

Inserting these definitions in formulas (5.10) we find that the functions

$$y^l(e)(\rho)=\rho^{(\kappa^2-e^2)^{1/2}}a^+ \tag{5.17a}$$

and

$$y^r(\rho)=\exp[-(\rho-1)]b^-, \tag{5.17b}$$

satisfy Eq. (5.7) on the interval $(0,1)$ and $(1,\infty)$, respectively. Finally define $k^l(e)$ and $k^r(e)$ by the requirements that they satisfy Eq. (5.7) over all of $\mathscr{R}_+$ and

$$k^l(e)(\rho)=y^l(e)(\rho),\quad \rho\in(0,1) \tag{5.18a}$$

and

$$k^r(e)(\rho)=y^r(\rho)\quad \rho\in(1,\infty). \tag{5.18b}$$

Then it is clear from these definitions that for these functions conclusions $(5.8)^l$ hold. To see that they are linearly independent we evaluate their determinant at the point $\rho=1$. Since the functions $k^{l,r}(e)$ are continuous, definitions (5.17) and (5.18) yield

$$\det|k^l(e),k^r(e)|=\det|a^+,b^-|.$$

It is an immediate consequence of definitions $(5.15)^+$, $(5.16)^-$ of formulas (5.11), (5.12), and of assumption (2.6) that

$$\det|a^a,b^-|\neq0.$$

Combining these two relations yields

$$\det|k^l(e),k^r(e)|\neq0. \tag{5.19}$$

Relation (5.19) completes the proof of Lemma 5.1.

The proof of Lemma 5.1 also shows that neither conclusion (5.8a) nor conclusion (5.8b) holds for every solution of Eq. (5.7). In other words, each of the two endpoints are of the limit ponit type.[19] Combining this fact with Lemma 5.1 we see that the operator $L(Q(e,\kappa))$ is one to one on $\mathfrak{a}_0^\infty(\mathscr{R}_+,\mathscr{C}_2)$. Combining this fact, in turn, with the Weyl lemma[24d] we see that the closure of this operator is also one to one.

To see that this operator is onto we use the Weyl–Weidmann[19] construction for the resolvent. First we need a notation. To any two given vectors a, b in $\mathscr{C}_2$ we assign the matrix $a\rangle\langle b$ defined by

$$(a\rangle\langle b)x=a\langle b,x\rangle,\quad \textit{for every } x \textit{ in } \mathscr{C}_2. \tag{5.20}$$

Here

$$\langle b,x\rangle=(b,\bar{x}),$$

where bar on the right denotes complex conjugate. Note that since our inner product is conjugate linear in the second argument, the left member is linear in both arguments. In short, it is a real inner product. Secondly, following the Weyl–Weidmann construction we set

$$R(e)(\xi,\eta)=\frac{1}{\det|k^r(e),k^l(e)|}\times\begin{cases}k^r(e)(\xi)\rangle\langle k^l(e)(\eta), & \eta<\xi,\\ k^l(e)(\xi)\rangle\langle k^r(e)(\eta), & \eta>\xi.\end{cases} \tag{5.21}$$

As usual, we define the operator corresponding to this kernel by

$$R(e)f(\xi)=\int R(e)(\xi,\eta)f(\eta)\,d\eta,\quad f\in\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2). \tag{5.22}$$

Thirdly, we claim that

$$L(Q(e,\kappa))R(e)=I\quad\text{on }\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2). \tag{5.23}$$

To see this we note that according to definition (4.2) the matrix J is antisymmetric with respect to the real inner product. Hence,

$$\langle a,Jb\rangle = -\langle b,Ja\rangle \quad\text{and}\quad \langle a,Ja\rangle = \langle b,Jb\rangle = 0.$$

These relations and another application of definition (4.2) yields

$$Jb\rangle\,\langle a - Ja\rangle\,\langle b = -\det|a,b|I \quad\text{on } \mathscr{C}_2.$$

This relation, in turn, together with definitions (5.21), (5.22), Eq. (5.7), and elementary algebra yields the validity of relation (5.23). Fourthly, we shall show that relation (5.23) holds on all of $\mathfrak{L}_2(\mathscr{R}_+\ \mathscr{C}_2)$ by showing that

$$R(e)\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)). \tag{5.24}$$

To prove relation (5.24) we need the Schur–Holmgren–Carleman[20] bound of a given integral operator R with reference to a given positive measurable function t. This is defined by

$$\|R\|(t) = \left(\sup t(\xi)^{-1}\int |R(\xi,\eta)|\,t(\eta)d\eta \cdot\sup t(\eta)^{-1}\int |R(\xi,\eta)|\,t(\xi)d\xi\right)^{1/2}, \tag{5.25}$$

where the supremum is taken over the support of t. According to their result,[20] if the support of $t(\xi)t(\eta)$ contains the support of $R(\xi,\eta)$, then

$$\|R\|\leqslant\|R\|(t), \tag{5.26}$$

where the left member is the operator norm. In view of relation (5.26) relation (5.24) is implied by the lemma which follows.

Lemma 5.2: *Suppose that the real number e satisfies assumption* (2.6). *Let the kernel* $R(e)(\xi,\eta)$ *be defined by relation* (5.21) *and let the function t be defined by*

$$t(\eta) = \eta^{-1/2}. \tag{5.27}$$

Then for each $\kappa = \pm 1, \pm 2,\cdots$

$$\|R(e)\|(t) < \infty. \tag{5.28}$$

To prove this lemma first define

$$z^l(e)(\rho) = \rho^{-(\kappa^2-e^2)^{1/2}}a^- \tag{5.29a}$$

and

$$z^r(\rho) = \exp[+(\rho-1)]b^+. \tag{5.29b}$$

Then according to definitions (5.15) $^-$, (5.16) $^-$ and formulas (5.10) these functions satisfy Eq. (5.7) on the intervals $(0,1)$ and $(1,\infty)$, respectively. At the same time, remembering definitions (5.17) we see that

$$|y^l(e)(\rho)|\leqslant|z^l(e)(\rho)|, \quad\text{for } \rho\in(0,1) \tag{5.30a}$$

and

$$|y^r(\rho)|\leqslant|z^r(\rho)|, \quad\text{for } \rho\in(1,\infty). \tag{5.30b}$$

Secondly, define

$$m^l(e)(\rho) = \begin{cases}|y^l(e)(\rho)|, & \rho\in(0,1)\\ |z^r(\rho)|, & \rho\in(1,\infty)\end{cases} \tag{5.31a}$$

and

$$m^r(e)(\rho) = \begin{cases}|z^l(e)(\rho)|, & \rho\in(0,1),\\ |y^r(\rho)|, & \rho\in(1,\infty).\end{cases} \tag{5.31b}$$

Since the general solution of Eq. (5.7) is a linear combination of these solutions we see from definitions (5.18) and (5.19) that to each κ there is a constant γ such that

$$|R(e)(\xi,\eta)|\leqslant\gamma\begin{cases}m^l(e)(\eta)m^r(e)(\xi), & \eta<\xi,\\ m^l(e)(\xi)m^r(e)(\eta), & \xi<\eta.\end{cases} \tag{5.32}$$

Combining this estimate with definition (5.27) yields

$$t(\xi)^{-1}|R(e)(\xi,\eta)|\,t(\eta) \leqslant\gamma\xi^{1/2}\eta^{-1/2}\begin{cases}m^l(e)(\eta)m^r(e)(\xi), & \eta<\xi,\\ m^l(e)(\xi)m^r(e)(\eta), & \xi<\eta.\end{cases} \tag{5.33}$$

Hence,

$$t(\xi)^{-1}\int |R(e)(\xi,\eta)|\,t(\eta)d\eta \leqslant\gamma\xi^{1/2}\Big(m^r(e)(\xi)\int_0^\xi \eta^{-1/2}m^l(e)(\eta)d\eta + m^l(e)(\kappa)(\xi)\int_\xi^\infty \eta^{-1/2}m^r(e)(\eta)d\eta\Big). \tag{5.34}$$

To estimate the right member first we make the additional assumption,

$$0<\xi<1. \tag{5.35}$$

Then we see from definitions (5.31), (5.29a), (5.17a), and assumption (2.6) that

$$\xi^{1/2}m^r(e)(\xi)\int_0^\xi \eta^{-1/2}m^l(e)(\eta)d\eta = \frac{\xi}{(\kappa^2-e^2)^{1/2}+\frac{1}{2}}. \tag{5.36}$$

Similarly, combining these definitions with definition (5.17b) we see that

$$\xi^{1/2}m^l(e)(\xi)\int_\xi^\infty \eta^{-1/2}m^r(e)(\eta)d\eta = \xi^{1/2+(\kappa^2-e^2)^{1/2}}\Big(\int_\xi^1 \eta^{-1/2-(\kappa^2-e^2)^{1/2}}d\eta + \int_\xi^\infty \eta^{-1/2}\exp[-(\eta-1)]d\eta\Big).$$

It is a matter of elementary algebra to estimate the integrals on the right. Because of assumption (2.6) this yields,

$$\xi^{1/2}m^l(e)(\xi)\int_\xi^\infty \eta^{-1/2}m^r(e)(\eta)\leqslant\xi\Big(\frac{1}{(\kappa^2-e^2)^{1/2}}+1\Big). \tag{5.37}$$

Secondly, we make the additional assumption

$$1<\xi. \tag{5.35$'$}$$

Then we see from definitions (5.31), (5.29b) and (5.17b) that

$$\xi^{1/2}m^l(e)(\xi)\int_\xi^\infty \eta^{-1/2}m^r(e)(\eta)d\eta\leqslant 1. \tag{5.37$'$}$$

Similarly, combining these definitions with definition (5.29b) we see that

$$\xi^{1/2}m^r(e)(\xi)\int_0^\xi \eta^{-1/2}m^l(e)(\eta)d\eta) = \xi^{1/2}\exp(-(\xi-1))\Big(\int_0^1 \eta^{-1/2+(\kappa^2-e^2)^{1/2}}d\eta$$

$$+\int_1^{\xi}\eta^{-1/2}\exp(\eta-1)d\eta\Big).$$

It is a matter of elementary algebra to estimate the integrals on the right. Because of assumption (2.6) this yields,

$$\xi^{1/2}m^r(e)(\xi)\int_0^{\xi}\eta^{-1/2}m^l(e)(\eta)d\eta\leqslant\frac{1}{(\kappa^2-e^2)^{1/2}+\frac{1}{2}}+2 \tag{5.36$'$}$$

Inserting estimates (5.36), (5.36)′, (5.37) and (5.37)′ in estimate (5.34) we obtain

$$\sup t(\xi)^{-1}\int|R(e)(\xi,\eta)|t(\eta)d\eta<\infty. \tag{5.38}$$

Here the supremum is taken in ξ over $\mathscr{R}_+$. According to definition (5.21) the kernel $R(e)(\xi,\eta)$ is symmetric. Inserting this fact and estimate (5.38) in definition (5.25) we arrive at the validity of conclusion (5.28). This completes the proof of Lemma 5.2.

Combining Lemma 5.2 with relations (5.23) and (5.26) proves conclusion (5.5) of Theorem 5.1. At the same time it follows that

$$R(e)=R(0,-L(Q(e,\kappa))). \tag{5.39}$$

To prove conclusion (5.6) define

$$K(e)=M^{-1}R(0,L(Q(e,\kappa)))\quad\text{on}\quad L(Q(e,\kappa))\mathfrak{C}_0^{\infty}(\mathscr{R}_+,\mathscr{C}_2). \tag{5.40}$$

Note that according to conclusion (5.5) the set on the right is dense. Similarly to the proof of conclusion (5.5) we show that this operator is bounded by showing that a Schur–Holmgren–Carleman bound is finite. This is the statement of the lemma which follows.

Lemma 5.3: *Suppose that the real number e satisfies assumption* (2.6). *Let the operator $K(e)$ be defined by relation* (5.40) *and let the function t be defined by relation* (5.27). *Then for each $\kappa=\pm 1,\pm 2,\cdots$*

$$\|K(e)\|(t)<\infty. \tag{5.41}$$

To prove this lemma first we note that relation (5.39) and definitions (5.40), (5.21), and (4.3) together show that

$$K(e)(\xi,\eta)=\frac{\xi^{-1}}{\det|(k^l(e),k^r(e)|}\times\begin{cases}k^r(e)(\xi))\rangle\langle k^l(e)(\eta), & \eta<\xi,\\ k^l(e)(\xi))\rangle\langle k^r(e)(\eta), & \xi<\eta.\end{cases}$$

Inserting estimate (5.32) in this relation yields

$$|K(e)(\xi,\eta)|\leqslant\gamma\xi^{-1}\begin{cases}m^l(e)(\eta)m^r(e)(\xi), & \eta<\xi,\\ m^l(e)(\xi)m^r(e)(\eta), & \xi<\eta.\end{cases} \tag{5.42}$$

Estimate (5.42) allows us to replace the kernel $R(e)(\xi,\eta)$ by $K(e)(\xi,\eta)$ in the proof of estimate (5.28) of Lemma 5.2. This way, we arrive at

$$\sup t(\xi)^{-1}\int|K(e)(\xi,\eta)|t(\eta)d\eta<\infty, \tag{5.43}$$

where the supremum in ξ is taken over $\mathscr{R}_+$.

Secondly we note that estimate (5.42) and definition (5.27) together yield

$$t(\eta)^{-1}|K(e)(\xi,\eta)|t(\xi)$$

$$<\gamma\xi^{-3/2}\eta^{1/2}\begin{cases}m^l(e)(\eta)m^r(e)(\xi), & \eta<\xi,\\ m^l(e)(\xi)m^r(e)(\eta), & \xi<\eta.\end{cases} \tag{5.44}$$

Hence

$$t(\eta)^{-1}\int|K(e)(\xi,\eta)|t(\xi)d\xi$$

$$\leqslant\gamma\eta^{1/2}(e)(\eta)\Big(m^r(e)(\eta)\int_0^{\eta}\xi^{-3/2}m^l(e)(\xi)d\xi+m^l(e)(\eta)\int_{\eta}^{\infty}\xi^{-3/2}m^r(e)(\xi)d\xi\Big). \tag{5.45}$$

To estimate the right member first we make the additional assumption

$$0<\eta<1. \tag{5.46}$$

Then, similarly to relations (5.36) and (5.37), we see that assumption (2.6) imiplies that

$$\eta^{1/2}m^r(e)(\eta)\int_0^{\eta}\xi^{-3/2}m^l(e)(\xi)d\xi=\frac{1}{(\kappa^2-e^2)^{1/2}-\frac{1}{2}} \tag{5.47}$$

and

$$\eta^{1/2}m^l(e)(\eta)\int_{\pi}^{\infty}\xi^{-3/2}m^r(e)(\xi)d\xi\leqslant\frac{1}{(\kappa^2-e^2)^{1/2}}+2. \tag{5.48}$$

Secondly, we make the additional assumption

$$1<\eta. \tag{5.46$'$}$$

Then, similarly to relations (5.36)′ and (5.37)′, we see that assumption (2.6) implies

$$\eta^{1/2}m^r(e)(\eta)\int_0^{\eta}\xi^{-3/2}m^l(e)(\xi)d\xi\leqslant\frac{1}{(\kappa^2-e^2)^{1/2}-\frac{1}{2}}+2, \tag{5.47$'$}$$

and

$$\eta^{1/2}m^l(e)(\eta)\int_{\eta}^{\infty}\xi^{-3/2}m^r(e)(\xi)d\xi\leqslant 2. \tag{5.48$'$}$$

Inserting estimates (5.47), (5.47)′, (5.48), and (5.48)′ in estimate (5.45) we obtain,

$$\sup t(\eta)^{-1}\int|K(e)(\xi,\eta)|t(\xi)d\xi<\infty. \tag{5.49}$$

Inserting estimates (5.43) and (5.49), in turn, in definition (5.25) we arrive at the validity of conclusion (5.41). This completes the proof of Lemma 5.3.

Combining Lemma 5.3 with definition (5.40) and relation (5.26) yields the validity of conclusion (5.6). This completes the proof of Theorem 5.1.

Next we derive relation (5.1) from Theorem 5.1. For this purpose first we note that conclusion (5.5) and the first resolvent equation[23] together yield

$$R(\pm i,L(Q(e,\kappa)))=R(0,L(Q(e,\kappa)))\pm iR(0,L(Q(e,\kappa)))\times R(\pm i,L(Q(e,\kappa))).$$

Multiplying this equation by M^{-1} and using conclusion (5.6) we obtain

$$M^{-1}R(\pm i,L(Q(e,\kappa)))\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)). \tag{5.50}$$

Secondly, we note that according to definitions (4.7) and (5.4)

$$\sup|P(e,\kappa)(\rho)-Q(e,\kappa)(\rho)|<\infty,$$

where the supremum is taken over ρ in $\mathscr{R}_+$. Hence

$$L(P(e,\kappa))-L(Q(e,\kappa))\in\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+,\mathscr{C}_2)). \tag{5.51}$$

Since the operator $L(Q(e,\kappa))$ is symmetric we see from conclusion (5.5) that $\pm i$ are in $\rho(L(Q(e,\kappa)))$.[24c,3d] This fact and relations (4.12) and (5.51) together allow us to apply the second resolvent equation to them.[26] This yields

$$\begin{aligned}R(\pm i,L(P(e,\kappa))) &= R(\pm i,L(Q(e,\kappa)))\\ &\quad + R(\pm i,L(Q(e,\kappa)))[L(P(e,\kappa))\\ &\quad - L(Q(e,\kappa))]R(\pm i,L(P(e,\kappa))).\end{aligned} \tag{5.52}$$

Multiplying this equation by M^{-1} and using relation (5.50) we arrive at the validity of relation (5.1). From this, in turn, we arrive at the validity of conclusion (4.13) of Theorem 4.1.

Since we defined the domains of our operators by closure, conclusion (4.13) clearly implies conclusion (4.14). This completes the proof of Theorem 4.1.

ACKNOWLEDGMENT

It is a pleasure to thank Professor Agmon for his valuable suggestions which have been incorporated in Secs. 2 and 3.

[1]F. Rellich, "Die zulässigen Randbedingungen bei den singulären Eigenwertproblemen der mathematischen Physik" Math. Z. **49**, 702–723 (1943–44).

[2]F. Rellich, *Eigenwerttheorie Partieller Differentialgleichungen*, Teil II (in German), University of Göttingen Lecture Notes (1953). See Chapter III.

[3]T. Kato, *Perturbation Theory for Linear Operators* (Springer–Verlag, 1966), a. Theorem V. 3.16; b. Theorem V. 4.4; c. Theorem V. 5.10; d. Problem V. 3.17.

[4]K.R. Parthasaraty, "On the exact bound of the Coulomb potential with respect to the Dirac operator," Lett. Math. Phys. **2**, 413–415 (1978).

[5]M. Schechter, *Spectra of Partial Differential Operators* (North-Holland, Elsevier, 1971).

[6]F. Stummel, "Singuläre elliptische Differential Operatoren in Hilbertschen Räumen," Math. Ann. **132**, 150–176 (1956).

[7]W.D. Evans, "On the unique self-adjoint extension of the Dirac operator and the existence of the Green matrix," Proc. London Math. Soc. (3) **20**, 537–557 (1970).

[8]M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, a. *Functional Analysis*, (Academic, 1972), Vol. I, Sec. II.4; b. *Fourier Analysis, Self-Adjointness* (Academic, 1975), Vol. II, p. 326, Notes to Sec. X.5.

[9]L. Gross, "The Cauchy Problem for the coupled Maxwell and Dirac equations," Commun. Pure Appl. Math. **19**, 1–15, (1966). See Lemma 1.4.

[10]R.T. Prosser, "Relativistic potential scattering," J. Math. Phys. **4**, 1048–54 (1963).

[11]M. Thompson, "Eigenfunction expansions and the associated scattering theory for potential perturbations of the Dirac equation," Q. J. Math. Oxford **23**, 17–55 (1972).

[12]K. Eckhardt, *Streutheorie für Dirac-operatoren*, Manuscripta Mathematica II (1974), 359–71. See Satz 1.

[13]J.C. Guillott and G. Schmitt, "Spectral and Scattering theory for Dirac operators," Arch. Rat. Mech. Anal. **55**, 139–206 (1974).

[14]R.A. Weder, "Spectral analysis of pseudodifferential operators," J. Funct. Anal. **20**, 319–37 (1975). See Theorem 2.1.

[15]P.A.M. Dirac, *Principles of Quantum Mechanics* (Clarendon, Oxford, 1936).

[16]A. Messiah, *Quantum Mechanics* (Wiley, New York, 1966), Vol. 2. See Eqs. 170(a), (b) in Sec. XX.25.

[17]P.A. Rejto, "On reducing subspaces for one electron Dirac operators," Israel J. Math. **9**, 111–143 (1971).

[18]W. Weyl, "Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugenhörigen Entwicklungen Willkürlicher Funktionen," Math. Ann. **68**, 220–69 (1909).

[19]Joachim Weidmann, "Oscillations-methoden für Systeme gewöhnlicher Differentialgleichungen," Math. Z. **119**, 349–73 (1971). See Satz 1.4 and Sec. 9.

[20]K.O. Friedrichs, *Spectral Theory of Operators in Hilbert Space*, Springer Lecture Notes in the Applied Mathematical Sciences, No. 9 (1973). See Sec. 20.

[21]U.W. Schmincke, "Essential self-adjointness of Dirac operators with a strongly singular potential," Math. Z. **126**, 71–81 (1972).

[22]Kalf–Schmincke–Walter–Wüst, *On the spectral theory of Schrödinger and Dirac operators with strongly singular potentials, in spectral theory and differential equations*, Springer Lecture Notes in Mathematics, No. 433 (1975).

[23]G. Nenciu, "Self-adjointness and invariance of the essential spectrum for Dirac operators defined as quadratic forms," Commun. Math. Phys. **48**, 235–247 (1976).

[24]N. Dunford and J.T. Schwartz, *Linear Operators Parts I and II*, (Wiley Interscience, 1957 and 1963). Chapters I–XIV. a. Definition VII 3.1; b. Lemma VII 3.6; c. Corollary XII 4.13; d. Lemma XIII 2.9.

[25]M.A. Naimark, *Normed Rings* (Walters, Nordhoff, 1972). See Subsection I.5.15.

[26]E. Hille and R.S. Phillips, *Functional Analysis and Semigroups*, Am. Math. Soc. Coll. XXXI (1957). See Subsection V.5.10.

[27]H. Weyl, *Theory of Groups and Quantum Mechanics* (Dover, New York, 1950). See. Sec. 4.7.

[28]P.R. Chernoff, "Essential self-adjointness of powers of generators of hyperbolic equations," J. Func. Anal. **12**, 401–14 (1973).

[29]K.E. Gustafson and P.A. Rejto, "Some essentially self-adjoint Dirac operators with spherically symmetric potentials," Israel J. Math. **14**, 63–75 (1973)

[30]J.J. Landgren, "Some essentially self-adjoint Dirac type operators I," Lett. Math. Phys. **2**, 61–70 (1977).

[31]J.J. Landgren, "Some essentially self-adjoint Dirac type operators II," Lett. Math. Phys. **2**, 145–153 (1977).

[32]B.M. Levitan and M. Otelbaev, "Conditions for the self-adjointness of Schrödinger and Dirac operators" Dokl. Akad. Nauk SSSR **235**, 768–771 (1977) [Sov. Math. Dokl. **18**, 1044–48 (1977)]; Math. Rev. **56**, 6142 (1978).

[33]Joachim Weidmann, *Linear Operatoren in Hilberträumen, Mathematische Leitfäden* (Teubner, Stuttgart, 1976). See. Sec. 10.6.

Cyclic pursuit in a plane

F. Behroozi
University of Wisconsin-Parkside, Kenosha, Wisconsin 53141

R. Gagnon
Hughes Aircraft Company, Fullerton, California 92634

(Received 12 January 1979; accepted for publication 17 May 1979)

The motion of an arbitrary set of points (or bugs) in a plane chasing one another in cyclic pursuit is studied. It is shown that for regular center-symmetric configurations, analytic solutions are easily obtained by going to an appropriate rotating frame of reference. A few cases of nonsymmetric configurations are discussed. In particular, it is shown that for three bugs in a triangular configuration, the center of the rotating coordinate system relative to which the bugs have no tangential velocity is the point of collapse and coincides with one of the two Brocard points of the triangle. For the case when all the bugs have the same speed, a theorem is proved that whenever a premature (i.e., nonmutual) capture occurs, the collision must be head on. This theorem is then applied to the case of three and four-bug configurations to show that these systems collapse to a point, i.e., the capture is mutual. Some aspects of these results are generalized to the case of the n-bug systems.

I. INTRODUCTION

A well-known problem that seems to have made its debut nearly a century ago,[1–3] and has since reappeared in numerous forms,[4–6] is the one about N bugs (or dogs) that are initially placed on the N vertices of a regular polygon. At a given time the bugs start pursuing one another along the instantaneous line of sight with constant speed either in a clockwise or counterclockwise direction. Given the speed of the bugs and the size of the polygon, one is normally asked for the length of the path covered, or the time elapsed before mutual capture occurs.

Watton and Kydon have described the analytical aspects of the problem for regular polygons.[7] Klamkin and Newman, using complex variables, have recently offered a more general and rigorous treatment of the problem, and have extended it to the case of nonsymmetric 3-bug configuration.[8] Here we describe a more intuitive solution for the regular polygons, and explore the problem further for nonsymmetric configurations by going to a rotating system of coordinates. In particular, we prove a theorem which is of general validity for the problem of pursuit. The application of this theorem to the case of 3-, 4-, and n-bug systems is considered.

II. REGULAR POLYGONS

Figure 1 shows the initial configuration for bugs 1, 2, 3, and n, placed on the vertices of a regular n-gon of side a. By the symmetry of the problem, the configuration of the system looks similar for all time, though shrinking in size, to an observer placed at the centroid of the polygon. The speed with which each bug approaches the center is simply $v \sin\phi = v \sin(\pi/n)$, so that

$$v_r = \dot{r} = -v \sin(\pi/n), \tag{1}$$

where r is the radial distance from the centroid.

Similarly, we obtain an equation for the tangential motion,

$$v_\theta = r\dot{\theta} = v \cos(\pi/n), \tag{2}$$

where θ is the polar angle from our reference line.

It is instructive to observe the system from a frame which rotates with variable angular speed $\dot{\theta} = [v \cos(\pi/n)]/r$ about the center of the polygon. To an observer in this frame, all the bugs move uniformly toward the center with speed $v \sin(\pi/n)$. Thus, the time for the bugs to reach the center is $T = r_0/[v \sin(\pi/n)]$, where r_0 is the initial radial distance. The path for each bug, however, always appears as a straight line independent of the speed. This means that the bugs may stop to rest on their way, or vary their speeds to suit the terrain (provided all the bugs do the same thing at a given time) without any change in their paths.

It is elementary to integrate Eq. (1) to obtain

$$r = r_0 - [v \sin(\pi/n)]t. \tag{3}$$

To find an expression for $\theta(t)$, we combine Eqs. (2) and (3) to get

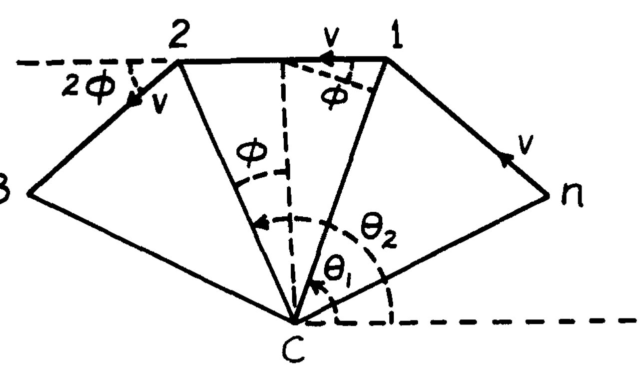


FIG. 1. Part of a regular polygon showing Bugs 1, 2, 3, and n. All bugs have the same speed v.

$$\dot\theta = (1/\{r_0 - [v\sin(\pi/n)]t\})v\cos(\pi/n). \tag{4}$$

Integration results in

$$\theta = \theta_0 - [\cot(\pi/n)]\ln\{[r_0 - vt\sin(\pi/n)]/r_0\}. \tag{5}$$

Finally, elimination of time between Eqs. (3) and (5) leads to the equation of the path,[9]

$$r = r_0\exp[(\theta_0 - \theta)\tan(\pi/n)]. \tag{6}$$

We note that $r(\theta)$ is independent of the speeds, as indeed should be the case. The speed only plays the role of a parameter which influences the time required for mutual capture. This time is given by

$$T = \int_{r_0}^{0} dr/\dot r = \int_{0}^{r_0} dr/[v(t)\sin(\pi/n)],$$

which for the case of constant speed results in

$$T = r_0/[v\sin(\pi/n)] = a/[2v\sin^2(\pi/n)], \tag{7}$$

where a is the length of a polygon side.

The distance traveled can be found by integrating the path element $dS = (dr^2 + r^2\,d\theta^2)^{1/2}$. Alternatively, we note that since Eq. (7) yields the time required for mutual capture, the path covered may be obtained through multiplication of the time of flight by the speed,

$$S = Tv = a/[2\sin^2(\pi/n)]. \tag{8}$$

The same results for T and S are obtained if we note that the approach speed of one bug with respect to its neighbor is

$$v_{ap} = v[1 - \cos(2\pi/n)] = v[2\sin^2(\pi/n)],$$

while the relative distance is just a. The time required for capture is therefore $T = a/v_{ap} = a/[2v\sin^2(\pi/n)]$, and the distance traveled is $S = Tv = a/[2\sin^2(\pi/n)]$, as before.

III. NONSYMMETRIC CONFIGURATIONS

We now proceed to consider the case of nonsymmetric convex polygons. The results of Sec. II lead us to ask whether there exists a point for the polygon such that an observer placed at this point will see the bugs approaching along straight lines. In other words, we wish to go to a rotational frame in which the bugs have no tangential velocity. In such a frame (should it exist), the center of rotation would be the point of collapse or mutual capture, if we arrange for the polygon to preserve its shape.

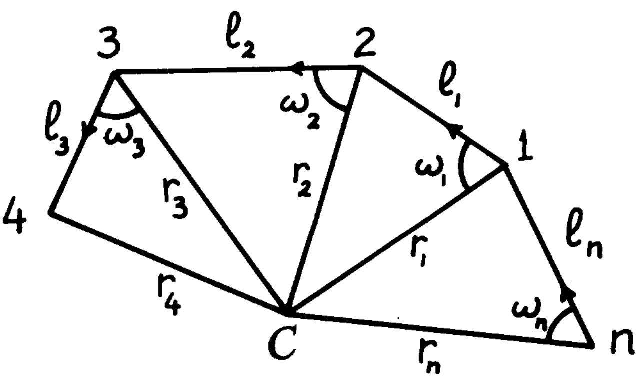


FIG. 2. Part of a nonsymmetric convex polygon. The symbols are self-explanatory.

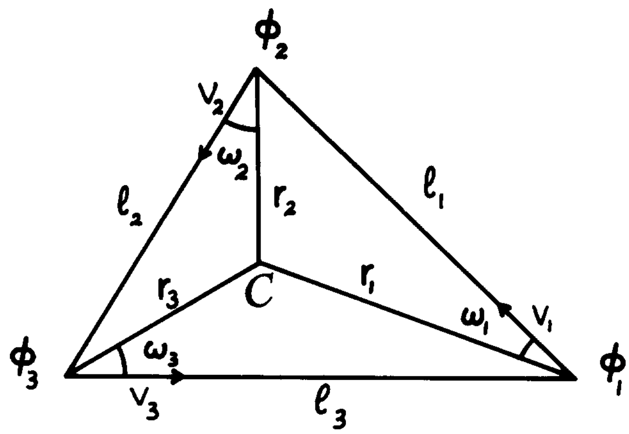


FIG. 3. An arbitrary (nondegenerate) triangle.

Consider the general convex polygon as shown in Fig. 2. The symbols are self-explanatory. For point C to be the point of collapse, we must have

$$1/T = (v_1\cos\omega_1)/r_1 = (v_2\cos\omega_2)/r_2 = \cdots$$
$$= (v_n\cos\omega_n)/r_n,$$

where T is the time required for mutual capture. Furthermore we must require

$$\dot\theta = (v_1\sin\omega_1)/r_1 = (v_2\sin\omega_2)/r_2 = \cdots$$
$$= (v_n\sin\omega_n)/r_n,$$

where $\dot\theta$ is the angular velocity with respect to point C.

Since $T\dot\theta = \tan\omega_1 = \tan\omega_2 = \cdots$, and $0 < \omega_i < \pi/2$ $(i = 1,2,\ldots,n)$, these two conditions are equivalent to

$$\omega_1 = \omega_2 = \omega_3 \cdots = \omega_n = \omega, \tag{A}$$

and hence

$$v_1/r_1 = v_2/r_2 = \cdots = v_n/r_n. \tag{B}$$

Condition (A) tells us that in order for the polygon to collapse to point C, lines drawn from C to the vertices of the polygon must all make the same angle ω with the sides. Then we only have to alter the speeds to each bug according to condition (B) for mutual capture to occur at point C.

In general, condition (A) cannot be met for arbitrary convex polygons. However special cases exist.[4] The arbitrary triangle is one such special case, which we now consider in more detail.

Consider Fig. 3. The symbols used are self-explanatory, and the chase is counterclockwise. The existence and uniqueness of point C satisying condition (A) is guaranteed in an arbitrary triangle through the work of the French mathematician Brocard[4,10] and others. In fact, point C is one of the two points in the triangle known as the Brocard points. The other Brocard point corresponds to the point of collapse for the clockwise chase.

Applying the law of sines to triangle $C\phi_2\phi_3$ results in

$$r_2/\sin\omega = r_1/\sin(\phi_2 - \omega).$$

Since $r_1/r_2 = v_1/v_2$, we have

$$v_1/v_2 = \sin(\phi_2 - \omega)/\sin\omega$$

resulting in

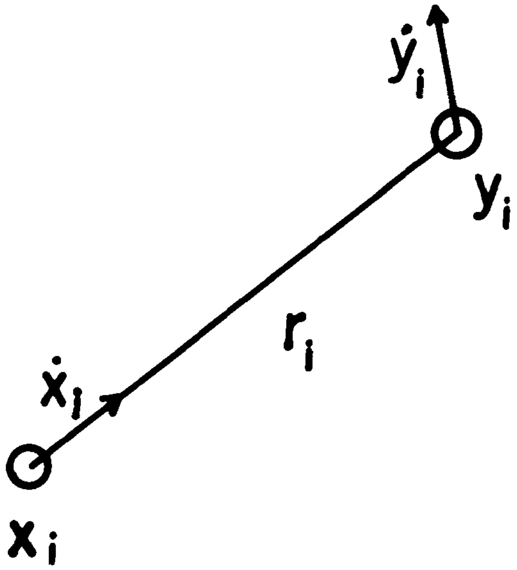


FIG. 4. Bug x_i is chasing Bug y_i.

$$\cot\omega = [(v_1/v_2) + \cos\phi_2]/\sin\phi_2. \quad (9)$$

The ratio of speeds v_1/v_2 is fixed by the requirement that the triangle remain similar to itself throughout the course of motion. This means that the approach speed of one bug with respect to the target bug must be proportional to the initial distance between them. Thus, we must have

$$T = l_1/(v_1 + v_2\cos\phi_2) = l_2/(v_2 + v_3\cos\phi_3) = l_3/(v_3 + v_1\cos\phi_1).$$

The above relations lead to

$$v_1/v_2 = (l_1 - l_2\cos\phi_2 + l_3\cos\phi_2\cos\phi_3)/(l_2 - l_3\cos\phi_3 + l_1\cos\phi_3\cos\phi_1)$$

which can be incorporated into Eq. (9) to give ω in terms of the constants of the triangle. Several equivalent expressions for ω exist,[4,11] for example, $\cot\omega = \cot\phi_1 + \cot\phi_2 + \cot\phi_3$.

The equation of the path is easily determined. Simple integration of the equations of motion for bug 2 of Fig. 3, results in

$$r_2 = r_{02} - (v_2\cos\omega)t,$$

and

$$\theta_2 = \theta_{02} - \tan\omega\,\ln[(r_{02} - v_2 t\cos\omega)/r_{02}],$$

leading to

$$r_2 = r_{02}\exp[(\theta_{02} - \theta)\cot\omega].$$

We note in passing that a general polygon does not remain similar to itself if the approach speeds are made proportional to the length of the corresponding sides. In other words, one can construct many polygons with the same side ratios; the triangle is unique in that given the side ratios, the angles are determined.

IV. HEAD-ON COLLISION THEOREM

For the general case of m bugs in n-dimensional space where the bugs have equal speeds, the following theorem is of general interest:

Theorem: A bug cannot capture a bug whose acceleration is bounded, except by a head-on collision.

This theorem can be proved geometrically, but a proof using tensor notation is more general because it avoids reference to a specific number of dimensions.

Let x_i be the coordinates of the bug that chases, y_i the coordinates of the bug being chased, $r_i = y_i - x_i$ the displacement vector, $r = (r_i r_i)^{1/2}$ the displacement,[12] and $v = (\dot{x}_i\dot{x}_i)^{1/2} = (\dot{y}_i\dot{y}_i)^{1/2}$ the constant velocity magnitude. See Fig. 4.

The approach speed, i.e., the rate at which the distance between the two bugs changes, is given by

$$\dot{r} = -v + \dot{y}_i\,\dot{x}_i/v. \quad (10)$$

Note that $\dot{x}_i/v = r_i/r$ is the unit vector along r_i. Since $\dot{y}_i\,\dot{y}_i = v^2$ and $r_i = y_i - x_i$, Eq. (10) reduces to

$$\dot{r} = -(\dot{y}_i\,\dot{y}_i - \dot{y}_i\,\dot{x}_i)/v = -(\dot{y}_i\,\dot{r}_i)/v. \quad (11)$$

Finally we write Eq. (11) as follows:

$$v\frac{dr}{dt} = -\frac{d}{dt}(r_i\,\dot{y}_i) + r_i\,\ddot{y}_i. \quad (12)$$

Now, consider the proposition that a capture is about to take place after a short time T, and integrate Eq. (12) over this time interval. We use the subscript 0 to represent initial values, and subscript f to represent final values. We obtain,

$$v(r_f - r_0) = -r_{if}\,\dot{y}_{if} + r_{i0}\,\dot{y}_{i0} + \int_{t_0}^{T+t_0} r_i\,\ddot{y}_i\,dt.$$

Since $r_f = r_{if} = 0$ at the instant of capture, we have

$$vr_0 + r_{i0}\,\dot{y}_{i0} = -\int_{t_0}^{T+t_0} r_i\,\ddot{y}_i\,dt. \quad (13)$$

Let A be an upper bound for $|\ddot{y}_i|$, and R be the upper bound for $|r_i/r_0|$ over the range of T. Then

$$|(r_{i0}\,\dot{y}_{i0}/r_0) + v| \leqslant TRA.$$

This equation holds throughout the time interval before capture. Taking the limit $T \to 0$, we obtain

$$\lim_{T\to 0} |(r_{i0}\,\dot{y}_{i0}/r_0) + v| = 0.$$

Since $r_{i0}/r_0 = \dot{x}_{i0}/v$, we have that

$$\lim_{T\to 0} (\dot{x}_{i0}\,\dot{y}_{i0}) = -v^2$$

which implies that $\dot{x}_i = -\dot{y}_i$ at the moment of capture. This proves the theorem.

As a corollary, we deduce that *a bug cannot capture a bug which is not capturing another bug, except by head-on collision*. This is because of the fact that a bug not capturing another bug cannot have infinite acceleration (it is chasing a distant bug with finite velocity). Its acceleration is bounded by v^2/r. And so a domino principle applies. The bug which cannot capture another bug, cannot itself be captured, and so on throughout all the bugs; except for the possibility of head-on capture. Thus except for this possibility, capture must be mutual. Of course, this theorem does not apply to the mutual capture on a regular polygon, because the acceleration becomes unbounded at the moment of capture.

We apply this theorem to the case of three bugs to show that the capture is mutual. First we note that a triangle con-

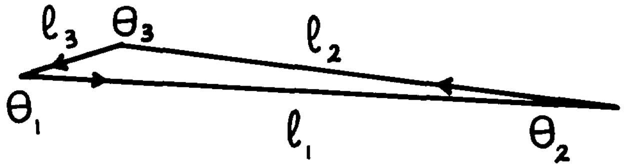


FIG. 5. Bug 3 is about to capture Bug 1.

figuration never becomes colinear, i.e., it always remains triangular. This is because a colinear configuration can only result from a colinear initial condition. (All bugs have their speeds along the same line, and so the configuration remains on a line). Thus a triangular configuration remains triangular throughout the course of motion.

Now suppose that a premature capture is about to take place in a three-bug system whose initial configuration was not colinear. From the preceding comments, it follows that the configuration remains noncolinear right up to the last moment before premature capture. Hence the configuration must resemble that of Fig. 5 just before the capture, where bug 3 is about to capture bug 1. By the assumption of imminent capture, we require that

$$l_3 \to 0 \tag{a}$$

and by the head-on collision theorem we deduce that

$$\theta_1 \to 0 \tag{b}$$

or that $\dot{\theta}_1$ must become negative.

On the other hand, by straightforward geometric arguments we may derive the equation

$$\dot{\theta}_1 = [(\sin\theta_1)/l_3] - [(\sin\theta_2)/l_1] \tag{14}$$

(without the loss of generality we have assumed unit velocity). By the law of sines we note that

$$\sin\theta_1 = (l_2 \sin\theta_2)/l_3.$$

It follows by Condition (a) that just before capture, it must be true that

$$\sin\theta_1 > \sin\theta_2. \tag{15}$$

Also by Condition (a) we see that ultimately

$$l_3 < l_1. \tag{16}$$

From Eqs. (14), (15), and (16) it follows ultimately that

$$\dot{\theta}_1 > 0,$$

which contradicts the deduction from Condition (b) that $\dot{\theta}_1$ must become negative. Hence premature capture cannot take place.

The impossibility of premature capture for the triangular configuration was originally proved by Klamkin and Newman.[8] We have reported it here to illustrate the utility of the head-on collision theorem. However, the theorem is of wider applicability. In what follows, we use the theorem to arrive at some useful results for the n-bug system.

V. THE N-BUG SYSTEMS

For some time we have entertained a conjecture to the effect that, in general

(a) a convex configuration remains convex, and

(b) premature capture is possible only in nonconvex configurations.

A configuration is convex if upon connecting the bug pairs $1\to2$, $2\to3$, $3\to4$, $\cdots$ a convex polygon results with the vertex angles less than π. In what follows we shall prove the conjecture for the 4-bug case and generalize some aspects of the proof for the n-bug systems. As expected, the head-on collision theorem plays an important part. We mention in passing that the bug system is capable of two types of motion: A forward motion in which the bugs chase one another, and a motion in reverse time in which the bugs run away from one another.

Consider Fig. 6 which shows a 4-bug configuration where bug 4 is about to capture bug 1. Let us assume that Fig. 6 represents the initial configuration at $t = 0$ and that the capture will take place at $t = T$. For unit velocity and a counterclockwise chase, the forward motion is governed by the following equations:

$$\dot{l}_i = -1 - \cos\theta_{i+1}, \tag{17}$$

$$\dot{\theta}_i = [(\sin\theta_i)/l_{i-1}] - [(\sin\theta_{i+1})/l_i]. \tag{17'}$$

In particular we have

$$\dot{\theta}_1 = [(\sin\theta_1)/l_4] - [(\sin\theta_2)/l_1], \tag{18}$$

$$\dot{l}_4 = -1 - \cos\theta_1. \tag{19}$$

The head-on collision theorem demands that at the moment of capture ($t = T$),

$$\lim_{t\to T}\theta_1 = 0. \tag{20}$$

Furthermore,

$$\lim_{t\to T} l_4 = 0. \tag{21}$$

Since $\Sigma\theta_i = 2\pi$, and $\theta_1 \to 0$, we have

$$\lim_{t\to T}\theta_4 = 2\pi - \theta_2(T) - \theta_3(T). \tag{22}$$

At $t = T$, the configuration is triangular with

$$[\theta_2(T) + \theta_3(T)] < \pi. \tag{23}$$

From Eqs. (22) and (23) we deduce the important result that

$$\lim_{t\to T}\theta_4 > \pi \tag{24}$$

which implies that the configuration must be nonconvex just before a premature capture. Could it have been convex at some earlier phase of its development?

Let us assume that the configuration, initially convex, becomes nonconvex at the time when θ_4 passes from the range $(0,\pi)$ to the range $(\pi,2\pi)$. Now if we start with this nonconvex configuration and retrace our steps, i.e., follow the motion in reverse time, θ_4 must pass from π^+ to π^- range.

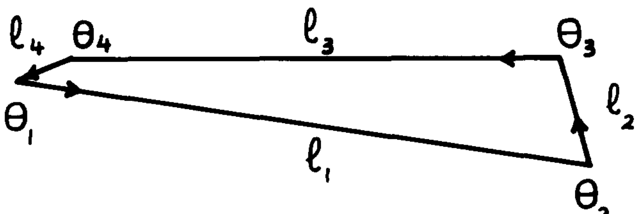


FIG. 6. Bug 4 is about to capture Bug 1.

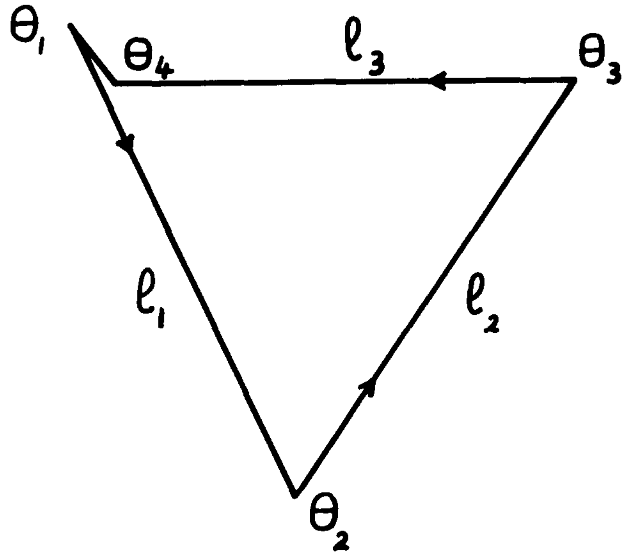


FIG. 7. A possible 4-bug nonconvex configuration leading to a premature capture.

The evolution of θ_4 in reverse time is governed by the equation

$$\dot{\theta}_4 = -[(\sin\theta_4)/l_3] + [(\sin\theta_1)/l_4], \quad (25)$$

which is the same as Eq. (18) except for a change in sign. When θ_4 is passing from π^+ to the π^- region, i.e., when θ_4 is transferring through the π barrier, $\dot{\theta}_4$ must be negative since θ_4 is diminishing. But Eq. (25) implies the contrary. This is so because when $\theta_4 = \pi$ we have $\theta_1 < \pi$ and hence $\sin\theta_1 > 0$, which results in $\dot{\theta}_4 > 0$. So θ_4 tends to bounce back from the π barrier and cannot pass through. The configuration remains nonconvex and cannot by time reversal reach its assumed convex initial condition. Conversely, a nonconvex configuration cannot evolve from a convex configuration. Furthermore, we conclude that only nonconvex 4-bug configurations can give rise to a premature capture.

The above arguments for the 4-bug case can be generalized to apply to the n-bug systems. One needs to replace 4 by n and follow a parallel argument. Also, Eqs. (22) and (23) require minor modifications and in Eq. (25) the subscript 3 must be incremented. There are nevertheless, a number of questions that need to be considered. First, can a nonconvex polygon evolve from a convex polygon by means of a crossover, when a bug crosses the line joining another bug pair. A little thought shows that the polygon must be nonconvex immediately after the crossover, at which time the same π barrier holds as before to prevent the emergence of a convex polygon. Second, would simultaneous nonmutual captures invalidate the preceeding argument? Perhaps. We do not have an answer to this question yet. Third, can the simultaneous transitions of more than one angle through the π invalidate the argument? Let us assume that θ_j and θ_{j+1} have made the transition from π^- to π^+. In reverse time they must retrace the steps and go from π^+ to π^-, after which the polygon is convex. But this transition cannot take place since θ_{j+2} is smaller than π and so by the same argument following Eq. (25), θ_{j+1} is unable to overcome the π barrier as before and the transition is forbidden. Fourth, is premature capture possible for nonconvex polygons? This possibility can be demonstrated by construction. In Fig. 7 we consider a 4-bug case. Assume that at the time of capture $l_1 = l_2 = l_3 = 1$, $l_4 = 0$, $\theta_1 = 0$, $\theta_2 = \theta_3 = 60°$, $\theta_4 = 240°$. Taking this configuration as our initial condition and working backward in time we will arrive at possible configurations, none of which are convex, which will yield premature capture in forward time. For the case considered, the initial values of the derivatives are given by Eqs. (17) and (17′), except for $\dot{\theta}_1$ which gives a problem because $l_4 = 0$. However by taking limiting values, we get

$$\lim[(\sin\theta_1)/l_4] = \lim[(\cos\theta_1)(\dot{\theta}_1)/\dot{l}_4] = -\dot{\theta}_1/2$$

and hence

$$\lim_{t \to 0} \dot{\theta}_1 = -(2/3)(\sin\theta_2)/l_1 = -\sqrt{3}/3.$$

In reverse time then $\dot{\theta}_1 = +\sqrt{3}/3$ initially, and one has no difficulty retrieving the configurations that led to the premature capture.

Although the above arguments do not constitute a proof of the conjecture, they make it plausible. Our recent computer-assisted analysis of cyclic pursuit[13] lends further support to our conjecture and leads us to believe that a mathematical proof is possible. However, we suspect that the computer will play a role in such a proof much in the same way that the four-color conjecture was proved.[14]

[1]E. Lucas, Nouvo Corres. Math. **3**, 175 (1877).
[2]H. Brocard, Nouvo Corres. Math. **3**, 280 (1877).
[3]H. Brocard, Nouvo Corres. Math. **6**, 211 (1880).
[4]In a series of four papers in Scripta Mathematics, Arthur Bernhart gives an excellent and extensively referenced review of the pursuit problems. These articles are listed here in chronological order: (a) Scripta Mathematica **20**, 125 (1954); (b) **23**, 49 (1957); (c) **24**, 23 (1959); (d) **24**, 189 (1959). The third paper in the list is of particular interest to readers of the present article.
[5]M. Gardner, Sci. Am. **213**, 101 (1965).
[6]H. Steinhaus, *Mathematical Snapshots* (Oxford U.P., New York, 1969), p. 136.
[7]A. Watton and D.W. Kydon, Am. J. Phys. **37**, 220 (1969).
[8]M.S. Klamkin and D.J. Newman, Am. Math. Mon. **78**, 631 (1971).
[9]It is also possible to get $r(\theta)$ directly by noticing
$$\frac{dr}{d\theta} = \frac{dr}{dt}\frac{dt}{d\theta} = -r\tan(\pi/n)$$
and thus $r = r_0 \exp[(-\tan\pi/n)(\theta - \theta_0)]$. However, as we need to refer to $\theta(t)$ and $r(t)$, later, we have obtained the parametric forms explicitly.
[10]R.A. Johnson, *Advanced Euclidean Geometry* (Dover, New York, 1960), pp. 263–70.
[11]P. Yff, Am. Math. Mon. **67**, 520 (1960).
[12]$r_i r_i$ is the same as $\Sigma_{i=1}^{N} r_i r_i$, the summation being understood for a repeated index in a product. This suppression of the summation sign is known as *Einstein's summation convention*.
[13]F. Behroozi and R. Gagnon, Am. Math. Mon. **82**, 804 (1975).
[14]K. Appel and W. Haken, Ill. J. Math. **21**, 429 (1977).

Stability and boundedness results of stochastic differential equations relative to a generalized norm

Marion E. Moore

Department of Mathematics, University of Texas at Arlington, Arlington, Texas 76019
(Received 28 February 1978; revised manuscript received 27 November 1978)

We consider results previously obtained in the context of a generalized norm. In particular, we develop a general comparison principle which permits us to give sufficient conditions for conditional stability and boundedness in the mean relative to a generalized norm. Finally, we provide a decomposition technique together with additional comments which demonstrate the applicability of our results.

1. INTRODUCTION

In this paper we extend some of the results obtained in Ref. 1 to a generalized norm setting. In this setting (cf., Ref. 2), we are able to develop a comparison principle and criteria for conditional stability and boundedness of solutions of a given stochastic differential system.

2. DEFINITIONS AND NOTATIONS

Let R^n denote n-dimensional Euclidean space, $\|\cdot\|_G$ a generalized norm defined from R^n to R^N_+ ($R_+ = [0,\infty)$). Additionally, let $\{\eta = \eta(t), t\in R_+\}$ be a Markov process defined on a probability space (Ω,F,P) which takes values in R. By $E[X/K]$ we shall mean the conditional expectation (mean) of the random vector X subject to the condition(s) K, and by $E_i[\|X(t)\|_G/x_0,\mu_0]$ we mean the ith component $(E[\|X(t)\|_G/x_0,\mu_0])_i$ of $E[\|X(t)\|_G/x_0,\mu_0]$. If $X=(X_1,\ldots,X_n)$ is a random vector, then by the expectation of X, $E(X)$, we naturally mean the vector of expectations of $X_1,\ldots,X_N$. For $k<n$, $M_{(n-k)}$ will denote a manifold of $(n-k)$ dimensions containing the origin.

We shall consider a system of stochastic differential equations

$$x' = f(t,x,\eta(t)), \quad x(t_0)=x_0, \quad \eta(t_0)=\eta_0, \qquad (2.1)$$

where $f\in[R_+\times R^n\times R, R^n]$, for which we shall assume existence of solutions $\{x(t,t_0,x_0),\eta(t,t_0,\eta_0)\}$ for all $t\geqslant t_0\geqslant 0$.

Further, we make the following definitions: The trivial solution of (2.1) is called,

(i) *Conditionally equistable in the mean*, if for each $\epsilon\in R^N_+ - \{0\}$, $t_0\in R_+$, there is a positive function $\delta=\delta(t_0,\epsilon)\in R^N_+ - \{0\}$, continuous in t_0 for each $\epsilon\in R^N_+ - \{0\}$, and such that $x_0\in M_{(n-k)}$ and the inequality $\|x_0\|_G \leqslant\delta$ implies that

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0] < \epsilon, \qquad t\geqslant t_0;$$

(ii) *Conditionally quasiequiasymptotically stable in the mean* if for each $\epsilon,\alpha\in R^N_+ - \{0\}$, $t_0\in R_+$, there exists a positive number $T=T(t_0,\epsilon,\alpha)$ such that

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0] < \epsilon, \quad t\geqslant t_0 + T$$

whenever $\|x_0\|_G\leqslant\alpha$ and $x_0\in M_{(n-k)}$;

(iii) *Conditionally equiasymptotically stable in the mean* if (i) and (ii) hold simultaneously.

Solutions of the stochastic differential system (2.1) are called conditionally,

(iv) *equibounded in the mean* if, for each $\alpha\in R^N_+ - \{0\}$, $t_0\in R_+$, there exists a positive function $\beta=\beta(t_0,\alpha)$, continuous in t_0 for each α, such that $x_0\in M_{(n-k)}$ and the inequality $\|x_0\|_G\leqslant\alpha$ implies that

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0] < \beta, \quad t\geqslant t_0;$$

(v) *quasiequiultimately bounded in the mean* if for each $\alpha\in R^N_+ - \{0\}$, $t_0\in R_+$, there exist positive numbers M and $T=T(t_0,\alpha)$ such that $x_0\in M_{(n-k)}$ and the inequality $\|x_0\|_G\leqslant\alpha$ implies that

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0] < M, \quad t\geqslant t_0+T;$$

(vi) *equiultimately bounded in the mean* if (iv) and (v) hold simultaneously.

Associated with the system (2.1) is the auxiliary differential system

$$u' = g(t,u), \quad u(t_0)=u_0, \qquad (2.2)$$

where $g\in C[R_+\times R^N_+, R^N]$, $g(t,u)$ is quasimonotone nondecreasing in u for fixed $t\in R_+$, and $u(t,t_0,u_0)$ is an arbitrary solution of (2.2).

One now formulates definitions for the system (2.2) corresponding to definitions (i)–(vi) formulated for (2.1). For instance, corresponding to (i) would be

(i′) the trivial solution of (2.2) is *conditionally equistable* if for each $\epsilon\in R^N_+ - \{0\}$, $t\in R_+$, there exists a $\delta=\delta(t_0,\epsilon)$ $\epsilon\in R^N_+ - \{0\}$, continuous in t_0 for each $\in R^N_+ - \{0\}$ such that $u_0=(u_{1_0},\ldots,u_{N_0})\leqslant\delta$ and $u_{i_0}=0$, $i=1,\ldots,k$ implies that $u(t,t_0,u_0)=(u_1,\ldots,u_N)<\epsilon$, $t\geqslant t_0$.

We shall say that a function $b(r)$ belongs to the class K if $b\in C[R^N_+, R^N_+]$, $b(r)=0$ if and only if $r=0$ and b is strictly increasing; a function $a(t,r)$ belongs to the class CK if $a\in C[R_+\times R^N_+, R^N_+]$, $a(t,0)\equiv 0$, and $a(t,r)$ is increasing in r for each $t\in R_+$.

If ϕ: $R^n \to R^m$, then we understand that ϕ is convex if ϕ_i is convex for each $1\leqslant i\leqslant m$; ϕ is concave if $-\phi$ is convex.

3. THE COMPARISON PRINCIPLE

Let $V\in C[R_+\times R^n\times R, R^N_+]$ and define

$$D^+E[V(t,x,\eta)] = \limsup_{h\to 0^+} \frac{1}{h}\{E[V(t+h, x+hf(t,x,\eta),\eta(t+h))/\eta] - V(t,x,\eta)\}.$$

We shall assume that the system (2.2) and the function V satisfy the following hypotheses:

(H_1): $g\in C[R_+\times R^N_+, R^N]$, $g(t,u)$ is concave and quasimonotone nondecreasing in u, for each fixed $t\in R_+$.

(H_2): Assume that $r(t,t_0,u)$ is the maximal solution of the auxiliary Eq. (2.2) existing for $t\geqslant t_0$, $t_0\in R_+$

(H_3): Assume that $g(t,0)\equiv 0$.

(H_4): Let $V\in C[R_+\times R^n\times R, R^N_+]$ and $V(t,x,\eta)$ be locally Lipschitzian in x for fixed t and uniformly in η. Also assume for $(t,x,\eta)\in R_+\times R^n\times R$, that $D^+E[V(t,x,\eta)] \leqslant g(t,V(t,x,\eta))$.

(H_5): $V_i(t,x,\eta)\equiv 0$, for $1\leqslant i<k<n$ iff $x\in M_{(n-k)}$, where $M_{(n-k)}$ is an $(n-k)$-dimensional manifold containing the origin.

(H_6): For $(t,x,\eta)\in R_+\times R^n\times R$, $b(\|x\|_G) \leqslant V(t,x,\eta)\leqslant a(t,\|x\|_G)$, where $b\in K$, b is convex and $a\in CK$.

(H_7): For $(t,x,\eta)\in R_+\times R^n\times R$, $b(\|x\|_G) \leqslant V(t,x,\eta)\leqslant a(t,\|x\|_G)$ where $a\in CK$, $b\in K$, b is convex and for each $1\leqslant i\leqslant N$, their exists j such that $\lim_{\psi_i\to\infty} b_j((0,...,0,\psi_i,0,...,0)) = \infty$.

We shall now include a basic comparison theorem for the stochastic differential system (2.1).

Theorem 3.1: *Suppose that hypotheses* (H_1), (H_2), *and* (H_4) *are satisfied, and that* $\{x(t)=x(t,t_0,x_0,\eta_0),\ \eta(t)=\eta(t,t_0,\eta_0)\}$ *is any solution of* (2.1) *with* $t\geqslant t_0\geqslant 0$ *and* $V(t_0,x_0,\eta_0)\leqslant u_0$. *Then* $E[V(t,x(t),n(t))/x_0,\eta_0]\leqslant r(t,t_0,u_0)$, $t\geqslant t_0$.

Proof: The essentials of the proof are contained in the proof of Theorem 3.1 of Ref. 1. Since the Lipschitz condition given in (H_4) is expressed in terms of a generalized norm, one makes this substitution instead of the usual norm. The Lebesgue convergence theorem, the concavity of g, and Jensen's inequality[3] lead to the inequality

$$D^+m(t)\leqslant g(t,m(t)),\quad t\geqslant t_0.$$

Theorem 4.1.1 of Ref. 4 now completes the proof.

4. CRITERIA FOR CONDITIONAL STABILITY AND BOUNDEDNESS IN THE MEAN

Equipped with the comparison principle (Theorem 3.1) we are now able to develop some criteria for conditional stability and boundedness of solutions of the stochastic differential system (2.1).

Theorem 4.1: *If* (H_i), $i=1,2,...,6$ *hold, and if the system* (2.1) *has the trivial solution, then*

(i) *conditional equistability of the trivial solution of* (2.2) *implies conditional equistability in the mean of the trivial solution of* (2.1);

(ii) *conditional quasiequiasymptotic stability of the solution* $u\equiv 0$ *of* (2.2) *implies conditional quasiequiasymptotic stability in the mean of the trivial solution of* (2.1);

(iii) *conditional equiasymptotic stability of the solution* $u\equiv 0$ *of* (2.2) *implies conditional equiasymptotic stability in the mean of the solution* $x\equiv 0$ *of* (2.1).

Proof of (i): Suppose that $\epsilon\in R^N_+-\{0\}$, $t_0\in R_+$ are given and that the trivial solution of (2.2) is conditionally equistable. Since the hypothesis of Theorem 3.1 is satisfied whenever $V(t_0,x_0,\eta_0)\leqslant u_0$, choose u_0 so that

$$a(t_0,\|x_0\|_G)\leqslant u_0,\quad u_{i_0}=0,\quad i=1,...,k.$$

Hence, by Theorem 3.1 we have

$$E[V(t,x(t),\eta(t))/x_0,\eta_0]\leqslant r(t,t_0,u_0),\quad t\geqslant t_0. \tag{4.1}$$

Suppose there is no $\delta\in R^N_+-\{0\}$ such that $\|x_0\|_G\leqslant\delta$ implies that

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0]<\epsilon.$$

Then for any δ there exists $t_1>t_0$ and a solution $\{x(t),\eta(t)\}$ such that

$$\epsilon_i\leqslant E_i[\|x(t_1)\|_G/x_0,\eta_0]\quad \text{for some } i,\quad 1\leqslant i\leqslant N.$$

Clearly then,

$$(0,...,0,\epsilon_i,0,...,0)\leqslant E[\|x(t_1)\|_G/x_0,\eta_0].$$

Let

$$\nu=\min\{b_j(0,...,\epsilon_i,...,0)\mid 1\leqslant i,j\leqslant N\}$$

and let $\hat\epsilon=(\nu,\nu,...,\nu)$. Then

$$\hat\epsilon\leqslant b((0,...,\epsilon_i,...,0))\leqslant b(E[\|x(t_1)\|_G/x_0,\eta_0]) \leqslant E[b(\|x(t_1)\|_G/x_0,\eta_0)].$$

By the conditional equistability hypothesis, there is δ_1 such that $u_0\leqslant\delta_1$, we have

$$\hat\epsilon\leqslant E[b(\|x(t_1)\|_G/x_0,\eta_0)] \leqslant E[V(t_1,x(t_1),\eta(t_1))/x_0,\eta_0] \leqslant r(t_1,t_0,u_0)<\hat\epsilon.$$

This contradiction together with the observation that $x_0\in M_{(n-k)}$ by (H_5), proves (i).

To prove (ii), let $\epsilon\in R^N_+-\{0\}$, $t_0\in R_+$, $\|x_0\|_G\leqslant\alpha$, and $x_0\in M_{(n-k)}$. By the hypothesis on $a(t,r)$, there is $\alpha_1=\alpha_1(t_0,\alpha)$ such that the inequalities $\|x_0\|_G\leqslant\alpha$ and $a(t_0,\|x_0\|_G)\leqslant\alpha_1$ hold concurrently. We choose u_0 as before, so that (4.1) holds.

Suppose that for $\|x_0\|_G\leqslant\alpha$ and $x_0\in M_{(n-k)}$, the inequality

$$E[\|x(t,t_0,x_0,\eta_0)\|_G/x_0,\eta_0]<\epsilon,\quad \text{for } t\geqslant t_0+T,$$

is false. Then there must exist a sequence $\{t_n\}$, $t_n\geqslant t_0+T$, $t_n\to\infty$ as $n\to\infty$, such that for some solution $\{x(t),\eta(t)\}$ of (2.1) we have

$$E_i[\|x(t_n,t_0,x_0,\eta_0)\|_G/x_0,\eta_0]\geqslant\epsilon_i>0,$$

for some i, $1\leqslant i\leqslant N$.

If we let $\hat\epsilon=(0,...,0,\epsilon_i,0,...,0)$, then

$$E[\|x(t_n,t_0,x_0,\eta_0)\|_G/x_0,\eta_0]\geqslant\hat\epsilon.$$

Since $\hat\epsilon\neq 0$, $b(\hat\epsilon)>0$ and

$$b(\hat\epsilon)\leqslant b(E[\|x(t_n,t_0,x_0,\eta_0)\|_G/x_0,\eta_0]).$$

Since the trivial solution of (2.2) is conditionally quasiequi asymptotically stable, there exists $T=T(t_0,\alpha_1,\epsilon)$ such that

$$u(t_n,t_0,u_0)=(u_1(t_n),...,u_N(t_n))<b(\hat\epsilon),$$

$t_n\geqslant t_0+T$, whenever, $u_0=(u_{1_0},...,u_{N_0})\leqslant\alpha_1$, $u_{i_0}=0$, $i=1,...,k$. By an argument similar to the one used in the proof of (i), we have

$$b(\hat{\epsilon}) \leqslant E\,[V(t_n,x(t_n),\eta(t_n))/x_0,\eta_0] \leqslant (r_1(t_n),\ldots,r_N(t_n)) < b(\hat{\epsilon});$$

a contradiction, which proves (ii).

The proof of (iii) follows from the proofs of (i) and (ii).

Theorem 4.2: *If* (H_1), (H_2), (H_4), (H_5), and (H_7) *are satisfied, then*

(i) *conditionally equiboundedness of solutions of* (2.2) *implies the conditionally equiboundedness in the mean of solutions of* (2.1);

(ii) *conditionally quasiequiultimately boundedness of solutions of* (2.2) *implies the conditionally quasiequiultimately boundedness in the mean of solutions of* (2.1);

(iii) *conditionally equiultimately boundedness of solutions of* (2.2) *implies conditionally equiultimately boundedness in the mean of solutions of* (2.1).

Proof of (i): Let $\{x(t),\eta(t)\}$ be a solution process of (2.1) with $\|x_0\|_G \leqslant \alpha$ and $x_0 \in M_{(n-k)}$. From Theorem 3.1 we have the inequality (4.1). By the hypothesis, given $\alpha_1 > 0$ and $t_0 \in R_+$, there is $\beta = \beta(t_0,\alpha_1)$ continuous in t_0 for each α_1, such that $u(t,t_0,u_0) = (u_1,\ldots,u_N) < \beta$, $t \geqslant t_0$, if $u_0 = (u_{1_0},\ldots u_{N_0}) \leqslant \alpha_1$, where α_1 is defined as in the proof of (ii) of Theorem 4.1.

Let

$$\gamma_i = E([\|x(t,t_0,\gamma_0,\eta_0)\|_G/x_0,\eta_0])_i = E_i[\|x(t)\|_G/x_0,\eta_0], \quad i = 1,\ldots,N$$

and let

$$\psi_i = \sup\{\gamma_i \mid b_j((0,\ldots,0,\gamma_i,0,\ldots,0)) < \beta_j, i = 1,\ldots,N\}$$

(see Ref. 5). We claim that $E\,[\|x(t)\|_G/x_0,\eta_0] < \psi = (\psi_0,\ldots,\psi_N)$. Otherwise there is a solution $\{x(t),\eta(t)\}$ of (2.1) that $\|x_0\|_G \leqslant \alpha$, $x_0 \in M_{(n-k)}$ and for $t_1 \geqslant t_0$, $\psi_i \leqslant E_i[\|x(t_1)\|_G/x_0,\eta_0]$, some i. Note that by (H_7) we have $b_j((0,\ldots,0,\psi_i,0,\ldots,0)) = \beta_j$ for some j. Hence,

$$\begin{aligned}\beta_j &\leqslant b_j(E\,[\|x(t_1)\|_G/x_0,\eta_0]) \\ &\leqslant E_j[b\,(\|x(t_1)\|_G/x_0,\eta_0)] \\ &\leqslant E_j[V(t_1,x(t_1),\eta(t_1))/x_0,\eta_0] \\ &\leqslant r_j(t_1,t_0,u_0) < \beta_j,\end{aligned}$$

a contradiction, which proves (i).

The proofs of (ii) and (iii) are similar in nature to that of (i). To illustrate our results, we consider the following decomposition technique.

Let $x \in \mathbb{R}^n$, $n = \Sigma_{i=1}^N n_i$, and let

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_{n_1} \\ x_{n_1+1} \\ \vdots \\ x_{n_2} \\ \vdots \\ x_{n_N} \end{bmatrix}.$$

Relabeling the x_i as

$$x_i = \begin{bmatrix} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^{n_i} \end{bmatrix},$$

x can then be written

$$\begin{bmatrix} \begin{bmatrix} x_1^1 \\ \vdots \\ x_1^{n_1} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} x_N^1 \\ \vdots \\ x_N^{n_N} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}.$$

Hence, the system

$$x' = f(t,x,\eta(t)), \quad x(t_0) = x_0, \quad \eta(t_0) = \eta_0$$

is equivalent to

$$\begin{aligned} x_1' &= f_1(t,x,\eta(t)) \\ x_2' &= f_2(t,x,\eta(t)) \\ &\vdots \\ x_N' &= f_N(t,x,\eta(t)). \end{aligned}$$

If we let $\alpha_i(x) = \|x_i\|^2 = x_i^T \cdot x_i$, where $\|x\|_G = (\alpha_1(x),\ldots,\alpha_N(x))$, and let $V_i(t,x) = x_i^T P_i x_i$, where P_i is an $n_i \times n_i$ positive definite matrix, and if m_i and M_i are the minimum and maximum eigenvalues of P_i respectively, then $m_i\|x_i\|^2 \leqslant V_i(t,x) \leqslant M_i\|x_i\|^2$, with

$$b_i(\|x\|_G) = m_i\|x_i\|^2, \quad a_i(\|x\|_G) = M_i\|x_i\|^2, \quad i = 1,\ldots,N.$$

Hence

$$b(\|x\|_G) = \begin{bmatrix} m_1\|x_1\|^2 \\ \vdots \\ m_N\|x_N\|^2 \end{bmatrix}, \quad a(\|x\|_G) = \begin{bmatrix} M_1\|x_1\|^2 \\ \vdots \\ M_N\|x_N\|^2 \end{bmatrix}.$$

Then

$$\begin{aligned} D^+E\,[V_i(t,x,\eta)] &= \limsup_{h\to 0} \frac{1}{h}\,[E\,(x_i + h f_i(t,x,\eta))^T \cdot P_i(x_i + h f_i(t,x,\eta)) - x_i^T P_i x_i] \\ &\leqslant \sum_{j=1}^N a_{ij}(t) V_j(t,x,\eta), \quad i = 1,\ldots,N, \end{aligned}$$

which is equivalent to

$$D^+E\,[V(t,x,\eta)] \leqslant A(t)V(t,x,\eta) \equiv g(t,V(t,x,\eta)), \quad A = (a_{ij}).$$

If A is quasimonotone and

$$a_{jj}(t) + \sum_{\substack{i=1 \\ i\neq j}}^N a_{ij}(t) < 0,$$

then the trivial solution of $u' = A(t)u$ is equistable,[6] therefore Theorem 4.1 assures the trivial solution of (2.1) is conditionally equistable in the mean. We remark that the above example is also valid by introducing a factor involving η as in Ref. 7, example 1.

Finally, it should be noted that stability concepts in the mean with regard to the generalized norm imply the corresponding usual stability results. Additionally, the preceding illustration would imply for example, that

$$\|x\|_G < \epsilon \quad \text{gives} \quad \sum_{i=1}^{N} \|x_i\|^2 < \sum_{i=1}^{N} \epsilon_i.$$

Hence, one concludes that conditional equistability in the mean in the context of the generalized norm includes the usual pth mean stability properties, thus unifying the usual stability concepts.

ACKNOWLEDGMENTS

The author wishes to thank Professor V. Lakshmikantham and Professor G. Ladde for their suggestions. Thanks are also due the referee for his comments.

[1]G.S. Ladde, "Systems of Differential Inequalities and Stochastic Differential Equations I," SUNY at Potsdam, preprint (1975).

[2]R.W. Mitchell and M.E. Moore, "Vector Lyapunov Functions and Perturbations of Nonlinear Systems," J. Math. Phys. Sci. **10**, 263–9 (1976).

[3]J.L. Doob, *Stochastic Processes* (Wiley, New York, 1953).

[4]V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities* (Academic, New York, 1969), Vols. I and II.

[5]R.W. Mitchell and D.A. Pace, "Generalized Stability of Motion and Vector Lyapunov Functions," Technical Report No. 42, University of Texas at Arlington, Arlington, Texas.

[6]G.S. Ladde, "Competitive Processes and Comparison Differential Systems," Trans. Am. Math. Soc. **221**, 391–402 (1976).

[7]G.S. Ladde, V. Lakshmikantham, and P.T. Liu, "Differential Inequalities and Stability and Boundedness of Stochastic Differential Equations," J. Math. Anal. Appl. **48**, 341–52 (1974).

Some applications of time-dependent canonical transformations to nonlinear nonconservative classical systems

A. Maharatna, Ranabir Dutt, and D. Chattarji

Department of Physics, Visva-Bharati, Santiniketan 731235, West Bengal, India
(Received 18 July 1978)

The recently proposed time-dependent canonical transformation method of Leach for the quadratic Hamiltonian has been extended to deal with nonlinear nonconservative classical systems. It is observed that the linear time dependent canonical transformations are not adequate to remove the time dependence of any arbitrary time-dependent nonlinear Hamiltonian. Alternatively, we propose that such a Hamiltonian may be transformed to a quadratic form by means of successive nonlinear canonical transformations. It is also shown that the canonical method is useful to obtain solutions for differential equations governing certain dissipative classical systems.

I. INTRODUCTION

Time-dependent classical and quantum oscillators have been studied recently by several authors.[1–7] One of the prime motivations of such studies is to examine whether dynamical symmetries are at all associated with explicitly time dependent systems. It was shown by Fradkin[8] that all classical dynamical problems involving central potentials, inherently possess both O_4 and SU_3 symmetries. Subsequently, Mukunda[9] proved in a very general way that any conservative n-dimensional classical system possesses invariance under O_{n+1} and SU_n algebras, independent of the fact whether the system is linear or nonlinear.

For further generalization of this observation, the problem has been extended to the time-dependent case by several authors. Invoking time-dependent canonical transformations, Günther and Leach[6,7] have recently shown that the Hamiltonian of a time dependent linear system can be transformed to a time independent form, which enables one to construct the generators of the underlying symmetry groups. This then implies that the dynamical symmetries associated with a linear conservative system remain unaltered even if the system becomes explicitly time dependent.

One question which then remains is whether such algebraic structures are preserved in the case of a more general situation, i.e., when the system is nonlinear as well as nonconservative. To determine it, one may proceed in either of the two ways: The time-dependent nonlinear Hamiltonian may be transformed to a time-independent form with the same order of nonlinearity by means of linear transformations as considered by Leach.[7] In the alternative approach, the nonlinear nonconservative Hamiltonian may be transformed to another time-dependent form but with a lower order of nonlinearity, by means of canonical transformations which are nonlinear in contrast to the linear ones considered in the first approach. By successive applications of such transformations one may then finally recast the Hamiltonian into a quadratic time-dependent form which consequently preserves invariance properties.[7]

In Sec. II, we shall show explicitly that in the first approach only a restricted class of time-dependent nonlinear Hamiltonians can be converted to time-dependent form with the help of linear canonical transformations. We shall show, however, that the elimination of time dependence from such Hamiltonians may be possible systematically by successive applications of time-dependent nonlinear canonical transformations. This will be discussed in the context of damped Duffing oscillator in Sec. III. In Sec. IV, we further show that the canonical method is useful to obtain solutions of differential equations governing a class of dissipative systems. To deal with such problems, Denman and Buch[3] proposed that the Hamilton–Jacobi technique of classical mechanics is applicable if the coordinate and temporal components are factored in the form of product instead of addition, as is usually done in the case of conservative system.[10] Our method may be considered as an alternative to the approach suggested by these authors.

II. LINEAR CANONICAL TRANSFORMATIONS FOR NONLINEAR NONCONSERVATIVE HAMILTONIAN

To apply linear canonical transformations of the type prescribed by Leach[7] to time-dependent nonlinear systems, we consider the Hamiltonian for a simple one-dimensional case with the lowest order nonlinearity (i.e., of cubic type) in addition to the quadratic part that describes a damped harmonic oscillator. We thus write

$$H(x,p,t) = \tfrac{1}{2}p^2 e^{-2\gamma t} + \tfrac{1}{2}\omega^2 x^2 e^{2\gamma t} + \tfrac{1}{3}A(t)\,x^3 + \tfrac{1}{3}B(t)\,p^3 + C(t)\,x^2 p + D(t)\,xp^2, \tag{2.1}$$

where A, B, C, and D are at the moment arbitrary functions of time. The canonical equations of motion obtained from (2.1) are

$$\dot{x} = \partial H/\partial p = pe^{-2\gamma t} + Bp^2 + Cx^2 + 2Dxp, \tag{2.2a}$$

$$\dot{p} = -\partial H/\partial x = -(\omega^2 x e^{2\gamma t} + Ax^2 + 2Cxp + Dp^2). \tag{2.2b}$$

We require that the Hamiltonian in (2.1) containing cubic nonlinear terms may be transformed into a time independent form having the same order of nonlinearity

$$\bar{H}(X,P) = \tfrac{1}{2}(P^2 + \Omega^2 X^2) + \tfrac{1}{3} aX^3 + \tfrac{1}{3}bP^3 + cX^2P + dXP^2 \tag{2.3}$$

with $\Omega^2 = \omega^2 - \gamma^2$, by means of the linear transformations

$$X = \Lambda^1_1(t)x + \Lambda^1_2(t)p, \tag{2.4a}$$

$$P = \Lambda^2_1(t)x + \Lambda^2_2(t)p. \tag{2.4b}$$

If the transformations (2.4) are to be canonical, the transformed variables X and P must satisfy the following equations:

$$\dot{X} = \partial\bar{H}/\partial P = P + bP^2 + cX^2 + 2dXP, \tag{2.5a}$$

$$\dot{P} = -\partial\bar{H}/\partial X = -(\Omega^2 X + aX^2 + 2cXP + dP^2). \tag{2.5b}$$

In order to determine the unknown time-dependent coefficients appearing in (2.3) and (2.4), the relations in (2.2), (2.4), and (2.5) are combined in such a manner that the new canonical variables are replaced by the old ones. This is effected by taking time derivative of the relations in (2.4), replacing $\dot{X}$ and $\dot{P}$ by the expressions (2.5), and then substituting $\dot{x}$ and $\dot{p}$ by the quantities given in (2.2). Finally we equate the coefficients of like powers of x, p, x^2, p^2, xp, etc. from either side of the equations and obtain the following relations among the parameters:

$$\begin{aligned}
\dot{\Lambda}^1_1 &= \Lambda^1_2\omega^2 e^{2\gamma t} + \Lambda^2_1,\\
\dot{\Lambda}^1_2 &= -\Lambda^1_1 e^{-2\gamma t} + \Lambda^2_2,\\
\dot{\Lambda}^2_1 &= \Lambda^2_2\omega^2 e^{2\gamma t} - \Omega^2\Lambda^1_1,\\
\dot{\Lambda}^2_2 &= -\Lambda^2_1 e^{-2\gamma t} - \Omega^2\Lambda^1_2.
\end{aligned} \tag{2.6}$$

and

$$\begin{aligned}
&c(\Lambda^1_1)^2 + 2d\Lambda^1_1\Lambda^2_1 + b(\Lambda^2_1)^2 - C\Lambda^1_1 + A\Lambda^1_2 = 0,\\
&c(\Lambda^1_2)^2 + 2d\Lambda^1_2\Lambda^2_2 + b(\Lambda^1_1)^2 - B\Lambda^1_1 + D\Lambda^1_2 = 0,\\
&c\Lambda^1_2\Lambda^1_1 + d(\Lambda^1_1\Lambda^2_2 + \Lambda^1_2\Lambda^2_1)\\
&\quad + b\Lambda^2_1\Lambda^2_2 - D\Lambda^1_1 + C\Lambda^1_2 = 0,\\
&a(\Lambda^1_1)^2 + 2c\Lambda^1_1\Lambda^2_1 + d(\Lambda^2_1)^2 - A\Lambda^2_2 + C\Lambda^2_1 = 0,\\
&a(\Lambda^1_2)^2 + 2c\Lambda^1_2\Lambda^2_2 + d(\Lambda^2_2)^2 - D\Lambda^2_2 + B\Lambda^2_1 = 0,\\
&a\Lambda^1_1\Lambda^1_2 + c(\Lambda^1_1\Lambda^2_2 + \Lambda^1_2\Lambda^2_1)\\
&\quad + d\Lambda^2_1\Lambda^2_2 - C\Lambda^2_2 + D\Lambda^2_1 = 0.
\end{aligned} \tag{2.7}$$

The time dependence of A, B, C, and D can be determined after evaluating the values of Λ's from Eqs. (2.6). The first order coupled equations (2.6) may be solved by the well-known matrix method.[11] Equations (2.6) may be written in a more convenient form

$$\frac{dZ}{dt} = AZ, \tag{2.8}$$

where Z is a column matrix given by

$$Z \equiv \begin{pmatrix} Z_1\\ Z_2\\ Z_3\\ Z_4 \end{pmatrix} = \begin{pmatrix} \Lambda^1_1 e^{-\gamma t}\\ \Lambda^1_2 e^{\gamma t}\\ \Lambda^2_1 e^{-\gamma t}\\ \Lambda^2_2 e^{\gamma t} \end{pmatrix} \tag{2.9}$$

and

$$A = \begin{pmatrix} -\gamma & \omega^2 & 1 & 0\\ -1 & \gamma & 0 & 1\\ -\Omega^2 & 0 & -\gamma & \omega^2\\ 0 & -\Omega^2 & -1 & \gamma \end{pmatrix}. \tag{2.10}$$

To get the solution for the vector Z, one needs to diagonalize the matrix A. For this purpose we define

$$Z \equiv Ty \tag{2.11}$$

so that

$$\frac{dy}{dt} = T^{-1}ATy. \tag{2.12}$$

The matrix T would be constructed such that $T^{-1}AT$ is a diagonal matrix

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_3 & 0\\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}, \tag{2.13}$$

where the λ's are the characteristic roots of the matrix A and are given by

$$\lambda_1 = \lambda_2 = 0,\quad \lambda_3 = 2i\Omega,\quad \lambda_4 = -2i\Omega. \tag{2.14}$$

Thus from (2.10), (2.13), and (2.14) we get

$$T = \begin{pmatrix} 1 & 0 & 2\omega^2 & 2\omega^2\\ 0 & 1 & 2(\gamma + i\Omega) & 2(\gamma - i\Omega)\\ \gamma & -\omega^2 & 2i\Omega\omega^2 & -2i\Omega\omega^2\\ 1 & -\gamma & 2i\Omega(\gamma + i\Omega) & -2i\Omega(\gamma - i\Omega) \end{pmatrix}. \tag{2.15}$$

Also from (2.12) and (2.13) we obtain

$$y = \begin{pmatrix} K_1\\ K_2\\ K_3 e^{2i\Omega t}\\ K_4 e^{-2i\Omega t} \end{pmatrix}, \tag{2.16}$$

where the K_i's are arbitrary constants. Combining Eqs. (2.9), (2.11), (2.15), and (2.16), it is then trivial to obtain the general solutions for Λ's which satisfy Eqs. (2.6). The solutions are

$$\begin{aligned}
\Lambda^1_1(t) &= K_1 e^{\gamma t} + 2\omega^2 K_3 e^{(\gamma + 2i\Omega)t} + 2\omega^2 K_4 e^{(\gamma - 2i\Omega)t},\\
\Lambda^1_2(t) &= K_2 e^{-\gamma t} + 2(\gamma + i\Omega)K_3 e^{-(\gamma - 2i\Omega)t}\\
&\quad + 2(\gamma - i\Omega)K_4 e^{-(\gamma + 2i\Omega)t},\\
\Lambda^2_1(t) &= \gamma K_1 e^{\gamma t} - \omega^2 K_2 e^{\gamma t} + 2i\omega^2\Omega K_3 e^{(\gamma + 2i\Omega)t}\\
&\quad - 2i\omega^2\Omega K_4 e^{(\gamma - 2i\Omega)t},\\
\Lambda^2_2(t) &= K_1 e^{-\gamma t} - \gamma K_2 e^{-\gamma t} + 2i\Omega(\gamma + i\Omega)K_3 e^{-(\gamma - 2i\Omega)t}\\
&\quad - 2i\Omega(\gamma - i\Omega)K_4 e^{-(\gamma + 2i\Omega)t}.
\end{aligned} \tag{2.17}$$

Instead of using these general expressions for Λ's, we shall use, however, particular sets of values that give simple results. To determine the explicit forms of the time-dependent coefficients, we use the following sets of simple solutions:

I. $K_1 = 1,\quad K_2 = K_3 = K_4 = 0$

$$\Lambda^1_1(t) = e^{\gamma t},\quad \Lambda^1_2(t) = 0,\qquad \Lambda^2_1(t) = \gamma e^{\gamma t},\quad \Lambda^2_2(t) = e^{-\gamma t}; \tag{2.18a}$$

$$\text{II.}\quad K_2 = 1/\omega,\quad K_1 = K_3 = K_4 = 0$$
$$\Lambda^1_1(t) = 0,\quad \Lambda^1_2(t) = e^{-\gamma t}/\omega, \qquad (2.18b)$$
$$\Lambda^2_1(t) = -\omega e^{\gamma t},\quad \Lambda^2_2(t) = -\gamma e^{-\gamma t}/\omega.$$

This amounts to considering only particular types of transformations out of a class of equivalent canonical transformations[12] for replacing one Hamiltonian by the other.

With the values of Λ 's in (2.18a), Eqs. (2.7) are found to be consistent only if the temporal dependence of A, B, C, D occur in the form:

$$\begin{aligned} A(t) &= (a + 3c\gamma + 3d\gamma^2 + b\gamma^3)\, e^{3\gamma t},\\ B(t) &= be^{-3\gamma t},\\ C(t) &= (c + 2d\gamma + b\gamma^2)\, e^{\gamma t},\\ D(t) &= (d + b\gamma)\, e^{-\gamma t}. \end{aligned} \qquad (2.19)$$

Similarly, for the set of values of Λ 's in (2.18b) we get from Eqs. (2.7)

$$\begin{aligned} A(t) &= -b\omega^3 e^{3\gamma t},\\ B(t) &= \frac{1}{\omega^3}(a - 3c\gamma + 3d\gamma^2 - b\gamma^3)\, e^{-3\gamma t},\\ C(t) &= \omega(d - b\gamma)\, K_1,\\ D(t) &= -\frac{1}{\omega}(c - 2d\gamma + b\gamma^2)\, e^{-\gamma t}. \end{aligned} \qquad (2.20)$$

It is then clear that although an arbitrary time-dependent nonlinear Hamiltonian cannot be recasted into a time-independent form through linear time-dependent canonical tansformations, only specific forms of time dependence of the coefficients as given in (2.19) and (2.20) may be dealt with by Leach's prescription.

III. NONLINEAR CANONICAL TRANSFORMATIONS

In this section, we attempt to show that the removal of explicit time dependence of a nonlinear Hamiltonian may be possible in successive steps by means of nonlinear canonical transformations. For illustrations of our method, we consider the Hamiltonian

$$H(x,p,t) = \tfrac{1}{2}p^2 e^{-2\gamma t} + \tfrac{1}{2}\omega^2 x^2 e^{2\gamma t} - \tfrac{1}{4}\epsilon x^4 e^{2\gamma t}, \qquad (3.1)$$

which describes a damped Duffing oscillator corresponding to the pair of canonical equations

$$\begin{aligned} \dot{x} &= p\, e^{-2\gamma t},\\ \dot{p} &= -(\omega^2 x - \epsilon x^3)\, e^{2\gamma t}. \end{aligned} \qquad (3.2)$$

We now require that the Hamiltonian in (3.1) containing quartic nonlinear term be transformed to another time-dependent form with the next lower order (cubic) nonlinearity

$$\bar{H}(X,P,t) = \tfrac{1}{2}(P^2 + \Omega^2 X^2) + a(t)X^2 P + b(t)XP^2 + \tfrac{1}{3}c(t)X^3 + \tfrac{1}{3}d(t)P^3. \qquad (3.3)$$

by means of the nonlinear transformations

$$\begin{aligned} X &= \Lambda^1_1(t)x + \Lambda^1_2(t)p + C_1(t)x^2 + D_1(t)p^2 + E_1(t)xp,\\ P &= \Lambda^2_1(t)x + \Lambda^2_2(t)p + C_2(t)x^2 + D_2(t)p^2 + E_2(t)xp, \end{aligned} \qquad (3.4)$$

If the transformations (3.4) are to be canonical, the transformed dynamical variables, X and P, must satisfy the canonical equations of motion

$$\begin{aligned} \dot{X} &= P + aX^2 + 2bXP + dP^2,\\ \dot{P} &= -(\Omega^2 X + 2aXP + bP^2 + cX^2). \end{aligned} \qquad (3.5)$$

Our task remains in determining the time-dependent coefficients in the transformed Hamiltonian and the transformation relations (3.4). We follow the same procedure as already discussed in the earlier section. Removing the dynamical variables and their derivatives from Eqs. (3.2), (3.4), and (3.5), we obtain the same set of four linear coupled differential equations among Λ 's [as given in Eqs. (2.6)] and 24 other relations involving Λ 's and various time-dependent parameters. Using the Λ 's given in (2.18a) in those relations, we find that for consistency one requires

$$\begin{aligned} &b = d = C_1 = D_1 = D_2 = E_1 = E_2 = 0,\\ &a(t) = \mu e^{-\gamma t},\\ &c(t) = -\mu\gamma e^{-\gamma t},\\ &C_2(t) = -\mu e^{\gamma t}, \end{aligned} \qquad (3.6)$$

where $\mu = (\epsilon/2)^{1/2}$. Thus the reduced Hamiltonian (3.3) and the appropriate time-dependent nonlinear transformations become

$$\bar{H} = \tfrac{1}{2}(P^2 + \Omega^2 X^2) + \mu e^{-\gamma t}X^2 P - \tfrac{1}{3}\mu\gamma e^{-\gamma t}X^3, \qquad (3.7)$$

$$\begin{aligned} X(t) &= xe^{\gamma t},\\ P(t) &= \gamma xe^{\gamma t} + pe^{-\gamma t} - \mu x^2 e^{\gamma t}. \end{aligned} \qquad (3.8)$$

Similar solutions for the time-dependent parameters corresponding to the other set of values of Λ 's in (2.18b) give the reduced Hamiltonian and the canonical transformations in the form

$$\begin{aligned} \bar{H}(X,P,t) = {}&\tfrac{1}{2}(P^2 + \Omega^2 X^2)\\ &+ (\mu/\omega^3)e^{-\gamma t}[-2\gamma\Omega^2 X^2 P\\ &+ (2\gamma^2 - \omega^2)XP^2\\ &- \tfrac{1}{3}\gamma^2(3\omega^2 - 2\gamma^2)X^3 + \tfrac{2}{3}\gamma P^3] \end{aligned} \qquad (3.9)$$

$$\begin{aligned} X(t) &= (1/\omega)p\, e^{-\gamma t} + (\mu/\omega)x^2 e^{\gamma t},\\ P(t) &= -\omega^2 xe^{\gamma t} - (\gamma/\omega)p\, e^{-\gamma t} - (\gamma/\omega)\mu x^2 e^{\gamma t}. \end{aligned} \qquad (3.10)$$

The results in (3.7)–(3.10) clearly indicate that a quartic nonlinear time-dependent Hamiltonian may be reduced to a number of time-dependent forms with cubic nonlinearity through the application of equivalent canonical transformations. It is important to mention here that if the general solutions fo Λ 's in (2.17) are used in (2.7), the time-dependent coefficients, such as a, b, C_1, D_2, etc. depend in general on the four unknown constants K 's. One is then free to choose the constants conveniently to reduce algebraic complexities of the calculations.

Since our ultimate aim is to convert the Hamiltonian (3.3) into a time-dependent quadratic one (which may be subsequently reduced to time-independent form by Leach's method), we repeat similar canonical procedure. The steps showing further reduction are presented in the Appendix. As we have considered here a simple Hamiltonian with quartic nonlinearity, two successive canonical transformations are enough to reduce the problem to a quadratic one. However, for a more complicated time-dependent nonlinear system, not only the number of successive canonical steps would be greater but also the form of the transformations will be more complex in general as compared to those in (3.8) and (3.10).

IV. CANONICAL METHOD OF SOLVING DIFFERENTIAL EQUATIONS FOR DISSIPATIVE SYSTEMS

The canonical approach may be used as well to obtain solutions for certain differential equations associated with harmonic Hamiltonian, which have been treated earlier by the Hamilton-Jacobi method.[3] We consider the equation

$$\ddot{x} + \gamma \dot{x} + g = 0, \tag{4.1}$$

which is canonical to the Hamiltonian

$$H(x,p,t) = \tfrac{1}{2} p^2 e^{-\gamma t} + gx\, e^{\gamma t}. \tag{4.2}$$

This may be transformed to the time-independent harmonic oscillator form

$$\bar{H}(X,P) = \tfrac{1}{2}(P^2 + X^2) \tag{4.3}$$

through the transformations

$$\begin{aligned} X &= A^1_1(t)x + A^1_2(t)p + r_1(t), \\ P &= A^2_1(t)x + A^2_2(t)p + r_2(t). \end{aligned} \tag{4.4}$$

It is important to mention that since the Hamiltonian in (4.2) and (4.3) are both quadratic, it is enough to take linear forms of the transformations relating the old and new canonical variables. Proceeding in the same manner discussed in the last two sections, we eliminate the dynamical variables to obtain the following equations for the time-dependent parameters:

$$\begin{aligned} \dot{A}^1_1 &= A^2_1, \\ \dot{A}^1_2 &= -A^1_1 e^{-\gamma t} + A^2_2, \\ \dot{A}^2_1 &= -A^1_1, \\ \dot{A}^2_2 &= -A^2_1 e^{-\gamma t} - A^1_2 \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} \dot{r}_1 &= r_2 + gA^1_2 e^{\gamma t}, \\ \dot{r}_2 &= -r_1 + gA^2_2 e^{\gamma t}. \end{aligned} \tag{4.6}$$

The solutions of Eqs. (4.5) are

$$\begin{aligned} A^1_1 &= A \sin(t+\alpha), \\ A^1_2 &= B \sin(t+\beta) + (A/\gamma)\, e^{-\gamma t} \sin(t+\alpha), \\ A^2_1 &= A \cos(t+\alpha), \\ A^2_2 &= B \cos(t+\beta) + (A/\gamma)\, e^{-\gamma t} \cos(t+\alpha), \end{aligned} \tag{4.7}$$

which when substituted in (4.6) give

$$\begin{aligned} r_1 &= C \sin(t+\delta) + (gA/\gamma)t \sin(t+\alpha) \\ &\quad + (gB/\gamma)e^{\gamma t} \sin(t+\beta) \\ r_2 &= C\cos(t+\delta) + (gA/\gamma)t\cos(t+\alpha) \\ &\quad + (gB/\gamma)e^{\gamma t}\cos(t+\beta), \end{aligned} \tag{4.8}$$

where A, B, α, β, etc. are arbitrary constants. It is clear from the Hamiltonian (4.3) that $X(t)$ is the coordinate of an undamped harmonic oscillator and hence

$$\begin{aligned} X(t) &= D \sin(t+\epsilon), \\ P(t) &= D\cos(t+\epsilon). \end{aligned} \tag{4.9}$$

Now from Eqs. (4.4), (4.7), (4.8), and (4.9) one may obtain

$$x(t) = E - (F/\gamma)e^{-\gamma t} - (g/\gamma)\, t, \tag{4.10}$$

which is the required solution of Eq. (4.1). The same result was obtained by Denman and Buch[3] by employing Hamilton–Jacobi technique.

We have checked also that other time-dependent problems considered by these authors are solvable by our canonical prescription. The advantage of this method is that one only needs to invoke appropriate canonical transformations to convert the time-depedent Hamiltonian to a suitable time-independent Hamiltonian for which either the exact or the near approximate solution is known. It is then trivial to revert the canonical transformations and to obtain the solution for the original nonconservative system.

V. DISCUSSIONS

In the present paper we have shown that linear time-dependent canonical prescription of Leach is useful to remove time dependence from only a restricted class of nonlinear Hamiltonians. However, it is observed that one may need to replace the linear transformation relations by appropriate nonlinear time-dependent canonical transformations for dealing with more general nonlinear nonconservative classical systems. Although we have restricted our discussion to one-dimensional cases, it is worthy to investigate the applicability of the ideas used here to general Hamiltonian systems. It has eben pointed out by Kohler[13] that classically every $2n$-dimensional Hamiltonian may be transformed to any other n-dimensional Hamiltonian by suitable canonical transformations. Based on this remark, we expect to develop a more rigorous prescription for reducing a general nonlinear time-dependent Hamiltonian to a time-independent quadratic form which is necessary for obtaining appropriate invariants[7] associated with the system. We finally observe that the canonical technique provides an elegant and straightforward method for obtaining solutions of the equation of motion describing certain dissipative classical systems.

APPENDIX: CONVERSION OF CUBIC NONLINEAR HAMILTONIAN TO THE QUADRATIC FORM

In Sec. III, we have explained how a given time-dependent Hamiltonian with quartic nonlinearity can be converted to another time-dependent form of the type (3.3) with cubic nonlinear terms. We here wish to show precisely the necessary steps for further reduction of this Hamiltonian to the quadratic time-dependent form

$$\bar{\bar{H}}(X_1,P_1,t) = \tfrac{1}{2}\alpha(t)P_1^2 + \tfrac{1}{2}\beta(t)X_1^2 + \delta(t)X_1P_1, \tag{A1}$$

in which α, β, and δ are to be determined. The new dynamical variables, X_1 and P_1 which obey the canonical equations

$$\begin{aligned} \dot{X}_1 &= \alpha P_1 + \delta X_1, \\ \dot{P}_1 &= \beta X_1 + \delta P_1, \end{aligned} \tag{A2}$$

are assumed to be connected to the old coordinates, X and P, by the relations

$$X_1(t) = S^1_1(t)X + S^1_2(t)P + L_1(t)X^2 + M_1(t)P^2 + N_1(t)XP,$$
$$P_1(t) = S^2_1(t)X + S^2_2(t)P + L_2(t)X^2 + M_2(t)P^2 + N_2(t)XP. \quad \text{(A3)}$$

Taking time derivative of the relations (A3) and replacing $(\dot{X},\dot{P})$ and $(\dot{X}_1,\dot{P}_1)$ by the Eqs. (3.5) and (A2), we equate the coefficients of the like powers of X, P, X^2, XP, etc. to get

$$\begin{aligned}
\dot{S}^1_1 &= \alpha S^2_1 + \delta S^1_1 + \Omega^2 S^1_2,\\
\dot{S}^1_2 &= \alpha S^2_2 + \delta S^1_2 - S^1_1,\\
\dot{S}^2_1 &= -\beta S^1_1 + \delta S^2_1 + \Omega^2 S^2_2,\\
\dot{S}^2_2 &= -\beta S^1_2 - \delta S^2_2 - S^2_1,\\
\dot{L}_1 &= \alpha L_2 + \delta L_1 - aS^1_1 + cS^1_2 + \Omega^2 N_1,\\
\dot{M}_1 &= \alpha M_2 + \delta M_1 - dS^1_1 + bS^1_2 - N_1,\\
\dot{N}_1 &= \alpha N_2 + \delta N_1 - 2bS^1_1 + 2aS^1_2 - 2L_1 + 2\Omega^2 M_1,\\
\dot{L}_2 &= -\beta L_1 - \delta L_2 - aS^2_1 + cS^2_2 + \Omega^2 N_2,\\
\dot{M}_2 &= -\beta M_1 - \delta M_2 - dS^2_1 + bS^2_2 - N_2,\\
\dot{N}_2 &= -\beta N_1 - \delta N_2 - 2bS^2_1 + 2aS^2_2 - 2L_2 + 2\Omega^2 M_2
\end{aligned} \quad \text{(A4)}$$

and

$$\begin{aligned}
&2aL_1 - cN_1 = 0,\\
&2bM_1 - dN_1 = 0,\\
&4bL_1 - 2cM_1 - aN_1 = 0,\\
&2dL_1 - 4aM_1 + bN_1 = 0,\\
&2aL_2 - cN_2 = 0,\\
&2bM_2 - dN_2 = 0,\\
&4bL_2 - 2cM_2 - aN_2 = 0,\\
&2dL_2 - 4aM_2 + bN_2 = 0.
\end{aligned} \quad \text{(A5)}$$

One may then check that the coupled equations (A4) are solvable if the conditions

$$\begin{aligned}
&2b - ad/b - a^2/c = 0,\\
&d - 2a^2d/bc + ab/c = 0,
\end{aligned} \quad \text{(A6)}$$

which are obtained from Eqs. (A5), are satisfied. As we have already mentioned that the time-dependent parameters, a, b, c, and d are in general dependent on the specific choice of the arbitrary constants K's appearing in (2.17), the conditions (A6) may be made valid by proper adjustments of these constants.

[1] H.R. Lewis, Jr., Phys. Rev. Lett. **18**, 510–12 (1967); H.R. Lewis, Jr., J. Math. Phys. **9**, 1976–86 (1968).

[2] B. Leaf, J. Math. Phys. **10**, 1971–79 (1969); B. Leaf, J. Math. Phys. **10**, 1980–87 (1969); M. Moshinsky, J. Math. Phys. **12**, 1772–83 (1971).

[3] H.H. Denman and L.M. Buch, J. Math. Phys. **14**, 326–29 (1973).

[4] L.H. Buch and H.H. Denman, Phys. Rev. **11**, 279–81 (1975).

[5] R. Dutt, Phys. Lett. A **59**, 345–47 (1976).

[6] N.J. Günther and P.G.L. Leach, J. Math. Phys. **8**, 572–76 (1977).

[7] P.G.L. Leach, J. Math. Phys. **18**, 1608–11 (1977); P.G.L. Leach, J. Math. Phys. **18**, 1902–07 (1977); P.G.L. Leach, J. Math. Phys. **19**, 446–51 (1978).

[8] D.M. Fradkin, Am. J. Phys. **33**, 798–812 (1967).

[9] N. Mukunda, Phys. Rev. **155**, 1383–86 (1967).

[10] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1950), Chap. 9.

[11] R. Bellman, *Introduction to Matrix Analysis* (McGraw-Hill, New York, 1965); M.C. Pease, *Methods of Matrix Algebra* (Academic, New York, 1970), Vol. III.

[12] E.C.G. Sudarshan and N. Mukunda, *Classical Dynamics—A Modern Perspective* (Wiley, New York, 1974), Chaps. 5 and 6; N. Mukunda, Pramana **2**, 1–14 (1974).

[13] R.H. Kohler, Found. Phys. **6**, 193–208 (1976).

Invariant *-product quantization of the one-dimensional Kepler problem

Christian Fronsdal

Department of Physics, University of California, Los Angeles, California 90024

(Received 6 December 1978)

It is shown that, with proper interpretation, phase space trajectories for the one-dimensional Kepler problem retain their significance after quantization. The classical and quantum theories are both entirely conventional; only the Wigner correspondence (Weyl quantization rule) is not. The usual quantum theory may be obtained by so(2,1)-invariant *-product quantization (generalized Moyal mechanics). [The analog in three dimensions is so(4,2) invariant *-quantization; it is believed that this should encounter no new difficulties of principle.] New results concerning invariant *-products on polynomials, in the case of so(2,1), are presented. The Kepler problem is the first known example of a nonanalytic *-representation. Piecewise analytic *-representations are defined and are shown to provide a general framework for invariant quantization on singular orbits of semisimple groups. When piecewise analytic *-representations are allowed, then the specification of an invariant *-product on polynomials is no longer sufficient to determine a unique quantum theory.

I. INTRODUCTION

Distinguished observables: Some very special physical systems possess the remarkable property that phase space trajectories continue to be meaningful after quantization. The harmonic oscillator is a well-known example[1]; the Kepler problem is another instance—as will be shown here. A related phenomenon is the high degree of accuracy of the WKB approximation for such systems. Recently it has been conjectured that these properties may be shared by a class of very interesting field theories; this would imply that solutions of the classical field equations would have a meaning in the corresponding quantum field theory. Let us agree to refer to such systems as "systems with distinguished Hamiltonians."

In order to decide whether a system has a distinguished Hamiltonian, it is necessary to known (i) the classical version, (ii) the quantum version, and (iii) the Wigner correspondence. The last maps operators in Hilbert space to functions on phase space: $\hat{F} \to \Xi(\hat{F}) = F$. The equation of motion $i\hbar(d/dt)\hat{F} = [\hat{F},\hat{H}]$ becomes

$$i\hbar(d/dt)F = [F\star H], \tag{1.1}$$

where the $\star$-product is defined by $F\star H = \Xi(\hat{F}\hat{H})$ and $[F\star H] = F\star H - H\star F$. This equation defines a trajectory in phase space if and only if the map $F \to [F\star H]$ is a derivation of the algebra of functions on phase space, with the ordinary product.[2] This happens if the bracket $[F\star H]$ coincides with the Poisson bracket $\{F,H\}$.

Definition 1: A system with classical Hamiltonian function H and Wigner correspondence Ξ is said to have a *distinguished Hamiltonian* if $\Xi([\hat{F},\hat{H}]) = i\hbar\{F,H\}$, for all operators $\hat{F}$ in the domain of Ξ. The exact form of the Wigner correspondence is not unique, even when both the classical and quantum descriptions are given, and the question of whether a system does or does not have a distinguished Hamiltonian depends on the choice of the Wigner correspondence.

A distinguished Hamiltonian defines trajectories in phase space that can be associated with the time development of the system, but these are not the only trajectories of interest. In particular, it is usually true that symmetries of the system are associated with trajectories. For example, the generators $(\hat{L}^A)$, $A = 1,\ldots,10$, of Poincaré transformations in conventional quantum field theories are bilinear expressions in the elementary field operators and the condition $\Xi([\hat{F},\hat{L}^A]) = i\hbar\{F,L^A\}$ is satisfied by the usual Wigner correspondence. This ensures that coordinate transformations retain their geometrical meaning after quantization. It is well known, however, that this correspondence of $[\hat{A},\hat{B}]$ with $\{A,B\}$ cannot extend to all operators $\hat{A}$, $\hat{B}$ in the domain of Ξ, and this leads to

Definition 2: The *distinguished observables* of a quantized system with Wigner correspondence Ξ are those associated with self-adjoint operators $\hat{L}$ such that $\Xi([\hat{L},\hat{F}]) = i\hbar\{L,F\}$ for all operators $\hat{F}$ in the domain Ξ.

It is evident that the distinguished observables form a Lie algebra, the $\hat{L}$'s with respect to [,] and the L's with respect to { }.

Invariant quantization: Let a classical mechanical system be given, and an algebra $\mathscr{A}$ of classical observables; that is, a linear space of functions on phase space, closed under the Poisson bracket; the corresponding abstract Lie algebra is also denoted $\mathscr{A}$. Let T be a representation of $\mathscr{A}$ by self-adjoint operators in a Hilbert space. Let $\hat{L} \equiv i\hbar T(L)$, for L in $\mathscr{A}$ and let Ξ be a Wigner correspondence. Then $\mathscr{A}$ will be an

 0022-2488/79/112226-07$01.00

algebra of distinguished observables if Ξ is invariant in the sense of the following[3]:

Definition 3: Ξ is said to be $\mathscr{A}$-invariant if (i) $\Xi(\hat{L}) = L$ for every L in $\mathscr{A}$ and (ii) $\Xi([\hat{L},\hat{F}]) = i\hbar\{L,F\}$ for every L in $\mathscr{A}$ and every $\hat{F}$ in the domain of Ξ. [$F \equiv \Xi(\hat{F})$.]

The problem of invariant (more precisely, $\mathscr{A}$-invariant) quantization of a classical mechanical system with an algebra $\mathscr{A}$ of classical observables is to find all possible pairs (T,Ξ) such that Ξ is $\mathscr{A}$-invariant in the sense of Definition 3. This is of particular interest when $\mathscr{A}$ is "large enough" to provide a system of coordinates for phase space, that is, when phase space can be identified with an orbit of the coadjoint action in the real vector space dual of $\mathscr{A}$.

The Kepler problem: The principal aim of this paper is to construct an so(4,2)-invariant Wigner correspondence for the Kepler problem, that is, an invariant quantization procedure that leads to the conventional quantum theory of the hydrogen atom. It should be emphasized that both classical and quantum systems are conventional; only the Wigner correspondence Ξ, and thus the associated $\star$-product and Moyal mechanics, are new. We shall simplify the calculations by dealing with the one-dimensional rather than the three-dimensional Kepler problem; this does not alleviate the principal difficulties. Our interest arises from the following considerations.

(i) The (regularized) Hamiltonian is one of the generators of $\mathscr{A} = \mathrm{so}(4,2)$ [4]; it follows that if an invariant Wigner correspondence exists, the hydrogen atom belongs to the class of systems for which phase space trajectories retain their meaning after quantization.

(ii) Invariant ($\star$-product) quantization seems to have more general applicability than the very much related geometric quantization procedures of Souriau[5] and Kostant.[6] Geometric quantization cannot handle the Kepler problem because of the nonexistence of "invariant polarizations," and this presents us with a challenge.

(iii) the phase space of the Kepler problem can be identified with an orbit in the dual of so(4,2). The fact that the orbit in question happens to be very singular gives rise to special difficulties, and it comes as no surprise that the associated "$\star$-representation" turns out to be nonanalytic. This is the first known example of a singular $\star$-representation.

Outline: The physical system is defined in Sec. II. In Sec. III it is shown that an invariant Wigner correspondence Ξ can be defined for a certain space of operators and the most general form is exhibited. The associated $\star$-product (generalized Moyal product) is the first known example of a nonanalytic $\star$-representation. In Sec. IV we present some new results concerning $\star$-products on polynomials. In Sec. V we study piecewise analytic $\star$-representations. When such $\star$-representations are allowed (and the Kepler problem tells us that they must be), then a $\star$-representation is not determined unambiguously by the $\star$-product on polynomials. In other words, a Weyl-type quantization rule for polynomials does not lead to a unique spectrum, even for polynomial operators.[7] An invariant Wigner correspondence and an associated piecewise analytic $\star$-representation is conjectured for the pair (T,W), where T is a unitary representation of a simple Lie group G and the phase space W is an orbit of the coadjoint action of G in the real vector space dual of the Lie algebra of G.

II. ONE-DIMENSIONAL KEPLER PROBLEM

The Kepler problem in one dimension presents the same difficulties as the realistic, three-dimensional case. In addition, it is of special interest because conventional quantization based on the Heisenberg algebra is not available.

Configuration space is the real half-line, $q > 0$, and phase space (W) is the real half-plane $0 < q,\ -\infty < p < \infty$. The symplectic structure is given by $\{q,q\} = 0 = \{p,p\}$ and $\{q,p\} = 1$. Conventional Weyl quantization would associate q and p with operators in $L^2(0,\infty;dr)$; in particular, $p \to (\hbar/i)(d/dr)$. Here one can object that this operator is not self-adjoint, and this remark casts doubts on the relevance of the whole procedure; indeed, q and p are supposed to be the most basic of observables, and the axioms of quantum mechanics require that they be represented by self-adjoint operators. On reflecting, however, on the fact that p, as a function on phase space, is not the generator of a group of canonical transformations (it comes to grief with a collision at $q = 0$), one understands that p is not a good observable even on the classical level. Consequently, it is illogical to attempt invariant quantization based on the Heisenberg algebra.[8] To express this in another way: If one insists on defining quantization in terms of representations of the Heisenberg algebra, then a nonintegrable representation must be used. Notice that we are not rejecting the quantum theory that is discovered by the conventional procedure; in fact we shall adopt it. But we choose our distinguished observables among those that are represented by self-adjoint operators.

In the case of the three-dimensional Kepler problem one is led by intuitive physical arguments to the idea of attempting invariant quantization based on so(4,2).[9] The one-dimensional analog is so(2,1) or sl(2,$\mathbb{R}$). The algebra $\mathscr{A}$ is spanned by the three functions q, qp, and qp^2. The Hamiltonian function is $H = p^2 - 1/q$, and this must be regularized. The mean anomaly s is the monotoneous function of t defined up to an additive constant by $ds/dt = 1/q$. The equation of motion $df/dt = \{E - H, f\}$ takes the form

$$df/ds = \{q(E-H), f\}, \tag{2.1}$$

where $q(E-H) = qE + 1 - qp^2$ is the regularized Hamiltonian. The main point is that *this regularized Hamiltonian is an element of* $\mathscr{A}$; *an invariant quantization procedure will therefore give us a system with a distinguished Hamiltonian.*

If the operator $\hat{F}$ is symmetric in $L^2(\mathbb{R}^3, d^3r)$, then the operator $r\hat{F}$ (where r is the radial coordinate) is symmetric in $L^2(\mathbb{R}^3, d^3r/r)$. This correspondence extends, in many cases, to self-adjoint operators and is familiar in connection with group theoretical treatments of the hydrogen atom.[10] The generators of so(4,2) are self-adjoint in d^3r/r, and this corresponds precisely to the self-adjointness of the operators $(\hbar/i)(d/dr^i)$ and $-\hbar^2\nabla$ in d^3r. The situation is different in one dimension. The operators $r[(\hbar/i)(d/dr)]^n$, $n=0,1,2$, are symmetric in $L^2(0,\infty;dr/r)$ and have unique self-adjoint extensions with a common, dense, invariant domain; the operators $[(\hbar/i)(d/dr)]^n$, $n=0,1,2$ are symmetric in $L^2(0,\infty;dr)$, but one of them, $(\hbar/i)(d/dr)$, has no self-adjoint extension. It is convenient to use the metric dr/r, rather than dr, and we take $L^2(0,\infty;dr/r)$ as the Hilbert space of one-dimensional hydrogen.

There is additional simplification to be gained by transforming to a momentum space description. We use Greek letters α,β to denote momentum space coordinates; the operators corresponding to the distinguished observables then take the form $(i\hbar d/d\alpha)\alpha^n$, $n=0,1,2$.

The mapping of phase space (W) into $\mathbb{R}^3$ given by $(q,p)\to(q,qp,qp^2)$ may be interpreted as a mapping of W to an orbit of the co-adjoint action of sl(2,$\mathbb{R}$) in the real vector space dual $\mathscr{A}^*$ of sl(2,$\mathbb{R}$). The orbit in question is a singular one; namely the upper cone

$$Q\equiv(L^1)^2-(L^2)^2-(L^3)^2=0,\quad L^1>0. \tag{2.2}$$

[see Eq. (3.1) below.] The unitary representation of SL(2,$\mathbb{R}$) defined in $L^2(0,\infty;dr/r)$ by the operators $r[(\hbar/i)(d/dr)]^n$, $n=0,1,2$, belongs to the discrete series and is characterized by

$$\hat{Q}\equiv(\hat{L}^1)^2-(\hat{L}^2)^2-(\hat{L}^3)^2=0, \tag{2.3}$$

$$\text{Spectrum }(\hat{L}^1/\hbar)=2,4,\cdots. \tag{2.4}$$

The operator that corresponds to regularized Hamiltonian is

$$\tfrac{1}{2}(E-1)\hat{L}^1+\tfrac{1}{2}(E+1)\hat{L}^3+1.$$

The spectrum of this operator is the same as that of $-(-E)^{1/2}\hat{L}^1+1$ if $E<0$ and that of $(-\sqrt{E})\hat{L}^3+1$ if $E>0$. The constraint $E-H=0$ thus gives the discrete spectrum $E=-\hbar^2/4n^2$, $n=1,2,\cdots$, and the continuous spectrum $E>0$.

III. INVARIANT WIGNER CORRESPONDENCE

A Wigner correspondence maps operators in Hilbert space to distributions on phase space.[11] The term "distinguished observable" is associated with the commutator Lie algebra $\hat{\mathscr{A}}$ of operators of the form $\hat{a}=a_A\hat{L}^A$ (sum $A=1,2,3$), and with the Poisson Lie algebra $\mathscr{A}$ of functions on phase space of the form $a=a_AL^A$, with a_A in $\mathbb{R}$ and

$$\begin{aligned}\hat{L}^1&=i\hbar(\partial/\partial\alpha)(\alpha^2+1), & L^1&=q(p^2+1),\\ \hat{L}^2&=i\hbar(\partial/\partial\alpha)(2\alpha), & L^2&=2qp,\\ \hat{L}^3&=i\hbar(\partial/\partial\alpha)(\alpha^2-1), & L^3&=q(p^2-1).\end{aligned} \tag{3.1}$$

A Wigner correspondence Ξ for the one-dimensional Kepler problem is called so(2,1)-invariant, or simply *invariant*, if for every a in $\mathscr{A}$ we have $\Xi(\hat{a})=a$ and

$$\Xi([\hat{a},\hat{F}])=i\hbar\{a,F\}, \tag{3.2}$$

where $\hat{F}$ is any operator in the domain of Ξ and $F=\Xi(\hat{F})$. (Definition 3.)

The linear operator Ξ acts on a space of operators defined by their integral kernels. [The integral kernel $\hat{F}_{\alpha,\beta}$ of an operator $\hat{F}$ is given heuristically by $(\hat{F}\psi)(\alpha)=\int_{\mathbb{R}}\hat{F}_{\alpha,\beta}\psi(\beta)d\beta$.] The precise domain of Ξ will not be determined; we shall assume that, on a proper topological subspace $\mathscr{S}$ of its domain, the linear map Ξ is a function-valued distribution that we shall write as

$$\Xi(\hat{F})=\int\hat{F}_{\alpha,\beta}\,\Xi_{\beta,\alpha}\,d\alpha\,d\beta,\quad \hat{F}\text{ in }\mathscr{S}. \tag{3.3}$$

Proposition 4: If Ξ is an invariant Wigner correspondence, then

$$\Xi_{\beta,\alpha}(q,p)=\int(e^{L\cdot\hat{L}/i\hbar\rho})_{\beta,\alpha}\,\omega(\rho)\,d\rho, \tag{3.4}$$

where ω is a complex distribution on $\mathbb{R}$ and

$$L\cdot\hat{L}\equiv g_{AB}L^A\hat{L}^B=L^1\hat{L}^1-L^2\hat{L}^2-L^3\hat{L}^3. \tag{3.5}$$

Proof: It is easily established that

$$(e^{\hat{a}/i\hbar})_{\beta,\alpha}=\frac{R(\beta')}{R(\beta)}\delta(\beta'-\alpha), \tag{3.6}$$

where R is the polynomial defined by $a=qR(p)$, namely $R(\alpha)=(a_1-a_3)+2a_2\alpha+(a_2+a_3)\alpha^2$ and

$$\beta'=[\lambda_2(\beta-\lambda_1)e^{2it}-\lambda_1(\beta-\lambda_2)]/[(\beta-\lambda_1)e^{2it}-\beta+\lambda_2], \tag{3.7}$$

where $\lambda_{1,2}$ are the zeros of R and

$$t^2\equiv a\cdot a\equiv g^{AB}a_Aa_B\equiv(a_1)^2-(a_2)^2-(a_3)^2. \tag{3.8}$$

Substituting $\hat{a}\to L\cdot\hat{L}/\rho$ in (3.6), we find

$$(e^{L\cdot\hat{L}/i\hbar\rho})_{\beta,\alpha}=\frac{1}{q(\beta-p)^2}\delta\Big(\frac{1}{q(\beta-p)}-\frac{1}{q(\alpha-p)}-\frac{2}{\rho}\Big), \tag{3.9}$$

so that the right-hand side of (3.4), after multiplication by $q(\beta-p)^2$, depends only on $[q(\beta-p)]^{-1}-[q(\alpha-p)]^{-1}$. To show that this is the most general form of $\Xi_{\beta,\alpha}$, we impose (3.2); taking $a=q$, we learn that it depends on q, $\alpha-p$, $\beta-p$; taking $a=qp$, we find that $q\Xi_{\beta,\alpha}$ depends only on $q(\alpha-p)$ and $q(\beta-p)$; finally we take $a=pq^2$ and discover that $q(\beta-p)^2\Xi_{\beta,\alpha}$ depends on $[q(\beta-p)]^{-1}-[q(\alpha-p)]^{-1}$ only.

By definition 3, the elements $\hat{a}$ of $\hat{\mathscr{A}}$ are in the domain of Ξ; however, it will not be taken for granted that these operators belong to the subspace $\mathscr{S}$ on which Ξ is determined by the formula (3.3). In fact, *we know of no choice of ω such that (3.3) makes sense for any operator $\hat{F}$ in the enveloping algebra $\hat{U}$ of $\hat{\mathscr{A}}$*. Nevertheless, we insist on including $\hat{U}$ in the domain of Ξ, although $\hat{U}$ is disjoint from the space $\mathscr{S}$ on which Ξ is defined by the formula (3.3). One of our main results will be to give a definition of Ξ on $\hat{U}$ that extends (3.3). To do this, we must place some restrictions on ω, and we begin by postulating that (3.3) applies to the unitary operators $\exp(\hat{a}/i\hbar)$ of the group representation.

Substitution of $\exp(\hat{a}/i\hbar)$ for $\hat{F}$ in (3.3) leads, with the help of (3.6) and a change of integration variables, to the following expression, for g_x in SL(2,$\mathbb{R}$):

g_x in SL(2,$\mathbb{R}$):

$$\Xi(\hat{T}_x) = \frac{x\cdot L}{2}\int_{-\infty}^{+\infty} \frac{dz}{(\cosh\zeta - x_0)^2}\,\omega\left(\frac{-x\cdot L}{\cosh\zeta - x_0}\right).$$

$$\equiv E(g_x). \tag{3.10}$$

Here $\cosh\zeta = (z^2+1)/2z$, $x\cdot L = x_A L^A$ (sum $A = 1,2,3$), and (x_μ), $\mu = 0,1,2,3$, are the parameters for SL(2,$\mathbb{R}$) defined by the standard homeomorphism between this group and the surface $x_0^2 + x_1^2 - x_2^2 - x_3^2 = 1$ in $\mathbb{R}^4$. In the domain of the exponential map $x_0 = \cos t$, $x_A = (a_A/t)\sin t$, and $\hat{T}_x = \exp(\hat{a}/i\hbar)$ is the unitary operator corresponding to $g_x = e^a$.

As usual, define the $\star$-product[12]

$$E(g_y)\star E(g_x) = \Xi(\hat{T}_y\hat{T}_x) = \Xi(\hat{T}_z) = E(g_z). \tag{3.11}$$

Differentiating with respect to y at the identity, we get

$$\Xi(\hat{b}\hat{T}_x) = b\star E(g_x) = [-i\hbar\ l(b)\ E](g_x), \tag{3.12}$$

where $l(b)$ is the vector field of left translations on SL(2,$\mathbb{R}$) generated by b. Higher derivatives give $c\star b\star E(g_x)$ and so on. Now, if $g \to E(g)$ were analytic at the identity, as was the case in all cases studied previously, then these expressions could be used to compute $\Xi(\hat{c}\hat{b}) = c\star b$ and higher $\star$-polynomials. However, *we know of no choice of ω for which E is analytic at* the identity and we believe that none exists for any reasonable choice of $\mathscr{S}$. The Kepler problem allows no analytic $\star$-representation.

Nevertheless, Ξ is uniquely defined on polynomials if $E(g_x)$ can be expanded as a power series in the coordinates L^A. This is the case if ω is an entire function, $\omega = \Sigma\omega_n\rho^n$, with coefficients that decrease fast enough to allow the exchange of the order of summation and integration. For convenience we introduce complex constants c_n and functions P_n by

$$\omega(\rho) = -\frac{1}{\pi\hbar}\sum_{n=1}^{\infty}\frac{(2n+1)!!}{(n-1)!(n+1)!}\,c_n\left(\frac{\rho}{i\hbar}\right)^{n-1}, \tag{3.13}$$

$$P_n(a) = \frac{(2n-1)!!}{n!}\,a^n, \quad n = 0,1,\cdots, \quad a \text{ in } \mathscr{A}. \tag{3.14}$$

Evaluation of the integral then gives for $0 < t < \pi$,

$$E(e^a) = \sum_{n=1}^{\infty} C_n(t)P_n(a), \tag{3.15}$$

$$C_n(t) = c_n\frac{i(2n+1)}{(n+1)!\,\sin t}(\hbar t)^{-n}e^{-i\pi}P_n^1(1/i\tan t). \tag{3.16}$$

According to (3.12), this function is annihilated by $b\star + i\hbar\, l(b)$. Applying $a\star + i\hbar\, l(a)$ to (3.15), we use the recursion relation

$$(2n+1)\left[(1-\mu^2)\frac{d}{d\mu} - \mu\right]P_n^\lambda(\mu)$$
$$= n(n+\lambda)P_{n-1}^\lambda(\mu) + (n+1)(\lambda - n - 1)P_{n+1}^\lambda(\mu) \tag{3.17}$$

and annul the coefficient of P_n^1, $n = 1,2,\cdots$. (Note that P_0^1 vanishes identically.) The result is

$$(2n+1)a\star c_nP_n(a) = (n+1)c_{n+1}P_{n+1}(a)$$
$$- n(1-n^2)(\hbar t)^2 c_{n-1}P_{n-1}(a) \tag{3.18}$$

for $n = 1,2,\cdots$. We shall normalize Ξ by taking

$$c_1 = 1, \quad \text{so that } c_1P_1(a) = a. \tag{3.19}$$

Now, solving (3.18) successively for $c_2P_2(a)$, $c_3P_3(a)$,$\cdots$, one obtains expressions for $c_nP_n(a)$ in the form of $\star$-polynomials in a:

$$c_nP_n(a) = \sum_{k=0}^{[n/2]} A_n^k(\hbar t)^{2k}(a\star)^{n-2k}. \tag{3.20}$$

The coefficients A_n^k are independent of the choice of the c_n; the $\star$-polynomials (3.20) satisfy (3.18) identically. In other words, (3.18) may be regarded as a recursion relation for a family of $\star$-polynomials, while (3.20) expresses each $\star$-polynomial as the ordinary function $c_nP_n(a)$. Since $A_n^0 = (2n-1)!!/n! \neq 0$, (3.20) determines the function $(a\star)^n$ for $n = 2,3,\cdots$, and we have proved:

Proposition 5: If an invariant Wigner correspondence Ξ is determined by (3.4), or equivalently by (3.10), and if ω is an entire function, then Ξ has a unique extension to the enveloping algebra $\hat{U}$ of $\hat{\mathscr{A}}$, given by $\Xi(\hat{a}\hat{b}\cdots) = a\star b\star\cdots$, the $\star$-product being defined for polynomials by (3.18). The $\star$-powers $(a\star)^n$ are polynomials of order n, with leading term c_na^n. The coefficients c_n are defined by (3.13) and must all be nonzero.[13]

Remark 6: The $\star$-product is defined for any two elements $\hat{F}$, $\hat{G}$ of $\hat{U}$ by $\Xi(\hat{F})\star\Xi(\hat{G}) = \Xi(\hat{F}\hat{G})$ and is evidently associative. The invariance requirement (Definition 3) tells us that, for a in $\mathscr{A}$,

$$[a\star F] = i\hbar\{a,F\} \tag{3.21}$$

so the $\star$-product is *invariant*.[14] This is why the $\star$-product on polynomials is determined by the $\star$-powers $(a\star)^n$ and hence by the $\star$-polynomials $c_nP_n(a)$.

Later we shall obtain a closed expression for the $c_nP_n(a)$ as $\star$-polynomials (Proposition 8), though not in the simple form (3.20). Direct solution of (3.18) leads to

$$c_na^n = (a\star)^n - \frac{(n+1)n(n-1)(n-2)}{6(2n-1)}(\hbar t)^2(a\star)^{n-2}$$
$$+ \frac{(n+1)\cdots(n-4)}{3\cdot 5!(2n-1)(2n-3)}$$
$$\times(5n^2 - 7n + 6)(\hbar t)^4(a\star)^{n-4} + \cdots$$
$$+ \frac{n}{(2n-1)!!}\begin{cases}(i\hbar t)^{n-1}a, & n \text{ odd}\\ (i\hbar t)^{n-2}\phi(n)a\star a, & n \text{ even}\end{cases}, \tag{3.22}$$

with $\phi(n) = \Sigma_1^n(1/k)$ and, if $c_n = 1$,

$$(a\star)^n = a^n + \frac{(n+1)n(n-1)(n-2)}{6(2n-1)}(\hbar t)^2a^{n-2}$$
$$+ \frac{(n+1)\cdots(n-4)}{3\cdot 5!(2n-3)(2n-5)}$$
$$\times(5n^2 - 23n + 30)(\hbar t)^4(a\star)^{n-4} + \cdots. \tag{3.23}$$

The numerical values of the c_n are unimportant; we favor the choice $c_n = 1$, for in this case, as (3.23) shows, the $\star$-product on polynomials is a deformation of the ordinary product, with deformation parameter $\hbar$.[15] If a is nilpotent, for example $a = q$, then $t = 0$ and $(a\star)^n = a^n$. Thus, if $c_n = 1$, and if $f(q)$, $g(q)$ are polynomials,[16]

$$f(q)\star g(q) = f(q)g(q). \tag{3.24}$$

This choice makes (3.15) convergent for $\sin t \neq 0$ and

$$\omega(\rho) = -(3/2\pi\hbar)\,{}_1F_1(5/2, 3, 2\rho/i\hbar) \tag{3.25}$$

an entire function. The right-hand side of (3.4) is now the integral kernel $\hat{K}_{\beta,\alpha}$ of the operator defined weakly by

$$\hat{K} = \int e^{L\cdot\hat{L}/i\hbar\rho}\omega(\rho)d\rho,$$

and (3.3) may be written

$$\Xi(\hat{F}) = \operatorname{tr}(\hat{F}\hat{K}),$$

provided $\hat{F}\hat{K}$ is trace-class.

IV. INVARIANT $\star$-PRODUCTS ON POLYNOMIALS

In this section, let (L^A), $A = 1,2,3$, denote, either (as earlier) a basis for $\mathscr{A} = \mathrm{sl}(2,\mathbb{R})$, or coordinates for its real vector space dual $\mathscr{A}^\star$. We recall[17] the construction of invariant $\star$-products on an orbit of the co-adjoint action, given by fixing the invariant

$$Q \equiv g_{AB}L^AL^B = (L^1)^2 - (L^2)^2 - (L^3)^2 \tag{4.1}$$

in $\mathbb{R}$. For the Kepler problem one has $Q = 0$. Let the solid spherical function $P_n(a)$ and the symmetric tensor T^n be defined by

$$P_n(a) = \sum_k (-)^k \frac{(2n-2k-1)!!}{(n-2k)!(2k)!!}(t^2Q)^k a^{n-2k}$$

$$= T^n_{A_1\cdots A_n}L^{A_1}\cdots L^{A_n} \tag{4.2}$$

—this agrees with (3.14) when $Q = 0$—and define

$$P_n(a\star) \equiv T^n_{A_1\cdots A_n}L^{A_1}\star L^{A_2}\star\cdots\star L^{A_n} \tag{4.3}$$

The function $\hat{Q} \equiv g_{AB}L^A\star L^B$ is an invariant of the co-adjoint action, and reduces to a constant on each orbit; we define λ and l by

$$\hat{Q} = \hbar^2(\lambda^2 - 1) = 4\hbar^2 l(l+1), \quad \lambda = 2l + 1, \tag{4.4}$$

Proposition 7: Any invariant $\star$-product on the space of polynomials in the linear coordinates (L^A), $A = 1,2,3$, of $\mathscr{A}^\star$, restricted to any orbit, is given in terms of complex parameters λ^2 and (d_n), $n = 2,3,\cdots$ by (4.4) and

$$P_n(a\star) = d_nP_n(a), \quad d_0 = d_1 = 1. \tag{4.5}$$

Proof: See Ref. 12, Example 41.

When $\lambda = 1$, then $P_n(a\star)$, defined in (4.3), coincides with the $\star$-polynomials (3.20), defined by (3.18) and (3.19). This will be shown below, Proposition 8.

The $\star$-exponential[18]: If an invariant $\star$-product on some orbit W is given, then we may consider the series

$$\mathrm{Exp}(a) = \sum_n \frac{1}{n!}(i\hbar)^{-n}(a\star)^n \tag{4.6}$$

Suppose, for the moment, that this series converges, for a in some neighborhood of the origin of $\mathscr{A}$, to a C^∞ function on $\mathscr{A}^\star$. Then, for a, b near the origin,

$$\mathrm{Exp}(a)\star\mathrm{Exp}(b) = \mathrm{Exp}(c), \tag{4.7}$$

where $c = c(a,b)$ is given by the Campbell–Hausdorff formula. Differentiation at the origin gives

$$[a\star + i\hbar\, l(a)]\mathrm{Exp}(b) = 0, \tag{4.8}$$

and from this it follows that $\mathrm{Exp}(b)$ is an invariant eigendistribution on the local group.[18] The general form is[19] (if we require regularity at $a_A = 0$)

$$\sum_n C_n(t)P_n(a),$$

with

$$C_n(t) = c_n\frac{i(2n+1)}{(n+\lambda)!\sin t}(\hbar t)^{-n}e^{-i\pi\lambda}Q_n^\lambda(1/i\tan t), \tag{4.9}$$

where Q_n^λ is an associated Legendre function, $c_0 = c_1 = 1$, and (c_n), $n = 2,3,\cdots$, are complex numbers. Applying $a\star + i\hbar \times l(a)$ to (4.9), making use of the recursion relations (3.17) that hold for Q_n^λ as they do for P_n^λ, and annulling the coefficient of Q_n^λ, we obtain

$$(2n+1)a\star c_nP_n(a) = (n+1)c_{n+1}P_{n+1}(a) - n(\lambda^2 - n^2)(\hbar t)^2c_{n-1}P_{n-1}(a) \tag{4.10}$$

for $n = 0,1,\cdots$.

If we compare terms of the same order in (a_A) in (4.6) and (4.9), we learn that $(a\star)^n = c_na^n$ + terms of lower order. The same conclusion is reached by solving (4.10), considering this equation as a recursion relation for a family of $\star$-polynomials. On the other hand we know, since the $\star$-product is invariant, that there are complex parameters d_n such that (4.5) holds, and this gives $(a\star)^n = d_na^n$ + lower order terms. Therefore, $c_n = d_n$, and we have the following:

Proposition 8: The $\star$-polynomials $P_n(a\star)$ defined by (4.3) are the unique solutions of the recursion relations

$$(2n+1)a\star P_n(a\star) = (n+1)P_{n+1}(a\star) - n(\lambda^2 - n^2)(\hbar t)^2P_{n-1}(a\star), \tag{4.11}$$

with the boundary condition $P_0(a\star) = 1$.

Proof: The above analysis assumes only the convergence of (4.9) near $a_A = 0$, which puts a limitation on the asymptotic behavior of c_n for large n. Otherwise, the c_n are arbitrary for $n = 2,3,\cdots$. The result (4.11) follows from (4.5), (4.10) and the equality $c_n = d_n$, and is completely independent of the values of $c_2, c_3,\cdots$; hence (4.11) is an identity for $n = 0,1,\cdots$.

Corollary: If the invariant $\star$-product is defined by $P_n(a\star) = c_nP_n(a)$, then the functions $c_nP_n(a)$ satisfy (4.10).

Now suppose that $c_N = 0$ for some positive integer N.

Then (4.10) gives $(N+1)c_{N+1}P_{N+1}(a) = (\hbar t)^2 N(\lambda^2 - N^2)c_{N-1}P_{N-1}(a)$, and thus we have the following:

Proposition 9: If an invariant $\star$-product on polynomials is defined by $P_n(a\star) = c_n P_n(a)$, $n = 0,1,\cdots$, then *either* all the c_n are nonzero, *or* λ is a positive integer and $c_n = 0$ for $n \geqslant \lambda$ and $c_n \neq 0$ for $n < \lambda$.

The integral case: When $\lambda = 2l + 1$ is a positive integer, there are two possibilities. (1) If $c_n = 0$ for $n \geqslant \lambda$ and $c_n \neq 0$ for $n < \lambda$, then every $\star$-polynomial is an ordinary polynomial of order $\leqslant 2l$ and the algebra of $\star$-polynomials is a linear space of finite dimension $(2l+1)^2$. This case is related to the finite-dimensional representations of sl(2,$\mathbb{R}$) and of su(2).[20] (2) If $c_n \neq 0$, $n \geqslant \lambda$, then the polynomials $\{P_n(a\star)\}$; $n \geqslant \lambda$, a in $\mathscr{A}$, span a linear space of polynomials that is invariant for the operators $b\star$ and $\star b$; this space is thus closed under $\star$-multiplication and forms a $\star$-product algebra. In this case one may ask whether the extension of this $\star$-product algebra to the lower $\star$-polynomials $P_n(a\star)$ with $n < \lambda$, which is always possible, is mandatory. It will be shown below that quantization can be carried out without such an extension, in one and only one way.

V. REPRESENTATIONS AND $\star$-REPRESENTATIONS

Let T be a unitary representation of G = the universal covering group of SL(2,$\mathbb{R}$), in some Hilbert space, and let Ξ be an invariant Wigner correspondence with values in $C^\infty(W)$. Assume that the domain of Ξ is invariant under the group action, then $\mathscr{E} = \Xi \circ T$ defines a $\star$-representation, which has been defined[21] as a distribution on G, with values in $C^\infty(W)$, such that (1) the domain and the kernel of $\mathscr{E}$ are closed under the convolution in the test function space and (2) $\mathscr{E}$ is invariant in the sense that, for a in sl(2,$\mathbb{R}$), $\mathscr{E} \circ ad(a) = ad^\star(a) \circ \mathscr{E}$. A $\star$-representation is said to be analytic if it coincides, in a neighborhood U of the identify of G, with $E(g^{-1}) d_r g$, where $d_r g$ is the right-invariant Haar measure and E is analytic in U.

If $\mathscr{E}$ is an analytic $\star$-representation, then there is a unique invariant $\star$-product on polynomials that is compatible with it. It is defined by expanding $E(e^a)$ in a power series in (a_A) and identifying with the series Exp(a), Eq. (4.6). Conversely, if an invariant $\star$-product on polynomials is given, then it is compatible with at most one analytic $\star$-representation.

A $\star$-representation $\mathscr{E}$ will be said to be associated with a unitary representation T if there exists an invariant Wigner correspondence such that $\mathscr{E} = \Xi \circ T$. If $\mathscr{E}$ is defined by integrating against $E(g^{-1}) d_r g$, then $E(g) = \Xi(T_g)$. In this paper we have selected a particular representation of SL(2,$\mathbb{R}$) and a particular (singular) orbit. We have found a general expression for Ξ that appears to rule out the existence of an analytic $\star$-representation in this case, though we have been unable to formulate a proof. We have found a type of Wigner correspondence, and hence a type of $\star$-representation, that, though not analytic, leads to a unique invariant $\star$-product on polynomials. We shall refer to this type as piecewise analytic $\star$-representations.

Definition 10: A $\star$-representation is said to be piecewise analytic if it coincides with $E(g^{-1}) d_r g$, where E is analytic almost everywhere.

As in the analytic case, we can determine the general form of $E(e^a)$ in the piecewise analytic case. The result is similar to (4.9), with some important differences. The function Q_n^λ is replaced by a linear combination $\mu Q_n^\lambda + \nu P_n^\lambda$, with μ, ν independent of n. The function $P_n^\lambda(1/i \tan t)$ is singular at $t = 0$; therefore, the representation (4.6) is not valid unless $\nu = 0$. Let us express this another way: the series (4.6) may possibly converge, in that case it defines Exp(a), but Exp$(a) \neq E(e^a)$ unless $\nu = 0$. Just as in the analytic case, a $\star$-product on polynomials exists only if the c_n satisfy the conditions of Proposition 9, and is then determined unambiguously. Therefore, every *invariant $\star$-product on polynomials is compatible with a family of $\star$-representations*, parametrized by ν/μ, precisely one of which (except in the integral case) is analytic.

The integral case: When $\lambda = 2l + 1$ is a positive integer, then $P_n^\lambda = 0$ for $n < \lambda$. The exception just referred to occurs when $\mu = 0$, $\nu = 1$, so that

$$E(e^a) = \sum_{n=\lambda}^{\infty} \frac{i(2n+1)c_n}{(n+\lambda)!\sin t} (\hbar t)^{-n} e^{-i\pi\lambda} P_n^\lambda\left(\frac{1}{i \tan t}\right) P_n(a). \quad (5.1)$$

Only the functions and $\star$-polynomials $P_n(a\star) = c_n P_n(a)$ for $n \geqslant \lambda$ appear in this case, so the lower $\star$-polynomials are irrelevant, as was hinted at near the end of the preceding section. Conversely, if we take the $\star$-product algebra to be spanned by the $P_n(a\star)$, $n \geqslant \lambda$, then we must take $\mu = 0$, and a unique $\star$-representation results. The $\star$-representation for the Kepler problem is recovered by taking $\lambda = 1$.

VI. CONCLUSIONS AND OUTLOOK

Distinguished Hamiltonians: It was shown that the one-dimensional Kepler problem is one of the very special physical systems that have distinguished Hamiltonians, systems for which phase space trajectories remain meaningful after quantization. In other words, if the phase space distribution function is supported on one point in phase space at $t = t_0$, then the same is true at all times. We expect to be able to show that the three-dimensional Kepler problem has the same property; a hint of what to expect will be given below. Next, it will be extremely interesting to investigate whether certain field theories may have distinguished Hamiltonians (an example: quantized fermion fields in interaction with classical boson fields), expecially in view of the importance that has recently been placed on exact solutions of nonlinear classical field equations. It is even possible to imagine that the future correct quantization of general relativity may be found by requiring this system to have meaningful trajectories, for in this way it might be possible to circumvent the difficulties associated with quantum fluctuations of the light cone.

$\star$*-representations*: The Kepler problem has led us to study nonanalytic $\star$-representations. So far it has not been proved that analytic $\star$-representations for the Kepler problem are impossible, but our many fruitless efforts to construct one is not our only reason for believing that none exist. It may turn out to be easier to prove the existence of systems that have no analytic $\star$-representation, instead of showing that the Kepler problem is such a system. Indeed, let us consider the case $\lambda = 1, \hat{Q} = 0$. There are two inequivalent, unitary, irreducible representations of SL(2,$\mathbb{R}$) with $\hat{Q} = 0$. One of them, T_+ say, is associated with the Kepler problem and is characterized by the positive spectrum for $\hat{L}^1$ given by Eq. (2.4); the other one, T_-, has a purely negative spectrum. Let us suppose that $\mathscr{E}_+ = \Xi_+ \circ T_+$ is analytic; hence given by (4.9) with $\lambda = 1$ and any choice of c_n's, subject only to the condition $c_n \neq 0$ and the requirement of convergence of (4.9). Now, if $\mathscr{E}_- = \Xi_- \circ T_-$ is also analytic, then we may choose the c_n's equal in $\mathscr{E}_\pm$, and reach the conclusion that $\mathscr{E}_\pm$ coincide although $T_\pm$ are inequivalent. It remains only to show that this apparently absurd conclusion constitutes a real contradiction, in order to complete the proof that $T_\pm$ cannot both have analytic $\star$-representations. A satisfactory treatment of these problems must perhaps wait for the development of the concept of equivalence classes of $\star$-representations.

The nonanalytic $\star$-representation that was constructed for the Kepler problem turned out to be piecewise analytic (Definition 10). The concept of piecewise analytic $\star$-representations may be wide enough to cover the general case; more precisely we venture the following:

Conjecture: Let G be a simple Lie group, $\mathscr{A}$ its Lie algebra, $\mathscr{A}^*$ the real vector space dual of $\mathscr{A}$; and W an orbit of the co-adjoint action of $\mathscr{A}$ in $\mathscr{A}^*$. Let T be any unitary representation of G and suppose that W is an orbit of maximal dimension.[22] Then there exists an invariant Wigner correspondence Ξ and a piecewise analytic $\star$-representation $\mathscr{E} = \Xi \circ T$.

There is some support for this conjecture. Let ϕ be the map from $\mathscr{A}^*$ to G defined by $\phi = \exp \circ K$, where K is the natural map from $\mathscr{A}^*$ to $\mathscr{A}$ defined by the Killing form and exp is the exponential map. Let χ_T denote the character of T, and consider

$$E_\xi(g) \equiv \int_{\mathbb{R}} \chi_T(g \cdot \phi(\xi/\rho))\omega(\rho)\, d\rho, \quad \xi \in W, \; g \in G,$$

where ω is an everywhere analytic function on $\mathbb{R}$. It is known that χ_T is locally integrable. If g is a regular element of G, then χ_T coincides, in a neighborhood of g, with an analytic function. The integrand, for regular g, is therefore piecewise analytic in ρ and analytic outside a compact interval. As $|\rho| \to \infty$ we have $\chi_T(g \cdot \phi(\xi/\rho)) \to \chi_T(g)$, which exists for g regular. Thus, mild growth conditions on ω are sufficient to guarantee the existence of the integral, for g regular. (It seems to be more difficult to prove that the integral defines a C^ω function on $\mathscr{A}^*$, for g regular.) Furthermore, E is invariant in the sense that

$$\begin{aligned} \mathrm{Ad}^*(g)E(g')\big|_\xi &= \int \chi_T(g' \cdot \mathrm{Ad}(g)\phi(\xi/\rho))\omega(\rho)\, d\rho \\ &= \int \chi_T(g' \cdot g \cdot \phi(\xi/\rho) \cdot g^{-1})\omega(\rho)\, d\rho \\ &= \mathrm{Ad}(g)E_\xi(g'). \end{aligned}$$

To prove the conjecture would be to establish the viability of the method of $\star$-quantization. Without assuming the truth of the conjecture it is in any case possible to use the above formula to construct many interesting $\star$-representations. This formula is valid in the case of the one-dimensional Kepler problem and will be tested in the three-dimensional case.

ACKNOWLEDGMENTS

It is a pleasure to thank Moshe Flato for the many ideas that he has contributed to this work. I also thank Daniel Sternheimer for his careful checking of the manuscript. This work was supported in part by the National Science Foundation.

[1] J.E. Moyal, Proc. Cambr. Phil. Soc. **45**, 99 (1949).
[2] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowiz, and D. Sternheimer, Ann. Phys. **111**, 111 (1977). See Sec. 1.b.
[3] The concepts are the same as in Ref. 2, but they are approached from another direction. In Ref. 2 we dealt with a purely classical system and attempted to identify the algebra of distinguished observables by intuitive physical arguments. We then proposed a quantization procedure that would in effect make this the algebra of distinguished observables in the sense of Definition 3.
[4] Ref. 2, Sec. 1.c.
[5] J.M. Souriau, *Structure des systemés dynamiques* (Dunod, Paris, 1970).
[6] B. Kostant, *Lectures in Modern Analysis and Applications III* Lecture Notes in Mathematics **170** (Springer-Verlag, Berlin, 1970), pp. 87–208.
[7] It was shown in Ref. 2 that, in the case of an analytic $\star$-representation, the spectra of quantum operators could be obtained from the $\star$-product on polynomials.
[8] C. Fronsdal, Rep. Math. Phys. (to be published). Here additional examples of physical systems that call for noncanonical quantization were given, including the case of compact phase space. See Sec. 4.
[9] Ref. 2, Sec. 1.c.
[11] E.P. Wigner, Phys. Rev. **40**, 749 (1932). The Wigner correspondence is the inverse of the correspondence invented by H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950), see p. 274.
[12] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Ann. Phys. **111**, 61 (1977). See Sec. 12.
[13] The last statement is proved below, Proposition 9.
[14] It follows from (3.21) that $\{a, F \star G\} = \{a, F\} \star G + F \star \{a, G\}$. The term "invariant $\star$-product" is used here with exactly the same meaning as in Refs. 2, 8, and 12.
[15] Compare Ref. 12, Sec. 3.
[16] The $\star$-product that was used in Ref. 2 to calculate the spectrum of the hydrogen atom Hamiltonian does not satisfy Eq. (3.24).
[17] Ref. 12, Example 41.
[18] Ref. 12, Sec. 12.
[19] Ref. 12, Example 50.
[20] For a general treatment of the compact case, see Ref. 8, Sec. 19.
[21] Ref. 8, Introduction.
[22] For orbits of smaller dimension there will be restrictions on T.

Explicit normalization of bound-state wave functions and the calculation of decay widths

R. S. Kaushal[a] and H. J. W. Müller-Kirsten
Department of Physics, University of Kaiserslautern, 6750 Kaiserslautern, Federal Republic of Germany

(Received 2 February 1979; accepted for publication 11 May 1979)

We present a general method for normalizing nontrivial bound-state wave functions. In particular, we consider explicit and formally complete asymptotic expansions of the solutions of the wave equation for an arbitrary quark confining power potential. The regular solution is identified and shown to be continuable to an eigensolution around a local minimum of the associated potential. The solutions are then normalized and used to derive explicit series expressions for certain decay rates.

1. INTRODUCTION

The quark model has been so successful as a means of interpreting the phenomena of particle physics that it is now a generally accepted belief that hadrons are composite structures of some basic quarklike constituents. An important aspect of the identification of a newly found hadron is the determination of its decay widths. The calculation of these quantities is therefore of some interest. It is well known that the most straightforward calculation of the leptonic and hadronic widths of a hadron requires a knowledge of the value of the wave function of the hadron at the origin. The determination of these widths thus reduces to a determination of this quantity.

In spite of considerable interest in decay widths, few attempts have been made in order to determine the analytic form of the wave function at the origin, except for the trivial cases of the Coulomb and harmonic oscillator potentials. The reason for this lack of activity is fairly obvious. The evaluation of the S-wave wave function at the origin requires the normalization of the latter and this is, in general, not a simple problem, since in most cases not even the necessary solutions are known. The following is an attempt to deal with this problem in a very general and straightforward way using as solutions the expansions derived in an earlier paper.[1] Throughout, our main emphasis is on the method rather than on specific applications.

Our procedure is the following: We assume that the quark–antiquark pair of a meson is bound together by a confinement potential $r^{\lambda}(\lambda \geqslant 1)$ which at short distances has to be corrected by the addition of a weak Coulomb component corresponding to exchange of massless gluons. In a previous paper[1] it was shown that the wave equation for such a potential possesses two pairs of solutions which are valid in complementary domains. In Sec. 2 we recapitulate these solutions and in the Appendix we rederive the WKB-like pair since this was given previously[1] only for a vanishing Coulomb component of the potential. In Sec. 3 we demonstrate that various branches of our solutions can be continued into one another in regions of common validity. In Sec. 4 we identify the regular solution. In Sec. 5 we normalize the bound-state solution and derive expressions for some decay rates. Finally, in Sec. 6 we summarize our conclusions.

2. RECAPITULATION OF BASIC RESULTS

We begin by recapitulating some results of previous work[1] which are indispensible for an understanding of the present investigation.

We are interested in the bound-state eigensolutions of the Schrödinger equation for an unscreened power potential gr^{λ} which is modified by the addition of a weak Coulomb component (corresponding to exchange of massless gluons), i.e.,

$$V(r) = g_1 r^{\lambda} - g_0/r + V_0\,, \tag{1}$$

where $\lambda \geqslant 1$ for confinement of the constituent quarks, $g_1 > 0$, and V_0 is a constant. Separating off the motion of the center of mass, we obtain the radial equation for the relative motion of the two particles of masses m_1, m_2, i.e.,

$$\frac{d^2\psi}{dr^2} + \frac{2\mu}{\hbar^2}\left[E - \frac{l(l+1)\hbar^2}{2\mu r^2} - V(r)\right]\psi = 0, \tag{2}$$

where, as usual,

$$\Psi = (1/r)\,\psi(r)\,P_l^m(\cos\theta)\,e^{im\varphi}\,,$$

$\mu = m_1 m_2/(m_1 + m_2)$ is the reduced mass of the two particles, and r is their separation. Inserting the potential (1) into (2) and setting

$$r = e^z(-\infty < z < \infty), \qquad \psi = e^{z/2}\phi(z), \tag{3}$$

we obtain our basic equation

$$(d^2\phi/dz^2) + (-L^2 + \widetilde{V}(z))\phi = 0, \tag{4}$$

where

$$\widetilde{V}(z) = \alpha\, e^{2z} - \beta\, e^{(2+\lambda)z} + \delta e^z \tag{5}$$

is the associated potential and

$$\begin{aligned} &\alpha = 2\mu(E - V_0)/\hbar^2, \quad \beta = 2\mu g_1/\hbar^2, \\ &\delta = 2\mu g_0/\hbar^2, \quad \gamma = l(l+1) \equiv L^2 - \tfrac{1}{4}. \end{aligned} \tag{6}$$

In Ref. 1 it was found convenient to introduce the following quantities:

[a] A.v. Humboldt foundation fellow on leave from Ramjas College, University of Delhi, Delhi, India.

$$\rho = \left[\frac{2\alpha}{(2+\lambda)\beta}\right]^{1/\lambda}, \tag{7}$$

$$z_0 = \ln\rho + \ln\left[1 + \frac{\delta}{2\alpha\rho\lambda} - \frac{(\lambda+1)\delta^2}{2(2\alpha\rho\lambda)^2} + \frac{(\lambda+1)(\lambda+2)\delta^3}{3(2\alpha\rho\lambda)^3} - \cdots\right], \tag{8}$$

and

$$h^4 = -2\widetilde{V}^{(2)}(z_0) = 4\alpha\lambda\rho^2 \times\left[1 + (\lambda+3)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + \left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\right],$$

$$h^2 = 2(\alpha\lambda)^{1/2}\rho\left[1 + (\lambda+3)\left(\frac{\delta}{4\alpha\rho\lambda}\right) - \frac{(\lambda+1)(\lambda+5)}{2}\left(\frac{\delta}{4\alpha\rho\lambda}\right)^2 + \cdots\right], \tag{9}$$

z_0 being defined by $\widetilde{V}'(z_0) = 0$. Then it was shown[1] that Eq. (2) or (4) possesses two similarly constructed pairs of discrete eigensolutions which are asymptotic expansions in powers of h.

The first pair of solutions involves parabolic cylinder functions $D_{(1/2)(q-1)}(\omega)$, where q is an odd integer, and the variable ω is defined as

$$\omega = h(z - z_0).$$

One of these solutions is

$$\phi{:}\phi_1(\omega,q,h^2) = N_1\Bigg[\phi_q(\omega) + \sum_{i=3}^{\infty}\frac{1}{h^{i-2}}\sum_{\substack{j=-2i\\ j\neq 0}}^{2i}\frac{[q,q+j]_i}{j}\phi_{q+j}(\omega) + \sum_{i=3}^{\infty}\frac{1}{h^{i-2}}\sum_{\substack{j=-2i\\ j\neq 0}}^{2i}\frac{[q,q+j]_i}{j} \times\sum_{i'=3}^{\infty}\frac{1}{h^{i'-2}}\sum_{\substack{j'=-2i'\\ j+j'\neq 0}}^{2i'}\frac{[q+j,q+j+j']_{i'}}{j+j'}\phi_{q+j+j'} + \cdots\Bigg], \tag{10}$$

where $\phi_q = D_{(q-1)/2}(\omega)$, N_1 is the appropriate overall normalization constant, and the coefficients $[q,q+j]_i$ are given in Ref. 1. The corresponding linearly independent solution is obtained by changing throughout the signs of ω and h [this change of signs leaves the differential equation invariant, i.e., Eq. (17) of Ref. 1, as well as the eigenvalue (as can be verified by reference to the explicit expansion given in Ref. 1)]. A further pair of solutions having the same eigenvalue follows in a similar way by the interchanges

$$\omega\to i\omega, \quad q\to -q, \quad h\to ih$$

and

$$\omega\to -i\omega, \quad q\to -q, \quad h\to -ih. \tag{11}$$

The connection between these four types of solutions is similar to (though, of course, more complicated than) the connection between the corresponding parabolic cylinder functions, i.e.,

$$D_{(1/2)(q-1)}(\omega) = \frac{[\frac{1}{2}(q-1)]!}{(2\pi)^{1/2}}\left[e^{i\pi(q-1)/4}D_{-(1/2)(q+1)}(i\omega) + e^{-i\pi(q-1)/4}D_{-(1/2)(q+1)}(-i\omega)\right].$$

The region of validity of the solutions is around $z = z_0$, i.e.,

$$z - z_0 = O\left(\frac{1}{h^\alpha}\right),$$

where $\alpha > 0$. In Sec. 5 we explain a more precise specification of this domain.

The second type of discrete eigensolutions obtained in Ref. 1 is WKB-like, and is expressed in terms of the variable

$$y = \omega/h. \tag{12}$$

One of these solution is

$$\phi{:}\phi_2(y,q,h^2) = N_2\chi(y)\exp\left\{\frac{h^2}{2\lambda^{1/2}}\int^y [v(y)]^{1/2}dy\right\}, \tag{13}$$

where N_2 is the appropriate overall normalization constant, $v(y)$ is

$$v(y) = \frac{2\lambda}{\widetilde{V}^{(2)}(z_0)}[\widetilde{V}(y+z_0) - \widetilde{V}(z_0)] \tag{14}$$

and

$$\chi(y) = \chi_q(y) + \frac{2}{h^2}\sum_{\substack{j=2,1,\cdots\\ j\neq 0}}^{-\infty}\frac{(q,q+2j)}{2j}\chi_{q+2j}(y) + \left(\frac{2}{h^2}\right)^2\sum_{\substack{j=2,1,\cdots\\ j\neq 0}}^{-\infty}\frac{(q,q+2j)}{2j} \times\sum_{\substack{j'=2,1,\cdots\\ j+j'\neq 0}}^{-\infty}\frac{(q+2j,q+2j+2j')}{2j+2j'}\chi_{q+2j+2j'}(y) + \cdots. \tag{15}$$

Here,

$$\chi_q(y) = \frac{1}{v^{1/4}} \exp\left\{ -\frac{q\lambda^{1/2}}{2} \int \frac{dy}{[v(y)]^{1/2}} \right\} \tag{16}$$

(with no further arbitrary constant), and the coefficients $(q,q+2j)$ are given in the Appendix. The expression for the discrete eigenvalue which is obtained together with the solution (15) is the same as the expression obtained in conjunction with the solution ϕ_1. The corresponding linearly independent solution $\tilde{\phi}_2$ (see the Appendix) is obtained by replacing $v^{1/2}$ by $-v^{1/2}$ or by making the substitutions $q \to -q$ and $h \to \pm ih$, leaving y unchanged [the relation (12) then requires correspondingly the substitution $\omega \to \pm i\omega$ in conformity with Eq. (11)]. The region of validity of this second pair of solutions ϕ_2, $\tilde{\phi}_2$ is the domain of large values of y, i.e.,

$$\exp\left[\mp \lambda^{1/2} \int^y \frac{dy}{[v(y)]^{1/2}} \right] < \tfrac{1}{2}h^2. \tag{17}$$

This domain excludes the region around $z = z_0$, where the first pair is valid.

An additional complication is provided by the Stokes singularities of the factor $v^{-1/4}$ occurring in Eq. (16). It is well known from WKB theory[2] that these singularities necessitate the derivation of connection formulas if the solution is required over the entire domain of the variable. It will be seen, however, that this problem can be circumvented in the present context as long as we are concerned only with asymptotic expansions which ignore exponentially decreasing contributions.

3. LINKAGE OF THE SOLUTIONS IN REGIONS OF COMMON VALIDITY

We now assume that there is a region in which the two pairs of solutions obtained above have a domain of common validity. Our objective here is to determine the constant of proportionality of the two types of solutions in their common range of validity. In fact, we demonstrate that ϕ_1 and $\tilde{\phi}_2$ are both essentially proportional to $y^{(1/2)(q-1)} \exp(-\frac{1}{4}h^2y^2)$ and therefore represent the same solution in the appropriate domain apart from a multiplying factor which involves q and h.

We consider first the solution ϕ_1. The large-ω asymptotic expansion of $\phi_q(\omega)$ is known to be[3]

$$\phi_q(\omega) = e^{-(1/4)\omega^2}\omega^{(1/2)(q-1)}\left[1 - \frac{(q-1)(q-3)}{8\omega^2} + \frac{(q-1)(q-3)(q-5)(q-7)}{2!(8\omega^2)^2} - \cdots\right] \tag{18}$$

for q a positive odd integer. Substituting this expansion in the solution (10), i.e.,

$$\phi_1 = N_1\left[\phi_q + \frac{1}{h}\left\{ \frac{[q,q+6]_3}{6}\phi_{q+6} + \frac{[q,q+2]_3}{2}\phi_{q+2} + \frac{[q,q-2]_3}{-2}\phi_{q-2} + \frac{[q,q-6]_3}{-6}\phi_{q-6}\right\} + \cdots\right] \tag{19}$$

(from the definition of the coefficients given in Ref. 1, it can be seen that $[q,q\pm 4]_3$ and other coefficients vanish) the coefficient of the dominant factor $e^{-(1/4)\omega^2}\omega^{(1/2)(q-1)}$ in the result is seen to be

$$\begin{aligned}
A = N_1\Bigg[1 &+ \frac{1}{h^2}\Bigg\{\frac{[q,q+8]_4}{8}\cdot\frac{(q+7)(q+5)(q+3)(q+1)}{2^7} - \frac{[q,q+4]_4}{4}\cdot\frac{(q+3)(q+1)}{2^3} \\
&- \frac{[q,q+6]_3}{6}\cdot\frac{[q+6,q+12]_3}{12}\frac{(q+11)(q+9)(q+7)(q+5)(q+3)(q+1)}{3\cdot 2^{10}} \\
&+ \left(\frac{[q,q+2]_3}{2}\cdot\frac{[q+2,q+8]_3}{8} + \frac{[q,q+4]_3}{4}\cdot\frac{[q+4,q+8]_3}{8} + \frac{[q,q+6]_3}{6}\cdot\frac{[q+6,q+8]_3}{8}\right) \\
&\times\frac{(q+7)(q+5)(q+3)(q+1)}{2^7} - \left(\frac{[q,q+2]_3}{2}\cdot\frac{[q+2,q+4]_3}{4} + \frac{[q,q+4]_3}{4}\cdot\frac{[q+4,q+4]_3}{4}\right. \\
&\left.+ \frac{[q,q+6]_3}{6}\cdot\frac{[q+6,q+4]_3}{4} + \frac{[q,q-2]_3}{-2}\cdot\frac{[q-2,q+4]_3}{4}\right)\cdot\frac{(q+3)(q+1)}{2^3}\Bigg\} + O\left(\frac{1}{h^3}\right)\Bigg] \\
= N_1\Bigg[1 &+ \frac{(q+1)(q+3)}{3^5\cdot 2^{15}h^2}\cdot\Bigg\{-(\lambda+4)^2(q^4 - 184q^3 + 3614q^2 + 15\,928q - 3231) + 648(\lambda+2) \\
&\times(q^2 - 52q - 93) + 2(\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) \\
&\cdot[(\lambda+4)(q^4 - 184q^3 + 3290q^2 + 32\,776q + 26\,901) + 162(\lambda+5) \\
&\times(q^2 - 52q - 93)] + (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 \\
&\cdot[-2(\lambda+2)(\lambda+5)(q^4 - 184q^3 + 3452q^2 + 24\,352q + 11\,835) \\
&+ (\lambda+3)(q^4 - 184q^3 + 3290q^2 + 32\,776q + 26\,901)]\Bigg\} + O\left(\frac{1}{h^4}\right)\Bigg].
\end{aligned} \tag{20}$$

The calculation of the corresponding factor in $\tilde{\phi}_2$ is more complicated. We consider first the exponential factors. Inserting the expansion

$$v(y)=\tfrac{1}{2}v^{(2)}y^2+\tfrac{1}{6}v^{(3)}y^3+\tfrac{1}{24}v^{(4)}y^4+\cdots=\lambda y^2+\frac{\lambda}{3}\Big[(\lambda+4)-(\lambda+1)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)+(\lambda+1)(\lambda+2)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)^2-\cdots\Big]y^3$$
$$+\frac{\lambda}{12}\Big[(\lambda^2+6\lambda+12)-(\lambda+1)(\lambda+5)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)+(\lambda+1)(\lambda+2)(\lambda+5)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)^2-\cdots\Big]y^4+\cdots \tag{21}$$

into the integrals

$$\int^y v^{1/2}\,dy,\qquad \int^y v^{-1/2}\,dy,$$

we obtain series expansions for the exponential factor in Eq. (13) and for χ_q [cf. Eq. (16)]. In particular, we have

$$\frac{1}{\lambda^{1/2}}\int^y v^{1/2}\,dy=\frac{1}{2}y^2+\frac{1}{2\cdot 3^2}\cdot\frac{v^{(3)}}{v^{(2)}}y^3+\frac{1}{2^5\cdot 3^2}\Big[\frac{3v^{(4)}}{v^{(2)}}-\Big(\frac{v^{(3)}}{v^{(2)}}\Big)^2\Big]y^4+\cdots. \tag{22}$$

Substituting these expressions into ϕ_2 [i.e., Eq. (13)], we obtain

$$\phi_2=N_2\,y^{-(1/2)(q+1)}e^{+(1/4)h^2y^2}\lambda^{-1/4}\Big\{1-\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}}y-\frac{1}{2^5\cdot 3^2}\Big[\frac{6v^{(4)}}{v^{(2)}}-5\Big(\frac{v^{(3)}}{v^{(2)}}\Big)^2\Big]y^2+\cdots$$
$$\times\exp\Big(+\frac{h^2}{2}\Big\{\frac{1}{2\cdot 3^2}\cdot\frac{v^{(3)}}{v^{(2)}}y^3+\frac{1}{2^5\cdot 3^2}\Big[\frac{3v^{(4)}}{v^{(2)}}-\Big(\frac{v^{(3)}}{v^{(2)}}\Big)^2\Big]y^2+\cdots\Big\}\Big)\Big(1+\frac{2}{h^2}\Big\{\frac{(q,q+4)}{4}\cdot\frac{1}{y^2}$$
$$\times\exp\Big[-4\Big(-\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}}y+\cdots\Big)\Big]+\frac{(q,q+2)}{2}\cdot\frac{1}{y}\exp\Big[-2\Big(-\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}}y+\cdots\Big)\Big]+\frac{(q,q-2)}{-2}$$
$$\cdot y\exp\Big[+2\Big(-\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}}y+\cdots\Big)\Big]+\frac{(q,q-4)}{-4}y^2\exp\Big[+4\Big(-\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}}y+\cdots\Big)\Big]+\cdots\Big\}+O\Big(\frac{1}{h^4}\Big)\Big). \tag{23}$$

Expanding the exponentials in the last bracket, we can rewrite this factor

$$1+\frac{2}{h^2}\Big\{\frac{a}{y^2}+\frac{b}{y}+c+dy+ey^2+\cdots\Big\}+O\Big(\frac{1}{h^4}\Big) \tag{24}$$

where

$$a=\tfrac{1}{4}(q,q+4)=\tfrac{1}{16}(q+1)(q+3),$$
$$b=\tfrac{1}{12}(q,q+4)\frac{v^{(3)}}{v^{(2)}}+\tfrac{1}{2}(q,q+2)$$
$$=-\frac{1}{2^4\cdot 3}(q+1)(3q+1)\Big[(\lambda+4)-(\lambda+1)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)+(\lambda+1)(\lambda+2)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)^2+\cdots\Big], \tag{25}$$
$$c=\frac{(q,q+4)}{2^5\cdot 3^2}\Big[3\frac{v^{(4)}}{v^{(2)}}+\Big(\frac{v^{(3)}}{v^{(2)}}\Big)^2\Big]+\frac{(q,q+2)}{12}\cdot\frac{v^{(3)}}{v^{(2)}}$$
$$=\frac{(q+1)}{2^7\cdot 3^2}\Big\{q\Big[-6(2\lambda^2+17\lambda+34)+(\lambda+1)(27\lambda+105)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)$$
$$-3(\lambda+1)(9\lambda^2+58\lambda+75)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)^2+\cdots\Big]$$
$$+\Big[-2(2\lambda^2+25\lambda+50)+(\lambda+1)(17\lambda+59)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)-(\lambda+1)(17\lambda^2+106\lambda+131)\Big(\frac{\delta}{2\alpha\rho\lambda}\Big)^2+\cdots\Big]\Big\}.$$

Next we expand the exponential factors in Eq. (23) and then express ϕ_2 as a series in rising powers of y by multiplying the expansion of the factors occurring in ϕ_2. In this way we obtain

$$\phi_2=N_2\,y^{-(1/2)(q+1)}e^{+(1/4)h^2y^2}\lambda^{-1/4}\Big[O\Big(\frac{1}{y^2},\frac{1}{y}\Big)+\gamma+O(y,y^2,y^3,\cdots)\Big], \tag{26}$$

where the coefficient of each power of y is an expansion in descending powers of h^2. In particular we have

$$\gamma(q,h^2)=1+\frac{2}{h^2}\Big[c+\frac{b(q-1)}{12}\cdot\frac{v^{(3)}}{v^{(2)}}+\frac{a(q-2)}{2^5\cdot 3}\cdot\frac{v^{(4)}}{v^{(2)}}+\frac{a(q^2-5q+5)}{2^5\cdot 3^2}\cdot\Big(\frac{v^{(3)}}{v^{(2)}}\Big)^2\Big]+O\Big(\frac{1}{h^4}\Big)$$
$$=1+\frac{(q+1)}{2^8\cdot 3^2h^2}\{(\lambda+4)^2(q^3-23q^2-39q-11)-6(\lambda+2)(q+2)(q+3)$$

$$+\frac{\delta(\lambda+1)}{2\alpha\rho\lambda}[-2(\lambda+4)(q^3-26q^2-54q-29)-3(\lambda+5)(q+2)(q+3)]$$
$$+(\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2[(\lambda+2)(\lambda+5)(2q^3-49q^2-93q-40)-(\lambda+3)(q^3-26q^2-54q-29]\Big\}+O\left(\frac{1}{h^4}\right). \quad (27)$$

Recalling that $y=\omega/h$ and that a second solution $\tilde{\phi}_2$ is obtained by changing the signs of q, ω^2, and h^2, we see that the coefficient of the dominant factor $e^{-(1/4)\omega^2}\omega^{(1/2)(q-1)}$ in $\tilde{\phi}_2(\omega,q,h)\equiv\phi_2(\omega\to i\omega,q\to-q,h\to ih)$ [see Eq. (11)] is

$$B=\frac{\tilde{N}_2}{\lambda^{1/4}h^{(1/2)(q-1)}}\gamma(-q,-h^2), \quad (28)$$

where

$$\tilde{N}_2(q,h^2)=N_2(-q,-h^2) \quad (29)$$

From (27) we obtain

$$B=\frac{\tilde{N}_2}{\lambda^{1/4}h^{(1/2)(q-1)}}\Big(1+\frac{(q-1)}{2^8 3^2 h^2}\Big\{[-(\lambda+4)^2(q^3+23q^2-39q+11)-6(\lambda+2)(q-2)(q-3)]$$
$$+\frac{(\lambda+1)\delta}{2\alpha\rho\lambda}[2(\lambda+4)(q^3+26q^2-54q+29)-3(\lambda+5)(q-2)(q-3)]+(\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2$$
$$\times[-(\lambda+2)(\lambda+5)(2q^3+49q^2-93q+40)+(\lambda+3)(q^3+26q^2-54q+29)]\Big\}+O\left(\frac{1}{h^4}\right)\Big). \quad (30)$$

It now follows that over their common range of validity

$$\phi_1=\tilde{\phi}_2, \quad (31)$$

with N_1 and $\tilde{N}_2$ related by $A=B$, i.e.,

$$\tilde{N}_2=N_1\cdot\lambda^{1/4}\cdot h^{(1/2)(q-1)}\left[1+\frac{m(q,\delta)}{3^5\cdot 2^{15}h^2}+O\left(\frac{1}{h^4}\right)\right], \quad (32)$$

where

$$m(q,\delta)=(\lambda+4)^2c_1(q)+6(\lambda+2)c_2(q)+(\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right)[2(\lambda+4)c_3(q)+3(\lambda+5)c_2(q)]$$
$$+(\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2[(\lambda+2)(\lambda+5)c_4(q)+(\lambda+3)c_3(q)] \quad (33)$$

and

$$\begin{aligned}c_1(q)&=-(q+1)(q+3)(q^4-184q^3+3614q^2+15\,928q-3231)+3456(q-1)(q^3+23q^2-39q+11),\\ c_2(q)&=108(q+1)(q+3)(q^2-52q-93)+3456(q-1)(q-2)(q-3),\\ c_3(q)&=(q+1)(q+3)(q^4-184q^3+3290q^2+32\,776q+26\,901)-3456(q-1)(q^3+26q^2-54q+29),\\ c_4(q)&=-2(q+1)(q+3)(q^4-184q^3+3452q^2+24\,352q+11\,835)+3456(q-1)(2q^3+49q^2-93q+40).\end{aligned} \quad (34)$$

N_2 then follows by changing the signs of q and h^2 in $\tilde{N}_2$.

4. IDENTIFICATION OF THE BOUND-STATE SOLUTION

Our aim is to determine explicitly the value of the S-wave bound-state wave function at the origin $r=0$ because this determines the various decay widths of the hadron. We are therefore interested in the solution which is regular at the origin. Since $r=0$ implies $z=-\infty$, this means that, in order to identify the regular solution, we must examine the behavior of $\phi_2(l)$, $\tilde{\phi}_2(l)$ in the limit $z\to-\infty$. We now show that $\tilde{\phi}_2(-l-1)$ possesses the required behavior, i.e., that (expressed in terms of r) $\tilde{\phi}_2(-l-1)\sim r^{l+(1/2)}$ for $r\to 0$, so that [cf. Eq. (3)] $\psi\sim r^{l+1}$. For reasons of simplicity we ignore the Coulomb perturbation in this identification (note that throughout our procedure is to treat the confinement force as that component which dominates the wave function, even at $r=0$, so that the limit $r\to 0$ is to be taken at the very end of a calculation).

We consider the solution ϕ_2, i.e., Eq. (13), and expand each of the factors in powers of e^y. For more transparency we quote a few intermediate steps. Thus, expanding in powers of e^y and using

$$e^y=e^{z-z_0}=r/\rho,$$

we find (apart from a constant)

$$\int^y[v(y)]^{1/2}dy=\left(\frac{\lambda}{\lambda+2}\right)^{1/2}\left\{\ln\frac{r}{\rho}+\left[\frac{(r/\rho)^{2+\lambda}}{\lambda(\lambda+2)}-\frac{(\lambda+2)}{4\lambda}\left(\frac{r}{\rho}\right)^2\right]+\cdots\right\}+\text{const.} \quad (35)$$

Here the dots imply terms containing higher powers of the last term in brackets. Similarly, we have (again apart from a constant)

$$\int^{y} \frac{dy}{[v(y)]^{1/2}} = \left(\frac{\lambda+2}{\lambda}\right)^{1/2} \left\{ \ln\frac{r}{\rho} - \left[\frac{(r/\rho)^{2+\lambda}}{\lambda(\lambda+2)} - \frac{(\lambda+2)}{4\lambda}\left(\frac{r}{\rho}\right)^2\right] + \cdots \right\} \tag{36}$$

and

$$v^{-1/4} = \left(\frac{\lambda+2}{\lambda}\right)^{1/4} \left\{ 1 - \left[\frac{1}{2\lambda}\left(\frac{r}{\rho}\right)^{2+\lambda} - \frac{\lambda+2}{4\lambda}\left(\frac{r}{\rho}\right)^2\right] + \cdots \right\}. \tag{37}$$

Substituting these expressions into Eq. (13), we obtain

$$\begin{aligned}\phi_2(y,q,h^2) = N_2 \left(\frac{r}{\rho}\right)^{(1/2)h^2/(\lambda+2)^{1/2}} &\exp\left(\frac{h^2}{2(\lambda+2)^{1/2}} \left\{\left[\frac{(r/\rho)^{2+\lambda}}{\lambda(\lambda+2)} - \frac{\lambda+2}{4\lambda}\left(\frac{r}{\rho}\right)^2\right] + \cdots\right\}\right) \sum_{i=0}^{\infty}\left(\frac{2}{h^2}\right)^i \sum_{j=2i}^{-\infty} (q,q+2j)_i \\ &\times \left(\frac{\lambda+2}{\lambda}\right)^{1/4} \left\{1 - \left[\frac{1}{2\lambda}\left(\frac{r}{\rho}\right)^{2+\lambda} - \frac{\lambda+2}{4\lambda}\left(\frac{r}{\rho}\right)^2\right] + \cdots\right\} \left(\frac{r}{\rho}\right)^{-(1/2)(\lambda+2)^{1/2}(q+2j)} \\ &\times \exp\left(-\tfrac{1}{2}(\lambda+2)^{1/2}(q+2j)\left\{\left[-\frac{(r/\rho)^{2+\lambda}}{\lambda(\lambda+2)} + \frac{(\lambda+2)}{4\lambda}\left(\frac{r}{\rho}\right)^2\right] + \cdots\right\}\right),\end{aligned}$$

where we have set

$$\chi(y) = \chi_q(y) + O\left(\frac{1}{h^2}\right) \equiv \sum_{i=0}^{\infty}\left(\frac{2}{h^2}\right)^i \sum_{j=2i}^{-\infty} (q,q+2j)_i\, \chi_{q+2j}(y) \tag{38}$$

[this defines the coefficients $(q,q+2j)_i$]. In the limit $r \to 0$, ϕ_2 is seen to be

$$\phi_2(y,q,h^2)\big|_{r\to 0} = N_2 \left(\frac{\lambda+2}{\lambda}\right)^{1/4} \left(\frac{r}{\rho}\right)^{+(1/2)h^2/(\lambda+2)^{1/2} - (1/2)(\lambda+2)^{1/2}q} \sum_{i=0}^{\infty}\left(\frac{2}{h^2}\right)^i \sum_{j=2i}^{0} (q,q+2j)_i \left(\frac{r}{\rho}\right)^{-(\lambda+2)^{1/2}j}. \tag{39}$$

We now recall the eigenvalue expansion derived in Ref. 1, i.e.,

$$(l+\tfrac{1}{2})^2 = \frac{h^4}{4(\lambda+2)} - \tfrac{1}{2}qh^2 + \cdots.$$

Taking the square root of this expansion, we obtain

$$l + \tfrac{1}{2} = \pm \frac{h^2}{2(\lambda+2)^{1/2}} \mp \tfrac{1}{2}(\lambda+2)^{1/2}q \pm \frac{\delta(q,h^2)}{h^2}, \tag{40}$$

where δ/h^2 is the remainder of $O(1/h^2)$, and we choose the upper signs. Thus, inserting Eq. (40) into (39) we have, for $l \geqslant 0$,

$$\phi_2(y,q,h^2)\big|_{r\to 0} = N_2 \left(\frac{\lambda+2}{\lambda}\right)^{1/4}\left(\frac{r}{\rho}\right)^{l+(1/2)}$$

and

$$\tilde{\phi}_2(y,q,h^2)\big|_{r\to 0} = \tilde{N}_2 \left(\frac{\lambda+2}{\lambda}\right)^{1/4}\left(\frac{r}{\rho}\right)^{-[l+(1/2)]}, \tag{41}$$

provided

$$\left(\frac{r}{\rho}\right)^{\delta(q,h^2)/h^2} = \sum_{i=0}^{\infty}\left(\frac{2}{h^2}\right)^i \times \sum_{j=2i}^{0} (q,q+2j)_i \left(\frac{r}{\rho}\right)^{-(\lambda+2)^{1/2}j}. \tag{42}$$

We are unable at present to give a proof of the last relation. For our purposes it suffices to know the limits (41). Of course, without our choice of the upper signs in Eq. (40), the left hand side of Eq. (42) could be the combination

$$C_1 \left(\frac{r}{\rho}\right)^{\delta(q,h^2)/h^2} + C_2 \left(\frac{r}{\rho}\right)^{[\delta(q,h^2)/h^2]+2l+1},$$

where C_1 and C_2 are constants. However, for $h^2 \to \infty$ (this is the asymptotic regime we are interested in), the right hand side of Eq. (42) is independent of r and approaches the value 1 [from Eq. (38) it can be seen that $(q,q)_0 = 1$]. Hence, $C_1 = 1$ and $C_2 = 0$. We now recall that, given one solution of the original wave equation, we can find another by replacing l by $-l-1$. Altogether, we then have the following two types of regular solutions near $r = 0$:

$$\begin{aligned}\phi_2(y,q,h^2,l)\big|_{r\to 0} &= N_2(l)\left(\frac{\lambda+2}{\lambda}\right)^{1/4}\left(\frac{r}{\rho}\right)^{l+(1/2)}, \\ \phi_{2R} &\equiv \tilde{\phi}_2(y,q,h^2,-l-1)\big|_{r\to 0} \\ &= \tilde{N}_2(-l-1)\left(\frac{\lambda+2}{\lambda}\right)^{1/4}\left(\frac{r}{\rho}\right)^{l+(1/2)}.\end{aligned} \tag{43}$$

We have seen in Sec. 3 that $\tilde{\phi}_2$ is the solution which goes over into the eigensolution ϕ_1 around a local minimum of the potential. We therefore identify ϕ_{2R} as that branch of our eigensolution which is regular at the origin (this is in fact the only possible choice since—as will be seen later—N_2 vanishes for positive integral values of q).

The continuation of the bound-state wave function to infinity is obtained in a way similar to its continuation to the origin. Again, $\tilde{\phi}_2$ is matched to ϕ_1, but now at the upper end of the region of validity of the latter. The asymptotic behavior for $r \to \infty$ is then obtained by investigating the behavior of $\tilde{\phi}_2$ in the limit $z \to \infty$, i.e.,

$$\tilde{\phi}_2 = \tilde{N}_2 \cdot \tilde{\chi} \cdot \exp\left\{-\frac{h^2}{2\lambda^{1/2}}\int^{y}[v(y)]^{1/2}\,dy\right\}$$

for (see the Appendix)

$$v(y) = -\frac{4\lambda}{h^4}[\tilde{V}(z) - \tilde{V}(z_0)]$$

$$\simeq \frac{4\lambda}{h^4}\beta\, e^{(2+\lambda)z}$$
$$= \frac{4\lambda}{h^4}\beta r^{2+\lambda}, \quad \text{for } z,r\rightarrow +\infty .$$

Considering for simplicity only the first term of the expansion of $\tilde{\chi}$, we obtain

$$\tilde{\phi}_2 \simeq \frac{\tilde{N}_2 h}{(4\lambda\beta)^{1/4} r^{(\lambda+2)/4}} \exp\left[-\frac{2\beta^{1/2}}{(\lambda+2)} r^{(\lambda+2)/2}\right] \times \exp\left[\frac{-qh^2}{2(\lambda+2)\beta^{1/2}} \frac{1}{r^{(1/2)(\lambda+2)}}\right], \quad \text{for } r\rightarrow\infty.$$

This behavior agrees with the behavior derived previously[1] by a different method and agrees in particular with well known expressions for the cases $\lambda = 1,2$.

Our bound-state eigensolution therefore consists of a central piece ϕ_1 (or pieces ϕ_1 matched together across Stokes singularities) and two pieces $\tilde{\phi}_2$ matched to ϕ_1 at the margins of its region of validity, one piece extending to $r = 0$, where it has the behavior of the regular solution, and the other piece extending to infinity, where it decreases exponentially. The behavior at infinity shows in addition that the confinement potential possesses only discrete eigenstates. Continuum states require the introduction of a cutoff.[1]

5. NORMALIZATION AND THE CALCULATION OF DECAY WIDTHS

We are now in a position to determine the normalization constants N_1 and N_2. The normalization constant of the overall wave function Ψ is defined for

$$\Psi = \frac{1}{r}\psi(r)\cdot\left[\frac{(2l+1)(l-m)!(-1)^m}{2(l+m)!}\right]^{1/2} \times P_l^m(\cos\theta)\frac{1}{(2\pi)^{1/2}}e^{im\varphi}. \tag{44}$$

The constant of normalization of $\psi(r)$ is then determined by

$$\int_0^\infty |\psi(r)|^2\, dr = 1.$$

Using Eq. (3), we can rewrite this integral as

$$\int_{-\infty}^{+\infty} dz\, e^{2z}\, |\phi_{2R}(z)|^2 = 1, \tag{45}$$

where ϕ_{2R} is the regular solution with appropriate continuation into regions where the expansion (13) does not apply.

We consider first the branch ϕ_1, i.e., Eq. (10). For ease of discussion we rewrite this solution as

$$\phi_1(\omega) \equiv N_1 \sum_{i=0}^{\infty}\sum_{j=-6i}^{6i}\frac{1}{h^i}(q,q+j)_i\, \phi_{q+j}(\omega), \tag{46}$$

and the coefficients $(q,q+j)_i$ follow by comparison with Eq. (10). Here, $\omega = h(z-z_0)$. We now set

$$z - z_0 = \frac{u}{h^\alpha} \tag{47}$$

and ask for which values of α and finite values of u is ϕ_1 an asymptotically decreasing expansion (i.e., for $h\rightarrow\infty$). The ratio of the $(i+1)$st term of Eq. (46) to the ith term is

$$R^{(i+1,i)} = \frac{1}{h}\cdot\frac{\Sigma_{j=-6(i+1)}^{6(i+1)}(q,q+j)_{i+1}\,\phi_{q+j}}{\Sigma_{j=-6i}^{6i}(q,q+j)_i\,\phi_{q+j}}. \tag{48}$$

Now, $\omega = h(z-z_0) = uh^{1-\alpha}$. We allow for large values of ω for $h\rightarrow\infty$ if $\alpha<1$. Next we recall that

$$\phi_q = D_{(q-1)/2}(\omega).$$

The asymptotic behavior of the parabolic cylinder function for large values of its argument (i.e., $h\rightarrow\infty$) is given by[2]

$$D_{(1/2)(q-1)}(uh^{1-\alpha}) = (uh^{1-\alpha})^{(1/2)(q-1)}e^{-(1/4)u^2h^{2-2\alpha}} \times \sum_{k=0}^{\infty}\frac{(-1)^k[(1/2)(q-1)]_{2k}}{k!(2^{1/2}uh^{1-\alpha})^{2k}}, \tag{49}$$

where

$$[\nu]_{2k} = \nu(\nu-1)(\nu-2k+1), \quad \text{for} \quad k>0,$$
$$= 1, \quad \text{for} \quad k = 0.$$
$$= 0, \quad \text{otherwise.}$$

Substituting Eq. (49) into (48), we have

$$R^{(i+1,i)} = \frac{1}{h}\sum_{j=-6(i+1)}^{6(i+1)}(q,q+j)_{i+1}(uh^{1-\alpha})^{(1/2)(q+j-1)} \times[1+O(1/h^{2-2\alpha})]\left(\sum_{j=-6i}^{6i}(q,q+j)_i \times(uh^{1-\alpha})^{(1/2)(q+j-1)}[1+O(1/h^{2-2\alpha})]\right)$$
$$\sim \frac{1}{h}\cdot\frac{(h^{1-\alpha})^{(1/2)(q+6i+5)}}{(h^{1-\alpha})^{(1/2)(q+6i-1)}}$$
$$\sim h^{2-3\alpha}.$$

This ratio decreases for $h\rightarrow\infty$ if $3\alpha>2$, i.e., $\alpha>\frac{2}{3}$. We have therefore

$$\tfrac{2}{3}<\alpha<1 \tag{50}$$

and so

$$|z-z_0| \lesssim O\left(\frac{1}{h^{2/3}}\right) \tag{51}$$

and

$$|\omega|\lesssim O(h^{1/3}). \tag{52}$$

We now return to the normalization integral (45). Changing the variable of integration to $y = z - z_0$, we have

$$e^{-2z_0} = \int_{-\infty}^{\infty} dy\, e^{2y}\, |\phi_{2R}|^2$$

and in terms of $\omega = hy$,

$$h\, e^{-2z_0} = \int_{-\infty}^{\infty} d\omega\, e^{2\omega/h}\, |\phi_{2R}(\omega)|^2. \tag{53}$$

The contribution of $\phi_1(\omega)$ to this integral is, according to Eqs. (31) and (52),

$$\int_{-h^{1/3}}^{h^{1/3}} e^{2\omega/h}\,|\phi_1(\omega)|^2\, d\omega = \left(\int_{-\infty}^{\infty} - \int_{-\infty}^{-h^{1/3}} - \int_{h^{1/3}}^{\infty}\right)e^{2\omega/h}\,|\phi_1(\omega)|^2\, d\omega. \tag{54}$$

Now it is clear that the integrand of these integrals can be expressed as a sum of terms such as

$$\omega^i D_{(1/2)(q-1)}(\omega)\, D_{(1/2)(p-1)}(\omega)$$

and the whole sum can be arranged in descending powers of h. Recalling the recurrence relation

$$\omega D_\nu(\omega) = \nu D_{\nu-1}(\omega) + D_{\nu+1}(\omega)\,,$$

we can evaluate the first integral in Eq. (54) with the help of the formula[2]

$$\int_{-\infty}^{\infty} [D_\nu(\omega)]^2\, d\omega = \nu!(2\pi)^{1/2}\,. \tag{55}$$

The magnitude of the second and third integrals in Eq. (54) can be estimated with the help of the asymptotic expansion (49). They are of $O(h^p e^{-h^{2/3}})$ for some finite p, and so are exponentially small. The contribution

$$\int_{h^{1/3}}^{\infty} e^{2\omega/h}\, |\phi(\omega)|^2\, d\omega \tag{56}$$

to the normalization integral is exponentially small, i.e., of $O(h^p e^{-h^{2/3}})$ in view of the exponential factor in the solution (13), i.e.,

$$\exp\left\{-\left|\frac{h^2}{2\lambda^{1/2}}\int^y [v(y)]^{1/2}\, dy\right|\right\} \sim \exp\left(-\frac{\omega^2}{4}\right)$$

for $h\to\infty$ [using the expansion (22)]. Thus, ignoring exponentially decreasing terms, we have finally

$$h\, e^{-2z_0} = \int_{-\infty}^{\infty} e^{2\omega/h}\, |\phi_1(\omega)|^2\, d\omega. \tag{57}$$

Inserting $\phi_1(\omega)$ and proceeding as explained above, we obtain

$$h\, e^{-2z_0} = (2\pi)^{1/2}\, [\tfrac{1}{2}(q-1)]!\, N_1^2 \left[1 + \frac{n(q,\delta)}{2^5\cdot 3^4 h^2} + O\left(\frac{1}{h^3}\right)\right], \tag{58}$$

where

$$n(q,\delta) = q\Bigg\{5184 + (41q^2 + 133)\Bigg[(\lambda+4)^2 - 2(\lambda+1)\times(\lambda+4)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(2\lambda^2 + 13\lambda + 17)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\Bigg]\Bigg\}. \tag{59}$$

We can now substitute N_1 of this expression into Eq. (32) and then $\tilde{N}_2$ into Eq. (43) and obtain for the value of the S-wave bound-state wave function at the origin [since $\tilde{N}_2(-l-1) = \tilde{N}_2(l)$]

$$\begin{aligned}|\Psi(0)|^2 &= \frac{1}{4\pi}\left(\frac{\lambda+2}{\lambda}\right)^{1/2}\frac{1}{\rho}\, |\tilde{N}_2|^2 \\ &= \frac{(\lambda+2)^{1/2} h^q e^{(-2z_0)}}{2(2\pi)^{3/2}\rho[(1/2)(q-1)]!} \\ &\quad\times\left[1 + \frac{m(q,\delta) - 3\cdot 2^9 n(q,\delta)}{2^{14}\cdot 3^5 h^2} + O\left(\frac{1}{h^3}\right)\right],\end{aligned} \tag{60}$$

where m and n are defined by Eqs. (34) and (59), respectively. Finally, inserting for z_0 the expansion (8), we have

$$\begin{aligned}|\Psi(0)|^2 &= \frac{(\lambda+2)^{1/2} h^q}{2(2\pi)^{3/2}\rho^3[\frac{1}{2}(q-1)]!} \\ &\quad\times\left[1 + \frac{m(q,\delta) - 3\cdot 2^9 n(q,\delta)}{2^{14}\cdot 3^5 h^2} + O\left(\frac{1}{h^3}\right)\right] \\ &\quad\times\left[1 + \frac{\delta}{2\alpha\rho\lambda} - \frac{(\lambda+1)\delta^2}{2(2\alpha\rho\lambda)^2} + \cdots\right]^{-2}.\end{aligned} \tag{61}$$

Since ρ has the dimension of a length and h is dimensionless, this expression is seen to have the required dimension of (length)$^{-3}$. For the linear potential Eq. (61) becomes

$$\begin{aligned}|\Psi(0)|^2 &= \frac{\beta\cdot 2^{(1/2)(q-4)} 3^{3/2}}{(2\pi)^{3/2}\,\Gamma[\frac{1}{2}(q+1)]} \\ &\quad\times\left\{\frac{3}{2}q + \left[f(q) + \frac{\delta}{\beta^{1/3}} g(q)\right]^{1/2}\right\}^{(1/2)(q-4)} \\ &\quad\times[1 + O(\delta/h^2)]\end{aligned} \tag{62}$$

where f and g are given by[1]

$$f(q) = \frac{3q^2+17}{3\cdot 2^3} - \frac{3q(19.85186q^2 + 0.29640)}{2^9\{\frac{3}{2}q + [(1/3\cdot 2^3)(3q^2+17)]^{1/2}\}},$$

$$g(q) = -\frac{3q + 6[(1/3\cdot 2^3)(3q^2+17)]^{1/2}}{2(\frac{3}{2})^{1/3}\{\frac{3}{2}q + [(1/3\cdot 2^3)(3q^2+17)]^{1/2}\}^{1/3}} + \frac{(6q^2+1)}{6(\frac{3}{2})^{1/3}\{\frac{3}{2}q + [1/3\cdot 2^3)(3q^2+17)]^{1/2}\}^{4/3}}.$$

Decay widths play an important role in exploring the origin of a newly found hadronic state. The leptonic and hadronic decay widths of a vector quark–antiquark bound state such as ψ and Υ can be expressed in terms of the S-wave bound-state wave function at the origin. Thus,[4]

$$\Gamma(\psi\to l\bar{l}) = \frac{16\pi\alpha^2 e_Q^2}{m_\psi^2}\, |\Psi(0)|^2 \tag{63}$$

and

$$\Gamma(\psi\to \text{ hadrons}) = \frac{160(\pi^2-9)}{81}\cdot\frac{\alpha_s^3}{m_\psi^2}\, |\Psi(0)|^2. \tag{64}$$

Here, α is the fine structure constant [not to be confused with α of Eq. (6)] and e_Q is the charge of the constituent quark of ψ and m_ψ the mass of ψ. Using Eq. (62), we obtain 2.02 and 3.28 keV for the leptonic decay rates of ψ (3.096 GeV) and ψ' (3.684 GeV), respectively. These values have to be compared with the experimentally determined[5] widths: 4.8 ± 0.6 keV for ψ and 2.1 ± 0.3 keV for ψ'. Including the factors $[1 + O(1/h^2)]/[1 + O(\delta)]$ given in Eq. (61) does not improve our prediction. This is due to the fact that the values of h^2 involved are not sufficiently large to ensure that $O(1/h^2)$ is considerably less than one. It does not suffice, therefore, to go only as far as the first nontrivial term in the expansion in descending powers of h. In all these calculations we used the value of β determined previously[1] by requiring the difference of the masses of the states $\psi(1\ {}^3S_1)$, and $\psi'(2\ {}^3S_1)$ to be 0.588 GeV. This value is $\beta = 0.3660$ GeV.[3] We have also chosen the same value of δ as before, i.e., $\delta = 0.1788$ GeV.

6. CONCLUSIONS

In the foregoing we have developed a general method for normalizing nontrivial bound-state wave functions. The reader with little experience in handling perturbation expansions may think that our expansions are complicated or even clumsy. A close look, however, will reveal that our expansions are the most natural and direct expansions one can derive, and that the notation has been devised in such a way that their complexity is reduced to a bare minimum. The perturbation method underlying our solutions is the only one we know of which places the WKB-like solution on the same footing as the Rayleigh–Schrodinger-like perturbation solution around a local minimum of the potential, and which is shown to yield explicitly one and the same eigenvalue expansion together with both types of solutions. We have demonstrated explicitly that various pieces of the overall bound-state solution can be continued into one another in regions of common validity.

In particular, we have shown that the regular solution can be identified and that, after matching to other branches, it can be continued to the expected exponentially decreasing behavior at infinity. We have then shown that as a consequence of our eigenvalue condition the dominant contribution to the normalization integral comes from fluctuations around a local minimum of the associated potential. The explicit demonstration of this connection is the main point of this paper and shows the validity of the eigenvalue expansion used for the calculation of the mass spectrum. Of course, since our expansion are asymptotic in h, they are particularly useful in cases where the binding energy α is large. The "usual WKB approximation" corresponds to the dominant terms of our expansions; it is therefore subject to the same restrictions, although few authors using the WKB approximation seem to realize this.

APPENDIX

In this Appendix we give the derivation of the solution ϕ_2. Only the special case $\delta = 0$ (no Coulomb component) has been given in Ref. 1. Our starting point is Eq. (4), in which we insert the expansion of $\tilde{V}(z)$ around its maximum $\tilde{V}(z_0)$ at z_0. Comparing the resulting equation with the equation of parabolic cylinder functions, we obtain (as shown in Ref. 1) the secular equation

$$-L^2 + \tilde{V}(z_0) = \tfrac{1}{2}qh^2 + \Delta h, \tag{A1}$$

where q is an odd integer and Δ remains to be determined. We substitute this relation for L^2 into Eq. (4) and obtain

$$\frac{d^2\phi}{dz^2} + [-\tilde{V}(z_0) + \tfrac{1}{2}qh^2 + \Delta h + \tilde{V}(z)]\,\phi = 0. \tag{A2}$$

Now,

$$\tilde{V}(z) = \alpha\, e^{2z} - \beta\, e^{(2+\lambda)z} + \delta\, e^{z} \tag{A3}$$

and z_0 is given by

$$\tilde{V}'(z_0) = 0. \tag{A4}$$

Next we change the variable in Eq. (A2) to

$$y = z - z_0 \tag{A5}$$

and recall that [cf. Eq. (9)]

$$h^4 = -2\tilde{V}^{(2)}(z_0). \tag{A6}$$

Then,

$$\frac{d^2\phi(y)}{dy^2} - \frac{h^4}{4\lambda}v(y)\,\phi(y) + (\tfrac{1}{2}qh^2 + \Delta h)\,\phi(y) = 0, \tag{A7}$$

where

$$v(y) = \frac{2\lambda}{\tilde{V}^{(2)}(z_0)}[\tilde{V}(y+z_0) - \tilde{V}(z_0)]. \tag{A8}$$

We observe that Eqs. (A4) and (A8) imply that

$$v(0) = 0, \quad v'(0) = 0. \tag{A9}$$

The function v therefore has the expansion

$$v(y) = \sum_{i=2}^{\infty} \frac{y^i}{i!} v^{(i)}(0), \tag{A10}$$

where

$$\begin{aligned} v^{(i)}(0) &= 2\lambda\frac{\tilde{V}^{(i)}(z_0)}{\tilde{V}^{(2)}(z_0)} \\ &= \left([2(2+\lambda)^{i-1} - 2^i] + (2^{i-1} - 1)\cdot\frac{\delta}{\alpha\rho}\cdot\zeta^{-(1+\lambda)}\right) \\ &\quad \times\left[1 + \frac{\delta}{2\alpha\rho\lambda}\zeta^{-(1+\lambda)}\right]^{-1} \end{aligned} \tag{A11}$$

and [see Eqs. (7) and (8)]

$$\zeta = \frac{1}{\rho}e^{z_0} = 1 + \frac{\delta}{2\alpha\rho\lambda} - \frac{(\lambda+1)\delta^2}{2(2\alpha\rho\lambda)^2} + \cdots. \tag{A12}$$

Substituting Eqs. (A12) into (A11), we obtain

$$\begin{aligned} v^{(i)}(0) &= [2(2+\lambda)^{i-1} - 2^i] \\ &\quad + \frac{\delta}{2\alpha\rho\lambda}[2^i(\lambda+1) - 2\lambda - 2(2+\lambda)^{i-1}] \\ &\quad + \frac{\delta^2}{(2\alpha\rho\lambda)^2}[2(2+\lambda)^i - 2^i(\lambda+1)(\lambda+2) \\ &\quad + 2\lambda(\lambda+2)] + \cdots. \end{aligned} \tag{A13}$$

We now proceed along the lines of Ref. 1. Setting in Eq. (A7)

$$\phi(y) = \chi(y)\exp\left\{\pm\frac{h^2}{2\lambda^{1/2}}\int^y [v(y)]^{1/2}dy\right\}, \tag{A14}$$

we find that $\chi(y)$ satisfies

$$\begin{aligned} &\frac{d^2\chi}{dy^2} \pm \frac{h^2}{\lambda^{1/2}}v^{1/2}(y)\frac{d\chi}{dy} \pm \frac{h^2}{4\lambda^{1/2}}\cdot\frac{v'(y)}{v^{1/2}(y)}\chi \\ &\quad + (\tfrac{1}{2}qh^2 + \Delta h)\chi = 0. \end{aligned} \tag{A15}$$

From now on we consider only the equation for the upper signs. The equation for the lower signs leads to another solution which can be obtained from the solution we shall derive by replacing $v^{1/2}$ by $-v^{1/2}$ or by changing the signs of q and h^2 throughout. Thus, choosing the upper signs in Eq. (A15), we can rewrite the equation in the form

$$D_q\chi = \frac{2}{h^2}\left(\frac{d^2\chi}{dy^2} + \Delta h\chi\right), \tag{A16}$$

$$D_q \equiv -\frac{2}{\lambda^{1/2}}v^{1/2}\frac{d\chi}{dy} - \frac{1}{2\lambda^{1/2}}\frac{v'}{v^{1/2}}\chi - q\chi. \tag{A17}$$

By construction, Δh is at most of $O(0)$ in h^2 for $h^2 \to \infty$. Hence, to a first approximation we can neglect the terms on the right hand side of Eq. (A16) and write for the solution to that order

$$\chi^{(0)} = \chi_q , \tag{A18}$$

where χ_q is the solution of

$$D_q \chi_q = 0,$$

i.e.,

$$\chi_q(y) = \frac{1}{v^{1/4}} \exp\left\{ - \frac{q\lambda^{1/2}}{2} \int \frac{dy}{[v(y)]^{1/2}} \right\}, \tag{A19}$$

apart from an overall multiplicative constant which we ignore in the following except in the context of normalization.

Proceeding as in previous applications of our perturbation method,[1] we evaluate $d^2\chi_q/dy^2$ and obtain

$$\frac{d^2\chi_q}{dy^2} + \Delta h \chi_q = \left(\Delta h + \frac{5}{16}\cdot\frac{v'^2}{v^2} + \frac{q\lambda^{1/2}v'}{2v^{3/2}} + \frac{q^2\lambda}{4v} - \frac{v''}{4v} \right) \chi_q . \tag{A20}$$

Looking at Eq. (A19), we observe the following relations which will be used in the following:

$$\frac{\chi_{q+j}}{\chi_q} = \left(\frac{\chi_{q+1}}{\chi_q} \right)^j, \quad \frac{\chi_{q+j}}{\chi_q} = \frac{\chi_q}{\chi_{q-j}} . \tag{A21}$$

The next step of our procedure is to re-express Eq. (A20) as a sum over various χ_{q+j}. It is not difficult to convince oneself that the series reversion which this step implies is possible only if $v(y)$ is expanded around a point $y = y_0$ for which both

$$v(y_0) = 0 \quad \text{and} \quad v'(y_0) = 0.$$

We have observed earlier [see Eq. (A9)] that this is the point $y = 0$. Hence, we now use Eqs. (A10) and (A13) in order to re-express Eq. (A20) as a sum over various χ_{q+j}. We then have (apart from an additive constant)

$$\frac{\lambda^{1/2}}{2} \int^y \frac{dy}{[v(y)]^{1/2}} = \tfrac{1}{2}\ln y + \sum_{i=1}^{\infty} \gamma_i y^i , \tag{A22}$$

where

$$\begin{aligned} \gamma_1 &= -\frac{1}{12}\cdot\frac{v^{(3)}}{v^{(2)}} \\ &= -\frac{1}{12}\left[(\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ \gamma_2 &= -\frac{1}{2^5\cdot 3}\left[\frac{v^{(4)}}{v^{(2)}} - \left(\frac{v^{(3)}}{v^{(2)}}\right)^2 \right] \\ &= \frac{1}{2^5\cdot 3}\left[2(\lambda+2) - (\lambda+1)(\lambda+3)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda^2+6\lambda+7)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ \gamma_3 &= \frac{1}{2^4\cdot 3^4\cdot 5}(\lambda+1)(\lambda-2)(\lambda+4) + O(\delta). \end{aligned} \tag{A23}$$

Expression (A22) can now be substituted into the relation

$$\frac{\chi_{q-1}}{\chi_q} = \exp\left[\frac{\lambda^{1/2}}{2} \int^y \frac{dy}{[v(y)]^{1/2}} \right] \tag{A24}$$

and the resulting expansion in power of y can be reserved. We then have

$$y^{1/2} = \sum_{i=0}^{\infty} d_{2i+1} \frac{\chi_{q-(2i+1)}}{\chi_q} , \tag{A25}$$

where

$$\begin{aligned} d_1 &= 1, \\ d_3 &= -\gamma_1 = \frac{1}{12}\left\{ (\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ d_5 &= -\gamma_2 + \tfrac{5}{2}\gamma_1^2 \\ &= \frac{1}{2^5\cdot 3^2}\left[(5\lambda^2+34\lambda+68) - (\lambda+1)(7\lambda+31) + \left(\frac{\delta}{2\alpha\rho\lambda}\right)(\lambda+1)(7\lambda^2+47\lambda+64)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ d_7 &= \frac{1}{2^7\cdot 3^3\cdot 5}(\lambda+4)(79\lambda^2+446\lambda+892) + O(\delta). \end{aligned} \tag{A26}$$

Inserting Eq. (A25) into (A10) and inverting the series, we obtain

$$\frac{1}{v(y)} = \frac{1}{\lambda} \sum_{i=2,1,0}^{-\infty} \delta_{2i} \frac{\chi_{q+2i}}{\chi_q} , \tag{A27}$$

where

$$\begin{aligned} \delta_4 &= 1, \\ \delta_2 &= -\tfrac{2}{3}\left[(\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ \delta_0 &= \frac{1}{2^3\cdot 3}\left[2(\lambda^2+11\lambda+22) - (\lambda+1)(7\lambda+25)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(7\lambda^2+44\lambda+55) \times \left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \\ \delta_{-2} &= -\frac{(\lambda+4)}{2^5\cdot 3^3\cdot 5}(727\lambda^2+4718\lambda+19\,232) + O(\delta). \end{aligned} \tag{A28}$$

In a similar way we find

$$\frac{v'^2}{v^2} = \sum_{i=2,1,0}^{-\infty} \tau_{2i} \frac{\chi_{q+2i}}{\chi_q} , \tag{A29}$$

where

$$\begin{aligned} \tau_4 &= 4, \qquad \tau_2 = 0, \\ \tau_0 &= \frac{(\lambda+1)}{2\cdot 3^2}\left[2(\lambda-2) + (5\lambda+11)\left(\frac{\delta}{2\alpha\rho\lambda}\right) - (5\lambda^2+28\lambda+29)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots \right], \end{aligned} \tag{A30}$$

$$\tau_{-2} = -\frac{(\lambda+4)}{2^3\cdot 3^3\cdot 5}(1495\lambda^2+9590\lambda+38\,772) + O(\delta),$$

etc., and

$$\frac{v''}{v} = \sum_{i=2,1,0}^{-\infty} \epsilon_{2i} \frac{\chi_{q+2i}}{\chi_q}, \tag{A31}$$

where

$$\epsilon_4 = 2,$$

$$\epsilon_2 = \frac{2}{3}\left[(\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\right],$$

$$\epsilon_0 = \frac{1}{2^2\cdot 3}\left[2(\lambda+1)(\lambda-2) + (\lambda+1)(5\lambda+11)\left(\frac{\delta}{2\alpha\rho\lambda}\right) - (\lambda+1)(5\lambda^2+28\lambda+29)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\right], \tag{A32}$$

$$\epsilon_{-2} = -\frac{(\lambda+4)}{2^4\cdot 3^3\cdot 5}(667\lambda^2 + 4778\lambda + 19\,352) + O(\delta),$$

etc., and finally

$$\frac{v'}{v^{3/2}} = \frac{1}{\lambda^{1/2}} \sum_{i=2,1,0}^{-\infty} \kappa_{2i} \frac{\chi_{q+2i}}{\chi_q}, \tag{A33}$$

where

$$\kappa_4 = 2,$$

$$\kappa_2 = -\frac{2}{3}\left[(\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\right], \tag{A34}$$

$$\kappa_0 = 0,$$

$$\kappa_{-2} = -\frac{(\lambda+4)}{2^4\cdot 3^3\cdot 5}(1151\lambda^2 + 7114\lambda + 28\,922) + O(\delta),$$

etc. These expansions can now be substituted in Eq. (A20). Then,

$$\frac{d^2\chi_q}{dy^2} + \Delta h \chi_q = \sum_{j=2,1,0}^{-\infty} (q,q+2j)\,\chi_{q+2j}, \tag{A35}$$

where for $j \neq 0$

$$(q,q+2j) = \frac{5}{16}\tau_{2j} + \frac{q}{2}\kappa_{2j} + \frac{q^2}{4}\delta_{2j} - \frac{1}{4}\epsilon_{2j}, \tag{A36}$$

and for $j = 0$

$$\begin{aligned}(q,q) &= \Delta h + \frac{5}{16}\tau_0 + \frac{q}{2}\kappa_0 + \frac{q^2}{4}\delta_0 - \frac{1}{4}\epsilon_0 \\ &= \Delta h + \frac{1}{2^5\cdot 3^2}\Big\{[6q^2(\lambda^2+11\lambda+22) - 2(\lambda+1) \\ &\quad \times(\lambda-2)] + (\lambda+1)[-3q^2(7\lambda+25) \\ &\quad -(5\lambda+11)]\frac{\delta}{2\alpha\rho\lambda} + (\lambda+1)\,[+3q^2(7\lambda^2 \\ &\quad + 44\lambda+55) + (5\lambda^2+28\lambda+29)]\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\Big\}.\end{aligned} \tag{A37}$$

For use in Sec. 3 we note the particular values

$$(q,q+4) = \tfrac{1}{4}(q+1)(q+3),$$

$$(q,q,+2) = -\tfrac{1}{6}(q+1)^2\left[(\lambda+4) - (\lambda+1)\left(\frac{\delta}{2\alpha\rho\lambda}\right) + (\lambda+1)(\lambda+2)\left(\frac{\delta}{2\alpha\rho\lambda}\right)^2 + \cdots\right]. \tag{A38}$$

Having determined these coefficients, we can now proceed with the calculation of the complete perturbation expansion. Thus, the first approximation $\chi^{(0)} = \chi_q$ leaves uncompensated on the right hand side of Eq. (A16) a sum of terms amounting to

$$R_q^{(0)} = \frac{2}{h^2}\sum_{j=2,1,0}^{-\infty}(q,q+2j)\,\chi_{q+2j}\,. \tag{A39}$$

Now since

$$D_{q+j} = D_q - j \quad \text{and} \quad D_q\,\chi_{q+j} = j\chi_{q+j}, \tag{A40}$$

we see that the terms of Eq. (A39) can be taken care of by adding to $\chi^{(0)}$ the next order contribution

$$\chi^{(1)} = \frac{2}{h^2}\sum_{\substack{j=2,1,-1\\ j\neq 0}}^{-\infty}\frac{(q,q+2j)}{2j}\chi_{q+2j}, \tag{A41}$$

excluding, of course, the term in χ_q. The coefficient of χ_q in Eq. (A39) set equal to zero, i.e.,

$$(q,q) = 0,$$

yields an expression for Δ which is identical with the expression obtained in Ref. 1 in connection with our other type of eigensolution. The present calculation therefore provides a check not only on the expression for Δ (to lowest order), but also on the coefficients τ_0, κ_0, δ_0, and ϵ_0 above.

The complete solution is obtained in our standard fashion,[1] leading to the sum

$$\chi = \chi^{(0)} + \chi^{(1)} + \chi^{(2)} + \cdots$$

in descending powers of h^2. The corresponding equation for Δ and thus the eigenvalues is

$$0 = (q,q) + \frac{2}{h^2}\sum_{\substack{j=2,1,\cdots\\ j\neq 0}}^{-\infty}\frac{(q,q+2j)}{2j}(q+2j,q) + \cdots.$$

Successive contribution $\chi^{(0)}, \chi^{(1)}, \ldots$ of χ form a rapidly decreasing sequence provided that y is large, i.e., z away from z_0.

[1]H.J.W. Müller-Kirsten, G.E. Hite, and S.K. Bose, "Explicit solution of the wave equation for arbitrary power potentials with application to charmonium spectroscopy," to appear in J. Math. Phys.

[2]R.B. Dingle, *Asymptotic Expansions: Their Derivation and Interpretation* (Academic, London, 1973).

[3]M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964).

[4]R.P. van Royen and V.F. Weisskopf, Nuovo Cimento A **50**, 617 (1967).

[5]P.A. Rapidis *et al.*, Phys. Rev. Lett. **39**, 526 (1977).

Quantum theory of infinite component fields

A. O. Barut

Department of Physics and Astrophysics, University of Colorado, Boulder, Colorado 80309

C. K. E. Schneider[a]

Department of Physics, New York University, 4 Washington Place, New York, New York 10003

Raj Wilson

Dublin Institute for Advanced Studies, 10 Burlington Road, Dublin 4, Ireland
and Department of Theoretical Physics, University of Madras, Madras 600025, India

(Received 12 July 1978)

The quantum theory of the infinite component SO(4,2) fields is formulated as a model for relativistic composite objects. We discuss three classes (timelike, lightlike, and spacelike) of physical solutions to a general class of infinite component wave equations. These solutions provide a definite physical interpretation to infinite component wave equations and are obtained by reducing SO(4,2) with respect to its orthogonal, pseudo-orthogonal, and Euclidean subgroups. The analytic continuations among these solutions are established. In the nonrelativistic limit the timelike physical states exactly reduce to the Schrödinger solution for the hydrogen atom—the simplest composite object. The wave equations for the three classes are studied in two different realizations. In one case the equations describe a three-dimensional internal Kepler motion with a discrete and a continuous energy spectrum and in the other case the equations describe a four-dimensional internal oscillatory motion with attractive as well as repulsive potentials. It is found that the Kustaanheimo–Steifel transformation of classical mechanics exactly relates these two internal motions also in the quantum case. Thus a completely relativistic theory of composite systems is established for which the internal dynamics is the generalization of the nonrelativistic two-body dynamics.

I. INTRODUCTION

There have been various approaches and attempts to formulate a relativistic quantum field theory of composite objects. In most of the attempts, models have been constructed to describe composite objects in terms of pointlike or elementary constitutents or fields which carry the basic guiding properties of the model.

In this paper we discuss models based on the concept of internal relativistic symmetries[1] on a representation space of $G \equiv SO(4,2)$. We do not assume any detailed internal dynamical mechanisms, nor the existence of any constituent pointlike objects. The internal dynamical group G and its Lie algebra serve the purpose of defining the spin and other internal quantum numbers, ensure relativistic covariance through the Lorentz subgroup and specify physical observables such as the mass, currents, etc. One of the reasons for seeking a group larger than the Lorentz group such as G is that the basis states belonging to the unitary irreducible representations (UIRs) of a group of rank 2 do not contain enough quantum numbers to describe a composite object. The relativistic theory based on the most degenerate UIR of G seems to be sufficient to describe many physical processes,[2] although one may use more general representations of G to describe more complex physical processes.

Here we attempt to give a self-consistent description of a quantum theory of a composite object using the SO(4,2) fields and discuss the solutions of a general class of wave equations for different regions of the four-momentum space. Similar investigations[3] have been made for the SL(2.C) and SO(3,2) fields. The emphasis is on the physical interpretation. The problem of introduction of internal coordinates into the infinite-component wave equations, hence, their physical interpretation in simpler cases was initiated earlier.[4] The material of this paper is arranged and summarized as follows: In Sec. II we give some general properties of the most degenerate UIR's of G that we shall need. We reduce G with respect to its orthogonal, pseudo-orthogonal, and Euclidean subgroups and give four lemmas, which enable us to determine the spectrum of certain generators. For the most degenerate UIR of G the generators are realized in terms of four real parameters. The choice of these internal parameters in fact prescribes the internal dynamical motion, e.g., Kepler motion, oscillatory motion, etc., in the theory. In (II.9) we have given three such realizations. Two of them are connected for $l_0 = 0$ by the Kustaanheimo–Stiefel (K–S) transformation. In (II.11) we have given two adjoint realizations of G.

We decompose G into irreducible components and construct orthonormal basis functions by diagonalizing a complete set of invariant operators. These basis functions are given in Table I. We expand all noncanonical bases in terms

[a]Supported by Deutscher Akademischer Austauschdienst (DAAD).

of the canonical one and the expansion coefficients are obtained explicity. In Appendix A we give the actions of the generators of G on the basis functions constructed in Sec. II.

In Sec. III we discuss the physical applications of Sec. II. We start with a general Lorentz invariant wave equation and classify its physical solutions in four-momentum space. We obtain the physical states from the basis states in Table I by applying the appropriate tilt transformation. The time-like physical solution reduces to the Schrödinger solution for the hydrogen atom in a particular limit.

We express the wave equation in terms of spherical coordinates $(\alpha,r,\theta,\varphi)$ and four real variables $(u_1,\ldots,u_4)$. We find that in the first case the wave equation describers a three-dimensional internal Kepler motion while in the second case the same wave equation describes a four-dimensional internal harmonic oscillator, as summarized in Table II. It is found that these two internal motions are related by the KS transformation. Also we establish analytic continuations among the physical solutions as shown in Table III. We have also calculated in closed form the matrix elements of the Lorentz boost between two canonical physical states. Similar matrix elements between noncanonical physical states can be found using the expansion coefficients.

Notation: Small letters are used to denote Lie algebras, big letters to denote groups. Example: so(4,2) and SO(4,2).

II. SOME PROPERTIES OF THE MOST DEGENERATE UIR'S OF SO(4,2)

A. Preliminaries

The Lie algebra so(4,2) of the pseudo-orthogonal group SO(4,2) contains the fifteen operators L_{AB},

$$\begin{aligned} &L_{AB} = -L_{BA}, \\ &[L_{AB},L_{CD}] = i(g_{AD}L_{BC} + g_{BC}L_{AD} - g_{AC}L_{BD} - g_{BD}L_{AC}), \\ &A,B,\cdots = 1,\ldots,6, \quad g_{AA} = (-\;-\;-\;-\;+\;+). \end{aligned} \tag{II.1}$$

The Minkowski space is $\{1,2,3;5=0\}$. Often we denote the set of operators contained in subgroups by subscripts, e.g., $\{L_{35},L_{36},L_{56}\}\equiv \mathrm{so}(1,2)_{356}$. The so(4,2) algebra contains the angular momentum vector $\mathbf{L}(L_{23},L_{31},L_{12})$, the Lenz–Runge vector $\mathbf{A}(L_{14},L_{24},L_{34})$, the Lorentz boost vector $\mathbf{M}(L_{15},L_{25},L_{35})$, the current vector operator $\mathbf{\Gamma}(L_{16},L_{26},L_{36})$, and $\mathrm{so}(1,2)_{456} = \{\Gamma_0 = L_{56},\ S = L_{46},\ T = L_{45}\}$.

The eigenvalues of the three Casimir invariants are expressed in terms of three parameters[5] (l_0,ρ_1,ρ_2):

$$\begin{aligned} I_2 &= \tfrac{1}{2}L_{AB}L^{AB} = l_0^2 + \rho_1^2 + \rho_2^2 - 5, \\ I_3 &= \tfrac{1}{24}\epsilon^{ABCDEF}L_{AB}L_{CD}L_{EF} = l_0\rho_1\rho_2, \\ I_4 &= \tfrac{1}{4}L_{AB}L^{BC}L_{CD}L^{DA} - \tfrac{1}{4}I_2^2 - 2I_2, \\ &= \tfrac{1}{4}[l_0^2 + \rho_1^2 + \rho_2^2 - 1]^2 - (l_0^2\rho_1^2 + l_0^2\rho_2^2 + \rho_1^2\rho_2^2), \end{aligned} \tag{II.2}$$

where $l_0 = 0,\frac{1}{2},1,\ldots;\ \rho_1,\rho_2\in\mathbb{C}$. However, for unitary representations, $\rho_1,\rho_2\in\mathbb{R}$. For $\rho_1=0, \rho_2=1$ (or $\rho_1=1,\rho_2=0$), one obtains a most degenerate UIR. Furthermore, Barut and Böhm[6] obtained a class of all UIR's, which fulfill the restrictions:

$$\rho_1^2+\rho_2^2 = 2(1+l_0^2), \quad \rho_1\rho_2 = 1-l_0^2, \qquad l_0 = 0,1,2,\cdots \quad \text{or} \quad \tfrac{1}{2},\tfrac{3}{2},\cdots.$$

These restrictions are equivalent to the "representation relation[6]"

$$\{L_{AB},L^{AC}\} = -2(1-l_0^2)\,g_B^C \tag{II.3}$$

and (II.2) becomes

$$\begin{aligned} &I_2 = -3(1-l_0^2), \quad I_3 = l_0(1-l_0^2), \\ &I_4 = -\tfrac{3}{4}(1-l_0^2)^2. \end{aligned}$$

In general, the basis of a UIR of SO(4,2) can be uniquely labeled by the eigenvalues of the nine invariant operators. In the case of the most degenerate UIR, one needs only one Casimir operator and the eigenvalues of the three mutually commuting operators to label the basis vectors. This UIR satisfies (II.3) and l_0 gives the lowest spin of the UIR. The choice of these three diagonal operators depends on the particular physical system (cf. Sec. III).

Considering reductions of $\mathrm{SO}(4,2)\equiv G$ with respect to orthogonal and pseudo-orthogonal subgroups and using (II.1) and (II.3), we can prove[7]:

Lemma 1: The Casimir invariants of a maximal noncompact subgroup $K = \mathrm{SO}(4,1)$ are

$$\begin{aligned} &\tfrac{1}{2}L_{\alpha\beta}L^{\alpha\beta} = -2(1-l_0^2), \\ &\tfrac{1}{64}\epsilon^{\alpha\beta\gamma\delta\epsilon}\epsilon_\alpha^{\beta'\gamma'\delta'\epsilon'}L_{\beta\gamma}L_{\delta\epsilon}L_{\beta'\gamma'}L_{\delta'\epsilon'} = -l_0^2(1-l_0^2), \\ &\alpha,\beta,\cdots\in I_K, \end{aligned} \tag{II.4}$$

$I_X\equiv$ set of indices for generators of X.

Lemma 2: The product of the elements of the coset G/K is

$$L_{\tau\alpha}L^{\tau\alpha} = -(1-l_0^2), \quad \alpha\in I_K, \quad \tau\notin I_K. \tag{II.5}$$

Lemma 3: Let H be a compact or noncompact stability subgroup of G.

Then

$$\begin{aligned} &L_{\mu\nu}L^{\mu\nu} - L_{\tau\tau'}L^{\tau\tau'} = -2(1-l_0^2), \\ &(\mu,\nu)\neq(\tau,\tau'), \quad \mu,\nu,\tau,\tau'\in I_H. \end{aligned} \tag{II.6}$$

The algebra of the iso-Poincaré subgroup $E(3,1)$ is obtained[8] from $\mathrm{Ad}\exp(i\alpha L_{35})L_{AB}$. In the limit $\alpha\to\pm\infty$ the translation operators $F_e^{(\pm)}\equiv L_{5e}\pm L_{3e}$ behave as

$$\mathrm{Ad}\exp(i\alpha L_{35})F_e^{(\pm)}\to F_e^{(\pm)}\exp(\pm\alpha), \quad F_e^{(\pm)}\in A^{(\pm)}$$

and

$$\mathrm{Ad}\exp(i\alpha L_{35})L_{ef}\to L_{ef}, \quad L_{ef}\in\mathrm{so}(3,1), \quad e,f = 1,2,4,6,$$

where

$A^{(\pm)}\oplus\mathrm{so}(3,1) = e^{(\pm)}(3,1)$ is the Lie algebra of $E^{(\pm)}(3,1)$. The operators of $e^{(\pm)}(3,1)$ satisfy the following commutation relations:

$$\begin{aligned} &[F_e^{(\pm)}, F_f^{(\pm)}] = 0, \\ &[L_{35},L_{ef}] = 0, \quad [L_{35}, F_e^{(\pm)}] = (\pm)iF_e^{(\pm)}, \\ &[F_e^{(\pm)},L_{fg}] = i[g_{ef}F_g^{(\pm)} - g_{eg}F_f^{(\pm)}], \end{aligned}$$

$$
\begin{aligned}
&[F_e^{(+)},F_f^{(-)}] = -2i[L_{ef}+g_{ef}L_{35}],\\
&[F_+^{(\pm)},F_-^{(\pm)}]=0, \qquad\qquad \text{(II.7)}\\
&[L_\pm,F_\pm^{(\pm)}]=0,\quad [L_+,L_-]=2L_{12},\\
&[L_{12},L_\pm] = \pm L_\pm,\quad [L_{12},F_\pm^{(\pm)}] = \pm F_\pm^{(\pm)},\\
&[L_\pm,F_\mp^{(\pm)}] = \pm 2F_4^{(\pm)},\\
&F_\pm^{(\pm)} \equiv F_1^{(\pm)} \pm iF_2^{(\pm)},\quad L_\pm \equiv L_{24} \pm iL_{41}.
\end{aligned}
$$

The two Casimir invariants of $E(3,1)$ are F_eF^e and W_eW^e, where $W_e = \frac{1}{2}\epsilon_e^{fgh}F_fL_{gh}$.

Considering the reduction of $E(3,1)$ to $E(3)$ or $E(2,1)$ and taking the action of Ad $\exp(i\alpha L_{35})$ on (II.6) in the limit $\alpha\to\mp\infty$, we can prove

Lemma 4: Let E be a Euclidean stability subgroup of G, then

$$
F_mF^m + F_lF^l = 0,\quad m\in I_E,\quad l\neq m. \qquad \text{(II.8)}
$$

The Lemmas 3 and 4 relate the spectrum of the square of certain operators $L\in so(4,2)$ to the Casimir invariants of subalgebras, which commute with L.

B. Some realizations of SO(4,2)

The generators of SO(4,2) may be realized in terms of six real variables x_A as $L_{AB} = i[g_{AA}x_A\partial_{x_B} - g_{BB}x_B\partial_{x_A}]$. In this space, the SO(4,2) transformations leave x_Ax^A invariant. However, for the most degenerate UIR satisfying (II.3), four real variables $(u_1,\ldots,u_4)$, or two complex variables (z_1,z_2), are sufficient. In Eq. (II.9) we have expressed the generators in terms of z_1,z_2, in terms of the three Cartesian coordinates and an angle α, and in terms of spin operators a_m, b_m; $m,n = 1,2$, satisfying Bose commutation relations $[a_m,a_n^+] = [b_m,b_n^+] = \delta_{mn}$:

$$
\begin{aligned}
L_{ij} &= \mathbf{L} = \tfrac{1}{2}\epsilon_{ijk}[a^+\sigma_k a + b^+\sigma_k b] = \tfrac{1}{2}\epsilon_{ijk}[z\sigma_k\partial_z - \partial_{\bar z}\sigma_k\bar z] = -i(\mathbf{r}\times\nabla)_k + \frac{|\mathbf{r}|}{\xi^2}l_0\xi_k\\
L_{i4} &= \mathbf{A} = \tfrac{1}{2}[a^+\sigma_i a - b^+\sigma_i b] = \tfrac{1}{2}[z\sigma_i\bar z - \partial_{\bar z}\sigma_i\partial_z] = \tfrac{1}{2}\Big[-x_i\nabla^2 - x_i + 2\partial_{x_i}(\mathbf{r}\cdot\nabla) + 2i\frac{|\mathbf{r}|}{\xi^2}(\boldsymbol{\xi}\times\nabla)_i - (-1)^{\delta_{i3}}\frac{l_0^2}{\xi^2}x_i\Big]\\
L_{i5} &= \mathbf{M} = \tfrac{1}{2}[a^+\sigma_i cb^+ - ac\sigma_i b] = -\tfrac{1}{2}[z\sigma_i\bar z + \partial_{\bar z}\sigma_i\partial_z]\\
&= \tfrac{1}{2}\Big[-x_i\nabla^2 + x_i + 2\partial_{x_i}(\mathbf{r}\cdot\nabla) + 2i\frac{|\mathbf{r}|}{\xi^2}(\boldsymbol{\xi}\times\nabla)_i - (-1)^{\delta_{i3}}\frac{l_0^2}{\xi^2}x_i\Big]\\
L_{i6} &= \boldsymbol{\Gamma} = \frac{i}{2}[a^+\sigma_i cb^+ + ac\sigma_i b] = \frac{i}{2}[z\sigma_i\partial_z + \partial_{\bar z}\sigma_i\bar z] = -i|\mathbf{r}|\partial_{x_i} - \frac{l_0}{\xi^2}(\mathbf{r}\times\boldsymbol{\xi})_i\\
L_{45} &= T = -\frac{i}{2}[a^+cb^+ - acb] = -\frac{i}{2}[z\partial_z + \bar z\partial_{\bar z} + 2] = -i(\mathbf{r}\times\nabla + 1),\\
L_{46} &= S = \tfrac{1}{2}[a^+cb^+ + acb] = -\tfrac{1}{2}[z\bar z + \partial_z\partial_{\bar z}] = \frac{1}{2}\Big[-|\mathbf{r}|\nabla^2 - |\mathbf{r}| + 2iz\frac{l_0}{\xi^2}(\mathbf{r}\times\nabla)_3 + \frac{|\mathbf{r}|}{\xi^2}l_0^2\Big] \qquad \text{(II.9)}\\
L_{56} &= \Gamma_0 = \tfrac{1}{2}[a^+a + b^+b + 2] = \tfrac{1}{2}[z\bar z - \partial_z\partial_{\bar z}] = \frac{1}{2}\Big[-|\mathbf{r}|\nabla^2 + |\mathbf{r}| + 2iz\frac{l_0}{\xi^2}(\mathbf{r}\times\nabla)_3 + \frac{|\mathbf{r}|}{\xi^2}l_0^2\Big]\\
L_0 &= -\tfrac{1}{2}[a^+a - b^+b] = -\tfrac{1}{2}[z\partial_z - \bar z\partial_{\bar z}] = i\partial_\alpha\\
a &= \begin{pmatrix}a_1\\a_2\end{pmatrix},\quad b = \begin{pmatrix}b_1\\b_2\end{pmatrix},\quad c = i\sigma_2,\quad z = \begin{pmatrix}z_1\\z_2\end{pmatrix},\quad \bar z = \begin{pmatrix}\bar z_1\\\bar z_2\end{pmatrix},\quad x_i = (x,y,z),\quad i = 1,2,3,\quad |\mathbf{r}| = (x^2+y^2+z^2)^{1/2},\\
\sigma_1 &= \begin{pmatrix}0&1\\1&0\end{pmatrix},\quad \sigma_2 = \begin{pmatrix}0&-i\\i&0\end{pmatrix},\quad \sigma_3 = \begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad \partial_z = \begin{pmatrix}\partial_{z_1}\\\partial_{z_2}\end{pmatrix},\quad \partial_{\bar z} = \begin{pmatrix}\partial_{\bar z_1}\\\partial_{\bar z_2}\end{pmatrix},\quad \xi_i = (x,y,0),\\
&[L_0,L_{AB}] = 0.
\end{aligned}
$$

The Bose realization is related to that in terms of z_1,z_2 by

$$
a = \frac{1}{\sqrt{2}}\begin{pmatrix}\bar z_1 + \partial_{z_1}\\ \bar z_2 + \partial_{z_2}\end{pmatrix},\quad b = \frac{1}{\sqrt{2}}\begin{pmatrix}z_2 + \partial_{\bar z_2}\\ -z_1 - \partial_{\bar z_1}\end{pmatrix}
$$

with

$$
z^+ = \bar z,\quad \partial_z^+ = -\partial_{\bar z},
$$

in order to satisfy Bose commutation relations.[9] Similarly, the bijective mapping

$$
-x_i = z\sigma\bar z,\quad i = 1,2,3 \quad -\infty < x_i < \infty,\quad \text{and}\quad \alpha = \arg(z_1) + \arg(z_2),\quad 0\leqslant\alpha < 4\pi
$$

relates the realization in (z_1,z_2) and in Cartesian coordinates. For $l_0 = 0$ the Cartesian realization becomes the usual Schrödinger realization. For $l_0\neq 0$ some operators become singular on the surface $x^2+y^2 = 0$. This singularity and the parameter l_0 are related to the Dirac string and the charge quantization condition, respectively.[10] A realization in terms of spherical polar coordinate (α,r,θ,ϕ) can be obtained from a Cayley–Klein parametrization:

$$z_1 = \sqrt{r}\sin\frac{\theta}{2}\exp\Big(\frac{i}{2}(\alpha+\phi)\Big), \quad z_2 = -\sqrt{r}\cos\frac{\theta}{2}\exp\Big(\frac{i}{2}(\alpha-\phi)\Big).$$

Also, one can obtain a realization in $u_1,...,u_4$, using $z_1 = u_1 + iu_4$, $z_2 = u_3 + iu_2$. Below we give the operators of so(1,2)$_{456}$ in terms of spherical polar coordinates and real variables, $t \equiv 2r$:

$$L_{56} = -t\Big[\frac{1}{t^2}\partial_t(t^2\partial_t) - \frac{1}{t^2}k(k-1) - \tfrac{1}{4}\Big] = \tfrac{1}{2}[u_\mu u_\mu - \tfrac{1}{4}\partial_{u_\mu}\partial_{u_\mu}],$$

$$L_{46} = -t\Big[\frac{1}{t^2}\partial_t(t^2\partial_t) - \frac{1}{t^2}k(k-1) + \tfrac{1}{4}\Big] = -\tfrac{1}{2}[u_\mu u_\mu + \tfrac{1}{4}\partial_{u_\mu}\partial_{u_\mu}],$$

$$L_{45} = -it\Big[\partial_t + \frac{1}{t}\Big] = -\frac{i}{2}[u_\mu\partial_{u_\mu} + 1], \quad L_0 = i\partial_\alpha = \frac{i}{2}[u_1\partial_{u_4} - u_4\partial_{u_1} + u_3\partial_{u_2} - u_2\partial_{u_3}],$$

where

$$-k(k-1) = \partial_\theta^2 + \cot\theta\partial_\theta + \frac{1}{\sin^2\theta}[\partial_\phi^2 + \partial_\alpha^2 + 2\cos\theta\partial_\phi\partial_\alpha] \qquad \text{(II.10)}$$

is the Casimir invariant of so(1,2) which is equal to the Casimir of the commuting so(3) (II.3). We will often consider this SO(1,2) realization in a Hilbert space of square integrable functions $f(t)$ with the scalar product

$$(f\cdot g) = \int_0^\infty dt\, t\, f^*(t)g(t) < \infty.$$

A realization of Ad SO(4,2) may be obtained from (II.9) using the definition of the translation operators $F_e^{(\pm)}$. In Eq. (II.11) we have given two such realizations, which are, respectively, the generalizations of the Bargmann–Wigner–Foldy–Shirokov realization and the Lomont–Moses realization for the Poincaré group. It is known that these two are realted by the Coester transformation[11]:

$$L_{\alpha\beta} = -i(\mathbf{r}\times\nabla)_\gamma + \frac{|\mathbf{r}|}{\xi^2}l_0\xi_\gamma = -i(\mathbf{r}\times\nabla)_\gamma + \eta_\gamma l_0, \quad L_{\alpha 6} = -i|\mathbf{r}|\partial_\alpha - \frac{(\mathbf{r}\times\boldsymbol{\xi})_\alpha}{\xi^2}l_0 = -i|\mathbf{r}|\partial_\alpha - (\boldsymbol{\eta}\times\hat{\eta})_\alpha l_0,$$

$$L_{35} = -i(\mathbf{r}\cdot\nabla + 1) = -i(\mathbf{r}\cdot\nabla + 1), \quad F_\alpha^- = x_\alpha = x_\alpha, \quad F_6^- = |\mathbf{r}| = |\mathbf{r}|, \qquad \text{(II.11)}$$

$$F_\alpha^+ = -x_\alpha\nabla^2 + 2\partial_\alpha(\mathbf{r}\cdot\nabla) + 2il_0\frac{|\mathbf{r}|}{\xi^2}(\boldsymbol{\xi}\times\nabla)_\alpha - (-1)^{\delta_{\alpha 4}}\frac{x_\alpha}{\xi^2}l_0^2$$

$$= -x_\alpha\nabla^2 + 2\partial_\alpha(\mathbf{r}\cdot\nabla) + 2il_0(\boldsymbol{\eta}\times\nabla)_\alpha - 2\hat{\eta}_\alpha\frac{l_0^2}{|\mathbf{r}|+z},$$

$$F_6^+ = -|\mathbf{r}|\nabla^2 + 2il_0\frac{z}{\xi^2}(\mathbf{r}\times\nabla)_4 + \frac{|\mathbf{r}|}{\xi^2}l_0^2 = -|\mathbf{r}|\nabla^2 - 2il_0(\boldsymbol{\eta}\times\nabla)_4 - 2\frac{l_0^2}{|\mathbf{r}|+z},$$

$$\alpha,\beta,\gamma = 1,2,4, \quad x_\alpha = (x,y,z), \quad \eta_\alpha = \Big(\frac{x}{|\mathbf{r}|+z}, \frac{y}{|\mathbf{r}|+z}, 1\Big), \quad \hat{\eta}_\alpha = (0,0,1).$$

In the next section we give the mathematical relation between the various coordinates.

C. Kustaanheimo–Stiefel transformation

The bijective mapping $(x,\alpha)\rightarrow(z_1z_2)\rightarrow(u_\mu)$ is a generalized version of the Kustaanheimo–Stiefel (KS) transformation[12] in $\mathbb{R}^4$. Let $u_1,...,u_n \in \mathbb{R}^n$ be Cartesian coordinates. According to a theorem of Hurwitz, one can construct an $n\times n$ linear orthonormal matrix A with the norm of each row $\|\cdot\| = u_1^2 + \cdots + u_n^2$ for $n = 1,2,4,8$. In $\mathbb{R}^4$,

$$A = \begin{pmatrix} u_1 & -u_2 & -u_3 & u_4 \\ u_2 & u_1 & -u_4 & -u_3 \\ u_3 & u_4 & u_1 & u_2 \\ u_4 & -u_3 & u_2 & -u_1 \end{pmatrix}.$$

The space $\mathbb{R}^4$ is mapped onto $\mathbb{R}^3 \ni (x,y,z)$ by A with the annihilation condition

$$u_4\,du_1 - u_3du_2 + u_2\,du_3 - u_1\,du_4 = 0$$

which corresponds to $l_0 = 0$ in (II.10). That is,

$$\begin{pmatrix} x \\ y \\ z \\ 0 \end{pmatrix} = A\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} u_1^2 - u_2^2 - u_3^2 + u_4^2 \\ 2(u_1u_2 - u_3u_4) \\ 2(u_1u_3 + u_2u_4) \\ 0 \end{pmatrix}.$$

In momentum space $\mathbf{p}=(-p_1,-p_2,-p_3)\in\mathbb{R}^3$ and $\tilde{u}\in\mathbb{R}^4$

$$\begin{pmatrix}-p_1\\-p_2\\-p_3\\0\end{pmatrix}=\frac{1}{2|\mathbf{r}|}A\begin{pmatrix}\tilde{u}_1\\\tilde{u}_2\\\tilde{u}_3\\\tilde{u}_4\end{pmatrix}=\frac{1}{2|\mathbf{r}|}\begin{pmatrix}u_1\tilde{u}_1-u_2\tilde{u}_2-u_3\tilde{u}_3+u_4\tilde{u}_4\\u_2\tilde{u}_1+u_1\tilde{u}_2-u_4\tilde{u}_3-u_3\tilde{u}_4\\u_3\tilde{u}_1+u_4\tilde{u}_2+u_1\tilde{u}_3+u_2\tilde{u}_4\\0\end{pmatrix},$$

where

$$\mathbf{r}^2=R^4\equiv[u_1^2+u_2^2+u_3^2+u_4^2]^2\quad\text{and}\quad 4R^2\mathbf{p}^2=\tilde{u}_1^2+\tilde{u}_2^2+\tilde{u}_3^2+\tilde{u}_4^2\equiv-\nabla_4^2.$$

D. The canonical and pseudocanonical bases

Let T_g be a UR of G on the Hilbert space $\mathscr{H}$. We shall decompose T_g and hence $\mathscr{H}$ into their irreducible components and construct[13] the basis vectors of $\mathscr{H}$.

Let H be a maximal compact [[noncompact]] subgroup of G. We diagonalize the invariant operators in the chain of cosets

$$\frac{G}{H},\quad\frac{H}{\mathrm{SO}(3)},\quad\frac{\mathrm{SO}(3)}{\mathrm{SO}(2)},\quad\mathrm{SO}(2)$$

and construct the common eigenfunctions Φ.

For the most degenerate UIR there is one invariant operator L_0 in G/H with $(L_0-l_0)\Phi^{l_0}(\cdot)=0$ and T_g reduces to $T_g^{l_0}$, and $\mathscr{H}$ to $\mathscr{H}(l_0)$. In $H/\mathrm{SO}(3)$ the invariant operator is L_{56}. The spectrum of L_{56} can be determined from Lemma 3 and is given[6] by the proper eigenvalues $\pm n$

$$(\tfrac{1}{2}L_{\mu\nu}L^{\mu\nu}=l_0^2+n^2-1,\quad n=|l_0|+n_0\geqslant 1,\quad n_0=1,\dots,\infty)$$

$$[[\mathrm{Sp}[L_{46}]=\pm\nu,\quad \mathrm{Sp}[L_{36}]=\pm\rho,\quad -\infty<(\nu,\rho)<+\infty]].$$

Thus the representation $T_g^{l_0}$ decomposes into $T_{g:\{n\}}^{l_0}[[T_{g:\{\nu\}}^{l_0},T_{g:\{\rho\}}^{l_0}]]$, $\mathscr{H}(l_0)$ into $\mathscr{H}(l_0,n)[[\mathscr{H}(l_0,\nu),\mathscr{H}(l_0,\rho)]]$, and the basis functions satisfy the eigenvalue equations

$$(L_{56}-n)\Phi_{\{n\}}^{l_0}(\cdots)=0,\quad [[(L_{46}-\nu)\Phi_{\{\nu\}}^{l_0}(\cdots)=0,\quad (L_{36}-\rho)\Phi_{\{\rho\}}^{l_0}(\cdots)=0]].$$

Further reductions to SO(3) and decompositions of $T_{g:\{n\}}^{l_0}$ and $\mathscr{H}(l_0,n)$ into irreducible components are well known due to the general branching theorem.[8] The basis functions satisfy the eigenvalue equations

$$(L^2-l(l+1))\Phi_{\{n,l\}}^{l_0}(\cdots)=0\quad\text{and}\quad(L_{12}-m)\Phi_{\{n,l,m\}}^{l_0}(\cdots)=0.$$

Thus T_g and $\mathscr{H}$ are decomposed as

$$T_g\supseteq T_g^{l_0}=\sum_{n=1+|l_0|}^{\infty}\sum_{l=|l_0|}^{n-1}\sum_{m=-l}^{+l}T_{g:\{n,l,m\}}^{l_0},\quad \mathscr{H}(l_0)=\sum_{n=1+|l_0|}^{\infty}\oplus\sum_{l=|l_0|}^{n-1}\oplus\sum_{m=-l}^{+l}\oplus\mathscr{H}(l_0,n,l,m).\tag{II.12}$$

The Hilbert space $\mathscr{H}(l_0)$ is spanned by eigenvectors $\chi_{\{n,l,m\}}^{l_0}$ obtained by the Fourier transform of the basis function $\Phi_{\{n,l,m\}}^{l_0}(\cdots)$ with respect to the vector $f(\cdots)\in\mathscr{D}(l_0)$ dense in $\mathscr{H}(l_0)$, i.e.,

$$\chi_{\{n,l,m\}}^{l_0}=\int d\mu(\cdots)\Phi_{\{n,l,m\}}^{l_0}(\cdots)f(\cdots).$$

We define the parity operation such that the Casimir invariant I_3 changes its sign under this operation and this change is incorporated into $l_0=0,\pm 1,\pm 2,\cdots$ or $\pm\frac{1}{2},\pm\frac{3}{2},\cdots$. The UR extended by parity, $T_{gp}^{l_0}$ and the Hilbert space $\mathscr{H}_p(l_0)$ become

$$T_{gp}^{l_0}=T_g^{l_0=|l_0|}+T_g^{l_0=-|l_0|},\quad \mathscr{H}_p(l_0)=\mathscr{H}(l_0=|l_0|)\oplus\mathscr{H}(l_0=-|l_0|).\tag{II.13}$$

Furthermore, considering

$$\mathrm{SO}(4)\sim\mathrm{SO}^{(1)}(3)\otimes\mathrm{SO}^{(1)}(3),\quad\text{with}\quad\tfrac{1}{2}(\mathbf{L}\pm\mathbf{A})\in\mathrm{so}^{\binom{1}{2}}(3),$$

and the reduction to

$$\mathrm{SO}^{(1)}(2)\otimes\mathrm{SO}^{(2)}(2)\quad\text{with}\quad\tfrac{1}{2}(L_3\pm A_2)\in\mathrm{so}^{\binom{1}{2}}(2),$$

one trivially constructs $\Phi_{\{l_1,l_2,m_1,m_2\}}(\cdots)$, where l_1,l_2 are eigenvalues of the Casimir invariant of $\mathrm{SO}^{(1)}(3)$ and $\mathrm{SO}^{(2)}(3)$ and m_1,m_2 are those of $\mathrm{SO}^{(1)}(2)$ and $\mathrm{SO}^{(2)}(2)$. Similarly[14] $\mathrm{SO}(2,2)\sim\mathrm{SO}^{(1)}(1,2)\otimes\mathrm{SO}^{(2)}(1,2)$ with $\frac{1}{2}(L_{56}\pm L_{34})$, $\frac{1}{2}(L_{46}\pm L_{36})$, $\frac{1}{2}(L_{45}\mp L_{36})\in\mathrm{so}^{\binom{1}{2}}(1,2)$ enables one to construct a parabolic basis function $\Phi_{\{n_1,n_2,m\}}^{l_0}(\cdots)$, which is equivalent to that obtianed by diagonalizing L_{56},L_{34},L_{12}.

TABLE I. Summary of the SO(4,2) basis states constructed in Sec. II and the corresponding physical states defined in Sec. III.

Stability subgroup	Diagonal operators	Decomposition of $\mathscr{H}(l_0)$	Basis state	Physical state
	$L_0(l_0),L_{56}(n),I_{SO(3)_{123}}(l),L_{12}(m)$	$\sum_{n=1+\lvert l_0\rvert}^{\infty}\oplus\sum_{l=\lvert l_0\rvert}^{n-1}\oplus\sum_{m=-l}^{l}\oplus\mathscr{H}(l_0,n,l,m)$	$\lvert l_0,nlm\rangle$	$\exp(-i\theta_n T)\Phi$
$SO(4)_{1234}\otimes SO(2)_{56}$	$L_0(l_0),L_{56}(n),I_{SO(3)_{124}}(\lambda),L_{12}(m)$	$\sum_{n=1+\lvert l_0\rvert}^{\infty}\oplus\sum_{\lambda=\lvert l_0\rvert}^{n-1}\oplus\sum_{m=-\lambda}^{+\lambda}\oplus\mathscr{H}(l_0,n,\lambda,m)$	$\lvert l_0;n\lambda m\rangle$	$\exp(-i\theta_n T)\Phi$
	$I_{SO^+(3)}(l_1),I_{SO^-(3)}(l_2),\frac{1}{2}(L_{12}\pm L_{34})\binom{m_1}{m_2}$	$\sum_{2l_1=\lvert l_0\rvert+l_0}^{\infty}\sum_{2l_2=\lvert l_0\rvert-l_0}^{\infty}\oplus\sum_{m_1=-l_1}^{+l_1}\sum_{m_2=-l_2}^{+l_2}\oplus\mathscr{H}(l_1,l_2,m_1,m_2)$	$\lvert l_1l_2m_1m_2\rangle$	$\exp(-i\theta_n T)\Phi$ $n=l_1+l_2+1$
	$L_0(l_0),L_{56}(n),L_{34}(n_1,n_2),L_{12}(m)$	$\sum_{n=1+\lvert l_0\rvert}^{\infty}\oplus\sum_{m}^{n-1}\oplus\sum_{n_1=0}^{n-k_1-k_2}\oplus\mathscr{H}(l_0,n_1,n_2,m)$	$\lvert l_0,n_1n_2m\rangle$	$\exp(-i\theta_n T)\Phi$
$SO(3,1)_{1235}\otimes SO(1,1)_{46}$	$L_0(l_0),L_{46}(\nu),I_{SO(3)_{123}}(l),L_{12}(m)$	$\int_{-\infty}^{+\infty}d\mu(\nu)\oplus\sum_{l=\lvert l_0\rvert}^{\infty}\oplus\sum_{m=-l}^{+l}\oplus\mathscr{H}(l_0,\nu,l,m)$	$\lvert l_0,\nu lm\rangle$	$\exp(-i\theta_\nu T)\Phi$ $\exp(-i\theta_\nu A_3)\Phi$
$SO(3,1)_{1245}\otimes SO(1,1)_{36}$	$L_0(l_0),L_{36}(\rho)I_{SO(3)_{124}}(\lambda),L_{12}(m)$	$\int_{-\infty}^{+\infty}d\mu(\rho)\oplus\sum_{\lambda=\lvert l_0\rvert}^{\infty}\oplus\sum_{m=-\lambda}^{+\lambda}\oplus\mathscr{H}(l_0,\rho,\lambda,m)$	$\lvert l_0\rho\lambda m\rangle$	$\exp(-i\theta_\rho A_3)\Phi$
	$L_0(l_0),F_6(\tau),I_{SO(3)_{124}}(\lambda),L_{12}(m)$	$\int_{-\infty}^{+\infty}d\mu(\tau)\oplus\sum_{\lambda=\lvert l_0\rvert}^{\infty}\oplus\sum_{m=-\lambda}^{+\lambda}\oplus\mathscr{H}(l_0,\tau,\lambda,m)$	$\lvert l_0,\tau\lambda m\rangle$	Φ
$E(3)_{124}\otimes E(1)_6$	$L_0(l_0),F_6(\tau),I_{E(2)_{12}}(\epsilon),L_{12}(m)$	$\int_{-\infty}^{+\infty}d\mu(\tau)\oplus\int_{-\infty}^{+\infty}d\mu(\epsilon)\oplus\sum_{m=\lvert l_0\rvert}^{\infty}\oplus\mathscr{H}(l_0,\tau,\epsilon,m)$	$\lvert l_0;\tau\epsilon m\rangle$	Φ
	$L_0(l_0),F_1(\tau_1),F_2(\tau_2),F_4(\tau_4)$	$\int_{-\infty}^{+\infty}d\mu(\tau_1)\oplus\int_{-\infty}^{+\infty}d\mu(\tau_2)\oplus\int_{-\infty}^{+\infty}d\mu(\tau_4)\oplus\mathscr{H}(l_0,\tau_1,\tau_2,\tau_4)$	$\lvert l_0,\tau_1\tau_2\tau_4\rangle$	Φ

In Table I we have summarized these decompositions and the basis states for discrete spin case that are relevant to physical applications. In the Appendix we have given the action[15] of the SO(4,2) generators on these basis states.

We can realize the diagonalized operators as in (II.9) and solve for the basis functions in spherical polar and parabolic coordinates:

$$\Phi^{l_0}_{\{n,l,m\}}(\alpha r\theta\phi)=\left[\frac{1}{(2l+1)!}\binom{n+l}{2l+1}\right](2r)^l e^{-r}\,{}_1F_1[l-n+1;2l+2;2r]\;Y^{l_0}_{lm}(\alpha\theta\phi),$$

$$(\Phi^{l_0'}_{\{n',l',m'\}},\Phi^{l_0}_{\{nlm\}})=\int_0^{4\pi}d\alpha\int_0^{\infty}t\,dt\int_0^{\pi}\sin\theta\,d\theta\int_0^{2\pi}d\phi\;\Phi^{l_0'*}_{\{n'l'm'\}}(\alpha r\theta\phi)\Phi^{l_0}_{\{nlm\}}(\alpha r\theta\phi)=\delta_{l_0'l_0}\delta_{n'n}\delta_{l'l}\delta_{m'm},\quad t=2r,$$

$$\Phi^{l_0}_{\{n\lambda m\}}(\alpha r\theta\phi)=\Phi^{l_0}_{\{nlm\}}(\alpha r\theta\phi)\,\big|_{l\to\lambda},$$

$$\Phi^{l_0}_{\{\nu lm\}}(\alpha r\theta\phi)=\frac{1}{\sqrt{2\pi}}\frac{e^{\pi\nu/2}}{(2l+1)!}\,\lvert(l+i\nu)!\rvert\,(2r)^l e^{-ir}\,{}_1F_1[l+i\nu+1;2l+2;2ir]\;Y^{l_0}_{lm}(\alpha\theta\phi)=\Phi^{l_0}_{\{nlm\}}(\alpha r\theta\phi)\Big|_{\substack{n\to\mp i\nu\\ r\to\pm ir}},$$

$$(\Phi^{l_0'}_{\{\nu'l'm'\}},\Phi^{l_0}_{\{\nu lm\}})=\delta_{l_0'l_0}\delta(\nu'-\nu)\delta_{l'l}\delta_{m'm},\quad \Phi^{l_0}_{\{\rho\lambda m\}}(\alpha r\theta\phi)=\Phi^{l_0}_{\{\nu lm\}}(\alpha r\theta\phi)\Big|_{\substack{\nu\to\rho\\ l\to\lambda}}=\Phi^{l_0}_{\{nlm\}}(\alpha r\theta\phi)\Big|_{\substack{n\to\mp i\rho\\ n\to\pm ir\\ l\to\lambda}},$$

$$\Phi^{l_0}_{\{mn_1n_2\}}(\alpha\xi\eta\phi)=\frac{1}{\sqrt{4\pi}}e^{-i\alpha l_0}\frac{1}{\sqrt{2\pi}}e^{im\phi}\frac{1}{(2k_1-1)!(2k_2-1)!}\frac{(n_1+2k_1-1)!(n_2+2k_2-1)!}{n_1!n_2!}\xi^{k_1-1/2}$$
$$\times\eta^{k_2-1/2}(\xi\eta)^{-1/2}e^{-(\xi+\eta)/2}\,{}_1F_1[-n_1;2k_1;\xi]\,{}_1F_1[-n_2;2k_2;\eta],\quad k_{\binom{1}{2}}=\tfrac{1}{2}[\,\lvert m(\pm)l_0\rvert+1],$$

$$(\Phi^{l_0'}_{\{m'n_1'n_2'\}},\Phi^{l_0}_{\{mn_1n_2\}})=\int_0^{4\pi}d\alpha\int_0^{\infty}\xi\,d\xi\int_0^{\infty}\eta\,d\eta\int_0^{2\pi}d\phi\;\Phi^{l_0'*}_{\{m'n_1'n_2'\}}(\alpha\xi\eta\phi)\Phi^{l_0}_{\{mn_1n_2\}}(\alpha\xi\eta\phi)=\delta_{l_0'l_0}\delta_{m'm}\delta_{n_1'n_1}\delta_{n_2'n_2},$$

$$Y^{l_0}_{lm}(\alpha\theta\phi)=\frac{1}{4\pi}e^{-i\alpha l_0}e^{im\phi}\sqrt{2l+1}\,P^l_{ml_0}(\cos\theta),$$

where[16]

$$P^l_{ml_0}(z)=(-1)^{l-l_0}2^{-l}\left(\frac{(l+m)!}{(l-l_0)!(l+l_0)!(l-m)!}\right)^{1/2}(1+z)^{-(m+l_0)/2}(1-z)^{-(m-l_0)/2}$$
$$\times\frac{d^{\,l-m}}{dz^{l-m}}[(1-z)^{l-l_0}(1+z)^{l+l_0}].\tag{II.14}$$

All the basis states in Table I are expanded in terms of the canonical basis states:

$$\lvert l_1l_2m_1m_2\rangle=\sum_{l=\lvert l_0\rvert}^{n-1}\langle l_0;nlm\mid l_1l_2m_1m_2\rangle\,\lvert l_0;nlm\rangle;\quad m=m_1+m_2,\quad n=l_1+l_2+1,\quad l_0=l_1-l_2,$$
$$\lvert l_0;mn_1n_2\rangle=\sum_{l=\lvert l_0\rvert}^{n-1}\langle l_0;nlm\mid l_0;mn_1n_2\rangle\mid l_0;nlm\rangle;\quad n=n_1+n_2+k_1+k_2,\quad k_{\binom{1}{2}}=\tfrac{1}{2}[\,\lvert m(\pm)l_0\rvert+1],\tag{II.15}$$
$$\lvert l_0;n\lambda m\rangle=\sum_{l=\lvert l_0\rvert}^{n-1}\langle l_0;nlm\mid l_0;n\lambda m\rangle\mid l_0;nlm\rangle,\quad \lvert l_0;\nu lm\rangle=\sum_{n=1+\lvert l_0\rvert}^{\infty}\langle l_0;nlm\mid l_0;\nu lm\rangle\mid l_0;nlm\rangle,$$

$$|l_0;\rho\lambda m\rangle = \sum_{n=1+|l_0|}^{\infty} \sum_{l=|l_0|}^{n-1} \langle l_0;n\lambda m \mid l_0;\lambda lm\rangle \langle l_0;nlm \mid l_0;n\lambda m\rangle \mid l_0;nlm\rangle.$$

The expansion coefficients $\langle l_0;nlm \mid l_1l_2m_1m_2\rangle$ and $\langle l_0;nlm \mid l_0;n_1n_2m\rangle$ are given by SO(3) C–G coefficients. The coefficients $\langle l_0;nlm \mid l_0;\nu lm\rangle$ and $\langle l_0;n\lambda m \mid l_0;\rho\lambda m\rangle$ are computed using the SO(1,2) realization (II.10), where we take the scalar product of the solutions of the corresponding operator equations. The results turn out to be Bargmann functions.[17]

The expansion coefficient $\langle l_0,nlm \mid l_0;n\lambda m\rangle$ is known to be a SO(4) rotation function. Thus,

$$\begin{gathered}\langle l_0,nlm \mid l_1l_2m_1m_2\rangle = \begin{bmatrix} l_1 & l_2 & l \\ m_1 & m_2 & m\end{bmatrix},\\ \langle l_0,nlm \mid l_0;mn_1n_2\rangle = \begin{bmatrix} \frac{1}{2}(n-1)+\frac{1}{2}(k_2-k_1) & \frac{1}{2}(n-1)+\frac{1}{2}(k_1-k_2) & l \\ \frac{1}{2}(n_2-n_1+2k_2-1) & \frac{1}{2}(n_1-n_2+2k_1-1) & k_1+k_2-1\end{bmatrix},\\ \langle l_0,nlm \mid l_0;n\lambda m\rangle = D_{lm\lambda}^{[n-1,l_0]}(-\pi/2),\quad \langle l_0,nLm \mid l_0;\mu Lm\rangle = V_{n\mu}^{L+1}\left(i\frac{\pi}{2}\right);\quad L=\{l,\lambda\},\quad \mu=\{\nu,\rho\},\end{gathered}\tag{II.16}$$

where

$$\begin{aligned} D_{lm\lambda}^{[n-1,l_0]}(\theta) &= \sum_{m_1m_2} \begin{bmatrix} \frac{1}{2}(n-1+l_0) & \frac{1}{2}(n-1-l_0) & l \\ m_1 & m_2 & m\end{bmatrix}\begin{bmatrix} \frac{1}{2}(n-1+l_0) & \frac{1}{2}(n-1+l_0) & \lambda \\ m_1 & m_2 & m\end{bmatrix} e^{-i\theta(m_1-m_2)},\\ V_{n\mu}^{L+1}(\theta) &= \frac{e^{i\pi/2}}{\sqrt{2}}\frac{e^{-i\pi(i\mu)/2}}{[\sin(\pi i\mu)]^{1/2}}\left[\binom{n+L}{n-L-1}\binom{i\mu+L}{i\mu-L-1}\right]^{1/2}(\tanh\theta/2)^{n+i\mu}(\sinh\theta/2)^{-2(L+1)}\\ &\quad\times {}_2F_1[L+1-i\mu,L+1-n;2L+2;-(\sinh\theta/2)^{-2}].\end{aligned}\tag{II.17}$$

E. The Euclidean basis

We reduce $\tilde G=\mathrm{Ad}\,\mathrm{SO}(4,2)$ with respect to its Euclidean and orthogonal subgroups as in the previous section. For the maximal noncompact Euclidean subgroup H, $H=\mathrm{E}(3)\otimes\mathrm{E}(1)$;

$$e(3)_{124}\ni F_1,F_2,F_4,L_{12},L_{24},L_{41};\quad e(1)\ni F_6.$$

The eigenvalue equation of the invariant operator in $\tilde G/H$, in the most degenerate UIR is given by $(L_0-l_0)\Phi^{l_0}(\cdot)=0$. In H/h where $h=\mathrm{SO}(3)_{124}$ or $\mathrm{E}(2)_{12}\otimes\mathrm{E}(1)_4$ or $\mathrm{E}(2)_{14}\otimes\mathrm{E}(1)_2$, the invariant operator is F_6 and by Lemma 4 the $\mathrm{Sp}(F_6)=\tau$, since $F_1^2+F_2^2+F_4^2=F_6^2=\tau^2$, $0\leqslant\tau^2<\infty$. The basis function satisfies $(F_6-\tau)\Phi^{l_0}_{\{\tau\}}(\cdots)=0$. Further reductions of h lead to the eigenvalue equations:

$$\begin{gathered}[L_{12}^2+L_{24}^2+L_{41}^2-\lambda(\lambda+1)]\Phi^{l_0}_{\{\tau\lambda\}}(\cdots)=0;\quad |l_0|\leqslant\lambda<\infty,\quad [L_{12}-m]\Phi^{l_0}_{\{\tau\lambda m\}}(\cdots)=0;\quad -\lambda\leqslant m\leqslant+\lambda,\\ [F_1^2+F_2^2-\epsilon^2]\Phi^{l_0}_{\{\tau\epsilon\}}(\cdots)=0;\quad 0\leqslant\epsilon^2<\infty,\quad [L_{12}-m]\Phi^{l_0}_{\{\tau\epsilon m\}}(\cdots)=0;\quad m=l_0,l_0\pm1,\cdots,\\ [F_1-\mu]\Phi^{l_0}_{\{\tau\epsilon\mu\}}(\cdots)=0;\quad 0\leqslant\mu<\infty,\quad [F_i-\tau_i]\Phi^{l_0}_{\{\tau_i\}}(\cdots)=0;\quad i=1,2,4,\quad 0\leqslant\tau_i<\infty,\quad \tau^2=\boldsymbol\tau^2=\tau_1^2+\tau_2^2+\tau_4^2.\end{gathered}$$

In order to obtain explicit expressions[18] for the eigenfunctions, we solve the eigenvalue equations in different coordinate systems. We find,

$$\begin{aligned} \Phi^{l_0}_{\{\tau\lambda m\}}(\alpha r\theta\phi) &= \frac{1}{2\sqrt{r}}J_{2\lambda+1}(2\sqrt{\tau_+ r})\,Y^{l_0}_{\lambda m}(\alpha,\theta,\phi),\quad \tau_\pm=\mathrm{Sp}(F_6^\pm),\\ \Phi^{l_0}_{\{\tau\lambda m\}}(\alpha r\theta\phi) &= \frac{1}{2\sqrt{\tau_-}}\delta(r-\tau_-)\,Y^{l_0}_{\lambda m}(\alpha\theta\phi),\\ \left(\Phi^{l_0'}_{\{\tau'\lambda'm'\}},\Phi^{l_0}_{\{\tau\lambda m\}}\right) &= \int_0^{4\pi}d\alpha\int_0^\infty t\,dt\int_0^\pi\sin\theta\,d\theta\int_0^{2\pi}d\phi\;\Phi^{l_0'*}_{\{\tau'\lambda'm'\}}(\alpha r\theta\phi)\Phi^{l_0}_{\{\tau\lambda m\}}(\alpha r\theta\phi)\\ &= \delta_{l_0'l_0}\delta(\tau'-\tau)\delta_{\lambda'\lambda}\delta_{m'm};\quad t=2r,\\ \Phi^{l_0}_{\{\tau_-\epsilon_-m\}}(\alpha\rho\phi z) &= \frac{e^{-i\alpha l_0}}{\sqrt{4\pi}}\frac{e^{im\phi}}{\sqrt{2\pi}}\frac{e^{ikz}}{\sqrt{2\pi}}J_{m-l_0}(\epsilon_-\rho);\quad k^2=\tau_-^2-\epsilon_-^2,\\ \left(\Phi^{l_0'}_{\{\tau'_-\epsilon'_-m'\}},\Phi^{l_0}_{\{\tau_-\epsilon_-m\}}\right) &= \int_0^{4\pi}d\alpha\int_0^\infty\rho\,d\rho\int_{-\infty}^{+\infty}dz\int_0^{2\pi}d\phi\;\Phi^{l_0'*}_{\{\tau'_-\epsilon'_-m\}}(\alpha\rho\phi z)\Phi^{l_0}_{\{\tau_-\epsilon_-m\}}(\alpha\rho\phi z)\\ &= \delta_{l_0'l_0}\delta(\epsilon'^2_- - \epsilon^2_-)\delta(k'-k)\delta_{m'm},\\ \Phi^{l_0}_{\{\tau_1\tau_2\tau_4\}}(\alpha xyz) &= \frac{e^{-i\alpha l_0}}{\sqrt{4\pi}}\frac{1}{(2\pi)^{3/2}}e^{i\boldsymbol\tau\mathbf{x}};\quad \boldsymbol\tau=(\tau_1,\tau_2,\tau_4);\quad \mathbf{x}=(x,y,z),\\ \left(\Phi^{l_0'}_{\{\tau_1'\tau_2'\tau_4'\}},\Phi^{l_0}_{\{\tau_1\tau_2\tau_4\}}\right) &= \int_0^{4\pi}d\alpha\int d\mathbf{x}\,\Phi^{l_0'*}_{\{\tau_1'\tau_2'\tau_4'\}}(\alpha\mathbf{x})\Phi^{l_0}_{\{\tau_1\tau_2\tau_4\}}(\alpha\mathbf{x})=\delta_{l_0'l_0}\delta(\boldsymbol\tau'-\boldsymbol\tau).\end{aligned}\tag{II.18}$$

Next we expand the above Euclidean bases in terms of the canonical states as follows:

$$\langle l_0,\tau\lambda m\mid = \sum_{n=1+|l_0|}^{\infty}\sum_{l=|l_0|}^{n-1}\langle l_0,n\lambda m\mid l_0,\tau\lambda m\rangle\langle l_0,nlm\mid l_0,n\lambda m\rangle\,|l_0,nlm\rangle,$$

$$\langle l_0,\tau\epsilon m\mid = \sum_{n=1+|l_0|}^{\infty}\sum_{\lambda=|l_0|}^{n-1}\sum_{l=|l_0|}^{n-1}\langle l_0;\tau\lambda m\mid l_0;\tau\epsilon m\rangle\langle l_0;n\lambda m\mid l_0;\tau\lambda m\rangle\langle l_0;nlm\mid l_0;n\lambda m\rangle\mid l_0;nlm\rangle,$$

$$\mid l_0;\tau_1\tau_2\tau_4\rangle = \sum_{n=1+|l_0|}^{\infty}\sum_{l=|l_0|}^{n-1}\sum_{m=-l}^{+l}\langle l_0;nlm\mid l_0,\tau_1\tau_2\tau_4\rangle\mid l_0;nlm\rangle. \tag{II.19}$$

The coefficient $\langle l_0;nlm\mid l_0;\tau_\pm\lambda m\rangle$ can be calculated by finding the solutions from (II.10) and computing the scalar product. It turns out[19] to be proportional to the radial part of the basis function $\Phi^{l_0}_{\{nlm\}}(ar\theta\phi)$ (II.14) with $r\to\tau_\pm$.

To compute the coefficient $\langle l_0;\tau\lambda m\mid l_0;\tau_-\epsilon_- m\rangle$ we express the operators of $e(3)$ in spherical polar coordinates[20] and take the scalar product of the solution of the invariant operator eigenvalue equation of SO(3) and E(2). The final expression turns out to be a Jacobi polynomial. To calculate $\langle l_0,nlm\mid l_0\tau_1\tau_2\tau_4\rangle$, we substitute $\Phi^{l_0}_{\{\tau_1\tau_2\tau_4\}}(\alpha xyz)$ and $\Phi^{l_0}_{\{nlm\}}(ar\theta\phi)$ and invert (II.19) by using the orthogonality properties of spherical harmonics and Laguerre polynomials. Then using Weisner's formula[21] we express the result in terms of Gegenbauer polynomials. Thus,

$$\langle l_0;nlm\mid l_0;\tau_\pm\lambda m\rangle=\sqrt{2}U_{NM}(\sqrt{2\tau_\pm}\,),\quad N=n+\lambda,\quad M=n-\lambda-1,\quad U_{NM}(\alpha)=U^*_{MN}(-\alpha),$$

$$\int_0^\infty\frac{d\alpha^2}{\pi}\,U^*_{N'M'}(\alpha)U_{NM}(\alpha)=\delta_{N'N}\delta_{M'M},\quad U_{NM}(\alpha)\equiv\left(\frac{M!}{N!}\right)^{1/2}\alpha^{N-M}e^{-\alpha\bar\alpha/2}L_M^{N-M}(\alpha\bar\alpha);$$

$$\langle l_0,\tau\lambda m\mid l_0;\tau_-\epsilon_- m\rangle=P^\lambda_{ml_0}\left(\frac{k}{\tau_-}\right),\quad k^2=\tau_-^2-\epsilon_-^2, \tag{II.20}$$

$$\langle l_0;nlm\mid l_0;\tau_1\tau_2\tau_4\rangle=\frac{1}{2\pi}\left(\frac{2}{\pi}\right)^{1/2}(2l+1)l!\left(\frac{(n-l-1)!}{(n+l)!}\right)^{1/2}(1+\cos\theta)(2i\sin\theta)^lC^{l+1}_{n-l-1}(-\cos\theta)\,Y^*_{lm}(\theta_\tau,\phi_\tau),$$

$$\cos\theta=(1-\tau^2)/(1+\tau^2),\quad \boldsymbol{\tau}=(\tau,\theta_\tau,\phi_\tau).$$

III. PHYSICAL REALIZATIONS OF THE DYNAMICAL GROUP

A. The wave equation

We shall obtain and classify the physical solutions of a class of wave equations using the basis states and their properties studied earlier (Sec. II).

A general wave equation can be constructed on the carrier space of SO(4,2)$\otimes T_4$ [22] as

$$[\alpha_1\Gamma_\mu p^\mu+\Lambda_s]\Psi(p)=0,\quad \Lambda_s=(\alpha_2+\alpha_3S)\,p_\mu p^\mu+(\gamma+\beta S). \tag{III.1}$$

The parameters $\alpha_1,\alpha_2,\alpha_3,\beta,\gamma$ are assumed to be real constants and they ensure equal dimension of each term in (III.1). In the case of the relativistic hydrogen atom $\alpha_1=1$, $\alpha_2=-\alpha/2m_p$, $\alpha_3=1/2m_p$, $\beta=\alpha_3(m_p^2-m_e^2)$, $\gamma=-\alpha_2(m_p^2+m_e^2)$.

We define an arbitrary Lorentz-boost

$$U(L_p)=\exp(-i\phi L_{12})\exp(-i\theta L_{31})\exp(i\phi L_{12})\exp(i\xi L_{35}), \tag{III.2}$$

$\xi\geqslant 0$, $0\leqslant\theta\leqslant\pi$, $-\pi\leqslant\phi\leqslant\pi$ such that $\Psi(p)=U(L_p)\Psi_R$, where Ψ_R are the physical states with certain restricted momenta– $|\mathbf{p}|=0$ (timelike $p_\mu p^\mu=m^2=M^2>0$, $p_\mu p^\mu=\Omega^2-\Omega^2=0$ or $|velocity|\equiv|\mathbf{v}|=1$ (lightlike) and $|\mathbf{v}|=\infty$ (spacelike, $p_\mu p^\mu=m^2=-\kappa^2\leqslant 0$).

Under (III.2), p_μ can be parametrized as

$$p^\mu=[p^0,p\cos\phi\sin\theta,p\sin\phi\sin\theta,p\cos\theta\,],\quad p^0=\theta(m^2)M\cosh\xi+\theta(-m^2)\kappa\sinh\xi+\delta_{m^20}\Omega e^\xi,$$
$$p=\theta(m^2)M\sinh\xi+\theta(-m^2)\kappa\cosh\xi+\delta_{m^20}\Omega e^\xi. \tag{III.3}$$

The rapidity ξ is defined by $\hat\xi\,(\tanh\xi)^{m^2/|m^2|}=\mathbf{p}/p^0$.

In terms of the restricted physical state Ψ_R (III.1) becomes

$$[\alpha_1\Gamma_0(p_0\cosh\xi-p\sinh\xi)+\alpha_1\Gamma_3(p_0\sinh\xi-p\cosh\xi)+(\alpha_3m^2+\beta)S+(\alpha_2m^2+\gamma)]\Psi_R=0 \tag{III.4}$$

In the next section we diagonalize this equation for different momentum orbits.

B. Solutions of class I: Timelike particles

Here with $p_0=M\cosh\xi$, $p=M\sinh\xi$, $M>0$, (III.4) becomes

$$[\alpha_1M\Gamma_0+(\alpha_3M^2+\beta)S+(\alpha_2M^2+\gamma)]\Psi(0)=0. \tag{III.5}$$

Since Γ_0 and S cannot be simultaneously diagonalized, we define the physical state

$$\Psi(0) = \frac{1}{N}\exp\{\theta\,[S,\Gamma_0]\}\Phi(0) = \frac{1}{N}\exp[-i\theta T]\Phi(0) \tag{III.6}$$

and rotate out either Γ_0 or S by appropriately choosing θ. $\Phi(0)$ are the basis states as given in Table I. Because of Eq. (III.6) the states $\Psi(0)$ have been called "tilted states" in the literature. If we choose $\theta = \theta_n$ such that

$$\begin{aligned} 1(\alpha_2 M^2 + \gamma)\cosh\theta_n &= -\alpha_1 M\Gamma_0,\\ 1(\alpha_2 M^2 + \gamma)\sinh\theta_n &= -(\alpha_3 M^2 + \beta)\Gamma_0, \end{aligned} \tag{III.7}$$

with $\Phi(0) = \Phi^{l_0}_{\{nlm\}}(\cdots)$, then the diagonalization of (III.4) gives the mass spectrum

$$n^2 = (\alpha_2 M^2 + \gamma)^2/[\alpha_1^2 M^2 - (\alpha_3 M^2 + \beta)^2]. \tag{III.8}$$

For the hydrogen atom, (III.7) gives

$$\theta_n \sim \ln(n)\ \left\{\text{using } \operatorname{arctanh}x = \ln\left(\frac{1+x}{1-x}\right)^{1/2}\right\}.$$

Using spherical polar coordinates, we obtain the physical state (III.6) as

$$\Psi_n(0) = \left[\frac{e^{-\theta_n}}{n}\frac{4}{(2l+1)!}\binom{n+l}{2l+1}\right]^{1/2} e^{-\theta_n}(2re^{-\theta_n})^l \exp[-(re^{-\theta_n})]\,{}_1F_1[-n+l+1;2l+2;2re^{-\theta_n}]\,Y^{l_0}_{lm}(\alpha\theta\phi), \tag{III.9}$$

where we have used (II.14) and the Bargmann function, $\langle n' \mid \exp(-i\theta_n T) \mid n\rangle = V^{l+1}_{n'n}(\theta_n)$, and Weisner's formula[21]

$$\begin{aligned}\sum_{n=0}^{\infty} L^{\gamma-1}_n(\omega)\,{}_2F_1[-n,\beta;\gamma;x]\,y^n = (1-y)^{\beta-\gamma}[1+(x-1)y]^{-\beta}\exp\left[-\frac{wy}{1-y}\right]{}_1F_1[\beta;\gamma;wxy/\{(1-y)\\ \times[1+(x-1)y]\}], \qquad |y| < \min(1,|1-x|^{-1}).\end{aligned}$$

For $\theta_n \sim \ln(n)$ and $l_0 = 0$ (III.9) reduces to the hydrogenic solution of the Schrödinger equation. Similarly the parabolic solution becomes after appropriate normalization:

$$\begin{aligned}\Psi_{n_1 n_2}(0) = {}&\left[\frac{2}{n^2}e^{-\theta_n}e^{\theta_n}\frac{1}{(2k_1-1)!}\frac{1}{(2k_2-1)!}\binom{n_1+2k_2-1}{2k_1-1}\binom{n_2+2k_2-1}{2k_2-1}\right]^{1/2}(\xi e^{-\theta_n})^{k_1-1/2}(\eta e^{-\theta_n})^{k_2-1}\\ &e^{-(\xi e^{-\theta_n})/2}e^{-(\eta e^{-\theta_n})/2}\,{}_1F_1[-n;2k_1;\xi e^{-\theta_n}]\,{}_1F_1[-n_2;2k_2;\eta e^{-\theta_n}]\frac{\sqrt{2}}{4\pi}\exp(-i\alpha l_0)\exp(im\phi),\end{aligned}$$

where we have used[23]

$$\langle mn_1' n_2' | e^{-i\theta_n T} | mn_1 n_2\rangle = V^{k_1}_{n_1'+k_1,n_1+k_1}(\theta_n)V^{k_2}_{n_2'+k_2,n_2+k_2}(-\theta_n), \quad k_{\binom{1}{2}} = \tfrac{1}{2}[|m(\pm)l_0| + 1].$$

The matrix element of the Lorentz–boost in the third direction between two physical states is given by[24]

$$\begin{aligned} N_{n'}N_n G_{n'n} &= \Psi_{n'}(0)e^{i\xi L_{35}}\Psi_n(0) = \langle l_0;n'l'm' \mid e^{i\theta_{n'}T}e^{i\xi L_{35}}e^{-i\theta_n T} \mid l_0;nlm\rangle\\ &= \langle l_0;n'l'm' \mid e^{-i\alpha L_{34}}e^{-i\beta L_{45}}e^{-i\gamma L_{34}} \mid l_0;nlm\rangle = \langle l_0;n'l'm' \mid e^{-i\alpha_1 L_{45}} - e^{-i\beta_1 L_{34}}e^{-\gamma_1 L_{45}} \mid l_0,nlm\rangle\\ &= \delta_{m'm}\sum_{L=|l_0|}^{\operatorname{Min}(n'-1,n-1)} D^{[n'-1,l_0]}_{l'mL}(\alpha)V^{L+1}_{n'n}(\beta)D^{[n-1,l_0]}_{Lml}(\gamma) = \delta_{m'm}\sum_{N=1+|l_0|}^{\infty} V^{l'+1}_{n'N}(\alpha_1)D^{[N-1,l_0]}_{l'ml}(\beta_1)V^{l+1}_{Nn}(\gamma_1),\end{aligned}$$

where we have made use of the 2×2 spinor representation of SO(2,1) with $L_{34} = \frac{1}{2}\sigma_3$, $L_{35} = (i/2)\sigma_2$, $L_{45} = -(i/2)\sigma_1$ and obtain:

$$\begin{aligned}
&2\cosh^2\frac{\beta}{2} = \cosh\theta_n\cosh\theta_{n'}\cosh\xi - \sinh\theta_n\sinh\theta_{n'} + 1,\\
&2\sinh^2\frac{\beta}{2} = \cosh\theta_n\cosh\theta_{n'}\cosh\xi - \sinh\theta_n\sinh\theta_{n'} - 1,\\
&\sin\alpha\sinh\beta = \cosh\theta_n\sinh\xi, \quad \sin\gamma\sinh\beta = -\cosh\theta_{n'}\sinh\xi,\\
&\cos\alpha\sinh\beta = \sinh\theta_n\cosh\theta_{n'} - \cosh\theta_n\sinh\theta_{n'}\cosh\xi,\\
&\cos\gamma\sinh\beta = \sinh\theta_n\cosh\theta_{n'}\cosh\xi - \cosh\theta_n\sinh\theta_{n'},\\
&\cos\left(\frac{\beta_1}{2}\right)\cosh\left(\frac{\alpha_1+\gamma_1}{2}\right) = \cosh\frac{\beta}{2}\cos\left(\frac{\alpha+\gamma}{2}\right),\\
&\cos\left(\frac{\beta_1}{2}\right)\sinh\left(\frac{\alpha_1+\gamma_1}{2}\right) = \sinh\frac{\beta}{2}\cos\left(\frac{\alpha-\gamma}{2}\right),\\
&\sin\left(\frac{\beta_1}{2}\right)\sinh\left(\frac{\alpha_1-\gamma_1}{2}\right) = -\sinh\frac{\beta}{2}\sin\left(\frac{\alpha-\gamma}{2}\right),
\end{aligned}$$

$$\sin\left(\frac{\beta_1}{2}\right)\cosh\left(\frac{\alpha_1-\gamma_1}{2}\right)=\cosh\frac{\beta}{2}\sin\left(\frac{\alpha+\gamma}{2}\right). \tag{III.11}$$

Next, if we choose $\theta=\theta_\nu$ such that

$$1(\alpha_1M^2+\gamma)\cosh\theta_\nu=-(\alpha_3M^2+\beta)S,\quad 1(\alpha_2M^2+\gamma)\sinh\theta_\nu=-\alpha_1MS, \tag{III.12}$$

with $\Phi(0)=\Phi^{l_0}_{\{\nu lm\}}$, then the diagonalization of (III.4) gives the mass spectrum

$$-\nu^2=(\alpha_2M^2+\gamma)^2\mid[\alpha_1^2M^2-(\alpha_3M^2+\beta)^2]. \tag{III.13}$$

And as before we obtain in terms of spherical polar coordinates the physical state

$$\Psi_\nu(0)=\frac{1}{\sqrt{2\pi}}\frac{e^{\pi\nu/2}}{(2l+1)!}\,|(l+i\nu)!|\,2e^{-\theta_\nu}(2re^{-\theta_\nu})^l\exp(-ire^{-\theta_\nu})\,{}_1F_1[l+1+i\nu;2l+2;2ir\exp(-\theta_\nu)]\,Y^{l_0}_{lm}(\alpha\theta\phi). \tag{III.14}$$

For $\theta_\nu\sim\ln\nu$ and $l_0=0$ (III.14) reduces to the Coulomb solution of the Schrödinger equation. Furthermore, $\Psi_\nu(0)$ can be expanded in terms of canonical basis states as

$$\Psi_n(0)=\frac{1}{N_\nu}\sum_{n=1+|l_0|}^{\infty}V^{l+1}_{n\nu}\left(\theta_\nu+i\frac{\pi}{2}\right)\Phi^{l_0}_{\{nlm\}}. \tag{III.15}$$

In order to obtain an internal dynamical picture, we first express (III.4) in terms of spherical polar coordinates:

$$\frac{1}{2\mu}\left[\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}+\frac{v_1}{r}+\frac{v_2}{r^2}+v_3\right]\Phi^{l_0}_{\{Nlm\}}(\alpha' r\theta\phi)=0, \tag{III.16}$$

where

$$N=n\text{ or }\nu,\quad r_0v_1=-2\alpha=2n\text{ or }-2\nu,\quad v_2=-l(l+1),$$

and $r_0^2v_3=-\alpha^2/n^2$ or α^2/ν^2 and $2\mu E=v_3$, the constant r_0 has the dimension of length. This equation describes a generalized three-dimensional internal Kepler motion with arbitrary lowest spin l_0 and total energy $\leqslant 0$.

On the other hand, in terms of four real variables $(u_1u_2u_3u_4)$, (III.4) becomes

$$\left(-\frac{1}{2\mu}\Delta^2_4\pm\tfrac{1}{2}\mu\omega^2u^2-E_N\right)\Phi^{l_0}_{\{Nlm\}}(u_1u_2u_3u_4)=0, \tag{III.17}$$

where μ is introduced as a mass parameter, $\omega=2/\mu u_0^2$ and $E_N=2\omega n$ or $2\omega\nu$, the constant u_0 has dimension of length. Equation (III.17) describes an internal four-dimensional harmonic oscillator with arbitrary spin l_0. For $l_0=0$, the K–S transformation (Sec. II.3) connects (III.16) and (III.17).

C. Solutions of class II: Lightlike states

In this case, $p_0=p=\Omega e^\xi$, $-\infty<\xi<\infty$ and (III.4) becomes

$$[\alpha_1\Omega F_6^\pm+\beta S+\gamma]\Psi(|\mathbf{v}|=1)=0. \tag{III.18}$$

Here $F_6^\pm$ and S cannot be simultaneously diagonalized and, furthermore, the only way to eliminate S is to take $\beta=0$. So the physical states and the basis states (II.18) are identical, e.g., $\Psi(|\mathbf{v}|=1)=\Phi^{l_0}_{\{\tau\lambda m\}}(\cdots)$.

In terms of spherical polar coordinates, (III.18) become

$$\frac{1}{2\mu}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}+\frac{v_1}{r}+\frac{v_2}{r^2}+v_3\right)\Psi^{l_0}_{\{\tau\lambda m\}}(\alpha' r\theta\phi)=0,$$
$$\left(\frac{r}{r_0}-\tau_-\right)\Psi^{l_0}_{\{\tau\lambda m\}}(\alpha' r\theta\phi)=0, \tag{III.19}$$

where

$$r_0v_1=\tau_+,\quad v_2=-\lambda(\lambda+1),\quad v_3=0=2E.$$

The first equation describes an internal three-dimensional Kepler motion with zero total energy. In terms of the four real variables $u_1,\ldots,u_4$, (III.18) become

$$\left(-\frac{1}{2\mu}\Delta^2_4-\omega\tau_+\right)\Psi^{l_0}_{\{\tau\lambda m\}}(u_1u_2u_3u_4)=0,$$
$$\left(\frac{u^2}{u_0^2}-\tau_-\right)\Psi^{l_0}_{\{\tau\lambda m\}}(u_1u_2u_3u_4)=0, \tag{III.20}$$

where μ is a mass parameter and $\omega=2/\mu u_0^2$ and the equation describes a free motion in four-space with kinetic energy $\omega\tau_+$. The second equations in (III.19) and (III.20) describe the position on a three-sphere and a 4-sphere, respectively. It is easy to see that the two internal motions (III.19)–(III.20) are connected by the K–S transformation.

D. Solutions of class III. Spacelike states

In this case, $p_0=\kappa\sinh\xi$, $p=\kappa\cosh\xi$ and (III.14) becomes

$$[\alpha_1\kappa\Gamma_3-(-\alpha_3\kappa^2+\beta)S-(-\alpha_2\kappa^2+\gamma)]\Psi(|\mathbf{v}|=\infty)=0. \tag{III.21}$$

We define the physical states

$$\Psi(|\mathbf{v}|=\infty)=\frac{1}{N}\exp\{\theta[S,\Gamma_3]\}\Phi(|\mathbf{v}|=\infty)$$
$$=\frac{1}{N}\exp(i\theta L_{34})\Phi(|\mathbf{v}|=\infty) \tag{III.22}$$

TABLE II. Comparison of the three classes of solutions of the infinite component wave equation (III.4) in two different realizations.

	Mass: M	Kepler motion $\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}+\frac{v_1}{r}+\frac{v_2}{r^2}+v_3\right)\Phi=0$			Oscillatory motion (kinetic energy + potential energy $-E)\Phi=0$		
		$r_0v_1=-2\alpha$	v_2	$v_3=2E\mu$	K.E.	P.E.	E
Class I $\|\mathbf{v}\|=0$	discrete $M^2>0$	$2n$ $n=\|l_0\|+1,\dots,\infty$	$-l(l+1)$ $l=\|l_0\|;\dots,(n-1)$	$-\frac{\alpha^2}{n^2r_0^2}$	$-\frac{1}{2}\mu\Delta_4^2$	$+\frac{1}{2}\mu\omega^2u$ $\omega=2/\mu u_0^2$	$2\omega n$
	continuum $M^2>0$	-2ν $-\infty<\nu<\infty$	$-l(l+1)$ $l=\|l_0\|,\dots,\infty$	$\frac{\alpha^2}{\nu^2r_0^2}$		$-\frac{1}{2}\mu\omega^2u^2$	$2\omega\nu$
Class II $\|\mathbf{v}\|=1$	$M^2=0$	τ_* $-\infty<\tau_*<\infty$	$-\lambda(\lambda+1)$ $\lambda=\|l_0\|,\dots,\infty$	0	$-\frac{1}{2}\mu\Delta_4^2$	0	$\omega\tau_*$
		Position $r/r_0=\tau_*$, $-\infty<\tau_*<\infty$			position $u^2/u_0^2=\tau$		
Class III $\|\mathbf{v}\|=\infty$	Imaginary $M^2=-\kappa^2$	-2ρ, -2ν $-\infty<(\rho,\nu)<\infty$	$-\lambda(\lambda+1)$ $-l(l+1)$	$\frac{\alpha^2}{\rho^2r_0^2}$ $\frac{\alpha^2}{\nu^2r_0^2}$	$-\frac{1}{2}\mu\Delta_4^2$	$-\frac{1}{2}\mu\omega^2u^2$	$2\omega\rho$ $2\omega\nu$

and choose $\theta=\theta_\rho$ such that

$$\begin{aligned}\mathbb{1}(-\alpha_2\kappa^2+\gamma)\cos\theta_\rho &= \alpha_1\kappa\Gamma_3,\\ \mathbb{1}(-\alpha_2\kappa^2+\gamma)\sin\theta_\rho &= (-\alpha_3\kappa^2+\beta)\Gamma_3.\end{aligned} \tag{III.23}$$

Therefore, $\Phi(|\mathbf{v}|=\infty)=\Phi^{l_0}_{\{\rho\lambda m\}}(\cdots)$ and the mass spectrum of the tachyons is given by

$$-\rho^2=[\alpha_2(-\kappa^2)+\gamma]^2/[\alpha_1^2(-\kappa^2)-(\alpha_3(-\kappa^2)+\beta)^2]. \tag{III.24}$$

The "physical" state $\Psi_\rho(|\mathbf{v}|=\infty)$ can be expanded in terms of canonical basis states as

$$\Psi_\rho(|\mathbf{v}|=\infty)=\frac{1}{N_\rho}\sum_{n=1+|l_0|}^{\infty}V^{\lambda+1}_{n\rho}\left(i\frac{\pi}{2}\right)\times\sum_{l=|l_0|}^{n-1}D^{[n-1,l_0]}_{lm\lambda}(-\theta_\rho-\pi/2)\times\Phi^{l_0}_{\{nlm\}}(\cdots). \tag{III.25}$$

If we choose $\theta=\theta_\nu$ such that

$$\begin{aligned}\mathbb{1}(-\alpha_2\kappa^2+\gamma)\cos\theta_\nu &= (-\alpha_3\kappa^2+\beta)S,\\ \mathbb{1}(-\alpha_2\kappa^2+\gamma)\sin\theta_\nu &= \alpha_1\kappa S,\end{aligned} \tag{III.26}$$

then $\Phi(|\mathbf{v}|=\infty)=\Phi^{l_0}_{\{\nu lm\}}(\cdots)$ and, correspondingly, the mass spectrum is

$$-\nu^2=[\alpha_2(-\kappa^2)+\gamma]^2/[\alpha_1^2(-\kappa^2)-(\alpha_3(-\kappa^2)+\beta)^2], \tag{III.27}$$

which is precisely identical to (III.24).

The solution of (III.21) can be written as

$$\Psi_\nu(|\mathbf{v}|=\infty)=\frac{1}{N_\nu}\sum_{n=1+|l_0|}^{\infty}V^{l+1}_{n\nu}\left(i\frac{\pi}{2}\right)\times\sum_{l'=|l_0|}^{n-1}D^{[n-1,l_0]}_{l'ml}(-\theta_\nu)\times\Phi^{l_0}_{\{nl'm\}}(\cdots). \tag{III.28}$$

In terms of spherical polar coordinates and the four real

TABLE III. Analytic continuations among timelike and spacelike solutions. ---- Weyl unitarity rotation; ···· SO(4) rotation; upper left to lower right diagonal: $\nu\leftrightarrow\nu$, $\theta_\nu\leftrightarrow -i\theta_\nu$, lower left to upper right diagonal $i\rho\leftrightarrow n$, $-i\theta_\rho\leftrightarrow\theta_n$.

Timelike $SO(4,2)\supset SO(3,1)_{1235}\otimes L_{46}$ $\Phi=\|l_0,\nu lm\rangle$ $\Psi(\|\mathbf{v}\|=0)=\exp(-i\theta_\nu T)\Phi$ $-\infty<\nu<\infty$ $l=\|l_0\|,\dots,\infty$ $m=-l,\dots,l$	------------ $i\nu\leftrightarrow n$ $\theta_\nu+i\pi/2\leftrightarrow\theta_n$	Timelike $SO(4,2)\supset SO(4)_{1234}\otimes L_{56}$ $\Phi=\|l_0,nlm\rangle$ $\Psi(\|\mathbf{v}\|=0)=\exp(-i\theta_n T)\Phi$ $n=1+\|l_0\|,\dots,\infty$ $l=\|l_0\|,\dots,(n-1)$ $m=-l,\dots,l$
$\nu\leftrightarrow\rho$ $\theta_\nu\leftrightarrow-i\theta_\rho-i\pi/2$		$n\leftrightarrow i\nu$ $\theta_n\leftrightarrow-i\theta_\nu+i\pi/2$
Spacelike $SO(4,2)\supset SO(3,1)_{1245}\otimes L_{36}$ $\Phi=\|l_0,\rho,\lambda m\rangle$ $\Psi(\|\mathbf{v}\|=\infty)=\exp(+i\theta_\rho A_3)\Phi$ $-\infty<\rho<\infty$ $\lambda=\|l_0\|,\dots,\infty$ $m=-\lambda,\dots,\lambda$	.... $\rho\leftrightarrow\nu$ $\theta_\rho\leftrightarrow\theta_\nu-\pi/2$	Spacelike $SO(4,2)\supset SO(3,1)_{1235})\otimes L_{46}$ $\Phi=\|l_0,\nu lm\rangle$ $\Psi(\|\mathbf{v}\|=\infty)=\exp(+i\theta_\nu A_3)\Phi$ $-\infty<\nu<\infty$ $l=\|l_0\|,\dots,\infty$ $m=-l,\dots,l$

variables, (III.21) becomes

$$\frac{1}{2\mu}\left(\frac{d^2}{dr^2}+\frac{2}{r}\frac{d}{dr}+\frac{v_1}{r}+\frac{v_2}{r^2}+v_3\right)\Phi^{l_0}_N(\alpha' r\theta\phi)=0, \tag{III.29}$$

$$\left(-\frac{1}{2\mu}\Delta^2_4-\tfrac{1}{2}\mu\omega^2u^2-E_N\right)\Phi^{l_0}_N(u_1u_2u_3u_4)=0, \tag{III.30}$$

where N denotes $\{\rho\lambda m\}$ or $\{\nu l m\}$, $r_0v_1=-2\alpha=-2\rho$ or -2ν, $v_2=-\lambda(\lambda+1)$ or $-l(l+1)$, $r_0^2v_3=\alpha^2/\rho^2$ or α^2/ν^2 and $E=v_3/2\mu$. In the second equation, $\omega=2/\mu u_0^2$, $E_N=2\omega\rho$ or $2\omega\nu$. Equation (III.29) describes scattering states of three-dimensional internal Kepler motion and (III.30) describes a four-dimensional harmonic oscillator with repulsive potential. As before, the K–S transformation relates these two internal motions. In both interpretations the constituents do not form bound states. This is in agreement with the interpretation of spacelike solutions as a composite system of two relativistic constituents, one of which has negative energy.[25]

In Table II we have summarized the three classes of solutions corresponding to the two types of internal motion. We could also establish analytic continuations among these solutions and they are detailed in Table III.

IV. CONCLUSION

Our conclusions are mathematical and physical. On the mathematical side we have developed a simple prescription to construct desired basis states using the most degenerate UIR's of SO(4,2). We have reduced SO(4,2) with respect to different subgroups and then constructed the complete set of orthonormal basis states as given in Table I. We have then expanded all noncanonical bases in terms of the canonical one.

On the physical side, we have formulated a relativistic quantum theory suitable to describe a composite object. We have diagonalized the Lorentz invariant wave operator by appropriately defining the physical states, which are the so-called tilted states. The analytic continuation among the different physical states have been established in Table III. In the nonrelativistic limit, the timelike physical wavefunction reduces to the Schrödinger solution for the simple composite object—the hydrogen atom. This relativistic theory can be used without making assumptions about any internal dynamical mechanism, or any internal constituents. However, internal pictures can be obtained by using different realizations of the wave operator. We have demonstrated (Table II) this for an internal three-dimensional Kepler motion and a four-dimensional oscillatory motion. Also we have given the matrix element of the boost transformation between two canonical physical states in terms of known functions. Similar matrix elements between noncanonical physical states can easily be obtained since we have expanded all noncanonical bases in terms of the canonical one.

APPENDIX

In this Appendix we are giving the actions of the group generators on the basis states constructed in Sec. II. We give below only the independent ones in each case; the rest can be obtained from the Lie product (II.1).

$$\begin{aligned}
&\text{(a) } L_{12}\,|\,l_0;nlm\rangle = m\,|\,l_0;nlm\rangle,\quad L_{23}\,|\,l_0;nlm\rangle=\alpha_l^m\,|\,l_0;nlm+1\rangle+\alpha_l^{-m}\,|\,l_0;nlm-1\rangle,\\
&\quad L_{34}\,|\,l_0;nlm\rangle=\beta^m_n(l+1)\,|\,l_0;nl+1m\rangle+b^m_n(l)\,|\,l_0;nlm\rangle+\beta^m_n(l)\,|\,l_0;nl-1m\rangle,\\
&\quad L_{46}\,|\,l_0;nlm\rangle=\alpha^l_n\,|\,l_0;n+1lm\rangle+\alpha^l_{-n}\,|\,l_0;n-1lm\rangle,\quad L_{56}\,|\,l_0;nlm\rangle=n\,|\,l_0;nlm\rangle.
\end{aligned} \tag{A1}$$

$$\begin{aligned}
&\text{(b) } L_{35}\,|\,l_0;\nu lm\rangle=-i\beta^m_{i\nu}(l+1)\,|\,l_0;\nu l+1m\rangle+b^m_{i\nu}(l)\,|\,l_0;\nu lm\rangle-i\beta^m_{i\nu}(l)\,|\,l_0;\nu l-1m\rangle,\\
&\quad L_{56}\,|\,l_0;\nu lm\rangle=\alpha^l_{i\nu}\,|\,l_0;i\nu-1lm\rangle+\alpha^l_{-i\nu}\,|\,l_0;i\nu+1lm\rangle,\\
&\quad L_{46}\,|\,l_0;\nu lm\rangle=\nu\,|\,l_0;\nu lm\rangle.
\end{aligned} \tag{A2}$$

$$\begin{aligned}
&\text{(c) } L_{24}\,|\,l_0;\rho lm\rangle=\alpha^m_\lambda\,|\,l_0;\rho\lambda m+1\rangle+\alpha_\lambda^{-m}\,|\,l_0;\rho lm-1\rangle,\\
&\quad L_{45}\,|\,l_0;\rho\lambda m\rangle=-i\beta^m_{i\rho}(\lambda+1)\,|\,l_0;\rho\lambda+1m\rangle+b^m_\rho(\lambda)\,|\,l_0;\rho\lambda m\rangle-i\beta^m_{i\rho}\,|\,l_0;\rho\lambda-1m\rangle,\\
&\quad L_{56}\,|\,l_0;\rho\lambda m\rangle=\alpha^\lambda_{i\rho}\,|\,l_0;i\rho-1\lambda m\rangle+\alpha^\lambda_{-i\rho}\,|\,l_0;i\rho+1\lambda m\rangle,\\
&\quad L_{36}\,|\,l_0;\rho\lambda m\rangle=\rho\,|\,l_0;\rho\lambda m\rangle.
\end{aligned} \tag{A3}$$

$$\begin{aligned}
&\text{(d) } L_{34}\,|\,l_0;n_1n_2m\rangle=(n_2+k_2-n_1-k_1)\,|\,l_0;n_1n_2m\rangle,\\
&\quad L_{56}\,|\,l_0;n_1n_2m\rangle=[(n_1+k_1+n_2+k_2)\equiv n]\,|\,l_0;n_1n_2m\rangle,\\
&\quad L_{14}\,|\,l_0;n_1n_2m\rangle=\tfrac{1}{2}[(n_2+k_2-k_2')(n_1+k_1-k_1')]^{1/2}\,|\,l_0;n_1-1n_2m+1\rangle\\
&\qquad\qquad +\tfrac{1}{2}[(n_2-1+k_2+k_2')(n_1+1+k_1-k_1')]^{1/2}\,|\,l_0;n_1+1n_2m-1\rangle\\
&\qquad\qquad -\tfrac{1}{2}[(n_2+k_2-k_2')(n_1+k_1+k_1')]^{1/2}\,|\,l_0;n_1n_2-1m+1\rangle\\
&\qquad\qquad -\tfrac{1}{2}[(n_2+1+k_2-k_2')(n_1-1+k_1+k_1')]^{1/2}\,|\,l_0;n_1n_2+1m-1\rangle,\\
&\quad L_{46}\,|\,l_0;n_1n_2m\rangle=\tfrac{1}{2}[(n_2+k_2+k_2')(n_2+1+k_2-k_2')]^{1/2}\,|\,l_0;n_1n_2+1m\rangle\\
&\qquad\qquad -\tfrac{1}{2}[(n_1+k_1+k_1')(n_1+1+k_1-k_1')]^{1/2}\,|\,l_0;n_1+1n_2m\rangle\\
&\qquad\qquad +\tfrac{1}{2}[(n_2+k_2-k_2')(n_2-1+k_2+k_2')]^{1/2}\,|\,l_0;n_1n_2-1m\rangle\\
&\qquad\qquad -\tfrac{1}{2}[(n_1+k_1-k_1')(n_1-1+k_1+k_1')]^{1/2}\,|\,l_0;n_1-1n_2m\rangle.
\end{aligned} \tag{A4}$$

$$\alpha_x^y = \tfrac{1}{2}[(x-y)(x+y+1)]^{1/2},$$

$$\beta_x^z(y) = \frac{1}{y}\left(\frac{1}{(2y+1)(2y-1)}(x-y)(x+y)(y+l_0)(y-l_0)(y-z)(y+z)\right)^{1/2},$$

$$b_z^x(y) = l_0 xz/y(y+1).$$

[1]A.O. Barut, P. Budini, and C. Fronsdal, Proc. R. Soc., Ser. A **291**, 106 (1966); Lectures by Y. Nambu, C. Fronsdal, A.O. Barut, I.T. Todorov, W. Rühl, and T. Takabayasi in *Proceedings of the International Conference on Particles and Fields*, C.R. Hagen, G. Guralnik, and V.A. Mathur (Eds.) (Wiley, New York, 1967).

[2]Y. Nambu, Progr. Theor Phys. Suppls. **37**, and **38**, 368 (1966); Phys. Rev. **160**, 1171 (1967); C. Fronsdal, Phys. Rev. **156**, 1653, 1665 (1967); **163**, 1835 (1967); **168**, 1845 (1968); **171**, 1811 (1968); H. Kleinert, Fortschr. Physik **16**, 1 (1968); H. Kleinert, Springer Tracts in Mod. Phys. **49**, 90 (1969); A.O. Barut, Springer Tracts in Mod. Phys. **50**, 1 (1969); A.O. Barut and R. Wilson, Phys. Rev. D **13**, 2629, 2647 (1976).

[3]W. Rühl, Commun. Math. Phys. **6**, 312 (1967); D.T. Stoyanov and I.T. Todorov, J. Math. Phys. **9**, 2146 (1968); H. Bebié, F. Ghielmetti, V. Gorgé, and H. Leutwyler, Phys. Rev. **177**, 2133, 2146 (1969); E.C.G. Sudarshan and N. Mukunda, Phys. Rev. D **1**, 571 (1970).

[4]A.O. Barut and I.H. Duru, Proc. R. Soc. (London), Ser. A **333**, 217 (1973); Lett. Nuovo Cimento **8**, 768 (1973); Phys. Rev. D **10**, 3448 (1974).

[5]T. Yao, J. Math. Phys. **8**, 1931 (1967); **9**, 1615 (1968); **12**, 315 (1971).

[6]A.O. Barut and A. Böhm, J. Math. Phys **11**, 2938 (1970).

[7]Special cases of these lemmas are given by A. Böhm, Boulder Lectures in Theor. Phys. B **10**, 483 (1968).

[8]R. Gilmore, *Lie Groups, Lie Algebras and Some of Their Applications* (Wiley, New York, 1974).

[9]G. Mack and I.T. Todorov, J. Math. Phys. **10**, 2078 (1969).

[10]H.J. Lipkin, W.I. Weisberger, and M. Peshkin, Ann. Phys. **53**, 203 (1969); A.O. Barut and G.L. Bornzin, J. Math. Phys. **12**, 841 (1971).

[11]M. Jaspers, Nuovo Cimento A **35**, 434 (1976); S.K. Bose and R. Parker, J. Math. Phys. **10**, 812 (1969).

[12]P. Kustaanheimo and E. Stiefel, J. Reine angew. Math. **218**, 204 (1965); M.M. Boiteux and M. René Lucas, C.R. Acad. Sci. Paris **274**, 867 (1972).

[13]N. Limić, J. Niederle, and R. Rączka, J. Math. Phys. **8**, 1079 (1967).

[14]A. Kihlberg, Arkiv för Fysik **30**, 121 (1965).

[15]S. Ström, Arkiv för Fysik **30**, 455 (1965); **40**, 1 (1969).

[16]N.J. Vilenkin, Trans. Moscow Math. Soc. **12**, 209 (1963); "Special Functions and the Theory of Group Representations," AMS Transl. of Math. Monograph, Vol. 22 (1968).

[17]G. Lindblad and B. Nagel, Ann. Inst. Henri Poincaré **13**, 27 (1970).

[18]E.G. Kalnins, J. Patera, R.T. Sharp, and P. Winternitz, Phys. Rev. D **8**, 2552 (1973).

[19]C. Itzykson, Commun. Math. Phys. **4**, 92 (1967).

[20]W. Miller, Jr., Commun. Pure Appl. Math. **17**, 527 (1964); J. Math. Phys. **9**, 1163, 1175, 1434 (1968).

[21]L. Weisner, Pacific J. Math. **5**, 1033 (1955); Can. J. Math. **11**, 141, 148 (1959).

[22]A.O. Barut, D. Corrigan, and H. Kleinert, Phys. Rev. Lett. **20**, 167 (1968); Phys. Rev. **167**, 1527 (1968). For a general axiomatic formulation of infinite-component wave equations, see A.O. Barut, Rep. Math. Phys. **11**, 401 (1977) and "dynamical groups for the Motion of Relativistic Composite Systems" (Proceedings of Tübingen Summer School on Group Theory, 1977).

[23]D. Corrigan and B. Hamprecht, Boulder Lectures in Theor. Phys. A **11**, 299 (1969).

[24]A.O. Barut and R. Wilson, J. Math. Phys. **17**, 900 (1976).

[25]A.O. Barut and J. Nagel, J. Phys. A **10**, 1233 (1977).

Borel multipliers for the Bondi–Metzner–Sachs group[a)]

U. Cattaneo
Institut de Physique, Université de Neuchâtel, CH-2000 Neuchâtel, Switzerland

(Received 16 February 1978)

It is shown that there exist only two cohomology classes of Borel multipliers for the Bondi–Metzner–Sachs group B when the subgroup of "supertranslations" is the additive group of a separable real Hilbert space. This implies that all Borel multipliers for the universal covering group $\tilde{B}$ of B are coboundaries and so every continuous unitary projective representation of B can be obtained from a continuous unitary (ordinary) representation of $\tilde{B}$.

I. INTRODUCTION

In recent years, the Bondi–Metzner–Sachs group (BMS group) has aroused some interest as a possible candidate for replacing the Poincaré group in an elementary particle theory taking gravity into account.[1–3] The BMS group was originally obtained as an asymptotic symmetry group of asymptotically flat (four-dimensional) space–times.[4–6] Apart from its connection with gravitation, the BMS group seems to be suited to elementary particle physics mainly because the "elementary particles" it defines have only discrete spins with a finite number of polarization states and because it leaves open the possibility of a coupling, which avoids the pitfall of O'Raifeartaigh's theorem, with an internal symmetry group.[3]

At present, there is some ambiguity in the term "BMS group" due to the fact that a definitive choice of the subgroup of the so-called "supertranslations" has not yet been made (see Ref. 7 for a discussion of this point). In this paper, we shall assume that the subgroup of "supertranslations" is the additive group of a separable real Hilbert space (which is a fundamental ingredient for the discreteness result quoted above).[2,8]

The BMS group B is a topological group which admits a universal covering group $\tilde{B}$. Irreducible continuous unitary representations of $\tilde{B}$ (on separable complex Hilbert spaces) have been studied by McCarthy, Crampin, and Piard.[2,8–10] However, from the viewpoint of quantum mechanics, this study is fully justified only when it has been shown that all continuous unitary projective representations (CUP-reps) of B can be obtained from the (ordinary) CU-reps of $\tilde{B}$, as in the case of the neutral component $\mathbf{P}_0$ of the Poincaré group and of its universal covering group.

The problem is complicated by the fact that B is not locally compact, and so many familiar techniques may not be applied. Yet, B is a Polish group (see Sec. IIA), i.e., it belongs to the class of all second countable metrically topologically complete groups (which can be considered as the most natural generalizations of second countable locally compact groups). Mackey's theory of CUP-reps can be extended from second countable locally compact to Polish groups,[11] and again Borel unitary multiplier representations (BUM-reps) play a prominent role. A generalization of the theory to include continuous unitary/antiunitary projective representations (CUAP-reps) of Polish groups is also possible.[11] However, since the BMS group is connected, every CUAP-rep of B is a CUP-rep. It should be emphasized that the problem of obtaining the relevant multipliers for the CUP-reps of a Polish group G is completely solved once the corresponding multipliers of the BUM-reps of G (or, shortly, the Borel multipliers for G) are determined. This is true, in particular, when we consider the symmetric multipliers for the subgroup of "supertranslations" (cf. Remark 5).

Our goal is to find the Borel multipliers for B; we shall show that there exist only two cohomology classes of such multipliers, as in the case of $\mathbf{P}_0$. In turn, this implies that there is only one cohomology class (namely, the neutral one) of Borel multipliers for $\tilde{B}$, so that every CUP-rep of B can be obtained from a CU-rep of $\tilde{B}$. McCarthy has tackled this last problem and given a partial result.[12] He got stuck trying to show that every continuous symmetric multiplier for the subgroup of "supertranslations" is cohomologous to the trivial one. In the present paper, the multiplier problem for B will be completely solved by applying techniques quite different from McCarthy's (and which can also be applied to a variety of groups other than B). We shall extensively use general results of the cohomology theory of Polish groups[13,14]; this theory has been presented[11] as the natural frame for the study of the multiplier problem for Polish groups.

In Sec. II, we shall first define the BMS group and show that it is a Polish group. Next, we shall determine the cohomology classes of Borel multipliers for B and obtain the result quoted above. The relation between this and the existence of a unique cohomology class of Borel multipliers for $\tilde{B}$ will be stressed in Sec. III. In the Appendix we shall prove a proposition needed in Sec. II.

The reader is referred to Ref. 11 (resp. Refs. 13, 14) for notation and results concerning CUP-reps (resp. cohomology of Polish groups). Departing from these references, we shall use the additive notation (with the neutral element denoted by 0) for the additive groups of vector spaces and the multiplicative notation (with the neutral element denoted by 1) for any other group considered. The symbol I will always stand for the trivial operation (of a group which should be clear from the context). We shall use the same character to denote a vector space and its additive group.

[a)]Supported by the Swiss National Science Foundation.

II. THE BOREL MULTIPLIERS FOR THE BMS GROUP

A. The structure of B

In this subsection we shall define the BMS group B and show that it is a Polish group.

Let τ be the normalized (positive) rotation invariant Radon measure on the two-dimensional Euclidean unit sphere $\mathbf{S}_2$, and let $\mathscr{L}^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$ be the real vector space of all real-valued τ-square-integrable functions on $\mathbf{S}_2$. We shall denote by $f^{\sim}$ the τ-equivalence class of $f\in\mathscr{L}^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$ and by $L^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$ the real Hilbert space of all τ-equivalence classes of elements of $\mathscr{L}^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$. Moreover, we shall write $\mathscr{L}^2_{\mathbf{R}}$ (resp. $L^2_{\mathbf{R}}$) short for $\mathscr{L}^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$ (resp. $L^2_{\mathbf{R}}(\mathbf{S}_2,\tau)$). For each $\Lambda\in\mathbf{L}_0$ (the neutral component of the Lorentz group) and each $f\in\mathscr{L}^2_{\mathbf{R}}$ we define a real-valued function f_Λ on $\mathbf{S}_2$ by

$$f_\Lambda(\mathbf{x}) = K_{\Lambda^{-1}}(\mathbf{x})f(\Lambda^{-1}\cdot\mathbf{x}). \tag{II.1}$$

Here $\mathbf{x}$ denotes a point of $\mathbf{S}_2$, the dot stands for the conformal operation of $\mathbf{L}_0$ on $\mathbf{S}_2$,[15,16] and

$$K_{\Lambda^{-1}}(\mathbf{x}) = (\Lambda^{-1})^0_\mu\, n^\mu, \tag{II.2}$$

where n is the lightlike four-vector $(1,\mathbf{x})$ and summation convention is used.

Remark 1: Let $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ be the canonical basis of $\mathbf{R}^3$. By means of the stereographic projection with pole $\mathbf{e}_3$, we identify the unit sphere $\mathbf{S}_2$ with the one-point compactification of $\mathbf{C}$.

For each $\Lambda\in\mathbf{L}_0$, we shall denote by $\gamma(\Lambda)\tau$ the measure on $\mathbf{S}_2$ which is the image of τ under the homeomorphism $\gamma(\Lambda):\mathbf{x}\mapsto\Lambda\cdot\mathbf{x}$ of $\mathbf{S}_2$ onto itself.[17]

Remark 2: Let χ be the continuous real-valued function on $\mathbf{L}_0\times\mathbf{S}_2$ defined by $\chi(\Lambda,\mathbf{x}) = K_\Lambda(\mathbf{x})$. Then, by virtue of (II.2), we have $\chi>0$ and

$$\chi(\Lambda'\Lambda,\mathbf{x}) = \chi(\Lambda,\mathbf{x})\chi(\Lambda',\Lambda\cdot\mathbf{x}) \tag{II.3}$$

for all Λ,Λ' in $\mathbf{L}_0$ and all $\mathbf{x}\in\mathbf{S}_2$. It can be easily checked that, for each $\Lambda\in\mathbf{L}_0$,

$$\gamma(\Lambda)\tau = (1/K^2_{\Lambda^{-1}})\cdot\tau \tag{II.4}$$

(the measure with density $1/K^2_{\Lambda^{-1}}$ relative to τ). Therefore, τ is an $\mathbf{L}_0$ quasi-invariant measure.[17]

We define a linear operation Φ of $\mathbf{L}_0$ on $L^2_{\mathbf{R}}$ by

$$\Phi(\Lambda)f^{\sim} = f^{\sim}_\Lambda, \tag{II.5}$$

where $f_\Lambda\in\mathscr{L}^2_{\mathbf{R}}$ is given by (II.1) (cf. Ref. 15, Sec. 3). The mapping $\Lambda\mapsto K_\Lambda$ of $\mathbf{L}_0$ into the set of all continuous real-valued functions on $\mathbf{S}_2$ is said to be the *conformal factor* associated with Φ.

Remark 3: Let $\rho_{\mathbf{L}_0}:\mathbf{SL}(2,\mathbf{C})\to\mathbf{L}_0$ be the covering mapping determined via conformal operation and stereographic projection. Consider the following closed subgroups of $\mathbf{L}_0$:

(a) $P = \left\{\rho_{\mathbf{L}_0}(A)\middle| A = \begin{pmatrix}\alpha & \beta\\ 0 & 1/\alpha\end{pmatrix}\quad \alpha,\beta \text{ in}\mathbf{C} \text{ and } \alpha\neq 0\right\}$,

(b) $N = \left\{\rho_{\mathbf{L}_0}(A)\middle| A = \begin{pmatrix}1 & \beta\\ 0 & 1\end{pmatrix}\quad \beta\in\mathbf{C}\right\}$,

(c) $C = \left\{\rho_{\mathbf{L}_0}(A)\middle| A = \begin{pmatrix}\alpha & 0\\ 0 & 1/\alpha\end{pmatrix}\quad \alpha\in\mathbf{C} \text{ and } \alpha\neq 0\right\}$.

Then $P = NC$ (which is a so-called minimal parabolic subgroup of $\mathbf{L}_0$) is the four-dimensional stabilizer of the pole $\mathbf{e}_3$ (Remark 1). We shall denote by π_P the canonical mapping of $\mathbf{L}_0$ onto the homogeneous space $\mathbf{L}_0/P$ of left cosets and identify this space with $\mathbf{S}_2$ by means of the homeomorphism $\Lambda P\mapsto\Lambda\cdot\mathbf{e}_3$ [Ref. 18, 16.10.8(i)]. We observe that Φ is a (linear) representation of $\mathbf{L}_0$ induced by the trivial representation of P on $\mathbf{R}$. However, it is not unitary because in (II.1) we have $K_{\Lambda^{-1}}(\mathbf{x})$ instead of $(1/K_{\Lambda^{-1}}(\mathbf{x}))$.

Let Φ_d be the contragredient (or dual) representation of Φ, and let ι be the canonical linear mapping of $L^2_{\mathbf{R}}$ onto its (topological) dual $(L^2_{\mathbf{R}})'$. Then $\Phi':\Lambda\mapsto\iota^{-1}\circ\Phi_d(\Lambda)\circ\iota$ is an operation of $\mathbf{L}_0$ on $L^2_{\mathbf{R}}$ and it is easy to show (cf. Ref. 2, Sec. 4), using (II.4), that for each $\Lambda\in\mathbf{L}_0$ and each $f\in\mathscr{L}^2_{\mathbf{R}}$ we have

$$\Phi'(\Lambda)f^{\sim} = f'^{\sim}_\Lambda, \tag{II.6}$$

where $f'_\Lambda\in\mathscr{L}^2_{\mathbf{R}}$ is defined by

$$f'_\Lambda(\mathbf{x}) = (1/K^3_{\Lambda^{-1}}(\mathbf{x}))f(\Lambda^{-1}\cdot\mathbf{x}). \tag{II.7}$$

Remark 4: Since (II.2) gives $K_\Lambda(\mathbf{x}) = 1$ for all $\Lambda\in\mathbf{SO}(3)$ and all $\mathbf{x}\in\mathbf{S}_2$, it follows from (II.1) and (II.7) that $\Phi|\mathbf{SO}(3) = \Phi'|\mathbf{SO}(3)$ and that $\Phi|\mathbf{SO}(3)$ is the quasiregular real representation of $\mathbf{SO}(3)$ [i.e., the one obtained from the quasiregular representation (Ref. 19, Chap. IX, § 2,1) by restriction to $\mathbf{R}$ of the field of scalars]. Therefore, $\Phi|\mathbf{SO}(3)$ is the direct sum of a family $(\phi_j)_{j\in\mathbf{N}}$ of irreducible subrepresentations such that the dimension of ϕ_j is $2j+1$, with each one appearing exactly once (Ref. 19, Chap. IX, § 2,7).

We shall now define the BMS group B and show that it is a topological group (actually a Polish one). A different proof of this has been previously given by Cantoni (Ref. 20, Appendix A).

Proposition 1: The operation Φ (resp. Φ') of $\mathbf{L}_0$ on $L^2_{\mathbf{R}}$ defined by (II.5) [resp. (II.6)] is topological.

Proof: We shall give a proof for Φ; the proof for Φ' goes along the same lines. Since $L^2_{\mathbf{R}}$ is separable, it is enough (Ref. 21, Lemma 1) to show that all partial mappings determined by $(\Lambda,f^{\sim})\mapsto\Phi(\Lambda)f^{\sim}$ are continuous. Now for each $f\in\mathscr{L}^2_{\mathbf{R}}$ the mapping $h:\Lambda\mapsto\Phi(\Lambda)f^{\sim} = f^{\sim}_\Lambda$ is continuous. For we have, by applying (II.3) and (II.4),

$$\begin{aligned}\|\Phi(\Lambda)f^{\sim}\|^2 &= \int_{\mathbf{S}_2}|f_\Lambda(\mathbf{x})|^2\,d\tau(\mathbf{x})\\ &= \int_{\mathbf{S}_2}K^2_{\Lambda^{-1}}(\Lambda\cdot\mathbf{x})|f(\mathbf{x})|^2\,d(\gamma(\Lambda^{-1})\tau)(\mathbf{x})\\ &= \int_{\mathbf{S}_2}(1/K^4_\Lambda(\mathbf{x}))|f(\mathbf{x})|^2\,d\tau(\mathbf{x})\end{aligned}$$

and

$$\begin{aligned}(\Phi(\Lambda)f^{\sim}|g^{\sim}) &= \int_{\mathbf{S}_2}f_\Lambda(\mathbf{x})g(\mathbf{x})\,d\tau(\mathbf{x})\\ &= \int_{\mathbf{S}_2}(1/K^3_\Lambda(\mathbf{x}))\,g(\Lambda\cdot\mathbf{x})f(\mathbf{x})\,d\tau(\mathbf{x})\end{aligned}$$

for all $\Lambda\in\mathbf{L}_0$ and all real-valued continuous functions g on $\mathbf{S}_2$. It follows that, if (Λ_n) is any sequence in $\mathbf{L}_0$ converging to Λ, then

$$\lim_{n\to\infty}\|\Phi(\Lambda_n)f^{\sim}\| = \|\Phi(\Lambda)f^{\sim}\|$$

and

$$\lim_{n\to\infty}(\Phi(\Lambda_n)f^{\sim}|g^{\sim}) = (\Phi(\Lambda)f^{\sim}|g^{\sim}),$$

by the dominated convergence theorem, whence the continuity of h. Moreover, for each $\Lambda \in \mathbf{L}_0$ and each $f \in \mathscr{L}^2_{\mathbf{R}}$, we have

$$\|\Phi(\Lambda) f^{\sim}\|^2 \leqslant \left\{ \sup_{\mathbf{x} \in \mathbf{S}_2} [1/K^4_\Lambda(\mathbf{x})] \right\} \|f^{\sim}\|^2,$$

so that the linear mapping $f^{\sim} \mapsto \Phi(\Lambda) f^{\sim}$ is continuous. ■

Now let B denote the external semidirect product of $\mathbf{L}_0$ by the additive group of $L^2_{\mathbf{R}}$ relative to Φ; i.e., put

$$B = L^2_{\mathbf{R}} \times_\Phi \mathbf{L}_0,$$

the set of all ordered pairs $(f^{\sim}, \Lambda)$ $(f \in \mathscr{L}^2_{\mathbf{R}}; \Lambda \in \mathbf{L}_0)$ equipped with the multiplication defined by

$$(f^{\sim}, \Lambda)(f'^{\sim}, \Lambda') = (f^{\sim} + \Phi(\Lambda) f'^{\sim}, \Lambda\Lambda').$$

By virtue of Proposition 1, the product topology is compatible with the group structure and makes B into a topological group, the BMS group. Since $L^2_{\mathbf{R}}$ is separable, B is Polish. We identify $L^2_{\mathbf{R}}$ (resp. $\mathbf{L}_0$) with the closed subgroup of "supertranslations" $L^2_{\mathbf{R}} \times \{1\}$ (resp. with the closed subgroup $\{0\} \times \mathbf{L}_0$) of B.

The BMS group admits a universal covering group,

$$\widetilde{B} = L^2_{\mathbf{R}} \times_{\Phi \circ \rho_{\mathbf{L}_0}} \mathbf{SL}(2,\mathbf{C}),$$

which is obviously Polish.

B. The Borel multipliers

Our goal in this subsection is to determine the group $H^2_{bm}(B,\mathbf{U}(1)_I)$ of all cohomology classes of multipliers of BUM-reps of B. First, let us recollect the meaning of Borel multipliers in the theory of CUP-reps of B.

We know that the study of CUP-reps of B is equivalent to the study of its BUM-reps and this, in turn, to the study of some (ordinary) CU-reps of groups obtained from topological extensions of B by $\mathbf{U}(1)_I$.[11] More precisely, if $\mathfrak{H}$ is a separable complex Hilbert space and $\tilde{u}$ a CUP-rep of B on $\mathbf{P}(\mathfrak{H})$ [namely, a continuous group homomorphism of B into the projective unitary group $\mathbf{PU}(\mathfrak{H})$], we can choose a normalized Borel section Σ associated with the canonical surjection $\Omega: \mathbf{U}(\mathfrak{H}) \to \mathbf{PU}(\mathfrak{H})$ and obtain a BUM-rep $u = \Sigma \circ \tilde{u}$ of B on $\mathfrak{H}$ (a lifting of $\tilde{u}$) with multiplier, say, μ. From u, we get a CU-rep on $\mathfrak{H}$

$$w: (\zeta, (f^{\sim}, \Lambda)) \mapsto \zeta u(f^{\sim}, \Lambda) \quad [\zeta \in \mathbf{U}(1); (f^{\sim}, \Lambda) \in B]$$

of a Polish group B^μ (the μ-extension of B). This representation satisfies $w(\zeta,(0,1)) = \zeta\,\mathrm{Id}_{\mathfrak{H}}$, i.e., is a $\mathbf{U}(1)$-split CU-rep. Conversely, if we start with w we obtain the CUP-rep $\tilde{u}$. It should be emphasized that it is not really the Borel mapping μ which is relevant in this context but its (Mackey–Moore) cohomology class $[\mu]$. Indeed, the choice of normalized Borel sections associated with Ω and different from Σ can lead to BUM-reps of B on $\mathfrak{H}$ different from u, but which are all liftings of $\tilde{u}$ with multipliers belonging to $[\mu]$. Moreover, every element of $[\mu]$ is a multiplier of a lifting of $\tilde{u}$. For this reason we also say that $\tilde{u}$ is a CU$[\mu]$-rep.

It should now be clear why the first step in the study of CUP-reps of B is the determination of $H^2_{bm}(B,\mathbf{U}(1)_I)$, the subgroup of $H^2_b(B,\mathbf{U}(1)_I)$, whose elements are equivalence classes of multipliers of BUM-reps of B (Ref. 11, Sec. IV). To apply the proposition of the Appendix to B, we need the following lemmas.

Lemma 1: Let P, N, C be the subgroups of $\mathbf{L}_0$ defined in Remark 3. Let ψ be a nontrivial operation of P on $\mathbf{R}$ such that $\psi(N) = \{\mathrm{Id}_{\mathbf{R}}\}$ and $\psi(c) = \epsilon(c)\mathrm{Id}_{\mathbf{R}}$ for all $c \in C$, where ϵ is a mapping of P into the set of all strictly positive real numbers. Then $H^1(P,\mathbf{R}_\psi) = \{0\}$.

Proof: Let $\mu_1 \in Z^1(N,\mathbf{R}_I)^C$, i.e., suppose that $\mu_1 \in Z^1(N,\mathbf{R}_I)$ and satisfies $\psi(c)\mu_1(n) = \mu_1(cnc^{-1})$ for all $c \in C$ and all $n \in N$. Then, if we choose

$$c' = \rho_{\mathbf{L}_0}\left(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\right),$$

we have

$$\psi(c')\mu_1(n) = \epsilon(c')\mu_1(n) = -\mu_1(n)$$

for all $n \in N$, hence $\mu_1 = 0$. Since $B^1(N,\mathbf{R}_I) = \{0\}$, it follows that $H^1(N,\mathbf{R}_I)^C = \{0\}$. Now, let $\mu_2 \in Z^1(C,\mathbf{R}_{\psi|C})$. As ψ is nontrivial, we can choose $c_0 \in C$ such that $\epsilon(c_0) \neq 1$. By the 1-cocycle identity and the commutativity of C, we have

$$\mu_2(c) = [-1+\epsilon(c)][-1+\epsilon(c_0)]^{-1}\mu_2(c_0)$$

for all $c \in C$, i.e., $\mu_2 = \delta\nu$, where $\nu = [-1+\epsilon(c_0)]^{-1}\mu_2(c_0) \in \mathbf{R}$. Therefore, $H^1(C,\mathbf{R}_{\psi|C}) = \{0\}$. By applying Proposition 1(ii) and Remark 2 of Ref. 13, Sec. 4 to the (internal) semidirect product P of C by N, we obtain

$$H^1(P,\mathbf{R}_\psi) \approx H^1(N,\mathbf{R}_I)^C \times H^1(C,\mathbf{R}_{\psi|C}) = \{0\}. \quad ■$$

The proof of the following lemma is the same as that of Ref. 22, Theorem 15.1.

Lemma 2: If Ψ is any topological operation of $\mathbf{SO}(3)$ on the additive group of $L^2_{\mathbf{R}}$, then $H^1_b(\mathbf{SO}(3),(L^2_{\mathbf{R}})_\Psi) = \{0\}$.

Lemma 3: Let χ be a real-valued continuous function on $\mathbf{L}_0 \times \mathbf{S}_2$ which is everywhere strictly greater than zero and satisfies $\gamma(\Lambda)\tau = \chi(\Lambda^{-1},\cdot)\cdot\tau$ for all $\Lambda \in \mathbf{L}_0$, as well as $\chi(\Lambda,\mathbf{x}) = 1$ for all $\Lambda \in \mathbf{SO}(3)$ and all $\mathbf{x} \in \mathbf{S}_2$. Suppose there exist a real number α and a topological operation Ψ of $\mathbf{L}_0$ on the additive group of $L^2_{\mathbf{R}}$ defined by $\Psi(\Lambda) f^{\sim} = f^{\sim}_{\Lambda,\alpha}$, where, for each $\mathbf{x} \in \mathbf{S}_2$,

$$f_{\Lambda,\alpha}(\mathbf{x}) = \chi(\Lambda^{-1},\mathbf{x})^\alpha f(\Lambda^{-1}\cdot\mathbf{x}) \quad (\Lambda \in \mathbf{L}_0; f \in \mathscr{L}^2_{\mathbf{R}}).$$

Then $H^1_b(\mathbf{L}_0,(L^2_{\mathbf{R}})_\Psi) = \{0\}$.

Proof: This demonstration will be modeled on that of Ref. 22, Theorem 14.1. Let P and π_P be as in Remark 3, let v be a Haar measure on $\mathbf{L}_0$, and let Δ_P be the modulus function on P. It is well known that $\mathbf{L}_0$ is unimodular (because it is a semisimple connected Lie group) and that P is not unimodular (cf. for instance Ref. 17, § 3, Exemple 4). Since $\mathbf{L}_0$ is countable at infinity, there exists a (locally v-integrable) real-valued function ρ on it, which never takes the value 0, satisfying

$$\chi(\Lambda'^{-1},\pi_P(\Lambda)) = \rho(\Lambda'^{-1}\Lambda)\rho(\Lambda)^{-1} \quad v(\Lambda)\text{-a.e.} \tag{II.8}$$

for all $\Lambda' \in \mathbf{L}_0$ and

$$\rho(\Lambda p) = \Delta_P(p)\rho(\Lambda) \quad v(\Lambda),\text{-a.e.} \tag{II.9}$$

for all $p \in P$ (Ref. 17, Chap. VII, § 2, Lemme 4 and 5; Ref. 23, Chap. IV, § 5, Cor. 3 to Prop. 5). Here, "$v(\Lambda)$-a.e." is short for "almost everywhere on $\mathbf{L}_0$ with respect to v." By virtue of (II.8), we can assume $\rho > 0$ because $\chi > 0$.

Now, let $\Lambda \mapsto f'_\Lambda$ be a mapping of $\mathbf{L}_0$ into $\mathscr{L}^2_\mathbf{R}$ such that $\Lambda \mapsto f'^{\sim}_\Lambda$ is an element of $Z^1_b(\mathbf{L}_0,(L^2_\mathbf{R})_\Psi)$. By Lemma 2, we can find a mapping $\Lambda \mapsto f_\Lambda$ of $\mathbf{L}_0$ into $\mathscr{L}^2_\mathbf{R}$ such that $\Lambda \mapsto f^{\sim}_\Lambda$ is a Borel 1-cocycle (actually continuous) cohomologous to $\Lambda \mapsto f'^{\sim}_\Lambda$ and satisfying $f^{\sim}_\Lambda = 0$ for all $\Lambda \in \mathbf{SO}(3)$. We define a real-valued function ω on $\mathbf{L}_0 \times \mathbf{S}_2$ by $\omega(\Lambda,\mathbf{x}) = f_\Lambda(\mathbf{x})$. By virtue of the 1-cocycle identity, ω satisfies the relation relation

$$\omega(\Lambda'\Lambda,\mathbf{x}) = \omega(\Lambda',\mathbf{x}) + \chi(\Lambda'^{-1},\mathbf{x})^\alpha \omega(\Lambda,\Lambda'^{-1}\cdot\mathbf{x}) \quad \tau(\mathbf{x})\text{-a.e.} \tag{II.10}$$

for all Λ,Λ' in $\mathbf{L}_0$.

If ξ is an arbitrary real-valued function on $\mathbf{L}_0 \times \mathbf{S}_2$, we shall denote by ξ^l the real-valued function on $\mathbf{L}_0 \times \mathbf{L}_0$ defined by $\xi^l(\Lambda,\Lambda') = \xi(\Lambda,\pi_P(\Lambda'))$. Since χ satisfies the identity (II.3) $\tau(\mathbf{x})$-a.e., we have

$$\chi(\Lambda'^{-1},\mathbf{x}) = \chi(\Lambda',\Lambda'^{-1}\cdot\mathbf{x})^{-1} \quad \tau(\mathbf{x})\text{-a.e.}$$

for all $\Lambda' \in \mathbf{L}_0$. Moreover, a subset A of $\mathbf{S}_2$ is τ-negligible if and only if $\pi_P^{-1}(A)$ is v-negligible (Ref. 17, Chap. VII, § 2, Lemme 4). It follows then from (II.10) that

$$\omega^l(\Lambda,\Lambda'^{-1}\Lambda'') = \chi^l(\Lambda',\Lambda'^{-1}\Lambda'')^\alpha[\omega^l(\Lambda'\Lambda,\Lambda'') - \omega^l(\Lambda',\Lambda'')] \quad v(\Lambda'')\text{-a.e.} \tag{II.11}$$

for all Λ,Λ' in $\mathbf{L}_0$. If we put $\Lambda'^{-1}\Lambda'' = \Lambda_1$ in (II.11), apply (II.8), and take account of a part of Fubini's theorem (Ref. 24, § 36, Theorem A), we obtain, for some fixed $\Lambda'' \in \mathbf{L}_0$,

$$\omega^l(\Lambda,\Lambda_1) = \rho(\Lambda_1)^{-\alpha}\rho(\Lambda'')^\alpha[\omega^l(\Lambda''\Lambda_1^{-1}\Lambda,\Lambda'') - \omega^l(\Lambda''\Lambda_1^{-1},\Lambda'')] \quad v(\Lambda)v(\Lambda_1)\text{-a.e.} \tag{II.12}$$

[where "$v(\Lambda)v(\Lambda_1)$-a.e." is short for "almost everywhere on $\mathbf{L}_0 \times \mathbf{L}_0$ with respect to the product measure $v \otimes v$"].

Keeping Λ'' fixed as above, we can define a real-valued function θ on $\mathbf{L}_0$ by

$$\theta(\Lambda) = \rho(\Lambda'')^\alpha \omega^l(\Lambda''\Lambda^{-1},\Lambda'')$$

and then write (II.12) as

$$\omega^l(\Lambda,\Lambda_1) = \rho(\Lambda_1)^{-\alpha}(\theta(\Lambda^{-1}\Lambda_1) - \theta(\Lambda_1)) \quad v(\Lambda)v(\Lambda_1)\text{-a.e.} \tag{II.13}$$

Since, for each pair Λ,Λ_1 of elements of $\mathbf{L}_0$ and each $p \in P$,

$$\omega^l(\Lambda,\Lambda_1 p) = \omega^l(\Lambda,\Lambda_1), \tag{II.14}$$

it follows that

$$\begin{aligned}\omega^l(\Lambda_1,\Lambda_1 p) &= \rho(\Lambda_1)^{-\alpha}\Delta_P(p)^{-\alpha}(\theta(p) - \theta(\Lambda_1 p))\\ &= \omega^l(\Lambda_1,\Lambda_1) = \rho(\Lambda_1)^{-\alpha}(\theta(1) - \theta(\Lambda_1))\end{aligned}$$
$$v(\Lambda_1)\text{-a.e.}$$

for all $p \in P$, where (II.9) has been used. Therefore, the real-valued function β on P defined by

$$\beta(p) = \theta(p^{-1}) - \Delta_p(p)^{-\alpha}\theta(1)$$

satisfies

$$\beta(p) = \theta(\Lambda_1 p^{-1}) - \Delta_P(p)^{-\alpha}\theta(\Lambda_1) \quad v(\Lambda_1)\text{-a.e.,} \tag{II.15}$$

for all $p \in P$. The operation ψ of P on $\mathbf{R}$ given by $\psi(p) = \Delta_p(p)^{-\alpha}\mathrm{Id}_\mathbf{R}$ is nontrivial. However, $\psi(N) = \{\mathrm{Id}_\mathbf{R}\}$ [by Ref. 17, Chap. VII, § 2, Prop. 10(b)] because N is Abelian. We check easily that $\beta \in Z^1(P,\mathbf{R}_\psi)$ by reason of (II.15) and infer the existence of a real number ζ such that

$$\beta(p) = -\zeta + \Delta_p(p)^{-\alpha}\zeta \tag{II.16}$$

for all $p \in P$ (Lemma 1).

Next, let p,p' be arbitrary elements of P and let ζ be as in (II.16). By applying successively (II.14), (II.13), (II.9), (II.15), (II.16), and once more (II.9), we have

$$\begin{aligned}\omega^l(\Lambda,\Lambda_1) = \omega^l(\Lambda,\Lambda_1 p)\\ &= \rho(\Lambda_1 p)^{-\alpha}(\theta(\Lambda^{-1}\Lambda_1 p) - \theta(\Lambda_1 p))\\ &= \rho(\Lambda_1 p)^{-\alpha}\theta(\Lambda^{-1}\Lambda_1 p)\\ &\quad - \rho(\Lambda_1 p')^{-\alpha}\Delta_P(p'^{-1}p)^{-\alpha}\theta(\Lambda_1 p)\\ &= \rho(\Lambda_1 p)^{-\alpha}\theta(\Lambda^{-1}\Lambda_1 p)\\ &\quad - \rho(\Lambda_1 p')^{-\alpha}\theta(\Lambda_1 p')\\ &\quad + \rho(\Lambda_1 p')^{-\alpha}\beta(p'^{-1}p)\\ &= \rho(\Lambda_1 p)^{-\alpha}(\theta(\Lambda^{-1}\Lambda_1 p) + \zeta)\\ &\quad - \rho(\Lambda_1 p')^{-\alpha}(\theta(\Lambda_1 p') + \zeta)\end{aligned}$$
$$v(\Lambda)v(\Lambda_1)\text{-a.e.} \tag{II.17}$$

On the other hand, by (II.9), (II.15), and (II.16),

$$\begin{aligned}\rho(\Lambda_1 p)^{-\alpha}(\theta(\Lambda_1 p) + \zeta)\\ &= \rho(\Lambda_1)^{-\alpha}(\theta(\Lambda_1) + \zeta) \quad v(\Lambda_1)\text{-a.e.}\end{aligned}$$

This implies the existence of a real-valued function h on $\mathbf{S}_2$ such that

$$h(\pi_P(\Lambda_1)) = \rho(\Lambda_1 p)^{-\alpha}(\theta(\Lambda_1 p) + \zeta) \quad v(\Lambda_1)\text{-a.e.}$$

for all $p \in P$. It follows from (II.17), with the help of (II.8), that

$$\omega(\Lambda,\mathbf{x}) = \chi(\Lambda^{-1},\mathbf{x})^\alpha h(\Lambda^{-1}\cdot\mathbf{x}) - h(\mathbf{x}) \quad v(\Lambda)\tau(\mathbf{x})\text{-a.e.} \tag{II.18}$$

We check [by applying (II.10) and (II.3) $\tau(\mathbf{x})$-a.e.] that the set L' of all $\Lambda \in \mathbf{L}_0$ such that

$$\omega(\Lambda,\mathbf{x}) = \chi(\Lambda^{-1},\mathbf{x})^\alpha h(\Lambda^{-1}\cdot\mathbf{x}) - h(\mathbf{x}) \quad \tau(\mathbf{x})\text{-a.e.} \tag{II.19}$$

is actually a subgroup of $\mathbf{L}_0$. Since $\mathbf{L}_0 - L'$ is v-negligible by (II.18) and Ref. 24, § 36, Theorem A, we have $L' = \mathbf{L}_0$, i.e., (II.19) is satisfied for all $\Lambda \in \mathbf{L}_0$. On the other hand, since $\omega(\Lambda,\mathbf{x}) = 0$ $\tau(\mathbf{x})$-a.e. for all $\Lambda \in \mathbf{SO}(3)$, there exists $\mathbf{x}' \in \mathbf{S}_2$ such that, when $\Lambda \in \mathbf{SO}(3)$, we have

$$\omega(\Lambda,\mathbf{x}') = 0 \quad \nu(\Lambda)\text{-a.e.,}$$

and [by (II.18)]

$$h(\Lambda^{-1}\cdot\mathbf{x}') = h(\mathbf{x}') \quad \nu(\Lambda)\text{-a.e.}$$

As

$$\{\Lambda^{-1}\cdot\mathbf{x}' \mid \Lambda \in \mathbf{SO}(3) \text{ and } h(\Lambda^{-1}\cdot\mathbf{x}') = h(\mathbf{x}')\} = \mathbf{S}_2 \quad \nu(\Lambda)\text{-a.e.,}$$

we can conclude that h is constant almost everywhere on $\mathbf{S}_2$ with respect to τ. In particular $h \in \mathscr{L}^2_\mathbf{R}$, and then (II.19) implies

$$f^{\sim}_\Lambda = -h^{\sim} + \Psi(\Lambda)h^{\sim} = \delta h^{\sim}(\Lambda)$$

for all $\Lambda \in \mathbf{L}_0$. ■

Let A be either $\mathbf{U}(1)$ or $\mathbf{R}$. According to Ref. 13, Sec. 4, we shall denote by $\hat{\Gamma}^2_Z$ the operation of $\mathbf{L}_0$ on $Z^2_b(L^2_\mathbf{R},A_l)$

such that

$$(\hat{I}^2_Z(\Lambda)h)(f^\sim,f'^\sim) = h(\Phi(\Lambda^{-1})f^\sim,\Phi(\Lambda^{-1})f'^\sim)$$

for all $\Lambda\in\mathbf{L}_0$, all $h\in Z^2_b(L^2_\mathbf{R},A_I)$ and all f,f' in $\mathscr{L}^2_\mathbf{R}$. Let $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathrm{U}(1))$ be as in the Appendix and define $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})$ likewise. Notice that

$$\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};A)\subseteq Z^2_b(L^2_\mathbf{R},A_I).$$

If $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};A)^{\mathbf{L}_0}$ stands for the subgroup of all elements of $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};A)$ invariant under $\hat{I}^2_Z(\Lambda)$ for all $\Lambda\in\mathbf{L}_0$, we have the following result (cf. Ref. 12b, Appendix A for a different proof).

Lemma 4: $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathrm{U}(1))^{\mathbf{L}_0}=\{1\}$.

Proof: To begin with, we remark that, because of Ref. 25, Corollary 1, it is enough to show that $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})^{\mathbf{L}_0}=\{0\}$. For if π is the covering mapping of $\mathbf{R}$ onto U(1), the mapping $\pi_a: f\mapsto\pi\circ f$ of $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})$ into $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathrm{U}(1))$ is a group isomorphism satisfying

$$\pi_a\circ\hat{I}^2_{Z,a}(\Lambda)=\hat{I}^2_{Z,a}(\Lambda)\circ\pi_a$$

for all $\Lambda\in\mathbf{L}_0$, where $\hat{I}^2_{Z,a}(\Lambda)=\hat{I}^2_Z(\Lambda)|\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};A)$ (with A as above). Since

$$\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})=\mathscr{L}(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})$$

(the vector space for all continuous bilinear forms on $L^2_\mathbf{R}\times L^2_\mathbf{R}$), there exists a canonical group isomorphism λ of $\mathscr{H}om(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})$ onto the additive group of the (completed) tensor product $L^2_\mathbf{R}\hat{\otimes}L^2_\mathbf{R}$. One checks easily that

$$\lambda_a\circ\hat{I}^2_{Z,a}(\Lambda)=(\Phi'\otimes\Phi')(\Lambda)\circ\lambda_a \tag{II.20}$$

for all $\Lambda\in\mathbf{L}_0$, where $\lambda_a=\lambda\,|\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})$. Now, any element f of $\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathbf{R})^{\mathbf{L}_0}$ satisfies $\hat{I}^2_{Z,a}(\Lambda)f=f$ for all $\Lambda\in\mathbf{L}_0$. By virtue of (II.20), this implies $(\Phi'\otimes\Phi')(\Lambda)\times\lambda_a(f)=\lambda_a(f)$; thus $\lambda_a(f)$ belongs to a one-dimensional subrepresentation of $\Phi'\otimes\Phi'$ or $f=0$. Moreover, $\lambda_a(f)$ is an antisymmetric tensor. But $\Phi'\otimes\Phi'|\mathbf{SO}(3)$ has no one-dimensional antisymmetric subrepresentation, for its one-dimensional subrepresentations are (unitarily equivalent to) subrepresentations of $\phi_j\otimes\phi_j$ $(j\geqslant0)$, hence symmetric (Remark 4). Therefore, $f=0$. ■

Proposition 2: $H^2_b(B,\mathrm{U}(1)_I)$ is a finite group of order 2.

Proof: Since $H^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)$ coincides with $H^1_c(L^2_\mathbf{R},\mathrm{U}(1)_I)$ and is trivially identified with $Z^1_c(L^2_\mathbf{R},\mathrm{U}(1)_I)$ (the group of all continuous unitary characters of the additive group of $L^2_\mathbf{R}$), there exists a canonical group isomorphism γ of $L^2_\mathbf{R}$ onto $H^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)$ (Ref. 26, 23.32). Let $\hat{I}^1_*$ be the operation of $\mathbf{L}_0$ on $H^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)$ such that, for each $\Lambda\in\mathbf{L}_0$ and each $\nu\in Z^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)$, $\hat{I}^1_*(\Lambda)[\nu]$ is (identified with) the 1-cocycle ν_Λ defined by $\nu_\Lambda(f^\sim)=\nu(\Phi(\Lambda^{-1})f^\sim)$ (Ref. 13, Sec. 4). Then γ satisfies

$$\gamma\circ\Phi'(\Lambda)=\hat{I}^1_*(\Lambda)\circ\gamma$$

for all $\Lambda\in\mathbf{L}_0$, i.e., it is an $\mathbf{L}_0$-module isomorphism [with $L^2_\mathbf{R}$ made into a topological $\mathbf{L}_0$-module relative to Φ' (Proposition 1)]. If we topologize $H^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)$ by transport of structure via γ, the operation $\hat{I}^1_*$ becomes topological and

$$H^1_b(\mathbf{L}_0,H^1_b(L^2_\mathbf{R},\mathrm{U}(1)_I)_{\hat{I}^1_*})\approx H^1_b(\mathbf{L}_0,(L^2_\mathbf{R})_{\Phi'}).$$

Moreover, the topology so defined is obviously finer than that of simple convergence. Therefore, by the proposition of the Appendix and Lemma 3, we have

$$H^2_b(B,\mathrm{U}(1)_I)\approx H^2_{b,s}(L^2_\mathbf{R},\mathrm{U}(1)_I)'\times\mathscr{H}om_a(L^2_\mathbf{R},L^2_\mathbf{R};\mathrm{U}(1))^{\mathbf{L}_0}\times H^2_b(\mathbf{L}_0,\mathrm{U}(1)_I).$$

Our thesis follows now from Lemma 4, from Ref. 25, Corollary 2, and from the well-known result that $H^2_b(\mathbf{L}_0,\mathrm{U}(1)_I)$ is a finite group of order 2 (this can be proved, for instance, by using the inflation-restriction sequence of Ref. 27, Chap. I, 5). ■

Remark 5: The relation $H^2_{b,s}(L^2_\mathbf{R},\mathrm{U}(1)_I)=\{1\}$ also implies $H^2_{c,s}(L^2_\mathbf{R},\mathbf{R}_I)=\{0\}$. In fact,

$$H^2_{b,s}(L^2_\mathbf{R},\mathrm{U}(1)_I)\approx H^2_{lr,s}(L^2_\mathbf{R},\mathrm{U}(1)_I)$$

by virtue of Ref. 14, Corollary 2; moreover

$$H^2_{lr,s}(L^2_\mathbf{R},\mathrm{U}(1)_I)\approx H^2_{lr,s}(L^2_\mathbf{R},\mathbf{R}_I)\approx H^2_{c,s}(L^2_\mathbf{R},\mathbf{R}_I)$$

on account of Ref. 25, Corollary, and of Ref. 14, Remark 5. This settles a comment of McCarthy (Ref. 12b, p. 147) on Ref. 28 (Sec. 3).

Remark 6: We can choose the following 2-cocycles as representatives of the two cohomology classes belonging to $H^2_b(B,\mathrm{U}(1)_I)$ [cf. Ref. 13, Proposition 1 (iii)]:

(a) the neutral element 1 of $Z^2_b(B,\mathrm{U}(1)_I)$ (i.e., the trivial 2-cocycle),

(b) μ_B defined by

$$\mu_B((f^\sim,\Lambda),(f'^\sim,\Lambda'))=\mu_{\mathbf{L}_0}(\Lambda,\Lambda')$$

for all f,f' in $\mathscr{L}^2_\mathbf{R}$ and all Λ,Λ' in $\mathbf{L}_0$, where $[\mu_{\mathbf{L}_0}]$ is the element of $H^2_b(\mathbf{L}_0,\mathrm{U}(1)_I)$ different from the neutral one. It is well known that $\mu_{\mathbf{L}_0}$ (and therefore also μ_B) can be chosen to be locally trivial, i.e., to take the value 1 in some neighborhood of 1×1 (this follows also from Ref. 14, Lemmas 2 and 3).

So far we have determined $H^2_b(B,\mathrm{U}(1)_I)$. However, it is obvious from our result that

$$H^2_{bm}(B,\mathrm{U}(1)_I)=H^2_b(B,\mathrm{U}(1)_I);$$

thus we are done. The physical meaning of $[\mu_B]$ is exactly the same as that of the corresponding element of $H^2_b(\mathbf{P}_0,\mathrm{U}(1)_I)$: It brings fermions into play.

III. THE BOREL MULTIPLIERS FOR $\widetilde{B}$

If we apply the general theory of CUP-reps of Polish groups,[11] we can now conclude that, by virtue of Proposition 2 and Remark 6, all CUP-reps of B arise from U(1)-split CU-reps of the Polish groups B^1 and B^{μ_B}. Since every U(1)-split CU-rep of B^1 is identifiable in a trivial way with a CU-rep of B, we have only to deal with CU-reps of B and U(1)-split CU-reps of B^{μ_B}. But, as in the case of $\mathbf{P}_0$, for the BMS group we have an even better result by considering the universal covering group.

Proposition 3: $H^2_b(\widetilde{B},\mathrm{U}(1)_I)=\{1\}$.

Proof: Let $\rho_B:\widetilde{B}\to B$ be the covering mapping and let $(\widetilde{E},\widetilde{\rho})$ be an arbitrary topological extension of $\widetilde{B}$ by $\mathbf{U}(1)_t$. If we put $\rho=\rho_B\circ\widetilde{\rho}$, we see that $\mathrm{Ker}\rho/\mathbf{U}(1)$ is a group of order 2 and $\mathrm{Ker}\rho$ is Abelian. Let σ be a normalized section associated with ρ and let $\rho':\mathrm{Ker}\rho\to\mathrm{Ker}\rho/\mathbf{U}(1)$ be the canonical surjection. The topological operation Ψ of B on $\mathrm{Ker}\rho$ such that, for each $b\in B$ and each $k\in\mathrm{Ker}\rho$, $\Psi(b)k=\sigma(b)k\sigma(b)^{-1}$ is trivial. For the operation Ψ_q of B on $\mathrm{Ker}\rho/\mathbf{U}(1)$ satisfying $\Psi_q(b)\circ\rho'=\rho'\circ\Psi(b)$ for all $b\in B$ is topological (Ref. 29, TGIII, § 2, Prop. 11), hence trivial because B is connected. On the other hand, we have $Z_c^1(B,\mathbf{U}(1)_t)=\{1\}$ by virtue of Ref. 13, Proposition 1 (ii) because $(L_{\mathbf{R}}^2)_{\Phi'}$ has no nontrivial $\mathbf{L}_0$-invariant element. It follows that $(\widetilde{E},\rho)$ is a topological extension of B by $(\mathrm{Ker}\rho)_t$. Moreover, as

$$H_b^2(\mathrm{Ker}\rho/\mathbf{U}(1),\mathbf{U}(1)_t)=\{1\},$$

$\mathrm{Ker}\rho$ is the (internal) direct product of $\mathbf{U}(1)$ and a finite subgroup C_2 of order 2.

The mapping $\rho_q:\widetilde{E}/C_2\to B$ deduced from ρ by passing to the quotient is continuous and open; so, if $\rho'':\widetilde{E}\to\widetilde{E}/C_2$ is the canonical surjection, then $(\widetilde{E}/C_2,\rho_q)$ is a topological extension of B by $\rho''(\mathbf{U}(1))_t$. Since $\rho''(\mathbf{U}(1))$ is topologically isomorphic to $\mathbf{U}(1)$, there exists, by Remark 6 and by a theorem of Brown,[30] a normalized Borel section σ_q associated with ρ_q whose restriction to some neighborhood of 1 is a local group homomorphism. Thus, $\widetilde{b}\mapsto\rho''^{-1}(\sigma_q(\rho_B(\widetilde{b})))$ is a local group homomorphism of $\widetilde{B}$ into $\widetilde{E}$ and can be extended in a unique way to a continuous group homomorphism $\widetilde{\sigma}:\widetilde{B}\to\widetilde{E}$ (Ref. 31, Chap. IV, Theorem 3.1, and Ref. 29, TGI, § 11, Prop. 11). Moreover, $\widetilde{\sigma}$ is a section associated with $\widetilde{\rho}$ because $\widetilde{\rho}(\widetilde{\sigma}(\widetilde{b}))=\widetilde{b}$ for all $\widetilde{b}$ in some neighborhood of 1 in $\widetilde{B}$. The result follows immediately from the theorem of Brown quoted above. ■

Remark 7: Proposition 3 may also be proved directly following the same steps of the proof of Proposition 2. Lemmas 1–4 are still true once we have made the appropriate changes [required by the substitution of $\mathbf{L}_0$ with $\mathbf{SL}(2,\mathbf{C})$], and we have in addition $H_b^2(\mathbf{SL}(2,\mathbf{C}),\mathbf{U}(1)_t)=\{1\}$.

Let $\mathfrak{H}$ be a separable complex Hilbert space and let ρ_B be as above. As pointed out by Bargmann (Ref. 32, 3e), if $\widetilde{u}$ is any CUP-rep of B on $\mathbf{P}(\mathfrak{H})$, then $\widetilde{u}\circ\rho_B$ is a CUP-rep of $\widetilde{B}$ on $\mathbf{P}(\mathfrak{H})$. It follows from Proposition 3 and from Ref. 11, Remark 2' that there exists a CU-rep w of $\widetilde{B}$ on $\mathfrak{H}$ such that

$$\widetilde{u}\circ\rho_B=\Omega\circ w; \tag{III.1}$$

this implies

$$w(\mathrm{Ker}\rho_B)\subset\{\zeta\,\mathrm{Id}_{\mathfrak{H}}\,|\,\zeta\in\mathbf{U}(1)\}. \tag{III.2}$$

Conversely, if w is any CU-rep of $\widetilde{B}$ on $\mathfrak{H}$ satisfying (III.2), the mapping $\widetilde{u}$ defined by (III.1) is a CUP-rep of B on $\mathbf{P}(\mathfrak{H})$. In other words, $\widetilde{B}$ is a *representation group for* B and is its unique one up to topological group isomorphisms [Ref. 33, Sec. IV, (A_1)]. The analogy with $\mathbf{P}_0$ and its universal covering group is now complete.

APPENDIX

Let K be a topological vector group, i.e., an Abelian topological group which is topologically isomorphic to the additive group of a topological vector space. The group of 2-coboundaries $B_b^2(K,\mathbf{U}(1)_t)$ is contained in the subgroup $Z_{b,s}^2(K,\mathbf{U}(1)_t)$ of all symmetric elements of $Z_b^2(K,\mathbf{U}(1)_t)$; therefore,

$$H_{b,s}^2(K,\mathbf{U}(1)_t)=Z_{b,s}^2(K,\mathbf{U}(1)_t)/B_b^2(K,\mathbf{U}(1)_t)$$

is a subgroup of $H_b^2(K,\mathbf{U}(1)_t)$. We denote by $\mathscr{H}om(K,K;\mathbf{U}(1))$ the (pointwise) Abelian group of all continuous mappings of $K\times K$ into $\mathbf{U}(1)$ such that all the partial mappings they determine are (continuous) group homomorphisms of K into A. Then $\mathscr{H}om_a(K,K;\mathbf{U}(1))$ is the subgroup of all antisymmetric elements of $\mathscr{H}om(K,K;\mathbf{U}(1))$.

If K is Polish, the group $H_b^1(K,\mathbf{U}(1)_t)$ is trivially identified with $Z_c^1(K,\mathbf{U}(1)_t)$; we transport on $H_b^1(K,\mathbf{U}(1)_t)$ the topology of simple (i.e., pointwise) convergence (which is Hausdorff and compatible with the group structure) and write $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{S}}$ for the topological group so obtained. For each $k\in K$, we denote by e_k the continuous mapping $[h]\mapsto h(k)$ of $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{S}}$ into $\mathbf{U}(1)$. The topology of $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{S}}$ is then generated by the set $\{e_k^{-1}(U)\}$ [U open set of $\mathbf{U}(1)$; $k\in K$].

In the following proposition, we shall use the particular notation of Ref. 13, Proposition 1.

Proposition: Let G be a Polish group which is the topological semidirect product of a topological group S by a topological vector group K. Suppose there exists a topology $\mathfrak{T}$ on $H_b^1(K,\mathbf{U}(1)_t)$, finer than that of simple convergence, making it into a Polish group such that the operation $\hat{I}_*^1$ of S on it is topological. If then $H_b^1(S,H_b^1(K,\mathbf{U}(1)_t)_{\hat{I}_*^1})=\{1\}$, we have

$$H_b^2(G,\mathbf{U}(1)_t)\approx H_{b,s}^2(K,\mathbf{U}(1)_t)'\times\mathscr{H}om_a(K,K;\mathbf{U}(1))^S\times H_b^2(S,\mathbf{U}(1)_t),$$

where $H_{b,s}^2(K,\mathbf{U}(1)_t)'=H_b^2(K,\mathbf{U}(1)_t)'\cap H_{b,s}^2(K,\mathbf{U}(1)_t)$.

Proof: The result follows immediately from Ref. 25, Proposition 2 and from Ref. 13, Proposition 1 (iii) (cf. also Ref. 14, Remark 1) once we have identified $H_b^2(K,\mathbf{U}(1)_t)$ with $H_{lr}^2(K,\mathbf{U}(1)_t)$ (Ref. 14, Corollary 2) and made the following remarks:

(a) If $[\mu]\in H_b^2(K,\mathbf{U}(1)_t)'$, then $[\mu_s]\in H_{b,s}^2(K,\mathbf{U}(1)_t)'$ and $\mu_a\in\mathscr{H}om_a(K,K;\mathbf{U}(1))^S$, where μ_s and μ_a are defined by

$$\mu_s(k,k')=\mu(k,k')\mu(k',k)$$

and $\qquad (k,k' \text{ in } K)$

$$\mu_a(k,k')=\mu(k,k')\,\mu(k',k)^{-1}.$$

(b) Let $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{T}}$ be the group $H_b^1(K,\mathbf{U}(1)_t)$ endowed with the Polish topology $\mathfrak{T}$. The identity mapping of $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{T}}$ onto $H_b^1(K,\mathbf{U}(1)_t)$ is Borel (Ref. 34, TGIX, § 6, Props. 11 and 14), so that the closed sets of these two topological spaces generate the same Borel structure. Now, the mapping $h\in Z^1(S,H_b^1(K,\mathbf{U}(1)_t)_{\hat{I}_*^1})$ defined by Ref. 13, (4.10), is Borel for the topology of simple convergence because $H_b^1(K,\mathbf{U}(1)_t)_{\mathfrak{S}}$ is Lindelöf (Ref. 34, TGIX, Appendice I, Cor. to Prop. 1) and, for each open set U of $\mathbf{U}(1)$ and each $k\in K$, $h^{-1}(e_k^{-1}(U))$ is a Borel set of S. Thus h is Borel for $\mathfrak{T}$. ■

ACKNOWLEDGMENTS

I am indebted to M. Romerio and W.F. Wreszinski for discussions.

[1]R. Sachs, Phys. Rev. **128**, 2851–64 (1962).
[2]P.J. McCarthy, Proc. R. Soc. London Ser. A **330**, 517–35 (1972).
[3]P.J. McCarthy, Phys. Rev. Lett. **29**, 817–9 (1972).
[4]H. Bondi, M.G.J. Van Der Burg, and A.W.K. Metzner, Proc. R. Soc. London Ser. A **269**, 21–52 (1962).
[5]R. Sachs, Proc. R. Soc. London Ser. A **270**, 103–26 (1962).
[6]R. Penrose, Phys. Rev. Lett. **10**, 66–8 (1963).
[7]M. Crampin and P.J. McCarthy, Phys. Rev. Lett. **33**, 547–50 (1974).
[8]P.J. McCarthy, Proc. R. Soc. London Ser. A **333**, 317–36 (1973).
[9]P.J. McCarthy and M. Crampin, Proc. R. Soc. London Ser. A **335**, 301–11 (1973).
[10]A. Piard, Rep. Math. Phys. **11**, 279–83 (1977).
[11]U. Cattaneo, Rep. Math. Phys. **9**, 31–53 (1976).
[12a]P.J. McCarthy, "Projective Representations of the Asymptotic Symmetry Group of General Relativity," in *Proceedings of the 2nd International Colloquium on Group Theoretical Methods in Physics* (Nijmegen, June 1973), pp. A137–A146.
[12b]P.J. McCarthy, Proc. R. Soc. London Ser. A **358**, 141–71 (1977).
[13]U. Cattaneo and A. Janner, J. Math. Phys. **15**, 1155–65 (1974).
[14]U. Cattaneo, Rep. Math. Phys. **12**, 125–32 (1977).
[15]P.J. McCarthy, J. Math. Phys. **13**, 1837–42 (1972).
[16]A. Held, E.T. Newman, and R. Posadas, J. Math. Phys. **11**, 3145–54 (1970).
[17]N. Bourbaki, *Éléments de mathématique. Livre VI: Intégration*, ASI 1306 (Hermann, Paris, 1963), Chaps. 7,8.
[18]J. Dieudonné, *Treatise on Analysis*, (Academic, New York, 1972), Vol. III.
[19]N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, Translations of Mathematical Monographs, Vol. 22 (Am. Math. Soc. (Providence, R.I., 1968).
[20]V. Cantoni, Ph.D. Thesis, King's College, University of London (1967).
[21]U. Cattaneo and A. Janner, J. Math. Phys. **15**, 1166–73 (1974).
[22]K.R. Parthasarathy and K. Schmidt, *Positive Definite Kernels, Continuous Tensor Products, and Central Limit Theorems of Probability Theory*, Lecture Notes in Mathematics **272** (Springer-Verlag, Berlin–Heidelberg–New York, 1972).
[23]N. Bourbaki, *Éléments de mathématique. Livre* VI: *Intégration*, ASI 1175 (Hermann, Paris, 1965), 2ème éd., Chaps. 1–4.
[24]P.R. Halmos, *Measure Theory* Van Nostrand, Princeton, New Jersey, 1950).
[25]U. Cattaneo, Math. Ann. **239**, 1–6 (1979).
[26]E. Hewitt and K.A. Ross, *Abstract Harmonic Analysis* (Springer-Verlag, Berlin–Heidelberg–New York, 1963), Vol. I.
[27]C.C. Moore, Trans. Am. Math. Soc. **113**, 40–63 (1964).
[28]U. Cattaneo, "Continuous unitary projective representations of Polish groups: The BMS-group," in *Group Theoretical Methods in Physics*, Lecture Notes in Physics **50**, edited by A. Janner, T. Janssen, and M. Boon (Springer-Verlag, Berlin-Heidelberg-New York, 1976).
[29]N. Bourbaki, *Éléments de mathématique. Topologie générale I* (Hermann, Paris, 1971), nouvelle éd., Chaps. 1–4.
[30]L.G. Brown, Pacific J. Math. **39**, 71–8 (1971).
[31]G. Hochschild, *The Structure of Lie Groups* (Holden–Day, San Francisco, 1965).
[32]V. Bargmann, Ann. Math. **59**, 1–46 (1954).
[33]U. Cattaneo, J. Math. Phys. **19**, 452–60 (1978).
[34]N. Bourbaki, *Éléments de mathématique. Topologie générale II* (Hermann, Paris, 1974), nouvelle éd., Chaps. 5–10.

Deformations of gauge groups. Gravitation[a)]

J. Fang and C. Fronsdal

Department of Physics, University of California, Los Angeles, California 90024

(Received 29 November 1978)

This is a review, and an attempt at completion, of the "Gupta program," the ultimate goal of which is *either* to show that Einstein's theory of gravitation is the only self-consistent field theory of interacting, massless, spin-2 particles in flat space *or* to discover interesting alternatives. It is useful to notice that the gauge group of general relativity is a deformation (in a mathematically precise sense) of the gauge group associated with the massless, spin-2 free field. The uniqueness of Einstein's theory depends on the stability of its gauge group with respect to a class of differentiable deformations. A generalized Gupta program for massless fields of arbitrary spins is proposed.

I. HISTORY OF THE GUPTA PROGRAM

A. Prehistory

Let us recall, in very brief outline, the main features of Einstein's 1915 theory of gravity.[1] Space–time is regarded as a four-dimensional Riemannian space endowed with a differentiable metric g, specifiable in terms of local coordinates by components $g_{\mu\nu}$. The metric is interpreted as a dynamical variable, to be determined by boundary conditions and a principle of least action. The action is required to be the integral over all space–time of a "Lagrangian" density constructed locally in terms of the $g_{\mu\nu}$ and their first order derivatives. The total Lagrangian density of the world is a sum of two terms,

$$\mathscr{L}(x) = \mathscr{L}_g\,[g_{\mu\nu}(x)] + \mathscr{L}_m\,[g_{\mu\nu}(x),\phi(x),\cdots]. \qquad (1.1)$$

Here $\phi,\cdots$ stands for matter fields (or, in a particle description, some other appropriate variables) and $\mathscr{L}_m$ is the Lagrangian density that would have been complete if g would have been fixed *a priori*. This part is not unique although general coordinate covariance, the equivalence principle and other considerations (less compelling and mostly aesthetic) often lead to definite expressions for it.

The "pure gravitational" part $\mathscr{L}_g$ of the Lagrangian density is essentially unique[1]:

$$\kappa^2\mathscr{L}_g/\sqrt{-g} = -g^{\mu\nu}(\Gamma^{\alpha}_{\mu\beta}\,\Gamma^{\beta}_{\nu\alpha} - \Gamma^{\alpha}_{\mu\nu}\,\Gamma^{\beta}_{\alpha\beta}) - 2\lambda, \qquad (1.2)$$

where κ is a constant subsequently related to the universal strength of gravitation, the "cosmological constant" λ is a parameter with the dimension of inverse length, $g\equiv\det(g_{\mu\nu})$ and $\Gamma^{\alpha}_{\mu\nu}$ are the components of the metric connection. If $R_{\mu\nu\lambda}{}^{\rho}$ denote the components of the curvature tensor, then

$$R_{\mu\nu}\equiv R^{\;\lambda}_{\mu\lambda\nu}, \quad R\equiv g^{\mu\nu}R_{\mu\nu}. \qquad (1.3)$$

The field equations are

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}\,R + \lambda g_{\mu\nu} = -\tfrac{1}{2}\kappa^2 T_{\mu\nu}, \qquad (1.4)$$

where $T_{\mu\nu}$ are the components of the stress–energy tensor of matter in the field g.

In order to find the physical interpretation of this theory, it is necessary to postulate that accessible regions of space–time can be endowed with local coordinates in terms of which the components $g_{\mu\nu}$ are almost constant,[2]

$$g_{\mu\nu}(x) = \delta_{\mu\nu} + \kappa h_{\mu\nu}(x). \qquad (1.5)$$

Here $\delta_{00}=1$, $\delta_{11}=\delta_{22}=\delta_{33}=-1$, and $\delta_{\mu\nu}=0$ for $\mu\neq\nu$. The second term is regarded as small and treated as a perturbation. In an appropriate physical context (slow, massive particles) one finds that the component h_{00} is directly related to the Newtonian gravitational potential.[1]

B. Preparation

Strangely, the amazing progress of theoretical physics in the years after 1915 owed very little to the general theory of relativity. By the end of the thirties it began to be felt that this beautiful theory was being left behind by advances in quantum mechanics and field theory. The possibility of discovering a unified field theory of the kind long sought by Einstein seemed to become increasingly remote with the appearance of Lorentz invariant field theories of several new forms of matter. Finally, it was suggested that a unified theory of matter and gravitation could be achieved only at the high price of setting aside the geometrical interpretation of gravity. The first published proposal to view Einstein's theory as a Lorentz invariant theory of a field in flat space, in the manner of electrodynamics and Dirac's electron fields, seems to have been made in 1939 by Fierz and Pauli, in a paper[3] that we shall have frequent occasion to refer to.

Fierz and Pauli[3] begin by formulating the requirements for a self-consistent (or at least not obviously inconsistent) field theory of massive particles of spin s, $2s=0,1,2,\cdots$. Briefly, they demand that the field equations be derived from an action principle and that the solutions carry an irreducible representation of the Poincaré group[4] with mass m and spin s. This can be done, even in the case of free fields, only at the cost of introducing certain auxiliary fields. In the case $s=2$ the field is a second rank symmetric tensor (with respect to Lorentz transformations); no auxiliary fields are needed if this tensor field is permitted to have a nonzero trace (that is,

[a)]Work supported in part by the National Science Foundation.

one must resist the temptation to impose $h' = \delta^{\mu\nu}h_{\mu\nu} = 0$ from the beginning). The free field equations obtained by Fierz and Pauli can be written as follows,[5]

$$(\partial^2 + m^2)h_{\mu\nu} - \partial_\mu \partial^\lambda h_{\lambda\nu} - \partial_\nu \partial^\lambda h_{\mu\lambda} + \partial_\mu \partial_\nu h' + \delta_{\mu\nu}[\partial^\lambda \partial^\rho h_{\lambda\rho} - (\partial^2 + m^2)h'] = 0. \tag{1.6}$$

Here $\partial_\mu \equiv \partial/\partial x^\mu$, $\partial^2 \equiv \delta^{\mu\nu}\partial_\mu \partial_\nu \equiv \partial_\mu \partial^\mu$, $h' \equiv \delta^{\mu\nu}h_{\mu\nu}$.

It is interesting to dwell on the extent to which this result is unique. In order that a tensor field ϕ be associated with an irreducible representation of the Poincaré group it must be symmetric, traceless, and divergence free. This last requirement means that $\partial^\mu \phi_{\mu\nu} = 0$ and this condition must be built into the action principle. To achieve that end, one is forced to introduce auxiliary fields besides $\phi_{\mu\nu}$, and here ambiguities arise. Economy suggests that the simplest possibilities be tried first; namely either a single auxiliary vector field or a single auxiliary scalar field. Both attempts suceed, so long as $m \neq 0$, but the first fails to give a satisfactory theory in the limit $m \to 0$. The second alternative is by far the simplest procedure for obtaining a consistent theory for the massless case, besides it has a very natural extension to higher integral spin values.[6,7] Adopting this scheme, one obtains an action that is unique up to a constant parameter that can be absorbed into the auxiliary scalar field, ψ say. To obtain the form (1.6) of the field equations, one sets $h_{\mu\nu} = \phi_{\mu\nu} + \delta_{\mu\nu}\psi$ and chooses the parameter (i.e., the normalization of ψ) so as to obtain the simplest possible expression. To recover the general form it is enough to substitute

$$h_{\mu\nu} \to h_{\mu\nu} + c\delta_{\mu\nu}h', \tag{1.7}$$

where c is real and, evidently, $c \neq -\frac{1}{4}$. *Thus* (1.6) *is the simplest form among a family of equivalent wave equations related to each other by transformations of the form* (1.7).

Next Fierz and Pauli[3] passed with m to zero. In this case it is no longr true that the space of solutions of (1.6) carries an irreducible representation of the Poincaré group. What saves the situation is the appearance of a new phenomanon, characteristic of massless fields, gauge invariance. We rewrite Eq. (1.6) with $m = 0$,

$$\partial^2 h_{\mu\nu} - \partial_\mu \partial^\lambda h_{\lambda\nu} - \partial_\nu \partial^\lambda h_{\mu\lambda} + \partial_\mu \partial_\nu h' + \delta_{\mu\nu}(\partial^\lambda \partial^\rho h_{\lambda\rho} - \partial^2 h') = 0. \tag{1.8}$$

This equation is solved identically by all fields of the form

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \tag{1.9}$$

with no restriction on the ξ_μ besides differentiablity. In view of the symmetry of the Lagrangian operator, this statement is equivalent to the observation that the left-hand side of (1.8) is identically divergence free. Let us add a fixed, external source to Eq. (1.8),

$$\partial^2 h_{\mu\nu} - \partial_\mu \partial^\lambda h_{\lambda\nu} - \partial_\nu \partial^\lambda h_{\mu\lambda} + \partial_\mu \partial_\nu h' + \delta_{\mu\nu}(\partial^\lambda \partial^\rho h_{\lambda\rho} - \partial^2 h') = -\kappa t_{\mu\nu}. \tag{1.10}$$

Since the left-hand side is identically divergence free it is necessary to choose t so that

$$\partial^\mu t_{\mu\nu} = 0. \tag{1.11}$$

The fact that (1.9) solves (1.8) identically is now expressed by the statement that (1.10) is invariant under the gauge transformation

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu. \tag{1.12}$$

It was pointed out by Fierz[8] that, as a direct result of (1.11) and the invariance of (1.10) under (1.12), the outgoing spherical waves excited by a point source have helicity ± 2. This highly satisfactory result seems somewhat miraculous; if one employs an auxiliary vector field instead of an auxiliary scalar field, then no gauge invariance appears in the massless limit and the theory is devoid of physical meaning.

Finally, Fierz and Pauli[3] calculated the linear approximation to Einstein's field equation (1.4), taking $\lambda = 0$.[2] That is, they inserted (1.5)into (1.4) and retained the terms linear in κ. All expansions are regarded as formal series in κ, and in particular it is supposed that $T_{\mu\nu} = t_{\mu\nu} + \kappa t_{\mu\nu}^{(1)} + \cdots$. The result is in full agreement with Eq. (1.10). If (1.10) is repalced by an equivalent form by effecting a transformation of the type (1.7), then agreement is restored by appropriate modifying (1.5),

$$g_{\mu\nu} = \delta_{\mu\nu} + \kappa(h_{\mu\nu} + c\delta_{\mu\nu} h'). \tag{1.13}$$

C. Genesis

The miraculous agreement between (1.10) and the linear approximation to (1.4) did not dispel the impression that these are two entirely unrelated theories. On the one hand, Eq. (1.10) is much like Maxwell's equations for massless, spin-1 or Dirac's equation for massive, spin-$\frac{1}{2}$ fields; it belongs to particle physics and quantum field theory. Einstein's theory, on the other hand, is highly nonlinear and founded on entirely different principles. In 1940 Rosen[9] did advocate a nongeometrical interpretation of Einstein's theory, as a nonlinear theory of a massless, spin-2 field in the flat space of special relativity, but the idea was not well received at first. In 1952 Gupta[10] adopted this interpretation when he applied his indefinite metric quantization technique in the first attempt to reconcile Einstein's theory of gravitation with quantum theory.

The "Gupta program" saw the light of day in Gupta's 1954 paper[11]: "Gravitation and Electromagnetism." We quote the opening paragraph: "Although Einstein's theory of the gravitational field is the most widely accepted theory of gravitation, it is rather disconcerting to note that Einstein's theory appears to be strikingly different from the present theories of the electromagnetic field and the meson fields. In fact, in the formulation of fundamental physical laws we always seek for harmony in nature, and we intuitively expect that there should be some uniformity in the description of various fields in nature. Therefore Einstein himself and others have tried to construct a unified theory of the gravitational and the electromagnetic fields, while some other authors have tried to find a linear theory of the gravitational field in flat (Minkowskian) space analogous to other existing field theories."

In a nutshell, Gupta's idea is that the nonlinearities of Einstein's theory of the metric field can be understood without founding it on a geometrical interpretation. Here is the argument.[11] Adopting the point of view of Lorentz invariant, flat space field theory we have been led to Eq. (1.10), and we have discovered the integrability condition (1.11). The source t must be constructed from the other fields with which we wish our spin-2 field to interact. *The only naturally available symmetric, divergence free source is the total energy–momentum tensor of a closed system of fields*, such as the fields of Maxwell–Dirac electrodynamics. *But such a system ceases to be closed the moment it provides the source for our spin-2 field*, unless it includes the spin-2 field itself. Therefore, we must replace the external source t by $t + \delta_1 t$, where $\delta_1 t$ is the energy-momentum tensor of the spin-2 field. However, this is not enough; we insist on an action principle formulation, then the new interaction must be introduced by adding a new term to the action, and this leads to an additional term $\delta_2 t$ in the expression for the energy–momentum tensor. This process appears to be endless and to lead ineluctably to a very nonlinear theory. That is, any self-consistent theory of an interacting, massless, spin-2 field in flat space is just as nonlinear as Einstein's theory. Gupta's own conclusion is: "Hence, not only do Maxwell's theory of the electromagnetic field and Einstein's theory of the gravitational field have many similarities, but the dissimilarities between these fields are a necessary consequence of the difference in their spins."[11] This is a model of restraint, for the idea that presents itself to the reader is, of course, much more speculative and exciting! Here is our interpretation of Gupta's idea:

Definition 1: The program that aims to classify all physically acceptable sources $S_{\mu\nu}$ for the Fierz-Pauli equation (1.10), as a formal power series in κ,

$$S_{\mu\nu} = \kappa t_{\mu\nu} + \sum_{n=1}^{\infty} \kappa^n \delta_n t_{\mu\nu} \tag{1.14}$$

will be called the *Gupta program*. (Notation: $t_{\mu\nu}$ and $\delta_n t_{\mu\nu}$, $n = 0,1,\cdots$, are polynomials in the $h_{\mu\nu}$ and in their first and second order derivatives, with coefficients constructed from other fields. The term $t_{\mu\nu}$ is independent of h and $\delta_1 t_{\mu\nu}$ has no constant term.)

Later, a more restircted program will be proposed that avoids the mathematically ill-defined qualification "physically acceptable." It will also be necessary to introduce a notion of equivalence and to call for the classification of equivalence classes of sources.

The problem evidently calls for an iterative approach. Let f_0 be the Lagrangian density for the free spin-2 field,

$$f_0 = \tfrac{1}{2}(\partial^\mu h^{\nu\lambda})(\partial_\mu h_{\nu\lambda}) - (\partial_\mu h^{\mu\nu})(\partial^\lambda h_{\lambda\nu})$$

$$- \tfrac{1}{2}(\partial^\mu h')(\partial_\mu h') + (\partial_\mu h^{\mu\nu})(\partial_\nu h'). \tag{1.15}$$

Let f^m be the Lagrangian density for a closed system of fields (excluding h, and closed only so long as not interacting with h), and let t denote the corresponding energy–momentum tensor (i.e., any tensor of rank 2 that deserves this name by virtue of being symmetric and divergence free). In accordance with the requirement of physical acceptability we postulate an action principle. Take the total Lagrangian density to be a formal series.

$$f^m + f_0 - \kappa h^{\mu\nu} t_{\mu\nu} + \sum_{n=1}^{\infty} \kappa^n f_n \equiv f, \tag{1.16}$$

where f_k, $k = 0,1,\cdots$ are polynomials in the components of h and their first order derivatives. [As in (1.14), we distinguish between the two terms of order κ: f_1 is assumed to have no constant term and no term linear in h.] This gives the field equations

$$-\frac{\delta f_0}{\delta h^{\mu\nu}} = -\kappa t_{\mu\nu} - \sum_{n=1}^{\infty} \kappa^n \delta_n t_{\mu\nu} \tag{1.17}$$

$$\delta_n t_{\mu\nu} \equiv -\frac{\delta f_n}{\delta h^{\mu\nu}}. \tag{1.18}$$

The left-hand side of (1.17) coincides with the left-hand side of Eq. (1.10), and is identically divergence free; whence the consistency condition

$$\partial^\mu (t_{\mu\nu} + \delta_1 t_{\mu\nu} + \sum_{n=2}^{\infty} \kappa^{n-1} \delta_n t_{\mu\nu}) = 0. \tag{1.19}$$

Here the first term is of order κ, since t is divergence free in the limit $\kappa \to 0$, by virtue of the matter field equations. It follows that the second term must also be of order κ, and thus it must have the form

$$\partial^\mu \delta_1 t_{\mu\nu} = -A_{\nu,\alpha\beta}\left(\frac{\delta f_0}{\delta h_{\alpha\beta}}\right) - D_\nu{}^K\left(\frac{\delta f^m}{\delta \phi^K}\right). \tag{1.20}$$

Here $A_{\nu,\alpha\beta}$ and $D_\nu{}^K$ are first order differential operators (including nonderivative terms) and $\{\phi^K\}$, $K = 1,2,\cdots$ stands for the complete set of matter fields. Thus, in order to determine f_1, one writes down a very general ansatz for it, with undetermined parameters. One evaluates the left-hand side of (1.20) and restricts the parameters so as to give it the structure of the right-hand side. Using the field equations to lowest order one obtains an expression for $\partial^\mu t_{\mu\nu} + \partial^\mu \delta_1 t_{\mu\nu}$ of order κ. Finally, one seeks to have this cancelled by the divergence of $\kappa^2 \delta_2 t_{\mu\nu}$, which leads to a partial determination of f_2.

D. Growing pains

Kraichnan,[12] in 1955, made the bold suggestion that implementation of Gupta's procedure may lead directly and unambiguously to Einstein's theory. Important advance was made by Thirring in 1961 and by Feynman in 1962, but at this point it is convenient to depart from the chronological ordering of contributions and to discuss the work of Wyss,[13] in 1965. We shall review a part of his paper here, but his main contribution will be reserved for subsection E.

Wyss used a simple, field theoretical matter model—a massive, free, real scalar field. Thus

$$f^m = \tfrac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \tfrac{1}{2} m^2 \phi^2, \tag{1.21}$$

$$t_{\mu\nu} = (\partial_\mu \phi)(\partial_\nu \phi) - \delta_{\mu\nu} f^m. \tag{1.22}$$

Using the matter field equations one finds, to order κ,

$$\partial^\mu t_{\mu\nu} = \kappa \partial_\nu \phi \, [(\partial_\mu h')(\partial^\mu \phi) - 2\partial_\mu (h^{\mu\nu} \partial_\nu \phi)]. \tag{1.23}$$

Wyss proved,[13] that in order to cancel this, one needs both f_1 and $f_2 \neq 0$. Note that $\delta_1 t = -\delta f_1 / \delta h$ is divergence free to lowest order and can be interpreted as the energy–momentum tensor of the free field h. Wyss finds a unique expression for f_2, bilinear in $h_{\mu\nu}$ and bilinear in $\phi,\partial_\mu \phi$. He does not actually calculate f_1, but obtains a unique expression for the divergence of $\delta_1 t_{\mu\nu}$ to order κ. From this he tries to determine the entire structure of Einstein's theory by a beautiful argument that we shall examine later.

It is necessary to point out the need for a minor improvement in Wyss' treatment of f_1. First of all, he does not seem to have considered the possibility of the appearance of the second term on the right-hand side of Eq. (1.20). It is not inconceivable that, when such a term is envisaged, a more general expression for f_2 might be obtained. We shall disregard this possibility for the moment, though it may be important in the wider context of a generalized Gupta program.[14] Accepting Wyss' formula for f_2, and thus also that $D = 0$, one finds that A must effectively have the form

$$A_{\nu,\alpha\beta} = \tfrac{1}{2}(\partial_\alpha h_{\beta\nu} + \partial_\beta h_{\alpha\nu} - \partial_\nu h_{\alpha\beta}). \qquad (1.24)$$

The qualification "effectively" means "up to terms that make no contribution to the problem of consistency of the field equations to order κ^2." As we said, Wyss did not actually use (1.24) to find f_1, but Feynman[15] had already arrived at (1.20), with $D = 0$ and A given by (1.24), and he had shown that these conditions give for f_1 a unique expression, trilinear in the $h_{\mu\nu}$ and their first order derivatives. This result is too strong. In fact, *the expression for f_1 cannot be unique*. To see this, let us accept Feynman's expression for f_1; it agrees with the comparable term in Einstein's nonlinear Lagrangian if the metric is related to $h_{\mu\nu}$ by $g_{\mu\nu} = \delta_{\mu\nu} + \kappa h_{\mu\nu}$. Now make the following substitution:

$$h_{\mu\nu} \to h_{\mu\nu} + \kappa(C_1 h' h_{\mu\nu} + C_2 \delta^{\alpha\beta} h_{\alpha\mu} h_{\beta\nu} + C_3 \delta_{\mu\nu} h' h' + C_4 \delta_{\mu\nu} h^{\alpha\beta} h_{\alpha\beta}), \qquad (1.25)$$

where the C_i are constants of order κ^0. Evidently, the Lagrangian (1.16) retains its general structure under this replacement; the first three terms also retain their explicit form, but the expression for f_1 will be modified. Thus one concludes that f_1 must contain at least four arbitrary real parameters (barrier accidental cancellations that, in fact, do not occur). The origin of this contradiction is not hard to find, it is hidden in the qualification "effectively" of the generality of (1.24). This problem will be cleared up in Sec. II.

Huggins, in his 1962 thesis,[16] attempted to construct f_1 by requiring (as proposed by Gupta) that $\delta_1 t_{\mu\nu}$ be the energy–momentum tensor of the free, spin-2 field. He concluded that this is a highly ill-defined procedure that can succeed only when the answer is known in advance. The idea that the spin-2 field should be coupled to the energy density is attractive, but "the trouble is that we need an extra condition to define the energy."[16] Ambiguities were also emphasized by Weinberg.[17]

E. Coming of age

The work reviewed in the preceding subsection was concerned with the problem of consistency of the wave equations. This problem had its origin in the gauge invariance of the free Fierz–Pauli wave equation for h, but *invariance under the gauge transformations* (1.12) *is not respected by the interactions*. Nevertheless, as pointed out by Wyss,[13] gauge invariance is the key to an elegant solution of the problem.

Fierz and Pauli[3] had already pointed out that the coordinate free feature of Einstein's geometrical theory could be understood as a type of gauge invariance. The effect on the components of the metric field, of an infinitesimal, local coordinate transformation, is of the form

$$g_{\mu\nu} \to g_{\mu\nu} + \kappa(\partial_\mu \xi_\nu + \partial_\mu \xi_\nu - 2\Gamma_{\mu\nu}{}^\lambda \xi_\lambda), \qquad (1.26)$$

where ξ_μ are differentiable functions and $\Gamma_{\mu\nu}{}^\lambda$ are the components of the metric connection. If the field component $h_{\mu\nu}$ are defined by $g_{\mu\nu} = \delta_{\mu\nu} + \kappa h_{\mu\nu}$, then (1.26) is equivalent to the nonlinear transformation[18]

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - \kappa \xi_\lambda \tilde{g}^{\lambda\rho} \times (\partial_\nu h_{\mu\rho} + \partial_\mu h_{\nu\rho} - \partial_\rho h_{\mu\nu}), \qquad (1.27)$$

where $\tilde{g}^{\lambda\rho}$ is the formal series defined by $\tilde{g}^{\lambda\rho}(\delta_{\rho\mu} + \kappa h_{\rho\mu}) = \delta_\mu{}^\lambda$. The full nonlinear field equations of Einstein are invariant under (1.27), while the linear approximation is invariant under (1.12). Thus gauge invariance is modified but not lost; in fact, loss of gauge invariance would signal a change in the number of degrees of freedom of the theory, the very catastrophy that the Fierz–Pauli program set out to avoid.

Thirring, in 1961,[19] pointed out that the terms of the series (1.16) can be calculated from the requirement that the sum be invariant under (1.27). Of course, this seems to beg the question, but the idea of making use of gauge invariance bore fruit later. Wyss, in the 1965 paper already referred to,[13] made the startling suggestion that *the structure of the Lie algebra of infinitesimal gauge transformations may be fixea by the requirement of consistency of the field equations to lowest order in κ*. He tried to determine this structure on the basis of (1.20), with $D = 0$ and A given by (1.24). Unfortunately, these restrictions on D and on A are, if not unjustifiable, at least unjustified. In the next section we shall attempt a more precise treatment.

Boulware and Deser[20] analyzed the problem by the methods of perturbative quantum field theory and S-matrix theory. Of course, they recognized that no quantum theory of gravity exists, and that even the linearized theory has no S matrix. However, all amplitudes that they examined are classical, and it would be possible to rewrite every step of their development in terms of purely classical concepts. It is possible, if all the assumptions made were carefully sorted out, and all reference to quantum theory expurgated, that it would be found that Boulware and Deser[20] anticipated some of our results, but to do so is a project in itself. We prefer to deal with the classical problem directly, by methods that avoid quantum extensions that are irrelevant to it.

II. A MORE PRECISE FORMULATION

A. Motivation

The Gupta program, if it is formulated and carried out with precision, should lead to one or the other of two conceivable conclusions: (i) Einstein's nonlinear theory of gravitation is the only consistent, Lorentz invariant theory of an interacting, massless, spin-2 field in flat space; (ii) there exists one or more such theories, besides that of Einstein, and these are possible alternative theories of gravitation. What could be gained by establishing the truth of either statement? In view of the difficulties of reconciling the general theory of relativity with the theory of quantized fields, it would certainly be important to have in hand any viable alternative theory of gravitation. But to establish the uniqueness of Einstein's theory may be even more fruitful, for the proof (that is, the detailed reconstruction of the full nonlinear theory from the flat space viewpoint) might serve as a paradigm for the construction of other fundamental interactions. Consider the case of interacting, massless, spin-$\frac{3}{2}$ fields. No consistent theory was known until the recent advent of supersymmetry and supergravity.[21] Perhaps supergravity could have been found by aplying Gupta's idea to the case of massless, spin-$\frac{3}{2}$ fields;[22] that would almost inevitably have spawned the whole of supersymmetry theory. At the present time, generalized supergravity[23] is limited by the lack of development of theories of massless fields with spin higher than 2.

Thus, we propose that a Gupta program for each spin may be interesting.[7] Our purpose in the rest of this paper is to make quite explicit the assumptions that are needed to carry it off in the spin-2 case and to arrive at a formula that is abstract enough to be applicable to other spins.

B. Eliminating the matter model

It is clearly not quite satisfying to have to base the reconstruction of the nonlinear theory on a specific model of matter. If it is true, as seems generally to be believed, that the result is independent of the choice of matter model, then it should be possible to eliminate this crutch. Therefore we propose a "restricted Gupta program."

Definition 2: The Gupta program (Definition 1, Sec. I C), restricted by the requirement that the sources $S_{\mu\nu}(x)$ be constructed entirely from the tensor $h_{\mu\nu}(x)$ and its first and second order derivatives, will be called the *restricted Gupta program*.[24]

We do not attempt to prove that an eventual completion of the restricted program automatically solves the general problem. However, it will be shown that the restricted program leads to Einstein's theory just as comfortably (and more rigorously) than previous arguments that relied on specific matter models. *A priori*, it could not have been known whether any nontrivial self-interaction exists, and we suspect that it does not in the spin-$\frac{3}{2}$ case;[22] nevertheless we propose that the restricted program is a logical, and even necessary, starting point.

C. Determination of f_1

In the absence of matter, the problem of consistency of the field equations

$$\frac{\delta f_0}{\delta h^{\mu\nu}} = \sum_{n=1}^{\infty} \kappa^n \delta_n t_{\mu\nu}, \tag{2.1}$$

to order κ, reduces to the requirement that $\partial^{\mu}\delta_1 t_{\mu\nu}$ vanish to order κ; thus, identically,[25]

$$\partial^{\mu}\delta_1 t_{\mu\nu} = -A_{\nu,\alpha\beta}\left(\frac{\delta f_0}{\delta h_{\alpha\beta}}\right). \tag{2.2}$$

In order to avoid any fortuitious assumptions on the structure of $A_{\nu,\alpha\beta}$ we prefer to replace (2.2) by

$$\left[\frac{\delta f_0}{\delta h_{\alpha\beta}} = 0\right] \Rightarrow [\partial^{\mu}\delta_1 t_{\mu\nu} = 0]. \tag{2.3}$$

We must use this requirement to determine f_1. Recall that f_1 is a polynomial in the $h_{\mu\nu}$ and their first order derivatives, with no constant term. A linear term will have the same effect as a cosmological term and will be excluded. Any second order term that is consistent with (2.3) can be absorbed into f_0 by means of a transformation of the type (1.7) and may thus be ignored. The result that we shall prove is limited by the restriction of f_1 to a homogeneous polynomial of order 3, but it is unlikely that any higher order polynomial can satisfy (2.3). To state our result we need the

Definition 3: Two formal series of the type

$$f(x) = \sum_{i=0}^{\infty} \kappa^n f_n(x), \tag{2.4}$$

where $f_n(x)$ is a polynomial in the $h_{\mu\nu}(x)$ and their first order derivatives, are said to be equivalent to order κ^n if one can be made to agree with the other up to terms of order κ^n, and up to an exact divergence, by making substitutions of the form

$$\kappa \to c\kappa, \quad h_{\mu\nu} \to h_{\mu\nu} + \sum_{i=1}^{\infty} \kappa^i H^i_{\mu\nu}, \tag{2.5}$$

where $c \neq 0$ and $H^i_{\mu\nu}$ is a polynomial in the $h_{\mu\nu}$ (with no constant terms and no derivatives).

Theorem 4: If f_0 is given by (1.15) and f_1 is a homogeneous polynomial of order 3, then Eq. (2.3) is satisfied if and only if f is equivalent to order κ *either* to the series defined by expanding Einstein's Lagrangian [with vanishing cosmological constant[2] and with the identification (1.5)] *or* to the series with $f_1 = 0$.

The proof is by direct computation. We begin with the most general ansatz for f_1 (ignoring exact divergences), calculate $\partial^{\mu}\delta_1 t_{\mu\nu}$, eliminate all terms that contains $\partial^{\mu}\partial_{\mu}$ by means of the field equations, and require that the resulting expression vanish identically. This leaves five undetermined constants. Finally, we verify that four of these can be adjusted by a substitution of the type (1.25) to turn f_1 into a constant multiple of the corresponding term in the expansion of Einstein's Lagrangian. (In other words, to order κ and in the absence of matter, Gupta's prescription for $\delta_1 t_{\mu\nu}$ is precise enough, any reasonable expression for the energy–momentum tensor will do, so long as there exists an f_1 such that $\delta f_1/\delta h^{\mu\nu} = -\delta_1 t_{\mu\nu}$.)

If f_1 is a homogeneous polynomial of order 3, then there is no loss of generality in adopting the form that appears in the expansion of Einstein's Lagrangian. This result was obtained without relying on a coupling to matter. It means that the choice (1.24) for the operators $A_{\nu,\alpha\beta}$, made by Feynman and Wyss and henceforth also by us, is unique up to equivalence.

D. Structure of the gauge algebra

It is evident that the Lagrangian that corresponds to the formal series (2.4) is not invariant under the gauge transformations (1.12) of the free field theory, since $\partial^{\mu}(\delta f/\delta h_{\mu\nu})$ does not vanish identically to order κ. We know that the complete Lagrangian must admit a gauge algebra, and try to determine the transformation law to order κ. By (1.18) and (2.2),

$$\partial^{\mu}\left(\frac{\delta f_1}{\delta h^{\mu\nu}}\right) + A_{\nu,\alpha\beta}\left(\frac{\delta f_0}{\delta h_{\alpha\beta}}\right) = 0 \tag{2.6}$$

or, to order κ,

$$\int (\partial_{\mu}\xi_{\nu} - \kappa\xi^{\lambda}\mathbf{A}_{\lambda,\mu\nu})\left(\frac{\delta f}{\delta h_{\mu\nu}}\right) dx = 0. \tag{2.7}$$

Here ξ is any differentiable vector field. From (2.7) it follows that the action is invariant to order κ under the infinitesimal transformations generated by

$$[\mathscr{L}_1(\xi,P)h]_{\mu\nu} = \partial_{\mu}\xi_{\nu} + \partial_{\nu}\xi_{\mu} + \kappa Q_{\mu\nu}, \tag{2.8}$$

where $Q_{\mu\nu}$ is determined by (2.7) to be of the form

$$Q_{\mu\nu} = -2A_{\lambda,\mu\nu}\xi^{\lambda} + \partial_{\mu}P_{\nu} + \partial_{\nu}P_{\mu}. \tag{2.9}$$

Here P is an arbitrary, differential vector field, both ξ and P may depend on h.

The idea of Wyss is that any complete action, satisfying the consistency conditions to all order in κ, must be invariant under a Lie algebra of infinitesimal transformations. The problem is whether the structure of this Lie algebra is determined by (2.8). Evidently, the question is too vague; we sharpen it by the following

Postulate 5: (i) The Lie algebra of infinitesimal transformations that we seek is defined by a formal power series in κ, given to order κ by (2.8), each term being determined by ξ and h. (ii) Its structure is independent of h.

The first part limits our attention to an algebra "of the same dimension" as the gauge algebra of the free field theory, without implying that it is in any sense maximal. The second part assures the gauge algebra of an existence of its own, independently of h. It is easy to see that this postulate implies the following: ξ is independent of h, and P is a fixed, bilinear, local function in ξ,h.[26] The operator $\mathscr{L}_1(\xi,P)$ is thus determined by ξ any may henceforth be denoted $\mathscr{L}_1(\xi)$.

Equations (2.8) and (2.9) give

$$[\mathscr{L}_1(\eta),\mathscr{L}_1(\xi)] = \kappa\mathscr{L}_1(\{\eta,\xi\}) + \kappa^2\mathscr{K}(\eta,\xi), \tag{2.10}$$

where the bracket $\{\cdot,\cdot\}$ is defined by

$$\{\eta,\xi\}_{\mu} \equiv P_{\mu}(\partial\xi,\eta) + \xi^{\lambda}\partial_{\mu}\eta_{\lambda} - (\xi\leftrightarrow\eta). \tag{2.11}$$

The requirement of consistency of the field equations, to higher orders in κ, produces corrections of order $\kappa^2, \kappa^3,\cdots$ to (2.8) and (2.9) and to (2.10).[27] Thus, (2.8) is replaced by a formal power series in κ that we denote $\mathscr{L}^{\kappa}(\xi)h$. The space $\{\mathscr{L}^{\kappa}(\xi)\}$ (ξ in the space N of differentiable vector fields on Minkowski space) has a Lie algebra structure with respect to the commutator of (formal series of) operators. We can write

$$[\mathscr{L}^{\kappa}(\eta),\mathscr{L}^{\kappa}(\xi)] = \kappa\mathscr{L}^{\kappa}(\{\eta,\xi\}^{\kappa}). \tag{2.12}$$

Here $\{\eta,\xi\}^{\kappa}$ is a formal power series in κ,

$$\{\eta,\xi\}^{\kappa} = \{\eta,\xi\} + \sum_{r=1}^{\infty} \kappa^r C_r(\eta,\xi), \tag{2.13}$$

where C_r is an antisymmetric, bidifferential operator. The bracket $\{\cdot,\cdot\}^{\kappa}$ plays the role of structure tensor for the gauge algebra.

Since we have a Lie algebra structure on the space $\{\mathscr{L}^{\kappa}(\xi)\}$, the bracket (2.13) must satisfy the Jacobi identity.

$$\sum_{\text{cycl.}} \{\psi,\{\eta,\xi\}^{\kappa}\}^{\kappa} = 0 \tag{2.14}$$

to every order in κ. To lowest order in κ this means that the bracket $\{\cdot,\cdot\}$ defined by (2.11) must satisfy the Jacobi identity:

$$\sum_{\text{cycl.}} \{\psi,\{\eta,\xi\}\} = 0. \tag{2.15}$$

In other words, the bracket $\{\cdot,\cdot\}$ confers on N the structure of a Lie algebra. The exact gauge algebra, defined by (2.13), is a formal deformation of this Lie algebra. Put more simply: $\{\cdot,\cdot\}^{\kappa}$ is a formal deformation of $\{\cdot,\cdot\}$.

Let us now study the bracket $\{\cdot,\cdot\}$. Write

$$P(h,\xi) = \sum_{r=0}^{n} P^r(h,\xi), \tag{2.16}$$

where P^r is a bidifferential operator of total order r. Then (2.11) can be written

$$\{\eta,\xi\} = \sum_{r=0}^{n} E^r(\eta,\xi), \tag{2.17}$$

where E^r is a bidifferential operator of total order $r+1$. The Jacobi identity (2.15) gives

$$\sum_{\text{cycl.}} \sum_{r+s=t} E^r(\psi,E^s(\eta,\xi)) = 0, \quad t = 0,1,\cdots. \tag{2.18}$$

This implies, in particular, that the bracket defined by

$$[\eta,\xi] \equiv E^0(\eta,\xi) \tag{2.19}$$

satisfies the Jacobi identity and defines a structure of Lie algebra on N; thus one sees that $\{\cdot,\cdot\}$ is a formal deformation of $[\cdot,\cdot]$.[28] This reduces the problem to the study of $[\cdot,\cdot]$ and its formal deformations.

The most general form for P^0 is given by[29]

$$[P^0(h,\xi)]_{\mu} = ch_{\mu\nu}\xi^{\nu}, \quad c \text{ in } \mathbb{R}. \tag{2.20}$$

We insert this expression into (2.11) and determine the first term $E^0(\eta,\xi)$ in (2.17). Direct calculation shows that (2.19) satisfies the Jacobi identity if and only if $c = 1$, in which case

$$[\eta,\xi]_{\mu} = \eta^{\lambda}\partial_{\lambda}\xi_{\mu} - \xi^{\lambda}\partial_{\lambda}\eta_{\mu}. \tag{2.21}$$

This is just the familiar Lie bracket for vector fields; thus we have proved

Theorem 6: The gauge algebra defined by (2.12) is a formal deformation of the usual Lie algebra of vector fields on Minkowski space; that is, $\{\cdot,\cdot\}^{\kappa}$ is defined by a formal power series,

$$\{\eta,\xi\}^{\kappa} = [\eta,\xi] + \sum_{r=1}^{\infty} \kappa^r D_r(\eta,\xi),$$

where the D_r are antisymmetric, bidifferential operators.

If the gauge algebra is precisely the usual Lie algebra of vector fields, then the exact Lagrangian is invariant under arbitrary, infinitesimal coordinate transformations. In this case, if one makes the usual assumption that no second order derivatives occur in the Lagrangian, one is led without fail to Einstein's nonlinear theory.

The deformation may be trivial; that is, an isomorphism.

Definition 7[30]: A deformation $\{\cdot,\cdot\}^\kappa$ of $[\cdot,\cdot]$ is called trivial if there exists a formal power series $T^\kappa = \text{id}. + \Sigma_{r=1}^\infty \kappa^r T_r$, where T_r are endomorphisms of the space N of vector fields, such that $T^\kappa\{\eta,\xi\}^\kappa = [T^\kappa\eta, T^\kappa\xi]$.

Evidently, if the deformation is trivial, then Einstein's theory may be recovered by a simple change of variables, given by $g^\kappa(\eta,\xi) \equiv g(T^\kappa\eta, T^\kappa\xi)$.

Lichnerowicz[31] has proved that *every differential deformation of* $[\cdot,\cdot]$ *is trivial*. [In fact, differentiably trivial; that is, one can restrict each T_r to be a differential operator.] That is, no nonisomorphic deformations exist. We conclude that *the restricted Gupta program leads unambiguously to Einstein's theory of gravitation*.

III. SUMMARY

In order to arrive at a precise mathematical problem we imposed a number of restrictions on the original Gupta program (Definition 1); namely:

(i) The restricted Gupta program (Definition 2) is concerned exclusively with the possible self-interactions of the spin-2 field. The self-interaction was assumed to be given by a power series in the coupling constant, and the first order term was assumed to be a homogeneous polynomial of degree three in the components of the spin-2 field and its first order derivatives. It was found (Theorem 4) that this first order term is unique up to equivalence (Definition 3).

(ii) The problem of calculating the exact form of the interaction was reduced to the problem of determining the structure of the gauge algebra of the interacting field theory (suggestion of Wyss). This structure was assumed (Postulate 5) to be independent of the spin-2 field.

It was shown that the gauge algebra is necessarily a formal deformation of the usual Lie algebra of vector fields on Minkowski space. But all such deformations are trivial (Licherowicz[31]); this means that the interacting theory can be transformed, by a redefinition of the field, to coincide with Einstein's theory of gravitation.

It is possible that the uniqueness of Einstein's theory of gravitation is provable in a much wider context than contemplated in our work. To begin with, the results ought to be improved by enlarging the concept of equivalence; one might try to modify Definition 3 by allowing the $H^{\;i}_{\mu\nu}$ in Eq. (2.5) to be polynomials in the $h_{\mu\nu}$ and in their derivatives of all orders. The physical interpretation of a theory with a Lagrangian that contains higher derivatives is less clear; but one should at least allow all Lagrangians belonging to any equivalence class that contains one conventionally acceptable member. The Yang–Mills type of approach to gravitation suggests a possible generalization of Einstein's Lagrangian; apparently it is not known whether or not this type of generalization is trivial. To attack the problem on a still deeper level, one could ask for a classification of all possible deformation of the original, Abelian gauge group of the linear theory, and then study the Lagrangian models associated with each.

It is hoped that our discussion of the Gupta program (for spin-2) may stimulate facilitate the search for theories of interacting, massless fields of spins $\frac{3}{2},\frac{5}{2},3,\cdots$. Perhaps this will lead to an understanding of the role played by the enigmatic neutrino.

ACKNOWLEDGMENTS

We wish to thank M. Flato for many stimulating conversations and helpful suggestions, L. Urrutia for numerous discussions, and D. Sternheimer for criticism of the manuscript.

[1]A. Einstein, Preuss. Akad. Wiss. Berlin, sitzber., 799 (1915); Ann. Phys. **49**, 769 (1916).

[2]The first terms on the right-hand side of Eq. (1.5) are the components of the metric tensor of the flat, Minkowski space–time. This is appropriate if λ, the cosmological constant in (1.2), vanishes; for then $g_{\mu\nu} = \delta_{\mu\nu}$ solves Eq. (1.4) to order κ^0. If $\lambda \neq 0$ it is necessary to replace $\delta_{\mu\nu}$ by the components of the de Sitter metric. The Fierz–Pauli program for this case was carried out for integer spins. C. Fronsdal, Phys. Rev. D **12**, 3819 (1975); J. Fang and C. Fronsdal, Lett. Math. Phys. **2**, 391 (1978); C. Fronsdal, UCLA preprint, UCLA/78/TEP/23, 1978.

[3]M. Fierz and W. Pauli, Proc. R. Soc. A **173**, 211 (1939).

[4]These are not the terms used in Ref. 3. Fierz and Pauli demanded positivity of the energy and irreducibility of the tensor field with respect to the Lorentz group.

[5]With the exception of Sec. I A we shall, throughout this paper, use the following conventions. The coordinate system of space–time (flat, Minkowski space) is fixed up to a Lorentz transformation. The components $\delta_{\mu\nu}$ of the metric are $+1$ for $\mu = \nu = 0$, -1 for $\mu = \nu = 1,2,3$ and zero otherwise, and $\delta^{\mu\nu} = \delta_{\mu\nu}$. Indices are raised and lowered exclusively by means of $\delta^{\mu\nu}$ and $\delta_{\mu\nu}$ and no distinction is made between vectors and covectors. The prime on a tensor indicates contraction on a pair of indices, for example $h' \equiv h^\mu{}_\mu \equiv \delta^{\mu\nu} h_{\mu\nu}$.

[6]L.P.S. Singh and C.R. Hagen, Phys. Rev. D **9**, 898 (1974).
[7]C. Fronsdal, Phys. Rev. D **18**, 3624 (1978).
[8]M. Fierz, Helv. Phys. Acta **13**, 45 (1940).
[9]N. Rosen, Phys. Rev. **57**, 147 (1940). See also A. Papapetrou, Proc. Roy. Irish Acad. **52A**, 11 (1947) and other papers cited there.
[10]S.N. Gupta, Proc. Phys. Soc. A **65**, 161, 608 (1952).
[11]S.N. Gupta, Phys. Rev. **96**, 1683 (1954).
[12]I. Kraichnan, Phys. Rev. **98**, 1118 (1955). This paper gives a very attractive algorithm that leads directly from the linear theory to Einstein's theory. See also S. Deser, J. Gen. Rel. Grav. **1**, 9 (1970), where another algorithm is proposed.
[13]W. Wyss, Helv. Phys. Acta **38**, 469 (1965).
[14]We have in mind the case of spin-$\frac{3}{2}$.
[15]R.P. Feynman, Lectures given at California Institute of Technology in 1962–1963.
[16]E.R. Huggins, Ph.D. Thesis, California Institute of Technology, 1962.
[17]S. Weinberg, Phys. Rev. **138**, B988 (1965).

[18]In view of our conventions (Ref. 5), we may refer to the set $\{\xi_\mu\}$, $\mu = 0,1,2,3$ as the components of a vector field. These conventions require that we use a different symbol ($\tilde{g}^{\mu\nu}$) for the contravarint components of g (usually denoted $g^{\mu\nu}$).

[19]W. Thirring, Ann. Phys. (N.Y.) **16**, 96 (1961). See also Fortschr. Physik **7**, 79 (1959).

[20]D.G. Boulware and S. Deser, Ann. Phys. **89**, 193 (1974).

[21]J. Wess and B. Zumino, Nucl. Phys. B **70**, 39 (1974); D.Z. Freedman, P. van Nieuwenhuizen, and S. Ferrara, Phys. Rev. D **13**, 3214 (1976); S. Deser and B. Zumino, Phys. Lett. B **62**, 325 (1976).

[22]A step in this direction has been made by M.T. Grisaru and H.N. Pendleton, Phys. Lett. B **67**, 323 (1977).

[23]S. Ferrara and P. van Nieuwenhuizen, Phys. Rev. Lett. **37**, 1669 (1976).

[24]We also insist on an action principle formulation; the action will be determined by the source up to a divergence and may be so chosen that only first order derivations appear.

[25]Since no matter fields are included, the second term in (1.20) does not appear.

[26]In principle, P may contain a term linear in ξ and independent of h, but this is of no consequence.

[27]Therefore, one would not be justified in requiring that the last term of (2.10) vanish. See, however, Ref. 29.

[28]If (2.18) is satisfied, then one can insert a factor σ^r under the summation in (2.17) without effecting the validity of the Jacobi identity.

[29]Wyss considered this as the most general expression for $P(h,\xi)$ and fixed c by requiring that the last term in (2.10) vanish. See Ref. 27.

[30]See M. Flato, A. Lichnerowicz, and D. Sternheimer, Compos. Math. **31**, 47 (1975); F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, Ann. Phys. **111**, 61, 111 (1978).

[31]A. Lichnerowicz, C.R. Acad. Sci. Paris **281**, 507 (1975); Comment. Math. Helv. **39**, 343 (1976).

Stationary charged C-metric

Hamid Farhoosh and Robert L. Zimmerman

Department of Physics, Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403
(Received 24 August 1978; revised manuscript received 27 November 1978)

The physical properties of the stationary charged C-metric are fully investigated. Tetrad components of curvature and Maxwell tensors are studied in a Bondi–Sachs coordinate system to show that this solution represents a uniformly accelerating and rotating charged particle with magnetic monopole and NUT parameter. The physical quantities—news function, mass loss, mass, charge, and multipole moments—are calculated. The different mass loss definitions are compared and it is shown that the definition given by Bondi is the more appropriate one. It is also shown that the magnetic monopole in the presence of rotation affects the electric charge.

I. INTRODUCTION

The interest in gravitational radiation has not been exhausted by the recent discovery of exact solutions. Most of the exact solution have been motivated by mathematical simplicity and do not correspond to physically interesting sources. The fact that energy acts as the source for gravitation makes the study of exact solutions very interesting and very difficult to solve. Gravitational radiation carries energy and consequently acts back on the field equations to serve as part of its own source and makes the field equations highly nonlinear. Linearization of the solutions fails to adequately describe the coupling between the radiation and the source, and consequently necessitates an exact solution.

A particularly interesting class of exact solutions is the type D vacuum metrics, which contain the familiar Kerr and Schwarzschild solutions. This class also contains the so-called charged C metric which has not been completely investigated.

Kinnersley and Walker[1] have shown that the charged C metric represents a uniformly accelerating Schwarzschild-like charged particle. Plebanski and Demianski[2] have generalized this solution to include rotation. This solution will be referred to as the stationary charged c metric. It is the main purpose of this paper to make a further analysis of this solution, and to obtain all the physically interesting quantities.

In all investigations of general relativity the work can be simplified by a suitable choice of coordinate systems. The mathematical solutions for a uniformly accelerating particle are simplified in an accelerated coordinate system where the particle is at rest. This is the case for both the static and stationary charged C-metrics. In this time-independent coordinate system, the line element can be obtained for a solution of Einstein–Maxwell equations. This line element is time independent, and only one of the tetrad components of curvature tensor is nonzero, i.e. ψ_2. It is in this coordinate system that mathematical simplicity is achieved and the formal properties of the metric are most easily studied, such as horizons, Killing vectors, etc. However, interesting physical properties such as the presence of radiation, mass loss, momentum, etc., can not be readily identified. In this paper we are not interested in studying the formal properties of the geometry but would like to explore physical properties of the solution. Therefore, we will transform the line element to a Bondi–Metzner–Sachs (BMS) coordinate system, which will be referred to as the inertial coordinates. It is in this coordinate system that radiation, mass loss, and the rest of the physically interesting quantities can be readily seen. In Sec. II the transformation to this coordinate system is constructed. Applying this transformation to the tetrad components of the Riemann and Maxwell tensors, from now on referred to as the Riemann and Maxwell scalars, we obtain these scalars in the BMS coordinates, which then can be interpreted, and consequently give the physical properties of the line element. However, it is not necessary to construct and make explicit the coordinate transformation in order to express the Riemann and Maxwell scalars in the inertial system for the rotating case. It is much easier and more constructive to derive these scalars using the asymptotic field equations along with some knowledge of the nonrotating results. Following this method, we obtain the Reimann and Maxwell scalars along with the news function in the inertial coordinate system. This is done in Sec. III.

In Sec. IV these results are discussd. The news function, mass loss, charge, linear and angular momenta, quadrupole and dipole moments are derived for small acceleration. It is shown that among the different definitions given for mass loss, Bondi's definition is the most appropriate one. It is also shown that in the presence of rotation, magnetic charge would affect the electric charge. To lowest order in the acceleration, the mass loss is proportional to A^2m^2, where A is the acceleration and m is the mass.

II. THE NONROTATING CASE

To explore the properties of a rotating, uniformly accelerating particle, we consider the most general form of type D vacuum metric. Plebanski and Demianski[2] have shown that this line element can be expressed as

$$ds^2 = \frac{1}{(p+q)^2}\left(\frac{Q}{1+(pq)^2}(d\tau - p^2 d\sigma)^2 - \frac{1+(pq)^2}{Q}dq^2\right.$$

$$-\frac{P}{1+(pq)^2}(d\sigma+q^2d\tau)^2-\frac{1+(pq)^2}{P}dp^2\Big), \quad (2.1)$$

where

$$P=-g_0^2+\gamma_0+2n_0p-\epsilon_0p^2+2m_0p^3-(e_0^2+\gamma_0)p^4, \quad (2.2)$$

$$Q=g_0^2-\gamma_0+2n_0q+\epsilon_0q^2+2m_0q^3+(e_0^2+\gamma_0)q^4.$$

The meaning of the parameters introduced in these equations will be made clear later. The nonvanishing tetrad components of the curvature and Maxwell tensors are given by

$$\phi_1=\frac{1}{\sqrt{2}}(e_0+ig_0)\frac{(p+q)^2}{(1-ipq)^2}, \quad (2.3)$$

$$\psi_2=(m_0+in_0)\frac{(p+q)^3}{(1-ipq)^3}-(e_0^2+g_0^2)$$

$$\times\frac{(p+q)^3}{(1-ipq)^3}\frac{p-q}{1+ipq}. \quad (2.4)$$

It is shown in the Appendix that performing a coordinate transformation to the null coordinates, the line element in (2.1) transforms to

$$ds^2=A^2r^2W^{-1}(Fa-a^2A^2y^4G)du^2+2dudr+2Ar^2dudx$$

$$-2aA^2r^2W^{-1}(Fx^2+Gy^2)dud\varphi-2ax^2drd\varphi$$

$$-r^2G^{-1}Wdx^2-2aAx^2r^2dxd\varphi$$

$$-r^2W^{-1}(G-a^2A^2x^4F)d\varphi^2, \quad (2.5)$$

and Eqs. (2.3) and (2.4) become

$$\phi_1=\frac{\bar e}{\sqrt{2}}\frac{1}{(1+iaAx^2)^2}r^{-2}+o(r^{-3}), \quad (2.6)$$

$$\psi_2=-\frac{1}{(1+iaAx^2)^3}Mr^{-3}+o(r^{-4}), \quad (2.7)$$

where

$$M\equiv m+iaAn+\frac{2A|\bar e|^2x}{1-iaAx^2}, \quad (2.8)$$

$$G(x)=\gamma-2Anx-\epsilon x^2-2Amx^3-(|\bar e|^2+a^2\gamma)A^2x^4, \quad (2.9)$$

$$F(y)=-\gamma-2Any+\epsilon y^2-2Amy^3+(|\bar e|^2+a^2\gamma)A^2y^4,$$

$$W\equiv 1+a^2A^2x^2y^2, \quad (2.10)$$

$$\gamma\equiv\frac{1}{1+a^2A^2},\quad \epsilon\equiv\frac{1-a^2A^2}{1+a^2A^2}, \quad (2.11)$$

$$\bar e\equiv e+ig,\quad |\bar e|^2\equiv e^2+g^2. \quad (2.12)$$

In these equations m is the mass, n is the NUT parameter, e and g are the electric and magnetic monopoles respectively, a is the rotation parameter,[2] and A is the acceleration.[1]

As can be seen, everything in this coordinate system is time independent. Also, there is only one Riemann scalar, ψ_2, similar to the non radiating Schwarzschild and Kerr solutions. The reason for this simplicity is that this is a uniformly accelerated coordinate system in which the particle is at rest. In order to identify the radiation and the physical properties of the line element, one has to go to the nonaccelerating BMS coordinate system.[3]

In order to transform to the inertial coordinate system we assume the transformation can be expanded in powers of r. Constructing this transformation for the nonrotating case, we use these results to serve as a guide for the rotating case. Guided by this information we solve the asymptotic field equations and obtain the rotating results. In principal, the coordinate transformation to BMS coordinates could be constructed for the rotating case; however, it is easier and more constructive to use the asymptotic field equations. The solution of these asymptotic equations gives us the Riemann and Maxwell scalars along with the news function, which otherwise could not have been easily obtained.

Let us now consider the transformation to inertial coordinates for the nonrotating case. This transformation has been discussed by Kinnersley and Walker. The nonrotating case is obtained by setting $a=0$ in Eq. (2.5),

$$ds^2=-A^2r^2G(x-A^{-1}r^{-1})du^2+2dudr+2Ar^2dudx$$

$$-r^2G^{-1}(x)dx^2-r^2G(x)d\varphi^2. \quad (2.13)$$

In order to find an inertial coordinate system, we assume the coordinate transformation can be expanded in powers of r as follows:

$$\bar u={}^0U(u,x)+\frac{{}^1U(u,x)}{r}+o(r^{-2}), \quad (2.14a)$$

$$\bar r={}^0R(u,x)r+{}^1R(u,x)+o(r^{-1}), \quad (2.14b)$$

$$\bar x={}^0X(u,x)+\frac{{}^1X(u,x)}{r}+o(r^{-2}), \quad (2.14c)$$

$$\bar\varphi=\varphi. \quad (2.14d)$$

Imposing the asymptotic flatness conditions (Newman–Unti conditions)[4]:

$$g^{\bar u\bar r}=1+o(\bar r^{-1}), \quad (2.15a)$$

$$g^{\bar r\bar r}=o(1), \quad (2.15b)$$

$$g^{\bar r\bar x}=o(\bar r^{-2}), \quad (2.15c)$$

$$g^{\bar u\bar u}=0, \quad (2.15d)$$

$$g^{\bar u\bar x}=0, \quad (2.15e)$$

we get a set of differential equations for ${}^0U$, ${}^0X$, and ${}^0R$. From Eq. (2.15a), (2.15b), and (2.15c), it follows that:

$${}^0U_{,u}+AG\,{}^0U_{,x}={}^0R^{-1}, \quad (2.16a)$$

$${}^0R_{,u}+AG\,{}^0R_{,x}=\tfrac{1}{2}AR\,{}^0G_{,x}, \quad (2.16b)$$

$${}^0X_{,u}+AG\,{}^0X_{,x}=0, \quad (2.16c)$$

where "," represents ordinary derivative.

The solution of (2.16) are

$${}^0X=-\tanh\chi, \quad (2.17a)$$

$${}^0R=G^{1/2}\cosh\chi, \quad (2.17b)$$

$$ {}^{0}U = \frac{1}{A\cosh\chi}\int_0^x G^{-3/2}dx + \alpha(\chi), \tag{2.17c} $$

where

$$ \chi = Au - \int_0^x G^{-1}dx \tag{2.18} $$

and $\alpha(\chi)$ is the solution of a homogeneous equation similar to Eq. (2.16c),

$$ \alpha_{,u} + AG\alpha_{,x} = 0, \tag{2.19} $$

so asymptotically we expect it to be a polynomial in $\bar{x}$, (cf. 2.14c). α is the supertranslation introduced by Bondi.

It also follows from Eqs. (2.15d) and (2.15e) that

$$ {}^{1}U = -\tfrac{1}{2}\,{}^{0}R\,({}^{0}U_{,x})^2 G, \tag{2.20a} $$

$$ {}^{1}X = -{}^{0}RG\,{}^{0}U_{,x}\,{}^{0}X_{,x}. \tag{2.20b} $$

The coordinate transformation given by Eqs. (2.14), (2.17), and (2.20) transforms the metric and its tetrads to an inertial coordinate system. In the coordinate system in which the line element (2.13) is written, one can construct a set of null tetrads given by

$$ l_\mu = (1,0,0,0), \quad l^\mu = (0,1,0,0), \tag{2.21a} $$

$$ n^\mu = (1, \tfrac{1}{2}A^2 r^2 G(x - A^{-1}r^{-1}), 0, 0), \tag{2.21b} $$

$$ n_\mu = (-\tfrac{1}{2}A^2 r^2 G(x - A^{-1}r^{-1}), 1, Ar^2, 0), $$

$$ m^\mu = \left(0, -\frac{1}{\sqrt{2}}AG^{1/2}r, \frac{1}{\sqrt{2}}\frac{G^{1/2}}{r}, \frac{i}{\sqrt{2}}\frac{1}{rG^{1/2}}\right). \tag{2.21c} $$

Applying the coordinate transformation given by Eqs. (2.14) and (2.17) on (2.21), we obtain

$$ l'_\mu = \frac{\partial x^\nu}{\partial \bar{x}^\mu} l_\nu = \left(\frac{\partial u}{\partial \bar{u}}, \frac{\partial u}{\partial \bar{r}}, \frac{\partial u}{\partial \bar{x}}, 0\right) \tag{2.22a} $$

$$ n'_\mu = \frac{\partial x^\nu}{\partial \bar{x}^\mu} n_\nu $$

$$ = \left(\frac{\partial u}{\partial \bar{u}} n_0 + \frac{\partial r}{\partial \bar{u}} n_1 + \frac{\partial x}{\partial \bar{u}} n_2, \frac{\partial u}{\partial \bar{r}} n_0 + \frac{\partial r}{\partial \bar{r}} n_1 + \frac{\partial x}{\partial \bar{r}} n_2, \frac{\partial u}{\partial \bar{x}} n_0 + \frac{\partial r}{\partial \bar{x}} n_1 + \frac{\partial x}{\partial x} n_2, 0\right), \tag{2.22b} $$

where

$$ \frac{\partial u}{\partial \bar{u}} = {}^{0}R + o(\bar{r}^{-1}) = \frac{G^{1/2}(x)}{\sin\bar{\theta}} + o(\bar{r}^{-1}), \tag{2.23a} $$

$$ \frac{\partial u}{\partial \bar{r}} = \frac{1}{2}\frac{G^{5/2}}{\sin^3\bar{\theta}}({}^{0}U_{,x})^2\bar{r}^{-2} + o(\bar{r}^{-3}), \tag{2.23b} $$

$$ \frac{\partial u}{\partial \bar{x}} = -\frac{G^{3/2}}{\sin^3\bar{\theta}}\,{}^{0}U_{,x} + o(\bar{r}^{-1}), \tag{2.23c} $$

$$ \frac{\partial r}{\partial \bar{u}} = -\tfrac{1}{2}AG_{,x}\bar{r} + o(1), \tag{2.23d} $$

$$ \frac{\partial r}{\partial \bar{r}} = G^{-1/2}\sin\bar{\theta} + o(\bar{r}^{-1}), \tag{2.23e} $$

$$ \frac{\partial r}{\partial \bar{x}} = -\frac{1}{\sin\bar{\theta}}\left[G^{-1/2}\cos\bar{\theta}\left(1 + \frac{A}{2}G_{,x}\frac{G^{1/2}}{\sin\bar{\theta}}(\bar{u} - \alpha)\right) + \frac{1}{2}\frac{G_{,x}}{\sin\bar{\theta}}\alpha_{,u}\right]\bar{r} + o(1), \tag{2.23f} $$

$$ \frac{\partial x}{\partial \bar{u}} = \frac{AG^{3/2}}{\sin\bar{\theta}} + o(\bar{r}^{-1}), \tag{2.23g} $$

$$ \frac{\partial x}{\partial \bar{r}} = -\frac{1}{2}\frac{G^{7/2}}{\sin^3\bar{\theta}}\left[A^{-1}G^{-3}\sin^3\bar{\theta}\left(1 - A^2\frac{\cos\bar{\theta}}{\sin^2\bar{\theta}}(\bar{u} - \alpha)^2\right) + 2AG^{-1}\cos\bar{\theta}\alpha_{,x}(\bar{u} - \alpha) - A\alpha_{,x}^2\right]\bar{r}^{-2} + o(\bar{r}^{-3}), \tag{2.23h} $$

$$ \frac{\partial x}{\partial \bar{x}} = \frac{G^{3/2}}{\sin^3\bar{\theta}}\,{}^{0}U_{,x}, \tag{2.23i} $$

where

$$ \cos\bar{\theta} \equiv \bar{x}. \tag{2.24} $$

Although this is a set of null tetrads in the inertial system, it is not the relevant one. In an asymptotically flat space–time one can construct a set of null tetrads given by

$$ \bar{l}_\mu = \bar{u}_{,\mu} = (1,0,0,0), \tag{2.25a} $$

$$ \bar{n}_\mu = (1, \tfrac{1}{2}, 0, 0) + o(\bar{r}^{-1}). \tag{2.25b} $$

Null tetrads, on the other hand, are defined through their orthonormality conditions which are preserved under Lorentz transformations and one can obtain Eq. (2.25) from Eq. (2.22) by applying a series of these transformations on (2.22). The required transformations are a null rotation about l'^μ followed by Lorentz transformation in the (l'^μ, n'^μ) plane, followed by a null rotation about n'^μ. Applying this on (2.22) we obtain the desired set of tetrads which satisfy Eq. (2.25). The two sets of tetrads are related by the following relations[5]:

$$ \bar{n}^\mu l'_\mu = A^{-1}, \tag{2.26a} $$

$$ \bar{n}^\mu n'_\mu = A^{-1}|\beta|^2, \tag{2.26b} $$

$$ \bar{l}^\mu l'_\mu = A^{-1}|\delta|^2. \tag{2.26c} $$

A corresponds to the Lorentz transformation, β to null rotation about l'^μ, and δ to the null rotation about n'^μ. These parameters must be determined in order to transform the Riemann and Maxwell scalars to the desirable form. They are obtained by substituting (2.22), (2.23), and (2.25) in (2.26). The result is

$$ A^{-1} = {}^{0}R + o(\bar{r}^{-1}), \tag{2.27a} $$

$$ \beta = -\frac{1}{\sqrt{2}}A\sin\bar{\theta}\bar{r} + o(1), \tag{2.27b} $$

$$ \delta = -\frac{1}{\sqrt{2}}A^{-1}G^{-1/2}(S - B - 1)\bar{r}^{-1} + o(\bar{r}^{-2}), \tag{2.27c} $$

where

$$S \equiv AG^{1/2}\cot\bar\theta\bar u, \tag{2.28a}$$

$$B \equiv \frac{AG^{1/2}}{\sin\bar\theta}(G\alpha_{,x} + \cos\bar\theta\alpha). \tag{2.28b}$$

The nonrotating results derived in this section can now be used in constructing and desired solution in the rotating case.

III. THE ROTATING CASE

In this section we will generalize the nonrotating results derived in the last section to the rotating case. We do this by multiplying the nonrotating results by some unknown functions and solving the asymptotic field equations for these unknown functions along with the news function which would not have been obtained easily if we had continued the methods of the last section for the rotating case.

Therefore, by analogy to (2.27) let

$$\delta = -\frac{1}{\sqrt{2}}A^{-1}G^{-1/2}V\bar r^{-1} + o(\bar r^{-2}), \tag{3.1}$$

$$\beta^* A^{-1} = -\frac{1}{\sqrt{2}}AG^{1/2}\omega\bar r + o(1), \tag{3.2}$$

V and ω being the unknown functions to be determined, and the asterisk representing the complex conjugate. Also in analogy to (2.14b) let

$$r = {}^0R^{-1}\Sigma^{-1}\bar r + o(1), \tag{3.3}$$

where Σ is to be determined.

Applying (3.3) on (2.7) and (2.6) and using (2.17a) and (2.17b) along with (2.24), we have

$$\phi'_1 = \frac{G}{\sqrt{2}\sin^2\bar\theta}\frac{\bar e\Sigma^2}{(1 + iaAx^2)^2}\bar r^{-2} + o(\bar r^{-3}), \tag{3.4}$$

$$\psi'_2 = -\frac{G^{3/2}}{\sin^3\bar\theta}\frac{M\Sigma^3}{(1 + iaAx^2)^3}\bar r^{-3} + o(\bar r^{-4}), \tag{3.5}$$

where M is defined by (2.8)

Now, the Riemann–Maxwell scalars as expressed by the barred tetrads are related to ϕ'_1 and ψ'_2 by[5]:

$$\bar\phi_0 = 2\delta(1 + \beta^* A^{-1}\delta)\phi'_1, \tag{3.6a}$$

$$\bar\phi_1 = (1 + 2\beta^* A^{-1}\delta)\phi'_1, \tag{3.6b}$$

$$\bar\phi_2 = 2\beta^* A^{-1}\phi'_1, \tag{3.6c}$$

$$\bar\psi_0 = 6\delta^2(1 + \beta^* A^{-1}\delta)^2\psi'_2, \tag{3.6d}$$

$$\bar\psi_1 = 3\delta[1 + \beta^* A^{-1}\delta(3 + 2\beta^* A^{-1}\delta)]\psi'_2, \tag{3.6e}$$

$$\bar\psi_2 = [1 + 6\beta^* A^{-1}\delta(1 + \beta^* A^{-1}\delta)]\psi'_2, \tag{3.6f}$$

$$\bar\psi_3 = 3\beta^* A^{-1}(1 + 2\beta^* A^{-1}\delta)\psi'_2, \tag{3.6g}$$

$$\bar\psi_4 = 6\beta^{*2}A^{-2}\psi'_2. \tag{3.6h}$$

These scalars have to satisfy the asymptotic field equations given by[6] (dropping the bars):

$$2\dot\psi^0_0 + \sqrt{2}\eth\psi^0_1 - 6\,{}^0\sigma\psi^0_2 - 6\phi^0_0\phi^{0*}_2 = 0, \tag{3.7a}$$

$$\sqrt{2}\dot\psi^1_0 - 8\eth^*({}^0\sigma\psi^0_1) + \sqrt{2}\eth^*\eth\psi^0_0 + 8\phi^{0*}_1\eth\phi^0_0 - 16\sqrt{2}\,{}^0\sigma\phi^0_1\phi^{*0}_1 - 8\sqrt{2}\phi^1_0\phi^{0*}_2 = 0, \tag{3.7b}$$

$$2\dot\psi^0_1 + \sqrt{2}\eth\psi^0_2 - 4\,{}^0\sigma\psi^0_3 - 4\phi^0_1\phi^{0*}_2 = 0, \tag{3.7c}$$

$$2\dot\psi^0_2 + \sqrt{2}\eth\psi^0_3 - 2\,{}^0\sigma\psi^0_4 - 2\phi^0_2\phi^{0*}_2 = 0, \tag{3.7d}$$

$$\sqrt{2}\dot\phi^0_0 + \eth\phi^0_1 - \sqrt{2}\,{}^0\sigma\phi^0_2 = 0, \tag{3.7e}$$

$$2\dot\psi^1_0 + \eth^*\eth\phi^0_0 - 2\sqrt{2}\eth^*({}^0\sigma\phi^0_1) = 0, \tag{3.7f}$$

$$\sqrt{2}\dot\phi^0_1 + \eth\phi^0_2 = 0, \tag{3.7g}$$

$$\psi^0_4 = -\ddot\sigma^{0*}, \tag{3.7h}$$

$$\psi^0_3 = \eth\dot\sigma^{0*}, \tag{3.7i}$$

where the dot represents the time derivative $\partial/\partial\bar u$ and $\eth$ is defined by

$$\eth\eta = -(\sin\bar\theta)^s\left\{\frac{\partial}{\partial\bar\theta} + \frac{i}{\sin\bar\theta}\frac{\partial}{\partial\bar\varphi}\right\}\{(\sin\bar\theta)^{-s}\eta\}, \tag{3.8}$$

s is the spin weight. We have used the following notation:

$$\phi_0 = \phi^0_0\bar r^{-3} + \phi^1_0\bar r^{-4} + \cdots, \tag{3.9a}$$

$$\phi_1 = \phi^0_1\bar r^{-2} + \phi^1_1\bar r^{-3} + \cdots, \tag{3.9b}$$

$$\phi_2 = \phi^0_2\bar r^{-1} + \phi^1_2\bar r^{-2} + \cdots, \tag{3.9c}$$

$$\psi_0 = \psi^0_0\bar r^{-5} + \psi^1_0\bar r^{-6} + \cdots, \tag{3.9d}$$

$$\psi_1 = \psi^0_1\bar r^{-4} + \psi^1_1\bar r^{-5} + \cdots, \tag{3.9e}$$

$$\psi_2 = \psi^0_2\bar r^{-3} + \psi^1_2\bar r^{-4} + \cdots, \tag{3.9f}$$

$$\psi_3 = \psi^0_3\bar r^{-2} + \psi^1_3\bar r^{-3} + \cdots, \tag{3.9g}$$

$$\psi_4 = \psi^0_4\bar r^{-1} + \psi^1_4\bar r^{-2} + \cdots. \tag{3.9h}$$

Again by analogy to (2.23g) and (2.23i) let

$$\frac{\partial x}{\partial\bar u} = \frac{AG^{3/2}}{\sin\bar\theta}\mathscr{E}_1, \tag{3.10a}$$

$$\frac{\partial x}{\partial\bar\theta} = -\frac{G(S - B)}{\sin\bar\theta}\mathscr{E}_2. \tag{3.10b}$$

Substituting (3.1), (3.2), (3.4), and (3.5) in (3.6), and using these results along with (3.8) and (3.10), one can show that (3.7) is satisfied, if

$$\mathscr{E}_1 = \mathscr{E}_2 = W_0^{-1} \equiv (1 + a^2A^2x^4)^{-1}, \tag{3.11}$$

$$\omega = \frac{1}{1 + iaAx^2}, \tag{3.12}$$

$$V = (S - B) - (1 + iaAx^2), \tag{3.13}$$

$$\Sigma = 1, \tag{3.14}$$

$$\frac{\partial}{\partial\bar u}(G\alpha_{,x} + \cos\bar\theta\alpha) = 0, \tag{3.15}$$

and

$${}^0\sigma = -\frac{A^{-1}G^{-1/2}}{4\sin\bar\theta}\left[\frac{1 + iaAx^2}{1 - iaAx^2}G_{,x} + 2\frac{S - B}{\cos\bar\theta} + 2AG^{1/2} \times\left(\frac{\partial}{\partial\bar\theta} + \tan\bar\theta\right)(G\alpha_{,x} + \cos\bar\theta\alpha) - \frac{4iaAxG}{1 - iaAx^2}\right). \tag{3.16}$$

S and B are defined by (2.28), and ${}^0\sigma$ is related to the news function by $N \equiv d\,{}^0\sigma/d\bar{u}$, and the news functionis then given by

$$N = -\frac{1}{2\sin^2\bar{\theta}} - \frac{1}{4\sin^2\bar{\theta}}\frac{1}{(1-iaAx^2)^3}$$

$$\times\Big\{(1-iaAx^2)\Big[GG_{,xx} - \frac{1}{2}(G_{,}x)^2\Big]$$

$$+ 2iaAG(xG_{,x} - 2G)\Big\}. \qquad (3.17)$$

Finally, substituting (3.11)–(3.14) together with (3.1)–(3.5) in (3.6), we get:

$$\phi_0^0 = -\frac{1}{2}\frac{A^{-1}G^{1/2}}{\sin^2\bar{\theta}}\bar{e}\frac{(S-B)^2-(1+iaAx^2)^2}{(1+iaAx^2)^3}, \qquad (3.18a)$$

$$\phi_1^0 = \frac{1}{\sqrt{2}}\frac{G}{\sin^2\bar{\theta}}\bar{e}\frac{(S-B)}{(1+iaAx^2)^3}, \qquad (3.18b)$$

$$\phi_2^0 = -\frac{AG^{3/2}}{\sin^2\bar{\theta}}\bar{e}\frac{1}{(1+iaAx^2)^3}, \qquad (3.18c)$$

$$\psi_0^0 = -\frac{3}{4}\frac{A^{-2}G^{1/2}}{\sin^3\bar{\theta}}M\frac{[(S-B)^2-(1+iaAx^2)^2]^2}{(1+iaAx^2)^5}, \qquad (3.18d)$$

$$\psi_1^0 = \frac{3}{2\sqrt{2}}\frac{A^{-1}}{\sin^3\bar{\theta}}M\frac{(S-B)[(S-B)^2-(1+iaAx^2)^2]}{(1+iaAx^2)^5}, \qquad (3.18e)$$

$$\psi_2^0 = -\frac{1}{2}\frac{G^{3/2}}{\sin^3\bar{\theta}}M\frac{[3(S-B)^2-(1+iaAx^2)^2]}{(1+iaAx^2)^5}, \qquad (3.18f)$$

$$\psi_3^0 = \frac{3}{\sqrt{2}}\frac{AG^2}{\sin^3\bar{\theta}}M\frac{(S-B)}{(1+iaAx^2)^5}, \qquad (3.18g)$$

$$\psi_4^0 = -3\frac{A^2G^{5/2}}{\sin^3\bar{\theta}}M\frac{1}{(1+iaAx^2)^5}. \qquad (3.18h)$$

We have now accomplished the transformation to the inertial coordinate system, and have obtained the Riemann–Maxwell scalars and the news function in this system. One can now discuss the physical properties of this metric, which will be done in the next section.

IV. DISCUSSION AND INTERPRETATION

Equation (3.18) gives the behavior of the field produced by the uniformly accelerated particle, as seen from an inertial coordinate system. In contrast with the accelerated coordinate system, fields are no longer time independent and none of the Riemann–Maxwell scalars are zero. The nonvanishing components, ϕ_2^0 and ψ_4^0, describe the behavior of the electromagnetic and gravitational radiation. The function G is similar to a Lorentz contracting factor whose zeros are correspondent to a horizon. This horizon describes a surface of infinite red shift beyond which the accelerated particle has no casual effect. Rotation, magnetic monopole, and the NUT parameter of the particle make the Riemann–Maxwell scalars complex.

Let us now consider the various physical properties of the source which can be defined in terms of these Riemann–Maxwell scalars. The quadrupole moment is defined in terms of ψ_0^0 by[7]

$$Q(\bar{u}) = -\tfrac{1}{2}\int_0^\pi \psi_0^0 P_2^2(\cos\bar{\theta})\sin\bar{\theta}d\bar{\theta}, \qquad (4.1)$$

where $P_m^n(\cos\bar{\theta})$ are the associated Legendre polynomials. The (complex) dipole moment is defined to be[7]

$$D(\bar{u}) + iL(\bar{u}) = -\tfrac{1}{2}\int_0^\pi \psi_1^0 P_1^1(\cos\bar{\theta})\sin\bar{\theta}d\bar{\theta}, \qquad (4.2)$$

where $D(\bar{u})$ is the mass dipole moment and $L(\bar{u})$ is the angular momentum.

The mass loss as defined by Newman and Unti[7] is

$$\frac{\partial M_{NU}}{\partial\bar{u}} = -\tfrac{1}{2}\int_0^\pi |N|^2\sin\bar{\theta}d\bar{\theta} + \tfrac{1}{4}\int_0^\pi \frac{\partial^2}{\partial\bar{u}^2}|{}^0\sigma|^2\sin\bar{\theta}d\bar{\theta}$$

$$-\tfrac{1}{2}\int_0^\pi |\phi_2^0|^2\sin\bar{\theta}d\bar{\theta}. \qquad (4.3a)$$

The last term was added in order to take into account the electromagnetic radiation.[5] Bondi's definition of the mass loss is

$$\frac{\partial M_B}{\partial\bar{u}} = -\tfrac{1}{2}\int_0^\pi |N|^2\sin\bar{\theta}d\bar{\theta} - \tfrac{1}{2}\int_0^\pi |\phi_2^0|^2\sin\bar{\theta}d\bar{\theta}. \qquad (4.3b)$$

The linear momentum along the axis of symmetry is given by[7]

$$P(\bar{u}) = -\tfrac{1}{4}\int_0^\pi (\psi_2^0 + \psi_2^{0*})P_1(\cos\bar{\theta})\sin\bar{\theta}d\bar{\theta}, \qquad (4.4)$$

and finally the (complex) charge is related to ϕ_1^0 by[6]

$$E = \frac{1}{\sqrt{4\pi}}\int_0^{2\pi}\int_0^\pi \phi_1^0\sin\bar{\theta}d\bar{\theta}d\bar{\varphi}. \qquad (4.5)$$

Using the results of the last section one can calculate these quantities for arbitrary uniform acceleration. However, this is not necessary nor even advisable, because of the mathematical complexity. The relevant behavior can be ascertained from the limit of small accelerations.

Keeping only the lowest order terms in A, Eq. (2.17c) gives

$$\bar{u} = u + \frac{1}{A}\{\bar{x} + A\,[n(1-(1-\bar{x}^2)^{1/2}) - m(\ln(1-\bar{x}^2)$$

$$+ 2 - 2(1-\bar{x}^2)^{1/2})]\} + \alpha + o(A). \qquad (4.6)$$

Although Eq. (2.17) is written for the nonrotating case, it can be shown that corrections due to the rotation involve terms of the second order in A. Therefore, up to the first order in A, (2.17) is valid for the rotating case as well.

In Eq. (4.6) we choose α so that

$$\bar{u} = u + o(A). \tag{4.7}$$

Comparing (4.7) and (4.6) gives

$$\alpha = -\frac{1}{A}\{\bar{x} + A[n(1-(1-\bar{x}^2)^{1/2}) - m(\ln(1-\bar{x}^2) + 2 - 2(1-\bar{x}^2)^{1/2})]\} \tag{4.8}$$

This choice of supertranslation α does indeed satisfy (2.19) and (3.15). This choice of α is also the one that makes the results reduce to the familiar form of the Schwartzschild case in the limit when $A \to 0$ and $a \to 0$.

Using (2.17a) it can be shown that

$$x = \bar{x} + A\left[\bar{u} - \frac{n\bar{x}^2}{1-\bar{x}^2} - m\left(\frac{\bar{x}^2}{1-\bar{x}^2} + \ln(1-\bar{x}^2)\right)\right] \times (1-\bar{x}^2) + o(A^2). \tag{4.9}$$

Substituting this result in (2.28) we have

$$(S - B) = 1 + o(A^2). \tag{4.10}$$

Using (4.9) and (4.10) in Eq. (3.18) we have

$$\phi_0^0 = ia\frac{\cos^2\bar{\theta}}{\sin\bar{\theta}}\bar{e} + o(A), \tag{4.11a}$$

$$\phi_1^0 = \frac{\bar{e}}{\sqrt{2}}(1 - 2A\cos\bar{\theta}\bar{u}) - \sqrt{2}\bar{e}An\cos\bar{\theta} - \frac{3}{\sqrt{2}}iaA\bar{e}\cos^2\bar{\theta} + 2\sqrt{2}\bar{e}Am\cos\bar{\theta}\ln(\sin\bar{\theta}) + o(A^2), \tag{4.11b}$$

$$\phi_2^0 = -A\bar{e}\sin\bar{\theta}(1 - 3A\cos\bar{\theta}\bar{u}) + 3A^2n\bar{e}\sin\bar{\theta}\cos\bar{\theta} + 3ia \times A^2\bar{e}\sin\bar{\theta}\cos^2\bar{\theta} - 6\bar{e}A^2m\sin\bar{\theta}\cos\bar{\theta}\ln(\sin\bar{\theta}) + o(A^3), \tag{4.11c}$$

$$\psi_0^0 = 3a^2\frac{\cos^4\bar{\theta}}{\sin^2\bar{\theta}}m + O(A^2), \tag{4.11d}$$

$$\psi_1^0 = -\frac{3}{\sqrt{2}}ia\frac{\cos^2\bar{\theta}}{\sin\bar{\theta}}m + o(A), \tag{4.11e}$$

$$\psi_2^0 = -m(1 - 3A\cos\bar{\theta}\bar{u}) - 2A|\bar{e}|^2\cos\bar{\theta} - (iaAn - 3Amn\cos\bar{\theta}) - 6iaAm\cos^2\bar{\theta} - 6Am^2\cos\bar{\theta}\ln(\sin\bar{\theta}) + o(A^2), \tag{4.11f}$$

$$\psi_3^0 = \frac{3}{2}Am\sin\bar{\theta}(1 - 4A\cos\bar{\theta}\bar{u}) + 3\sqrt{2}A^2|\bar{e}|^2\sin\bar{\theta}\cos\bar{\theta} + \frac{3}{\sqrt{2}}An\sin\bar{\theta}(iaA - 4Am\cos\bar{\theta}) - \frac{15}{\sqrt{2}}iaA^2m \times \sin\bar{\theta}\cos^2\bar{\theta} + 12\sqrt{2}A^2m^2\sin\bar{\theta}\cos\bar{\theta}\ln(\sin\bar{\theta}) + o(A^3), \tag{4.11g}$$

$$\psi_4^0 = -3A^2m\sin^2\bar{\theta}(1 - 5A\cos\bar{\theta}\bar{u}) - 6A^3|\bar{e}|^2\sin^2\bar{\theta}\cos\bar{\theta} - 3A^2n\sin^2\bar{\theta}(iaA - 5Am\cos\bar{\theta}) + 15iaA^3m\sin^2\bar{\theta} \times \cos^2\bar{\theta} - 30A^3m^2\sin^2\bar{\theta}\cos\bar{\theta}\ln(\sin\bar{\theta}) + o(A^4), \tag{4.11h}$$

$$^0\sigma = ia\frac{\cos\bar{\theta}}{\sin^2\bar{\theta}} + o(A), \tag{4.11i}$$

$$N = Am\cos\bar{\theta}\left(1 + \frac{2}{\sin^2\bar{\theta}}\right) + \frac{iaA}{\sin^2\bar{\theta}} + o(A^2). \tag{4.11j}$$

Equation (4.11) has basically three different types of terms as denoted by the parentheses. The first term in the brackets on the right-hand side represents the effect of a Lorentz contraction term due to the instantaneous velocity $A\bar{u}\cos\bar{\theta}$. The second kind of term represents the effect of the charge, rotation, and NUT parameter. The third kind of term contains a line singularity along the axis of symmetry. This singularity is a result of the force that produces the uniform acceleration and feeds energy into the particle. These singular terms lead to divergent integrals and consequently make an infinite contribution to the mass loss. In the complete solution these different terms can't be distinguished and are contained in the function G.

In order to get a better feeling for the interpretation of these scalars let us compare ψ_2 in Eq. (4.11f) to that for a Schwarzschild particle that has been Lorentz transformed to a moving frame. Bondi[3] has applied a K transformation to the Schwarzschild solution and showed that the mass aspects in the two coordinate systems are related by

$$(\psi_2^0)_r = \frac{(\psi_2^0)_M}{(\cosh\nu + \cos\theta\sinh\nu)^3}. \tag{4.12}$$

$(\psi_2^0)_r$ is the mass aspect in a frame which is at rest, and $(\psi_2^0)_M$ is that in the frame moving with the particle along the axis of symmetry with velocity $\tanh\nu$.

If we let $\tanh\nu = V = A\bar{u}$ and expand Eq. (4.12) for small A we have

$$(\psi_2^0)_r = -m(1 - 3A\bar{u}\cos\bar{\theta}). \tag{4.13}$$

This is identical to Eq. (4.11f) if we let $n = g = e = a = 0$. Comparison with (3.18f) shows that G is similar to the Lorentz contracting factor. This relation between the mass aspects is called the Doppler shift in the mass aspect. Also notice that in the limit $A \to 0$ and $a \to 0$, Eq. (4.11) reduces to the familiar Schwarzschild case.

Substituting (4.11) in (4.1)–(4.5), we obtain the following results. The quadrupole moment, mass dipole moment, and angular momentum become

$$Q(\bar{u}) = \tfrac{9}{5}a^2m + o(A^2), \tag{4.14}$$

$$D(\bar{u}) = -\tfrac{1}{2}\int_0^\pi \mathrm{Re}\psi_1^0 P_1^1(\cos\bar{\theta})\sin\bar{\theta}d\bar{\theta} = o(A), \tag{4.15}$$

$$L(\bar{u}) = -\tfrac{1}{2}\int_0^\pi \mathrm{Im}\psi_1^0 P_1^1(\cos\bar{\theta})\sin\bar{\theta}d\bar{\theta} = -\frac{1}{\sqrt{2}}am, \tag{4.16}$$

respectively. Equation (4.16) shows that the particle is rotating and a is the angular momentum per unit mass.

We will calculate the mass loss using both definitions in Eq. (4.3a) and (4.3b). We will then compare their results. From (4.3b) we have

$$\frac{\partial \dot{M}_B}{\partial \bar{u}} = -\tfrac{1}{3}A^2m^2 - \tfrac{2}{3}A^2|\bar{e}|^2 - 2A^2m^2\int_0^\pi \frac{\cos^2\bar{\theta}}{\sin\bar{\theta}} \times\left(1+\frac{1}{\sin^2\bar{\theta}}\right)d\bar{\theta} - \frac{a^2A^2}{2}\int_0^\pi \frac{d\bar{\theta}}{\sin^3\bar{\theta}} + o(A^4). \tag{4.17a}$$

The first term represents the mass loss through gravitational radiation, while the second term is the familiar electromagnetic radiated power. The last two divergent integrals are due to the line singularity discussed before. The rotation and the NUT parameter have no effect up to this order. They contribute to the terms of the order of A^4 and higher.

On the other hand, using the fact that $N \equiv \partial\, {}^0\sigma/\partial\bar{u}$ and $\psi_4^0 = -\partial N/\partial\bar{u}$ the second term in Eq. (4.3a) can be written as

$$\begin{aligned}\tfrac{1}{4}\int_0^\pi \frac{\partial^2}{\partial\bar{u}^2}|{}^0\sigma|^2\sin\bar{\theta}d\bar{\theta} &= \tfrac{1}{4}\int_0^\pi \frac{\partial^2}{\partial\bar{u}^2}({}^0\sigma{}^0\sigma^*)\sin\bar{\theta}d\bar{\theta}\\ &= \tfrac{1}{4}\int_0^\pi \frac{\partial}{\partial\bar{u}}(N\,{}^0\sigma^* + {}^0\sigma N^*)\sin\bar{\theta}d\bar{\theta}\\ &= \tfrac{1}{2}\int_0^\pi |N|^2\sin\bar{\theta}d\bar{\theta} - \tfrac{1}{4}\int_0^\pi(\psi_4^0\,{}^0\sigma^*\\ &\quad + {}^0\sigma\psi_4^{0*})\sin\bar{\theta}d\bar{\theta}.\end{aligned} \tag{4.17a'}$$

Therefore, Newman–Unti's mass loss is given by

$$\frac{\partial M_{NU}}{\partial\bar{u}} = -\tfrac{1}{4}\int_0^\pi(\psi_4^0\,{}^0\sigma^* + {}^0\sigma\psi_4^{0*})\sin\bar{\theta}d\bar{\theta} - \tfrac{1}{2}\int_0^\pi|\phi_2^0|^2\sin\bar{\theta}d\bar{\theta}. \tag{4.17b}$$

The lowest order term in ψ_4^0 is of the order of A^2 and is real, while ${}^0\sigma$ is of the order of 1 and is pure imaginary; therefore aside from the electromagnetic part, $\partial M_{Nu}/\partial\bar{u}$, unlike the Bondi's mass loss, has no $o(A^2)$ term, i.e.,

$$\frac{\partial M_{NU}}{\partial\bar{u}} = -\tfrac{2}{3}A^2|\bar{e}|^2 + o(A^3). \tag{4.17b'}$$

Therefore, one can conclude that the definition given by Newman and Unti is not consistent with the result expected from perturbation methods. The Newman-Unti's definition predicts no mass loss for a uniformly accelerated particle due to its emission of gravitational radiation up to the order of A^2.

The linear momentum in Eq. (4.4) becomes

$$P(\bar{u}) = Am(\bar{u}+n) - \tfrac{2}{3}A(\bar{e})^2 + 3m^2\int_0^\pi \sin\bar{\theta}\cos^2\bar{\theta} \times \ln(\sin\bar{\theta})d\bar{\theta}. \tag{4.18}$$

The first term represents the instantaneous momentum of a particle that is undergoing a constant acceleration. The last term represents the contribution of the line source. The NUT parameter contributes to the retarded time but up to this order it does not affect the mass.

Finally the complex charge is

$$E = \sqrt{2\pi}[(e + aAg) + i(g - aAe)]. \tag{4.19}$$

Equation (4.19) suggests that in the presence of rotation the magnetic monopole contributes to the electric charge through a term proportional to aA. Likewise, the magnetic charge is influenced by the electric monopole. This is a result expected from the electromagnetic theory because a moving or rotating electric charge produces a magnetic field and therefore it can serve as a magnetic charge. This may suggest an experimental way to see if there exists any magnetic monopole or not. Basically what one can do is measure the charge before and after the spinning particle is accelerated. If there is any difference in these measurements, one can conclude the existence of magnetic monopoles.

APPENDIX

In Eqs. (2.1) and (2.2) make the following change of scale:

$$\begin{aligned}&(p,q,\sigma,a)\to l^{-1}(q',p',\sigma',a'),\\ &n_0\to ln',\quad \epsilon_0\to l^2\epsilon',\quad m_0\to l^3m',\\ &e_0+ig_0\to l^2(e'+ig'),\quad \gamma_0\to\gamma'+l^4g'^2,\end{aligned} \tag{A1}$$

where

$$\gamma' \equiv \frac{1}{a^2+b^2},\quad \epsilon' \equiv -\frac{1}{ab}\frac{a^2-b^2}{a^2+b^2},\quad l^{-2} = ab, \tag{A2}$$

a being the rotation parameter and b is the inverse of the acceleration. Performing another change of scale transformation

$$\begin{aligned}&b\to\frac{1}{A},\quad \tau'\to\frac{t}{A},\quad \sigma'\to\frac{z}{A},\quad p'\to -Ax,\quad q'\to -Ay,\\ &P'\to A^2G,\quad Q'\to A^2F,\quad \gamma'\to A^2\gamma,\quad n'\to A^2n,\end{aligned} \tag{A3}$$

followed by the coordinate transformation

$$\begin{aligned}&Au = t + \int_y F^{-1}(y)dy,\quad A(x+y) = r^{-1},\\ &\varphi = z - aA\int_y y^2F^{-1}(y)dy,\end{aligned} \tag{A4}$$

we obtain (2.5).

[1]W. Kinnersley and M. Walker, Phys. Rev. D **2**, 1359 (1970).
[2]J.F. Plebanski and M. Demianski, Ann. Phys. **98**, 98 (1976).
[3]H. Bondi, M.G.J. van der Burg, and A.W.K. Metzner, Proc. Soc. London, Ser.A **269**, 21 (1963).
[4]E.T. Newman and T.W.J. Unti, J. Math. Phys. **3**, 891 (1962).
[5]A.I. Janis and E.T. Newman, J. Math. Phys. **6**, 902 (1965).
[6]A.R. Exton, E.T. Newman, and R. Penrose, J. Math. Phys. **10**, 1566 (1969).
[7]E.T. Newman and T.W.J. Unti, J. Math. Phys. **6**, 1806 (1965).

Generalized static electromagnetic fields in Brans–Dicke theory

T. Singh and L. N. Rai

Applied Mathematics Section, Institute of Technology, B. H. U. Varanasi-221005, India
(Received 29 September 1978)

A general class of static, axially symmetric solutions of the Brans–Dicke–Maxwell equations is obtained under the assumption $-r^2 \cdot g_{rr} = g_{tt} \cdot g_{\phi\phi} \cdot g_{zz}$. These solutions have been subjected to conformal transformation and they are found to be new solutions of the static, axially symmetric Einstein–Maxwell scalar fields. We have also developed a more general set of solutions under the unit transformations given by Morganstern.

1. INTRODUCTION

Brans and Dicke have proposed a new theory of gravitation,[1-3] which differs from Einstein's theory by the presence of a scalar field Φ in the field equations. One of the important consequences of this new theory is that the gravitational "constant" G is inversely proportional to the strength of the scalar field Φ. This new theory incorporates Mach's principle in contrast to Einstein's theory.[1,4] Einstein's theory is a special case of the Brans-Dicke theory, if we let $\Phi =$ const and let the coupling constant $\omega \rightarrow \infty$.

According to Dicke,[3] by a proper choice of the conformal factor and scaling the metric as well as field quantities suitably, the Brans-Dicke coupled scalar electromagnetic fields reduce to the coupled zero-mass scalar and electromagnetic fields of Einstein's gravitation theory. The corresponding result for vacuum Brans-Dicke fields have been established by Peters[5] and Tabensky and Taub.[6] Many authors have obtained solutions of the Brans-Dicke fields coupled to electromagnetic fields.[7-13] Some cylindrically symmetric Einstein–Maxwell scalar fields have been discussed by Rao *et al.*,[14,15] and Teixeira *et al.*[16] In this paper we have considered the Brans–Dicke fields coupled to source-free electromagnetic fields in a static, axially symmetric space–time. In Sec. 2 we have written Brans–Dicke–Maxwell field equations for a static axially symmetric line element. In Sec. 3 we have obtained the solutions of Brans–Dicke field equations for three types of electromagnetic field (i.e., azimuthal, radial, and longitudinal fields). In Sec. 4 we have developed a more general set of solutions for the said field equations under the unit transformation given by Morganstern.[17] Since the Brans–Dicke scalar electromagnetic fields are conformal to the coupled zero-mass and source-free electromagnetic fields of Einstein's gravitational theory, an application of this method is made in Sec. 5.

2. FIELD EQUATIONS

The field equations of the Brans–Dicke theory coupled to electromagnetic fields are

$$R_{ij} = -\frac{8\pi}{\Phi}E_{ij} - \frac{\omega}{\Phi^2}\Phi_{,i}\Phi_{,j} - \frac{\Phi_{;ij}}{\Phi} \tag{2.1}$$

and

$$(3+2\omega)\Phi^{\ k}_{;k} = 0 \tag{2.2}$$

with

$$E_{ij} = (F_i^{\ \alpha}F_{j\alpha} - \tfrac{1}{4}g_{ij}\, F_\alpha^{\ \beta}F^\alpha_{\ \beta}), \tag{2.3}$$

where E_{ij} is the electromagnetic energy–momentum tensor, $\Phi_{,i} = (\partial\Phi/\partial x^i)$ and ω is the coupling constant. F_{ij} is the skew-symmetric electromagnetic field tensor which satisfies Maxwell's equations for empty space,

$$F_{[ij;k]} = 0 \tag{2.4}$$

and

$$F_{;j}^{\ ij} = 0, \tag{2.5}$$

the semicolon denoting covariant differentiation. In (2.2) we take $\omega \neq -\frac{3}{2}$ and therefore

$$\Phi^{\ k}_{;k} = 0. \tag{2.6}$$

We take the line element for a static system with axial symmetry as

$$ds^2 = e^{2\eta}dt^2 - e^{2\lambda}dr^2 - r^2e^{2\beta}d\phi^2 - e^{2\gamma}dz^2, \tag{2.7}$$

where η, λ, β, and γ are functions of r only. It is known[18] that, for any static, axially symmetric system, the metric components can be described by only three functions of radial coordinates. So we choose a relation

$$\lambda = \eta + \beta + \gamma. \tag{2.8}$$

The nonvanishing components of the Ricci tensor for the metric (2.7) are

$$R^0_0 = -e^{-2(\eta+\beta+\gamma)}\cdot(\eta_{11} + \eta_1/r), \tag{2.9}$$

$$R^1_1 = -e^{-2(\eta+\beta+\gamma)}\cdot[\eta_{11} + \beta_{11} + \gamma_{11} + (\beta_1 - \eta_1 - \gamma_1)/r - 2(\eta_1\beta_1 + \eta_1\gamma_1 + \beta_1\gamma_1)], \tag{2.10}$$

$$R^2_2 = -e^{-2(\eta+\beta+\gamma)}\cdot(\beta_{11} + \beta_1/r), \tag{2.11}$$

and

$$R^3_3 = -e^{-2(\eta+\beta+\gamma)}\cdot(\gamma_{11} + \gamma_1/r). \tag{2.12}$$

Since the electromagnetic field $F_{\mu\nu}$ depends only on r, from

(2.4) we have

$$F_{02} = \mathscr{E}_\phi, \quad (2.13)$$

$$F_{03} = \mathscr{E}_z, \quad (2.14)$$

and

$$F_{23} = k\mathscr{B}_r, \quad (2.15)$$

where $\mathscr{E}_\phi, \mathscr{E}_z$, and $\mathscr{B}_r$ are constants, and from (2.5) we get

$$F_{01} = \mathscr{E}_r \cdot r^{-1} \cdot e^{2\eta}, \quad (2.16)$$

$$F_{12} = k\mathscr{B}_z \cdot re^{2\beta}, \quad (2.17)$$

and

$$F_{13} = k\mathscr{B}_\phi \cdot r^{-1} e^{2\gamma}, \quad (2.18)$$

where $\mathscr{E}_r, \mathscr{B}_z$, and $\mathscr{B}_\phi$ are constants. Also the scalar field depends on r only so from (2.6)

$$\Phi = C \log r, \quad (2.19)$$

where C is a constant. If we define the constants

$$b_r = 8\pi(\mathscr{E}_r^2 + k^2\mathscr{B}_r^2)/2, \quad (2.20)$$

$$b_\phi = 8\pi(\mathscr{E}_\phi^2 + k^2\mathscr{B}_\phi^2)/2 \quad (2.21)$$

and

$$b_z = 8\pi(\mathscr{E}_z^2 + k^2\mathscr{B}_z^2)/2, \quad (2.22)$$

then the Brans–Dicke equations can be written as

$$r^2\eta_{11} + r\eta_1 = \frac{1}{C\log r}[b_r e^{2\eta} + b_\phi e^{2\gamma} + b_z \cdot r^2 e^{2\beta}] - \frac{r\eta_1}{\log r}, \quad (2.23)$$

$$r^2\beta_{11} + r\beta_1 = \frac{1}{C\log r}\cdot[-b_r e^{2\eta} + b_\phi e^{2\gamma} - b_z \cdot r^2 e^{2\beta}] - \frac{(1 + r\beta_1)}{\log r}, \quad (2.24)$$

$$r^2\gamma_{11} + r\gamma_1 = \frac{1}{C\log r}\cdot[-b_r e^{2\eta} - b_\phi e^{2\gamma} + b_z \cdot r^2 e^{2\beta}] - \frac{r\gamma_1}{\log r}, \quad (2.25)$$

and

$$r^2\left[\eta_1\gamma_1 + (\eta_1 + \gamma_1)\left(\beta_1 + \frac{1}{r}\right)\right] = -\frac{1}{\log r}\left(1 + \frac{\omega}{2\log r}\right) - \frac{r\lambda_1}{\log r} + \frac{1}{C\log r} \times [-b_r e^{2\eta} + b_\phi e^{2\gamma} + b_z \cdot r^2 e^{2\beta}] \quad (2.26)$$

3. SOLUTIONS OF BRANS–DICKE FIELD EQUATIONS

The method for obtaining the solution in all the three cases (radial, aximuthal, and longitudinal) is identical. First we solve the first three equations and get the values of η, β, and γ [λ is obtained from (2.8)] with a total of six constants of integration; then Eq. (2.26) gives a relation which reduces six constants to five.

A. Azimuthal fields

In this case ($b_r = b_z = 0$) the field equations are

$$r^2\eta_{11} + r\eta_1 = \frac{1}{C\log r}\cdot(b_\phi e^{2\gamma}) - \frac{r\eta_1}{\log r}, \quad (3.1)$$

$$r^2\beta_{11} + r\beta_1 = \frac{1}{C\log r}\cdot(b_\phi e^{2\gamma}) - \frac{(1 + r\beta_1)}{\log r}, \quad (3.2)$$

$$r^2\gamma_{11} + r\gamma_1 = \frac{1}{C\log r}\cdot(-b_\phi e^{2\gamma}) - \frac{r\gamma_1}{\log r}, \quad (3.3)$$

and

$$r^2\left[\eta_1\gamma_1 + (\eta_1 + \gamma_1)\cdot\left(\beta_1 + \frac{1}{r}\right)\right] = -\frac{1}{\log r}\left(1 + \frac{\omega}{2\log r}\right) + \frac{1}{C\log r}(b_\phi e^{2\gamma}) - \frac{r\lambda_1}{\log r} \quad (3.4)$$

along with (2.19),

Equation (3.3) gives

$$\gamma = -\log\left[2\left(\frac{b_\phi}{B'C}\right)^{1/2}\right]\cdot(\log r)^{1/2}\cosh\left(\frac{\sqrt{B'}}{2}\cdot\log(\log r^D)\right) = -\log X, \quad (3.5)$$

B' and D being constants of integration.

The sum of (3.1) and (3.3) gives

$$\eta = -\gamma + h\log\left(\frac{\log r}{d}\right), \quad (3.6)$$

where h and d are constants of integration.

The sum of (3.2) and (3.3) gives

$$\beta = -\gamma + p\log\left\{\frac{r^{-1/p}\cdot\log r}{f}\right\} \quad (3.7)$$

with p and f as constants of integration.

Finally substitution of (3.5)–(3.7) into (3.4) gives

$$ph + (p + h) + \omega/2 = (a^2 - 1)/4 \quad (3.8)$$

with a as a constant ($\sqrt{B'} = a$).

We thus obtain

$$g_{00} = \left(\frac{\log r}{d}\right)^{2h}\cdot X^2, \quad (3.9)$$

$$g_{11} = -r^{-2}\cdot\left(\frac{\log r}{d}\right)^{2h}\cdot\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)}\cdot X^2, \quad (3.10)$$

$$g_{22} = -\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)}\cdot X^2, \quad (3.11)$$

and

$$g_{33} = -X^{-2} \quad (3.12)$$

where

$$X = \left[2\cdot\left(\frac{b_\phi}{a^2c}\right)^{1/2}\cdot(\log r)^{1/2}\cdot\cosh\left(\frac{a}{2}\cdot\log(\log r^D)\right)\right].$$

B. Radial fields

In this case ($b_\phi = b_z = 0$) the field equations are

$$r^2\eta_{11} + r\eta_1 = \frac{1}{C\log r}\cdot(b_r e^{2\eta}) - \frac{r\eta_1}{\log r}, \quad (3.13)$$

$$r^2\beta_{11} + r\beta_1 = \frac{1}{C\log r}\cdot(-b_r e^{2\eta}) - \frac{(1 + r\beta_1)}{\log r}, \quad (3.14)$$

$$r^2\gamma_{11} + r\gamma_1 = \frac{1}{C\log r}\cdot(-b_r e^{2\eta}) - \frac{r\gamma_1}{\log r}, \quad (3.15)$$

and

$$r^2\left[\eta_1\gamma_1 + (\eta_1 + \gamma_1)\left(\beta_1 + \frac{1}{r}\right)\right]$$
$$= -\frac{1}{\log r}\left(1 + \frac{\omega}{2\log r}\right) + \frac{1}{C\log r}(-b_r e^{2\eta}) - \frac{r\lambda_1}{\log r} \tag{3.16}$$

along with (2.19).

Equation (3.13) gives

$$\eta = -\log\left[2\left(\frac{b_r}{B'C}\right)^{1/2}\cdot(\log r)^{1/2}\cdot\sinh\left(\frac{\sqrt{B'}}{2}\cdot\log(\log r^D)\right)\right]$$
$$= -\log X', \tag{3.17}$$

where B' and D are constants of integration.

The sum of (3.13) and (3.15) gives

$$\gamma = -\eta + h\log\left(\frac{\log r}{d}\right), \tag{3.18}$$

where h and d are constants of integration.

The sum of (3.13) and (3.14) gives

$$\beta = -\eta + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right), \tag{3.19}$$

p and f being constants of integration.

Finally substituting (3.17)–(3.19) into (3.16) gives

$$ph + (p+h) + \frac{\omega}{2} = \frac{(a^2-1)}{4}, \tag{3.20}$$

where a is a constant ($\sqrt{B'} = a$).

We thus obtain

$$g_{00} = (X')^{-2}, \tag{3.21}$$

$$g_{11} = -r^{-2}\cdot\left(\frac{\log r}{d}\right)^{2h} \cdot\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)}\cdot(X')^2 \tag{3.22}$$

$$g_{22} = -\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)\cdot(X')^2} \tag{3.23}$$

and

$$g_{33} = -\left(\frac{\log r}{d}\right)^{2h}\cdot(X')^2, \tag{3.24}$$

where

$$X' = \left[2\left(\frac{b_r}{a^2C}\right)^{1/2}\cdot(\log r)^{1/2}\cdot\sinh\left(\frac{a}{2}\cdot\log(\log r^D)\right)\right].$$

C. Longitudinal fields

In this case ($b_r = b_\phi = 0$) the field equations are

$$r^2\eta_{11} + r\eta_1 = \frac{1}{C\log r}\cdot(b_z r^2 e^{2\beta}) - \frac{r\eta_1}{\log r}, \tag{3.25}$$

$$r^2\beta_{11} + r\beta_1 = \frac{1}{C\log r}\cdot(-b_z r^2 e^{2\beta}) - \frac{(1+r\beta_1)}{\log r}, \tag{3.26}$$

$$r^2\gamma_{11} + r\gamma_1 = \frac{1}{C\log r}\cdot(b_z r^2 e^{2\beta}) - \frac{r\gamma_1}{\log r}, \tag{3.27}$$

and

$$r^2\left[\eta_1\gamma_1 + (\eta_1 + \gamma_1)\left(\beta_1 + \frac{1}{r}\right)\right]$$
$$= -\frac{1}{\log r}\left(1 + \frac{\omega}{2\log r}\right) - \frac{r\lambda_1}{\log r} + \frac{1}{C\log r}(b_r\cdot r^2 e^{2\beta}) \tag{3.28}$$

along with (2.19).

Equation (3.26) gives

$$\beta = -\log\left[2\left(\frac{b_z}{B'C}\right)^{1/2}\cdot r(\log r)^{1/2}\cdot\cosh\left(\frac{\sqrt{B'}}{2}\cdot\log(\log r^D)\right)\right]$$
$$= -\log X'', \tag{3.29}$$

where B' and D are constants of integration.

The sum of (3.25) and (3.26) gives

$$\eta = -\beta + h\log\left(\frac{r^{-1/h}\cdot\log r}{d}\right) \tag{3.30}$$

with h and d as constants of integration.

The sum of (3.26) and (3.27) gives

$$\gamma = -\beta + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right), \tag{3.31}$$

p and f being constants of integration.

Finally substitution of (3.29)–(3.31) into (3.28) gives

$$ph + (p+h) + \omega/2 = (a^2-1)/4, \tag{3.32}$$

where a is a constant and ($\sqrt{B'} = a$).

We thus obtain

$$g_{00} = r^{-2}\cdot\left(\frac{\log r}{d}\right)^{2h}\cdot(X'')^2, \tag{3.33}$$

$$g_{11} = -r^{-4}\cdot\left(\frac{\log r}{d}\right)^{2h} \cdot\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)}\cdot(X'')^2, \tag{3.34}$$

$$g_{22} = -r^2\cdot(X'')^{-2}, \tag{3.35}$$

and

$$g_{33} = -r^{-2}\left(\frac{\log r}{f}\right)^{[(a^2-1)/2-\omega-2h]/(1+h)}\cdot(X'')^2, \tag{3.36}$$

where

$$X'' = \left[2\left(\frac{b_z}{a^2C}\right)^{1/2}\cdot r(\log r)^{1/2}\cdot\cosh\left(\frac{a}{2}\log(\log r^D)\right)\right].$$

4. CONSTRUCTION OF NEW SOLUTIONS

We consider, without loss of generality, the scalar Φ as a specific function Λ, viz., $\Phi = \Phi_0\Lambda$ in the original Brans–Dicke equations and subsequently scaling length, time, and reciprocal mass by the common factor $[\Lambda(x)]^{(1-\theta)/2}$ as follows:

$$L \to \bar{L} = \Lambda^{(1-\theta)/2}\cdot L, \tag{4.1a}$$

$$M \to \bar{M} = \Lambda^{-(1-\theta)/2}\cdot M, \tag{4.1b}$$

θ being a parameter. Under the transformation (4.1), we have

$$g_{ij} \to \bar{g}_{ij} = \Lambda^{(1-\theta)} \cdot g_{ij}, \tag{4.2a}$$

$$\Phi \to \bar{\Phi} = \bar{\Phi}_0 \cdot \Lambda = \Phi_0 \cdot \Lambda^{-(1-\theta)} \cdot \Lambda = \Phi_0 \cdot \Lambda^{\theta}. \tag{4.2b}$$

The above transformation can also be considered as a means for obtaining new solutions from the old ones.

We consider a new static, axially symmetric metric

$$ds^2 = e^{2\bar{\eta}} dt^2 - e^{2\bar{\lambda}} dr^2 - r^2 e^{2\bar{\beta}} d\phi^2 - e^{2\bar{\gamma}} dz^2 \tag{4.3}$$

with $\bar{\eta}$, $\bar{\lambda}$, $\bar{\beta}$, and $\bar{\gamma}$ as function of r only. Under the unit transformation (4.2) we get the values of $\bar{\eta}$, $\bar{\lambda}$, $\bar{\beta}$, and $\bar{\gamma}$.

A. Azimuthal fields

When we consider the solution (3.9)–(3.12) and (2.19) and apply the unit transformation (4.2), we obtain

$$\Phi \to \bar{\Phi} = \Phi_0 \Lambda^{\theta} = (C \log r)^{\theta},$$

where $\Lambda = C \log r$ and we have assumed $\Phi_0 = 1$ for simplicity. Finally, the solution in the present case is given by

$$\bar{\Phi} = (C \log r)^{\theta},$$

$$\bar{\eta} = \frac{1}{2} \log\left\{\left(\frac{2b_\phi}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + h \log\left(\frac{\log r}{d}\right), \tag{4.5}$$

$$\bar{\beta} = \frac{1}{2} \log\left\{\left(\frac{2b_\phi}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + p \log\left(\frac{r^{-1/p} \cdot \log r}{f}\right), \tag{4.6}$$

$$\bar{\gamma} = -\frac{1}{2} \log\left\{\left(\frac{2b_\phi}{BC}\right) \cdot \log r \cdot (C \log r)^{\theta-1} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\}, \tag{4.7}$$

and

$$\bar{\lambda} = \frac{1}{2} \log\left[\left(\frac{2b_\phi}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left\{\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right\}\right] + h \log\left(\frac{\log r}{d}\right) + p \log\left(\frac{r^{-1/p} \cdot \log r}{f}\right) \tag{4.8}$$

for the metrix (4.3). Clearly $\bar{\eta} + \bar{\beta} + \bar{\gamma} = \bar{\lambda} + (1-\theta) \log(C \log r)$, i.e. $\bar{\eta} + \bar{\beta} + \bar{\gamma} \neq \bar{\lambda}$ unless $\theta = 1$ (1961 form of Brans–Dicke theory).

B. Radial fields

In a similar manner the solution (3.21–(3.24) after unit transformation gives

$$\bar{\Phi} = (C \log r)^{\theta}, \tag{4.9}$$

$$\bar{\eta} = -\frac{1}{2} \log\left\{\left(\frac{2b_r}{BC}\right) \cdot \log r \cdot (C \log r)^{\theta-1} \cdot \sinh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\}, \tag{4.10}$$

$$\bar{\gamma} = \frac{1}{2} \log\left\{\left(\frac{2b_r}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \sinh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + h \log\left(\frac{\log r}{d}\right), \tag{4.11}$$

$$\bar{\beta} = \frac{1}{2} \log\left\{\left(\frac{2b_r}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \sinh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + p \log\left(\frac{r^{1/p} \cdot \log r}{f}\right), \tag{4.12}$$

and

$$\bar{\lambda} = \frac{1}{2} \log\left[\left(\frac{2b_r}{BC}\right) \cdot \log r \cdot (C \log r)^{1-\theta} \cdot \sinh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right] + h \log\left(\frac{\log r}{d}\right) + p \log\left(\frac{r^{1/p} \cdot \log r}{f}\right) \tag{4.13}$$

for the metric (4.3). Here again

$$\bar{\eta} + \bar{\beta} + \bar{\gamma} \neq \bar{\lambda}.$$

C. longitudinal fields

Through similar steps the solution (3.33)–(3.36) after unit transformation becomes

$$\bar{\Phi} = (C \log r)(C \log r)^{\theta-1}, \tag{4.14}$$

$$\bar{\beta} = -\frac{1}{2} \log\left\{\left(\frac{2b_z}{BC}\right) \cdot (r^2 \log r)(C \log r)^{\theta-1} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\}, \tag{4.15}$$

$$\bar{\eta} = \frac{1}{2} \log\left\{\left(\frac{2b_z}{BC}\right) \cdot (r^2 \log r) \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + h \log\left(\frac{r^{-1/h} \cdot \log r}{d}\right), \tag{4.16}$$

$$\bar{\gamma} = \frac{1}{2} \log\left\{\left(\frac{2b_z}{BC}\right) \cdot (r^2 \log r) \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + p \log\left(\frac{r^{-1/p} \cdot \log r}{f}\right), \tag{4.17}$$

and

$$\bar{\lambda} = \frac{1}{2} \log\left\{\left(\frac{2b_z}{BC}\right) \cdot (r^2 \log r) \cdot (C \log r)^{1-\theta} \cdot \cosh^2\left[\left(\frac{B}{2}\right)^{1/2} \cdot \log(\log r^D)\right]\right\} + h \log\left\{\frac{r^{-1/h} \log r}{d}\right\} + p \log\left\{\frac{r^{-1/p} \cdot \log r}{f}\right\} \tag{4.18}$$

for metric (4.3). Here again $\bar{\eta} + \bar{\beta} + \bar{\gamma} + \neq \bar{\lambda}$.

5. DISCUSSION OF THE SOLUTIONS

As pointed out in the introduction, the Brans–Dicke scalar electromagnetic fields are conformal to the coupled zero-mass scalar and source-free electromagnetic field of Einstein's gravitational theory. A similar result in the case of vacuum Brans–Dicke fields has been established by Peters[5] and Tabensky and Taub.[6] It has been shown that the conformal transformation

$$\exp\left\{\frac{\sqrt{2}V}{(\omega+\frac{3}{2})^{1/2}}\right\} = {}_{\rm BD}\Phi, \tag{5.1a}$$

$$g_{\mu\nu} = {}_{\rm BD}\Phi\cdot{}_{\rm BD}\,g_{\mu\nu}, \tag{5.1b}$$

where ${}_{\rm BD}\,\Phi$ and ${}_{\rm BD}\,g_{\mu\nu}$ are the quantities occurring in the Brans–Dicke theory, reduces the Brans–Dicke vacuum fields to zero-mass scalar fields of Einstein's gravitational theory and vice versa. The above transformation works again when the source-free electromagnetic field is also present in addition to the scalar fields.

In this case we again consider a new static, axially symmetric metric

$$ds^2 = e^{2\bar{\bar\eta}}dt^2 - e^{2\bar{\bar\lambda}}dr^2 - r^2e^{2\bar{\bar\beta}}d\phi^2 - e^{2\bar{\bar\gamma}}dz^2, \tag{5.2}$$

with $\bar{\bar\eta}$, $\bar{\bar\lambda}$, $\bar{\bar\beta}$, and $\bar{\bar\gamma}$ as functions of r only. After conformal transformation we get the values of $\bar{\bar\eta}$, $\bar{\bar\lambda}$, $\bar{\bar\beta}$, and $\bar{\bar\gamma}$.

A. Azimuthal fields

We consider the solution (3.9)–(3.12). Applying the conformal transformation (5.1), we obtain

$$V=\frac{(\omega+\frac{3}{2})^{1/2}}{\sqrt{2}}\log(C\log r).$$

Finally, the solution in the present case is given by

$$V=\frac{(\omega+\frac{3}{2})^{1/2}}{\sqrt{2}}\cdot\log(C\log r) \tag{5.3}$$

$$\bar{\bar\eta}=\frac{1}{2}\log\left[\left(\frac{2b_\phi}{B}\right)\cdot(\log r)^2\cdot\cosh^2\left\{\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right\}\right] + h\log\left(\frac{\log r}{d}\right), \tag{5.4}$$

$$\bar{\bar\beta}=\frac{1}{2}\log\left[\left(\frac{2b_\phi}{B}\right)(\log r)^2\cdot\cosh^2\cdot\left\{\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right\}\right] + p\log\left\{\frac{r^{-1/p}\cdot\log r}{f}\right\}, \tag{5.5}$$

$$\bar{\bar\gamma}= -\frac{1}{2}\log\left\{\left(\frac{2b_\phi}{BC^2}\right)\cdot\cosh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\}, \tag{5.6}$$

$$\bar{\bar\lambda}=\frac{1}{2}\log\left\{\frac{2b_\phi}{B}\right\}\cdot(\log r)^2\cdot\cosh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right] + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right) + h\log\left(\frac{\log r}{d}\right) \tag{5.7}$$

for the metric (5.2). Here $\bar{\bar\eta}+\bar{\bar\beta}+\bar{\bar\gamma}=\bar{\bar\lambda}+\log(C\log r)$, i.e., $\bar{\bar\eta}+\bar{\bar\beta}+\bar{\bar\gamma}\neq\bar{\bar\lambda}$.

B. Radial fields

Similarly for the solution (3.21)–(3.24), after conformal transformation we have

$$V=\frac{(\omega+\frac{3}{2})^{1/2}}{\sqrt{2}}\cdot\log(C\log r), \tag{5.8}$$

$$\bar{\bar\eta}= -\frac{1}{2}\log\left\{\left(\frac{2b_r}{BC^2}\right)\cdot\sinh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\}, \tag{5.9}$$

$$\bar{\bar\chi}=\frac{1}{2}\log\left\{\left(\frac{2b_r}{B}\right)(\log r)^2\cdot\sinh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} \times h\log\left(\frac{\log r}{d}\right), \tag{5.10}$$

$$\bar{\bar\beta}=\frac{1}{2}\log\left\{\left(\frac{2b_r}{B}\right)\cdot(\log r)^2\cdot\sinh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right), \tag{5.11}$$

$$\bar{\bar\lambda}=\frac{1}{2}\log\left\{\left(\frac{2b_r}{B}\right)\cdot(\log r)^2\cdot\sinh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right) + h\log\left(\frac{\log r}{d}\right) \tag{5.12}$$

for the metric (5.2). Here $\bar{\bar\eta}+\bar{\bar\beta}+\bar{\bar\gamma}\neq\bar{\bar\lambda}$

C. Longitudinal fields

Through similar steps the solution (3.33)–(3.36) after conformal transformation give

$$V=\frac{(\omega+\frac{3}{2})^{1/2}}{\sqrt{2}}\cdot\log(C\log r), \tag{5.13}$$

$$\bar{\bar\beta}= -\frac{1}{2}\log\left[\left(\frac{2b_z}{BC^2}\right)\cdot r^2\cdot\cosh^2\left\{\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right\}\right], \tag{5.14}$$

$$\bar{\bar\eta}=\frac{1}{2}\log\left\{\left(\frac{2b_z}{B}\right)\cdot(r\log r)^2\cdot\cosh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} + h\log\left(\frac{r^{-1/h}\cdot\log r}{d}\right), \tag{5.15}$$

$$\bar{\bar\gamma}=\frac{1}{2}\log\left\{\left(\frac{2b_z}{B}\right)\cdot(r\log r)^2\cdot\cosh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right), \tag{5.16}$$

and

$$\bar{\bar\lambda}=\frac{1}{2}\log\left\{\left(\frac{2b_z}{B}\right)\cdot(r\log r)^2\cdot\cosh^2\left[\left(\frac{B}{2}\right)^{1/2}\cdot\log(\log r^D)\right]\right\} + h\log\left(\frac{r^{-1/h}\cdot\log r}{d}\right) + p\log\left(\frac{r^{-1/p}\cdot\log r}{f}\right) \tag{5.17}$$

for the metric (5.2). Clearly $\bar{\bar\eta}+\bar{\bar\beta}+\bar{\bar\gamma}\neq\bar{\bar\lambda}$ again.

We observe that the solutions given by (5.2)–(5.17) for coupled zero-mass and electromagnetic fields in Einstein's theory are different from those of Teixeira *et al.*[16] Thus the solutions for Brans–Dicke–Maxwell fields in the present work are new, and they cannot be obtained from the known solutions of Einstein–Maxwell scalar fields by conformal transformation.

6. CONCLUSIONS

In Secs. 2 and 3 we have found that static cylindrically symmetric systems may contain exclusively radial, azimuthal, or longitudinal electromagnetic fields and not combination of these fields. The choice of coordinates $r^2 g_{rr} = -g_{tt} \cdot g_{\phi\phi} \cdot g_{zz}$ has an advantage over other more frequently used such as $g_{rr} = g_{zz}$ or $g_{\phi\phi} = r^2 g_{rr}$ in the sense that it gives the field equations in a form which is easy to solve.

Janis *et al.* [19] have given a method of obtaining some Brans–Dicke–Maxwell fields from the Einstein scalar field solutions irrespective of any symmetry, but this method does not admit combination of electric and magnetic fields. However, our solutions allow combinations of electric and magnetic fields.

[1]C.H. Brans and R.H. Dicke, Phys. Rev. **124**, 925 (1961).

[2]C.H. Brans, Phys. Rev. **125**, 2194 (1962).

[3]R.H. Dicke, Phys. Rev. **125**, 2163 (1962).

[4]H.Y. Chiu and W.F. Hoffman, *Gravitation and Relativity* (Benjamin, New York, 1964), p. 121.

[5]P.C. Peters, J. Math. Phys. **10**, 1029 (1969).

[6]R. Tabensky and A.H. Taub, Commun. Math. Phys. **29**, 61 (1973).

[7]M.N. Mahanta and D.R.K. Reddy, J. Math. Phys. **13**, 708 (1972).

[8]S.K. Luke and G. Szamosi, Phys. Rev. D **6**, 3359 (1972).

[9]J.R. Rao, R.N. Tiwari, and K.S. Bhamra, Ann. Phys. (N.Y.) **87**, 480 (1974).

[10]J.R. Rao, R.N. Tiwari, and K.S. Bhamra, Ann. Phys. (N.Y.) **90**, 573 (1975).

[11]H.A. Buchdahl, Nuovo Cimento B **12**, 269 (1972).

[12]A.K. Raychoudhari and N. Bandyopadhyay, Prog. Theor. Phys. (Japan) **59**, 414 (1978).

[13]T. Singh and L.N. Rai, "Stationary axially symmetric Brans–Dicke–Maxwell fields," Gen. Rel. Grav. J. (to be published).

[14]J.R. Rao, A.R. Roy, and R.N. Tiwari, Ann. Phys. (N.Y.) **69**, 473 (1972).

[15]J.R. Rao, R.N. Tiwari, and A.R. Roy, Ann. Phys. (N.Y.), **78**, 553 (1973).

[16]A.F. da F. Teixeira, Idel Wolk, and M.M. Som, J. Math. Phys. **15**, 1756 (1974).

[17]R.E. Morganstern, Phys. Rev. D **3**, 2946 (1971).

[18]J.L. Synge, *Relativity: The General Theory* (North-Holland, Amersterdam, 1960), Chap. 8.

[19]A.I. Janis, D.C. Robinson, and J. Winicour, Phys. Rev. **186**, 1729 (1969).

Dissipative operators for infinite classical systems and equilibrium

J. V. Pulè and A. Verbeure
Instituut voor Theoretische Fysica, K. U. Leuven, B-3030 Leuven, Belgium

(Received 26 February 1979; accepted for publication 8 June 1979)

Dissipative differential operators for infinite classical systems are characterized and used to obtain correlation inequalities for equilibrium states.

I. INTRODUCTION

Recently there has been intense activity in the context of irreversibility of quantum mechanical systems and particularly in the study of semigroups of completely positive maps[1–5] to describe irreversible evolutions. In addition to this work, these semigroups have been used to construct explicit perturbations around the equilibrium states and to formulate variational principles.[6–7] Also, from the mathematical point of view the generators of semigroups of completely positive maps have been characterized completely.[2–3]

Much less work has been done for infinite classical systems. In this case the algebra of observables is Abelian and the notion of complete positivity coincides with the notion of positivity. For convolution semigroups on measures the generators have been characterized by Feller.[8] In this paper we study dissipative operators on the Abelian algebra for infinite classical systems and we concentrate on differential operators as these seem to be the most relevant ones to physics, e.g., the Fokker–Planck equation and the classical limit of the quantum mechanical dissipative operators.[9]

In Sec. 2 the dissipative differential operators for these infinite systems are characterized and shown to be, as for finite systems, of the elliptic and parabolic type, while in Sec. 3 these operators are used to obtain correlation inequalities characterizing equilibrium states. Finally we indicate how these correlation inequalities can also be obtained from thermodynamic stability.

The algebra of observables of the infinite classical system is similar to that of Ref. 10. We denote by Λ a bounded open set of R^ν, $\nu \in N$. Let $\mathscr{X}$ be the set of infinite countable subsets $x=(x_i, i\in N)$ of $R^\nu \times R^\nu$, $x_i=(q_i, p_i)$, such that $x \cap \Lambda \times R^\nu$ contains only a finite number of elements. $\mathscr{X}$ is called the set of configurations of the infinite system, and $\mathscr{X}$ is made into a measurable space as in Ref. 11.

Denote by $\widetilde{\mathscr{A}}$ the real algebra under pointwise multiplication of unbounded observables generated by the functions Sf of $\mathscr{X}$ into R which are described by a sequence $(f^{(m)})_{m=0}^\infty$ of C^∞-functions $f^{(m)}$ from $(R^\nu \times R^\nu)^m$ in R which are symmetric, have compact support, and with a finite number of components different from zero. Then for any $x \in \mathscr{X}$ and $Sf \in \widetilde{\mathscr{A}}$,

$$(Sf)(x) = \sum_{m=0}^{\infty} \sum_{i_1,\cdots,i_m} f^{(m)}(x_{i_1},\cdots,x_{i_m}).$$

For simplicity, if $g \in C_0^\infty[(R^\nu \times R^\nu)^m]$ then by Sg we denote Sf where $f^{(n)}=0$ for $n\neq m$ and $f^{(m)}=g$.

We define the Poisson bracket $\{\cdot,\cdot\}$ as the bilinear map of $\widetilde{\mathscr{A}}\times\widetilde{\mathscr{A}}$ into $\widetilde{\mathscr{A}}$ given by, for all ϕ and ψ elements of $\widetilde{\mathscr{A}}$,

$$\{\phi,\psi\}(x) = \sum_i \frac{\partial\phi}{\partial q_i}\cdot\frac{\partial\psi}{\partial p_i}(x) - \frac{\partial\phi}{\partial p_i}\cdot\frac{\partial\psi}{\partial q_i}(x),$$

where $\partial/\partial q_i \cdot \partial/\partial p_i$ stands for the scalar product of the gradients.

Furthermore, let $\mathscr{F}$ be the set of states on $\widetilde{\mathscr{A}}$, i.e., the positive linear functionals on $\widetilde{\mathscr{A}}$ given by the probability measures ω on $\mathscr{X}$, such that all elements of $\widetilde{\mathscr{A}}$ belong to $L_1(\mathscr{X},\omega)$.

2. CHARACTERIZATION OF DISSIPATIVE DIFFERENTIAL OPERATORS

For simplicity of notation we restrict ourselves to one-dimensional systems, i.e., $\nu=1$.

Let us denote

$$D^q_{i_1,\cdots,i_t} = \frac{\partial^t}{\partial q_{i_1}\cdots\partial q_{i_t}},$$
$$D^p_{i_1,\cdots,i_t} = \frac{\partial^t}{\partial p_{i_1}\cdots\partial p_{i_t}},$$

and consider the following differential operator Δ from $\widetilde{\mathscr{A}}$ into $\widetilde{\mathscr{A}}$ defined as follows. For all $x=(x_i)\in\mathscr{X}$ and $f\in\widetilde{\mathscr{A}}$,

$$(\Delta f)(x) = \sum_{n=0}^{N} (\Delta_{(n)} f)(x), \tag{1}$$

where

$$(\Delta_{(0)} f)(x) = A_{(0)}(x) f(x),$$
$$(\Delta_{(n)} f)(x) = \sum_{i_1,\cdots,i_n} \sum_{t=0}^{n} A^t_n(x;x_{i_1},\cdots x_{i_n}) \times D^q_{i_1,\cdots,i_t} D^p_{i_{t+1},\cdots i_n} f,$$

where the functions $x \to A^t_n(x;y_1\cdots y_n)$ for fixed $y_1,\ldots,y_n$ are elements of $\widetilde{\mathscr{A}}$ and the functions $(y_1,\cdots,y_n) \to A^t_n(x;y_1,\cdots,y_n)$ for fixed $x\in\mathscr{X}$ are C^∞-functions of compact support on $(R^{2\nu})^n$, which are symmetric in the first t arguments and in the last $n-t$ arguments.

Definition 2.1: The differential operator Δ given in (1) is called dissipative if for all $f\in\widetilde{\mathscr{A}}$,

$$\mathscr{D}(\Delta,f) = \Delta(f^2) - 2f\Delta f \geqslant 0.$$

Theorem 2.2: If Δ is a differential operator as in (1), then Δ is dissipative if and only if

(i) $\Delta_{(n)} = 0$ for $n > 2$,

(ii) $A_0(x) \leqslant 0$,

(iii) for any indices r,s

$$0 \leqslant \sum_{i,j=r,s} \{A_2^0(x;x_i,x_j)Q_iQ_j + A_2^1(x;x_i;x_j)Q_iP_j + A_2^2(x;x_i;x_j)P_iP_j\},$$

where the Q_i and P_i are arbitrary numbers.

Proof: Suppose first that (i)–(iii) are satisfied, then since

$$\mathscr{D}(\Delta,f) = -A_0(x)f(x) + \sum_{i,j} A_2^0(x;x_i,x_j)\frac{\partial f}{\partial q_i}\frac{\partial f}{\partial q_j} + A_2^1(x;x_i,x_j)\frac{\partial f}{\partial q_i}\frac{\partial f}{\partial p_j} + A_2^2(x;x_i,x_j)\frac{\partial f}{p_i}\frac{\partial f}{p_j}, \quad (2)$$

it follows that Δ is dissipative.

Suppose now that Δ is dissipative. For $N=2$, for any fixed $x \in \mathscr{K}$, choose first a function $g \in C_0^\infty(R^2)$ such that $g(t)=1$ in a neighborhood of x_1 and for $i \neq 1$ x_i is not in support of g. In this case, formula (2) reduces to

$$\mathscr{D}(\Delta,Sg)(x) = -A_0(x) \geqslant 0$$

which yields (ii).

For any pair of points x_r, x_s in x, and any choice of real numbers Q_r, Q_s, P_r, P_s choose a function $g \in C_0^\infty(R^2 \times R^2)$ such that the points (x_i,x_j), $i,j \neq r,s$ are not in the support of g,

$$\frac{\partial g(x_r,x_s)}{\partial q_r} = Q_r, \quad \frac{\partial g(x_r,x_s)}{\partial q_s} = Q_s,$$

$$\frac{\partial g(x_r,x_s)}{\partial p_r} = P_r, \quad \frac{\partial g(x_r,x_s)}{\partial p_s} = P_s,$$

and such that $g(x_r,x_s)$ vanishes. In this case using formula (2), the dissipativity is expressed by (iii). Hence for $N=2$, (i), (ii), and (iii) follow from dissipativity. Suppose now $N>2$. We shall prove that $\Delta_N = 0$, $\Delta_{(N-1)} = 0, \cdots$, up to $\Delta_{(3)} = 0$ consecutively. As $\Delta = \Sigma_{n=0}^N \Delta_{(n)}$,

$$\mathscr{D}(\Delta,f) = \sum_{n=0}^N \mathscr{D}(\Delta_{(n)},f).$$

To prove that $\Delta_{(N)} = 0$ we show that for each $x \in \mathscr{K}$ and for each subset of elements $(x_{s_1},\cdots,x_{s_N})$ of x not necessarily distinct,

$$A_N^r(x;x_{s_1},x_{s_2},\cdots,x_{s_N}) = 0 \quad \text{for } r = 0,\cdots,N. \quad (3)$$

Suppose first that $r>0$ and that there are k different elements in the set $\{x_{s_1},\cdots x_{s_N}\}$ which we shall denote by $y_1,\cdots,y_k$. Take a function $g \in C_0(R^{2k})$, such that $(x_{i_1},\cdots,x_{i_k})$ does not belong to the support of g unless $x_{i_t} = y_t$, $t = 1,\cdots,k$ and such that $(\partial g/\partial q_{s_1})(y_1,\cdots y_k) = 1$ and the other first derivatives with respect to the q's and p's vanish, together with all higher derivatives except

$$\frac{\partial^{N-1} g(y_1,\cdots,y_k)}{\partial q_{s_2}\cdots\partial q_{s_r}\,\partial p_{s_{r+1}}\cdots\partial p_{s_N}} = v \neq 0. \quad (4)$$

Then

$$\mathscr{D}(\Delta_{(N)},Sg)(x) = c\,A_N^r(x;x_{s_1},\cdots,x_{s_N})v,$$

where c is some nonzero integer. The only contribution from the t-summation in formula (1) comes from $t=r$ since this is fixed by the number of p-derivatives in formula (4), while the other summation gives rise to the constant c. Furthermore, $\mathscr{D}(\Delta_{(n)},Sg)(x) = 0$, except for $n=N$ and $n=0,2$. In particular $\mathscr{D}(\Delta_{(0)},Sg)(x) = -A_0(x)\,Sg(x)$ and hence is independent of v. Since $\mathscr{D}(\Delta_{(2)},Sg)(x)$ is also independent of v,

$$\mathscr{D}(\Delta,Sg)(x) = -A_0(x)\,Sg(x) + A_2^0(x;x_{s_1},x_{s_1}) + c\,A_N^r(x;x_{s_1},\cdots,x_{s_N})v.$$

As v is arbitrary (3) follows from dissipativity. The case $r=0$ can be treated similarly by interchanging the role of the p's and q's. By repeating this argument we show that $\Delta_{N-1} = \cdots = \Delta_3 = 0$. □

For systems with one degree of freedom the phase space is R^2 and we denote its elements by $x = (q,p) \in R^2$. In this case, dissipative differential operators are at most of second order, and if

$$(\Delta f)(x) = \alpha(x)f(x) + \beta(x)\frac{\partial f(x)}{\partial q} + \gamma(x)\frac{\partial f(x)}{\partial p} + a(x)\frac{\partial^2 f(x)}{\partial q^2} + b(x)\frac{\partial^2 f}{\partial q\partial p} + c(x)\frac{\partial^2 f}{\partial p^2}, \quad (5)$$

the conditions for dissipativity are

$\alpha \leqslant 0$, $0 \geqslant a$, $0 \geqslant c$, $b^2 \leqslant 4ac$.

Equation (5) can be written in terms of Poisson brackets,

$$(\Delta f)(x) = \alpha(x)f(x) + \beta(x)\{-p,f\}(x) + \gamma\{q,f\} + a(x)\{p,\{p,f\}\}(x) + b(x)\{-p,\{q,f\}\}(x) + c(x)\{q,\{q,f\}\}(x). \quad (6)$$

For the infinite systems we cannot express the most general dissipative operator in terms of Poisson brackets as is done in (6). In the next section we discuss dissipative perturbations of KMS states. As the KMS equation is usually defined through the Poisson bracket we are led to consider only a suitable subclass of dissipative differential operators of the type, for $\phi \in \widetilde{\mathscr{A}}$, $x \in \mathscr{K}$,

$$(\Delta\phi)(x) = A(x)\,\phi(x) + \sum_{i=1}^n A_i(x)\,\{B_i,\phi\}(x) + \sum_{i,j=1}^n C_{ij}(x)\,\{D_i,\{D_j,\phi\}\}(x), \quad (7)$$

where A, A_i, B_i, C_{ij}, D_i, $i,j = 1,\cdots,n$ are all elements of $\widetilde{\mathscr{A}}$. The following conditions are sufficient for dissipativity: For each $x \in \mathscr{K}$

(i) $A(x) \leqslant 0$,

(ii) the matrix $(C_{ij}(x))_{i,j=1,\cdots n}$ is positive.

Before proceeding we want to point out that differential operators of the type (7) have been obtained as the classical limit of quantum mechanical dissipative generators in Ref. 9 where as a result a generalized Fokker–Planck differential operator is obtained.

3. KMS AND CORRELATION INEQUALITIES

Let us first introduce the set of interactions which will be considered. Suppose that for all $\Lambda \subset R^\nu$, the local Hamiltonian H_Λ with open boundary conditions is given by: for any configuration $X = (x_1,\cdots,x_n)$, $x_i \in \Lambda \times R^\nu$,

$$H_\Lambda(x) = \int_\Lambda da\,(Sh)(X+a),$$

where

$$Sh(X) = (St)(X) + (Sv)(X),$$

$$t(q,p) = \epsilon(q)\, p^2/2m,$$

ϵ is a C_0^∞-function on R^ν which support the unit cube centered around zero and such that,

$$\int \epsilon(q)\, dq = 1, \quad Sv \text{ is some element of } \widetilde{\mathscr{A}}.$$

The motivation for this form of the Hamiltonian is found in Refs. 12 and 13. However, here we have the more restrictive condition of finite range interaction. Because of this condition for each local observable $\phi \in \widetilde{\mathscr{A}}$ the limit

$$\lim_{\Lambda \to \infty} \{H_\Lambda, \phi\}$$

exists and is denoted by $L\phi \equiv \lim_{\Lambda \to \infty} \{H_\Lambda, \phi\}$.

Definition 3.1: A state $\omega \in \mathscr{F}$ satisfies the static KMS condition[14,15] at inverse temperature β for a Hamiltonian H if for all $F,G \in \widetilde{\mathscr{A}}$

$$\beta\,\omega(FLG) = \omega(\{F,G\}).$$

Theorem 3.2: With the Hamiltonian of above, let ω be any state of $\mathscr{F}$, then the following are equivalent:

(i) ω is a KMS state at inverse temperature β;

(ii) for all dissipative differential operators Δ_0 as in (7) with $A = 0, A_i, B_i, C_{ij}, D_i (i,j = 1,\cdots,n) \in \widetilde{\mathscr{A}}$:

$$\beta\,\omega(\Delta_0 H) \geqslant \omega\Big(\sum_{i=1}^n \{B_i, A_i\}\Big) - \omega\Big(\sum_{i,j} \{D_j,\{D_i;C_{ij}\}\}\Big); \quad (8)$$

(iii) for all dissipative differential operators as in (7) with $C_{ij}, D_i (i,j = 1,\cdots,n)$ elements of $\widetilde{\mathscr{A}}$:

$$\beta\,\omega\Big(\sum_{i,j=1}^n C_{ij}\{D_i\{D_j,H\}\}\Big) \geqslant -\omega\Big(\sum_{i,j}\{D_j,\{D_i,C_{ij}\}\}\Big). \quad (9)$$

If these conditions are satisfied, then for all complex observables $X,Y \in \widetilde{\mathscr{A}} + i\,\widetilde{\mathscr{A}}$,

$$\beta\,\omega(\{\bar{Y},\{Y,H\}\})\,\omega(|X|^2) \geqslant |\omega(\{X,Y\})|^2. \quad (10)$$

Proof (i) implies (ii): Since $(C_{ij})_i, j = 1,\cdots,n$ is a positive matrix

$$0 \leqslant \beta^2\omega\Big(\sum_{i,j=1}^n C_{ij}\{H,D_i\}\{H,D_j\}\Big).$$

Applying (i),

$$0 \leqslant \beta\,\omega\Big(\sum_{i,j=1}^n \{C_{ij}\,\{H,D_j\},D_i\}\Big).$$

Hence again using (i):

$$\begin{aligned} 0 &\leqslant \beta\,\omega\Big(\sum_{i,j=1}^n \{C_{ij},D_i\}\,\{H,D_j\}\Big) \\ &\quad + \beta\,\omega\Big(\sum_{i,j=1}^n C_{ij}\{\{H,D_j\},D_i\}\Big) \\ &= \omega\Big(\sum_{i,j=1}^n \{\{C_{ij},D_i\},D_j\}\Big) \\ &\quad + \beta\,\omega\Big(\sum_{i,j=1}^n C_{ij}\{D_i,\{D_j,H\}\}\Big). \end{aligned}$$

Therefore,

$$\beta\,\omega\Big(\sum_{i,j=1}^n C_{ij}\{D_i,\{D_j,H\}\}\Big) \geqslant -\omega\Big(\sum_{i,j=1}^n \{D_j,\{D_i,C_{ij}\}\}\Big).$$

Also,

$$\beta\,\omega\Big(\sum_i A_i\{B_i,H\}\Big) = \omega\Big(\sum_i \{B_i,A_i\}\Big)$$

yielding (ii).

That (ii) implies (iii) is trivial by putting $A_i = 0$ $(i = 1,\cdots,n)$.

Now (iii) implies (i): Take $n = 2$ in (iii) and

$$\begin{pmatrix} C_{11}(x) & C_{12}(x) \\ C_{21}(x) & C_{22}(x) \end{pmatrix} = \tilde{C}(x)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

where $\tilde{C}(x) \in \widetilde{\mathscr{A}}$, then

$$\begin{aligned} &\beta\,\omega(\tilde{C}\{D_1,\{D_2,H\}\}) + \omega(\{D_2,\{D_1,\tilde{C}\}\}) \\ &\quad - \beta\,\omega(\tilde{C}\{D_2\{D_1,H\}\}) - \omega(\{D_1,\{D_2,\tilde{C}\}\}) \geqslant 0. \end{aligned}$$

Using the Jacobi identity for the Poisson brackets,

$$\beta\,\omega(\tilde{C}\{H,\{D_1,D_2\}\}) - \omega(\{\tilde{C},\{D_1,D_2\}\}) \leqslant 0.$$

As this inequality also holds for $-D_1$ instead of D_1 one gets the equality

$$\beta\,\omega(\tilde{C}\{H,\{D_1,D_2\}\}) = \omega(\{\tilde{C},\{D_1,D_2\}\}).$$

This is the KMS equation for the arbitrary observable $\tilde{C}$ and the particular ones of the type $\{D_1,D_2\}$. From this we will deduce that the KMS equation holds for all observables. From linearity it is sufficient to prove it for observables of the type Sg where $g \in C_0^\infty((R^{2\nu})^n)$. Suppose the compact set Λ contains the support of $\tilde{C}$ and the support of $\{H,\tilde{C}\}$. Take $D_1 = Sf, f \in C_0^\infty(R^{2\nu})$ and such that $f(q,p) = \Sigma_{\alpha=1}^{\nu} p^\alpha$ for $(q,p) \in \Lambda$; take $D_2 = Sh$, $h \in C_0^\infty((R^{2\nu})^n)$ and such that

$$h(x_1,\cdots,x_n) = \int_{-\infty}^{0} g(q_1 + se, p_1, q_2 + se, p_2, \cdots, q_n + se, p_n)\, ds$$

for $x_1,\cdots,x_n \in \Lambda$, where $e = (1,1,\cdots,1) \in R^\nu$. Then on Λ

$$g(x_1,\cdots,x_n) = \sum_{i=1}^n \sum_{\alpha=1}^{\nu} \frac{\partial h}{\partial q_i^\alpha}(x_1,\cdots,x_n)$$

and therefore, for any $A \in \widetilde{\mathscr{A}}$ with support in Λ,

$$A\{D_1,D_2\} = A\,Sg. \quad (*)$$

Furthermore, applying $\{H,\cdot\}$ to $(*)$ with $A = \tilde{C}$,

$$\begin{aligned} &\{H,\tilde{C}\}\{D_1,D_2\} + \tilde{C}\{H,\{D_1,D_2\}\} \\ &\quad = \{H,\tilde{C}\}Sg + \tilde{C}\{H,Sg\}. \end{aligned}$$

Using again $(*)$ with $A = \{H,\tilde{C}\}$ gives

$$\tilde{C}\{H,\{D_1,D_2\}\} = \tilde{C}\{H,Sg\}$$

yielding the KMS equation

$$\beta\,\omega(\tilde{C}\{H,Sg\}) = \omega(\{\tilde{C},Sg\}).$$

Finally, we derive inequality (10) from (i). Let $(\cdot,\cdot)$ be the inner product on $\widetilde{\mathscr{A}} + i\widetilde{\mathscr{A}}$ defined by the state ω: for all

$$X,Y \in \widetilde{\mathscr{A}} + i\widetilde{\mathscr{A}},$$
$$(X,Y) = \omega(\bar{X},Y).$$

By (i) and Schwartz's inequality,

$$|\omega(\{X,Y\})|^2 = \beta^2|(\bar{X},\{Y,H\})|^2$$

$$\leqslant \beta^2 (\bar{X},X)\,(\{Y,H\},\{Y,H\})$$
$$= \beta^2\,\omega(|X|^2)\,\omega(\{\bar{Y},H\}\ \{Y,H\})$$
$$= \beta\,\omega\,(|X|^2\,\omega(\{\bar{Y},\{Y,H\}\}). \qquad \square$$

At first sight the inequalities (8) and (9) appear fundamentally different in the sense that (8) contains a terms with a single and double Poisson bracket, whereas (9) contains only a double Poisson bracket term. However, written out in terms of differential operators, both (8) and (9) contain first order differential operators, which are essential to establish the equivalence with the KMS equation. This can be seen clearly in the case of a free system with one degree of freedom. The inequality containing only second derivatives is then, using (5),

$$\beta\,\omega\left(a\frac{\partial^2 H}{\partial q^2} + b\frac{\partial^2 H}{\partial q \partial p} + c\frac{\partial^2 H}{\partial p^2}\right)$$
$$\geqslant -\,\omega\left(\frac{\partial^2 a}{\partial q^2} + \frac{\partial^2 b}{\partial q \partial p} + \frac{\partial^2 c}{\partial p^2}\right),$$

where $H = p^2/2$. This inequality is clearly satisfied by any density matrix state ω, given by $\omega(f) = \int_{R^2} dq\,dp\,C$, $e^{-\alpha p^2/2} f(q,p)$, where $\alpha \leqslant \beta$ and C is a normalization constant.

Note that inequality (10) by itself does not imply the KMS property, e.g., the state with constant density matrix satisfies (10), though it does imply time invariance, as is immediately seen by putting $Y = H$. On the other hand, by putting X, a constant, in (10), we recover

$$\omega\,(\{\bar{Y},\{Y,H\}\}) \geqslant 0,$$

which is the condition for energetic stability.[16] Therefore, inequality (10) supplemented by some cluster property implies the KMS condition.[16]

Above we have proved the equivalence between the KMS property and the correlation inequalities (8) and (9). As in the quantum mechanical case,[7] one expects that these inequalities are a way of expressing notions of (thermodynamic) stability. In principle we could proceed as in Ref. 17 and obtain these inequalities from stability, but here the dissipative generators are second order differential operators instead of first order, and their solutions cannot be approximated by phase space perturbations. This makes it technically much more involved. Therefore, we limit ourselves to finite systems. The method can be generalized to the infinite case if enough technical conditions are imposed on the dynamical systems, but this does not yield more physical insight.

For finite systems with n degrees of freedom the phase space $\mathscr{K}$ is now R^{2n}, as algebra of observables we take $\mathscr{D}(R^{2n})$, and we limit ourselves to the set $\mathscr{E}$ of density matrix states $\omega(\cdot) = \int dx\,\rho(x)$ where $\rho(x) \geqslant 0$, $\int \rho(x)\,dx = 1$, and $|\int dx\,\rho(x) \log \rho(x)| < \infty$. Let H be the Hamiltonian of the system which is twice continuously differentiable. Let

$$F(\rho) = \beta\,\omega\,(H) + \omega\,(\log \rho)$$

be the free energy. Then we have the following result.

Theorem 3.3: Suppose that $\rho_0 \in \mathscr{E}$, and $x \to \rho_0(x)$ is twice continuously differentiable, and ρ_0 satisfies the variational principle, i.e., for all $\rho \to \mathscr{E}$, $F(\rho) \geqslant F(\rho_0)$. Then,

(i) for all dissipative differential operators Δ_0 as in (7) with $A = 0, A_i, B_i, C_{ij}, D_i$ $(i,j = 1,\cdots,n)$ in $\mathscr{D}(R^{2n})$,

$$\beta\,\omega_0(\Delta_0 H) \geqslant \omega_0\Big(\sum_{i=1}^{n} \{B_i,A_i\}\Big) - \omega_0\Big(\sum_{ij} \{D_j,\{D_i,C_{ij}\}\}\Big)$$

where $\omega_0(\cdot) = \int dx\,\rho_0(x)$;

(ii) for all complex observables $X,Y \in \mathscr{D}(R^{2n}) + i\mathscr{D}(R^{2n})$,

$$\beta\,\omega_0(\{\bar{Y},\{Y,H\}\})\omega_0\,(|X|^2) \geqslant |\omega_0(\{X,Y\})|^2.$$

Proof: First we prove that any density matrix satisfying the variational principle is strictly positive. Suppose that ρ_0 vanishes at a point, which without loss of generality we take as $x = 0$. Let σ be the density matrix with support in a compact set Λ containing the origin as an interior point.

We define now for $1 > \epsilon > 0$

$$\rho_\epsilon(x) = (1-\epsilon)\rho_0(x) + \epsilon\,\sigma_\epsilon(x),$$

where $\sigma_\epsilon(x) = \epsilon^{-1/2}\sigma(\epsilon^{-1/(4n)}x)$. One calculates (put $\beta = 1$),

$$\frac{F(\rho_\epsilon) - F(\rho_0)}{\epsilon}$$
$$= \int [\sigma_\epsilon(x) - \rho_0(x)]\,H(x)\,dx + \frac{1-\epsilon}{\epsilon}\int dx\,\rho_0(x)$$
$$\times [\log \rho_\epsilon(x) - \log \rho_0(x)] - \int dx \rho_0(x) \log \rho_0(x)$$
$$+ \int dx\,\sigma_\epsilon(x) \log \rho_\epsilon(x).$$

Now we consider the limit $\epsilon \to 0$. The first term tends to

$$H(0) - \int dx\,\rho_0(x)\,H(x),$$

which is bounded above as a consequence of the variational principle and the continuity of H. Using the inequality $\log x \leqslant x - 1$, the second term is bounded above by zero. The third term is bounded. The last term is equal to

$$\int_\Lambda du\,\sigma(u) \log\,[(1-\epsilon)\rho_0(\epsilon^{1/4n}u) + \epsilon^{1/2}\,\sigma(u)],$$

which is bounded above by

$$\log\Big[\sup_{u\in\Lambda} \rho_0(\epsilon^{1/4n}u) + \epsilon^{1/2} \sup_{u\in\Lambda} \sigma(u)\Big],$$

which tends to minus infinity as $\epsilon \to 0$.

Therefore, for ϵ small enough

$$F(\rho_\epsilon - F(\rho_0) < 0,$$

contradicting the variational principle. Hence ρ_0 is strictly positive.

Now we proceed with the perturbation of the state ρ_0 satisfying the variational principle. Suppose that the compact set Λ contains the supports of the functions A_i, B_i, D_i, and C_{ij} $(i,j = 1,\cdots,n)$. From above there exists a $\delta > 0$ such that $\rho(x) > \delta$ for all $x \in \Lambda$. Hence there exists a $\lambda_0 > 0$ such that for all $\lambda \in [0,\lambda_0]$,

$$\rho_\lambda = \rho_0 + \lambda \Delta_0^{*} \rho_0$$

is strictly positive, where

$$\Delta_0^{*}\rho_0 = \sum_i \{A_i\rho_0,B_i\} + \sum_{i,j} \{D_j,\{D_i,C_{ij}\rho_0\}\}.$$

Also $\int\rho_\lambda(x) = 1$ and $|\int dx\, \rho_\lambda(x) \log \rho_\lambda(x)| < \infty$. Hence $\rho_\lambda \in \mathscr{E}$ and therefore

$$\lim_{\lambda \to 0} \frac{F(\rho_\lambda) - F(\rho_0)}{\lambda} \geqslant 0,$$

yielding

$$\int dx\, \Delta_0^* \rho_0(x) [H(x) + \log\rho_0(x)] \geqslant 0,$$

or

$$\omega_0(\Delta_0 H) + \omega_0(\Delta_0 \log \rho_0) \geqslant 0, \qquad (\alpha)$$

where $\omega_0(f) = \int dx\, \rho_0(x) f(x)$, $f \in \mathscr{D}$.

Let

$$f(x,y) = \{\rho(x) [\ln \rho(x) - \ln \rho(y)] - \rho(x) + \rho(y)\}^{1/2}$$

and consider

$$F(x,y) \equiv \Delta_0^y f(x,y)^2,$$

where Δ_0^y denotes Δ_0 applied to $y \to f(x,y)^2$.
Since Δ_0 is dissipative

$$\Delta_0^y f(x,y)^2 \geqslant 2f(x,y)\, \Delta_0^y f(x,y),$$

and therefore $F(x,x) \geqslant 0$ because $f(x,x) = 0$,
and hence $\int dx\, F(x,x) \geqslant 0$.
But on the other hand,

$$F(x,x) = -\rho_0(x)\, \Delta_0 \log \rho_0(x) + \Delta_0 \rho(x),$$

so that by integrating

$$\int dx\ F(x,x) = -\omega_0(\Delta_0 \log \rho_0) + \omega_0(\Delta_0^* 1) \geqslant 0. \quad (\beta)$$

Combining (α) and (β) one gets (i).

To derive (ii), use Schwartz inequality to obtain

$$\begin{aligned}
&\left| \int \rho_0(x) \{Y,X\}(x)\, dk \right|^2 \\
&= \left| \int \rho_0(x) \frac{\{\rho_0, Y\}}{\rho_0(x)} X(x)\, dx \right|^2 \\
&\leqslant \int \rho_0(x) \, |X|^2(x)\, dx \int \frac{\{\rho_0, Y\}\{\rho_0, \bar{Y}\}(x)}{\rho_0(x)} dx \\
&= \omega_0(|X|^2)\, \omega_0(\widetilde{\Delta}_0 \log \rho_0), \qquad (\gamma)
\end{aligned}$$

where $\widetilde{\Delta}_0 f = \{\bar{Y}, \{Y, f\}\} = \{\mathrm{Re}Y, \{\mathrm{Re}Y, f\}\} + \{\mathrm{Im}Y, \{\mathrm{Im}Y, f\}\}$. Using (α) and (γ) one gets (ii).

[1]K. Kraus, Ann. Phys. **64**, 311 (1971).

[2]V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, J. Math. Phys. **17**, 821 (1976).

[3]G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[4]E.B. Davies, *Quantum Theory of Open System* (Academic, New York, 1976).

[5]D.E. Evans and J.T. Lewis, "Dilations of Irreversible Evolutions in Algebraic Quantum Theory," Comm. DIAS n°24, (1977).

[6]G. Sewell, Commun. Math. Phys. **55**, 53 (1977).

[7]M. Fannes and A. Verbeure, J. Math. Phys. **19**, 558 (1978).

[8]W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1971), Vol. 2.

[9]J.V. Pulè and A. Verbeure, J. Math. Phys. **20**, 733 (1979).

[10]D. Ruelle, *Statistical Mechanics* (Benjamin, Amsterdam, 1969).

[11]M. Aizenman, S. Goldstein, C. Gruber, J. Lebowitz, and Ph. Martin, Commun. Math. Phys. **53**, 209 (1977).

[12]G. Gallavotti and S. Miracle-Sole, Ann. Inst. H. Poincaré **8**, 287 (1968).

[13]A. Gerardi, N. Ianiro, C. Maffei, and C. Marchiori, J. Math. Phys. **16**, 438 (1975).

[14]R. Kubo, Phys. Soc. Japan **12**, 570 (1957).

[15]G. Gallavotti and E. Verboven, Nuovo Cimento **28**, 274 (1975).

[16]B. Demoen, P. Vanheuverzwijn, and A. Verbeure, J. Math. Phys. **19**, 2256 (1978).

[17]B. Demoen and A. Verbeure, Commun. Math. Phys. **62**, 191 (1978).

One particle reduced density matrix of impenetrable bosons in one dimension at zero temperature

H. G. Vaidya[a] and C. A. Tracy[b]
Institute for Theoretical Physics, State University of New York at Stony Brook, Stony Brook, L.I., New York 11794

(Received 11 April 1979; accepted for publication 18 May 1979)

We compute exactly the one particle reduced density matrix $\rho(r)$ of a system of impenetrable bosons in one dimension at zero temperature. We do this by relating $\rho(r)$ to a certain double scaling limit of the transverse correlation function of the one-dimensional spin 1/2 $X-Y$ model. We study the asymptotic behavior of $\rho(r)$ for large r. This expansion contains oscillatory terms which arise due to the intrinsic quantum mechanical nature of the problem. We use these results to discuss the analytic structure of the momentum density function $n(k)$.

I. INTRODUCTION

One of the model systems that has generated considerable interest is the system of bosons in one dimension interacting with the potential $c\delta(x_i - x_j)$. In particular, the limit $c \to \infty$ corresponds to a gas of impenetrable bosons. The ground state many particle wave function of this system of impenetrable bosons was first derived by Girardeau.[1] The study of the one particle reduced density matrix (which we refer to simply as the density matrix) was initiated by Schultz[2] and by Lenard.[3]

In this paper, we report an exact calculation of the reduced density matrix for this system at zero temperature. Let $\psi_{N,L}(x_1,x_2,\dots,x_N,\tau)$ be the normalized ground state wave function of N impenetrable bosons on a chain of length L at time τ. The density matrix $\rho_{N,L}(x - x',\tau)$ is defined by

$$\rho_{N,L}(x - x',\tau) = N\int_0^L dx_1 \cdots \int_0^L dx_{N-1} \times \psi_{N,L}(x_1,x_2,\cdots,x_{N-1},x,\tau) \times \psi^*_{N,L}(x_1,x_2,\cdots,x_{N-1},x',0). \tag{1.1}$$

In particular, we study $\rho_{N,L}(x - x',\tau)$ in the thermodynamic limit: $N\to\infty$, $L\to\infty$ such that $\rho = N/L$ is the constant particle density, and we write

$$\rho(x - x',\tau) = \lim_{N\to\infty, L\to\infty} \rho_{N,L}(x - x',\tau). \tag{1.2}$$

In the following sections we will derive an exact answer for $\rho(r) \equiv \rho(r,0)$. In the remainder of this section we will describe the main features of the result.

Let $x = k_F r$ (k_F = Fermi wave vector). (Note that in our paper $k_F = 1$, while in Ref. 3, $k_F = \pi$.) Then $\rho(x)$ has the following asymptotic expansion:

$$\rho(x) = \frac{\rho_\infty}{|x|^{1/2}}\left\{1 + \frac{1}{8x^2}\left(\cos 2x - \frac{1}{4}\right) + \frac{3\sin 2x}{16x^3} + \frac{3}{256x^4}\left(\frac{11}{8} - 31\cos 2x\right) + O(x^{-5})\right\} \tag{1.3}$$

where

$$\rho_\infty = \pi e^{1/2} 2^{-1/3} A^{-6} = 0.92418\ldots \tag{1.4}$$

[in Eq. (1.4), $A = 1.2824\ldots$ is Glaisher's constant].

Lenard[3] derived an expansion of $\rho(x)$ for small x [see Eq. (56) and (57) of Ref. 3]. We have used these results to extend the expansion to order x^9 [Lenard expanded $\rho(x)$ to order x^4]. The result is

$$\rho(x) = 1 - \frac{x^2}{6} + \frac{|x|^3}{9\pi} + \frac{x^4}{120} - \frac{11|x|^5}{1350\pi} - \frac{x^6}{5040} + \frac{122|x|^7}{105\pi\times 7!} + \left(\frac{1}{24\,300\pi^2} + \frac{1}{9!}\right)x^8 - \frac{253|x|^9}{98\,000\times 27^2\times\pi} + O(x^{10}). \tag{1.5}$$

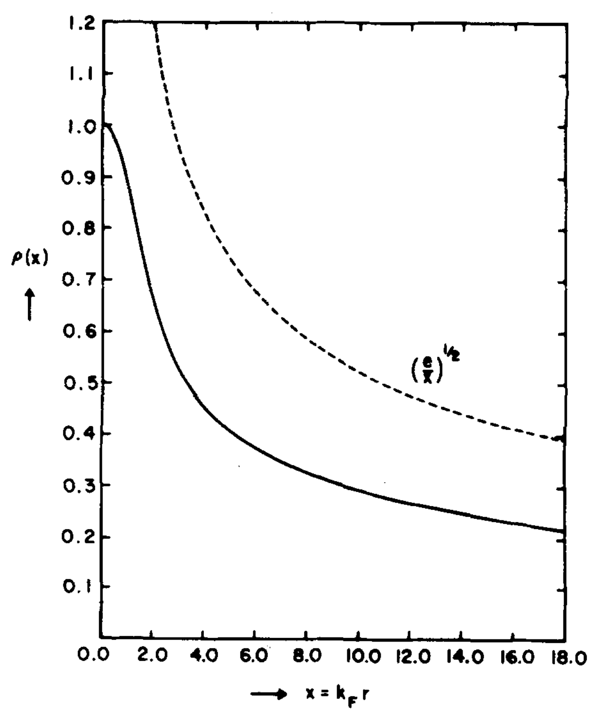


FIG. 1. $\rho(x = k_F r)$ as a function of x. The dotted line is a plot of Lenard's upper bound.

[a] Address as of 15 September, 78: Department of Physics, Columbia University, New York, NY 10027.
[b] Present address: Dept. of Mathematics, Dartmouth College, Hanover, NH 03755.

 0022-2488/79/112291-22$01.00

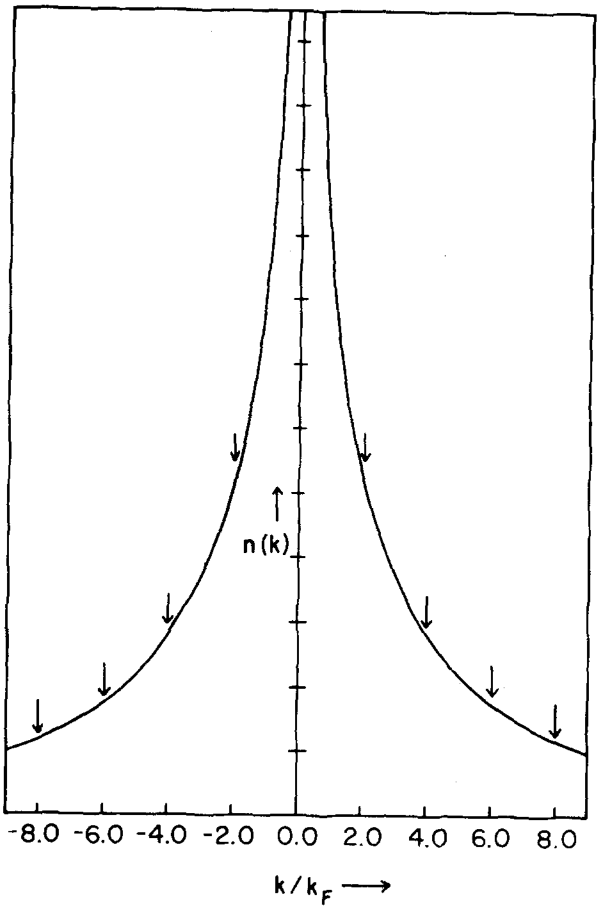


FIG. 2. Schematic plot of $n(k)$ as a function of (k/k_f), showing the $k^{-1/2}$ singularity at the origin. The arrows mark the points of nonanalyticity of $n(k)$ where a derivative of $n(k)$ diverges.

We use the expansions (1.3) and (1.5) to plot $\rho(x)$ as a function of x in Fig. 1. We also plot Lenard's upper bound $\rho(x) \leqslant (e/x)^{1/2}$.

The expansion of $\rho(x)$ for large x has the following general structure:

$$\rho(x) = \frac{\rho_\infty}{|x|^{1/2}}\left[1 + \sum_{n=1}^{\infty}\frac{c_{2n}}{x^{2n}} + \sum_{m=1}^{\infty}\frac{\cos 2mx}{x^{2m}} \times\left(\sum_{n=0}^{\infty}\frac{c'_{2n,m}}{x^{2n}}\right) + \sum_{m=1}^{\infty}\frac{\sin 2mx}{x^{2m+1}}\left(\sum_{n=0}^{\infty}\frac{c''_{2n,m}}{x^{2n}}\right)\right]. \tag{1.6}$$

This expansion enables us to study the singularity structure of the one particle momentum density function $n(k)$. $n(k)$ is defined by

$$n(k) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} dr e^{-ikr}\rho(r). \tag{1.7}$$

The $|x|^{-1/2}$ falloff of $\rho(x)$ for large x leads to a $|k|^{-1/2}$ singularity in $n(k)$ at the origin. The terms with $\sin 2mx$ and $\cos 2mx$ in Eq. (1.6) lead to additional points of nonanalyticity for $n(k)$ at $k = \pm 2mk_F (m = 1,2,...)$. At these points some higher derivative of $n(k)$ diverges. For example, at $k = \pm 2k_F$, $d^2n(k)/dk^2$ is divergent. Note that a system of free fermions has a sharp Fermi surface at zero temperature, while for the system of impenetrable bosons only the second derivative of the momentum distribution function diverges at $k = \pm 2k_F$.

In Fig. 2, we show schematically the behavior of $n(k)$ as a function of (k/k_F). The arrows mark the points of nonanalyticity of $n(k)$. Figure 3 shows the branch cut structure of $n(k)$ in the k plane. All the branch points in Fig. 3 are square root branch points.

The full answer for $\rho(r)$ is written out explicitly in Sec. 7. The behavior of $\rho(r)$ for nonzero temperatures and for the case when c is finite has been discussed in Ref. 4.

II. FORMULATION OF THE PROBLEM

A. Relation between the boson problem and the *X–Y* model

Schultz[2] approached the boson problem by replacing the continuum by a lattice of evenly spaced lattice points, with lattice spacing ϵ. He expressed the density matrix on this lattice as a determinant. Lenard[3] showed that one can

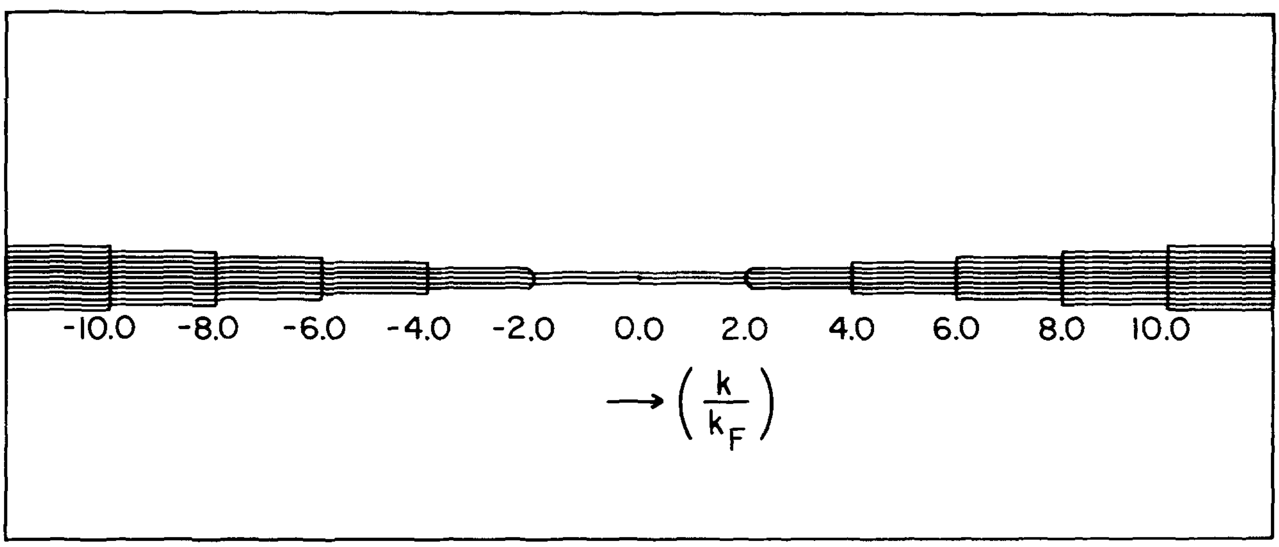


FIG. 3. Analytic structure of $n(k)$ in the k plane showing square root branch points at $k = \pm 2mk_F$, $m = 0,1,2,...$.

take the thermodynamic limit before the continuum limit to obtain $\rho(r)$. Therefore, from Eq. 80 and 87 of Ref. 3 we have

$$\rho(r) = \lim_{\substack{\epsilon\to 0, s\to\infty \\ s\epsilon = r \text{ fixed}}} (\rho\epsilon)^{-1}\rho_s\,, \tag{2.1}$$

where

$$\rho_s = \frac{1}{2}\begin{vmatrix} G_1 & G_2 & \cdots & G_s \\ G_0 & G_1 & \cdots & G_{s-1} \\ \vdots & & & \\ G_{-(s-2)} & G_{-(s-1)} & \cdots & G_1 \end{vmatrix}, \tag{2.2}$$

with

$$G_m = \begin{cases} \dfrac{2}{\pi m}\sin\pi\rho m\epsilon, & \text{for } m\neq 0, \\ -1 + 2\epsilon\rho, & \text{for } m = 0. \end{cases} \tag{2.3}$$

We follow Schultz and compute $\rho(r,\tau)$ by relating it to a certain double scaling limit of the XX correlation function of the spin $\frac{1}{2}XY$ model in one dimension. This model is defined by the Hamiltonian

$$H_N = -\sum_{i=1}^{N}\left[(1+\gamma)S_i^x S_{i+1}^x + (1-\gamma)S_i^y S_{i+1}^y + hS_i^z\right], \tag{2.4}$$

where $S_i^\alpha = \frac{1}{2}\sigma_i^\alpha$, $\alpha = x,y,z$, and σ_i^α are the usual Pauli matrices. In Eq. (2.4), $\gamma(0\leqslant\gamma\leqslant 1)$, is the anisotropy parameter and h is the magnetic field in the z direction. We impose cyclic boundary conditions $S_{N+1}\equiv S_1$. We denote the ground state transverse correlation function in the thermodynamic limit $(N\to\infty)$ by $\rho_{xx}(R,t;\gamma)$ defined by

$$\rho_{xx}(R,t;\gamma)\equiv\langle S_1^x(0)S_{R+1}^x(t)\rangle\,. \tag{2.5}$$

In Eq. (2.5) the brackets denote the gound state expectation value.

For the anisotropic X–Y model, $\rho_{xx}(R,t=0;\gamma)\equiv\rho_{xx}(R;\gamma)$ can be written as[5]

$$\rho_{xx}(R;\gamma) = \frac{1}{4}\begin{vmatrix} \overline{G}_1 & \overline{G}_2 & \cdots & \overline{G}_R \\ \overline{G}_0 & \overline{G}_1 & \cdots & \overline{G}_{R-1} \\ \vdots & & & \\ \overline{G}_{-(R-2)} & \overline{G}_{-(R-1)} & \cdots & \overline{G}_1 \end{vmatrix}, \tag{2.6}$$

where

$$\overline{G}_m = -(2\pi)^{-1}\int_{-\pi}^{\pi} d\varphi\,\frac{e^{i\varphi m}(h-\cos\varphi + i\gamma\sin\varphi)}{[(h-\cos\varphi)^2+\gamma^2\sin^2\varphi]^{1/2}}\,. \tag{2.7}$$

for $\gamma = 0$ this reduces to

$$\overline{G}_m = \begin{cases} \dfrac{2}{\pi m}\sin m\varphi_0, & \text{for } m\neq 0, \\ -1+\dfrac{2}{\pi}\varphi_0, & \text{for } m = 0, \end{cases} \tag{2.8}$$

where

$$\cos\phi_0 = h. \tag{2.9}$$

Comparing Eqs. (2.3) and (2.8), we see that the determinants in Eqs. (2.2) and (2.6) are identical when

$$\rho = \frac{1}{\pi} \text{ and } \epsilon = \arccos(h). \tag{2.10}$$

Under this identification

$$\rho_s = 2\rho_{xx}(s;0). \tag{2.11}$$

As $\epsilon\to 0$, $h\to 1^-$ and we can write $\epsilon = (1-h^2)^{-1/2}$. The equivalence can be established in a similar manner for the time dependent case.

To use Eq. (2.6) we need to evaluate determinants of very large dimension. This in general is very difficult to do. However, when $\gamma\neq 0$, the determinant (2.6) has been studied in great detail in Ref. 6. There it was shown (for the general case $t\neq 0$) that for $h<1$

$$\rho_{xx}(R,t;\gamma) = \rho_{xx}(\infty)\exp\Big[-\sum_{n=1}^{\infty}F^{(2n)}(R,t;\gamma)\Big], \tag{2.12}$$

where

$$\rho_{xx}(\infty) = [2(1+\gamma)]^{-1}[\gamma^2(1-h^2)]^{1/4} \tag{2.13}$$

and

$$F^{(2n)}(R,t;\gamma) = (2n)^{-1}2^{-2n}(2\pi)^{-2n}\int_{-\pi}^{\pi}d\varphi_1\cdots\int_{-\pi}^{\pi}d\varphi_{2n} \times\prod_{j=1}^{2n}\left[\frac{e^{-iR\varphi_j - itA_j}(A_j - A_{j+1})}{A_j\sin(1/2)(\varphi_j+\varphi_{j+1})}\right], \tag{2.14}$$

with $\phi_{2n+1}\equiv\phi_1$; $\mathrm{Im}\phi_j<0, j = 1,2,\cdots,2n$; and

$$A_j\equiv A(\varphi_j) = [(\cos\varphi_j - h)^2+\gamma^2\sin^2\varphi_j]^{1/2}. \tag{2.15}$$

All reference to the original determinant has been eliminated in this formula.

We define the double scaling limit of $\rho_{xx}(R,t;\gamma)$ (denoted by $\lim_C$ following Ref. 7) by $h\to 1^-$, $\gamma\to 0$, $R\to\infty$, $t\to\infty$ such that

$$r = (1-h^2)^{1/2}R, \tag{2.16a}$$

$$\tau = \tfrac{1}{2}(1-h^2)t, \tag{2.16b}$$

and

$$g = \gamma(1-h^2)^{-1/2} \tag{2.16c}$$

are held fixed.

We now use Eqs. (2.1) and (2.11) to recover $\rho(r,\tau)$ as

$$\rho(r,\tau) = \lim_C\lim_{g\to 0} 2\pi(1-h^2)^{-1/2}\rho_{xx}(R,t;\gamma). \tag{2.17}$$

The calculation of this paper is based on the assumption that the two limits in Eq. (2.17) can be interchanged. We thus have

$$\rho(r,\tau) = \lim_{g\to 0}\lim_C 2\pi(1-h^2)^{-1/2}\rho_{xx}(R,t;\gamma). \tag{2.18}$$

B. Double scaling limit of $\rho_{xx}(R,t;\gamma)$

In $\lim_C$ we may expand $A(\phi_j)$ around $\phi_j = 0$ and rescale ϕ_j by

$$\varphi_j = (1-h^2)^{1/2}k_j \tag{2.19}$$

to obtain

$$A(\varphi_j)\simeq\tfrac{1}{2}(1-h^2)\epsilon(k_j), \tag{2.20}$$

where

$$\epsilon(k_j)\equiv\epsilon_j = [(k_j^2+\mu^2)(k_j^2+\mu^{*2})]^{1/2}, \tag{2.21}$$

with

$$\mu = g + i(1-g^2)^{1/2} \tag{2.22}$$

and μ^* is the complex conjugate of μ, and g is given by Eq. (2.16c). Similarly,

$$\sin\tfrac{1}{2}(\varphi_j+\varphi_{j+1})\simeq\tfrac{1}{2}(1-h^2)^{1/2}(k_j+k_{j+1}). \tag{2.23}$$

We scale the integrals in Eq. (2.14) using the formulas (2.19) – (2.23) and denote the scaled functions by $h^{(2n)}(r,\tau,g)$. Substituting in Eq. (2.18), we get

$$\rho(r,\tau) = \lim_{g\to 0} \pi g^{1/2}\exp[-H(\pi^2;r,\tau,g)], \tag{2.24}$$

where

$$H(z;r,\tau,g) = \sum_{n=1}^{\infty} n^{-1} z^{2n} h^{(2n)}(r,\tau,g) \tag{2.25}$$

and

$$h^{(2n)}(r,\tau,g) = \frac{1}{2}\int_{-\infty}^{\infty} dk_1 \cdots \int_{-\infty}^{\infty} dk_{2n} \times \prod_{j=1}^{2n}\left[\frac{e^{-irk_j - i\tau\epsilon_j}(\epsilon_j-\epsilon_{j+1})}{\epsilon_j(k_j+k_{j+1})}\right],$$

$$k_{2n+1}\equiv k_1 . \tag{2.26}$$

We will now restrict our attention to the case $\tau = 0$ and denote the above functions by $H(z,r,g)$ and $h^{(2n)}(r,g)$, respectively.

III. ANALYSIS OF $h^{(2n)}(r,g)$

The integrands in Eq. (2.26) have branch points at $k_j = \pm i\mu, \pm i\mu^*$. Their analytic structure is shown in Fig. 4. In the limit $g\to 0$ the branch points pinch the real axis in pairs leading to logarithmic divergences in g. The general structure can be seen to be the following for $g\ll 1$:

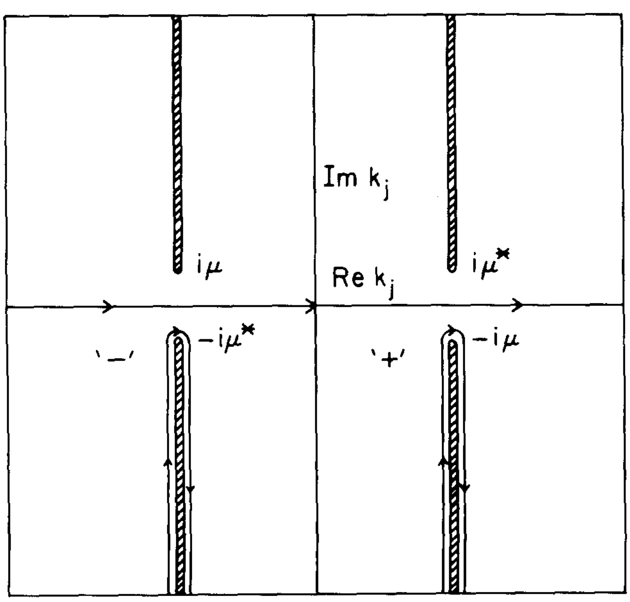


FIG. 4. Analytic structure of the integrand in Eq. (2.26) in the k_j plane. We show the branch cuts + and − around which the contour of integration over k_j is bent.

$$h^{(2n)}(r,g) = C_{2n}^{(2n)}\ln^{2n}(gr) + \sum_{m=1}^{2n} C_{2n-m}^{(2n)}(r)\ln^{2n-m}(gr) + o(1), \tag{3.1}$$

where $C_{2n}^{(2n)}$ are constants and $C_m^{(2n)}(r)$ $(m\neq 2n)$ are functions of r independent g. It is clear that we need to sum these logarithmic divergences systematically to take the limit $g\to 0$.

We can show that the constant $C_{2n}^{(2n)}$ (with the factor n^{-1} of 2.25 included) in Eq. (3.1) is twice the constant C_{2n} appearing in the small variable expansion of the functions $f_{2n}(t)$ of the two dimensional Ising model [see Eq. (3.146) of Ref. 8]. In the Ising model the coefficients multiplying the divergent terms satisfy certain relations so that on summing up the divergences to all orders one gets a simple 'ln$t$' divergence. Something similar has to happen in Eq. (3.1) if we are to get a finite result in the $g\to 0$ limit. The relation between the constants multiplying the leading divergent terms suggests the possibility of expressing the divergent parts of $h^{(2n)}(r,g)$ in terms of the Ising model functions.

The functions $f_{2n}(t)$ and the related functions $g_{2n+1}(t)$ are given by[8]

$$f_{2n}(t) = (-1)^n n^{-1}\int_1^{\infty} dy_1 \cdots \int_1^{\infty} dy_n \times \prod_{j=1}^{2n}\frac{e^{-ty_j}}{(y_j^2-1)^{1/2}(y_j+y_{j+1})}\prod_{j=1}^{n}(y_{2j}^2-1), \tag{3.2}$$

where $y_{2n+1}\equiv y_1$ and

$$g_{2n+1}(t) = (-1)^n\int_1^{\infty} dy_1\cdots\int_1^{\infty} dy_{2n+1} \times \prod_{j=1}^{2n+1}\frac{e^{-ty_j}}{(y_j^2-1)^{1/2}}\prod_{j=1}^{2n}(y_j+y_{j+1})^{-1} \times \prod_{j=1}^{n}(y_{2j}^2-1). \tag{3.3}$$

The result which we derive in Sec. 4–6 is, for $g\ll 1$,

$$H(z;r,g) = 2F(z;gr) + \bar{H}[z;r,G(z;gr)] + o(1), \tag{3.4}$$

where

$$F(z;gr) = \sum_{n=1}^{\infty} z^{2n} f_{2n}(gr) \tag{3.5a}$$

$$G(z;gr) = \sum_{n=0}^{\infty} z^{2n+1} g_{2n+1}(gr) \tag{3.5b}$$

and $\bar{H}[z;r,G(z;gr)]$ is given by Eq. (6.43). We assume that the error estimate remains $o(1)$ after carrying out the sum over n in Eq. (2.25).

Note that in the right hand side of Eq. (3.4), the g dependence is only through the functions $F(z;gr)$ and $G(z;gr)$. These functions can be expressed in terms of Painlevé functions of the third kind[9] and their behavior as $g\to 0$ has been studied in detail.[9] In this limit $F(\pi^{-1};gr)$ has the correct divergence to cancel out the factor $g^{1/2}$ in Eq. (2.24) and $\lim_{g\to 0}G(\pi^{-1};gr) = 1$.

Our strategy for deriving Eq. (3.4) is the following: We bend the contours of integration in Eq. (2.26) into the lower half plane as illustrated in Fig. 4. Each integral now separates into a sum of two parts, one each around the ' + ' and ' − ' branch cuts. Now, $h^{(2n)}(r,g)$ is a sum of 2^{2n} integrals.

We then separate each integral such that the divergent parts are expressed in terms of $f_{2n}(gr)$ and $g_{2n+1}(gr)$ with coefficients that are functions of r alone. Then summing up $h^{(2n)}(r,g)$ to all orders leads to Eq. (3.4).

In the next section, we will systematize the evaluation of the 2^{2n} integrals in $h^{(2n)}(r,g)$, using a transfer matrix formulation. In Appendix A we give a detailed evaluation of $h^{(2)}(r,g)$ as $g\to 0$. A study of Appendix A is instructive in understanding the basic strategy involved in the solution of the problem. The transfer matrix formulation enables us to analyze the 2^{2n} integrals in $h^{(2n)}(r,g)$ in a systematic manner.

IV. TRANSFER MATRIX FORMULATION

As illustrated in Fig. 4 and Appendix A, each integral separates into a sum of two pieces; one coming from the $+$ branch cut and the other coming from the $-$ branch cut. We denote this by using a state vector $\sigma_j \equiv \pm$ for each "y_j" variable of integration. [$\sigma_j \equiv +(-)$ implies that the y_j integration is carried out around the $+(-)$ branch cut.] All the 2^{2n} integrals in $h^{(2n)}(r,g)$ can be written out systematically as follows:

$$h^{(2n)}(r,g) = \sum_P h_P^{(2n)}(r,g), \tag{4.1}$$

where P is a permulation of (l) $+$ signs and $(2n-l)$ $-$ signs $l=0,1,2,\cdots,2n$, representing the branch cut around which each y_j integral $(j=1,2,\cdots,2n)$ is evaluated.

We introduce the symbol $\doteq$ with the following meaning: The two sides of $\doteq$ are equal only after carrying out integrations satisfying the conditions (i) each y_j variable is integrated on $\int_g^\infty dy_j e^{-ry_j}$ and (ii) the integrand has an overall factor

$$\prod_{j\,\text{odd}} (y_j^2-g^2)^{1/2} \prod_{j\,\text{even}} (y_j^2-g^2)^{-1/2},$$

for an integral over $y_1, y_2,\ldots$.

We introduce the notation $\int dy$ with the following meaning:

$$u(r,g) = \int dy f(y;r) \tag{4.2}$$

means that the equality in Eq. (4.2) holds on integrating over all the y_j variables in $f(y,r)(y\equiv y_1,y_2,\cdots,y_m)$ with the conditions (i) and (ii) above.

On changing the integration variables as in Appendix A we can write

$$h^{(2n)}(r,g) = \int dy \sum_{\sigma_j=\pm} \langle\sigma_1|M_0(1,2)|\sigma_2\rangle \times\langle\sigma_2|M_e(2,3)|\sigma_3\rangle\cdots\langle\sigma_{2n}|M_e(2n,1)|\sigma_1\rangle \tag{4.3}$$

$$= \int dy\,\mathrm{Tr}\Big[\prod_{j=1}^{n} M_0(2j-1,2j)M_e(2j,2j+1)\Big], \tag{4.4}$$

where $2n+1\equiv 1$ and "Tr" denotes the trace of the 2×2 matrix on the right. Here,

$$M_0(2j-1,2j) = \begin{array}{c|cc} & + & - \\ \hline + & \dfrac{e^{-ir}(y_{2j-1}+2i)}{y_{2j-1}+y_{2j}+2i} & \dfrac{y_{2j-1}+2i}{y_{2j-1}+y_{2j}} \\ - & \dfrac{y_{2j-1}-2i}{y_{2j-1}+y_{2j}} & \dfrac{e^{+ir}(y_{2j-1}-2i)}{y_{2j-1}+y_{2j}-2i} \end{array} \tag{4.5}$$

and

$$M_e(2j,2j+1) = \begin{array}{c|cc} & + & - \\ \hline + & \dfrac{e^{-ir}}{(y_{2j}+2i)(y_{2j}+y_{2j+1}+2i)} & \dfrac{1}{(y_{2j}+2i)(y_{2j}+y_{2j+1})} \\ - & \dfrac{1}{(y_{2j}-2i)(y_{2j}+y_{2j+1})} & \dfrac{e^{+ir}}{(y_{2j}-2i)(y_{2j}+y_{2j+1}-2i)} \end{array}. \tag{4.6}$$

It is clear that in the odd variable integrals we can set $g=0$. The even variable integrals diverge in the limit $g\to 0$. The aim is to get enough y_{2j}'s in the numerator to make these integrals converge and at the same time factor out integrals identical to $f_{2n}(gr)$ or $g_{2n+1}(gr)$ (after rescaling as in Appendix A).

The matrices in Eqs. (4.5) and (4.6) have the structure

$$\begin{bmatrix} (11) & (12) \\ (12)^* & (11)^* \end{bmatrix}$$

and will be denoted by $[(11),(12)]$ for the sake of compactness. In this notation

$$M_0(2j-1,2j) = \left[\frac{e^{-ir}(y_{2j-1}+2i)}{y_{2j-1}+y_{2j}+2i},\ \frac{y_{2j-1}+2i}{y_{2j-1}+y_{2j}}\right]. \tag{4.5'}$$

We now use the identity

$$(y_{2j-1}+y_{2j}\pm 2i)^{-1} = (y_{2j-1}\pm 2i)^{-1}[1-y_{2j}(y_{2j-1}+y_{2j}\pm 2i)^{-1}] \tag{4.7}$$

to write

$$M_0(2j-1,2j) = M_0^{(1)}(2j-1,2j) + M_0^{(2)}(2j-1,2j) \tag{4.8a}$$

and

$$M_e(2j,2j+1) = M_e^{(1)}(2j,2j+1) + M_e^{(2)}(2j,2j+1), \tag{4.8b}$$

where

$$M_0^{(1)}(2j-1,2j) = [e^{-ir}, (y_{2j-1}+2i)(y_{2j-1}+y_{2j})^{-1}], \tag{4.9a}$$

$$M_0^{(2)}(2j-1,2j) = -y_{2j}\,[e^{-ir}(y_{2j-1}+y_{2j}+2i)^{-1},0]\,, \tag{4.9b}$$

$$M_e^{(1)}(2j,2j+1) = [e^{-ir}[(y_{2j}+2i)(y_{2j+1}+2i)]^{-1}, [(y_{2j}+2i)(y_{2j}+y_{2j+1})]^{-1}], \tag{4.10a}$$

and

$$M_e^{(2)}(2j,2j+1) = -y_{2j}\,[e^{-ir}[(y_{2j}+2i)(y_{2j+1}+2i) \times(y_{2j}+y_{2j+1}+2i)]^{-1},0]. \tag{4.10b}$$

Hence,

$$H(z;r,g) = \sum_{n=1}^{\infty} n^{-1}z^{2n}h^{(2n)}(r,g) \tag{4.11a}$$

$$= \sum_{n=1}^{\infty} n^{-1}z^{2n}\sum_{l=0}^{2n}\int dy\,\mathrm{Tr}G_l(1\to 2n)\,, \tag{4.11b}$$

where $G_l(1\to 2n)\equiv G_l(y_1,y_2,\cdots,y_{2n})$ enumerates all possible products of $2n$ matrices with l $M_{(y)}^{(2)}$'s and $(2n-l)$ $M_{(y)}^{(1)}$'s. In particular,

$$G_0(1\to 2n) \doteq \prod_{j=1}^{2n} M_0^{(1)}(2j-1,2j)M_e^{(1)}(2j,2j+1) \tag{4.12a}$$

and

$$G_{2n}(1\to 2n) \doteq \prod_{j=1}^{2n} M_0^{(2)}(2j-1,2j)M_e^{(2)}(2j,2j+1). \tag{4.12b}$$

In Eqs. (4.12), $y_{2n+1}\equiv y_1$.

Separating out $G_0(2n)$ and changing the order of summation in the remaining sum, we have

$$H(z;r,g) = \int dy\Big[\sum_{n=1}^{\infty} n^{-1}z^{2n}\,\mathrm{Tr}G_0(1\to 2n) + \sum_{l=1}^{\infty}\sum_{n=[(l+1)/2]}^{\infty} n^{-1}z^{2n}\,\mathrm{Tr}G_l(1\to 2n)\Big]\,, \tag{4.13}$$

where $[x]$ is the largest integer $\leqslant x$.

Now let us look at the structure of $G_l(2n)$, $l\neq 0$. Each term has $l\,M_{(y)}^{(2)}$'s seperated by a string of $M_{(y)}^{(1)}$'s. There are four such distinct strings given by

$$F_{2k}^{(1)}(2\to 2k+1) \doteq\Big[\prod_{j=1}^{k-1} M_e^{(1)}(2j,2j+1)M_0^{(1)} \times(2j+1,2j+2)\Big]M_e^{(1)}(2k,2k+1)\,, \tag{4.14}$$

$$F_{2k+1}^{(2)}(2\to 2k+2) \doteq \prod_{j=1}^{k} M_e^{(1)}(2j,2j+1)M_0^{(1)}(2j+1,2j+2), \tag{4.15}$$

$$F_{2k+1}^{(3)}(1\to 2k+1) \doteq \prod_{j=1}^{k} M_0^{(1)}(2j-1,2j)M_e^{(1)}(2j,2j+1)\,, \tag{4.16}$$

and

$$F_{2k}^{(4)}(1\to 2k) \doteq\Big[\prod_{j=1}^{k-1} M_0^{(1)}(2j-1,2j)M_e^{(1)}(2j,2j+1)\Big] \times M_0^{(1)}(2k-1,2k)\,. \tag{4.17}$$

In Eqs. (4.14) and (4.17) the product is replaced by 1 for $k=1$. In Eq. (4.14) to (4.17), $k=1,2,3,\ldots$ and the "r" and "g" dependence is understood. In the above equations the starting variables is either y_1 or y_2, but it is understood that in the actual product the starting and ending variables are chosen so as to match the variables of $M^{(2)}$'s on either side. For examples, a typical string involving $F_{2k}^{(1)}(y)$ might be

$$M_0^{(2)}(7,8)F_{2k}^{(1)}(8\to 2k+7)M_0^{(2)}(2k+7,2k+8).$$

It is also clear that the value of k in $F_k^{(j)}(y)$ ($j=1,2,3$, or 4) is chosen to fit the missing variables between two successive $M_{(y)}^{(2)}$'s. The case where the two $M_{(y)}^{(2)}$'s are adjacent [either $M_0^{(2)}(2j-1,2j)M_e^{(2)}(2j,2j+1)$ or $M_e^{(2)}(2j,2j+1) \times M_0^{(2)}(2j+1,2j+2)$], is accounted for by choosing

$$F_1^{(2)}=1, \qquad F_1^{(3)}=1. \tag{4.18}$$

We define the generating functions

$$F^{(j)}(y;z)\equiv F^{(j)}(y,z;r,g)\equiv\sum_{n=1}^{\infty} z^{2n}F_{2n}^{(j)}(y), \quad \text{for } j=1,4, \tag{4.19a}$$

and

$$F^{(j)}(y;z)\equiv F^{(j)}(y,z;r,g) \equiv\sum_{n=0}^{\infty} z^{2n+1}F_{2n+1}^{(j)}(y), \quad \text{for } j=2,3. \tag{4.19b}$$

In Eq. (4.19), y symbolically represents all the y_j variables appearing in $F^{(j)}(y,z;r,g)$, $F_{2n}^{(j)}(y)$, and $F_{2n+1}^{(j)}(y)$. The choice of the y_j's is determined by factors on either side, as illustrated in the discussion following Eq. (4.17).

Before introducing a transfer matrix notation for the second term in Eq. (4.13) let us write down $G_1(2n)$ and $G_2(2n)$ explicitly:

$$G_1(1\to 2n)\doteq\sum_{j=1}^{n}[M_0^{(2)}(2j-1,2j)F_{2n}^{(1)}(2j\to 2n,1\to 2j-1) + M_e^{(2)}(2j,2j+1)F_{2n}^{(4)}(2j+1\to 2n,1\to 2j)] \tag{4.20a}$$

$$\doteq n\big[M_0^{(2)}(1,2)F_{2n}^{(1)}(2\to 2n,1) + M_e^{(2)}(2n,1)F_{2n}^{(4)}(1\to 2n)\big] \tag{4.20b}$$

since in Eq. (4.19a) all the n terms are identical on integration (they are equivalent to a cyclic relabeling of the variables which does not change the value of the integral). Similarly,

$$\begin{aligned}G_2(1\to 2n)\doteq \frac{n}{2}\Big\{&\sum_{k=1}^{n-1}[M_0^{(2)}(1,2)F_{2k}^{(1)}(2\to 2k+1)M_0^{(2)}(2k+1,2k+2)F_{2n-2k}^{(1)}(2k+2\to 2n,1)\\ &+M_e^{(2)}(2n,1)F_{2k}^{(4)}(1\to 2k)M_e^{(2)}(2k,2k+1)F_{2n-2k}^{(4)}(2k+1\to 2n)]\\ &+\sum_{k=1}^{n}[M_0^{(2)}(1,2)F_{2k-1}^{(2)}(2\to 2k)M_e^{(2)}(2k,2k+1)F_{2n-2k+1}^{(3)}(2k+1\to 2n,1)\\ &+M_e^{(2)}(2n,1)F_{2k-1}^{(3)}(1\to 2k-1)M_0^{(2)}(2k-1,2k)F_{2n-2k+1}^{(2)}(2k\to 2n)]\Big\}.\end{aligned} \tag{4.21}$$

We now use a transfer matrix notation to enumerate the matrix products in $G_l(2n)$. $G_l(2n)$ has $l\,M^{(2)}$'s each of which is either even ($M_e^{(2)}$) or odd ($M_0^{(2)}$). We construct the compound transfer matrix

$$\mathcal{N}(y;z)\equiv\mathcal{N}(y,z;r,g)$$
$$\begin{array}{cc} & \begin{array}{cc} \quad o & \qquad\qquad e \end{array} \\ \begin{array}{c} o \\ e \end{array} & \begin{bmatrix} M_0^{(2)}(y)F^{(1)}(y;z) & M_0^{(2)}(y)F^{(2)}(y;z) \\ M_e^{(2)}(y)F^{(3)}(y;z) & M_e^{(2)}(y)F^{(4)}(y;z) \end{bmatrix}. \end{array} \tag{4.22}$$

Then symbolically

$$G_l(1\to 2n) \doteq \frac{n}{l}\,[\mathrm{Tr}^{(2)}\mathcal{N}(y;z)^l\,]_{2n}\,. \tag{4.23}$$

The subscript "$2n$" denotes the power of z^{2n} in the expansion in power of z in the right hand side of Eq. (4.23). $\mathrm{Tr}^{(2)}$ denotes the trace over $\mathcal{N}^l(y;z)$ which is treated as a 2×2 matrix as in Eq. (4.22). For example,

$$\mathrm{Tr}^{(2)}\mathcal{N}(y;z) = M_0^{(2)}(y)F^{(1)}(y;z) + M_e^{(2)}(y)F^{(4)}(y;z), \tag{4.24}$$

$$\begin{aligned} &= M_0^{(2)}(y)F^{(1)}(y;z)M_0^{(2)}(y)F^{(1)}(y;z) \\ &\quad + M_e^{(2)}(y)F^{(4)}(y;z)M_e^{(2)}(y)F^{(4)}(y;z) \\ &\quad + M_0^{(2)}(y)F^{(2)}(y;z)M_e^{(2)}(y)F^{(3)}(y;z) \\ &\quad + M_e^{(2)}(y)F^{(3)}(y;z)M_0^{(2)}(y)F^{(2)}(y;z). \end{aligned} \tag{4.25}$$

It can be readily seen that, together with Eqs. (4.19) and (4.23), Eq. (4.24) leads to (4.20b) and Eq. (4.25) leads to (4.21) on introducing the appropriate integration variables. The notation of Eqs. (4.22) and (4.23) is introduced for book-keeping purposes and it is to be intepreted only in terms of sums of the form (4.21). The reader is advised to start with Eq. (4.4), substitute Eq. (4.8), and write out some terms explicitly to see that Eqs. (4.12a), (4.13), and (4.23) generate all the contributions to $h^{(2n)}(r,g)$.

Thus, we have

$$\begin{aligned} H(z;r,g) = \int dy\Big[&\sum_{n=1}^{\infty} n^{-1} z^{2n}\,\mathrm{Tr}G_0(1\to 2n) \\ &+ \sum_{l=1}^{\infty}\;\sum_{n=[(l+1)/2]}^{\infty} l^{-1} z^{2n}\,\mathrm{Tr}[\mathcal{N}(y;z)^l\,]_{2n}\Big]. \end{aligned} \tag{4.26}$$

We can do the sum over n in the second term to write

$$\begin{aligned} H(z;r,g) = \int dy\Big[&\sum_{n=1}^{\infty} n^{-1} z^{2n}\,\mathrm{Tr}G_0(1\to 2n) \\ &+ \sum_{l=}^{\infty} l^{-1}\,\mathrm{Tr}\mathcal{N}(y;z)^l\Big]. \end{aligned} \tag{4.27}$$

We write

$$\lim_{g\to 0}\mathcal{N}(y,\pi r^{-1};r,g)\equiv\mathcal{N}(y;r). \tag{4.28}$$

It should be kept in mind that $\mathrm{Tr}\mathcal{N}(y;z)^l$ has an overall cyclic structure, namely, the last variable in the trace = the first variable [see for example Eqs. (4.20) and (4.21)].

V. ANALYSIS OF $\mathcal{N}^l(y;z)$

A. Factorization of $F^{(j)}(y;z)$, $j = 1,2,3,4$

We first introduce the notion of connectedness. Two matrices are connected if they have at least one variable in common; otherwise they are disconnected. Thus, if two factors are disconnected in the matrix form, the corresponding integral is a product of two integrals. We denote the fact that the two matrices are disconnected by introducing a bar (|) between them. We will aim to reduce $F^{(1)}(y;z)$ to $F^{(4)}(y;z)$ in a form such that the divergent part (as $g\to 0$) is disconnected from the matrices on both sides. We can then integrate over the variables of this factor independently of the other variables in the matrix product.

We now use the identity

$$M_0^{(1)}(2j-1,2j) = M_0^{(1)'}(2j-1) + y_{2j}M_0^{(1)''}(2j-1,2j), \tag{5.1}$$

where

$$M_0^{(1)'}(2j-1) = [e^{-ir},y_{2j-1}^{-1}(y_{2j-1}+2i)] \tag{5.2a}$$

and

$$\begin{aligned} &M_0^{(1)''}(2j-1,2j) \\ &\quad = -y_{2j-1}^{-1}\,[0,(y_{2j-1}+2i)(y_{2j-1}+y_{2j})^{-1}], \end{aligned} \tag{5.2b}$$

to write

$$\begin{aligned} &F_{2k}^{(1)}(2\to 2k+1) \\ &\doteq M_e^{(1)}(2,3)M_0^{(1)'}(3)\Big[\prod_{j=2}^{k-1} M_e^{(1)}(2j,2j+1) \\ &\quad\times M_0^{(1)}(2j+1,2j+2)\Big]M_e^{(1)}(2k,2k+1) \\ &\quad + y_4M_e^{(1)}(2,3)M_0^{(1)''}(3,4) \\ &\quad\times\Big[\prod_{j=2}^{k-1} M_e^{(1)}(2j,2j+1)M_0^{(1)}(2j+1,2j+2)\Big] \\ &\quad\times M_e^{(1)}(2k,2k+1). \end{aligned} \tag{5.3}$$

The product is replaced by 1 for $k = 1,2$ and $k = 1,2,\ldots$.

Carrying out the same procedure repeatedly and defining

$$\begin{aligned} &E_{2n}^{(1)}(2\to 2n+1) \\ &\quad = \Big[\prod_{j=1}^{n-1} y_{2j+2}M_e^{(1)}(2j,2j+1)M_0^{(1)''}(2j+1,2j+2)\Big] \\ &\qquad\times M_e^{(1)}(2n,2n+1), \\ &\qquad\qquad (\text{product}\equiv 1 \text{ for } n=1),\quad n = 1,2,3,\ldots, \end{aligned} \tag{5.4}$$

we have

$$F_2^{(1)}(2,3) \doteq E_2^{(1)}(2,3) \tag{5.5a}$$

and

$$\begin{aligned} &F_{2k}^{(1)}(2\to 2k+1) \\ &\quad\doteq \sum_{l=1}^{k-1} E_{2l}^{(1)}(2\to 2l+1)M_0^{(1)'}(2l+1) \\ &\qquad\times |F_{2k-2l}^{(1)}(2l+2\to 2k+1) + E_{2k}^{(1)}(2\to 2k+1), \\ &\qquad\qquad \text{for } k = 2,3,\ldots . \end{aligned} \tag{5.5b}$$

In the integrals involving the y_j variables of $E_{2n}^{(1)}(y)$, we can set $g = 0$. We adopt the notation that we can set $g = 0$ in all the integrals with the $E^{(j)}(y)'$ as integrands.

Defining the generating function

$$E^{(1)}(y;z)\equiv E^{(1)}(y,z;r)\equiv\sum_{n=1}^{\infty} z^{2n}E_{2n}^{(1)}(y), \tag{5.6}$$

we can write the recursion relations (5.5) in a compact form

$$F^{(1)}(y;z) \doteq E^{(1)}(y;z) M_0^{(1)'}(y) \mid F^{(1)}(y;z) + E^{(1)}(y;z). \tag{5.7}$$

Matching the power of z on the two sides of Eq. (5.7) reproduces Eq. (5.5) for $k = 1,2,\ldots$, after introducing the appropriate integration variables. Note that we cannot algebraically solve Eq. (5.7) for $F^{(1)}(y;z)$ since it is an equality only under the integral sign.

Next we use the identity

$$M_e^{(1)'}(2j,2j+1) = [(y_{2j}+2i)^{-1},0] M_e^{(1)'}(2j+1) + y_{2j} M_e^{(1)''}(2j,2j+1), \tag{5.8}$$

where

$$M_e^{(1)'}(2j+1) = [e^{-ir}(y_{2j+1}+2i)^{-1}, y_{2j+1}^{-1}] \tag{5.9a}$$

and

$$M_e^{(1)''}(2j,2j+1) = -y_{2j+1}^{-1}(y_{2j}+y_{2j+1})^{-1}[0,(y_{2j}+2i)^{-1}]. \tag{5.9b}$$

Using this identity repeatedly, we can write

$$|F_{2k}^{(1)}(2\to 2k+1) \doteq \sum_{l=0}^{k-1} |F_{2k-2l-1}^{(5)}(2\to 2k-2l)| \times E_{2l+1}^{(2)}(2k-2l+1\to 2k+1) + |E_{2k}^{(3)}(2\to 2k+1), \quad k=1,2,\ldots. \tag{5.10}$$

The bars represent the fact that these factors are not connected to the matrices on the side of the bar when introduced in the integral. In Eq. (5.10)

$$F_{2k-1}^{(5)}(2\to 2k) = \left[\prod_{j=1}^{k-1} M_e^{(1)}(2j,2j+1) M_0^{(1)}(2j+1,2j+2)\right] \times [(y_{2k}+2i)^{-1},0], \quad (\text{product}\equiv 1 \text{ for } k=1), \quad k=1,2,\ldots, \tag{5.11}$$

$$E_{2l+1}^{(2)}(1\to 2l+1) = M_e^{(1)'}(1)\left[\prod_{j=1}^{l} y_{2j} M_0^{(1)}(2j-1,2j) M_e^{(1)''}(2j,2j+1)\right], \quad (\text{product}\equiv 1 \text{ for } l=0), \quad l=0,1,2,\ldots, \tag{5.12}$$

and

$$E_{2k}^{(3)}(2\to 2k+1) = y_2 M_e^{(1)''}(2,3)\left[\prod_{j=2}^{k} y_{2j} M_0^{(1)}(2j-1,2j) M_e^{(1)''}(2j,2j+1)\right], \quad (\text{product}\equiv 1 \text{ for } k=1), \quad k=1,2,\ldots. \tag{5.13}$$

Defining the generating functions

$$F^{(5)}(y;z) \equiv F^{(5)}(y,z;r,g) \equiv \sum_{n=0}^{\infty} z^{2n+1} F_{2n+1}^{(5)}(y), \tag{5.14a}$$

$$E^{(2)}(y;z) \equiv E^{(2)}(y,z;r) \equiv \sum_{n=0}^{\infty} z^{2n+1} E_{2n+1}^{(2)}(y), \tag{5.14b}$$

and

$$E^{(3)}(y;z) \equiv E^{(3)}(y,z;r) \equiv \sum_{n=1}^{\infty} z^{2n} E_{2n}^{(3)}(y), \tag{5.14c}$$

and using the recursion relations (5.10) and Eq. (5.7), we can write

$$F^{(1)}(y;z) \doteq E^{(1)}(y;z) M_0^{(1)'}(y) | F^{(5)}(y;z) | E^{(2)}(y;z) + E^{(1)}(y;z) M_0^{(1)'}(y) | E^{(3)}(y;z) + E^{(1)}(y;z). \tag{5.15}$$

A similar analysis of $F^{(2)}(y;z)$, $F^{(3)}(y;z)$, and $F^{(4)}(y;z)$ results in

$$F^{(2)}(y;z) \doteq zE^{(1)}(y;z) M_0^{(1)'}(y) | F^{(5)}(y;z) | E^{(2)}(y;z) M_0^{(1)}(y) + zE^{(1)}(y;z) M_0^{(1)'}(y) | E^{(3)}(y;z) M_0^{(1)}(y) + E^{(4)}(y;z), \tag{5.16}$$

$$F^{(3)}(y;z) \doteq E^{(5)}(y;z) | F^{(5)}(y;z) | E^{(2)}(y;z) + E^{(5)}(y;z) | \times E^{(3)}(y;z) + E^{(6)}(y;z), \tag{5.17}$$

and

$$F^{(4)}(y;z) \doteq zE^{(5)}(y;z) | F^{(5)}(y;z) | E^{(2)}(y;z) M_0^{(1)}(y) + zE^{(5)}(y;z) | E^{(3)}(y;z) M_0^{(1)}(y) + zE^{(6)}(y;z) M_0^{(1)}(y), \tag{5.18}$$

respectively.

In Eqs. (5.16)–(5.18)

$$E^{(j)}(y;z) \equiv E^{(j)}(y,z;r) \equiv \sum_{n=0}^{\infty} z^{2n+1} E_{2n+1}^{(j)}(y), \quad j=4,5,6, \tag{5.19}$$

with

$$E_1^{(4)} = 1,$$

$$E_{2n+1}^{(4)}(2\to 2n+2) = E_{2n}^{(1)}(2\to 2n+1) M_0^{(1)}(2n+1,2n+2), \quad \text{for } n\geqslant 1, \tag{5.20}$$

$$E_1^{(5)} = M_0^{(1)'}(1),$$

$$E_{2n+1}^{(5)}(1\to 2n+1) = y_2 M_0^{(1)''}(1,2) E^{(1)}(2\to 2n+1) M_0^{(1)'}(2n+1), \quad \text{for } n\geqslant 1, \tag{5.21}$$

and

$$E_1^{(6)} = 1,$$

$$E_{2n+1}^{(6)}(1\to 2n+1) = y_2 M_0^{(1)''}(1,2) E_{2n}^{(1)}(2\to 2n+1), \quad \text{for } n\geqslant 1. \tag{5.22}$$

On the right hand side of Eqs. (5.15)–(5.18) all the divergences (as $g\to 0$) are in the $F^{(5)}(y;z)$ term. $F_{(y;z)}^{(5)}$ is disconnected from the factors (which are convergent) on either side. The integrals over the y_j variables in $F_{2n+1}^{(5)}(y)$ factor out, and these can be done independently of the other integrations. We will analyze $F^{(5)}(y;z)$ in the next subsection.

We will summarize our notation here for reference: (1) In $E_n^{(j)}(y)$ or $F_n^{(j)}(y)$, y stands for the n y_j variables appearing in $E_n^{(j)}(y)$ or $F_n^{(j)}(y)$. (2) In the generating functions $E^{(j)}(y;z)$ or $F^{(j)}(y;z)$, y stands for an arbitrary number of y_j variables. (3) The choice of the y_j variables in a factor is determined by the y_j variables on either side. For examples,

$$E_n^{(j)}(y) F_m^{(k)}(y) E_p^{(l)}(y) \equiv E_n^{(j)}(y_1\to y_n) F_m^{(k)}(y_n\to y_{n+m-1}) \times E_p^{(l)}(y_{n+m-1}\to y_{n+m+p-2}). \tag{5.23}$$

DIAGRAM	TERM IN INTEGRAND
$\underset{j}{\times}$	dy_j
$\overset{+}{\underset{j}{\times}}$	$\begin{cases}(y_j+2i) & \text{for } j \text{ odd}\\ (y_j+2i)^{-1} & \text{for } j \text{ even}\end{cases}$
$\overset{-}{\underset{j}{\times}}$	$\begin{cases}(y_j-2i) & \text{for } j \text{ odd}\\ (y_j-2i)^{-1} & \text{for } j \text{ even}\end{cases}$
$\underset{j}{\times}$ ---- $\underset{j+1}{\times}$	$(y_j+y_{j+1})^{-1}$
Example	
$\overset{+}{\underset{4}{\times}}$ ---- $\overset{-}{\underset{5}{\times}}$	$\dfrac{(y_5-2i)}{(y_4+2i)(y_4+y_5)}$

FIG. 5. Diagrammatic representation for the different functions of y_j occuring in the integrand.

(4) A bar separating two factors indicates that these factors do not have any y_j variables in common. For example,

$$E_n^{(j)}(y)\,|\,F_m^{(k)}(y)\equiv E_n^{(j)}(y_1\to y_n)F_m^{(k)}(y_{n+1}\to y_{n+m}). \tag{5.24}$$

(5) A product of two generating functions is understood in the following sense: Let

$$F^{(k)}(y;z)\equiv\sum_{n=1}^{\infty}z^nF_n^{(k)}(y) \tag{5.25a}$$

and

$$F^{(m)}(y;z)\equiv\sum_{n=1}^{\infty}z^nF_n^{(m)}(y). \tag{5.25b}$$

Then,

$$F^{(k)}(y;z)F^{(m)}(y;z)\equiv\sum_{n=2}^{\infty}z^n\sum_{l=1}^{n-1}F_l^{(k)}(y_1\to y_l)F_{n-l}^{(m)}(y_l\to y_{n-1}) \tag{5.26}$$

and

$$F^{(k)}(y;z)\,|\,F^{(m)}(y;z)\equiv\sum_{n=2}^{\infty}z^n\sum_{l=1}^{n-1}F_l^{(k)}(y_1\to y_l)F_{n-l}^{(m)}(y_{l+1}\to y_n). \tag{5.27}$$

B. Analysis of $F^{(5)}(y;z)$

At this stage we will intorduce the diagrammatic notation of Fig. 5 to represent the different terms in the integrand. We can write $F_{2k+1}^{(5)}(2\to 2k+2)$ in the following form by multiplying the pairs of matrices $M_e^{(1)}(2j,2j+1)M_0^{(1)}(2j+1,2j+2)$ in Eq. (5.11):

$$F_{2k+1}^{(5)}(2\to 2k+2)\doteq\prod_{j=1}^{k}[a(2j\to 2j+2),b(2j\to 2j+2)]\times[(y_{2k+2}+2i)^{-1},0]\quad(\text{product}\equiv 1\quad\text{for } k=0), \tag{5.28}$$

where

$$a(2j\to 2j+2)=\frac{(y_{2j+1}-2i)}{(y_{2j}+2i)(y_{2j}+y_{2j+1})(y_{2j+1}+y_{2j+2})}+\frac{e^{-2ir}}{(y_{2j}+2i)(y_{2j+1}+2i)} \tag{5.29a}$$

$$\equiv\overset{+}{\underset{2j}{\times}}\text{----}\overset{-}{\underset{2j+1}{\times}}\text{----}\underset{2j+2}{\times}+\frac{e^{-2ir}}{(y_{2j+1}+2i)}\overset{+}{\underset{2j}{\times}}\ \underset{2j+1}{\times}\ \underset{2j+2}{\times} \tag{5.29b}$$

and

$$b(2j\to 2j+2)=e^{-ir}\overset{+}{\underset{2j}{\times}}\ \underset{2j+1}{\times}\text{----}\underset{2j+2}{\times}+e^{+ir}\overset{+}{\underset{2j}{\times}}\text{----}\underset{2j+1}{\times}\ \underset{2j+2}{\times}. \tag{5.30}$$

In the second term of Eq. (5.29b), we can integrate over y_{2j+1} (and set $g=0$). Thus,

$$a(2j\to 2j+2)=\overset{+}{\underset{2j}{\times}}\text{----}\overset{-}{\underset{2j+1}{\times}}\text{----}\underset{2j+2}{\times}+e(r)\overset{+}{\underset{2j}{\times}}\ \underset{2j+2}{\times}, \tag{5.29c}$$

where

$$e(r)=e^{-2ir}\int_0^{\infty}dx\,e^{-rx}x(x+2i)^{-1}. \tag{5.31}$$

We can write Eq. (5.28) in the following form:

$$F_{2k+1}^{(5)}(2\to 2k+2)\doteq F_{2k+1}^{(6)}(2\to 2k+2)+\sum_{l=1}^{k}F_{2k+1,2l}^{(7)}(2\to 2k+2),\quad(\text{sum}\equiv 0\text{ for }k=0),\quad k=0,1,2,\ldots, \tag{5.32}$$

where

$$F_{2k+1}^{(6)}(2\to 2k+2)\doteq\left[\left(\prod_{j=1}^{k}a(2j\to 2j+2)\right)\overset{+}{\underset{2k+2}{\times}},0\right],\quad(\text{product}\equiv 1\text{ for }k=0),\quad k=0,1,2,\ldots \tag{5.33}$$

is a diagonal matrix and

$$F_{2k+1,2l}^{(7)}(2\to 2k+2)\doteq\left[\prod_{j=1}^{l-1}a(2j\to 2j+2),0\right][0,b(2l\to 2l+2)]\times\left\{\prod_{j=l+1}^{k}[a(2j\to 2j+2),b(2j\to 2j+2)]\right\}\times\left[\overset{+}{\underset{2k+2}{\times}},0\right],$$

(first product$\equiv$1 for $l=1$, second product$\equiv$1 for $l=k$),

$$l\leqslant k,\ k=1,2,\ldots. \tag{5.34}$$

Using the identity

$$b(2j\to 2j+2) = \overset{+}{\underset{2j}{\times}} b_1(2j+1,2j+2) + b_2(2j,2j+1)\underset{2j+2}{\times}, \tag{5.35}$$

where

$$b_1(2j+1,2j+2) = e^{-ir}\underset{2j+1}{\times}\cdots\cdots\underset{2j+2}{\times} \tag{5.36a}$$

and

$$b_2(2j,2j+1) = e^{+ir}\overset{+}{\underset{2j}{\times}}\cdots\cdots\underset{2j+1}{\times} \tag{5.36b}$$

we can factorize $F^{(7)}_{2k+1,2l}(2\to 2k+2)$ as

$$\begin{aligned}&F^{(7)}_{2k+1,2l}(2\to 2k+2)\\ &\doteq F^{(6)}_{2l-1}(2\to 2l)\,|S^{(1)}_{2k-2l+2}(2l+1\to 2k+2)\\ &\quad + F^{(8)}_{2l}(2\to 2l+1)\,|F^{(5)}_{2k-2l+1}(2l+2\to 2k+2),\\ &\qquad l\leqslant k,\quad k=1,2,3,\ldots,\end{aligned} \tag{5.37}$$

where

$$\begin{aligned}&S^{(1)}_{2n}(1\to 2n)\\ &\doteq [0,b_1(1,2)]\prod_{j=1}^{n-1}[a(2j\to 2j+2),b(2j\to 2j+2)]\\ &\quad\times\Big[\overset{+}{\underset{2n}{\times}},0\Big],\\ &\qquad(\text{product}\equiv 1 \text{ for } n=1),\quad n=1,2,\ldots,\end{aligned} \tag{5.38}$$

$$\begin{aligned}F^{(8)}_{2n}(1\to 2n+1)&\doteq\prod_{j=1}^{n-1}[a(2j\to 2j+2),0]\,[0,b_2(2n,2n+1)]\\ &(\text{product}\equiv 1 \text{ for } n=1),\quad n=1,2,\ldots.\end{aligned} \tag{5.39}$$

Carrying out a similar factorization for $S^{(1)}_{2n}(1\to 2n)$, we get the relations

$$S^{(1)}_2(1,2)\doteq F^{(8)^T}_2(1,2) \tag{5.40a}$$

and

$$\begin{aligned}&S^{(1)}_{2n}(1\to 2n)\\ &\doteq\sum_{j=1}^{n-1}F^{(8)^T}_{2j}(1\to 2j)\,|S^{(1)}_{2n-2j}(2j+1\to 2n)\\ &\quad+\sum_{j=1}^{n-1}F^{(9)}_{2j+1}(1\to 2j+1)\,|F^{(5)}_{2n-2j-1}(2j+2\to 2n)\\ &\quad+F^{(8)^T}_{2n}(1\to 2n),\\ &\qquad n=2,3,\ldots,\end{aligned} \tag{5.40b}$$

where

$$\begin{aligned}F^{(8)^T}_{2j}(1\to 2j)&\doteq[0,b_1(1,2)]\prod_{l=1}^{j-1}[a(2l\to 2l+2),0]\Big[\overset{+}{\underset{2j}{\times}},0\Big],\\ &(\text{product}\equiv 1 \text{ for } j=1),\quad j=1,2,\ldots\end{aligned} \tag{5.41}$$

is the transpose of $F^{(8)}_{2j}(y)$ (under the integral) and

$$\begin{aligned}F^{(9)}_{2j+1}(1\to 2j+1)&\doteq[0,b_1(1,2)]\prod_{l=1}^{j-1}[a(2l\to 2l+2),0]\\ &\quad\times[0,b_2(2j,2j+1)],\\ &(\text{product}\equiv 1 \text{ for } j=1),\quad j=1,2,\ldots.\end{aligned} \tag{5.42}$$

We define the generating functions

$$F^{(6)}(y;z)\equiv F^{(6)}(y,z;r,g)\equiv\sum_{n=0}^{\infty}z^{2n+1}F^{(6)}_{2n+1}(y), \tag{5.43a}$$

$$S^{(1)}(y;z)\equiv S^{(1)}(y,z;r,g)\equiv\sum_{n=1}^{\infty}z^{2n}S^{(1)}_{2n}(y), \tag{5.43b}$$

$$F^{(8)}(y;z)\equiv F^{(8)}(y,z;r,g)\equiv\sum_{n=1}^{\infty}z^{2n}F^{(8)}_{2n}(y), \tag{5.43c}$$

and

$$F^{(9)}(y;z)\equiv F^{(9)}(y,z;r,g)\equiv\sum_{n=1}^{\infty}z^{2n+1}F^{(9)}_{2n+1}(y). \tag{5.43d}$$

Then Eqs. (5.40) imply that

$$S^{(1)}(y;z)\doteq F^{(8)^T}(y;z)\,|S^{(1)}(y;z)+F^{(9)}(y;z)\,|F^{(5)}(y;z)+F^{(8)^T}(y;z). \tag{5.44}$$

Solving for $S^{(1)}(y;z)$, we have

$$S^{(1)}(y;z)\doteq[1-F^{(8)^T}(y;z)]^{-1}\{F^{(8)^T}(y;z)+F^{(9)}(y;z)\,|F^{(5)}(y;z)\}. \tag{5.45}$$

Using the relations (5.32) and (5.37), we get

$$F^{(5)}(y;z)\doteq[1-F^{(8)}(y;z)-F^{(6)}(y;z)[1-F^{(8)^T}(y;z)]^{-1}F^{(9)}(y;z)]^{-1}F^{(6)}(y;z)[1-F^{(8)^T}(y;z)]^{-1}. \tag{5.46}$$

What Eq. (5.46) says is the following: $F^{(j)}(y;z)$ ($j=5,6,8,$ and 9) are 2×2 matrices each term of which is a power series in z. The coefficient of z^n in the power series is an integrand with n variables. If we evaluate all the integrals and sum up the power series, we will get a function of z, r, and g. On substituting these functions in the elements of $F^{(j)}(y;z)$ and carrying out the operation on the right hand side of Eq. (5.46) we get an equality. At various stages of the above procedure we have factored out integrals so that they can be done for the different $F^{(j)}_n(y)$ s independently of one another.

Carrying out the matrix multiplications in Eq. (5.33), (5.39), and (5.42), we can write

$$\int dy F^{(6)}(y;z)=[q^{(1)}(z),0], \tag{5.47a}$$

$$\int dy F^{(8)}(y;z)=[0,q^{(2)}(z)], \tag{5.47b}$$

and

$$\int dy F^{(9)}(y;z)=[q^{(3)*}(z),0], \tag{5.47c}$$

where

$$q^{(1)}(z) \equiv q^{(1)}(z;r,g) = \int dy \sum_{n=0}^{\infty} z^{2n+1} q^{(1)}_{2n+1}(y), \tag{5.48}$$

$$q^{(2)}(z) \equiv q^{(2)}(z;r,g) = \int dy \sum_{n=1}^{\infty} z^{2n} q^{(2)}_{2n}(y), \tag{5.49}$$

and

$$q^{(3)}(z) \equiv q^{(3)}(z;r,g) = \int dy \sum_{n=1}^{\infty} z^{2n+1} q^{(3)}_{2n+1}(y)\,. \tag{5.50}$$

In Eqs. (5.48)–(5.50)

$$q^{(1)}_{2n+1}(2\to 2n+2) \doteq \left[\prod_{j=1}^{n} a(2j\to 2j+2)\right] \overset{+}{\underset{2n+2}{\times}}, \quad n=0,1,2,\ldots, \tag{5.51a}$$

$$q^{(2)}_{2n}(1\to 2n) \doteq b^{*}_{1}(1,2)\left[\prod_{j=1}^{n-1} a(2j\to 2j+2)\right] \overset{+}{\underset{2n}{\times}}, \quad n=1,2,3,\ldots, \tag{5.51b}$$

and

$$q^{(3)}_{2n+1}(1\to 2n+1) \doteq b^{*}_{1}(1,2)\left[\prod_{j=1}^{n-1} a(2j\to 2j+2)\right] b_2(2n,2n+1), \quad n=1,2,\ldots\,. \tag{5.51c}$$

In Eq. (5.51a) the product$\equiv 1$ for $n=0$ and in Eqs. (5.51b) and (5.51c) for $n=1$. Substituting Eq. (5.51) in (5.46) and carrying out the matrix multiplications leads to Eq. (7.15), where

$$\lim_{g\to 0} q^{(j)}(\pi^{-1};r,g) \equiv q^{(j)}(r). \tag{5.52}$$

C. Analysis of $q^{(j)}(z)$, $j=1,2,3$

We first show that

$$q^{(1)}(z) = \frac{A(z)}{1-ze(r)A(z)}, \tag{5.53}$$

where

$$A(z;r,g) \equiv A(z) = \int dy \sum_{k=0}^{\infty} z^{2k+1} \alpha_{2k+1}(y), \tag{5.54a}$$

with

$$\alpha_{2k+1}(2\to 2k+2) \doteq \overset{+}{\underset{2}{\times}} - - - - \overset{-}{\underset{3}{\times}} - - - - \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2k+1}{\times}} - - - \overset{+}{\underset{2k+2}{\times}}, \quad k=0,1,\ldots\,. \tag{5.54b}$$

Clearly,

$$q^{(1)}_{1}(2) \doteq \alpha_1(2), \tag{5.55}$$

$$q^{(1)}_{3}(2\to 4) \doteq \overset{+}{\underset{2}{\times}} - - - - \overset{-}{\underset{3}{\times}} - - - - \overset{+}{\underset{4}{\times}} + e(r) \overset{+}{\underset{2}{\times}} \overset{+}{\underset{4}{\times}} \doteq \alpha_3(2\to 4) + e(r)\alpha_1(2)\alpha_1(4) \tag{5.56a,b}$$

and satisfies Eq. (5.53).

We will now prove Eq. (5.53) by induction. Let

$$q^{(1)}_{2k-1}(4\to 2k+2) \doteq \left[\frac{A(z)}{1-ze(r)A(z)}\right]_{2k-1}, \tag{5.57}$$

where the subscript $2k-1$ denotes the coefficients of z^{2k-1} in the expansion on the right hand side (without the integration over y). Then,

$$q^{(1)}_{2k+1}(2\to 2k+1) \doteq a(2\to 4) q^{(1)}_{2k-1}(4\to 2k+1) \tag{5.58}$$

$$\doteq \left[\overset{+}{\underset{2}{\times}} - - - - \overset{-}{\underset{3}{\times}} - - - - \underset{4}{\times} + e(r) \overset{+}{\underset{2}{\times}} \underset{4}{\times}\right] \times \left\{\alpha_{2k-1} + e(r)[A^{2}(z)]_{2k-2} + \cdots + e^{k-1}(r)\alpha_1^{k}\right\}, \tag{5.59}$$

where inside the curly bracket the variables of integration start from y_4. Therefore,

$$q^{(1)}_{2k+1}(2\to 2k+2) \doteq \alpha_{(2k+1}(2\to 2k+2) + e^{k}(r)\alpha_1^{k+1} + \sum_{l=1}^{k-1} e^{l}(r) \times \left\{\alpha_1(2)[A^{l}(z)]_{2k-l} + \overset{+}{\underset{2}{\times}} - - - - \overset{-}{\underset{3}{\times}} - - - - \underset{4}{\times}\ [A^{l+1}(z)]_{2k-l+1}\right\}. \tag{5.60}$$

Now

$$\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\overset{-}{\times}}\text{----}\underset{4}{\times}\,[A^2(z)]_{2k-2}\doteq\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\overset{-}{\times}}\text{----}\underset{4}{\times}\sum_{l=1}^{k-1}\alpha_{2l-1}(4\to 2l+2)\alpha_{2k-2l-1}(2l+4\to 2k+2) \tag{5.61}$$

$$\doteq\sum_{l=1}^{k-1}\alpha_{2l+1}(2\to 2l+2)\alpha_{2k-2l-1}(2l+4\to 2k+2) \tag{5.62}$$

$$\doteq[A^2(z)]_{2k}-\alpha_1(2)\alpha_{2k-1}(4\to 2k+2)\,. \tag{5.63}$$

Similarly, it can be shown that

$$\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\overset{-}{\times}}\text{----}\underset{4}{\times}\,[A^{l+1}(z)]_{2k-l-1}\doteq(A^{l+1}(z))_{2k-l+1}-\alpha_1(2)[A^l(z)]_{2k-l}\,,\quad\text{for } l=2,3,\cdots,k-1. \tag{5.64}$$

Using the relations (5.63) and (5.64) in Eq. (5.60), we get

$$q^{(1)}_{2k+1}(2\to 2k+2)\doteq\left[\frac{A(z)}{1-ze(r)A(z)}\right]_{2k+1} \tag{5.65}$$

Hence, by induction, Eq. (5.53) holds.

In $\alpha_{2k+1}(2\to 2k+2)$ we use the identity

$$\underset{2j-1}{\overset{-}{\times}}\text{----}\underset{2j}{\overset{+}{\times}}=-\underset{2j-1}{\times}\text{----}\underset{2j}{\times}+\underset{2j-1}{\times}\underset{2j}{\overset{+}{\times}} \tag{5.66}$$

to write

$$\begin{aligned}\alpha_{2k+1}(2\to 2k+2)\doteq&\left(\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\times}\right)\alpha_{2k-1}(4\to 2k+2)\\&-\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\times}\text{----}\underset{4}{\times}\text{----}\underset{5}{\overset{-}{\times}}\text{----}\underset{6}{\overset{+}{\times}}\cdots\underset{2k+1}{\overset{-}{\times}}\text{----}\underset{2k+2}{\overset{+}{\times}}\,.\end{aligned} \tag{5.67}$$

Using the identity (5.66) for $j=3,4,\cdots,k+1$ and defining

$$\gamma_k(2\to k+1)\doteq\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\times}\text{----}\underset{4}{\times}\cdots\underset{k}{\times}\text{----}\underset{k+1}{\times}\,,\quad k=1,2,\ldots, \tag{5.68}$$

we obtain the recursion relations

$$\alpha_1(2)\doteq\gamma_1(2) \tag{5.69a}$$

and

$$\begin{aligned}\alpha_{2k+1}(2\to 2k+2)\doteq&(-1)^k\gamma_{2k+1}(2\to 2k+2)\\&+\sum_{l=1}^{k}(-1)^{l+1}\gamma_{2l}(2\to 2l+1)\alpha_{2k-2l+1}(2l+2\to 2k+2)\,,\quad k=1,2,\ldots\,.\end{aligned} \tag{5.69b}$$

We define the generating functions

$$\Gamma^{(1)}(z)\equiv\Gamma^{(1)}(z;r,g)=\int dy\sum_{k=0}^{\infty}(-1)^k z^{2k+1}\gamma_{2k+1}(y), \tag{5.70a}$$

$$\Gamma^{(2)}(z)\equiv\Gamma^{(2)}(z;r,g)=\int dy\sum_{k=1}^{\infty}(-1)^{k+1}z^{2k}\gamma_{2k}(y)\,. \tag{5.70b}$$

Then Eq. (5.69) implies that

$$A(z)=\frac{\Gamma^{(1)}(z)}{1-\Gamma^{(2)}(z)}\,. \tag{5.71}$$

We can rewrite $g_{2k+1}(gr)$ (see Eq. (3.3)] as

$$g_{2k+1}(gr)=(-1)^k\int_g^\infty dy_2\cdots\int_g^\infty dy_{2k+2}\prod_{j=2}^{2k+2}e^{-ry_j}\prod_{j=2}^{2k+1}(y_j+y_{j+1})^{-1}(y_{2k+2}^2-g^2)^{-1/2}\prod_{j=1}^{k}\frac{(y_{2j+1}^2-g^2)^{1/2}}{(y_{2j}^2-g^2)^{1/2}}\,, \tag{5.72}$$

$$=(-1)^k\int dy\left[\underset{2}{\times}\text{----}\underset{3}{\times}\text{----}\underset{4}{\times}\cdots\underset{2k+1}{\times}\text{----}\underset{2k+2}{\times}\right]. \tag{5.73}$$

Now in $\gamma_{2k+1}(2\to 2k+2)$ we use the identity

$$\underset{2}{\overset{+}{\times}}=(2i)^{-1}\left[\underset{2}{\times}-y_2\underset{2}{\overset{+}{\times}}\right] \tag{5.74}$$

to write

$$\int dy\gamma_{2k+1}(2\to 2k+2)=(2i)^{-1}\left[(-1)^k g_{2k+1}(gr)-\int dyy_2\underset{2}{\overset{+}{\times}}\text{----}\underset{3}{\times}\cdots\underset{2k+1}{\times}\text{----}\underset{2k+2}{\times}\right]. \tag{5.75}$$

Next we use the identity

$$\underset{2j+1}{\times} \text{----} \underset{2j+2}{\times} = y_{2j+1}^{-1}\left[\underset{2j+1}{\times}\underset{2j+2}{\times} - y_{2j+2}\underset{2j+1}{\times} \text{----} \underset{2j+2}{\times}\right] \tag{5.76}$$

for $j = 1,2,\cdots,k$ to write

$$\int dy\gamma_1(2) = (2i)^{-1}\left[g_1(gr) - \int dy y_2 \overset{+}{\underset{2}{\times}}\right] \tag{5.77a}$$

and

$$\int dy\gamma_{2k+1}(2\to 2k+2)$$

$$= (-1)^k(2i)^{-1}\Bigg\{g_{2k+1}(gr) + \sum_{l=1}^{k} g_{2k-2l+1}(gr)\int dy\left[\prod_{j=1}^{l}\frac{y_{2j}}{y_{2j+1}}\overset{+}{\underset{2}{\times}}\text{----}\underset{3}{\times}\text{----}\underset{4}{\times}\cdots\underset{2l}{\times}\text{----}\underset{2l+1}{\times}\right]$$

$$-\int dy\left(\frac{\Pi_{j=1}^{k+1}y_{2j}}{\Pi_{j=1}^{k}y_{2j+1}}\right)\overset{+}{\underset{2}{\times}}\text{----}\underset{3}{\times}\text{----}\underset{4}{\times}\cdots\underset{2k+1}{\times}\text{----}\underset{2k+1}{\times}\Bigg\},\quad k=1,2,\ldots. \tag{5.77b}$$

It is clear that in the last term and in the integrals multiplying the $g_{2m+1}(gr)$'s in the second term we can set $g = 0$. We do this and define

$$e^{(1)}(z;r) = \sum_{l=0}^{\infty} z^{2l}e_{2l}^{(1)}(r) \tag{5.78a}$$

and

$$e^{(2)}(z;r) = \sum_{l=0}^{\infty} z^{2l+1}e_{2l+1}^{(1)}(r), \tag{5.78b}$$

where

$$e_0^{(1)}(r) = 1 \tag{5.79a}$$

and

$$e_l^{(1)}(r) = \int_0^\infty dy_1\cdots\int_0^\infty dy_l(y_1+2i)^{-1}\prod_{j=1}^{l}e^{-ry_j}\prod_{j=1}^{l-1}(y_j+y_{j+1})^{-1},\quad l\geqslant 1. \tag{5.79b}$$

Then the recursion relation (5.77) implies that, as $g\to 0$,

$$\Gamma^{(1)}(z) = (2i)^{-1}[G(z;gr)e^{(1)}(z;r) - e^{(2)}(z;r)] + O(g^2). \tag{5.80}$$

Similarly, we can show that

$$\Gamma^{(2)}(z) = 1 - e^{(1)}(z;r) + \Gamma^{(2)}(z)e^{(3)}(z;r) + O(g^2), \tag{5.81}$$

where

$$e^{(3)}(z;r) = \sum_{l=0}^{\infty} z^{2l+1}e_{2l+1}^{(3)}(r), \tag{5.82}$$

with

$$e_l^{(3)}(r) = \int_0^\infty dy_1\cdots\int_0^\infty dy_l\prod_{j=1}^{l}e^{-ry_j}\prod_{j=1}^{l-1}(y_j+y_{j+1})^{-1}. \tag{5.83}$$

This completes the factorization of $q^{(1)}(z)$. $q^{(2)}(z)$ and $q^{(3)}(z)$ can be analysed in a similar manner by relating them to $q^{(1)}(z)$ and functions of r. These functions are in turn related to $e^{(j)}(z;r)$ $(j = 1,2,3)$. We will not write down the details of this calculation.

Note that the g dependence of the functions $q^{(j)}(z)$ $(j = 1,2,3)$ is only through $G(gr;z)$. We use the fact

$$\lim_{g\to 0} G(gr,\pi^{-1}) = 1$$

to take this limit in $q^{(j)}(z)$. This leads to Eq. (7.7)–(7.9) on using $e^{(3)}(r) = (2r)^{-1}$ (which is shown later in this section).

We will next reduce $e^{(j)}(\pi^{-1};r)$ $(j = 1,2,3)$ to their final form [Eqs. (7.3), (7.4) and (5.96)]. The procedure is the same in all three cases and will be illustrated for $e^{(1)}(\pi^{-1};r)$.

Scaling out r in Eq. (5.79b) $(ry_j\to y_j)$, we can write $e_l^{(1)}(r)$ in the form of an integral over iterated kernels

$$e_l^{(1)}(r) = \int_0^\infty d\sigma_1\cdots\int_0^\infty d\sigma_l(y_1+2ir)^{-1}K(1,2)K(2,3)\cdots K(l-1,l),\quad l\geqslant 1, \tag{5.84}$$

where

$$d\sigma_j = e^{-y_j}dy_j,\quad K(j,j+1) = (y_j+y_{j+1})^{-1}. \tag{5.85}$$

The kernel $(y_j + y_{j+1})^{-1}$ has been studied in Ref. 9, where the eigenvalues and eigenfunctions are explicitly written down. We use the results in Sec. E of this reference (for the special case $\nu = 0$)

$$\int_0^\infty d\sigma_2 K(1,2)\chi_p(2) = \lambda_p \chi_p(1), \tag{5.86}$$

where

$$\lambda_p = \pi \operatorname{sech}\pi p \tag{5.87}$$

are the eigenvalues and

$$\chi_p(x) = (2\lambda_p)^{-1/2} \int_1^\infty d\xi \exp[-(\xi - 1)x/2]\varphi_p(\xi) \tag{5.88}$$

are the eigenfunctions of the integral equation (5.86).

In Eq. (5.86)

$$\varphi_p(\xi) = C_p F(\tfrac{1}{2} + ip, \tfrac{1}{2} - ip, 1; \tfrac{1}{2} - \tfrac{1}{2}\xi) \tag{5.89}$$

where

$$C_p = (p \tanh\pi p)^{1/2} \tag{5.90}$$

and $F(a,b,c;z)$ is the hypergeometric function. $\phi_p(\xi)$ satisfies the integral equation

$$\int_1^\infty d\xi \frac{\varphi_p(\xi)}{\xi + x} = \lambda_p \varphi_p(x). \tag{5.91}$$

Writing the eigenfunction expansion for $K(j,j+1)$ as

$$K(j,j+1) = \int_0^\infty dp\, \lambda_p \chi_p(j)\chi_p(j+1) \tag{5.92}$$

and doing the integrations over $y_2 \to y_l$ in Eq. (5.84), we can write

$$e_l^{(1)}(r) = \int_0^\infty dp\, \lambda_p^{l-1}\left[\int_0^\infty dx\, e^{-x}\chi_p(x)\right] \times \left[\int_0^\infty dy\, e^{-y}(y + 2ir)^{-1}\chi_p(y)\right]. \tag{5.93}$$

Using Eq. (5.88) and (5.91), we can show that

$$\int_0^\infty dx e^{-x}\chi_p(x) = (2p\lambda_p \tanh\pi p)^{1/2}. \tag{5.94}$$

Using this in Eq. (5.92), doing the sum in Eq. (5.78a) (which in a geometric progression is $z^2\lambda^2 p$), and setting $z = \pi^{-1}$, we get Eq. (7.3). A similar analysis leads to Eq. (7.4).

We can similarly show that

$$e_l^{(3)}(r) = \frac{2}{r}\int_0^\infty dp p \lambda_p^l \tanh\pi p. \tag{5.95}$$

Substituting in Eq. (5.82), summing up the series, setting $z = \pi^{-1}$, and doing the final integral over p, we get

$$e^{(3)}(\pi^{-1}; r) = (2r)^{-1}. \tag{5.96}$$

This completes our analysis of the second term in Eq. (4.26) in the limit $g \to 0$. On taking this limit we get the expression for $H^{(1)}(r)$ in the final result (see Sec. VII).

VI. ANALYSIS OF $G_0(1 \to 2n)$

A. Factorization of $G_0(1 \to 2n)$

We can write

$$G_0(1 \to 2n) \doteqdot \prod_{j=1}^n [a(2j \to 2j+2), b(2j \to 2j+2)]. \tag{6.1}$$

The product in Eq. (6.1) has a cyclic structure, i.e., $y_{2n+1} \equiv y_1$ and $y_{2n+2} \equiv y_2$. Separating the a's as in Eq. (5.32), we have

$$G_0(1 \to 2n) \doteqdot \left[\prod_{j=1}^n a(2j \to 2j+2), 0\right] + \sum_{l=1}^n [F_{2n,2l}^{(10)}(1 \to 2n) + F_{2n,2l}^{(11)}(2 \to 2n, 1)], \tag{6.2}$$

where

$$F_{2n,2l}^{(10)}(1 \to 2n) \doteqdot [0, b_1(1,2)] \prod_{j=1}^{n-l} [a(2j \to 2j+2), b(2j \to 2j+2)] \times \prod_{j=n-l+1}^{n-1} [a(2j \to 2j+2), 0]\left[\overset{+}{\underset{2n}{\times}}, 0\right] \tag{6.3}$$

and

$$F_{2n,2l}^{(11)}(2 \to 2n, 1) \doteqdot \prod_{j=1}^{n-l} [a(2j \to 2j+2), b(2j \to 2j+2)] \times \prod_{j=n-l+1}^{n-1} [a(2j \to 2j+2), 0] \times [0, b_2(2n, 1)] \tag{6.4}$$

with $l \leqslant n$, $n = 1, 2, \ldots$. In Eq. (6.2) only the first term has a cyclic structure. In Eqs. (6.3) and (6.4) the first product $\equiv 1$ for $l = n$ and the second product $\equiv 1$ for $l = 1$.

By methods similar to those used in Sec. V we can derive the following recursion relations:

$$F_{2n,2l}^{(10)}(1 \to 2n) \doteqdot F_{2n}^{(8)^T}(1 \to 2n) + \sum_{k=1}^{n-l} \{F_{2k}^{(8)^T}(1 \to 2k) | F_{2n-2k,2l}^{(10)}(2k+1 \to 2n) + F_{2k+1}^{(9)}(1 \to 2k+1) | F_{2n-2k-1,2l-1}^{(12)}(2k+2 \to 2n)\}, \tag{6.5}$$

where

$$F_{2n-1,2l-1}^{(12)}(2 \to 2n) \doteqdot \prod_{j=1}^{n-l} [a(2j \to 2j+2), b(2j \to 2j+2)] \prod_{j=n-l+1}^{n-1} [a(2j \to 2j+2), 0]\left[\overset{+}{\underset{2n}{\times}}, 0\right], \tag{6.6}$$

and

$$F^{(11)}_{2n,2l}(2\to 2n,1) \doteq F^{(8)}_{2n}(2\to 2n,1) + \sum_{k=1}^{n-l} \{F^{(6)}_{2k-1}(2\to 2k) | F^{(13)}_{2n-2k+1,2l}(2k+1\to 2n,1)$$
$$+ F^{(8)}_{2k}(2\to 2k+1) | F^{(11)}_{2n-2k,2l}(2k+2\to 2n,1)\}, \tag{6.7}$$

where

$$F^{(13)}_{2n-1,2l}(1\to 2n-1) \doteq [0,b_1(1,2)] \prod_{j=1}^{n-l-1} [a(2j\to 2j+2), b(2j\to 2j+2)]$$
$$\times \prod_{j=n-l}^{n-2} [a(2j\to 2j+2),0] [0,b_2(2n-2,2n-1)], \quad n\geqslant l+1, \ l\geqslant 1. \tag{6.8}$$

In Eqs. (6.5)–(6.7), $l\leqslant n$ and $n = 1,2,\ldots$.

Defining the generating functions

$$F^{(k)}(z,\omega;y) \equiv F^{(k)}(z,\omega;y,r,g) \equiv \sum_{n=1}^{\infty} \sum_{l=1}^{n} z^{2n}\omega^{2l} F^{(k)}_{2n,2l}(y), \quad \text{for } k = 10 \text{ and } 11, \tag{6.9a}$$

$$F^{(12)}(z,\omega;y) \equiv F^{(12)}(z,\omega;y,r,g) \equiv \sum_{n=1}^{\infty} \sum_{l=1}^{n} z^{2n-1}\omega^{2l-1} F^{(12)}_{2n-1,2l-1}(y), \tag{6.9b}$$

and

$$F^{(13)}(z,\omega;y) \equiv F^{(13)}(z,\omega;y,r,g) \equiv \sum_{n=2}^{\infty} \sum_{l=1}^{n-1} z^{2n-1}\omega^{2l} F^{(13)}_{2n-1,2l}(y), \tag{6.9c}$$

we have

$$F^{(10)}(z,1;y) \doteq F^{(8)^T}(y;z) \left| F^{(10)}(z,1;y) + \frac{1}{2} z \frac{d}{dz} F^{(8)^T}(y;z) + F^{(9)}(y;z) \right| F^{(12)}(z,1;y) \tag{6.10}$$

and

$$F^{(11)}(z,1;y) \doteq \frac{1}{2} z \frac{d}{dz} F^{(8)}(y;z) + F^{(6)}(y;z) | F^{(13)}(z,1;y) + F^{(8)}(y;z) | F^{(11)}(z,1;y). \tag{6.11}$$

The reason for evaluating these functions at $\omega = 1$ will be clear later on. We can now reduce $F^{(12)}(z,1;y)$ and $F^{(13)}(z,1;y)$ by deriving recursion relations to get

$$F^{(12)}(z,1;y) \doteq \frac{1}{2} \frac{d}{dz} [zF^{(6)}(y;z)] + F^{(6)}(y;z) | F^{(10)}(z,1;y) + F^{(8)}(y;z) | F^{(12)}(z,1;y) \tag{6.12}$$

and

$$F^{(13)}(z,1;y) \doteq \frac{1}{2} \left(z \frac{d}{dz} - 1\right) F^{(9)}(y;z) + F^{(8)^T}(y;z) | F^{(13)}(z,1;y) + F^{(9)}(y;z) | F^{(11)}(z,1;y). \tag{6.13}$$

Solving for $F^{(12)}(z,1;y)$ and $F^{(13)}(z,1;y)$ and substituting in Eqs. (6.10) and (6.11), we get

$$F^{(10)}(z,1;y) \doteq \frac{1}{2} [1 - F^{(8)^T}(y;z) - F^{(9)}(y;z)[1 - F^{(8)}(y;z)]^{-1} F^{(6)}(y;z)]^{-1}$$
$$\times \left\{ z \frac{d}{dz} F^{(8)^T}(y;z) + F^{(9)}(y;z)[1 - F^{(8)}(y;z)]^{-1} \left(z \frac{d}{dz} + 1\right) F^{(6)}(y;z) \right\} \tag{6.14}$$

and

$$F^{(11)}(z,1;y) \doteq \frac{1}{2} [1 - F^{(8)}(y;z) - F^{(6)}(y;z)[1 - F^{(8)^T}(y;z)]^{-1} F^{(9)}(y;z)]^{-1}$$
$$\times \left\{ z \frac{d}{dz} F^{(8)}(y;z) + F^{(6)}(y;z)[1 - F^{(8)^T}(y;z)]^{-1} \left(z \frac{d}{dz} - 1\right) F^{(9)}(y;z) \right\}. \tag{6.15}$$

$F^{(10)}(z,1;y)$ and $F^{(11)}(z,1;y)$ each have one term without any derivative. On taking the trace the contribution of these nonderivative terms to $\mathrm{Tr}[F^{(10)}(z,1;y) + F^{(11)}(z,1;y)]$ is of the form

$$\mathrm{Tr}[A^{-1}B - (A^T)^{-1}B^T] = \mathrm{Tr}[A^{-1}B - (A^{-1}B)^T] = 0, \tag{6.16}$$

where

$$A = 1 - F^{(8)^T}(y;z) - F^{(9)}(y;z)[1 - F^{(8)}(y;z)]^{-1} F^{(6)}(y;z) \tag{6.17a}$$

and

$$B = \tfrac{1}{2} F^{(9)}(y;z)[1 - F^{(8)}(y;z)]^{-1} F^{(6)}(y;z) \tag{6.17b}$$

since

$$F^{(6)^T}(y;z) \doteq F^{(6)}(y;z) \quad \text{and} \quad F^{(9)^T}(y;z) = F^{(9)}(y;z). \tag{6.18}$$

Thus, the nonderivative terms cancel out on taking the trace.

Substrituting the expressions (5.47) for $F^{(6)}(y;z)$, $F^{(8)}(y;z)$, and $F^{(9)}(y;z)$ in Eqs. (6.14) and (6.15) and carrying out the matrix multiplications, we can show that

$$\int dy\, \mathrm{Tr}[F^{(10)}(z,1;y) + F^{(11)}(z,1;y)] = -\frac{1}{2} z \frac{d}{dz} \ln q(z)\,, \tag{6.19}$$

where $q(z)$ is given by

$$q(z) \equiv q(z;r,g) = [1 - q^{(2)}(z)q^{(2)*}(z)]^2 - q^{(1)}(z)q^{(3)*}(z) - q^{(1)*}(z)q^{(3)}(z) + q^{(1)}(z)q^{(1)*}(z)q^{(3)}(z)q^{(3)*}(z)$$
$$- q^{(1)}(z)q^{(2)*^2}(z)q^{(3)}(z) - q^{(1)*}(z)q^{(2)^2}(z)q^{(3)*}(z)\,. \tag{6.20}$$

We need to evaluate $\Sigma_{n=1}^{\infty} n^{-1} z^{2n}\, \mathrm{Tr} G_0(1\to 2n)$. Using Eq. (6.2), we have

$$\sum_{n=1}^{\infty} n^{-1} z^{2n}\, \mathrm{Tr} G_0(1\to 2n) \doteq 2\,\mathrm{Re} \sum_{n=1}^{\infty} n^{-1} z^{2n} \prod_{j=1}^{n} a(2j\to 2j+2) + 2\int_0^z \frac{dz'}{z'} \mathrm{Tr}[F^{(10)}(z',1;y) + F^{(11)}(z',1;y)]\,. \tag{6.21}$$

The second term can be evaluated using Eq. (6.19). The first term will be analyzed in the next subsection.

B. Evaluation of $\Pi_{j=1}^{n} a(2j\to 2j+2)$

Because of its cyclic structure this term requires a different factorization than that of $q^{(1)}(z)$. Using the equation (5.29c) for $a(2j\to 2j+2)$, we can write

$$\prod_{j=1}^{n} a(2j\to 2j+2) \doteq \bar{a}_{2n}(2\to 2n,1) + e(r) \sum_{l=1}^{n} d_{2n-1,2l-1}(2\to 2n)\,, \qquad n=1,2,\ldots \tag{6.22}$$

where

$$\bar{a}_{2n}(2\to 2n,1) \doteq (y_1+y_2)^{-1} \overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} ---- \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2n-1}{\times}} ---- \overset{+}{\underset{2n}{\times}} ---- \overset{-}{\underset{1}{\times}} \tag{6.23}$$

and

$$d_{2n-1,2l-1}(2\to 2n) \doteq \left[\prod_{j=1}^{n-1} a(2j\to 2j+2)\right] \overset{+}{\underset{2n-2l+2}{\times}} ---- \overset{-}{\underset{2n-2l+3}{\times}} ---- \overset{+}{\underset{2n-2l+4}{\times}} \cdots \overset{-}{\underset{2n-1}{\times}} ---- \overset{+}{\underset{2n}{\times}}\,,$$
$$(\text{product} \equiv 1 \ \text{for}\ l=n)\,. \tag{6.24}$$

We can write down the recursion relations

$$d_{2n-1,2n-1}(2\to 2n) \doteq \alpha_{2n-1}(2\to 2n) \tag{6.25a}$$

and

$$d_{2n-1,2l-1}(2\to 2n) \doteq \alpha_{2n-1}(2\to 2n) + e(r) \sum_{m=0}^{n-l-1} \alpha_{2m+1}(2\to 2m+2) d_{2n-2m-3,2l-1}(2m+4\to 2n),$$
$$\text{for}\ l<n;\quad n=2,3,\ldots\,. \tag{6.25b}$$

Define

$$D(z,\omega) \equiv D(z,w;r,g) = \int dy \sum_{n=1}^{\infty} \sum_{l=1}^{n} z^{2n-1} \omega^{2l-1} d_{2n-1,2l-1}(y)\,. \tag{6.26}$$

Then,

$$D(z,\omega) = \int dy \sum_{n=1}^{\infty} \sum_{l=1}^{n} z^{2n-1} \omega^{2l-1} \alpha_{2n-1}(y) + z\, e(r) A(z) D(z,\omega)\,. \tag{6.27}$$

Hence,

$$D(z,1) = \frac{1}{2}\frac{d}{dz}[zA(z)] + z\, e(r) A(z) D(z,1)\,. \tag{6.28}$$

Solving for $D(z,1)$, we get

$$D(z,1) = -[2e(r)]^{-1} \frac{d}{dz}\{\ln[1 - z\, e(r) A(z)]\}\,. \tag{6.29}$$

Now we use the identity (5.66),

$$\overset{-}{\underset{2j+1}{\times}} ---- \overset{+}{\underset{2j+2}{\times}} \doteq - \underset{2j+1}{\times} ---- \underset{2j+2}{\times} + \underset{2j+1}{\times}\ \overset{+}{\underset{2j+2}{\times}}\,, \tag{5.66}$$

for $j=0,1,\cdots,n-1$ in Eq. (6.23) to get

$$\bar{a}_{2n}(2\to 2n,1) \doteq (-1)^n (y_1+y_2)^{-1} \underset{2}{\times} ---- \underset{3}{\times} ---- \underset{4}{\times} \cdots \underset{2n}{\times} ---- \underset{1}{\times} + \sum_{k=0}^{n-1} (-1)^k p_{2n,2n-2k}(2\to 2n,1)\,, \tag{6.30}$$

where

$$p_{2n,2k}(2\to 2n,1) \doteqdot \underset{2}{\overset{+}{\times}} ---- \underset{3}{\overset{-}{\times}} ---- \underset{4}{\overset{+}{\times}} \cdots \underset{2k-1}{\overset{-}{\times}} ---- \underset{2k}{\overset{+}{\times}} ---- \underset{2k+1}{\times} \cdots \underset{2n}{\times} ---- \underset{1}{\times}\,. \tag{6.31}$$

In Eq. (6.30) the first term is equal to $nf_{2n}(gr)$ (on integrating over y). Define

$$P(z,\omega) \equiv P(z,\omega;r,g) + \int dy \sum_{n=1}^{\infty} \sum_{k=1}^{n} (-1)^{n+k} z^{2n} \omega^{2k} p_{2n,2k}(y)\,. \tag{6.32}$$

Then Eq. (6.30) leads to

$$\int dy \sum_{n=1}^{\infty} n^{-1} z^{2n} \bar{a}_{2n} = F(z;gr) + 2\int_0^z \frac{dz'}{z'} P(z',1)\,. \tag{6.33}$$

We now use Eq. (5.66) for $j = k-1$ in Eq. (6.31) to get

$$p_{2n,2k}(2\to 2n,1) \doteqdot -p_{2n,2k-2}(2\to 2n,1) + p_{2k-2,2k-2}(2\to 2k-1)\gamma_{2n-2k+2}(2k\to 2n,1)\,, \quad \text{for } 2\leqslant k\leqslant n, \quad n = 2,3,... \tag{6.34}$$

and

$$p_{2n,2}(y) = \gamma_{2n}(y), \quad n = 1,2,...\,. \tag{6.35}$$

Thus,

$$\sum_{n=2}^{\infty} \sum_{k=2}^{n} (-1)^{n+k} z^{2n}\omega^{2k} p_{2n,2k}(y) \doteqdot \sum_{n=2}^{\infty} \sum_{k=2}^{n} (-1)^{n+k} z^{2n}\omega^{2k}\left[-p_{2n,2k-2}(y) + p_{2k-2,2k-2}(y)\gamma_{2n-2k+2}(y)\right] \tag{6.36}$$

Adding and subtracting the missing terms on the left and right leads to

$$P(z,\omega) - \omega^2 \Gamma^{(2)}(z) = \omega^2\left\{P(z,\omega) - [1-\Gamma^{(2)}(z)]\int dy \sum_{n=1}^{\infty} (z\omega)^{2n} p_{2n,2n}(y)\right\}. \tag{6.37}$$

We can show that

$$\int dy \sum_{n=1}^{\infty} z^{2n} p_{2n,2n}(y) = \Gamma^{(2)}(z)[1-\Gamma^{(2)}(z)]^{-1}. \tag{6.38}$$

Substituting Eq. (6.38) in (6.37) and solving for $P(z,\omega)$, we get

$$P(z,\omega) = \frac{\omega^2[\Gamma^{(2)}(z) - \Gamma^{(2)}(z\omega)]}{(1-\omega^2)[1-\Gamma^{(2)}(z\omega)]}\,. \tag{6.39}$$

For $\omega = 1$, both the numerator and the denominator in Eq. (6.39) are equal to zero. Using L'Hospital's rule, we have

$$P(z,1) = -\frac{z}{2}\frac{d}{dz}\ln[1-\Gamma^{(2)}(z)]\,. \tag{6.40}$$

Equations (6.22), (6.29), (6.33), and (6.40) lead to

$$\int dy \sum_{n=1}^{\infty} n^{-1} z^{2n} \sum_{j=1}^{n} a(2j\to 2j+2) = F(z;gr) - \ln[1-\Gamma^{(2)}(z) - z\,e(r)\Gamma^{(1)}(z)]\,. \tag{6.41}$$

Sustituting Eq. (6.41) in (6.21) and doing the integral using Eq. (6.19) leads to

$$\int dy \sum_{n=1}^{\infty} n^{-1} z^{2n}\, \mathrm{Tr} G_0(1\to 2n) = 2F(z;gr) - 2\,\mathrm{Re}\{\ln[1-\Gamma^{(2)}(z) - z\,e(r)\Gamma^{(1)}(z)]\} - \ln q(z)\,. \tag{6.42}$$

We can now summarize the results of Secs. IV–VI to write down $\bar{H}[z;r,G(z;gr)]$ [see Eq. (3.4)]

$$\bar{H}[z;r,G(z;gr)] = \int dy \sum_{l=1}^{\infty} l^{-1}\, \mathrm{Tr}\mathscr{N}^l(y;z) - 2\,\mathrm{Re}\ln[1-\Gamma^{(2)}(z) - z\,e(r)\Gamma^{(1)}(z)] - \ln q(z)\,. \tag{6.43}$$

The lower limit of integration in the first term of Eq. (6.43) is now zero. We have shown that in all the terms on the right hand side of Eq. (6.43) the g dependence is only through $G(z;gr)$.

We can now use Eq. (2.23), (2.26), and (2.32) of Ref. 7, which say that

$$\lim_{g\to 0} (gr)^{1/2} \exp[-2F(\pi^{-1},gr)] = \pi^{-1}\rho_\infty\,, \tag{6.44}$$

where we have expressed the constants in Eq. (2.32) in terms of ρ_∞ given by Eq. (1.7). This completes the analysis of the $g\to 0$ limit of $g^{1/2}\exp[-H(\pi^{-1};r,g)]$. Setting

$$\lim_{g\to 0} \Gamma^{(1)}(\pi^{-1};r,g) = \Gamma^{(1)}(r), \tag{6.45}$$

$$\lim_{g\to 0} q(\pi^{-1};r,g) = q(r), \tag{6.46}$$

we get the final answer for $q(r)$ which is written out in detail in the next section.

VII. RESULTS

Using the results of Secs. IV–VI, we can write down the final form of $\rho(r)$:

$$\rho(r) = \frac{\rho_\infty}{r^{1/2}} (q(r)\{e^{(1)}(r) - [(2r)^{-1} + \pi^{-1}e(r)]\Gamma^{(1)}(r)\}$$
$$\times\{e^{(1)*}(r) - [(2r)^{-1} + \pi^{-1}e^*(r)]\Gamma^{(1)*}(r)\})$$
$$\times \exp[-H^{(1)}(r)]), \tag{7.1}$$

where ρ_∞ is given by Eq. (1.4) and

$$e(r) = e^{-2ir}\int_0^\infty dx\, e^{-rx}x(x+2i)^{-1}, \tag{7.2}$$

$$e^{(1)}(r) = 1 + \left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty dp\left(\frac{p}{\sinh^3\pi p}\right)^{1/2}$$
$$\times\int_0^\infty dy\, e^{-y}\frac{\chi_p(y)}{y+2ir}, \tag{7.3}$$

$$e^{(2)}(r) = \left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty dp\left(\frac{p\,\mathrm{sech}\pi p}{\tanh^3\pi p}\right)^{1/2}$$
$$\times\int_0^\infty dy e^{-y}\frac{\chi_p(y)}{y+2ir}. \tag{7.4}$$

$\chi_p(y)$ satisfies Eq. (5.86) and is explicitly written down in Eqs. (5.88) and (5.89):

$$\Gamma^{(1)}(r) = (2i)^{-1}[e^{(1)}(r) - e^{(2)}(r)] \tag{7.5}$$

and

$$q(r) = [1 - q^{(2)}(r)q^{(2)*}(r)]^2 - q^{(1)}(r)q^{(3)*}(r)$$
$$- q^{(1)*}(r)q^{(3)}(r) + q^{(1)}(r)q^{(1)*}(r)q^{(3)}(r)q^{(3)*}(r)$$
$$- q^{(1)}(r)q^{(2)*^2}(r)q^{(3)}(r) - q^{(1)*}(r)q^{(2)^2}(r)q^{(3)*}(r), \tag{7.6}$$

with

$$q^{(1)}(r) = \frac{\Gamma^{(1)}(r)}{e^{(1)}(r) - [(2r)^{-1} + \pi^{-1}e(r)]\Gamma^{(1)}(r)}, \tag{7.7}$$

$$q^{(2)}(r) = \frac{e^{ir}\{(2r)^{-1} - \pi^{-1}e(r)[e^{(1)}(r) - 1]\}}{e^{(1)}(r)}q^{(1)}(r)$$
$$- \frac{e^{ir}[e^{(1)}(r) - 1]}{e^{(1)}(r)}, \tag{7.8}$$

and

$$q^{(3)}(r) = \frac{e^{ir}\{(2r)^{-1} - \pi^{-1}e(r)[e^{(1)}(r) - 1]\}}{e^{(1)}(r)}q^{(2)}(r)$$
$$+ e^{2ir}[e^{(1)}(r)]^{-1}[(2r)^{-1} + 2ie^{(2)*}(r)$$
$$- \pi^{-1}e^{-2ir}e^*(r)e^{(1)*}(r)]. \tag{7.9}$$

In Eq. (7.1), $H^{(1)}(r)$ is given by

$$H^{(1)}(r) = \sum_{n=2}^\infty H_n^{(1)}(r), \tag{7.10}$$

where

$$H_n^{(1)}(r) = \int_0^\infty dy_1 \cdots \int_0^\infty dy_n$$
$$\times \prod_{j=1}^n e^{-ry_j}\prod_{j=0} y_{2j+1}\prod_{j=1} y_{2j}^{-1}f_n(y;r) \tag{7.11}$$

and

$$f_n(y;r) \equiv f_n(y_1, y_2, \cdots, y_n; r)$$
$$= \left[\sum_{l=1}^\infty l^{-1}\,\mathrm{Tr}\mathscr{N}(y;r))^l\right]_n, \tag{7.12}$$

where $\mathscr{N}(y;r)$ is a 4×4 matrix of functions of y and r. The summation inside the bracket on the right hand side of Eq. (7.12) generates terms with m y_j's ($m = 2,3,4...$). The subscript n denotes that we keep all the terms with exactly n y_j's in $f_n(y;r)$. The discussion following Eqs. (4.21) and (4.22) explains the method of generating these terms in detail.

In Eq. (7.12)

$$\mathscr{N}(y;r) = \begin{bmatrix} M_0^{(2)}(y;r)F^{(1)}(y;r) & M_0^{(2)}(y;r)F^{(2)}(y;r) \\ M_e^{(2)}(y;r)F^{(3)}(y;r) & M_e^{(2)}(y;r)F^{(4)}(y;r)\end{bmatrix}, \tag{7.13}$$

where $M_0^{(2)}(y;r)$ and $M_e^{(2)}(y;r)$ are given by Eqs. (4.8b) and (4.9b), respectively:

$$F^{(j)}(y;r) = E^{(j,1)}(y;r)F^{(5)}(r)E^{(j,2)}(y;r) + E^{(j,3)}(y;r),$$
$$j = 1,2,3, \text{ and } 4, \tag{7.14}$$

where

$$F^{(5)}(r) = \frac{1}{q(r)}\begin{bmatrix} q^{(1)}(r)[1 - q^{(1)*}(r)q^{(3)}(r)] + q^{(1)*}(r)q^{(2)^2}(r) & q^{(1)}(r)q^{(2)*}(r) + q^{(1)*}(r)q^{(2)}(r) \\ q^{(1)}(r)q^{(2)*}(r) + q^{(1)*}(r)q^{(2)}(r) & q^{(1)*}(r)[1 - q^{(1)}(r)q^{(3)*}(r)] + q^{(1)}(r)q^{(2)*^2}(r)\end{bmatrix}, \tag{7.15}$$

$$E^{(1,1)}(y;r) = E^{(2,1)}(y;r) = E^{(1)}(y;r)M_0^{(1)'}(y;r),$$

$$[M_0^{(1)'}(y;r) \text{ is given by Eq. (5.2a) and } E^{(1)}(y;r) \text{ by Eqs. (5.4) and (5.6)}], \tag{7.16}$$

$$E^{(3,1)}(y;r) = E^{(4,1)}(y;r) = E^{(5)}(y;r), \quad [E^{(5)}(y;r) \text{ is given by Eqs. (5.19) and (5.21)}], \tag{7.17}$$

$$E^{(1,2)}(y;r) = E^{(3,2)}(y;r) = E^{(2)}(y;r), \quad [E^{(2)}(y;r) \text{ is given by Eqs. (5.12) and (5.14b)}], \tag{7.18}$$

$$E^{(2,2)}(y;r) = E^{(4,2)}(y;r) = \pi^{-1}E^{(2)}(y;r)M_0^{(1)}(y;r), \quad [M_0^{(1)}(y;r) \text{ is given by Eq. (4.8a)}], \tag{7.19}$$

$$E^{(1,3)}(y;r) = E^{(1)}(y;r)M_0^{(1)'}(y;r)E^{(3)}(y;r) + E^{(1)}(y;r), \quad [E^{(3)}(y;r) \text{ is given by Eqs. (5.13) and (5.14c)}], \tag{7.20}$$

$$E^{(2,3)}(y;r) = \pi^{-1}E^{(1,3)}(y;r)M_0^{(1)}(y), \tag{7.21}$$

$$E^{(3,3)}(y;r) = E^{(5)}(y;r)E^{(3)}(y;r) + E^{(6)}(y;r), \quad [E^{(6)}(y;r) \text{ is given by Eqs. (5.19) and (5.22)}], \tag{7.22}$$

and

$$E^{(4,3)}(y;r) = E^{(3,3)}(y;r)M_0^{(1)}(y;r). \tag{7.23}$$

The 2×2 matrices $E^{(j)}$ have functions of the y_k's. Substituting these in Eq. (7.13), we carry out the trace in Eq. (7.12) and pick up all the terms with n variables to obtain, $f_n(y;r)$. In Eq. (7.16)–(7.29)

$$E^{(j)}(y;r) \equiv E^{(j)}(\pi^{-1}, y; r), \quad \text{for } j = 1,2,3,5, \text{ and } 6. \tag{7.24}$$

VIII. ASYMPTOTIC EXPANSION OF $\rho(r)$ FOR LARGE r

All the functions inside the curly bracket in Eq. (7.1) can be expressed in terms of $e^{(1)}(r)$, $e^{(2)}(r)$, and $e(r)$. It may be shown that, for $r \gg 1$,

$$e(r) = \frac{e^{-2ir}}{2ir^2}\left[1 - \frac{1}{ir} - \frac{3}{2r^2} + 0(r^{-3})\right]. \tag{8.1}$$

To expand $e^{(1)}(r)$ [Eq. (7.3)] and $e^{(2)}(r)$ [Eq. (7.4)], we expand the $(y + 2ir)^{-1}$ factor in the denominators in powers of r^{-1} to write

$$e^{(1)}(r) = 1 + \sum_{n=1}^{\infty} C_n^{(1)} r^{-n} \tag{8.2a}$$

and

$$e^{(2)}(r) = \sum_{n=1}^{\infty} C_n^{(2)} r^{-n}, \tag{8.2b}$$

where

$$C_n^{(1)} = (-1)^{n-1}(2i)^{-n}\left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty dp\left(\frac{p}{\sinh^3\pi p}\right)^{1/2} \times \int_0^\infty dy\, e^{-y} y^{n-1}\chi_p(y) \tag{8.3a}$$

and

$$C_n^{(2)} = (-1)^{n-1}(2i)^{-n}\left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty dp\left(\frac{p\,\mathrm{sech}\pi p}{\tanh^3\pi p}\right)^{1/2} \times \int_0^\infty dy\, e^{-y} y^{n-1}\chi_p(y). \tag{8.3b}$$

Using Eq. (5.88), we can write

$$\int_0^\infty dy\, e^{-y} y^{n-1}\chi_p(y) = \frac{(n-1)!2^n}{(2\lambda_p)^{1/2}}\int_1^\infty d\xi \frac{\phi_p(\xi)}{(\xi+1)^n}, \tag{8.4}$$

where $\phi_p(\xi)$ is defined by Eq. (5.89). Using Eq. (5.91), we have

$$\int_1^\infty d\xi \frac{\varphi_p(\xi)}{(\xi+1)^n} = \frac{(-1)^{n-1}\lambda_p}{(n-1)!}\frac{d^{n-1}\varphi_p(x)}{dx^{n-1}}\bigg|_{x=1} \tag{8.5a}$$

$$= \frac{2^{1-n}\lambda_p C_p}{[(n-1)!]^2}\frac{\Gamma(n-\frac{1}{2}+ip)\Gamma(n-\frac{1}{2}-ip)}{\Gamma(\frac{1}{2}+ip)\Gamma(\frac{1}{2}-ip)}, \tag{8.5b}$$

where, in Eq. (8.5b), $\Gamma(z)$ is the gamma function. Substituting in Eqs. (8.4) and (8.3) and doing the final integration over p, we can evaluate $C_n^{(1)}$ and $C_n^{(2)}$. We write down the first few terms in Eq. (8.2). For $r \gg 1$,

$$e^{(1)}(r) = 1 + \frac{1}{8ir} + \frac{3}{128r^2} - \frac{15}{1024ir^3} - \frac{525}{2\times 4^7 r^4} + O(r^{-5}) \tag{8.6a}$$

and

$$e^{(2)}(r) = \frac{1}{4ir} + \frac{3}{32r^2} - \frac{45}{512ir^3} - \frac{525}{4^6 r^4} + O(r^{-5}) \tag{8.6b}$$

Using the expansions (8.1) and (8.6), we can derive an asymptic expansion for large r for the curly bracket in Eq. (7.1). The expansion of $H^{(1)}(r)$ is more difficult. We first note the following: An integral of the form

$$\int_0^\infty dy_1 \cdots \int_0^\infty dy_n \prod_{j=}^{n} e^{-ry_j}\frac{f_n(y)}{\Pi(y_j + y_{j+1})}$$

falls off as r^{-m}, where

$m = n +$ number of y_j factors in the numerator of $f_n(y)$
 $-$ number of y_j factors in the denominator of $f_n(y)$
 $-$ number of $(y_j + y_{j+1})$ factors in the denominator. (8.7)

For example,

$$\int_0^\infty dy_1 \int_0^\infty dy_2\, e^{-r(y_1 y_2)}\frac{y_1 y_2}{(y_1+y_2)} \sim r^{-3}. \tag{8.8}$$

A study of the matrices in $\mathcal{N}(y;r)$ shows that

$$\int dy\ \mathrm{Tr}\mathcal{N}(y;r)^l \sim \begin{cases} r^{-2l-1} & \text{for } l \text{ odd}, \\ r^{-2l} & \text{for } l \text{ even}. \end{cases} \tag{8.9}$$

Thus, to obtain the expansion to $0(x^{-4})$ in the Introduction, we only need to study $\mathrm{Tr}\mathcal{N}(y;r)$ and $\mathrm{Tr}\mathcal{N}(y;r)^2$. The first term needs to be evaluated to the leading and next leading order in x^{-1}, while the second term needs to be evaluated only to the leading order. Now,

$$\mathrm{Tr}\mathcal{N}(y;r) = \mathrm{Tr}[M_0^{(2)}(y;r)F^{(1)}(y;r) + M_e^{(2)}(y;r)F^{(4)}(y;r)]. \tag{8.10}$$

We can use the expressions (7.14) for $F^{(1)}(y;r)$ and $F^{(4)}(y;r)$ to write out the right hand side of Eq. (8.10).

We will illustrate the method of deriving the asymptotic expansion by a typical term in Eq. (8.10), namely,

$$\mathrm{Tr}M_0^{(2)}(y;r)E^{(1)}(y;r) = \mathrm{Tr}\sum_{n=1}^{\infty}\pi^{-2n}M_0^{(2)}(y_1,y_2)E_{2n}^{(1)}(y_2,y_3,\cdots,y_{2n},y_1). \tag{8.11}$$

Expanding the $(y_1 + y_2 + 2i)^{-1}$ term in $M_0^{(2)}(y_1,y_2)$ [see Eq. (4.8b)], we have

$$M_0^{(2)}(y_1,y_2) = -y_2 e^{-ir}[(2i)^{-1} + \tfrac{1}{4}(y_1+y_2),0] + \text{ higher order terms}, \tag{8.12}$$

where the higher order terms contribute to $O(x^{-5})$ and higher. Next we multiply out the matrices in $E_{2n}^{(1)}(y_2 \to y_{2n}, y_1)$ [see Eq. (5.4)] keeping the terms only to leading and next leading order. This leads to

$$E^{(2)}(2\to 2n,1) = \left\{e^{-ir}\left[y_1(y_1+2i)^{-1}t_{2n-1}^{(1)}(2\to 2n) - \sum_{m=1}^{n-1} t_{2m-1}^{(1)}(2\to 2m)t_{2n-2m+1}^{(2)*}(2m+1\to 2n,1)\right],\ t_{2n}^{(3)}(2\to 2n,1)\right\} + \text{ higher order terms}, \tag{8.13}$$

where

$$t^{(1)}_{2k-1}(2\to 2k) = (-1)^{k+1}\Big(\prod_{j=2}^{k} y_{2j}\Big/\prod_{j=1}^{k-1} y_{2j+1}\Big)\overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} ---- \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2k-1}{\times}} ---- \overset{+}{\underset{2k}{\times}}, \quad k=1,2,..., \tag{8.14}$$

$$t^{(2)}_{2k+1}(1\to 2k+1) = (-1)^{k+1} y_{2k+1}\Big(\prod_{j=1}^{k} y_{2j}\Big/\prod_{j=0}^{k} y_{2j+1}\Big)\underset{1}{\times} ---- \overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} \cdots \overset{-}{\underset{2k-1}{\times}} ---- \overset{+}{\underset{2k}{\times}} ---- \underset{2k+1}{\times}, \quad k=1,2,..., \tag{8.15}$$

and

$$t^{(3)}_{2k}(1\to 2k+1) = (-1)^{k+1} y_{2k+1}\prod_{j=1}^{k}\left(\frac{y_{2j}}{y_{2j+1}}\right)\overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} ---- \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2k-1}{\times}} ---- \overset{+}{\underset{2k}{\times}} ---- \underset{2k+1}{\times}, \quad k=1,2,\ldots . \tag{8.16}$$

Substituting in Eq. (8.11) and making use of Eqs. (7.10) to (7.12), we see that

$$\text{Contribution from } \operatorname{Tr}M_0^{(2)}(y;r)E^{(1)}(y;r) \text{ to } H^{(1)}(r) = \sum_{n=1}^{\infty} 2\pi^{-2n}\operatorname{Re}\left[I^{(1)}_{2n}(r) + I^{(2)}_{2n}(r)\right] + O(r^{-5}), \tag{8.17}$$

where

$$I^{(1)}_{2n}(r) = (-1)^n e^{-2ir}\int_0^\infty dy_1\cdots\int_0^\infty dy_{2n}\prod_{j=1}^{2n} e^{-ry_j} y_1(y_1+2i)^{-1}\left[(2i)^{-1} + \tfrac14(y_1+y_2)\right]$$
$$\times\left\{\overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} ---- \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2n-1}{\times}} ---- \overset{+}{\underset{2n}{\times}}\right\} \tag{8.18}$$

and

$$I^{(2)}_{2n}(r) = (-1)^n e^{-2ir}\int_0^\infty dy_1\cdots\int_0^\infty dy_{2n}\prod_{j=1}^{2n} e^{-ry_j} y_1\left[(2i)^{-1} + \tfrac14(y_1+y_2)\right]$$
$$\times\sum_{m=1}^{n-1}\left\{\overset{+}{\underset{2}{\times}} ---- \overset{-}{\underset{3}{\times}} ---- \overset{+}{\underset{4}{\times}} \cdots \overset{-}{\underset{2m-1}{\times}} ---- \overset{+}{\underset{2m}{\times}}\right\}\left\{\underset{2n+1}{\times} ---- \overset{-}{\underset{2m+2}{\times}} ---- \overset{+}{\underset{2m+3}{\times}} \cdots \overset{+}{\underset{2n-1}{\times}} ---- \overset{-}{\underset{2n}{\times}} ---- \underset{1}{\times}\right\}. \tag{8.19}$$

Expanding the $(y_{2j+1} - 2i)$ and $(y_{2j} + 2i)$ factors in Eq. (8.18) and making use of Eq. (8.7), we see that

$$I^{(1)}_{2n}(r) = \frac14 e(r)\left(1 + \frac{1}{2i}\frac{d}{dr}\right)\left[\int_0^\infty dy_2\cdots\int_0^\infty dy_{2n}\prod_{j=2}^{2n} e^{-ry_j}\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1}\right]$$
$$+\frac{e^{-2ir}}{16}\int_0^\infty dy_1\cdots\int_0^\infty dy_{2n}\prod_{j=1}^{2n} e^{-ry_j} y_1(y_1+y_2)\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1} + O(r^{-5}). \tag{8.20}$$

Using the expansion (8.1) for $e(r)$ and scaling r in the integrals in Eq. (8.20) $(r\,y_j\to y_j)$, we can write

$$I^{(1)}_{2n}(r) = \frac{e^{-2ir}}{8r^2}\left[\left(\frac{1}{ir} + \frac{5}{2r^2}\right)\int_0^\infty dy_2\cdots\int_0^\infty dy_{2n}\prod_{j=2}^{2n} e^{-y_j}\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1}\right.$$
$$\left.+\frac{1}{2r^2}\int_0^\infty dy_2\cdots\int_0^\infty dy_{2n}\prod_{j=2}^{2n} e^{-y_j} y_2\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1}\right] + O(r^{-5}). \tag{8.21}$$

The integrals in Eq. (8.21) can be analyzed by techniques similar to those used in the analysis of $e^{(1)}(r)$ and $e^{(2)}(r)$. In particular, we can show that

$$\sum_{n=1}^{\infty}\pi^{-2n}\int_0^\infty dy_2\cdots\int_0^\infty dy_{2n}\prod_{j=2}^{2n} e^{-y_j}\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1} = 2i\pi^{-1}C_1^{(2)} = \frac{1}{2\pi} \tag{8.22}$$

and

$$\sum_{n=1}^{\infty}\pi^{-2n}\int_0^\infty dy_2\cdots\int_0^\infty dy_{2n}\prod_{j=2}^{2n} e^{-y_j}\prod_{j=2}^{2n-1}{}'(y_j+y_{j+1})^{-1} y_2 = 4\pi^{-1}C_2^{(2)} = \frac{3}{8\pi}. \tag{8.23}$$

Thus we have

$$\sum_{n=1}^{\infty}\pi^{-2n} I^{(1)}_{2n}(r) = \frac{e^{-2ir}}{16\pi r^3}\left(-i + \frac{23}{8r}\right) + O(r^{-5}). \tag{8.24}$$

The integral in $I^{(2)}_{2n}(r)$ factorizes into two factors, each of which can be analyzed in a similar manner. Adding the two contributions and using Eq. (8.17), we get

$$\text{contribution from } \operatorname{Tr}M_0^{(2)}(y;r)E^{(1)}(y;r) \text{ to } H^{(1)}(r) = -\frac{3\sin 2r}{64r^3} + \frac{63\cos 2r}{512\,r^4} + O(r^{-5}). \tag{8.25}$$

The remaining terms in $\operatorname{Tr}\mathcal{N}(y;r)$ and $\operatorname{Tr}\mathcal{N}(y;r)^2$ can be analyzed similarly. The calculation is tedious but straightforward on using the integral equation techniques developed in the analysis of $e^{(1)}(r)$. The final result is

$$H^{(1)}(r) = -\frac{\sin 2r}{32r^3} - \frac{1}{1024r^4} + \left(\frac{31}{8} - 119\cos 2r\right) + O(r^{-5}). \tag{8.26}$$

Substituting $H^{(1)}(r)$ and the expansion for the curly bracket in Eq. (7.1) and identifying $r\equiv x$ (since $k_f = 1$), we get the large x expansion in Eq. (1.3).

ACKNOWLEDGMENTS

We wish to acknowledge the many interesting and helpful discussion that we have had with Professor B.M. McCoy during the course of this calculation. This work was supported in part by National Science Foundation Grants Numbers PHY-76-15328 and DMR 77-07863-A01. We wish to thank Professor M. Sato, Dr. M. Jimbo, and Dr. T. Miwa for pointing out a calculational error in Eq. (1.3).

APPENDIX A: FACTORIZATION OF $h^{(2)}(r,g)$

In this Appendix we will carry out a factorization of $h^{(2)}(r,g)$ as $g\to 0$ in terms of $f_2(gr)$, $g_1(gr)$, and factors that are functions of r alone.

Along the branch cut $-$, we make a change of variables

$$k_j = -\sqrt{1-g^2} - iy_j, \quad g\leqslant y_j \leqslant \infty, \tag{A1a}$$

so that

$$[(k_j^2+\mu^2)(k_j^2+\mu^{*2})]^{1/2} \to (y_j - 2i)(y_j^2-g^2)^{1/2} + O(g^2), \quad \text{as } g\to 0. \tag{A1b}$$

Along the branch cut $+$

$$k_j = \sqrt{1-g^2} - iy_j, \quad g\leqslant y_j \leqslant \infty, \tag{A2a}$$

so that

$$[(k_j^2+\mu^2)(k_j^2+\mu^{*2})]^{1/2} \to (y_j + 2i)(y_j^2-g^2)^{1/2} + O(g^2), \quad \text{as } g\to 0. \tag{A2b}$$

Thus, using the obvious notation, as $g\to 0$

$$h^{(2)}(r,g) = \tfrac{1}{2}[h^{(2)}_{++}(r,g) + h^{(2)}_{+-}(r,g) + h^{(2)}_{-+}(r,g) + h^{(2)}_{--}(r,g)] + o(1), \tag{A3}$$

where

$$h^{(2)}_{++}(r,g) = 2e^{-2ir}\int_g^\infty dy_1 \int_g^\infty dy_2 \frac{e^{-r(y_1+y_2)}(y_1+2i)(y_1^2-g^2)^{1/2}}{(y_2+2i)(y_2^2-g^2)^{1/2}(y_1+y_2+2i)^2}, \tag{A4}$$

$$\begin{aligned} h^{(2)}_{+-}(r,g) = &\int_g^\infty dy_1 \int_g^\infty dy_2 \frac{e^{-r(y_1+y_2)}}{(y_1+2i)(y_1^2-g^2)^{1/2}(y_2-2i)(y_2^2-g^2)^{1/2}(y_1+y_2)^2} \\ &\times [(y_1+2i)^2(y_1^2-g^2) + (y_2-2i)^2(y_2^2-g^2)], \end{aligned} \tag{A5}$$

and

$$h^{(2)}_{-+}(r,g) = h^{(2)}_{+-}(r,g), \quad h^{(2)}_{--}(r,g) = h^{(2)*}_{++}(r,g). \tag{A6}$$

Therefore,

$$h^{(2)}(r,g) = \mathrm{Re}\, h^{(2)}_{++}(r,g) + h^{(2)}_{+-}(r,g), \tag{A7}$$

$$\begin{aligned} h^{(2)}_{++}(r,g) = 2\,e^{-2ir}&\int_g^\infty dy_1 \int_g^\infty dy_2 \frac{e^{-r(y_1+y_2)}(y_1^2-g^2)^{1/2}}{(y_2^2-g^2)^{1/2}} \\ &\times\left[\frac{(y_1+2i)}{(y_2+2i)(y_1+y_2+2i)^2} - \frac{1}{2i(y_1+2i)} + \frac{1}{2i(y_1+2i)}\right] \end{aligned} \tag{A8}$$

$$\begin{aligned} = &\frac{e^{-2ir}}{i}\int_g^\infty \frac{dy_1 e^{-ry_1}(y_1^2-g^2)^{1/2}}{y_1+2i}\int_g^\infty dy_2 \frac{e^{-ry_2}}{(y_2^2-g^2)^{1/2}} - \frac{e^{-2ir}}{i}\int_g^\infty dy_1 \int_g^\infty dy_2 \\ &\times \frac{e^{-r(y_1+y_2)}(y_1^2-g^2)^{1/2}y_2[(y_1+2i)^2 + 4y_2(y_1+2i) + y_2^2 + 4i(y_1+2i) + 2iy_2]}{(y_2^2-g^2)^{1/2}(y_1+2i)(y_1+y_2+2i)^2}. \end{aligned} \tag{A9}$$

In the second integral we can set $g=0$. In the first integral in the y_1 integration we can set $g=0$, and in the y_2 integration we rescale $y_2\to gy_2$ to get, for $g\to 0$,

$$\begin{aligned} h^{(2)}_{++}(r,g) = g_1(gr)&\frac{e^{-2ir}}{i}\int_0^\infty \frac{dy_1 e^{-ry_1}y_1}{y_1+2i} - \frac{e^{-2ir}}{i} \\ &\times \int_0^\infty dy_1 \int_0^\infty dy_2 \frac{e^{-r(y_1+y_2)}}{(y_1+y_2+2i)}\left[\frac{y_1+y_2}{y_1+y_2+2i} + \frac{y_1}{2(y_2+2i)} + \frac{y_2}{2(y_1+2i)}\right]. \end{aligned} \tag{A10}$$

Now

$$h^{(2)}_{+-}(r,g) = \int_g^\infty dy_1 \int_g^\infty dy_2 \, \frac{e^{-r(y_1+y_2)}(y_1^2-g^2)^{1/2}(y_1y_2-4)}{(y_2^2-g^2)^{1/2}(y_1+y_2)^2(y_2^2+4)}, \tag{A11}$$

$$= \int_g^\infty dy_1 \int_g^\infty dy_2 \, \frac{e^{-r(y_1+y_2)}(y_1^2-g^2)^{1/2}}{(y_2^2-g^2)^{1/2}(y_1+y_2)^2}\left(\frac{y_1y_2-4}{y_2^2+4}+1-1\right). \tag{A12}$$

Rescaling and setting $g=0$ wherever allowed, we have

$$h^{(2)}_{+-}(r,g) = -\int_1^\infty dy_1 \int_1^\infty dy_2 \, \frac{e^{-gr(y_1+y_2)}(y_1^2-1)^{1/2}}{(y_2^2-1)^{1/2}(y_1+y_2)^2} + \int_0^\infty dy_1 \int_0^\infty dy_2 \, \frac{e^{-r(y_1+y_2)}y_1}{(y_1+y_2)(y_2^2+4)}. \tag{A13}$$

We identify the first term with $f_2(gr)$, and substituting in Eq. (A7) we get the result, as $g \to 0$,

$$h^{(2)}(r,g) = 2f_2(gr) + g_1(gr)\,\mathrm{Re}\left(\frac{e^{-2ir}}{i}\int_0^\infty \frac{dx\, x\, e^{-rx}}{x+2i}\right) + 2\int_0^\infty dy_1 \int_0^\infty dy_2 \, \frac{e^{-r(y_1+y_2)}y_1}{(y_1+y_2)(y_2^2+4)}$$
$$-\,\mathrm{Re}\left\{\frac{e^{-2ir}}{i}\int_0^\infty dy_1 \int_0^\infty dy_2 \, \frac{e^{-r(y_1+y_2)}}{(y_1+y_2+2i)}\left[\frac{y_1+y_2}{(y_1+y_2+2i)} + \frac{y_1}{2(y_2+2i)} + \frac{y_2}{2(y_1+2i)}\right]\right\}. \tag{A14}$$

Thus, our objective of expressing the g dependence only through the functions $f_2(gr)$ and $g_1(gr)$ is achieved.

[1]M. Girardeau, J. Math. Phys. **1**, 516 (1960).
[2]T.D. Schultz, J. Math. Phys. **4**, 666 (1963).
[3]A. Lenard, J. Math. Phys. **5**, 930 (1964).
[4]H.G. Vaidya and C.A. Tracy, Phys. Rev. Lett. **42**, 3 (1979).
[5]E. Barouch and B.M. McCoy. Phys. Rev. A **3**, 786 (1971).
[6]H.G. Vaidya and C.A. Tracy, Physica (Utrecht) A **92**, 1 (1978).
[7]H.G. Vaidya and C.A. Tracy, Phys. Lett. A **68**, 378 (1978).
[8]T.T. Wu, B.M. McCoy, C.A. Tracy, and E. Barouch. Phys. Rev. B **13**, 316 (1976).
[9]B.M. McCoy, C.A. Tracy, and T.T. Wu, J. Math. Phys. **18**, 1058 (1977).

Transmission and reflection coefficients in the slowing down region

S. A. El Wakil[a] and M. H. Haggag
Physics Department, Faculty of Science, Mansoura University, Mansoura, Egypt
E. A. Saad
Atomic Energy Authority, Cairo, Egypt

(Received 10 May 1978; accepted for publication 17 April 1979)

The Green's function approach which was formulated by Case is used to obtain the reflection and transmission coefficients for slab geometry in the slowing down region using the Greuling–Goertzel approximation for the scattering kernel. We obtain a set of linear equations for the coefficients using the principles of invariance.

1. INTRODUCTION

In this paper, we derive a set of linear algebraic equations whose solution provides the energy dependent reflection and transmission coefficients for a slab under the assumption that the energy transfer kernel is given by the Greuling–Goertzel approximation.[1] The method of computation is based on the Green's function approach introduced by Case[2] and utilized by other authors[3–6] for different problems in transport theory. In addition to the Green's function technique, the invariant embedding principles of Ambarzumian[7] and Chandrasekhar[8] are used.

The plan of this paper is as follows. In Sec. 2 the time dependent integrodifferential Boltzmann equation is transformed to a time independent equation by Laplace transformation. The solution of this equation can be expressed in terms of the infinite medium Green's function by adding a surface distribution on the boundary of the medium.[2] This surface distribution can be obtained from the volume flux by $x \to x_s$ (where x_s is a point on the boundary of the medium). In Sec. 3 these equations of the surface distribution as expressed in terms of the infinite medium Green's function, are converted to a set of equations for the angular distribution on the surface of the medium, at the same time, we introduce the relation between the angular distribution on the surface and the reflection and transmission function of the medium. In Sec. 4 we calculate the infinite medium Green's function in the slowing down region by using the Fourier transformation and the technique quoted from Ref. 9. The calculated Green's function can now be used in the equations for the angular distribution. However, to facilitate numerical computations, it is rather convenient to transform these equations in matrix form by expanding the angular distribution and the Green's function in orthogonal sets with respect to lethargy u and angular dependence μ. We use the Laguerre polynomial for lethargy and half range Legendre polynomial for angular dependence. The resulting equations for the reflection and transmission matrices are expressible in terms of the matrix elements of the infinite medium Green's function and the incident intensity. If one uses for the incident density an azimuthally symmetric delta function, the equations reduce to a relation between the reflection and transmission matrices and the Green's matrix. The reflection and transmission coefficients for other incident distributions $S(u,\mu)$ can be obtained by integrating the result of delta distributions over $S(u,\mu)$. Finally in Sec. 6 we calculate the matrix elements of the Green's function by the contour integration method.[10]

2. FORMULATION OF THE BOLTZMANN EQUATION

The time dependent slowing down equation for plane geometry asssuming isotropic scattering is written as:

$$\left(\frac{1}{v}\frac{\partial}{\partial t}+\mu\frac{\partial}{\partial x}+1\right)\psi(x,u,\mu,t) = \frac{1}{2}\int_{-1}^{1}\int_{u_0}^{u} du'\,d\mu'\,c(u')f(u'\to u) \times\psi(x,u',\mu',t)+q(x,u,\mu,t), \tag{1}$$

where $\psi(x,u,\mu,t)$ is the angular flux in lethargy domain, i.e., $\psi(u)=\psi(E)|dE/du|$, and $f(u'\to u)$ is a rotationally invariant scattering kernel, $q(x,u,\mu,t)$ is a given neutron source, and $c(u)=\sigma(u)/\sigma_t(u)$ is the mean number of secondaries per collision. In the above equation the distance is measured in units of mean free path.

Define the Laplace transform of ψ and q by

$$\psi(x,u,\mu,s)=\int_0^\infty e^{-st}\,\psi(x,u,\mu,t)\,dt \tag{2}$$

and

$$q(x,u,\mu,s)=\int_0^\infty e^{-st}\,q(x,u,\mu,t)\,dt. \tag{3}$$

Then Eq. (1) becomes

$$\left(B+\mu\frac{\partial}{\partial x}\right)\psi(x,u,\mu,s) = \frac{1}{2}\int_{u_0}^{u} du'\,c(u')f(u'\to u) \times\int_{-1}^{1}\psi(x,u',\mu',s)\,d\mu'+q(x,u,\mu,s) \tag{4}$$

with $B=s/v+1$. The solution of this equation in terms of the infinite medium Green's function is[2]:

$$\psi(x,u,\mu,s)=\int_V\int G(x,u,\mu,s|\,x',u',\mu') \times q(x,u',\mu')\,dx'\,du'\,d\mu'$$

[a]Present address: Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061.

 0022-2488/79/112313-06$01.00

$$+\int_s\int G(x,u,\mu,s|\, x_s',u',\mu')\times\mu'\psi(x_s',u',\mu',s)\, dS'\, du'\, d\mu', \qquad (5)$$

where V is the volume in which the angular density is to be determined, S is the boundary of V, and x_s varies over the surface of the medium. The integral equation for the surface distribution $\psi(x_s,u,\mu,s)$ is obtained from Eq. (5) as $x\rightarrow x_s$ from within (without). The result is

$$\begin{pmatrix}\psi(x_s,u,\mu,s)\\ 0\end{pmatrix} = \int\int_V G(s_x,u,\mu,s/x',u',\mu')\times q(x',u',\mu',s)\, dx'\, du'\, d\mu' + \int_S\int G_{\mp}(x_s,u,\mu,s/x_s',\mu',u')\times\mu'\psi(x_s,u',\mu',s)\, du'\, d\mu'\, ds', \qquad (6)$$

where $G(x,u,\mu,s/x',u',\mu')$ is the infinite medium Green's function satisfying the equation

$$\left(B+\mu\frac{\partial}{\partial x}\right)G(x,u,\mu,s/x_0,u_0,\mu_0) = \delta(x-x_0)\,\delta(u-u_0)\,\delta(\mu-\mu_0) + \frac{1}{2}\int_{u_0}^{u} du'\, c(u')f(u'\rightarrow u)\times\int_{-1}^{1} G(x,u',\mu',s/x_0,u_0,\mu_0)\, d\mu' \qquad (7)$$

and the jump condition

$$\mu[G_+(x_s+\epsilon,u,\mu,s/x_0,u_0,\mu_0) - G_-(x_s-\epsilon,u,\mu,s/x_0,u_0,\mu_0)] = \delta(u-u_0)\,\delta(\mu-\mu_0). \qquad (8)$$

Equation (6) can be used to calculate the surface angular distributions once the infinite medium Green's function is known. We shall next reformulate this equation in order to obtain from the infinite medium Green's function, the angular distribution at the boundaries of the medium. These in turn will be related to the reflection and transmission coefficients. For the remainder of this paper we will restrict our calculations to slab geometry.

3. REFLECTION AND TRANSMISSION COEFFICIENTS OF SLAB GEOMETRY

Consider a homogeneous, source-free, slab of thickness L. For a given angular distribution $I^+(0,u,\mu,t)$ incident from the left-hand side of the slab, the reflected and transmission angular distributions $I^-(0,u,\mu,t)$ and $I^+(x,u,\mu,t)$ respectively can be related to the infinite medium Green's function by use of Eq. (6); namely,

$$I^-(0,u,\mu,s) = \int_0^u du'\int_0^1 G(\epsilon-0,u,\mu,s/0,u',\mu')\, I^+(0,u',\mu',s)\,\mu'\, d\mu' + \int_0^u du'\int_{-1}^0 G(\epsilon-0,u,\mu,s/0,u',\mu')I^-(0,u',\mu',s)\,\mu'\, d\mu' - \int_0^u du'\int_0^1 G(\epsilon-0,u,\mu,s/L,u',\mu')I^+(L,u',\mu',s)\,\mu\, d'\mu',$$

$$\mu<0, \qquad (9)$$

$$0 = \int_0^u du'\int_0^1 G(\epsilon+0,u,\mu,s/0,u',u')I^+(0,u',\mu',s)\,\mu'\, d\mu' + \int_0^u du'\int_{-1}^0 G(\epsilon+0,u,\mu,s/0,u',\mu')I^-(0,u',\mu',s)\,\mu'\, d\mu' - \int_0^u du'\int_0^1 G(\epsilon+0,u,\mu,s/L,u',\mu')I^+(L,u',\mu',s)\,\mu'\, d\mu'$$

$$\mu>0, \qquad (10)$$

at $x=0$, and

$$0 = \int_0^u du'\int_0^1 G(L+\epsilon,u,\mu,s/0,u',\mu')\, I^+(0,u',\mu',s)\,\mu'\, d\mu' + \int_0^u du'\int_{-1}^0 G(L,u,\mu,s/0,u',\mu')\, I^-(0,u',\mu',s)\,\mu'\, d\mu' - \int_0^u du'\int_0^1 G(\epsilon-0,u,\mu,s/L,u',\mu')I^+(L,u',\mu',s)\,\mu'\, d\mu',$$

$$\mu<0 \qquad (11)$$

$$I^+(L,u,\mu,s) = \int_0^u du'\int_0^1 G(L-\epsilon,u,\mu,s/0,u',\mu')\, I^+(0,u',\mu',s)\,\mu'\, d\mu' + \int_0^u du'\int_0^1 G(L-\epsilon,u,\mu,s/0,u',\mu')I^-(0,u',\mu',s)\,\mu'\, d\mu' - \int_0^u du'\int_0^1 G(\epsilon-0,u,\mu,s/0,u',\mu')I^+(L,u',\mu',s)\,\mu'\, d\mu',$$

$$\mu>0 \qquad (12)$$

at $x=L$.

The reflection function of the slab $R(x,u,\mu/u',\mu',t-t')$ is defined as the probability density of a neutron incident on the slab at time t' in the state (u',μ') to be reflected at time t in the state (u,μ). Likewise, the transmission function $T(x,u,\mu/u',t-t')$ is defined as the probability density for the same neutron to be transmitted by the slab in state (u,μ) at time t (see Fig. 1). These functions are related to the reflected and transmitted angular distribution by the equations[11]

$$I^-(0,u,\mu,t) = \int_0^t dt'\int_0^u du'\int_0^1 R(x,u,\mu/u',\mu',t-t')\times I^+(0,u',\mu',t')\, d\mu', \quad \mu<0 \qquad (13)$$

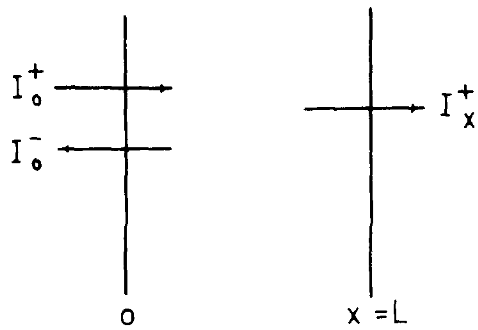


FIG. 1. The reflection and transmission distribution.

and

$$I^{+}(x,u,\mu,t)=\int_0^t dt'\int_0^u du'\int_0^1 T(x,u,\mu/u',\mu',t-t')\times I^{+}(0,u',\mu',t')\,d\mu',\qquad \mu>0. \tag{14}$$

By taking the Laplace transform of Eqs. (13) and (14) one gets:

$$I^{-}(0,u,\mu,s)=\int_0^u du'\int_0^1 R(x,u,\mu/u',\mu',s)\times I^{+}(0,u',\mu',s)\,d\mu',\qquad \mu<0 \tag{15}$$

and

$$I^{+}(x,u,\mu,s)=\int_0^u du'\int_0^1 T(x,u,\mu/,u',\mu',s)\times I^{+}(0,u',\mu',s)d\mu',\qquad \mu>0. \tag{16}$$

The Eqs. (9)–(12) and (13)–(16) represent two different sets of equations for the surface angular distributions. By substituting (15)–(16) into (9)–(12) the surface angular distribution can be eliminated and we get equations for the transmission and reflection coefficients in terms of the infinite medium Green's function. In Sec. 4 we will compute the infinite medium Green's function by using the Fourier transform method and the method adapted from Ref. 9.

4. CALCULATION OF THE GREEN FUNCTION

Let us define the Fourier transform of G by:

$$G(x-x_0,u,\mu;u_0,\mu_0;s)=\int_{-\infty}^{\infty} e^{-ik(x-x_0)}\,G(k,u,\mu;u_0,\mu_0,s)\,dk, \tag{17}$$

then Eq. (7) gives

$$[B-ik\mu]\,G(k,u,\mu,s/u_0,\mu_0)=\delta(u-u_0)\,\delta(\mu-\mu_0)+\tfrac{1}{2}\int_{u_0}^u du'\,c(u')f(u'\to u)\times\int_{-1}^1 G(k,u',\mu',s/u_0,\mu_0)\,d\mu'. \tag{18}$$

A simple calculation shows that

$$G(k,u,\mu,s/u_0,\mu_0)=\frac{\delta(u-u_0)\,\delta(\mu-\mu_0)}{B-ik\mu}+\frac{1}{2}\cdot\frac{1}{B-ik\mu}\cdot\frac{1}{\alpha(k,B)}\cdot\frac{1}{B-ik\mu_0}(B-ik\mu_0)\,\phi(k,u,s/u_0,\mu_0)-\delta(u-u_0) \tag{19}$$

with

$$\alpha(k,B)=\frac{1}{2}\int_{-1}^1\frac{d\mu}{B-ik\mu}=\frac{1}{ik}Q_0(B/ik) \tag{20}$$

and ϕ is the solution of

$$\phi(k,u,s/u_0,\mu_0)=\alpha(k,B)\int_{u_0}^u du'\,c(u')f(u'\to u)\times\phi(k,u',s/u_0,\mu_0)+\frac{\delta(u-u_0)}{B-ik\mu_0}. \tag{21}$$

In order to solve Eq. (21), the G–G approximation for $f(u'\to u)$ is used

$$f(u'\to u)=\frac{\xi}{\gamma^2}e^{(u'-u)/\gamma}+\left(1-\frac{\xi}{\gamma}\right)(u'-u) \tag{22}$$

and we use a method adapted from Ref. 9 to get

$$\begin{aligned}\phi(k,u,s/u_0,u_0)&=-\frac{c\xi/\gamma^2\alpha(k,B)}{[1-c(1-\xi/\gamma)\,\alpha(k,B)]^2}-\frac{1}{B-ik\mu_0}\exp\left(\frac{-1/\gamma[1-c\alpha(k,B)](u-u_0)}{1-c(1-\xi/\gamma)\,\alpha(k,B)}\right)\\&\quad+\frac{\delta(u-u_0)}{[1-c(1-\xi/\gamma)\,\alpha(k,B)]}\cdot\frac{1}{B-ik\mu_0}.\end{aligned} \tag{23}$$

Upon neglecting the contribution from the delta sources, the Green's function is obtained from Eqs. (23) and (19) as

$$\begin{aligned}G(x,u,\mu,s/u_0,\mu_0)&=\frac{1}{4\pi}\int_{-\infty}^{+\infty}dk\,e^{ik(x-x_0)}\frac{1}{B-ik\mu}\cdot\frac{1}{B-ik\mu_0}\cdot\frac{c\xi/\gamma^2}{[1-c(1-\xi/\gamma)\,\alpha(k,B)]^2}\\&\quad\times\exp\left(\frac{-1/\gamma[1-c\alpha(k,B)](u-u_0)}{1-c(1-\xi/\gamma)\,\alpha(k,B)}\right).\end{aligned} \tag{24}$$

In principle Eq. (24) may be put into Eqs. (9)–(12) to compute the angular distribution on the surface, then using (15) and (16) to calculate the reflection and transmission coeffiients. By expanding the angular densities and the Green's function to orthogonal sets with respect to both lethargy u and angular dependence μ we get a set of equations which give the reflection and transmission coefficients.

5. EXPANSION IN ORTHOGONAL POLYNOMIALS

Let us return to Eqs. (9)–(12), (15), and (16) and expand the angular distribution into Laguerre polynomials with respect to lethargy u

$$I^{+}(0,u,\mu,s)=M(u)\sum_{j=1}^N I_j^{\pm}(0,\mu,s)\,L_j(u)I_j(u), \tag{25}$$

$$I^{+}(L,u,\mu,s)=M(u)\sum_{j=1}^N I_j^{+}(L,\mu,s)\,L_j(u), \tag{26}$$

and

$$G(x,u,\mu,s/x',u',\mu')=M(u)\sum_{pl}L_p(u)\,L_l(u')\,G_{pl}(x,\mu,s/x',\mu'), \tag{27}$$

where $M(u)=u\,e^{-u}$, is the Maxwell's distribution. Inserting these expansions into Eqs. (9)–(12), (15), and (16) and using the orthogonality behavior of Laguerre polynomials, the resulting equations can be put in a matrix form as

$$\mathbf{I}^{-}(0,u,s)=\int_0^1\mathbf{G}(\epsilon-0,\mu,s/0,\mu')\,\mathbf{I}^{+}(0,\mu',s)\mu'\,d\mu'+\int_0^1\mathbf{G}(\epsilon-0,\mu,s/0,\mu')\,\mathbf{I}^{-}(0,\mu',s)\mu'\,d\mu'$$

$$-\int_0^1 \mathbf{G}(0,\mu,s/L,\mu')\,\mathbf{I}^+(L,\mu',s)\mu'\,d\mu', \qquad \mu<0, \tag{28}$$

$$\begin{aligned} 0 = &\int_0^1 \mathbf{G}(\epsilon+0,\mu,s/0,\mu')\,\mathbf{I}^+(0,\mu',s)\mu'\,d\mu' \\ &+\int_0^1 \mathbf{G}(\epsilon+0,\mu,s/0,\mu')\,\mathbf{I}^-(0,\mu',s)\mu'\,d\mu' \\ &-\int_0^1 \mathbf{G}(\epsilon+0,\mu,s/0,\mu')\,\mathbf{I}^+(L,\mu',s)\mu'\,d\mu', \qquad \mu>0 \end{aligned} \tag{29}$$

$$\begin{aligned} 0 = &\int_0^1 \mathbf{G}(L+\epsilon,\mu,s/0,\mu')\,\mathbf{I}^+(0,\mu',s)\mu'\,d\mu', \\ &+\int_{-1}^0 \mathbf{G}(L,\mu,s/0,\mu')\,\mathbf{I}^-(0,\mu',s)\mu'\,d\mu', \\ &-\int_0^1 \mathbf{G}(\epsilon-0,\mu,s/0,\mu')\,\mathbf{I}^+(L,\mu',s)\mu'\,d\mu', \qquad \mu<0 \end{aligned} \tag{30}$$

$$\begin{aligned} \mathbf{I}^+&(L,\mu,s) \\ &=\int_0^1 \mathbf{G}(L-\epsilon,\mu,s/0,\mu')\,\mathbf{I}^+(0,\mu',s)\mu'\,d\mu' \\ &\quad+\int_{-1}^0 \mathbf{G}(L-\epsilon,\mu,s/0,\mu')\,\mathbf{I}^-(0,\mu',s)\mu'\,d\mu', \\ &\quad-\int_0^1 \mathbf{G}(\epsilon-0,\mu,s/0,\mu')\,\mathbf{I}^+(L,\mu',s)\mu'\,d\mu', \qquad \mu>0 \end{aligned} \tag{31}$$

$$\mathbf{I}^-(0,\mu,s) = \int_0^1 \mathbf{R}(x,\mu/\mu',s)\,\mathbf{I}^+(0,\mu',s)\,d\mu', \qquad \mu<0 \tag{32}$$

and

$$\mathbf{I}^+(x,\mu,s) = \int_0^1 \mathbf{T}(x,\mu/\mu',s)\,d\mu', \qquad \mu>0, \tag{33}$$

where $\mathbf{I}^-$ and $\mathbf{I}^+$ are column matrices, and $\mathbf{G}(x,\mu;x',\mu')$ is a square matrix satisfying the equation

$$\begin{aligned} &\left(B+\mu\frac{\partial}{\partial x}\right)\mathbf{G}(x,\mu;x_0,\mu_0) \\ &\quad=\mathbf{c}\int_1^1 \mathbf{G}(x,\mu',s/x_0,\mu_0)\,d\mu' + \mathbf{I}\delta(x-x_0)\,\delta(\mu-\mu_0). \end{aligned} \tag{34}$$

This equation is obtained from Eq. (7) upon using the kernel (22) and the expansion (27). The matrix $\mathbf{c}$ is given by,

$$\begin{aligned} c_{jl} = &\left(1-\frac{\xi}{\gamma}\right)\delta_{jl} + \frac{\Gamma(2+j)\,\Gamma(2+1)}{2j!\,l!b\,a^2} \\ &\times\frac{\xi}{\gamma^2}F(-j,1,2,1/b)\,F(-l,2,2,1/b), \end{aligned} \tag{35}$$

where F is the hypergeometric function $b=1/\gamma$ and $a=(1-1/\gamma)$. Expanding $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{I}^\pm$ into half range Legendre polynomial for $\mu<0$ and $\mu>0$ gives

$$\mathbf{I}^-(0,\mu,s) = \tfrac{1}{4}\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^-(0,s)\,\mathbf{R}_{mn}(0,s)\,P_{m-1}(2\mu+1), \tag{36}$$

$$\mathbf{I}^+(L,\mu,s) = \tfrac{1}{4}\sum_{m,n} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{T}_{mn}(L,s)\,P_{m-1}(2\mu-1), \tag{37}$$

where $d_{mn} = \surd(2m-1)(2n-1)$.

Substituting in Eqs. (28)–(31), and using the orthogonality behavior of Legendre polynomial one gets

$$\begin{aligned} &\sum_n\left(\frac{2n-1}{2m-1}\right)^{1/2}\mathbf{I}_n^+(0,s)\,\mathbf{R}_{nm'}(0,s) \\ &\quad=2\sum_{j=1}^{\infty}(2j-1)\,\mathbf{I}_j^+(0,s)\,\overline{\mathbf{G}}_{m'j}^{-+}(0,s/0) \\ &\qquad+\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{R}_{mn}(0,s)\,\mathbf{G}_{nm'}^{--}(0,s/0) \\ &\qquad-\sum_{m,n} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{T}_{mn}(L,s)\,\overline{G}_{nm'}^{-+}(0,s/L), \qquad \mu<0, \end{aligned} \tag{38}$$

$$\begin{aligned} 0 = &\,2\sum_{j=1}^{\infty}(2j-1)\,\mathbf{I}_j^+(0,s)\,\overline{\mathbf{G}}_{m'j}^{-+}(L,s/0) \\ &+\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{R}_{mn}(0,s/0)\,\mathbf{G}_{nm'}^{--}(0,s/0) \\ &-\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{T}_{nm}(L,s)\,\mathbf{G}_{nm'}^{-+}(0,s/L), \\ &\mu>0, \end{aligned} \tag{39}$$

$$\begin{aligned} 0 = &\,2\sum_{j=1}^{\infty}(2j-1)\,\mathbf{I}_j^+(0,s)\,\mathbf{G}_{m'j}^{++}(L,s/0) \\ &+\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{R}_{mn}(0,s/0)\,\mathbf{G}_{m'n}^{+-}(L,s/0) \\ &-\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{T}_{mn}(L,s)\,\mathbf{G}_{m'n}^{++}(0,s/L), \qquad \mu<0, \end{aligned} \tag{40}$$

$$\begin{aligned} &\sum_n\left(\frac{2n-1}{2m'-1}\right)^{1/2}\mathbf{T}_{nm'}(L,s)\,\mathbf{I}_n^+(0,s) \\ &=2\sum_{j=1}(2j-1)\,\mathbf{I}_j^+(0,s)\,\mathbf{G}_{m'j}^{++}(L,s/0) \\ &\quad+\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{R}_{mn}(0,s)\,\mathbf{G}_{nm'}^{+-}(L,s/0) \\ &\quad-\sum_{m,n=1}^{\infty} d_{mn}\,\mathbf{I}_n^+(0,s)\,\mathbf{T}_{mn}(L,s)\,\mathbf{G}_{m'n}^{++}(0,s/0), \qquad \mu>0, \end{aligned} \tag{41}$$

where

$$\begin{aligned} G_{mn}^{\pm\pm}(x,s/x') = &\int_{0(-1)}^{1(0)} P_{n-1}(2\mu\mp1)\,d\mu \\ &\times\int_{0(-1)}^{1(0)} G(x,\mu,s/x',\mu') \\ &\times P_{m-1}(2\mu'\mp1)\,\mu'\,d\mu'. \end{aligned} \tag{42}$$

The set of Eqs. (38)–(41) can now be used to calculate the reflection and transmission matrices, in terms of $G^{\pm\pm}$ and the given distribution. However, if the incident distribution is chosen to be an azimuthally symmetric angular flux in the form

$$I^+ = \delta(\mu-\mu_0)\,\delta(u-u_0), \tag{43}$$

then

$$I^-(0,u,\mu,s) = R\,(0;u,\mu,/u_0,\mu_0,s), \qquad \mu<0 \tag{44}$$

and

$$I^{+}(x,u,\mu,s) = T(x,u,\mu,/u_0,\mu_0,s), \qquad \mu > 0. \tag{45}$$

These equations may be considered as boundary conditions on the surface of the medium.[12] If this representation of $I^{\pm}$ is used, and R and T are expanded into Laguerre polynomial with respect to lethargy u and half range Legendre polynomials for $\mu > 0$ and $\mu < 0$, i.e.,

$$R(0,u,\mu/u_0,\mu_0,s) = \tfrac{1}{2}M(u) \sum_{pj} \sum_{m,n} d_{mn} R_{pj}^{m,n}(0,s) P_{m-1}(2\mu - 1) \times P_{n-1}(2\mu_0 - 1) L_p(u) L_j(u_0) \tag{46}$$

and

$$T(L,u,\mu/u_0,\mu_0,s) = \tfrac{1}{2}M(u) \sum_{pj} \sum_{m,n} d_{mn} T_{pj}^{m,n}(L,s) P_{m-1}(2\mu - 1) \times P_{n-1}(2\mu_0 - 1) L_p(u) L_j(u_0). \tag{47}$$

Then Eqs. (28)–(31) and orthogonality of both the Laguerre and Legendre polynomials may be used to obtain:

$$\frac{1}{\sqrt{(2m-1)(2n-1)}} \mathbf{R}_{mn}(0,s) = G_{mn}^{-+}(0,s|0) + \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \cdot \mathbf{G}_{mk}^{--}(0,s|0) R_{kn}(0,s) - \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \mathbf{G}_{mk}^{-+}(0,s|L) \mathbf{T}_{kn}(L,s), \tag{48}$$

$$0 = \mathbf{G}_{mn}^{-+}(0,s|0) + \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \mathbf{G}_{mk}^{--}(0,s|0) \cdot \mathbf{R}_{kn}(0,s) - \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \mathbf{G}_{mk}^{-+}(0,s|L) \mathbf{T}_{kn}(L,s), \tag{49}$$

$$0 = \mathbf{G}_{mn}^{++}(L,s|0) + \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \mathbf{G}_{mk}^{+-}(L,s|0) \cdot \mathbf{R}_{kn}(0,s) - \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \mathbf{G}_{mk}^{++}(0,s|L) \mathbf{T}_{kn}(L,s), \tag{50}$$

and

$$\sqrt{(2m-1)(2n-1)}\, \mathbf{T}_{mn}(L,s) = \mathbf{G}_{mn}^{++}(L,s|0) + \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \times \mathbf{G}_{mk}^{+-}(L,s|0) \mathbf{R}_{kn}(0,s) - \sum_k \left(\frac{2k-1}{2n-1}\right)^{1/2} \times \mathbf{G}_{mk}^{++}(0,s|0) \mathbf{T}_{kn}(L,s), \tag{51}$$

where $G_{mn}^{\pm\pm}$ is given by Eq. (42) and may be evaluated by contour integration.

6. EVALUATION OF THE GREEN'S MATRIX $\mathbf{G}_{mn}^{\pm\pm}$

The evaluation of reflection and transmission coefficients has been reduced to the calculation of the Green's function matrix $\mathbf{G}_{mn}^{\pm\pm}$. In this section we will obtain an integral representation for $\mathbf{G}_{mn}^{\pm\pm}$ using the contour integral method. Substituting Eq. (24) into Eq. (27), transforming the variables μ,μ' in order to extend the range of integration to $[-1,1]$, and utilizing the recurrence relation of Legendre polynomial and using the equation

$$Q_n(x) = \int_{-1}^{1} \frac{P_n(v)}{z - v} dv \tag{52}$$

for the second kind Legendre polynomial, then Eq. (42) becomes

$$\frac{B\gamma^2}{2c} G_{\substack{n+1\\m+1}}^{\pm\pm} = \frac{1}{2\pi i} \int_{i\infty}^{-i\infty} \frac{H^{\pm\pm}(z)\,dz}{[(1 - vzQ_0(z)]^2} \times \int_0^\infty du \int_0^u du'\, M(u') L_p(u) L_j(u'). \tag{53}$$

In the above equation we put $Z = B/ik$, $v = (c/B)(1 - \xi/\gamma)$, and

$$H^{\pm\pm}(Z) = e^{-B/Z(x - x_0)} \times \exp\left(\frac{-1/\gamma[1 - cBZQ_0(Z)](u - u_0)}{1 - ZQ_0(Z)}\right) \cdot Q_m(2Z \mp 1) \cdot Q_n(2Z \mp 1) Z. \tag{54}$$

The function under the integral sign is analytic in the complex plane cut from -1 to 1 along the real axis except for poles of the second order at $\pm k_0$ where k_0 is the positive solution of

$$Q_0(k_0^{-1}) = \tanh^{-1} K_0 = k_0/v. \tag{55}$$

For $x > 0$, we evaluate the integration along the imaginary axis by extending the integration around the closed contour $C = \Sigma_{i=1}^{6} C_i$ as shown in Fig. 2. Let

$$\phi(x) = \frac{1}{2\pi i} \int_c \frac{H^{\pm\pm}(Z)\,dZ}{[1 - ZQ_0(Z)]^2} = \lim_{z \to k_0^{-1}} \frac{d}{dZ} (Z - k_0^{-1}) \left(\frac{H^{\pm\pm}}{[1 - ZQ_0(Z)]^2}\right) \tag{56}$$

which is, of course, the residue at $Z = k_0^{-1}$. Since this residue is equal to zero, therefore

$$\phi(x) = \frac{1}{2\pi i} \int_c \frac{H^{\pm\pm}(Z)\,dZ}{[1 - ZQ_0(Z)]^2} = 0. \tag{57}$$

Also the integral over C_6 approaches zero as the contour

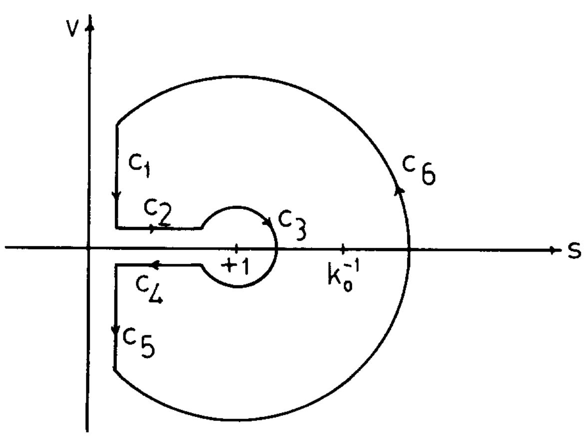


FIG. 2. Contour integration path.

aproaches infinity, accordingly

$$\int_{-i\infty}^{i\infty} \frac{H^{\pm\pm}(Z)\,dZ}{[1 - ZQ_0(Z)]^2} = -\int_{c_2+c_3+c_4} \frac{H^{\pm\pm}(Z)\,dZ}{[1 - ZQ_0(Z)]^2}. \tag{58}$$

Since $Q_n(Z)$ is discontinuous on the branch cut $(-1, +1)$, then

$Q_n(s \pm 0i) = Q_n(s) \mp i\frac{1}{2}\pi P_n(s)$, with $Z = s + iv$,

and

$$G^{\pm\pm}_{\substack{n+1\\ m+1}} = -\frac{2c\xi}{B\gamma^2} \int_0^\infty du \int_0^u du' M(u')\, L_p(u)\, L_j(u') \times \int_0^1 s ds\, e^{-B/s(x-x_0)}(H^+ + H^-), \tag{59}$$

where

$$H^{\pm} = \exp\left(\frac{-1/\gamma[1 - (cs/B)\,Q_0(s) \pm ics\pi/2B\,](u - u_0)}{1 - vsQ_0(s) \pm ivs\pi/2}\right) \cdot [Q_m(2s \mp 1) \mp i\pi/2P_m(2s + 1)] \times [Q_n(2s \mp 1) \mp i\pi/2P_n(2s \mp 1)]/[1 - vsQ_0(s) \pm i\pi vs/2]^2. \tag{60}$$

We have used the fact that the limit of the integral over C_3 is zero, since $Z = 1$ is a branch point but not a pole. For $x < 0$

$$G_{nm}(-x, x_0) = (-1)^{n+m} G_{nm}(x, x_0). \tag{61}$$

It is found that

$$G^{+-}_{nm}(x; x) = G^{-+}_{nm}(x; x). \tag{62}$$

7. CONCLUSION

The time dependent Boltzmann integrodifferential equation is transformed to an integral form by using the boundary source method developed by Case. This method permits one to solve the finite medium problems in terms of the infinite medium Green's function. The reflection and transmission functions are calculated in the slowing down region. The *G–G* approximation is used for the scattering, and the energy dependence is separated by expansion in Laguerre polynomial. The Green's function is calculated by Fourier transform and by a method adapted from Ref. 9, and the Fourier inverse is calculated by a contour integration. A set of algebraic equations for the reflection and transmission coefficients are obtained by use of the invariant embedding principles. If the incident distribution is assumed to be a delta function then the resultant set of equations is considerably simplified.

ACKNOWLEDGMENTS

The authors are much indebted to Professor P.F. Zweifel and Professor W. Greenberg for critical reading and for their comments.

[1] G. Goertzel and E. Greuling, Nucl. Sci. Eng. **7**, 236 (1964).
[2] K.M. Case, "On the boundary value problem of linear transport theory," Univ. of Michigan, Ann Arbor, Report (1967).
[3] K.M. Case, R. Zelazny, and M. Kanal, J. Math. Phys. **11**, 223 (1970).
[4] K.M. Case and R.D. Hazeltine, J. Math. Phys. **11**, 1126 (1970).
[5] M. Kanal, J. Math. Phys. **10**, 3042 (1970).
[6] S.A. El Wakil, ATKE **18**, 215 (1971); **17**, 241 (1971).
[7] V.A. Ambarzumian, C.R. Dokl. Akad Sci. USSR **38**, 229 (1943).
[8] S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960).
[9] M.M.R. Williams, *The Slowing Down and Thermalization of Neutrons* (North Holland, Amsterdam, 1966).
[10] K.M. Case, F. De Hoffmann, and G. Placzek, "Introduction to the theory of neutron diffusion 1," Los Alamos (1953).
[11] A. Mokel, Nucl. Sci. Eng. **29**, 44 (1967).
[12] A.F. Henry and P.C. Kalambokas, Nucl. Sci. Eng. **61**, 181 (1976).

Unitarity and renormalized 't Hooft identities

Seichi Naito

Department of Physics, Osaka City University, Sumiyoshiku, Osaka, Japan

(Received 1 November 1978)

With the help of the generalized Slavnov identities in a spontaneously broken gauge model, we compactly and explicitly prove renormalized 't Hooft identities which lead to the unitarity of the renormalized S matrix.

I. INTRODUCTION

Unitarity in a pure Yang–Mills field was first investigated by 't Hooft [1] by proving ('t Hooft) identities [i.e., (6.18) in Ref. 1] *on the perturbational ground*. Furthermore, his perturbational method has been applied [2] to a spontaneously broken gauge theory, without taking account of the mass renormalization effect. However, it is difficult in this approach to compactly and explicitly show the validity of 't Hooft identities to any order of coupling constants. In this paper, we give nonperturbational proof of *renormalized* 't Hooft identities. In order to show our method concretely, we consider a spontaneously broken gauge model

$$\mathscr{L}(x) = -\tfrac{1}{4}[f^{\gamma}_{\mu\nu}(x)]^2 - \tfrac{1}{2}[\partial^x_\mu \phi^{\gamma}(x) + g\epsilon_{\gamma\delta\epsilon} A^{\delta}_{\mu}(x) \times \phi^{\epsilon}(x)]^2 - \tfrac{1}{2}\mu^2[\phi^{\gamma}(x)]^2 - (\lambda/4)\{[\phi^{\gamma}(x)]^2\}^2,$$

with

$$f^{\gamma}_{\mu\nu}(x) \equiv \partial^x_\mu A^{\gamma}_{\nu}(x) - \partial^x_\nu A^{\gamma}_{\mu}(x) + g\epsilon_{\gamma\delta\epsilon} A^{\delta}_{\mu}(x) A^{\epsilon}_{\nu}(x),$$

which has been quantized [3] in a gauge-independent path-dependent formalism. [4–5]. [Hereafter, we refer various equations in Ref. 3 with a prefix R, i.e., (1.1) will be referred to as (R1.1).]

We give Sec. II as a preliminary to Sec. III; We introduce in Sec. II Faddeev–Popov ghosts' operators acting on linear space of covariant Green's functions, and we give partial differential equations satisfied by covariant Green's functions in the spontaneously broken case. These differential equations are used for obtaining generalized Slavnov identities and for constructing an effective Lagrangian which can be quantized in a canonical formalism. Furthermore, we give commutation and anticommutation relations among asymptotic quantum fields. With the help of results in Sec. II, we give in Sec. III a *nonperturbational proof* of *renormalized* 't Hooft identities (which lead to the unitarity of the S matrix); this proof is more compact and explicit than our previous proof [cf. (3.23) in Ref. 6]. In the Appendix, we give pole structures of two point Green's functions.

II. GENERALIZED SLAVNOV IDENTITIES AND ASYMPTOTIC FIELDS

In this section, various relations (which are in need for Sec. III) will be derived in our formalism. [3–6]

In order to introduce Faddeev–Popov ghosts' operators $\tilde{c}$, $\tilde{d}$ acting on a linear space of covariant Green's functions, we replace $\chi_{\alpha\beta}(x,y)$ in Ref. 3 with $i\tilde{c}^{\alpha}(x)\tilde{d}^{\beta}(y)$, so that we find [from (R4.14), (R5.6), (R4.19), and (R4.20)]

$$[\Box^x \tilde{c}^{\alpha}(x) + g\epsilon_{\alpha\beta\gamma} \tilde{A}^{\beta}_{\mu}(x)\partial^x_\mu \tilde{c}^{\gamma}(x) + (1/\xi)\, g\epsilon_{\alpha\beta\gamma} (\tilde{\varphi}^{\beta}(x) + v^{\beta}) g\epsilon_{\gamma\delta\epsilon} v^{\delta}\tilde{c}^{\epsilon}(x) + i\theta^{\alpha}_{d}(x)]|G) = 0, \tag{2.1}$$

$$(H|\theta^{\alpha}_{d}(x) = 0, \tag{2.2}$$

$$[\Box^x \tilde{d}^{\alpha}(x) + g\epsilon_{\alpha\beta\gamma}\, \partial^x_\mu (\tilde{A}^{\beta}_{\mu}(x)\, \tilde{d}^{\gamma}(x)) + (1/\xi) g\epsilon_{\alpha\beta\gamma} v^{\beta} g \times \epsilon_{\gamma\delta\epsilon} (\tilde{\varphi}^{\delta}(x) + v^{\delta})\tilde{d}^{\epsilon}(x) - i\theta^{\alpha}_{c}(x)]|G) = 0, \tag{2.3}$$

$$(H|\theta^{\alpha}_{c}(x) = 0, \tag{2.4}$$

$$\begin{aligned}[\Box^x \tilde{A}^{\alpha}_{\nu}(x) &- (1-\xi)\partial^x_\nu\partial^x_\mu \tilde{A}^{\alpha}_{\mu}(x) - g\epsilon_{\alpha\beta\gamma} v^{\beta} \\ &\times g\epsilon_{\gamma\zeta\eta} v^{\eta} \tilde{A}^{\zeta}(x) + g j^{\alpha}_{\nu}(x) \\ &- g\epsilon_{\alpha\beta\gamma} \tilde{\varphi}^{\beta}(x)\, \partial^x_\nu \tilde{\varphi}^{\gamma}(x) - g\epsilon_{\alpha\beta\gamma} \\ &\times \tilde{\varphi}^{\beta}(x) g\epsilon_{\gamma\zeta\eta} \tilde{A}^{\zeta}_{\nu}(x)\, \tilde{\varphi}^{\eta}(x) - g\epsilon_{\alpha\beta\gamma} (\partial^x_\nu \tilde{c}^{\beta}(x)) \\ &\times \tilde{d}^{\gamma}(x) - g\epsilon_{\alpha\beta\gamma} v^{\beta} g\epsilon_{\gamma\zeta\eta} \tilde{A}^{\zeta}_{\nu}(x)\, \tilde{\varphi}^{\eta}(x) - g\epsilon_{\alpha\beta\gamma} \\ &\times \tilde{\varphi}^{\beta}(x)\, g\epsilon_{\gamma\zeta\eta} v^{\eta} \tilde{A}^{\zeta}_{\nu}(x) - i\eta^{\alpha}_{\nu}(x)]|G) = 0,\end{aligned} \tag{2.5}$$

$$(H|\eta^{\alpha}_{\nu}(x) = 0, \tag{2.6}$$

and

$$\begin{aligned}[\Box^x \tilde{\varphi}^{\alpha}(x) &- (1/\xi)\, g\epsilon_{\alpha\beta\gamma} v^{\gamma} g\epsilon_{\beta\delta\epsilon} v^{\delta} \tilde{\varphi}^{\epsilon}(x) + g\epsilon_{\alpha\beta\gamma} \\ &\times (\partial^x_\mu \tilde{A}^{\beta}_{\mu}(x))\, \tilde{\varphi}^{\gamma}(x) + 2g\epsilon_{\alpha\beta\gamma} \tilde{A}^{\beta}_{\mu}(x)\, \partial^x_\mu \tilde{\varphi}^{\gamma}(x) + g \\ \times \epsilon_{\alpha\zeta\beta} \tilde{A}^{\zeta}_{\mu}(x)\, g\epsilon_{\beta\eta\delta} &\tilde{A}^{\eta}_{\mu}(x)\, \tilde{\varphi}^{\delta}(x) - \lambda (\tilde{\varphi}^{\gamma}(x))^2 \tilde{\varphi}^{\alpha}(x) \\ &+ g\epsilon_{\alpha\zeta\beta} \tilde{A}^{\zeta}_{\mu}(x)\, g\epsilon_{\beta\eta\delta} v^{\delta} \tilde{A}^{\eta}_{\mu}(x) - \mu^2 \tilde{\varphi}^{\alpha}(x) - \mu^2 v^{\alpha} \\ &- \lambda ((\tilde{\varphi}^{\gamma}(x) + v^{\gamma})^2(\tilde{\varphi}^{\alpha}(x) + v^{\alpha}) - (\tilde{\varphi}^{\gamma}(x))^2 \tilde{\varphi}^{\alpha}(x)) \\ - (1/\xi)\, g\epsilon_{\gamma\alpha\beta}\, g\epsilon_{\beta\delta\zeta} &v^{\delta}\tilde{c}^{\zeta}(x)\tilde{d}^{\gamma}(x) - i\zeta^{\alpha}(x)]|G) = 0,\end{aligned} \tag{2.7}$$

$$(H|\zeta^{\alpha}(x) = 0. \tag{2.8}$$

Bose-like operators $[\tilde{A}^{\alpha}_{\mu}(x), \eta^{\alpha}_{\mu}(x), \tilde{\varphi}^{\alpha}(x), \zeta^{\alpha}(x)]$ and Fermi-like operators $[\tilde{c}^{\alpha}(x), \theta^{\alpha}_{c}(x), \tilde{d}^{\alpha}(x)\, \theta^{\alpha}_{d}(x)]$ in (2.1)–(2.8) satisfy the following commutation or anticommutation relations:

$$[\tilde{A}^{\alpha}_{\mu}(x), \eta^{\beta}_{\nu}(y)] = \delta_{\alpha\beta}\, \delta_{\mu\nu}\delta^4(x-y),$$

$$[\tilde{\varphi}^{\alpha}(x), \zeta^{\beta}(y)] = \{\tilde{c}^{\alpha}(x), \theta^{\beta}_{c}(y)\} = \{\tilde{d}^{\alpha}(x), \theta^{\beta}_{d}(y)\} = \delta_{\alpha\beta}\, \delta^4(x-y), \tag{2.9}$$

and all other commutators or anticommutators vanish identically provided anticommutators are to be used between two Fermi-like operators and commutators are to be used otherwise). By comparing (2.1) and (2.3) with (R4.14) and (R5.6), $i\tilde{c}^{\alpha}(x)\tilde{d}^{\beta}(y)$ and $\chi_{\alpha\beta}(x,y)$ are found to be the same in the absence of external $\tilde{c}$ and $\tilde{d}$ fields, as it should be. This correspondence between $\chi_{\alpha\beta}(x,y)$ and $i\tilde{c}^{\alpha}(x)\tilde{d}^{\beta}(y)$ is very useful in deriving generalized Slavnov identities. In Ref. 3, generalized Ward–Takahashi identities (R4.18) have been

 0022-2488/79/112319-08$01.00

derived by the following calculations:

$$i\int d^4y\,[\eta_\mu^\delta(y)(\delta|D_\mu^y|\gamma)+\zeta^\delta(y)\,g\epsilon_{\delta\epsilon\gamma}\,\tilde\phi^\epsilon(y)]\,\chi_{\alpha\gamma}(x,y)|G) \tag{2.10}$$

$$\begin{aligned}&= -i\int d^4y\,\chi_{\alpha\gamma}(x,y)[(\gamma|D_\mu^y|\delta)\,\eta_\mu^\delta(y)+g\epsilon_{\gamma\epsilon\delta}\\&\times\tilde\phi^\epsilon(y)\,\zeta^\delta(y)]|G)+i\int d^4y\,[-\chi_{\alpha\eta}(x,y)\\&\times(\eta|\,D_\mu^z|\delta)\,g\epsilon_{\delta\gamma\phi}(\partial_\mu^z\,\chi_{\phi\gamma}(z,y))_{z=y}+\chi_{\alpha\eta}(x,y)\\&\times g\epsilon_{\eta\gamma\delta}((\delta|D_\mu^z|\phi)\,\partial_\mu^z\,\chi_{\phi\gamma}(z,y))_{z=y}]|G)\\&-i\int d^4y\,g\epsilon_{\gamma\epsilon\delta}\,\tilde\phi^\epsilon(y)\,\chi_{\alpha\eta}(x,y)(-1/\xi)\,g\epsilon_{\eta\delta\zeta}\\&\times g\epsilon_{\zeta\beta\phi}v^\beta\chi_{\phi\gamma}(y,y)|G)\end{aligned} \tag{2.11}$$

$$\begin{aligned}&=i\int d^4y\,\chi_{\alpha\eta}(x,y)\,g\epsilon_{\eta\gamma\delta}((\delta|E(z)|\phi)\,\chi_{\phi\gamma}(z,y))_{z=y}|G)\\&-\int d^4y\,[(\delta_{\gamma\delta}\,\Box^y-g\epsilon_{\gamma\beta\delta}\,\partial_\mu^y\,\tilde A_\mu^\beta(y)-g\epsilon_{\gamma\beta\delta}\,\tilde A_\mu^\beta(y)\\&\times\partial_\mu^y-(1/\xi)\,g\epsilon_{\gamma\epsilon\theta}\,\tilde\phi^\epsilon(y)\,g\epsilon_{\theta\delta\zeta}\,v^\zeta)\,\chi_{\alpha\gamma}(x,y)]\\&\times(\xi\partial_\nu^y\tilde A_\nu^\delta(y)+g\epsilon_{\delta\beta\gamma}\,v^\beta\,\tilde\varphi^\gamma(y))|G)\end{aligned} \tag{2.12}$$

$$=(\xi\partial_\nu^x\,\tilde A_\nu^\alpha(x)+g\epsilon_{\alpha\beta\gamma}\,v^\beta\,\tilde\varphi^\gamma(x))|G), \tag{2.13}$$

where we have used

$$\tilde\phi^\alpha\equiv\tilde\varphi^\alpha(x)+v^\alpha.$$

In transforming (2.10) [(2.12)] into (2.11) [(2.13)], we have used (R5.1)–(R5.3) [(R4.14) and (R5.6)]. On the other hand, in transforming (2.11) into (2.12), we have only used (R4.19), (R4.20), (R5.4) and (R5.5), which are valid *even if* we replace $\chi_{\alpha\beta}(x,y)$ with $i\tilde c^\alpha(x)\tilde d^\beta(y)$. Therefore, we have the equality between (2.11)′ and (2.12)′, where the mathematical expression (2.11)′ [(2.12)′] is to be obtained from (2.11) [(2.12)] by replacing $\chi_{\alpha\beta}(x,y)$ with $i\tilde c^\alpha(x)\tilde d^\beta(y)$.

Thus, we find

$$\begin{aligned}&(\xi\,\partial_\mu^x\,\tilde A_\mu^\alpha(x)+g\epsilon_{\alpha\beta\gamma}\,v^\beta\,\tilde\varphi^\gamma(x))|G)\\&\quad=\int d^4y[\eta_\mu^\delta(y)(\delta\,|D_\mu^y|\gamma)+\zeta^\delta(y)\,g\epsilon_{\delta\beta\gamma}\tilde\phi^\beta(y)]\\&\quad\times\tilde d^\gamma(y)\tilde c^\alpha(x)|G)+\int d^4y\,\theta_c^\delta(y)(\xi\partial_\mu^y\,\tilde A_\mu^\delta(y)\\&\quad+g\epsilon_{\delta\beta\gamma}v^\beta\tilde\varphi^\gamma(y))\tilde c^\alpha(x)|G)+\tfrac12\int d^4y\theta_d^\delta(y)\\&\quad\times g\epsilon_{\delta\beta\gamma}\tilde d^\beta(y)\tilde d^\gamma(y)\tilde c^\alpha(x)|G)\\&\equiv\int d^4y\,[\eta_\mu^\delta(y)\,\tilde\Sigma_\mu^\delta(y)+\zeta^\delta(y)\,\tilde\Lambda^\delta(y)\\&\quad+\theta_c^\delta(y)\,\tilde\Gamma^\delta(y)+\theta_d^\delta(y)\,\tilde\Delta^\delta(y)]\tilde c^\alpha(x)|G),\end{aligned} \tag{2.14}$$

where we have used (2.1), (2.3), (2.9), and (R4.11). As emphasized in Ref. 3, it should be noticed that generalized Slavnov identities (2.14) have been proved by using differential equations (2.1)–(2.8), and we did not need to use boundary conditions in the spontaneously broken case [i.e., (R4.21)]. In Sec. 3 we shall extensively use generalized Slavnov identities (2.14) in proving renormalized 't Hooft identities which lead to the unitarity of the S matrix. In that proof, the following separability property of (2.14) plays essential roles: The variables (x,α) which appear in the l.h.s. of (2.14) appear only in the term $\tilde c^\alpha(x)$. Furthermore, the operators $\tilde\Sigma_\mu^\delta(y)$, $\tilde\Lambda^\delta(y)$, $\tilde\Gamma^\delta(y)$ and $\tilde\Delta^\delta(y)$ depend only on the attached indices (μ,δ) and the space–time variable y.

In order to obtain S matrix elements by using Lehmann–Symanzik–Zimmerman (LSZ) formalism,[7] it is convenient to introduce quantum field opertors $[A_\mu^\alpha,\phi^\beta,c^\gamma,d^\delta]$ which are obtained by quantizing the Langrangian system,

$$\begin{aligned}\mathscr L(x)=&-\tfrac14[f_{\mu\nu}^\gamma(x)]^2-\tfrac12[\partial_\mu^x\phi^\gamma(x)+g\epsilon_{\gamma\delta\epsilon}A_\mu^\delta(x)\\&\phi^\epsilon(x)]^2-\tfrac12\mu^2[\phi^\gamma(x)]^2-\tfrac14\lambda\,[(\phi^\gamma(x))^2]^2\\&-\frac{1}{2\xi}\,[\xi\,\partial_\mu^x\,A_\mu^\gamma(x)+g\epsilon_{\gamma\delta\epsilon}v^\delta\phi^\epsilon(x)]^2\\&-\partial_\mu^x d^\gamma(x)\partial_\mu^x c^\gamma(x)\\&+d^\gamma(x)\,[g\epsilon_{\gamma\delta\epsilon}\,A_\mu^\delta(x)\partial_\mu^x+(1/\xi)\,g\epsilon_{\gamma\delta\eta}\phi^\delta(x)\\&\times g\epsilon_{\eta\theta\epsilon}v^\theta\,]c^\epsilon(x),\end{aligned} \tag{2.15}$$

in the canonical formalism. The quantum system corresponding to (2.15) is described by the Bose fields $(A_\mu^\alpha,\phi^\alpha)$ and the Fermi fields (c^α,d^α) together with their canonical conjugates $(\Pi_{A_\mu}^\alpha,\Pi_\phi^\alpha)$ and $(\Pi_c^\alpha,\Pi_d^\alpha)$. In order to investigate our quantum system, it is more convenient to treat the (covariant) T^* products (among $A_\mu^\alpha,\Pi_{A_\nu}^\beta,\phi^\gamma,\Pi_\phi^\delta,c^\epsilon,\Pi_c^\zeta,d^\eta,\Pi_d^\theta$) than the (noncovariant) time-ordered products. (We shall not discuss anti-time-ordered products in detail, since they can be treated in similar fashion to time-ordered products). The general definitions of T^* products in terms of T products may be easily imagined from the following special examples:

$$\begin{aligned}&\langle 0|T^*[\Pi_{A_4}^\alpha(x)\Pi_{A_4}^\beta(y)]|0\rangle\\&\quad\equiv\langle 0|T\,[\Pi_{A_4}^\alpha(x)\,\Pi_{A_4}^\beta(y)]|0\rangle-i\xi\delta_{\alpha\beta}\,\delta^4(x-y),\end{aligned} \tag{2.16}$$

where we have taken account of an equal-time commutation relation,

$$[A_4^\alpha(x),\Pi_{A_4}^\beta(y)]_{x_0=y_0}=\delta_{\alpha\beta}\,\delta^3(\mathbf x-\mathbf y), \tag{2.17}$$

and an equation of motion,

$$\Pi_{A_4}^\beta(y)=\xi\partial_4^x\,A_4^\beta(x)+\xi\partial_i^x\,A_i^\beta(x)+g\epsilon_{\beta\gamma\delta}v^\gamma\phi^\delta(x). \tag{2.18}$$

Since T^* products are easily proved[3] to satisfy the same differential equations as (2.1)–(2.8), we have identities

$$\begin{aligned}&(H\,|\,\tilde A_\mu^\alpha(s)\,\tilde\Pi_{A_\nu}^\beta(t)\tilde\phi^\gamma(u)\,\tilde\Pi_\phi^\delta(v)\tilde c^\epsilon(w)\tilde\Pi_c^\zeta(x)\\&\quad\times\tilde d^\eta(y)\,\tilde\Pi_d^\theta(z)\cdots|G)_-\\&\quad\equiv\langle 0|T^*\,[A_\mu^\alpha(s)\,\Pi_{A_\nu}^\beta(t)\phi^\gamma(u)\,\Pi_\phi^\delta(v)c^\epsilon(w)\,\Pi_c^\zeta(x)\\&\quad\times d^\eta(y)\,\Pi_d^\theta(x)\cdots]|0\rangle.\end{aligned} \tag{2.19}$$

Operators $\tilde\Pi$'s in the l.h.s. of (2.19) have been defined by

$$\tilde\Pi_{A_4}^\beta(x)\equiv\xi\partial_4^x\,\tilde A_4^\beta(x)+\xi\partial_i^x\,\tilde A_i^\beta(x)+g\epsilon_{\beta\gamma\delta}v^\gamma\tilde\phi^\delta(x),\quad\text{etc.,}$$

corresponding to quantum equations (2.18), etc. As well as identities (2.19), we can derive other identities between $|G)_+$ and anti-T^* products, where states $|G)_\pm$ are obtained from (R4.22) and (R4.23) by specifying Feynman propagators $D_{\mu\nu}^{\alpha\beta}$ and $D^{\alpha\beta}$: For example,

$$D_{\mu\nu}^{\pm 11}(x-y)=D_{\mu\nu}^{\pm 22}(x-y)$$

$$= \mp \frac{1}{(2\pi)^4 i} \int d^4p \, e^{ip(x-y)} (\mu | g_{\pm}(p;\xi) | \nu), \qquad (2.20)$$

with

$$(\mu | g_{\pm}(p;\xi) | \nu) \equiv \frac{\delta_{\mu\nu}}{p^2 + v^2 g^2 \pm i\epsilon} + \left(\frac{1}{\xi} - 1\right) \frac{p_\mu p_\nu}{(p^2 + v^2 g^2/\xi \pm i\epsilon)(p^2 + v^2 g^2 \pm i\epsilon)}.$$

Finally, we shall give commutation and anticommutation relations among asymptotic fields, by using the pole structures of the covariant two-point Green's functions in the special gauge $Z_A^0(\xi_0) = 1/\xi_0$ [cf. (A1) in the Appendix]. Since the renormalization S matrix elements among physical particles have been proved [3] to be independent of ξ, unitarity proof in the special gauge is sufficient for physical purposes. For simplicity of notations, various constants $Z_A^0(\xi_0), \cdots$ in the Appendix will be referred to as $Z_A^0, \cdots$.

Assuming LSZ asymptotic conditions

$$\psi(x) \to \psi(x; {}^{\text{out}}_{\text{in}}) \quad \text{as} \quad x_0 \to {}^{\infty}_{-\infty} \qquad (2.21)$$

in the sense of weak convergence, we shall derive commutation or anticommutation relations between creation operators $[c_\psi]$ and annihilation operators $[a_\psi]$ for asymptotic fields ψ, where ψ denotes any one of A_μ^α, $\varphi^\alpha (\equiv \phi^\alpha - v^\alpha)$, c^α and d^α; First, c_ψ's and a_ψ's are given by

$$\Psi^3(x;) = \int \frac{d^3k}{(2\pi)^3 \sqrt{2k_0}} [e^{-ikx} c_\Psi^0(k;) + e^{ikx} a_\Psi^0(k;)], \qquad (2.22)$$

$$\varphi^3(x;) = \int \frac{d^3q}{(2\pi)^3 \sqrt{2q_0}} [e^{-iqx} c_\varphi^0(q;) + e^{iqx} a_\varphi^0(q;)], \qquad (2.23)$$

$$\begin{aligned} A_\mu^{\pm}(x;) &= \int \frac{d^3p}{(2\pi)^3 \sqrt{2p_0}} [e^{-ipx} c_{A_\mu}^{\mp}(p;) + e^{ipx} a_{A_\mu}^{\pm}(p;)], \\ &\quad + \int \frac{d^3l}{(2\pi)^3 \sqrt{2l_0}} l_\mu [e^{-ilx} c_\chi^{\mp}(l;) + e^{ilx} a_\chi^{\pm}(l;)], \\ &\equiv V_\mu^{\pm}(x;) + \partial_\mu^x \chi^{\pm}(x;), \end{aligned} \qquad (2.24)$$

and

$$\Phi^{\pm}(x;) = \int \frac{d^3l}{(2\pi)^3 \sqrt{2l_0}} [e^{-ilx} c_\Phi^{\mp}(l;) + e^{ilx} a_\Phi^{\pm}(l;)], \qquad (2.25)$$

where the symbol $\Psi(\Phi)$ denotes any one of A_μ, c, and d (φ, c, and d) and energies are given by

$$k_0 = (\mathbf{k}^2)^{1/2}, \quad q_0 = (\mathbf{q}^2 + \mu_H^2)^{1/2},$$
$$p_0 = (\mathbf{p}^2 + M^2)^{1/2}, \quad l_0 = (\mathbf{l}^2 + m^2)^{1/2}. \qquad (2.26)$$

Then applying the Greenberg–Robinson theorem [8] to clothed two-point Green's functions (A1)–(A8) in the Appendix, we obtain the following commutation and anticommutation relations among a_ψ's and c_ψ's:

$$\begin{aligned}
[a_{A_\mu}^0(k;), c_{A_\nu}^0(k';)] &= Z_A^0 \, \delta_{\mu\nu} (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}'), \\
[a_\varphi^0(q;), c_\varphi^0(q';)] &= Z_\varphi^0 \, (2\pi)^3 \delta^3(\mathbf{q} - \mathbf{q}'), \\
\{a_c^0(k;), c_d^0(k';)\} &= -\{a_d^0(k), c_c^0(k')\} \\
&= Z^0 (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}'), \\
[a_{A_\mu}^{\pm}(p;), c_{A_\nu}^{\pm}(p';)] &= Z_A^c (\delta_{\mu\nu} + p_\mu p_\nu / M^2) \\
&\quad \times (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'), \\
\{a_c^{\pm}(l;), c_d^{\pm}(l';)\} &= -\{a_d^{\pm}(l;), c_c^{\pm}(l';)\} \\
&= Z^c (2\pi)^3 \delta^3(\mathbf{l} - \mathbf{l}'), \\
[a_\chi^{\pm}(l;), c_\chi^{\pm}(l';)] &= \alpha (2\pi)^3 \delta^3(\mathbf{l} - \mathbf{l}'), \\
[a_\chi^{\pm}(l;), c_\varphi^{\pm}(l';)] &= [a_\varphi^{\pm}(l;), c_\chi^{\pm}(l';)] \\
&= \mp \beta (2\pi)^3 \delta^3(\mathbf{l} - \mathbf{l}'), \\
[a_\varphi^{\pm}(l;), c_\varphi^{\pm}(l';)] &= \gamma (2\pi)^3 \delta^3(\mathbf{l} - \mathbf{l}'),
\end{aligned} \qquad (2.27)$$

and all other commutators or anticommutators vanish identically (provided anticommutators are to be used between two Fermi fields and commutators are to be used otherwise). In contrast to (2.21), we have not explicitly written "in" or "out" in (2.22)–(2.27), and we shall use these abbreviated expressions, unless we want to explicitly distinguish "in" from "out."

III. UNITARITY AND PROOF OF RENORMALIZED 't HOOFT IDENTITIES

S matrix elements are defined by

$$S_{\beta\alpha} = \langle \beta \, \text{out} | \alpha \, \text{in} \rangle, \qquad (3.1)$$

where "in" and "out" are, respectively, "in" and "out" states in which there exist nonphysical particles as well as physical particles. Then unitarity relations are expressed in the form

$$\langle \alpha \, \text{out} | \beta \, \text{out} \rangle = \sum_n \langle \alpha \, \text{out} | P_n \, (\text{in}) | \beta \, \text{out} \rangle, \qquad (3.2)$$

where P_n (in) is the projection operator onto the subspace containing n-particle "in" states. If P_1 (in) is obtained in the form

$$P_1 \, (\text{in}) = \sum_{\gamma\delta} c_\gamma \, (;\text{in}) |0\rangle \, f_{\gamma\delta} \langle 0 | a_\delta \, (;\text{in}), \qquad (3.3)$$

P_n (in) is given by

$$\begin{aligned} P_n \, (\text{in}) &= \frac{1}{n!} \sum_{\gamma_n \delta_n} \cdots \sum_{\gamma_2 \delta_2} \sum_{\gamma_1 \delta_1} c_{\gamma_n}(;\text{in}) \cdots c_{\gamma_2}(;\text{in}) c_{\gamma_1}(;\text{in}) |0\rangle \\ &\quad \times f_{\gamma_n \delta_n} \cdots f_{\gamma_2 \delta_2} f_{\gamma_1 \delta_1} \langle 0 | a_{\delta_1}(;\text{in}) \\ &\quad \times a_{\delta_2}(;\text{in}) \cdots a_{\delta_n}(;\text{in}). \end{aligned} \qquad (3.4)$$

An explicit expression for (3.3) is found by using (2.27) to be

$$\begin{aligned} P_1 \, (\text{in}) = \Bigg[&\iint \frac{d^3k}{(2\pi)^3} c_{A_\mu}^0(k;) |0\rangle \\ &\times \frac{(\delta_{\mu\nu} - k_\mu k_\nu / \mathbf{k}^2)(1 - \delta_{\mu 4})(1 - \delta_{\nu 4})}{Z_A^0} \\ &\times \langle 0 | a_{A_\nu}^0(k;) \\ &+ \int \frac{d^3q}{(2\pi)^3} c_\varphi^0(q;) |0\rangle \\ &\times \frac{1}{Z_\varphi^0} \langle 0 | a_\varphi^0(q;) + \int \frac{d^3p}{(2\pi)^3} c_{A_\mu}^{\pm}(p;) |0\rangle \end{aligned}$$

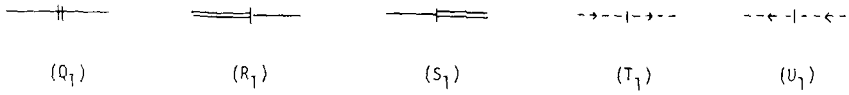


FIG. 1. Graphical representation of Q_1, R_1, S_1, T_1, and U_1 defined in (3.5).

$$\times \frac{1}{Z_A^c}\langle 0|a^{\pm}_{A_\mu}(p;) + \int \frac{d^3 l}{(2\pi)^3} c_S^{\pm}(l;)|0\rangle$$

$$\times \frac{-\gamma}{(-\xi_0 m^2\beta + vg\gamma)^2}\langle 0|a_S^{\pm}(l;)\quad\Big]$$

$$\Big[+\int \frac{d^3 k}{(2\pi)^3} c_S^0(k;)|0\rangle \frac{ik_\nu^*}{2\mathbf{k}^2}\langle 0|a_{A_\nu}(k;)$$

$$+\int \frac{d^3 l}{(2\pi)^3} c_S^{\pm}(l;)|0\rangle \frac{\mp i}{-\xi_0 m^2\beta + vg\gamma}$$

$$\times \langle 0|a_\varphi^{\pm}(l;)\quad\Big]$$

$$\Big[+\int \frac{d^3 k}{(2\pi)^3} c_{A_\mu}^0(k;)|0\rangle \frac{-ik_\mu^*}{2\mathbf{k}^2}\langle 0|a_S^0(k;)$$

$$+\int \frac{d^3 l}{(2\pi)^3} c_\varphi^{\pm}(l;)|0\rangle \frac{\pm i}{-\xi_0 m^2\beta + vg\gamma}$$

$$\times \langle 0|a_S^{\pm}(l;)\quad\Big]$$

$$\Big[+\int \frac{d^3 k}{(2\pi)^3} c_d^0(k;)|0\rangle \frac{1}{Z^0}\langle 0|a_c^0(k;)$$

$$+\int \frac{d^3 l}{(2\pi)^3} c_d^{\pm}(l;)|0\rangle \frac{1}{Z^c}\langle 0|a_c^{\pm}(l;)\Big]$$

$$\Big[+\int \frac{d^3 k}{(2\pi)^3} c_c^0(k;)|0\rangle \frac{-1}{Z^0}\langle 0|a_d^0(k;)$$

$$+\int \frac{d^3 l}{(2\pi)^3} c_c^{\pm}(l;)|0\rangle \frac{-1}{Z^c}\langle 0|a_d^{\pm}(l;)\Big]$$

$$\equiv Q_1 + R_1 + S_1 + T_1 + U_1. \tag{3.5}$$

Operators c_S^0 and a_S^0 [$c_S^{\pm}$ and $a_S^{\pm}$] in (3.5) are defined by (2.22) [(2.25)], provided $S^3(x;)\equiv\xi_0\partial_\mu^x A_\mu^3(x;)$ [$S^{\pm}(x;)\equiv\xi_0\,\partial_\mu^x A_\mu^{\pm}(x;) \mp ivg\,\varphi^{\pm}(x;)$]. In deriving (3.5), we have used (anti) commutation relations (2.27) and the relation (A10). As shown in Ref. 3, physical particles in our quantum system are massless transverse vector particles, massive particles of spin one and massive Higgs' scalar particle. Therefore, Q_1 represents (except for the last term in Q_1) the projection operator onto the subspace containing one physical particle.

For later convenience, we graphically represent Q_1, R_1, S_1, T_1, and U_1 by Fig. 1 [(Q_1), (R_1), (S_1), (T_1), and (U_1), respectively]. Furthermore, we represent the sum of Q_1, R_1, and S_1 by Fig. 2.

In the unitarity relations (3.2), there appear terms

$$\langle 0|a_{\alpha_j}(;\text{out})\cdots a_{\alpha_1}(;\text{out})\,c_{\gamma_n}(;\text{in})\cdots c_{\gamma_1}(;\text{in})|0\rangle$$
$$\times\langle 0|a_{\delta_1}(;\text{in})\cdots a_{\delta_n}(;\text{in})c_{\beta_1}(;\text{out})\cdots c_{\beta_k}(;\text{out})|0\rangle. \tag{3.6}$$

Hereafter, quantities (3.6) will be graphically represented by Fig. 3.

With the help of (3.3) (Figs. 1,2) and (3.6) (Fig. 3), we can represent $n!\langle 0|a_{\alpha_j}(;\text{out})\cdots a_{\alpha_1}(;\text{out})\,P_n\,(\text{in})\,c_{\beta_1}(;\text{out})\cdots c_{\beta_k}(;\text{out})|0\rangle$ graphically by Fig. 4. In Fig. 4, we have used notations drawn in Fig. 5.

Hereafter we shall always investigate the physical case when all external α's and β's in Fig. 4 are physical particles, so that they will not be explicitly drawn hereafter. Corresponding to shorthand notations used in Fig. 4, we introduce the following shorthand notations drawn in Fig. 6.

In the following, we can prove the identities drawn in Fig. 7. In other words, we shall prove

$$\langle\alpha\ \text{out}|Q_n\ (\text{in})|\beta\ \text{out}\rangle = \langle\alpha\ \text{out}|P_n\ (\text{in})|\ \beta\ \text{out}\rangle, \tag{3.7}$$

where Q_n (in) and P_n (in) are projection operators onto n-particle states given by Q_1 (in) and P_1 (in) [cf. (3.5)].

In order to prove Fig. 7 by mathematical induction, we first assume that Fig. 7 is valid for n. Then we obtain Fig.8.

In order to investigate the r.h.s. of Fig. 8, we use LSZ formulas [7] giving S matrix elements in terms of covariant Green's functions [i.e., (2.19) with (2.16)]. Then we can apply the generalized Slavnov identities (2.14) to the bottom line $\tilde{S}^{\pm(3)}$ in the l.h.s. of Fig. 9. By moving the term in the r.h.s. of (2.14) to the left of all factors in Θ blob, operators η, ζ, θ_c, and θ_d disappear on account of (2.2), (2.4), (2.6), and (2.8).

We indicate by the graphs (a), (a)', (b), and (c) in Fig. 9 the special case when η, ζ, θ_c, and θ_d in (2.14) do not commute or anticommute with the top line of each graph. As an illustration, we shall prove that the l.h.s. of Fig. 9(a) is given by the r.h.s. of Fig. 9 (a). For this purpose, we derive the following identities, with the help of which we can calculate the third term in the l.h.s. of Fig. 9(a):

$$\lim_{k^2\to 0} k^2\int d^4x e^{(\pm)ikx}\,\tilde{\Sigma}_\mu^3(x)|G)_{(\mp)}$$
$$= \lim_{k^2\to 0}\frac{ik_\mu}{-Z^0}k^2\int d^4x\, e^{(\pm)ikx}\,\tilde{d}^3(x)|G)_{(\mp)}, \tag{3.8}$$

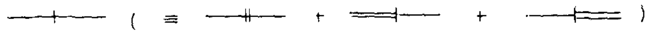


Fig. 2. Graphical representation of $Q_1 + R_1 + S_1$ defined in (3.5).

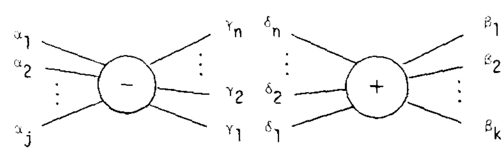


Fig. 3. Graphical representation of (3.6). [Since Faddeev–Popov ghosts' operators $\bar{c}$ and $\bar{d}$ are anticommuting, the ordering of creation and annihilation operators in (3.6) is important.]

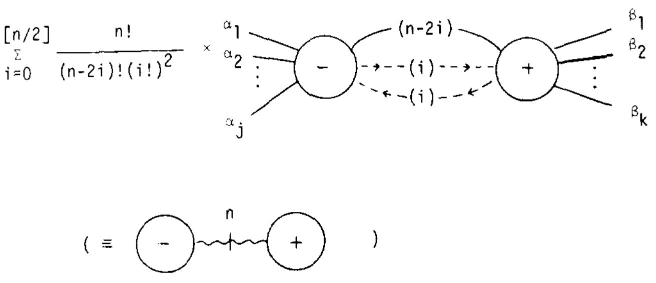


FIG. 4. Graphical representation of $n!\langle 0|a_{\alpha_j}(;\text{out})\cdots a_{\alpha_1}(;\text{out})\, P_n(;\text{in})\, c_{\beta_1}(;\text{out})\cdots c_{\beta_k}(;\text{out})|0\rangle$.

and

$$\lim_{l^2+m^2\to 0}(l^2+m^2)\int d^4x\, e^{(\pm)ilx}\tilde{A}^{\pm}(x)|G)_{(\mp)}$$
$$=\lim_{l^2+m^2\to 0}\frac{\mp i(-\xi_0 m^2\beta+vg\gamma)}{Z^c}\times(l^2+m^2)\int d^4x\, e^{(\pm)ilx}\tilde{d}^{\pm}(x)|G)_{(\mp)}. \qquad (3.9)$$

In order to show the equalities (3.8) and (3.9), we have used the separability of generalized Slavnov identities, i.e., the fact that the factors in the l.h.s. of (3.8) and (3.9) appear also in applying the generalized Slavnov identities to $(H|\tilde{A}^3_\mu(x)\tilde{S}^3(y)|G)_{(\mp)}$ and $(H|\tilde{\varphi}^{\pm}(x)\tilde{S}^{\pm}(y)|G)_{(\mp)}$. On the other hand, these Green's functions can be calculated by (A1) and (A6) (A7). Since the l.h.s. of (3.8) [(3.9)] is easily found to be proportional to the r.h.s. of (3.8) [(3.9)], proportional constant can be determined by using (A2) [(A8)], so that we obtain (3.8) [(3.9)]. In much the same way, the first and the second term in the l.h.s. of Fig. 1(a) can be investigated and they are found to vanish by using following three properties (I)–(III):

(I) There do not exist pole terms in the Green's functions among S particles [cf. (A9)].

(II) Transition between particle A^3_μ and physical particle φ^3 does not occur [cf. (A3)].

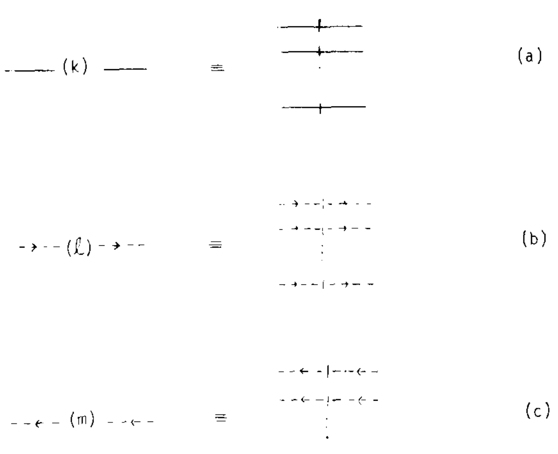


FIG. 5. In the r.h.s. of (a), (b), and (c), there are k lines (drawn in Fig. 2), l lines [drawn in Fig. 1(T_1)] and m lines [drawn in Fig. 1 (U_1)] respectively.

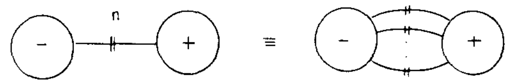


FIG. 6. The number n in the l.h.s. represents that in the r.h.s. there exist n lines drawn in Fig. 1 (Q_1).

(III) For massless (massive) physical particles A^3_μ, $[A^{\pm}_\mu]$, polarization vectors $\epsilon_\mu(k)$ $[\epsilon_\mu(p)]$ are required to satisfy

$$k_\mu\epsilon_\mu(k)=0,\ [\text{cf. }(3.5)]\quad p_\mu\epsilon_\mu(p)=0\quad [\text{cf. }(2.27)]. \qquad (3.10)$$

In conclusion, we have obtained Fig. 9(a). Similarly, we obtain Fig. 9(a)′ by using (3.8) and (3.9) for $|G)_+$. Graphical notations (V_1), (W_1), (X_1) and (Y_1) in Fig. 9(a) and (a)′ represent respectively V_1, W_1, X_1 and Y_1, which are defined by

$$\begin{aligned}
V_1+W_1+X_1+Y_1
&\equiv\Big[\int\frac{d^3k}{(2\pi)^3}c^0_d(k)|0\rangle\frac{-i}{Z^0}\langle 0|a^0_S(k)\\
&\quad+\int\frac{d^3l}{(2\pi)^3}c^{\pm}_d(l)|0\rangle\frac{-i}{Z^c}\langle 0|a^{\pm}_S(l)\Big]\\
&\Big[+\int\frac{d^3k}{(2\pi)^3}c^0_S(k)|0\rangle\frac{i}{Z^0}\langle 0|a^0_d(k)\\
&\quad+\int\frac{d^3l}{(2\pi)^3}c^{\pm}_S(l)|0\rangle\frac{i}{Z^c}\langle 0|a^{\pm}_d(l)\Big]\\
&\Big[+\int\frac{d^3k}{(2\pi)^3}c^0_c(k)|0\rangle\frac{-k^*_\nu}{2\mathbf{k}^2}\langle 0|a^0_{A_\nu}(k)\\
&\quad+\int\frac{d^3l}{(2\pi)^3}c^{\pm}_c(l)|0\rangle\frac{\pm 1}{-\xi_0m^2\beta+vg\gamma}\\
&\quad\times\langle 0|a^{\pm}_\varphi(l)\ \Big]\\
&\Big[+\int\frac{d^3k}{(2\pi)^3}c^0_{A_\mu}(k)|0\rangle\frac{-k^*_\mu}{2\mathbf{k}^2}\langle 0|a^0_c(k)\\
&\quad+\int\frac{d^3l}{(2\pi)^3}c^{\pm}_\varphi(l)|0\rangle\frac{\mp 1}{-\xi_0m^2\beta+vg\gamma}\\
&\quad\times\langle 0|a^{\pm}_c(l)\ \Big],
\end{aligned} \qquad (3.11)$$

respectively. {Hereafter we shall not explicitly give results for ⊕ blob [like Fig. 9 (a)′], since they are the ones obtained by reflecting results for ⊖ blob [like Fig. 9(a)] with respect to vertical bars.} In the similar fashion, we can prove Fig. 9(b) and Fig. 9(c), where we have used

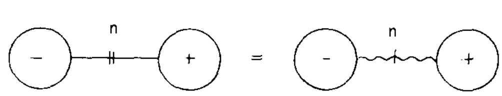


FIG. 7. The l.h.s. (r.h.s.) of Fig. 7 is defined by Fig. 6 [Fig. 4]. We shall prove that these identities hold in the case when all undrawn external lines represent physical particles or $\tilde{S}$ particles.

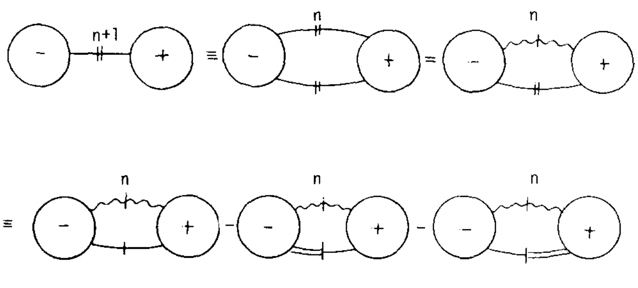


Fig. 8. Mathematical identities obtained by assuming Fig. 7 is valid for n [cf. Fig. 2].

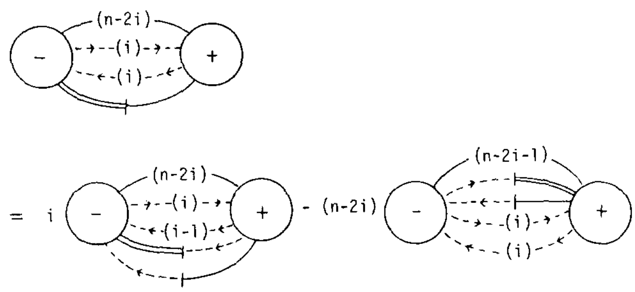


Fig. 10. Identities obtained by applying the generalized Slavnov identities (2.14) to the left bottom line $\tilde{S}$.

$$\lim_{l^2+m^2\to 0}(l^2+m^2)\int d^4x\, e^{ilx}\tilde{\Delta}^{\,\delta}(x)|G)_- = 0. \tag{3.12}$$

In Fig. 9(a), (a)', (b) and (c), it should be noticed that the positions of $\tilde{c}$ lines are different from the original positions of $\tilde{S}$ lines. In moving $\tilde{c}$ lines to their original positions, we must take account of the anticommuting property of $\tilde{c}$ lines.

Applying the generalized Slavnov identities (2.14) to the left bottom line and moving the term in the r.h.s. of (2.14) to the left of all factors in O blob, we obtain Fig. 10, where we have used (3.8) and (3.9) together with (3.5), (3.11), and Fig. 9.

Furthermore, the last graph in Fig. 10 can be similarly calculated and we obtain Fig. 11. By using anticommuting property of Faddeev–Popov ghosts's operators, the second graph in the r.h.s. of Fig. 11 is found to be of opposite value of the graph in the l.h.s. of Fig. 11. Then we find Fig. 12(a) from Fig. 10, Fig. 11 and the definition in Fig. 4. Similarly we find Fig. 12(b). Then Fig. 8, Fig. 12(a), (b) and the definition in Fig. 4 lead to the result that we obtain Fig. 7 for $n+1$, provided Fig. 7 is valid for n. On the other hand, Fig. 7 is valid for $n=1$ on account of the properties of external physical particles [i.e., (II) and (III) in (3.10)]. Therefore, we conclude by the mathematical induduction that Fig. 7 [i.e., (3.7)] is valid for general n.

In order to prove the unitarity of the renormalized S matrix by using Fig. 7, we split Q_1 in (3.5) into two terms, i.e., into the projection operator onto one physical particle states and into the projection operator onto one S particle states. Correspondingly to this splitting, we shall use graphical notations in Fig. 13(a) and (b), and then we obtain identities drawn in Fig. 14 by using properties (I)–(III) in (3.10) and Fig. 7. Thus we have proved the unitarity of the renormalized S matrix among physical particles, especially in a spontaneously broken case.

APPENDIX

With the help of the generalized Slavnov identities (2.14), we find the following pole structures of two point Green's functions:

$$(H|\tilde{A}^3_\mu(x)\tilde{A}^3_\nu(y)|G)_\pm = \mp\frac{1}{(2\pi)^4 i}\int d^4p\, e^{ip(x-y)}\Big[\frac{Z^0_A(\xi)\delta_{\mu\nu}}{p^2\pm i\epsilon} \times + \Big(\frac{1}{\xi}-Z^0_A(\xi)\Big)\frac{p_\mu p_\nu}{(p^2\pm i\epsilon)^2} + (\text{regular terms at } p^2=0)\Big], \tag{A1}$$

which will be expressed hereafter as

$$(H|\tilde{A}^3_\mu(x)\tilde{A}^3_\nu(y)|G)_\pm \to \frac{Z^0_A(\xi)\delta_{\mu\nu}}{p^2\pm i\epsilon} + \Big(\frac{1}{\xi}-Z^0_A(\xi)\Big)\times\frac{p_\mu p_\nu}{(p^2\pm i\epsilon)^2} + (\text{R.T. at } p^2=0).$$

In the same notation, we have

$$(H|\tilde{c}^3(x)\tilde{d}^{\,3}(y)|G)_\pm \to \frac{Z^0(\xi)\delta_{\mu\nu}}{p^2\pm i\epsilon} + (\text{R.T. at } p^2=0), \tag{A2}$$

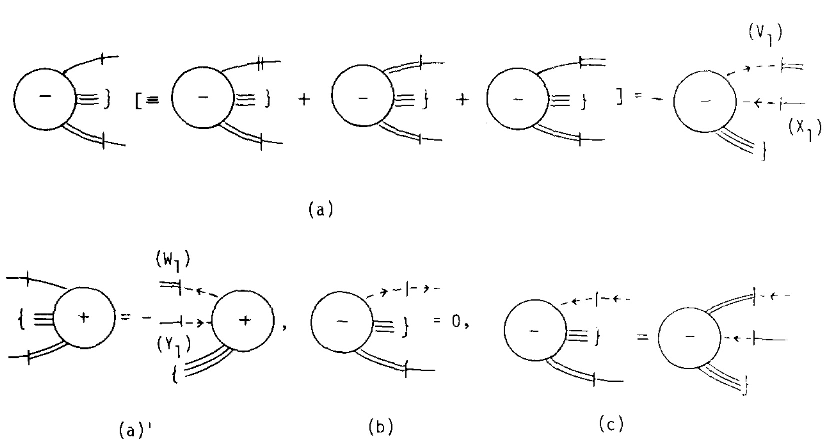


Fig. 9. Lines with (V_1), (W_1), (X_1) and (Y_1) in (a) and (a)' represent respectively V_1, W_1, X_1, and Y_1, in (3.11). Identities (a)–(c) will be used in deriving Fig. 10 and Fig. 11.

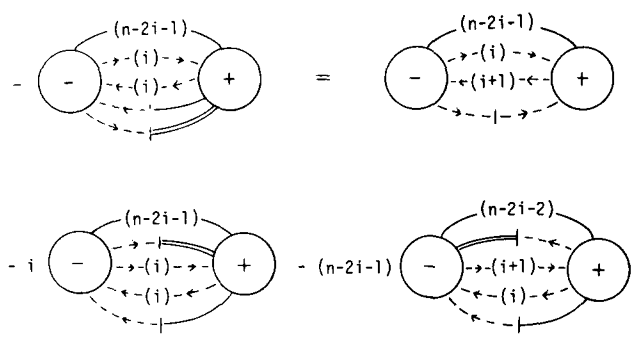


FIG. 11. Identities obtained by applying the generalized Slavnov identities (2.14) to the right bottom line $\tilde{S}$.

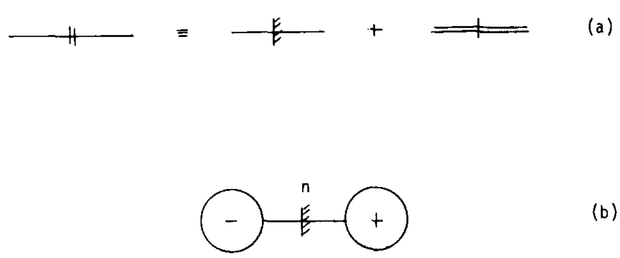


FIG. 13. The first (second) graph in the r.h.s. of (a) represents projection operators onto physical particles (nonphysical particles $\tilde{S}$). The number n in (b) represents that there exists n lines representing physical particle states.

$(H|\tilde{A}^3_\mu(x)\,\tilde{\varphi}^3(y)|G)_\pm \to 0$ (from charge conjugation invariance), (A3)

$$(H|\tilde{\varphi}^3(x)\,\tilde{\varphi}^3(y)|G)_\pm \to \frac{Z^0_\varphi(\xi)}{p^2+\mu_H^2\pm i\epsilon} + (\text{R.T. at } p^2+\mu_H^2=0), \tag{A4}$$

$$(H|\tilde{A}^+_\mu(x)\tilde{A}^-_\nu(y)|G)_\pm \to Z^c_A(\xi)\frac{\delta_{\mu\nu}+p_\mu p_\nu/M^2}{p^2+M^2\pm i\epsilon} + p_\mu p_\nu\frac{\alpha(\xi)}{p^2+m^2(\xi)\pm i\epsilon} + (\text{R.T. at } p^2+M^2=0 \text{ and } p^2+m^2(\xi)=0). \tag{A5}$$

$$(H|\tilde{A}^-_\mu(x)\tilde{\varphi}^+(y)|G)_\pm = -(H|\tilde{A}^+_\mu(x)\tilde{\varphi}^-(y)|G)_\pm \to \frac{p_\mu\beta(\xi)}{p^2+m^2(\xi)\pm i\epsilon} + (\text{R.T. at } p^2+m^2(\xi)=0). \tag{A6}$$

$$(H|\tilde{\varphi}^+(x)\tilde{\varphi}^-(y)|G)_\pm \to \frac{\gamma(\xi)}{p^2+m^2(\xi)\pm i\epsilon} + (\text{R.T. at } p^2+m^2(\xi)=0), \tag{A7}$$

and

$$(H|\tilde{c}^-(x)\tilde{d}^+(y)|G)_\pm \to \frac{Z^c(\xi)}{p^2+m^2(\xi)\pm i\epsilon} + (\text{R.T. at } p^2+m^2(\xi)=0), \tag{A8}$$

where masses μ_H in (A4) and M in (A5) have been proved to be gauge-independent[3] and fields like $A^\mp_\mu(x)$ are defined by

$$A^\pm_\mu(x) \equiv \frac{1}{\sqrt{2}}[A^1_\mu(x) \pm iA^2_\mu(x)], \text{ etc,.}$$

Applying identities [which can be derived from (2.14) and (2.3)]

$$(H|[\xi\partial^x_\mu\tilde{A}^\alpha_\mu(x) + g\epsilon_{\alpha\beta\gamma}v^\beta\tilde{\varphi}^\gamma(x)][\xi\partial^y_\nu\tilde{A}^\delta_\nu(y) + g\epsilon_{\delta\epsilon\zeta}v^\epsilon\tilde{\varphi}^\zeta(y)]|G)_\pm = -i\xi\delta_{\alpha\beta}\delta^4(x-y) \tag{A9}$$

to (A6), (A7), and (A8), we find the important relation

$$\xi^2m^4(\xi)\alpha(\xi) - 2\xi m^2(\xi)vg\beta(\xi) + v^2g^2\gamma(\xi) = 0. \tag{A10}$$

The expression (A5) means that there do not appear scalar particles having gauge-invariant mass M. This interesting fact of R_ξ gauges results from calculations which are something like

$$(H|\tilde{A}^+_\mu(x)\tilde{A}^-_\nu(y)|G)_\pm \to (\mu|g(p;\xi) + g(p;\xi)\Pi(p;\xi)g(p;\xi) + \cdots|\nu)$$

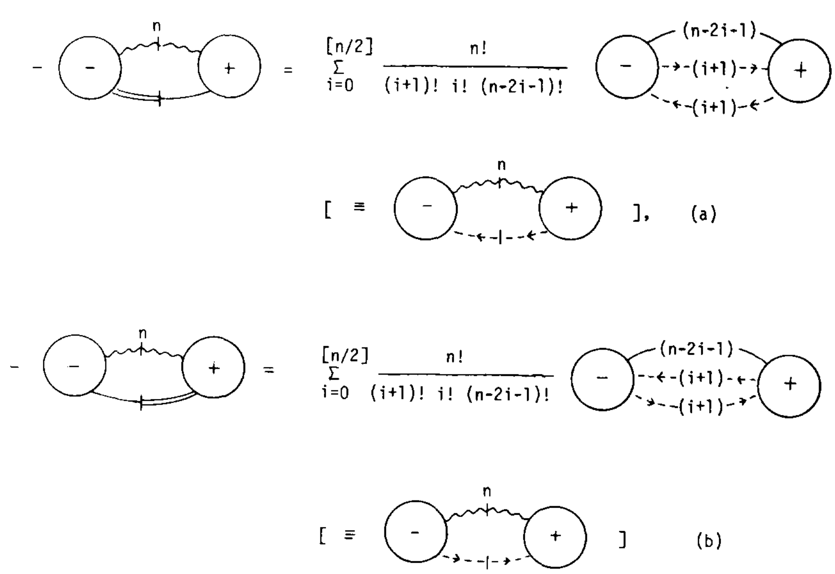


FIG. 12. Renormalized 't Hooft identities.

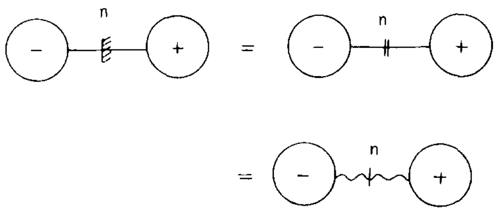


FIG. 14. These identities show that only physical particles contribute to unitarity relations.

$$= (\delta_{\mu\nu} + p_\mu p_\nu [1/\xi - 1 + B(p^2;\xi)/\xi] / \{p^2 + [v^2 g^2 - A(p^2;\xi) - p^2 B(p^2;\xi)] / \xi \pm i\epsilon\}) \times [p^2 + v^2 g^2 - A(p^2;\xi) \pm i\epsilon]^{-1},$$

where $(\mu|g(p;\xi)|\nu)$ is defined by (2.20), while $(\rho|\Pi(p;\xi)|\sigma)$ is the proper self-energy part

$$(\rho|\Pi(p;\xi)|\sigma) \equiv \delta_{\rho\sigma} A(p^2;\xi) + p_\rho p_\sigma B(p^2;\xi).$$

[1]'t Hooft, Nucl. Phys. B **33**, 173 (1971).
[2]'t Hooft, Nucl. Phys. B **35**, 167 (1971); C.M. Van Der Kolk, Nucl. Phys. B **125**, 513 (1977).
[3]S. Naito and S. Akahori, Phys. Rev. D **18**, 1230 (1978).
[4]S. Mandelstam, Phys. Rev. **175**, 1580 (1968).
[5]S.Naito, Phys. Rev. D **14**, 3584 (1976).
[6]S.Naito, J. Math. Phys **19**, 567 (1978).
[7]J.D. Bjorken and S.D. Drell, *Relativistic Quantum Fields* (McGraw–Hill, New York, 1965), Chapters 16 and 17.
[8]O.W. Greenberg, J. Math. Phys. **3**, 859 (1962); D.W. Robinson, Helv. Phys. Acta **35**, 403 (1962); G.F. Dell'Antonio, J. Math. Phys. **2**, 759 (1961).

Electromagnetism without monopoles is possible in nontrivial U(1)-fibre bundles[a)]

M. Asorey
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, Spain

L. J. Boya
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Salamanca, Spain

(Received 18 May 1979; accepted for publication 22 June 1979)

If the first homotopy group of a spacetime manifold M is p cyclic with $p > 1$, we prove that there exists $(p - 1)$ nontrivial and nonequivalent fibre bundles on M with gauge group U(1) in which all the connections describe electromagnetic gauge fields without monopoles. Some spacetimes verifying this condition are introduced. This stresses the physical relevance of the difference between the first Chern classes with integer coefficients and those with real coefficients.

1. INTRODUCTION

In recent years principal fibre bundles and their connections became a useful and powerful tool for the study of gauge fields. In a famous paper Wu and Yang[1] established a clarifying table of correspondence between the most important elements of gauge field theories and the differential geometric concepts related with them. From the publication of this table it has been used as a dictionary by the growing number of people working on gauge field theory.

The main aim of this paper is to analyze the character of the relation established in that table between electromagnetic gauge fields without monopoles and connections defined in trivial principal fibre bundles with structural group U(1). In Sec. 2 we prove that when the spacetime manifold M has its first homotopy group p-cyclic with $p \neq 1$, there are electromagnetic gauge fields without magnetic charges defined in nontrivial principal bundles on M with structural group U(1), which implies that in this case the relation between trivial principal fibre bundles and electromagnetism without monopoles is not so direct as the Wu–Yang's dictionary seems to indicate. In Sec. 3 we find some examples of spacetime manifolds with the topological conditions studied in Sec. 2. We see in Sec. 4 that the phenomenon found in Sec. 2 is a direct consequence of the non-one-to-one character of the correspondence between the first Chern classes with integer coefficients and those with real coefficients. We finish the paper by setting the main conclusions in Sec. 5.

2. NONTRIVIAL U(1)-FIBRE BUNDLES WITH ELECTROMAGNETIC FIELDS WITHOUT MONOPOLES.

In a previous paper[2] we have seen that in quantum mechanics the electromagnetic field is described by equivalence classes of gauge fields with gauge group U(1) which are defined on the maximal, open, and connected submanifold M of the spacetime manifold in whose points there are not magnetic sources or singular electric sources.

In that paper we have proved also that fixing a point $x_0 \in M$, each electromagnetic field $\mathscr{A}$ defined on M is univocally characterized by the H-homomorphism A from the H-group ΩM_{x_0} (of loops beginning and ending at x_0) into U(1) which maps each $\gamma \in \Omega M_{x_0}$ in its correspondent element of the holonomy group of any gauge field belonging to the class $\mathscr{A}$. Therefore the map κ from the set of electromagnetic fields defined on M into H-Hom$(\Omega M_{x_0},\mathrm{U}(1))$ given by

$$\kappa(\mathscr{A}) = A$$

for each $\mathscr{A}$, is one-to-one.

On the other hand Konstant[3] has proved that the image through κ of the electromagnetic fields having the same strength F on M is an orbit for the action of Hom$(\pi_1(M),\mathrm{U}(1))$ on H-Hom$(\Omega M_{x_0},\mathrm{U}(1))$ defined by

$$l\Pi(\gamma) = l(\gamma)\Pi([\gamma])$$

for every $l \in H$-Hom$(\Omega M_{x_0},\mathrm{U}(1))$, $\Pi \in$ Hom$(\pi_1(M),\mathrm{U}(1))$ and $\gamma \in \Omega M_{x_0}$. Hence, since this action is free, there is a biunivocal correspondence between the electromagnetic fields defined on M which give rise to the same field strength, and the group Hom$(\pi_1(M),\mathrm{U}(1))$. Now, we have the following lemma:

Lemma: For any gauge fields Γ and $\bar{\Gamma}$ defined in the same principal fibre bundle $P(M,\mathrm{U}(1))$ with the same field strength, there exists a homomorphism h from $\pi_1(M)$ into $\mathbb{R}$ such that

$$\kappa([\bar{\Gamma}]) = \kappa([\Gamma])e^{-2\pi ih}, \tag{1}$$

where $e^{-2\pi ih}$ is the homomorphism from $\pi_1(M)$ into U(1) defined by $e^{-2\pi ih}(\eta) = e^{-2\pi ih(\eta)}$ for every $\eta \in \pi_1(M)$.

Proof: Since U(1) is an Abelian group, if ω and $\bar{\omega}$ are the connection forms of Γ and $\bar{\Gamma}$, there will be on M a real 1-form α such that

$$\bar{\omega} = \omega + (\Pi_P^*\alpha)\chi, \tag{2}$$

where Π_P is the projection from P on M, and χ is the element of the Lie algebra of U(1) which is defined by the tangent vectors to the closed curve c of U(1) with $c(\theta) = e^{2\pi i\theta}$ for all $\theta \in [0,1]$.

On the other hand, since $d\bar{\omega} = d\omega + (\Pi_P^* d\alpha)\chi$ and because U(1) is Abelian, for the curvatures Ω and $\bar{\Omega}$ of Γ and $\bar{\Gamma}$ the equality $\bar{\Omega} = \Omega + (\Pi_P^* d\alpha)\chi$ holds. But, since Γ and $\bar{\Gamma}$ have the same field strength F, $\Omega = \bar{\Omega} = e/hc(\Pi_P^* F)\chi$, and therefore $d\alpha = 0$; i.e., the 1-form α is closed. Now, this implies that the map associating the real number $\int_\gamma \alpha$ to each closed curve γ in M, only depends of the homotopy class of γ, because $\int_{\gamma'} \alpha = \int_\gamma \alpha$ for every γ' with $[\gamma] = [\gamma']$.

[a)]With partial financial support of I.E.N.

If we call h the homomorphism from $\pi_1(M)$ into $\mathbb{R}$ induced by this map, we shall see that h verifies equality (1). Of course, it suffices to prove that for each closed curve γ in M, if c^γ is an horizontal lift of γ with respect to Γ, the curve $\bar{c}^\gamma$ defined by $\bar{c}^\gamma(t) = c^\gamma(t)\exp\{-2\pi i\int_0^t \alpha(\dot\gamma_t)\,dt\}$ for each $t\in[0,1]$ [where $\dot\gamma_t$ is the tangent vector at γ in $\gamma(t)$] is an horizontal lift of γ with respect to $\bar\Gamma$. According with Leibniz's formula for all $t\in[0,1]$

$$\dot{\bar{c}}_t^\gamma = \dot c_t^\gamma \exp\left\{-2\pi i\int_0^t \alpha(\dot\gamma_t)\,dt\right\} - \alpha(\dot\gamma_t)\chi^*_{\bar c^\gamma(t)},$$

where χ^* is the vector field associated to χ by the U(1) action on P, and $\dot c_t^\gamma\exp\{-2\pi i\int_0^t\alpha(\dot\gamma_t)\,dt\}$ is the image of $\dot c_t^\gamma$ by the action of $\exp\{-2\pi i\int_0^t\alpha(\dot\gamma_t)dt\}$ on P. Therefore, since $\omega(\chi^*) = \chi$ and $\omega(\dot c_t^\gamma) = 0$, it follows that

$$\bar\omega(\dot{\bar c}_t^\gamma) = 0;$$

i.e., $\bar c^\gamma$ is an horizontal lift of γ with respect to $\bar\Gamma$.

When M has $\pi_1(M) = \mathbb{Z}_p$ with $p > 1$, since $\mathrm{Hom}(\mathbb{Z}_p,\mathrm{U}(1)) = \mathbb{Z}_p$, only p different electromagnetic fields exist having the same field strength F: In particular p different electromagnetic fields exist having null field strength, and therefore having no monopoles, because the magnetic charge enclosed by any closed spacelike surface σ is given by $\int_\sigma F$, which is always null in this case.

On the other hand, we have $\mathrm{Hom}(\mathbb{Z}_p,\mathbb{R}) = 0$, which implies, by the lemma, that all gauge fields defined in a principal fibre bundle $P(M,\mathrm{U}(1))$ having null field strength are equivalent; i.e., they represent the same electromagnetic field. Hence, there are p nonequivalent U(1)-fibre bundles on M having a gauge field with null field strength, and such that every other U(1)-fibre bundle on M having such a gauge field is equivalent only to one of them. In particular, because the trivial U(1)-fibre bundle $M\times\mathrm{U}(1)$ has the standard flat connection (which obviously has null curvature), it is equivalent to one of those principal fibre bundles, but not to the others, which hence are nontrivial. Furthermore, each one of those nontrivial fibre bundles $P(M,\mathrm{U}(1))$ verifies that if ω is the connection form of one of the gauge fields defined in P with null F, equality (2) holds for every gauge field $\bar\Gamma$ defined in P which connection form is $\bar\omega$, and therefore

$$\bar F = d\alpha,$$

which implies that neither has monopoles, because $\int_\sigma d\alpha = 0$ for each closed surface σ in M. Summarizing, the following theorem holds.

Theorem: If M has $\pi_1(M) = \mathbb{Z}_p$ with $p > 1$, there are $(p-1)$ nonequivalent and nontrivial U(1)-fibre bundles on M such that all electromagnetic gauge fields defined in them do not have monopoles.

3. SOME SPACETIME MANIFOLDS WHOSE FIRST HOMOTOPY GROUP IS $\mathbb{Z}_p$ $(p\neq 1)$

Most of the spacetime manifolds considered in general relativity are connected and simply connected $[\pi_1(\mathscr{M}) = 0]$, but if we give a free action of $\mathbb{Z}_p$ $(p > 1)$ on them, the quotient manifolds $M = \mathscr{M}/\mathbb{Z}_p$ have $\pi_1(M) = \mathbb{Z}_p$. Indeed, in this case we have the following short exact sequence of homotopy groups [note that here $\pi_0(\mathscr{M}),\pi_0(M)$, and $\pi_0(\mathbb{Z}_p)$ are also groups]

$$\cdots\to\pi_1(\mathscr{M})\to\pi_1(M)\to\pi_0(\mathbb{Z}_p)\to\pi_0(\mathscr{M})\to\cdots,$$

and since $\pi_1(\mathscr{M}) = \pi_0(\mathscr{M}) = 0, \pi_1(M)\approx\pi_0(\mathbb{Z}_p) = \mathbb{Z}_p$. If such an action on $\mathscr{M}$ exists, according with the theorem there are electromagnetic fields without monopoles defined in nontrivial U(1)-fibre bundles on M.

The simplest 4-manifold where such a $\mathbb{Z}_2$-action exists is S^4, because the $\mathbb{Z}_2$-action defined by the antipodal map is obviously free ($\mathbb{Z}_2$ is the unique compact Lie group acting freely on even spheres[4]). But, since $S^4/\mathbb{Z}_2 = \mathbb{R}P^4$ is compact, for every Lorentzian metric defined in $\mathbb{R}P^4$ the causality condition does not hold, because there always exists closed nonspacelike curves. Hence $\mathbb{R}P^4$ is not physically acceptable as a spacetime manifold. Furthermore $\mathbb{R}P^4$ is not orientable because its Euler characteristic is 2, and therefore no spacetime structure defined on it exists.

In general relativity a very usual 4-manifold where there is also such a $\mathbb{Z}_2$-action is $\mathbb{R}^2\times S^2$. Indeed, the $\mathbb{Z}_2$-action induced by the antipodal map of S^2 is free. In this case it is easy to define a Lorentzian metric on $\mathbb{R}^2\times S^2/\mathbb{Z}_2 \approx\mathbb{R}^2\times\mathbb{R}P^2$ not having closed nonspacelike curves; e.g., since this $\mathbb{Z}_2$-action on $\mathbb{R}^2\times S^2$ is isometric with respect to internal and external Schwarzchild, Kruskal, and singularity free Reissner–Nordström metrics defined in $\mathbb{R}^2\times S^2$, each one of these metrics induces on $\mathbb{R}^2\times\mathbb{R}P^2$ a Lorentzian structure without closed nonspacelike curves through the canonical projection from $\mathbb{R}^2\times S^2$. But, because the manifold $\mathbb{R}^2\times\mathbb{R}P^2$ is not orientable, there is no spacetime structure on it.

The two previous cases are examples of 4-manifolds with $\pi_1(M) = \mathbb{Z}_2$, but without any spacetime structure. We shall see that there exist four-dimensional manifolds with $\pi_1(M) = \mathbb{Z}_p$ for every integer $p > 2$ and with spacetime structures. In order to find them let us consider the 4-manifold $\mathbb{R}\times S^3$. This is a very usual manifold in relativistic cosmology, because there are many cosmologically interesting spacetimes defined on it: Einstein's static universe, de Sitter's, Friedmann's (or Robertson–Walker's) with positive spatial curvature, etc. For each $p > 1$ we define the action of $\mathbb{Z}_p = \{a,a^2,\ldots,a^p = e\}$ on $\mathbb{R}\times S^3$ induced from the action of $\mathbb{Z}_p$ on S^3 given by

$$(x^1,x^2,x^3,x^4)a^r = (x^1\cos2\pi r/p - x^2\sin2\pi r/p,$$
$$x^1\sin2\pi r/p + x^2\cos2\pi r/p, x^3\cos2\pi r/p - x^4\sin2\pi r/p,$$
$$x^3\sin2\pi r/p + x^4\cos2\pi r/p),$$

for $r = 1,2,\ldots,p$ and for all $(x^1,x^2,x^3,x^4)\in S^3\subset\mathbb{R}^4$. Let us observe that when $p = 2$ this action is that induced by the antipodal map of S^3. For each $p > 1$ this action is obviously free and therefore the first homotopy group of its quotient manifold $\mathbb{R}\times S^3/\mathbb{Z}_p$ is $\mathbb{Z}_p$. This manifold is the *lens space* designated by $L(p,1)$. Now, since that $\mathbb{Z}_p$-action preserves any orientation on $\mathbb{R}\times S^3$, and is isometric and isochronous with respect to Einstein static universe, de Sitter, and Friedmann (with positive spacial curvature) metrics of $\mathbb{R}\times S^3$, each one of this metrics induces on $L(p,1)$ a spacetime structure verifying the causality condition. The structures induced in that way by the de Sitter spacetime are *elementary spacetime* *forms* according to Calabi and Markus.[5] The ones induced from the Friedmann metric have spatial anisotropy when $p > 2$.[6]

According to the theorem, on all the spacetimes built above there exist nontrivial U(1)-fibre bundles in which there are electromagnetic gauge fields not having magnetic charges.

4. INTEGRAL VERSUS REAL CHERN CLASSES

Another important characteristic which have those manifolds with $\pi_1(M) = \mathbb{Z}_p$ $(p > 2)$ is that the first Chern classes of the U(1)-fibre bundles defined on them with integer and real coefficients are not in one-to-one correspondence. Indeed, the exact sequence of Abelian groups

$$0 \to \mathbb{Z} \to \mathbb{R} \to \mathrm{U}(1) \to 0$$

induces the following commutative diagram of homomorphisms of Čech cohomology groups whose rows are exact sequences

$$\begin{array}{ccccccccccc}
\cdots \to & \check{H}^1(M,\mathbb{Z}) & \to & \check{H}^1(M,\mathbb{R}) & \to & \check{H}^1(M,\mathrm{U}(1)) & \to & \check{H}^2(M,\mathbb{Z}) & \to & \check{H}^2(M,\mathbb{R}) & \cdots \\
 & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
\cdots \to & \check{H}^1(M,\underline{\mathbb{Z}}) & \to & \check{H}^1(M,\underline{\mathbb{R}}) & \to & \check{H}^1(M,\underline{\mathrm{U}(1)}) & \to & \check{H}^2(M,\underline{\mathbb{Z}}) & \to & \check{H}^2(M,\underline{\mathbb{R}}) & \cdots
\end{array} \tag{3}$$

where $\underline{\mathbb{Z}}$, $\underline{\mathbb{R}}$ and $\underline{\mathrm{U}(1)}$ are, respectively, the sheaves of germs of local $\mathbb{Z}$, $\mathbb{R}$, and U(1) valued functions (differentiable). Since $\mathbb{Z}$ is discrete and M connected, the homomorphisms $\check{H}^q(M,\mathbb{Z}) \to \check{H}^q(M,\underline{\mathbb{Z}})$ of diagram (3) are isomorphisms. Now, since M is a spacetime manifold, it is paracompact[7] and therefore $\check{H}^q(M,\underline{\mathbb{R}}) = 0$ for every $q > 1$. This implies that the homomorphism $\check{H}^1(M,\underline{\mathrm{U}(1)}) \to \check{H}^2(M,\underline{\mathbb{Z}})$ of diagram (3) is an isomorphism, and therefore we have also an isomorphism between $\check{H}^1(M,\underline{\mathrm{U}(1)})$ and $\check{H}^2(M,\mathbb{Z})$. Since $\check{H}^1(M,\underline{\mathrm{U}(1)})$ can be identified with the classes of U(1)-fibre bundles on M, the latter isomorphism established another isomorphism between these classes and $\check{H}^2(M,\mathbb{Z})$. The element of $\check{H}^2(M,\mathbb{Z})$ associated in this way to each principal fibre bundle $P(M,\mathrm{U}(1))$ is its first Chern class with integer coefficients.[8] Therefore first Chern classes with integer coefficients classify completely the U(1)-fibre bundle classes.

On the other hand, because M is paracompact, the first Chern class with real coefficients of a principal fibre bundle $P(M,\mathrm{U}(1))$ is the element of $\check{H}^2(M,\mathbb{R}) \simeq H^2_{dR}(M)$ (second de Rham cohomology group) corresponding to the class of the closed 2-form F of M such that $(\Pi^*_P F)\chi$ is the curvature of any connection in P. This class is not dependent on the chosen connection and is the image of the first Chern class of P with integer coefficients through the homomorphism $i^2 : \check{H}^2(M,\mathbb{Z}) \to \check{H}^2(M,\mathbb{R})$ of diagram (3).[3,9] Therefore when this homomorphism is non-one-to-one, the correspondence between the first Chern classes with integer and with real coefficients of U(1)-fibre bundles is not one-to-one, i.e., few U(1)-fibre bundles can be the same first Chern real class and different integral class. Now then, when $\pi_1(M) = \mathbb{Z}_p$ with $p > 2$, since $\check{H}^1(M,\mathbb{R}) \simeq H^1_{dR}(M) = 0$ and $\check{H}^1(M,\mathrm{U}(1)) = \mathbb{Z}_p$, we have $\operatorname{Ker} i^2 = \mathbb{Z}_p$, which implies that i^2 is not one-to-one.

This result provides us with other proof of the theorem of Sec. 2, because if $\operatorname{Ker} i^2 = \mathbb{Z}_p$, there are p nonequivalent U(1)-fibre bundles in which all gauge fields give a strength that is an exact 2-form, and hence do not have magnetic charges.

Furthermore, this result implies that the Chern–Weil construction of the first Chern class of an U(1)-fibre bundle[10] builds its first Chern class with real coefficients, but does not determine that with integer coefficients.

5. CONCLUSIONS

Consequently, we might assure that the relation between electromagnetic fields without monopoles and trivial U(1)-fibre bundles is not so simple as Wu–Yang's dictionary of equivalences seems at first sight to indicate, because although all electromagnetic gauge fields defined in trivial U(1)-fibre bundles do not contain monopoles, also it is true that when the topology of the spacetime manifold is too complicate (and we have analyzed some concrete such cases) there can be nontrivial U(1)-fibre bundles in which all electromagnetic fields also have no monopoles.

On the other hand, we have seen that this result is a consequence of the existence of U(1)-fibre bundles having different first Chern classes with integer coefficients but the same with real ones, which implies the uselessness of the Chern–Weil method for building the first Chern class with integer coefficients of those fibre bundles.

ACKNOWLEDGMENTS

We thank Dr. J.F. Cariñena and Dr. M.T. Lozano for their very useful discussions about subjects directly related with this work.

[1]T.T. Wu and C.N. Yang, Phys. Rev. D **12**, 3845 (1975).

[2]M. Asorey and L.J. Boya, "Difference of observability between classical electromagnetic and gravitational gauge fields," Int. J. Theor. Phys. (to be published).

[3]B. Konstant, Springer Lecture Notes in Mathematics **170** (1970).

[4]G.E. Bredon, *Introduction to Compact Transformation Groups* (Academic, New York, 1972).

[5]E. Calabi and L. Markus, Ann. Math. **75**, 63 (1962).

[6]S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge U.P., Cambridge, England, 1973).

[7]According to K. Nomizu, "Lie groups and differential geometry," Math. Soc. Japan (1956), this result was shown by Hano and Ozeki. However the first published proof of the same is in R. Geroch, J. Math. Phys. **9**, 1739 (1968). A very simple proof is given in A. Marathé, J. Diff. Geom. **7**, 571 (1972). See also Ref. 9.

[8]F. Hirzebruch, *Topological Methods in Algebraic Geometry* (Springer, Berlin, 1966).

[9]M. Asorey, Ph.D. thesis University of Zaragoza, Spain (1978).

[10]J.W. Milnor and J.D. Stasheff, *Characteristic Classes* (Princeton U.P., Princeton, N.J., 1974).

Generalized method of a resolvent operator expansion. II.

M. Znojil

Institute of Nuclear Physics, Czechoslovak Academy of Sciences, 250 68 Řež, Czechoslovakia

(Received 27 September 1978)

The second part of the present series is devoted to the rearrangement of the Born series in the model space. An application is found in pion–nucleus scattering. We present a generalization of the "extended schematical model" suggested recently by Hirata *et al.*

1. INTRODUCTION

Inversion of a linear operator which is not necessarily Hermitian is a typical mathematical operation used in the scattering theory. The well-known Born expansion of a resolvent operator being not applicable in the whole region of energies, other approximation methods should be used in the intermediate energy region. In the present paper, we are inspired by the recent development in pion–nucleus scattering physics and describe the method which, within a model space, performs an efficient analytic continuation of high energy expansion by representing it in an equivalent matrix continued fraction form.

The method is obtained from the abstract parametric expansion of a resolvent $R(E) = (E - H)^{-1}$ as described in Ref. 1 (denoted I). As a new feature, the knowledge of the matrix elements of powers of H (moments) is required between the bra and ket vectors of the model spaces $d \subset V^+$ and $D \subset V$, respectively. This enables one to calculate explicitly the input parameters for definition of the expansion I.

In Sec. 2, we define the basis in the whole Hilbert space by applying the general parametrization[1] to H and H^+ separately and requiring block biorthogonality. In the special one-dimensional case, the resulting biorthogonal basis coincides with the one often used in numerical practice (Lanczos[2]) and physical calculations (Haydock and Kelly[3]).

Although the whole resolvent expansion I may be written immediately, the knowledge of scalar product (not employed in I) makes possible an efficient approximation technique. Due to the variability of the model spaces and their dimension M, all the necessary matrix elements of $R(E)$ may be taken between d and D without loss of generality. All the remaining basis vectors $\notin d, D$ are eliminated in Sec. 3, and their use is replaced by the knowledge of moments. This algebraic rearrangement leads to an efficient algorithm representing $R(E)$ between d and D by the matrix continued fraction formula. In the special case $M = 1$, we obtain the classical continued fraction results.[4]

The physical and numerically tested example of applicability of the present method is the extended schematical model (ESM) of the elastic π–^{4}He and π–^{16}O scattering given by Hirata *et al.*[5,6] It is described in Sec. 4 and shown to be equivalent to the one-dimensional version of the present method. This is the heuristic basis for the coupled channel generalization of ESM: We suggest that due to the analogy of the physical (isobar doorway) picture, the similar cutoff ($N \sim 3$–4) of the continued fraction (with $M > 1$) should work in the description of inelastic pion–nucleus scattering, with new indices representing the different coupled channels. The detailed numerical calculation should be given elsewhere, as it may vary in dependence on additional physical assumptions.

2. GENERALIZED LANCZOS BASIS

We shall work in the vector space V and its dual V^+ and employ small or capital letters for denoting bras ($\langle y| \in V^+$) or kets ($|Y\rangle \in V$), respectively. The kets (bras) will always occur in groups numbered by a pair of indices, e.g., $|X_k^i\rangle$, $i = 1,2...,M_k$ $(i \in \overline{1,M_k})$, $k = 1,2,\cdots$. The standard matrix conventions with respect to these indices will be accepted (kets or bras being ordered in rows or columns, respectively) so that we shall be allowed to omit either upper indices or both of them [compare respectively Eqs. (2) and (3) or Eqs. (3) and (4) below].

Let us consider a linear operator H on V. We shall formulate our task as the evaluation of the matrix elements of the type

$$T^{ij} = \langle x_1^i | R(E) | X_1^j \rangle, \qquad R(E) = (E - H)^{-1}, \quad i,j \in \overline{1,M_1} \tag{1}$$

for some arbitrarily chosen finite set of bras $\langle x_1^i|$ and an independent set of kets $|X_1^j\rangle$ spanning the model spaces $d \subset V^+$ and $D \subset V$ respectively, $d^+ \neq D$ in general. In the present section we describe the introduction of the block-biorthogonal basis generated from $\langle x_1|$ and $|X_1\rangle$.

We employ the formalism for expanding the resolvent operator $R(E)$ as described in Ref. 1. First, the two sets of groups of vectors

$$|X\rangle := \{\{|X_k^i\rangle, \ i \in \overline{1,M_k}\}, \ k = 1,2,\cdots\}$$

and

$$\langle x| := \{\{\langle x_k^i|, \ i \in \overline{1,m_k}\}, \ k = 1,2,\cdots\}$$

will be defined in a recurrent way from the initial groups $|X_1\rangle$ and $\langle x_1|$. According to I, the group $|X_{k+1}\rangle$ will be determined for any $k \geq 1$ by the k-fold action of H on $|X_1\rangle$. The partial subtraction of the linear combinations of "old group" vectors is also allowed by the corresponding recurrent defi-

 0022-2488/79/112330-04$01.00

nition (we slightly change the notation of I)

$$H|X_k^i\rangle = \sum_{j=1}^{M_{k-1}} |X_{k-1}^j\rangle C_k^{ji} + \sum_{j=1}^{M_k} |X_k^j\rangle A_k^{ji}$$

$$+ \sum_{j=1}^{M_{k+1}} |X_{k+1}^j\rangle B_k^{ji} \tag{2}$$

with arbitrary groups of parameters A, B, C. The shortened notation enables us to abbreviate Eq. (2) as

$$H|X_k\rangle = |X_{k-1}\rangle C_k + |X_k\rangle A_k + |X_{k+1}\rangle B_k \tag{3}$$

or even

$$H|X\rangle = |X\rangle Q. \tag{4}$$

The same sort of parametrization will be used for the action of the Hermitian conjugate operator H^+ ($\neq H$ in general) on another initial group $|x_1\rangle \in V$, $H^+|x\rangle = |x\rangle q^+$, i.e., $\langle x|H = q\langle x|$, i.e.,

$$\langle x_k^i|H = \sum c_k^{ij}\langle x_{k-1}^j| + \sum a_k^{ij}\langle x_k^i| + \sum b_k^{ij}\langle x_{k+1}^j|. \tag{5}$$

Employing formulas of I, we shall be able to write formally the expansion of $R(E)$ or $R^+(E)$. Once the scalar products $\langle x_k^i|X_l^j\rangle = D_{kl}^{ij}$ are known, both parametrizations of Eqs. (4) and (5) are not independent and may be used together as an input for I due to the identities

$$\langle x|H|X\rangle = DQ = qD. \tag{6}$$

In our paper, we reduce the complexity of the general expansion by demanding the block-diagonal form of D:

$$\langle x_k^i|X_l^j\rangle = D_k^{ij}, \quad k=l, \quad m_k = M_k,$$

$$= 0, \quad k \neq l. \tag{7}$$

This assumption (block-biorthogonality) will naturally restrict the freedom in the parametrizations (4) and (5) since Eq. (6) implies

$$\begin{aligned} \langle x_{k-1}|H|X_k\rangle &= D_{k-1}C_k = b_{k-1}D_k, \\ \langle x_k|H|X_k\rangle &= D_kA_k = a_kD_k, \\ \langle x_{k+1}|H|X_k\rangle &= D_{k+1}B_k = c_{k+1}D_k. \end{aligned} \tag{8}$$

This is very analogous to the one-dimensional case[7] ($M=1$) where the sets $|X\rangle$ and $\langle x|$ will form the well-known Lanczos biorthogonal basis.[2]

To restrict the freedom in the parametrizations (4), (5) allowed by Eq. (8) and hidden in the arbitrary decomposition of the linear superposition $b_k\langle x_{k+1}|$ and $|X_{k+1}\rangle B_k$, the modified vectors may be introduced in the form

$$\begin{aligned} |\Xi_{k+1}\rangle &= |X_{k+1}\rangle \Delta_k, \\ \Delta_k &= B_kB_{k-1}\cdots B_1, \quad \Delta_0 = I, \\ \langle \xi_{k+1}| &= \delta_k\langle x_{k+1}|, \\ \delta_k &= b_1b_2\cdots b_k, \quad \delta_0 = I. \end{aligned} \tag{9}$$

Assuming

$$\det B_k \neq 0 \neq \det b_k, \quad 0 \neq \det D_k,$$

$$M_k = M, \quad k = 1,2,\cdots \tag{10}$$

the vectors of Eq. (9) are defined uniquely by the recurrences

$$\begin{aligned} H|\Xi_k\rangle &= |\Xi_{k-1}\rangle g_{k-1}\beta_k + |\Xi_k\rangle g_k\alpha_k + |\Xi_{k+1}\rangle, \\ \langle \xi_k|H &= \beta_k g_{k-1}\langle \xi_{k-1}| + \alpha_k g_k\langle \xi_k| + \langle \xi_{k+1}|, \end{aligned} \tag{11}$$

$$\begin{aligned} \beta_k &= \delta_{k-1}D_k\Delta_{k-1} = (g_k)^{-1}, \quad g_0 = 0, \\ \alpha_k &= \delta_{k-1}a_kD_k\Delta_{k-1}, \end{aligned}$$

which follow from Eqs. (4), (5), (10). From the formal point of view Eqs. (11) may be considered to be special cases of parametrization (4), (5) with $B = I = b$. The resolvent operator expansion formulas I may thus be summarized by the expression

$$\langle \xi_k|R(E)|\Xi_m\rangle = \sum_{l=1}^{\min(k,m)} [(\varphi_k g_{k-1})(\varphi_{k-1}g_{k-2})\cdots(\varphi_{l+1}g_l)]$$

$$\times \varphi_l[(g_l\varphi_{l+1})\cdots(g_{m-1}\varphi_m)], \tag{12}$$

where $[\cdots] = I$ for $k=l$ or $l=m$ and the backward matrix recurrence

$$\varphi_k = (Eg_k - g_k\alpha_k g_k - g_k\varphi_{k+1}g_k)^{-1}, \quad \varphi_{N+1} = 0, \tag{13}$$

defines the auxiliary sequence φ_k in terms of α,β and $N \gg 1$ ($N \to \infty$).

3. MODEL SPACE DEFINITION OF PARAMETERS

In this section, we shall generalize some relations known for $M=1$ in the classical theory of moments.[4] We shall eliminate the higher groups of vectors ($k>1$) from the definitions of type (8) and arrive at analogs of the classical relations expressing, e.g., B_k of the one-dimensional case as a square root of the ratio of determinants formed from the Hankel matrices of momenta.[4]

By inserting the vectors (9) given by Eq. (11) into the matrices

$$\begin{aligned} \langle \xi_{k+1}|H^l|\Xi_{k+1}\rangle &= \alpha_{k+1}^{[l]}, \quad \langle \xi_{k+1}|H^l|\Xi_k\rangle = \gamma_k^{[l]}, \\ \langle \xi_k|H^l|\Xi_{k+1}\rangle &= \beta_{k+1}^{[l]}, \end{aligned} \tag{14}$$

we obtain the recurrence relations for matrices (14) initialized at $k=0$ by the (generalized) moments ($M_1 \times M_1$ matrices) $\alpha_1^{[l]}$ denoted as

$$\langle \xi_1|H^l|\Xi_1\rangle = \langle x_1|H^l|X_1\rangle = \langle l\rangle, \quad l = 0,1,2,\cdots. \tag{15}$$

Hence the two sets of matrices $\alpha_k = \alpha_k^{[1]}$, $\beta_k = \beta_k^{[1]} = \gamma_{k-1}^{[1]}$ are uniquely defined by H and by the choice of groups $|X_1\rangle$ and $\langle x_1|$ spanning model spaces D and d, respectively. It is not surprising, since the simple expansion of $R(E)$ within the model space $[\equiv \varphi_1(E)]$ into powers of H provides the same conclusion.

The insertion of α_k,β_k into Eq. (13) defines φ_1, i.e., the rearrangement of the standard power series in form of the matrix continued fraction. Although the convergence is not investigated now, we shall assume that it is sufficiently quick in the practical calculations (cf., e.g., Sec. 4 or Ref. 8).

The recurrences resulting from Eq. (14) do not give the optimal form of the desired matrices since many cancellations occur. In spite of the cancellations, they may be used for $k = 1,2$, and 3, given

$$\beta_1 = g_1^{-1} = D_1 = \langle 0\rangle, \quad \alpha_1 = \langle 1\rangle,$$

$$\begin{aligned}
\beta_2 &= g_2^{-1} = \langle 2\rangle - \langle 1\rangle g_1\langle 1\rangle,\\
\alpha_2 &= \langle 3\rangle - \langle 2\rangle g_1\langle 1\rangle - \langle 1\rangle g_1\langle 2\rangle + \langle 1\rangle g_1\langle 1\rangle g_1\langle 1\rangle,\\
\beta_3 &= \langle 4\rangle - \langle 2\rangle g_1\langle 2\rangle - (\langle 3\rangle - \langle 2\rangle g_1\langle 1\rangle)\\
&\quad\times g_2(\langle 3\rangle - \langle 1\rangle g_1\langle 2\rangle),\\
\alpha_3 &= \langle 5\rangle - \langle 2\rangle g_1\langle 1\rangle g_1\langle 2\rangle - [(\langle 3\rangle - \langle 2\rangle g_1\langle 1\rangle)g_2(\langle 4\rangle\\
&\quad - \langle 1\rangle g_1\langle 1\rangle g_1\langle 2\rangle] - [\text{h.c.}] + (\langle 3\rangle - \langle 2\rangle g_1\langle 1\rangle)\\
&\quad\times g_2(\langle 3\rangle - \langle 1\rangle g_1\langle 1\rangle g_1\langle 1\rangle)g_2(\langle 3\rangle - \langle 1\rangle g_1\langle 2\rangle).
\end{aligned} \tag{16}$$

Because of the rapidly increasing complexity of these expressions, we prefer to simplify them by introducing some auxiliary subexpressions for higher k.

We start with the definition of projectors and operators

$$\begin{aligned}
&P_k = |\Xi_k\rangle g_k\langle\xi_k|, \quad S_k = S_{k-1} - P_k,\\
&k = 1,2,\cdots, \quad S_0 = I,\\
&\Omega_k = 1 - |\Xi_{k+1}\rangle g_{k+1}\langle\xi_1|H^k,\\
&\omega_k = 1 - H^k|\Xi_1\rangle g_{k+1}\langle\xi_{k+1}|.
\end{aligned} \tag{17}$$

Exploiting the block tridiagonality of Q in Eq. (4) and employing the obvious properties

$$\begin{aligned}
|\Xi_k\rangle &= S_{k-1}H|\Xi_{k-1}\rangle = S_{k-1}HS_{k-2}H|\Xi_{k-2}\rangle\\
&= S_{k-1}H^{k-1}|\Xi_1\rangle,\\
\langle\xi_k| &= \langle\xi_1|H^{k-1}S_{k-1}, \quad \Omega_k|\Xi_{k+1}\rangle = 0 = \langle\xi_{k+1}|\omega_k,\\
S_{k+1} &= \Omega_k S_k = S_k\omega_k,
\end{aligned} \tag{18}$$

we arrive at the fundamental implicit form of the recurrent definitions of parameters

$$\begin{aligned}
\beta_k &= \langle\xi_k|\Xi_k\rangle = \langle\xi_1|H^{k-1}S_{k-1}H^{k-1}|\Xi_1\rangle,\\
\alpha_k &= \langle\xi_1|H^{k-1}S_{k-1}HS_{k-1}H^{k-1}|\Xi_1\rangle,\\
S_{k-1} &= S_{k-2} - S_{k-2}H^{k-2}|\Xi_1\rangle g_{k-1}\langle\xi_1|H^{k-2}S_{k-2}.
\end{aligned} \tag{19}$$

To derive the explicit algorithm which should give an iterative definition of β_k, based on the initialization (15), we introduce the auxialiary sequence $\eta_{k+1}^{[l,r]}$, $l,r \geqslant k \geqslant 1$, defined by the recurrence

$$\eta_{k+1}^{[l,r]} = \eta_k^{[l,r]} - \eta_k^{[l,k-1]}g_k\eta_k^{[k-1,r]}, \quad \eta_1^{[l,r]} = \langle l+r\rangle, \tag{20}$$

which coincides with the required sequence β_{k+1} for $l = r = k$ since $\eta_{k+1}^{[l,r]} = \langle\xi_1|H^lS_kH^r|\Xi_1\rangle$, $k \geqslant 0$, $\eta_{k+1}^{[k,k]} = \beta_{k+1}$. Thus, Eq. (20) is the first part of the desired algorithm.

The form of α_k will be more complex since the definition (19) containes two projectors. By the explicit use of the projection properties of Ω_k and ω_k, it is a rather trivial block-matrix multiplication manipulation to show that the ansatz

$$\begin{aligned}
\alpha_{k-1} &= \langle\xi_1|H^{k-2}\Omega_{k-3}Z_{k-3}\\
&\quad\times\omega_{k-3}H^{k-2}|\Xi_1\rangle, \quad k = 3,4,\cdots,
\end{aligned} \tag{21}$$

$$Z_0 = H, \quad Z_1 = H - P_1HP_1, \cdots,$$

implies

$$\begin{aligned}
\alpha_{k+1} &= \langle\xi_1|H^k\Omega_{k-1}[\Omega_{k-3}Z_{k-3}\omega_{k-3}\\
&\quad - P_{k-1}HP_{k-1}]\omega_{k-1}H^k|\Xi_1\rangle, \quad k = 3,4,\cdots.
\end{aligned}$$

It enables us to write

$$\begin{aligned}
&\alpha_{k+1}\\
&= (\langle\xi_1|H^k - \langle\xi_1|H^kS_{k-1}H^{k-1}|\Xi_1\rangle g_k\langle\xi_1|H^{k-1})Z_{k-1}\\
&\quad\times(H^k|\Xi_1\rangle - H^{k-1}|\Xi_1\rangle g_k\langle\xi_1|H^{k-1}S_{k-1}H^k|\Xi_1\rangle
\end{aligned} \tag{22}$$

[cf. Eq. (21), $k > 1$] and derive the recurrence for operators Z

$$\begin{aligned}
Z_k &= \Omega_{k-2}(Z_{k-2} - S_{k-2}H^{k-1}|\Xi_1\rangle g_k\\
&\quad\times\alpha_kg_k\langle\xi_1|H^{k-1}S_{k-2})\omega_{k-2}, \quad k = 2,3,\cdots
\end{aligned} \tag{23}$$

From this we define the matrices $\vartheta_k^{[l,r]} = \langle\xi_1|H^lZ_kH^r|\Xi_1\rangle$, $l,r \geqslant k$ which, due to Eq. (23), satisfy the recurrent definition

$$\begin{aligned}
\vartheta_k^{[l,r]} &= \mu_{k-2}^{[l,r]} - \eta_{k-1}^{[l,k-2]}g_{k-1}\mu_{k-2}^{[k-2,r]}\\
&\quad - \mu_{k-2}^{[l,k-2]}g_{k-1}\eta_{k-1}^{[k-2,r]} + \eta_{k-1}^{[l,k-1]}g_{k-1}\\
&\quad\times\mu_{k-2}^{[k-2,k-2]}g_{k-1}\eta_{k-1}^{[k-2,r]}, \quad k = 3,4,\cdots,
\end{aligned} \tag{24}$$

where $\mu_{k-2}^{[l,r]} = \vartheta_{k-2}^{[l,r]} - \eta_{k-1}^{[l,k-1]}g_k\alpha_kg_k\eta_{k-1}^{[k-1,r]}$ and the initializations are

$$\begin{aligned}
&\vartheta_0^{[l,r]} = \langle l+r+1\rangle,\\
&\vartheta_1^{[l,r]} = \langle l+r+1\rangle - \langle l\rangle g_1\alpha_1g_1\langle r\rangle,\\
&\vartheta_2^{[l,r]} = \langle\xi_1|H^lS_1HS_1H^r|\Xi_1\rangle - \eta_2^{[l,1]}g_2\alpha_2g_2\eta_2^{[1,r]}.
\end{aligned} \tag{25}$$

From Eq. (22) we obtain the definition

$$\begin{aligned}
\alpha_{k+1} &= \vartheta_{k-1}^{[k,k]} - \eta_k^{[k,k-1]}g_k\vartheta_{k-1}^{[k-1,k]}\\
&\quad - \vartheta_{k-1}^{[k,k-1]}g_k\eta_k^{[k-1,k]} + \eta_k^{[k,k-1]}g_k\vartheta_{k-1}^{[k-1,k-1]}\\
&\quad g_k\eta_k^{[k-1,k]}, \quad k = 1,2,\cdots, \quad \alpha_1 = \langle 1\rangle,
\end{aligned} \tag{26}$$

of the parameter matrices α_{k+1}, i.e., the second part of the algorithm based on the second auxiliary sequence $\vartheta_k^{[l,r]}$ defined by the recurrence (24). Thus, Eq. (13) with $k = N$, $N-1,\ldots,1$ and Eqs. (20), (24), and (26) define the matrix elements of $R(E)$ between the vectors from d and D model subspaces in terms of moments $\langle l\rangle$, $l = 0,1,\ldots,2N-1$.

4. AN EXAMPLE OF APPLICATION: ESM EXTENDED

The role of relativistic effects in the pion–nucleus scattering is probably not negligible, and some of the recent attempts try therefore to introduce the isobar degrees of freedom into the nucleus and employ the dominant role of the π-$N(1232)$ resonance in the microscopic description of the π-^{4}He [5,9] and π-^{16}O [6] elastic scattering at intermediate energies. We shall briefly review the algorithm suggested by Hirata *et al.*[6] (ESM) since it considers the pion–nucleus transition operator just in the form of Eq. (1) with $M=1$. The $|X_1\rangle$ or $\langle x_1|$ state is assumed to be a coherent superposition of the isobar–hole (Δh) states representing the response of the nuclear ground state to the incoming pion of momentum k, $|X_1\rangle = \Sigma\, |\Delta h\rangle\, \langle \Delta h\, |H\, |k,K\rangle$, $K=0$, or outgoing pion of momentum k', $\langle x_1| = \Sigma\, \langle k',K\, |H\, |\Delta h\rangle\, \langle \Delta h\, |$, $K=0$, respectively. We might therefore apply our method with $M=1$. In Ref. 6 the equivalent numerical evaluation is fulfilled along the standard lines: The complete basis $|e_1\rangle, |e_2\rangle, |e_3\rangle, \ldots, |e_{\bar N}\rangle$ in V is constructed by using the Lanczos method[7] and diagonalizing the realistic semiphenomenological Hamiltonian H. The inversion (Green function) is then given by summing all the contributions from the $\bar N$ separate states.

The numerical conclusions of Ref. 6 are surprising and nontrivial: The cutoff parameter $N \ll \bar N$ sufficient to accurate description of π-^{16}O elastic scattering turned out to be as low as 3. This defines ESM and restricts strongly the necessary number N of collective doorway states $|e_k\rangle$ (linear combinations of $|X_1\rangle$, $|X_2\rangle, \ldots, |X_N\rangle$!) belonging to the whole isobar–hole space V.

The same result has been obtained for different nuclei.[5] This inspires us to the assumption that also the nonelastic scattering may be considered along the same lines. Using the same physical picture of the isobar excitations in the resonance region, the adequate method should be based on our matrix continued fraction formula (13) with $k=1$ which is entirely equivalent to diagonalization due to the use of the same truncated part of the vector space V spanned by the groups $|X_1\rangle, \ldots, |X_N\rangle$.

Even in the simplest one-dimensional version, our method represents further simplification of ESM since we have eliminated the redundant intermediate step of diagonalization which is an infinite numerical process in principle.[7] Using $M>1$, we may take into account the possibility of channel coupling reflected by interaction within the k th group $|X_k^i\rangle$, $k \geqslant 1$, with the upper index numbering the channels. Assuming that the physical mechanism (isobar creation) is not modified too much in the inelastic processes, we are entitled to use even the same level of approximation ($N=3$–4), although more involved physical structure is described by more states, channel coupling, and $M>1$. The algebraic utilization of the block tridiagonality of H in block-biorthogonal basis is related also to certain numerical advantages of the continued fraction form.[4] The detailed evaluations of the pi-nucleus cross sections are to be published elsewhere after clarifying the problems connected with the proper choice of the effective interaction H.[5,9] This is out of the scope of the present paper.

5. SUMMARY

While in I the series expansion of the vectors $R(E)\,|X_1^i\rangle$ in terms of the expansion set $|X_k\rangle \sim H^{k-1}|X_1\rangle + \cdots$, $k=1,2,\cdots$, was derived, the present paper is oriented to the expansion of the model space matrix elements $\langle x_1^i|R(E)|X_1^j\rangle$ into the matrix continued fraction by eliminating all the groups of vectors $\langle x_k|$, $|X_k\rangle$ with $k>1$. The recurrences are given which define the parameters of the continued fraction in terms of moment matrices $\langle x_1^i|H^l|X_1^j\rangle$ in the most compact way. The groups of $M \geqslant 1$ vectors $\langle x_1|$ and $|X_1\rangle$ span two independent bra and ket model spaces d and D, respectively. This guarantees the broad and flexible applicability of the method which was illustrated by simplifying and generalizing ESM[6] in pion–nucleus scattering.

1. M. Znojil, J. Math. Phys. **18**, 717 (1977).
2. C. Lanczos, J. Res. NBS **45**, 255 (1950).
3. R. Haydock and M.J. Kelly, J. Phys. C **8**, L290 (1975).
4. N.I. Achiezer, *Klassicheskaya problemma momentov* ... (GIFML, Moscow, 1961); H.S. Wall, *Analytic Theory of Continued Fractions* (Van Nostrand, London, 1948).
5. M. Hirata, F. Lenz, and K. Yazaki, Ann. Phys. (N.Y.) **108**, 116 (1977).
6. M. Hirata, J.H. Koch, F. Lenz, and E.J. Moniz, Phys. Lett. B **70**, 281 (1977).
7. J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1965).
8. M. Znojil, J. Phys. A: Math. and Gen. **9**, 1 (1976).
9. W. Weise, Nucl. Phys. A **278**, 402 (1977).

On the Lorentz–Dirac equation

J. L. Sanz
Departamento de Física Teórica, Universidad Autónoma de Madrid, Canto Blanco, Madrid-34 Spain

(Received 14 May 1979; accepted for publication 15 June 1979)

We study the problem of a charged particle affected by an external electromagnetic field, adopting the point of view that its evolution is governed by an autonomous ordinary second order system of differential equations on M_4 whose acceleration can be expanded into a power series of the charge e and assuming the acceleration satisfies the Lorentz–Dirac equation. Thus, we locate the dynamics of the system in a symplectic structure up to order e^3 for external fields satisfying certain weak conditions. This is applied to obtain conserved quantities associated to symmetries of the external field that are also isometries.

I. INTRODUCTION

One of the most important and interesting problems in classical electrodynamics concerns the equations of motion of a single charged particle moving in an external electromagnetic field. The usual Lorentz force represents a direct action of the field on the particle and does not contain the action of the charge on itself (self-action). When this self-field is taken into account one can derive as equations of motion of a charged particle the well-known Lorentz-Dirac equation[1] (LDE) which contains the Lorentz force and a radiation reaction term involving derivatives of the acceleration.

There are many difficulties involved in the LDE if one attempts to consider it as an exact equation of motion of third differential order in the position of the particle: (i) there is no "Newtonian causality" in the sense that initial position and velocity do not suffice to predict the motion; (ii) there exist the so-called "runaway solutions," i.e., exact solutions such that the acceleration increases without limit. To avoid this type of unphysical solution, one can assume that the initial acceleration must be such that the acceleration, in the limit as time approaches infinity, must remain finite. Thus, one arrives at an integro-differential equation whose kernel contains the external forces in such a way that the particle experiences the unphysical phenomenon of preacceleration. Moreover, "Newtonian causality" is not included in this approach because the usual constraint on the acceleration allows two or more solutions to some problems.[2]

A different approach can be followed in order to eliminate these unphysical results: one can consider the radiation reaction force as a perturbation compared with the usual Lorentz force, following the suggestions of some authors.[3] More precisely, we shall assume that the evolution of a charged particle affected by an external electromagnetic field is governed by an autonomous ordinary second order system of differential equations on M_4; in this sense, the LDE needs to be viewed as a condition that must satisfy the acceleration. We also assume that these functions can be expanded into a power series of the electric charge e. These assumptions incorporate "Newtonian causality," eliminate the "runaway solutions," and also the possibility of "advanced effects" (preacceleration) and furnish a recurrent method to calculate the acceleration to any order.

We define the notions of "Hamiltonian form" in the past (resp. future) and give a proof of a formal theorem stating the uniqueness of the Hamiltonian form in the past (resp. future) by assuming its existence at any order in the perturbative framework we have adopted. We also give the explicit analytical form of both "Hamiltonian forms" up to order e^3 for external electromagnetic fields satisfying certain weak conditions and obtain the exact analytical form of the generating function associated to both "Hamiltonian forms." Summing up, we locate the dynamics of the single charged particle in a symplectic structure up to order e^3.

There are physical reasons, not only mathematical ones, for introducing this symplectic structure: In this paper we show an application concerning the construction (without any ambiguity) of conserved quantities associated to certain symmetries of the external electromagnetic field. In fact, we give the general form of the conserved quantity associated to a symmetry of the field that is also an isometry.

II. THE LORENTZ–DIRAC EQUATION (LDE)

A. The covariant formalism[4]

Consider the autonomous ordinary second order system of differential equations on M_4

$$\frac{dx^\alpha}{d\tau} = u^\alpha, \quad \frac{du^\alpha}{d\tau} = \xi^\alpha(x^\beta, u^\gamma), \tag{1a}$$

where the u^α's satisfy

$$u^2 \equiv -(uu) = +1, \quad 0 < u^0 < +\infty. \tag{1b}$$

Its general solution depends on six essential parameters. In the framework of special relativity, this covariant formalism is frequently used to study the nonisolated systems which are constituted by one structureless point particle. ξ^α is the so-called acceleration of the particle and may depend on certain characteristic parameters of the particle (e.g., the mass m, the charge e, etc.); we shall write the explicit dependence on the mass of the particle: $\xi^\alpha(x^\beta, u^\gamma; m)$. The following system on M_4,

$$\frac{dx^\alpha}{d\lambda} = \pi^\alpha, \quad \frac{d\pi^\alpha}{d\lambda} = \theta^\alpha(x^\beta, \pi^\gamma), \tag{2a}$$

 0022-2488/79/112334-05$01.00

where

$$\theta^{\alpha}(x^{\beta},\pi^{\gamma}) = \pi^2\xi^{\alpha}(x^{\beta},u^{\gamma}\equiv\pi^{-1}\pi^{\gamma};m\equiv\pi),$$

$$\pi\equiv + [-(\pi\pi)]^{1/2}, \quad \pi^2\equiv(\pi)^2, \tag{2b}$$

has the same integral curves as Eqs. (1). The geometrical view of the mass in this formulation is clear. Hereafter, we shall call the function θ^{α} the dynamics which, obviously, satisfies $(\pi\theta) = 0$. The dynamic system (2) can be represented by the vector field on $\widehat{M}_4$

$$\vec{H} = \pi^{\alpha}\partial_{\alpha} + \theta^{\alpha}\partial_{\underline{\alpha}}. \tag{3}$$

B. The Lorentz–Dirac equation

If $F_{\alpha\beta}(x^{\lambda})$ is the external electromagnetic field acting on a particle with electric charge e and mass m, it is usually assumed that the motion of such a particle is governed by[1] the LDE

$$m\dot{u}^{\alpha} = eF^{\alpha\rho}(x^{\sigma})u_{\rho} + \gamma e^2(\ddot{u}^{\alpha} - \dot{u}^2u^{\alpha}), \tag{4}$$

where $\cdot\equiv d/d\tau$, $\gamma\equiv 2/3$, $\dot{u}^2\equiv(\dot{u}\dot{u})$. The second term in the rhs is due to radiation reaction, and it represents an action of the particle on itself. We have remarked in the Introduction some nonphysical phenomena involved in the LDE: existence of "runaway solutions" and violation of "Newtonian causality" if one attempts to consider it as an exact equation of motion of third differential order, existence of "advanced effects" (preacceleration) and violation of "Newtonian causality," if one attempts to consider it as an exact integro-differential equation. Thus, we shall follow another approach to eliminate these nonphysical results, adopting the assumptions:

(i) The equations of motion of the charged particle constitute a dynamic system on M_4 of type (1),

$$\dot{u}^{\alpha} = \xi^{\alpha}(x^{\beta},u^{\gamma};m). \tag{5}$$

(ii) The acceleration ξ^{α} must satisfy the LDE in the sense that

$$m\xi^{\alpha} = eF^{\alpha\rho}(x^{\sigma})u_{\rho} + \gamma e^2\left[u^{\rho}\partial_{\rho}\xi^{\alpha} + \xi^{\rho}\frac{\partial\xi^{\alpha}}{\partial u^{\rho}} - (\xi\xi)u^{\alpha}\right]. \tag{6}$$

(iii) The functions ξ^{α} can be expanded into a power series of the charge e,

$$\xi^{\alpha} = \sum_{r=1}^{\infty} e^r\xi^{(r)\alpha}, \tag{7}$$

where the function $\xi^{(r)\alpha}$ are charge–independent.

Thus, we have incorporated "Newtonian causality," i.e., the system possesses three degrees of freedom, and we eliminate unphysical results because the previous assumptions constitute the basis of a recurrent method which allows us to compute the acceleration to any order as a function of the corresponding one to lower orders. The introduction of expansion (7) in (6) yields

$$\xi^{(1)\alpha} = m^{-1}F^{\alpha\rho}(x^{\sigma})u_{\rho}, \quad \xi^{(2)\alpha} = 0,$$

$$\xi^{(r)\alpha} = \gamma m^{-1}\Big\{u^{\rho}\partial_{\rho}\xi^{(r-2)\alpha} + \sum_{\substack{p+q=r-2\\ p,q=1,\dots,r-3}}\left[\xi^{(p)\rho}\frac{\partial\xi^{(q)\alpha}}{\partial u^{\rho}} - \xi^{(p)\rho}\xi^{(q)}_{\rho}u^{\alpha}\right]\Big\}, \quad r>2. \tag{8}$$

The problem regarding convergence of the series obtained by this method remains unsolved. In particular, we have up to order e^3

$$\xi^{\alpha} = em^{-1}u_{\rho}\{F^{\alpha\rho}(x^{\sigma}) + \gamma e^2m^{-1}u^{\sigma}\partial_{\sigma}F^{\alpha\rho}\} + O(e^4) \tag{9}$$

and the dynamics, according to (2b), is

$$\theta^{\alpha} = e\pi_{\rho}\{F^{\alpha\rho}(x^{\sigma}) + \gamma e^2\pi^{-2}\pi^{\sigma}\partial_{\sigma}F^{\alpha\rho}\} + O(e^4). \tag{10}$$

It is evident that the e^3-term vanishes in the particular case of a constant external field and thus radiation reaction does not appear until order e^4

$$\xi^{\alpha} = em^{-1}u_{\rho}\{F^{\alpha\rho} + \gamma e^3m^{-2}(\eta^{\alpha\lambda} + u^{\alpha}u^{\lambda})F_{\lambda\sigma}F^{\sigma\rho}\} + O(e^7), \quad F_{\alpha\beta} = Ct_{\alpha\beta}. \tag{11}$$

III. SYMPLECTIC STRUCTURE ASSOCIATED TO THE DYNAMICS

A. The antiquity condition

1. Definitions

(i) Consider a tensor function f on $\widehat{M}_4$. We shall say that this function approaches zero in the infinite past (resp. future), and we shall write $\lim_{x\to\epsilon\infty} f = 0$ with $\epsilon = -1$ in the past and $\epsilon = +1$ in the future, if the following is satisfied

$$\lim_{\mu\to\epsilon\infty} f(x^{\rho} + \mu\pi^{\rho},\pi^{\sigma}) = 0 \quad \forall(x,\pi)\in\widehat{M}_4. \tag{12}$$

The physical idea leading to the preceding concept is: f vanishes at $x^0\equiv t = -\infty$ (resp. $t = +\infty$) in the rest frame of the observer $\pi^{-1}\pi^{\alpha}$.

(ii) Let us assume that the tensor function f on $\widehat{M}_4$ approaches zero in the infinite past (resp. future). We shall say that the "antiquity index" of f is $s\geqslant 0$, and we shall write $\mathrm{ind}_{\epsilon}(f) = s$, if s is the smaller of the superior bounds of the numbers p such that

$$\lim_{\mu\to\epsilon\infty} \mu^pf(x^{\rho} + \mu\pi^{\rho},\pi^{\sigma}) = 0 \quad \forall(x,\pi)\in\widehat{M}_4. \tag{13}$$

The physical idea leading to this antiquity concept is the following: f decreases like t^{-s} at $t = -\infty$ (resp. $t = +\infty$) in the infinite past (resp. future), in the rest frame of the observer $\pi^{-1}\pi^{\alpha}$.

(iii) We shall say that $\vec{H}$ is a regular dynamic system in the past (resp. future) if

$$\lim_{x\to\epsilon\infty} \theta^{\alpha} = 0. \tag{14}$$

(iv) Consider a 1-form on $\widehat{M}_4$

$$\Phi = \Phi_{\alpha}dx^{\alpha} + \Phi_{\underline{\alpha}}d\pi^{\alpha} \quad (\alpha = \underline{\alpha}\ \text{numerically}) \tag{15}$$

where $\Phi\cdot$ are vector functions on $\widehat{M}_4$. We shall say that Φ is a regular 1-form in the past (resp. future), and we shall write $\lim_{x\to\epsilon\infty}(\Phi + \frac{1}{2}d\pi^2) = 0$, if

$$\lim_{x\to\epsilon\infty} \Phi_{\alpha} = \lim_{x\to\epsilon\infty} (\Phi_{\underline{\alpha}} - \pi_{\alpha}) = 0. \tag{16}$$

(v) Consider a 2-form on $\widehat{M}_4$

$$\sigma = \tfrac{1}{2}\sigma_{\alpha\beta}dx^\alpha \wedge dx^\beta + \sigma_{\alpha\underline{\beta}}dx^\alpha \wedge d\pi^\beta + \tfrac{1}{2}\sigma_{\underline{\alpha\beta}}d\pi^\alpha \wedge d\pi^\beta \quad (17)$$

where $\sigma_{\cdot\cdot}$ are tensor functions on $\widehat{M}_4$ and $\sigma_{\beta\alpha} = -\sigma_{\beta\alpha}$, $\sigma_{\underline{\alpha\beta}} = -\sigma_{\underline{\beta\alpha}}$. We shall say that σ is a regular 2-form in the past (resp. future), and we shall write $\lim_{x\to\epsilon\infty}(\sigma - dx^\alpha \wedge d\pi_\alpha) = 0$, if

$$\lim_{x\to\epsilon\infty}\sigma_{\alpha\beta} = \lim_{x\to\epsilon\infty}(\sigma_{\alpha\underline{\beta}} - \eta_{\alpha\beta}) = \lim_{x\to\epsilon\infty}\sigma_{\underline{\alpha\beta}} = 0. \quad (18)$$

2. Lemma 1

Consider the differential systems[5]

$$£(\vec{D})f = 0,\quad £(\vec{D})\Phi = 0,\quad £(\vec{D})\omega = 0 \quad (\vec{D} \equiv \pi^\rho\partial_\rho), \quad (19\text{a,b,c})$$

where f, Φ, and ω are a tensor function, a 1-form and a 2-form on $\widehat{M}_4$, respectively. Then, the general solution of (17a), (17b), or (17c) satisfying the condition $\lim_{x\to\epsilon\infty} f = 0$, $\lim_{x\to\epsilon\infty}\Phi = 0$ or $\lim_{x\to\epsilon\infty}\omega = 0$ ($\epsilon = -1$ or $+1$), respectively, is

$$f = 0,\quad \Phi = 0,\quad \text{or } \omega = 0. \quad (20\text{a,b,c})$$

Proof: (i) The general solution of the homogeneous system $\pi^\rho\partial_\rho f = 0$ is an arbitrary function of seven independent solutions, for example, (h^i, π^σ): $h^i \equiv x^i + \pi^{-2}(x\pi)\pi^i$. By considering the corresponding limit condition $0 = \lim_{x\to\epsilon\infty} f(h^i,\pi^\sigma) = f(h^i,\pi^\sigma) \Rightarrow f \equiv 0$.

(ii) Adopting the general form (15) for Φ, we can obtain from $£(\vec{D})\Phi = 0$

$$\pi^\rho\partial_\rho\Phi_\alpha = 0,\quad \pi^\rho\partial_\rho\Phi_{\underline{\alpha}} = -\Phi_\alpha.$$

Taking into account the limit condition on Φ and applying the first part of the lemma repeatedly, we obtain $\Phi \equiv 0$.

(iii) Adopting the general form (17) for ω, we can obtain from $£(\vec{D})\omega = 0$

$$\pi^\rho\partial_\rho\omega_{\alpha\beta} = 0,\quad \pi^\rho\partial_\rho\omega_{\alpha\underline{\beta}} = -\omega_{\alpha\beta},$$

$$\pi^\rho\partial_\rho\omega_{\underline{\alpha\beta}} = -\omega_{\alpha\underline{\beta}} + \omega_{\beta\underline{\alpha}}.$$

Taking into account the limit condition on ω and applying the first part of the Lemma repeatedly, we obtain $\omega \equiv 0$.

3. Regularity of the dynamics

The external electromagnetic field $F_{\alpha\beta}(x^\sigma)$ that we are considering must satisfy the Maxwell equations. Moreover, in all practical physical situations the field acts on the test charge in a certain region of space–time. We shall impose in this sense, the following weak conditions on the field and derivatives,

$$\text{ind}_\epsilon(F_{\alpha\beta}) > 1,\quad \text{ind}_\epsilon(\partial_\rho F_{\alpha\beta}) > 1 \quad (\epsilon = -1 \text{ or } +1), \quad (21)$$

that express the physical idea: the field and derivatives decrease faster than t^{-1} in the infinite past (resp. future).

If we also assume that the field derivatives of any order approach zero in the infinite past (resp. future),

$$\lim_{x\to\epsilon\infty}\partial_{\rho_1\cdots\rho_n\cdots}F_{\alpha\beta} = 0, \quad (22)$$

a close study of Eq. (8) yields $\lim_{x\to\epsilon\infty}\theta^\alpha = 0$, i.e., $\vec{H}$ is a regular dynamic system in the past (resp. future).

B. Hamiltonian forms

1. Definitions

Let us consider the dynamic system constituted by a single charged particle, whose acceleration is given by the series (8), represented by the vector field on $\widehat{M}_4$: $\vec{H} = \pi^\alpha\partial_\alpha + \theta^\alpha\partial_{\underline{\alpha}}$. We shall assume that the external field $F_{\alpha\beta}(x^\sigma)$ satisfies the weak conditions (21) and also (22) (i.e., $\vec{H}$ is a regular dynamic system in the past or in the future).

A "Hamiltonian form" associated to the dynamics θ^α is a symplectic[6] form σ on $\widehat{M}_4$ that satisfies:

(i) It is invariant under $\vec{H}$,

$$£(\vec{H})\sigma = 0; \quad (23)$$

(ii) σ can be expanded into power series of the electric charge e,

$$\sigma = \sum_{r=0}^{\infty} e^r\sigma^{(r)}, \quad (24)$$

where $\sigma^{(r)}$ are e-independent;

(iii) it is a regular 2-form in the past ($\epsilon = -1$) or in the future ($\epsilon = +1$),

$$\lim_{x\to\epsilon\infty}(\sigma - dx^\alpha \wedge d\pi_\alpha) = 0. \quad (25)$$

2. Theorem 1 (on the uniqueness of σ)

Let us assume that there exists a σ satisfying properties (23), (24), and (25) in the past case [or (25) in the future case]. Then, σ is unique.

Proof: Let us assume the existence of two symplectic forms σ_1 and σ_2 possessing properties (23)–(25). It is clear that $\omega \equiv \sigma_1 - \sigma_2$ must satisfy

$$£(\vec{H})\omega = 0,\quad \omega = \sum_{r=0}^{\infty} e^r\omega^{(r)},\quad \lim_{x\to\epsilon\infty}\omega = 0$$

$$(\epsilon = -1 \text{ or } +1).$$

By introducing the developments for $\vec{H}$ and ω into the first equation and proceeding inductively, i.e., supposing $\omega^{(p)} \equiv 0$ $\forall p < r$, one obtains up to order e^r: $£(\vec{D})\omega^{(r)} = 0$. Then taking into account the limit condition $\lim_{x\to\epsilon\infty}\omega^{(r)} = 0$, we can apply Lemma 1 to obtain $\omega^{(r)} \equiv 0$; hence we have proved the uniqueness of σ.

In the past case (resp. future case) we shall call σ the "Hamiltonian form" in the past and write σ_{-1} (resp. the "Hamiltonian form" in the future and write σ_{+1}); in general $\sigma_{-1} \neq \sigma_{+1}$.

3. Construction of σ_ϵ

Let us consider the two approximated symplectic forms

$$\sigma_\epsilon = \sigma^{(0)} - eF - \gamma e^3 G_\epsilon + O(e^4), \quad (26\text{a})$$

where

$$\sigma^{(0)} \equiv dx^\alpha \wedge d\pi_\alpha,\quad F \equiv \tfrac{1}{2}F_{\alpha\beta}(x^\sigma)dx^\alpha \wedge dx^\beta,$$

$$G_\epsilon \equiv dx^\alpha \wedge d(\pi^{-2}\pi^\rho F_{\alpha\rho}) + d\left(\pi^{-2}\pi^\rho\int_{\epsilon\infty}^{0} dy\,\tilde{F}_{\alpha\rho}\right) \wedge d\pi^\alpha, \quad (26\text{b})$$

$$\tilde{F}_{\alpha\beta}(x^\rho,\pi^\sigma;y)\equiv F_{\alpha\beta}(x^\rho+y\pi^\rho,\pi^\sigma).$$

After a lengthy but easy calculation, one can prove that σ_ϵ defined by (26), satisfies (23)–(25) identically up to order e^3. Therefore, according to Theorem 1, σ_ϵ is the "Hamiltonian form" in the past ($\epsilon=-1$) and the "Hamiltonian form" in the future ($\epsilon=+1$) up to order e^3.

Let us assume that $e=0$. In this case we have a free particle ($\theta^\alpha=0$) and $\sigma_\epsilon=\sigma^{(0)}$, which is the simplest of the symplectic structures that can be associated to that dynamics.

4. The generating function associated to $\vec{H}$

As σ satisfies (23), there exists a function H such that[5]

$$i(\vec{H})\sigma=-dH. \qquad (27)$$

H is the generating function associated to the generator $\vec{H}$ and is defined only up to an additive constant.

Theorem 2 (on the exact form of H): Let us assume that there exists a σ satisfying properties (23)–(25) and that $\vec{H}$ is a regular dynamic system in the past (or in the future). Then, the generating function associated to $\vec{H}$ by means of (27) is $H=\frac{1}{2}\pi^2+ct$.

Proof: By defining $\Phi\equiv-dH$ we have $£(\vec{H})\Phi=-d£(\vec{H})H=0$ and, if we use (15) and (17) as expressions for Φ and σ, (27) yields

$$\Phi_\alpha=\pi^\rho\sigma_{\rho\alpha}-\theta^\rho\sigma_{\alpha\rho},\quad \Phi_\alpha=\pi^\rho\sigma_{\rho\underline{\alpha}}+\theta^\rho\sigma_{\underline{\rho}\underline{\alpha}},$$

and the regularity of $\vec{H}$ and σ furnishes the limit condition $\lim_{x\to\epsilon\infty}(\Phi+\frac{1}{2}d\pi^2)=0$.

Next we shall prove the uniqueness of Φ satisfying $£(\vec{H})\Phi=0$, $\lim_{x\to\epsilon\infty}(\Phi+\frac{1}{2}d\pi^2)=0$. Let us assume the existence of two 1-forms, Φ_1 and Φ_2, possessing those two properties; then it is clear that $\psi\equiv\Phi_1-\Phi_2$ must satisfy

$$£(\vec{H})\psi=0,\quad \psi=\sum_{r=0}^{\infty}e^r\psi^{(r)},$$

$$\lim_{x\to\epsilon\infty}\psi=0\quad(\epsilon=-1\ \text{or}\ +1).$$

By introducing the developments for $\vec{H}$ and ψ into the first equation and proceeding inductively, i.e., supposing $\psi^{(p)}\equiv0$ $\forall p<r$, one obtains $£(\vec{D})\psi^{(r)}=0$ up to order e^r. Taking into account the limit condition $\lim_{x\to\epsilon\infty}\psi^{(r)}=0$, we can apply Lemma 1 to obtain $\psi^{(r)}\equiv0$; hence the uniqueness of $\Phi\equiv dH$ has been proved.

Let us consider the 1-form $\Phi=-\frac{1}{2}d\pi^2$ that obviously satisfies

$$£(\vec{H})\Phi=-\tfrac{1}{2}d£(\vec{H})\pi^2=(\pi\theta)=0$$

and

$$\lim_{x\to\epsilon\infty}(\Phi+\tfrac{1}{2}d\pi^2)=0.$$

Then, according to the uniqueness of Φ, $H=\frac{1}{2}\pi^2+ct$ is the exact analytical expression of the generating function associated to $\vec{H}$ in the two cases (past and future).

IV. SYMMETRIES AND CONSERVED QUANTITIES

A. Conseved quantities associated to isometries

Let G_r ($r\leqslant10$) be an r-parameter group of motions of M_4, i.e., each infinitesimal generator $\vec{\Xi}$ of the group leaves the Minkowski metric η invariant:

$$£(\vec{\Xi})\eta=0,\quad \vec{\Xi}=\Xi^\alpha(x^\sigma)\partial_\alpha$$
$$(\Longleftrightarrow\partial_\alpha\Xi_\beta+\partial_\beta\Xi_\alpha=0). \qquad (28)$$

Obviously G_r must be a subgroup of the Poincaré group, possibly this group itself. G_r induces a natural realization on $\widehat{M}_4$ defined by

$$\vec{\Xi}=\Xi^\alpha\partial_\alpha+\pi^\rho\partial_\rho\Xi^\alpha\partial_{\underline{\alpha}}. \qquad (29)$$

Theorem 3: Let $\vec{H}$ be a dynamical system on M_4 and σ a symplectic form on $\widehat{M}_4$ invariant under $\vec{H}$ such that the generating function associated to $\vec{H}$ by means of (27) is $H=H(\pi^2)$. Let σ be invariant under the natural realization $\vec{\Xi}$ of a group of motions of M_4. Then, the functions Ξ associated with each generator $\vec{\Xi}$ by means of $i(\vec{\Xi})\sigma=d\Xi$ are conserved quantities.

Proof: As σ is symplectic, according to the Darboux theorem,[7] coordinates (q^α,P_β) exist such that σ can be written $\sigma=dq^\alpha\wedge dP_\alpha$. Then, by developing $i(\vec{H})\sigma=-dH$ and $i(\vec{\Xi})\sigma=d\Xi$ one obtains

$$£(\vec{H})q^\alpha=-\frac{\partial H}{\partial p_\alpha},\quad £(\vec{H})p_\alpha=\frac{\partial H}{\partial q^\alpha},$$
$$£(\vec{\Xi})q^\alpha=\frac{\partial\Xi}{\partial p_\alpha},\quad £(\vec{\Xi})p_\alpha=-\frac{\partial\Xi}{\partial q^\alpha}.$$

It is clear that $£(\vec{H})\Xi=-£(\vec{\Xi})H$, and on the other hand, (26) gives

$$£(\vec{\Xi})H=\frac{\partial H}{\partial\pi^2}£(\vec{\Xi})\pi^2=-2\frac{\partial H}{\partial\pi^2}\pi_\alpha\pi^\rho\partial_\rho\Xi^\alpha=0;$$

hence $£(\vec{H})\Xi=0$, and the theorem is proved.

B. Construction of conserved quantities

Let F be an external electromagnetic field invariant under a group of motions (isometries) of M_4 generated by $\vec{\Xi}_a$ ($a=1,\ldots,r\leqslant10$):

$$£(\vec{\Xi}_a)F=0,\quad F\equiv\tfrac{1}{2}F_{\alpha\beta}(x^\sigma)dx^\alpha\wedge dx^\beta. \qquad (30)$$

As F satisfies the equation $dF=0$ there exists a 1-form A (defined only up to $A\to A+dS$) such that $F=dA$. Thus, from (30) we conclude the existence of r functions χ_a on M_4 such that

$$£(\vec{\Xi}_a)A=d\chi_a. \qquad (31)$$

$A=A_\alpha dx^\alpha$, $A_\alpha(x^\sigma)$ being the electromagnetic potential. In general, the ambiguity in A is not sufficient for the elimination of all the χ_a's, but in the particular case of $r=1$, for example, an electromagnetic potential invariant under $\vec{\Xi}$ can always be found.

Theorem 4: Let σ_ϵ be the "Hamiltonian form" in the past (resp. future) up to order e^3 given by (26). If $\vec{\Xi}_a$ are the generators of isometries of M_4 that leave the external electromagnetic field F invariant, then

$$\Xi_{a(\epsilon)} = (\Xi_a \pi) + e\{(\Xi_a A) - \chi_a\} + \gamma e^3 \pi^{-2} \pi^\rho \left\{ -\Xi_a^\alpha F_{\alpha\rho} + (\pi^\sigma \partial_\sigma \Xi_a^\alpha) \int_{\epsilon\infty}^0 dy \tilde{F}_{\alpha\rho} \right\} + O(e^4), \tag{32}$$

$$\vec{\Xi}_a = \Xi_a^\alpha \partial_\alpha,$$

are the conserved quantities associated to the symmetries $\vec{\Xi}_a$.

Proof: We can apply Theorem 3 because $i(\vec{H})\sigma_\epsilon = -d(\frac{1}{2}\pi^2)$ and $\pounds(\vec{\Xi}_a)F = 0 \Rightarrow \pounds(\vec{\Xi}_a)\sigma_\epsilon = 0$ at least up to order e^3. An easy calculation of $i(\vec{\Xi}_a)\sigma_\epsilon$ yields (32).

Note that there is no contradiction between the existence of some conserved quantity and the fact that the charged particle radiates energy. The action of this loss is to modify the equations of motion through a radiation rection term, and then the new equations of motion including this self-action admit conserved quantities that are associated with the symmetries of the external field. Of course, these conserved quantities differ from those that are obtained if the Lorentz force is the only one acting on the particle ($\Xi_a = (\Xi_a \pi + e\{(\Xi_a A) - \chi_a\})$; they include corrective terms due to the radiation reaction force. Finally, it is clear that in the particular case $F_{\alpha\beta} \equiv 0$ we reobtain the well-known expression $\Xi_a = (\Xi_a \pi)$.

[1]H.A. Lorentz, *The Theory of Electrons* (Taubner, Leipzig,, 1909), 1st ed., pp. 49, 253; P.A.M. Dirac, Proc. Roy. Soc. (London) A **167**, 148 (1938); J.A. Wheeler and R.P. Feynman, Rev. Mod. Phys. **17**, 157 (1945).

[2]W.E. Baylis and J. Huschilt, Phys. Rev. D **13**, 3237 (1976).

[3]L.D. Landau and E.M. Lifshitz, *Teoría Clásica de Campos* (Reverté, Barcelona, 1966), p. 277 (1966); V. Ginzburg, *Physique Théorique et Astrophysique* (Mir, Moscow, 1978), p. 65.

[4]$\alpha,\beta,\gamma,\rho,\sigma,\cdots = 0,1,2,3$; $i,j,k,\cdots = 1,2,3$; all indices follow the summation convention; we shall take the Minkowski metric $\eta_{\alpha\beta}$ with signature $+2$ and the units will be so chosen that the velocity of light in vacuum will be $c = 1$; $(ab) \equiv a^\rho b_\rho$; $\partial_\alpha \equiv \partial/\partial x^\alpha$, $\partial_{\underline{\alpha}} \equiv \partial/\partial \pi^\alpha$. $\hat{M}_4$ is the fiber vector bundle of all vectors π^σ tangent to M_4 such that $(\pi\pi) < 0$, $0 < \pi^0 < +\infty$.

[5]$\pounds(\vec{V})$ is the Lie derivative operator associated with $\vec{V}$ and $i(\vec{V})$ its interior product. $\pounds(\vec{V}) = i(\vec{V})d + di(\vec{V})$.

[6]A symplectic form is a 2-form on a variety of even dimension that is closed and has maximum rank.

[7]G. Darboux, Bull. Sci. Math. **6** (2), 14 (1882).

Matrix representation of the kinetic theory differential operator

Ulrich Weinert[a)]

Institute for Theoretical Physics, University of Düsseldorf, Federal Republic of Germany

(Received 27 March 1979; accepted for publication 18 May 1979)

A general scheme is given to determine the matrix representation of the differential operator of the kinetic theory of gas mixtures, i.e., the streaming operator in the one-particle phase space. The matrix elements are evaluated explicitly for the Burnett basis functions and a Lorentz type external field. The density, temperature, and mean velocity are treated as field quantities in order to assure the applicability of the results to inhomogeneous problems.

INTRODUCTION

The velocity distribution functions of the components of a gas mixture or a plasma are determined by a system of "kinetic equations," which—under certain assumptions[1]—may be considered as Boltzmann type integrodifferential equations. A generalized moment method can be introduced by representing the kinetic equations algebraically as systems of "balance equations" for the generalized moments, which similar to quantum mechanics form a vector in a function space.[2]

There are two operators involved in the system of kinetic equations, the nonlinear collision operator describing the influence of the molecular interactions, and the differential operator containing the streaming terms in the phase space under the influence of external fields. So both types of operators have to be represented algebraically. While the collision terms have been discussed extensively,[3–6] general expressions for the differential operators are rarer.[7] However, they are necessary with respect to a more general formalism for the solution of the transport equations.

In this paper, we are concerned with the procedure of evaluating the differential matrix elements under the assumption that the external force is of the Lorentz type. Our general results, given at the end of Sec. III, are suitable for treating all kinds of problems of transport theory; in Sec. IV, we restrict ourselves, however, as a special example to the case of drift tube experiments, as there are already comparable expressions for the differential matrix elements.[8]

The evaluation is based upon the choice of the Burnett basis functions, Sec. II. In comparison with other possible systems of basis functions, e.g., the Hermite functions in the tensorial[9] or scalar[10] notation, they have a rather involved algebraic structure. The diagonality properties, however, of the linearized collision matrix show the advantages of the Burnett basis functions as irreducible representations of the group of three-dimensional rotations. On the other hand, the differential matrix elements for the system of Hermite functions can easily be evaluated, since the general method, described in this paper, and their algebraic properties are given.[11]

Subsequent papers will show some special examples for the use of the differential matrix elements. Therefore, here we restrict ourselves to a purely formal deduction of the general results without further applications.

I. GENERAL FORMALISM

Consider a mixture of gaseous components of species $i, j, \cdots$, which is sufficiently dilute so that only simultaneous two-particle collisions have to be taken into account. Then the velocity distribution function $f(i)$ for a component i is determined by the "kinetic equation"

$$D(i)[f(i)] = \sum_j B(i,j)[f(i),f(j)]. \tag{1.1}$$

The bilinear collision operators $B(i,j)$ describe the influence of the molecular interactions; a special example is the Boltzmann operator. The differential operator $D(i)$ depends on the external fields and can be written as the total time derivative:

$$D(i) \equiv \frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{c}_i \cdot \frac{\partial}{\partial \mathbf{x}} + \dot{\mathbf{c}}_i \cdot \frac{\partial}{\partial \mathbf{c}_i}, \tag{1.2}$$

where we assume that the distribution function $f(i)$ depends on the particle velocity $\mathbf{c}_i$, the spatial coordinate $\mathbf{x}$, and the time t.

It is convenient to introduce a generalized moment method by passing from the kinetic equation (1.1) to its algebraic representation.[5,7] For that reason, we consider a system of basis functions $\{\varphi_\lambda(i)\}$ and expand the distribution function as

$$f(i) = a_i^{\lambda}\, \varphi_\lambda(i), \tag{1.3}$$

which determines the "generalized moments" a_i^{λ}. (We use the summation convention to sum over indices appearing once and only once as super- and subscript.) The index λ stands for an n- tuple of indices, where n is the number of variables taken into account for the algebraization. In our case [cf. Eq. (1.2)] we have $n \leqslant 7$; actually we will restrict ourselves to $n = 3$ as only the velocity dependence is algebraized. Moreover λ can contain continuous indices so that the sum of Eq. (1.3) may turn into an integral.

We introduce the system of dual basis functions $\{\varphi^{\lambda}(i)\}$ by the requirement

$$\langle \varphi^{\lambda'}(i) \mid \varphi_\lambda(i) \rangle = \delta_\lambda^{\ \lambda'}, \tag{1.4}$$

provided the "scalar product" $\langle \ | \ \rangle$ is defined. By means of the dual basis functions we express the generalized moments

[a)] On leave at Department of Chemistry, Brown University, Providence, Rhode Island, till September 1979.

as

$$a_i^{\lambda} = \langle \varphi^{\lambda}(i) \,|\, f(i) \rangle. \tag{1.5}$$

We insert the expansion (1.3) in the kinetic equation (1.1) and obtain

$$D(i)[\varphi_{\lambda}(i)a_i^{\lambda}] = \sum_j B(i,j)[\varphi_{\lambda}(i)a_i^{\lambda}; \varphi_{\mu}(j)a_j^{\mu}], \tag{1.6}$$

which in the case of $B(i,j)$ being local operators yields

$$D(i)[\varphi_{\lambda}(i)a_i^{\lambda}] = \sum_j B(i,j)[\varphi_{\lambda}(i); \varphi_{\mu}(j)]a_i^{\lambda}a_j^{\mu}, \tag{1.7}$$

because of their bilinearity. We multiply by the dual basis functions $\varphi^{\lambda'}(i)$ and obtain

$$\langle \varphi^{\lambda'}(i) \,|\, D(i)\varphi_{\lambda}(i)\rangle a_i^{\lambda} = \sum_j \langle \varphi^{\lambda'}(i) \,|\, B(i,j) \times [\varphi_{\lambda}(i),\varphi_{\mu}(j)]\rangle a_i^{\lambda}a_j^{\mu}, \tag{1.8a}$$

or abbreviated as

$$D_{\lambda}^{\lambda'}(i)a_i^{\lambda} = \sum_j B_{\lambda,\mu}^{\lambda'}(i,j)a_i^{\lambda}a_j^{\mu}. \tag{1.8b}$$

We recall the fact that in general the differential matrix $D_{\lambda}^{\lambda'}(i)$ still contains differentiations with respect to the variables, which the generalized moments depend on.

While the matrix representation of a linear operator is well known from quantum mechanics, the "collision elements" $B_{\lambda,\mu}^{\lambda'}(i,j)$, which represent the bilinear collision operator, can be considered as elements of a third-order tensor.[2]

In this paper, we are especially concerned with the differential matrix elements. Consider the differential operator on the left-hand side of the kinetic equation (1.1) acting upon the expansion (1.3). We have

$$D[f] \equiv \frac{d}{dt}\varphi_{\lambda}a^{\lambda} = a^{\lambda}\frac{d}{dt}\varphi_{\lambda} + \varphi_{\lambda}\frac{d}{dt}a^{\lambda}, \tag{1.9}$$

where we omitted the index i. Equation (1.9) yields

$$D_{\lambda}^{\lambda'}a^{\lambda} = \left\langle \varphi^{\lambda'} \,\middle|\, \frac{d}{dt}\varphi_{\lambda}\right\rangle a^{\lambda} + \left\langle \varphi^{\lambda'} \,\middle|\, \varphi_{\lambda}\frac{d}{dt}\right\rangle a^{\lambda}, \tag{1.10}$$

showing the separation of the differential matrix into a purely multiplicative matrix, and a part which effects additional differentiations of the generalized moments. Further evaluations require the specification of the basis functions.

In the following section we omit the particle species index i, because it is redundant in the context of this paper.

II. BURNETT BASIS FUNCTIONS

We choose a space- and time-dependent vector field $\mathbf{c}_0(\mathbf{x},t)$, the "reference velocity," and a temperature field $T(\mathbf{x},t)$, and define the "peculiar velocity" $\mathbf{C}$[12] shifting the particle velocity $\mathbf{c}$ by $\mathbf{c}_0$

$$\mathbf{C} := \mathbf{c} - \mathbf{c}_0. \tag{2.1}$$

By means of the temperature field the peculiar velocity can be normalized to the physically dimensionless quantity

$$\check{\mathbf{C}} := \mathbf{C}/\gamma \tag{2.2a}$$

with

$$\gamma = (k_B T/\mu)^{1/2}, \tag{2.2b}$$

μ being the mass of a particle of the species under consideration. To facilitate the notation we define a normalized kinetic energy

$$\epsilon := \tfrac{1}{2}\check{C}^2. \tag{2.3}$$

Then the special system of basis functions that we want to use can be written as

$$\varphi_{\lambda} \equiv \Phi_{n,l,m} := (-1)^{(n-l)/2} 2^{(l+1)/2} \frac{(2l+1)^{1/2}}{2\pi} \frac{1}{(n+l+1)!!} \times \gamma^{-n-3}\epsilon^{l/2} L_{(n-l)/2}^{(l+1/2)}(\epsilon) e^{-\epsilon} Y_{l,m}(\vartheta,\psi). \tag{2.4}$$

The spherical harmonics $Y_{l,m}$ depend on the direction (ϑ,ψ) of the peculiar velocity and are defined in the quantum mechanical way as the eigenfunctions of the squared angular momentum operator and its projection onto the 3-axis.[13]

We define the "scalar product" as

$$\langle \;|\; \rangle := \int d^3c \equiv \int d^3C \tag{2.5}$$

and obtain, by means of the orthogonality relations of the Laguerre polynomials[14] $L_{(n-l)/2}^{(l+1/2)}(\epsilon)$ and the spherical harmonics,

$$\varphi^{\lambda} \equiv \Phi^{n,l,m} := (-1)^{(n-l)/2} 2^{(l+1)/2} (n-l)!! \left(\frac{2\pi}{2l+1}\right)^{1/2} \times \gamma^{n}\epsilon^{l/2} L_{(n-l)/2}^{(l+1/2)}(\epsilon) Y_{l,m}^{*}(\vartheta,\psi), \tag{2.6}$$

which indeed yields the duality relation

$$\langle \Phi^{n',l',m'} \,|\, \Phi_{n,l,m}\rangle = \delta_n^{n'}\delta_l^{l'}\delta_m^{m'}. \tag{2.7}$$

The generalized moments

$$A^{n,l,m} = \langle \Phi^{n,l,m} \,|\, f\rangle \tag{2.8}$$

are the spherical components of a completely symmetric nth order tensor.[15] Consistently the definition (2.4) yields the restriction

$$0 \leqslant n - l \quad \text{even}, \tag{2.9}$$

and the spherical properties imply the familiar inequality

$$-l \leqslant m \leqslant l. \tag{2.10}$$

For $n' = l' = 0$ we have $\Phi^{0,0,0} = 1$, hence the only nonvanishing zeroth-order moment is

$$A^{0,0,0} = \langle 1 \,|\, f\rangle = n(\mathbf{x},t), \tag{2.11}$$

provided the distribution function is normalized to the particle density $n(\mathbf{x},t)$. For $n' = 1$ we obtain, with

$$\Phi^{1,1,m} = \gamma(\tfrac{8}{3}\pi\epsilon)^{1/2} Y_{1,m}^{*}(\vartheta,\psi) \tag{2.12}$$

the spherical components of the diffusion flux vector $\mathbf{J}$,

$$A^{1,1,m} = \langle C^m \,|\, f\rangle = J^m. \tag{2.13}$$

For $n' = 2$, $l' = 0$ we have

$$\Phi^{2,0,0} = -2\gamma^2 L_1^{(1/2)}(\epsilon) = C^2 - 3\gamma^2. \tag{2.14}$$

The related generalized moment is

$$A^{2,0,0} = \langle C^2 - 3k_B T/\mu \,|\, f\rangle. \tag{2.15}$$

If we specify the temperature by the definition

$$\tfrac{1}{2}\mu C^2 =: \tfrac{3}{2}k_B T, \tag{2.16}$$

we obtain in equivalence to the ideal gas law

$$A^{2,0,0} = 0. \tag{2.17}$$

For further discussions of the generalized moments related to our Burnett basis functions, we refer to a previous paper.[15]

III. EVALUATION OF THE DIFFERENTIAL MATRIX ELEMENTS

The Burnett basis functions (2.4) actually depend upon the normalized peculiar velocity $\check{\mathbf{C}}$ and the temperature factor γ. So the total time derivative is

$$\frac{d}{dt}\Phi_{n,l,m} = \frac{d\gamma}{dt}\frac{\partial}{\partial\gamma}\Phi_{n,l,m} + \left(\frac{d}{dt}\check{\mathbf{C}}\right)\cdot\frac{\partial}{\partial\check{\mathbf{C}}}\Phi_{n,l,m}. \tag{3.1}$$

According to Eq. (1.2) the total time derivative on the right-hand side can be written as

$$\frac{d}{dt} = \frac{\partial}{\partial t} + (\mathbf{C}+\mathbf{c}_0)\cdot\frac{\partial}{\partial\mathbf{x}} + [\boldsymbol{\psi} + \boldsymbol{\Omega}\times(\mathbf{C}+\mathbf{c}_0)]\cdot\frac{\partial}{\partial\mathbf{c}}, \tag{3.2}$$

provided a particle of the species under consideration is accelerated by a Lorentz type force

$$\mu\dot{\mathbf{c}} = \mu\boldsymbol{\psi}(\mathbf{x},t) + \mu\boldsymbol{\Omega}(\mathbf{x},t)\times\mathbf{c}, \tag{3.3}$$

where in general $\boldsymbol{\psi}(\mathbf{x},t)$ and $\boldsymbol{\Omega}(\mathbf{x},t)$ are space- and time-dependenct vector fields.

For the further evaluation of Eq. (3.1), we write

$$\frac{d}{dt}\check{\mathbf{C}} = \frac{d}{dt}\frac{1}{\gamma}\mathbf{C} = \frac{1}{\gamma}\frac{d}{dt}\mathbf{C} - \frac{1}{2}\check{\mathbf{C}}\frac{d}{dt}\ln T, \tag{3.4}$$

where we used the definition (2.2b), and obtain, by means of the relation

$$\frac{d}{dt}\mathbf{C} = -\left(\frac{\partial}{\partial t} + (\mathbf{C}+\mathbf{c}_0)\cdot\frac{\partial}{\partial\mathbf{x}}\right)\mathbf{c}_0 + [\boldsymbol{\psi} + \boldsymbol{\Omega}\times(\mathbf{C}+\mathbf{c}_0)], \tag{3.5}$$

the total time derivative of the basis functions,

$$\begin{aligned}\frac{d}{dt}\Phi_{n,l,m} = &-\tfrac{1}{2}(n+3)\Phi_{n,l,m}\left(\frac{\partial\ln T}{\partial t} + (\mathbf{C}+\mathbf{c}_0)\cdot\frac{\partial\ln T}{\partial\mathbf{x}}\right) - \left(\frac{\partial}{\partial t}\mathbf{c}_0 + (\mathbf{C}+\mathbf{c}_0)\cdot\frac{\partial}{\partial\mathbf{x}}\mathbf{c}_0\right)\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\\ &+ [\boldsymbol{\psi} + \boldsymbol{\Omega}\times(\mathbf{C}+\mathbf{c}_0)]\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m} - \frac{1}{2}\left(\frac{\partial\ln T}{\partial t} + (\mathbf{C}+\mathbf{c}_0)\cdot\frac{\partial\ln T}{\partial\mathbf{x}}\right)\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}.\end{aligned} \tag{3.6}$$

We insert this relation in the general expression (1.10) and obtain

$$\begin{aligned}D^{n',l',m'}_{n,l,m}A^{n,l,m} = &\left(\frac{\partial}{\partial t} + \mathbf{c}_0\cdot\frac{\partial}{\partial\mathbf{x}}\right)A^{n',l',m'} + \langle\Phi^{n',l',m'}|\mathbf{C}\Phi_{n,l,m}\rangle\cdot\frac{\partial}{\partial\mathbf{x}}A^{n,l,m} - \tfrac{1}{2}(n'+3)\\ &\times\left(\frac{\partial\ln T}{\partial t} + \mathbf{c}_0\cdot\frac{\partial\ln T}{\partial\mathbf{x}}\right)A^{n',l',m'} - \tfrac{1}{2}\sum_n (n+3)\langle\Phi^{n',l',m'}|\mathbf{C}\Phi_{n,l,m}\rangle\cdot\frac{\partial\ln T}{\partial\mathbf{x}}A^{n,l,m}\\ &-\left(\frac{\partial}{\partial t} + \mathbf{c}_0\cdot\frac{\partial}{\partial\mathbf{x}}\right)\mathbf{c}_0\cdot\left\langle\Phi^{n',l',m'}\middle|\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m} - \left(\frac{\partial}{\partial\mathbf{x}}\mathbf{c}_0\right)^{\sim}:\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m}\\ &+(\boldsymbol{\psi} + \boldsymbol{\Omega}\times\mathbf{c}_0)\cdot\left\langle\Phi^{n',l',m'}\middle|\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m} + \boldsymbol{\Omega}\times\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m}\\ &-\frac{1}{2}\left(\frac{\partial\ln T}{\partial t} + \mathbf{c}_0\cdot\frac{\partial\ln T}{\partial\mathbf{x}}\right)\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m} - \frac{1}{2}\frac{\partial\ln T}{\partial\mathbf{x}}\cdot\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m},\end{aligned} \tag{3.7}$$

where the tilde indicates that the tensor has to be transposed before the scalar multiplication. We use the identity

$$\boldsymbol{\Omega}\times\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}} = \boldsymbol{\Omega}\cdot\mathbf{C}\times\frac{\partial}{\partial\mathbf{C}} \tag{3.8}$$

and define, in analogy to the angular momentum operator of quantum mechanics,

$$\mathbf{L} := -i\mathbf{C}\times\frac{\partial}{\partial\mathbf{C}}, \tag{3.9}$$

which enables us to write

$$\boldsymbol{\Omega}\times\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle A^{n,l,m} = i\boldsymbol{\Omega}\cdot\langle\Phi^{n',l',m'}|\mathbf{L}\Phi_{n,l,m}\rangle A^{n,l,m}. \tag{3.10}$$

Now we are left with a set of "elementary" matrix elements,

$$\langle\Phi^{n',l',m'}|\mathbf{C}\Phi_{n,l,m}\rangle, \tag{3.11a}$$

$$\left\langle \Phi^{n',l',m'} \middle| \frac{\partial}{\partial \mathbf{C}} \Phi_{n,l,m} \right\rangle, \tag{3.11b}$$

$$\left\langle \Phi^{n',l',m'} \middle| \mathbf{C}\frac{\partial}{\partial \mathbf{C}} \Phi_{n,l,m} \right\rangle, \tag{3.11c}$$

$$\langle \Phi^{n',l',m'} | \mathbf{L}\Phi_{n,l,m} \rangle, \tag{3.11d}$$

$$\left\langle \Phi^{n',l',m'} \middle| \mathbf{C}\cdot\frac{\partial}{\partial \mathbf{C}} \Phi_{n,l,m} \right\rangle, \tag{3.11e}$$

and

$$\left\langle \Phi^{n',l',m'} \middle| \mathbf{CC}\cdot\frac{\partial}{\partial \mathbf{C}} \Phi_{n,l,m} \right\rangle. \tag{3.11f}$$

The most convenient way to evaluate these matrix elements is to express the scalar products, which each of them is part of, in terms of spherical tensors, i.e., the matrix elements (3.11a)–(3.11f) are represented in their spherical components. If this has been done, we can choose the simplest one for explicit evaluation and then use the Wigner–Eckart theorem to express the whole set.

To begin with the simplest nonscalar case, we discuss the right-hand side of Eq. (3.10) or the matrix elements (3.11d) of the operator **L**. Because of the eigenvalue equations of the scalar spherical harmonics we have

$$L_3 \Phi_{n,l,m} = m\Phi_{n,l,m}. \tag{3.12}$$

Expressing the scalar product of $\mathbf{\Omega}$ and **L** in terms of their spherical components we have

$$\mathbf{\Omega}\cdot\mathbf{L} = \Omega^M L_M \tag{3.13a}$$

or

$$i\mathbf{\Omega}\cdot\langle \Phi^{n',l',m'} | \mathbf{L}\Phi_{n,l,m} \rangle = i\Omega^M \langle \Phi^{n',l',m'} | L_M \Phi_{n,l,m} \rangle, \tag{3.13b}$$

respectively. Now we can apply the Wigner–Eckart theorem and obtain

$$\langle \Phi^{n',l',m'} | L_M \Phi_{n,l,m} \rangle = \mathscr{L}_{n,l}^{n',l'} \langle l',m' \mid 1,M;l,m\rangle, \tag{3.14}$$

where because of $L_{M=0} = L_3$ the "reduced matrix element" $\mathscr{L}_{n,l}^{n',l'}$ can be obtained by choosing $M = 0$ and for example $m = m' = 1$:

$$\mathscr{L}_{n,l}^{n',l'} \langle l',1 \mid 1,0;l,1\rangle = \delta_l^{\,l'} \delta_n^{\,n'}. \tag{3.15}$$

Thus Eq. (3.13b) becomes

$$i\mathbf{\Omega}\cdot\langle \Phi^{n',l',m'} | \mathbf{L}\Phi_{n,l,m} \rangle = i\Omega^M \delta_l^{\,l'} \delta_n^{\,n'} \frac{\langle l',m' \mid 1,M;l,m\rangle}{\langle l',1 \mid 1,0;l',1\rangle}. \tag{3.16}$$

For the evaluation of the other elementary matrix elements we need the following algebraic relations for the Burnett basis functions:

$$C_3 \Phi_{n,l,0} = \frac{l+1}{2l+3}\left[(n+l+3)\gamma^2 \Phi_{n+1,l+1,0} + \Phi_{n-1,l+1,0}\right] + \frac{l}{2l-1}\left[(n-l+2)\gamma^2 \Phi_{n+1,l-1,0} + \Phi_{n-1,l-1,0}\right], \tag{3.17}$$

$$\frac{\partial}{\partial C_3}\Phi_{n,l,0} = -\frac{l+1}{2l+3}(n+l+3)\Phi_{n+1,l+1,0} - \frac{l}{2l-1}(n-l+2)\Phi_{n+1,l-1,0}, \tag{3.18}$$

$$\mathbf{C}\cdot\frac{\partial}{\partial \mathbf{C}}\Phi_{n,l,m} = -(n+3)\Phi_{n,l,m} - (n+l+3)(n-l+2)\gamma^2 \Phi_{n+2,l,m}, \tag{3.19}$$

and

$$\begin{aligned}\left(C_3\frac{\partial}{\partial C_3} - \frac{1}{3}\mathbf{C}\cdot\frac{\partial}{\partial \mathbf{C}}\right)\Phi_{n,l,0} = &-\frac{(l+2)(l+1)}{(2l+3)(2l+5)}(n+l+3)\left[(n+l+5)\gamma^2\Phi_{n+2,l+2,0} + \Phi_{n,l+2,0}\right] \\ &-\frac{2}{3}\frac{l(l+1)}{(2l+3)(2l-1)}\left[(n+l+3)(n-l+2)\gamma^2\Phi_{n+2,l,0} + (n+\tfrac{3}{2})\Phi_{n,l,0}\right] \\ &-\frac{l(l-1)}{(2l-1)(2l-3)}(n-l+2)\left[(n-l+4)\gamma^2\Phi_{n+2,l-2,0} + \Phi_{n,l-2,0}\right].\end{aligned} \tag{3.20}$$

These relations can be deduced straightforwardly by means of several algebraic properties of the generalized Laguerre and Legendre polynomials.[16]

By means of Eq. (3.17) we have for the spherical components of the matrix elements (3.11a):

$$\langle \Phi^{n',l',m'} | C_M \Phi_{n,l,m} \rangle = \frac{\langle l',m' | 1,M;l,m\rangle}{\langle l',0 | 1,0;l,0\rangle}\left[\frac{l'}{2l'+1}\delta_l^{\,l'-1}\{(n'+l'+1)\gamma^2\delta_n^{\,n'-1} + \delta_n^{\,n'+1}\}\right.$$

$$+\frac{l'+1}{2l'+1}\delta_l^{\,l'+1}\{n'-l')\gamma^2\delta_n^{\,n'-1}+\delta_n^{\,n'+1}\}\Big]. \tag{3.21}$$

Here we again used the Wigner–Eckart theorem. Similarly, Eq. (3.18) yields the following expressions for the matrix elements (3.11b):

$$\left\langle\Phi^{n',l',m'}\middle|\frac{\partial}{\partial C^M}\Phi_{n,l,m}\right\rangle$$
$$=-\delta_n^{\,n'-1}\frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\left(\frac{l'}{2l'+1}(n'+l'+1)\delta_l^{\,l'-1}+\frac{l'+1}{2l'+1}(n'-l')\,\delta_l^{\,l'+1}\right). \tag{3.22}$$

Equation (3.19) yields the scalar matrix elements (3.11e),

$$\left\langle\Phi^{n',l',m'}\middle|\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}\Phi_{n,l,m}\right\rangle=-\delta_l^{\,l'}\delta_m^{\,m'}[(n'+l'+1)(n'-l')\gamma^2\delta_n^{\,n'-2}+(n'+3)\delta_n^{\,n'}], \tag{3.23}$$

and by straightforward application of the relations (3.19) and (3.17) we obtain

$$\left\langle\Phi^{n',l',m'}\middle|C_M\left(\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}+n+3\right)\Phi_{n,l,m}\right\rangle$$
$$=-\gamma^2\frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\left(\frac{l'}{2l'+1}(n'+l'+1)\delta_l^{\,l'-1}\{(n'+l'-1)(n'-l')\gamma^2\delta_n^{\,n'-3}+(n'-l'+2)\delta_n^{\,n'-1}\}\right.$$
$$\left.+\frac{l'+1}{2l'+1}(n'-l')\delta_l^{\,l'+1}\{(n'+l'+1)(n'-l'-2)\gamma^2\delta_n^{\,n'-3}+(n'+l'+3)\delta_n^{\,n'-1}\}\right). \tag{3.24}$$

Now we are left with the matrix elements (3.11c); their evaluation requires a short discussion, because they are elements of a second-order tensor, which has to be decomposed into its parts of rank 0 (trace), rank 1 (antisymmetric part), and rank 2 (symmetric and traceless part).[17] Consider the double scalar product of two dyads **ab** and **cd**. Denoting by "$\frown 0$" the symmetrized and traceless tensor we can write

$$\mathbf{ab}:\mathbf{cd}=\overset{\frown 0}{(\mathbf{ab})}:\overset{\frown 0}{(\mathbf{cd})}-\tfrac{1}{2}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})+\tfrac{1}{3}\,\mathbf{a}\cdot\mathbf{b}\;\mathbf{c}\cdot\mathbf{d}. \tag{3.25}$$

Identifying **ab** with

$$\left(\frac{\partial}{\partial\mathbf{x}}\mathbf{c}_0\right)$$

and **cd** with

$$\mathbf{C}\frac{\partial}{\partial\mathbf{C}}$$

we thus have

$$\left(\frac{\partial}{\partial\mathbf{x}}\mathbf{c}_0\right)^{\sim}:\mathbf{C}\frac{\partial}{\partial\mathbf{C}}=\left(\frac{\partial}{\partial\mathbf{x}}\mathbf{c}_0\right)^{\frown 0}:\left(\mathbf{C}\frac{\partial}{\partial\mathbf{C}}\right)^{\frown 0}+\tfrac{1}{2}i(\mathrm{curl}\,\mathbf{c}_0)\cdot\mathbf{L}+\tfrac{1}{3}(\mathrm{div}\,\mathbf{c}_0)\mathbf{C}\cdot\frac{\partial}{\partial\mathbf{C}}, \tag{3.26}$$

where we used Eq. (3.9). Hence the matrix elements (3.11c) consist of the three parts mentioned above. The parts of rank 0 and of rank 1 are already known, so we have to evaluate the rank 2 part, which may be represented in terms of its spherical components

$$\left\langle\Phi^{n',l',m'}\middle|\left[\mathbf{C}\frac{\partial}{\partial\mathbf{C}}\right]_{2,M}\Phi_{n,l,m}\right\rangle$$
$$=-\frac{\langle l',m'|2,M;l,m\rangle}{\langle l',0|2,0;l,0\rangle}\left[\frac{l'(l'-1)}{(2l'-1)(2l'+1)}(n'+l'+1)\delta_l^{\,l'-2}\{(n'+l'-1)\gamma^2\delta_n^{\,n'-2}+\delta_n^{\,n'}\}\right.$$
$$+\frac{2}{3}\frac{l'(l'+1)}{(2l'+3)(2l'-1)}\delta_l^{\,l'}\{(n'+l'+1)(n'-l')\gamma^2\delta_n^{\,n'-2}+(n'+\tfrac{3}{2})\delta_n^{\,n'}\}$$
$$\left.+(n'-l')\frac{(l'+2)(l'+1)}{(2l'+3)(2l'+1)}\delta_l^{\,l'+2}\{(n'-l'-2)\gamma^2\delta_n^{\,n'-2}+\delta_n^{\,n'}\}\right], \tag{3.27}$$

where we made use of the algebraic relation (3.20).

As now all the elementary matrix elements are evaluated, we can proceed to write down the general result for the differential matrix by substituting in Eq. (3.7):

$$D_{n,l,m}^{n',l',m'}A^{n,l,m}=\left(\frac{\partial}{\partial t}+\mathbf{c}_0\cdot\frac{\partial}{\partial\mathbf{x}}\right)A^{n',l',m'}+\tfrac{1}{3}(n'+3)(\mathrm{div}\,\mathbf{c}_0)A^{n',l',m'}+\frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\left[\frac{l'}{2l'+1}\delta_l^{\,l'-1}\right.$$
$$\left.\times\left\{(n'+l'+1)\frac{k_BT}{\mu}\delta_n^{\,n'-1}+\delta_n^{\,n'+1}\right\}+\frac{l'+1}{2l'+1}\delta_l^{\,l'+1}\left\{(n'-l')\frac{k_BT}{\mu}\delta_n^{\,n'-1}+\delta_n^{\,n'+1}\right\}\right]$$

$$\times\left(\frac{\partial}{\partial \mathbf{x}}\right)^M A^{n,l,m} + \frac{1}{2}\frac{k_BT}{\mu}\left(\frac{\partial \ln T}{\partial \mathbf{x}}\right)^M \frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\left[(n'+l'+1)\frac{l'}{2l'+1}\delta_l^{\,l'-1}\right.$$
$$\times\left\{\frac{k_BT}{\mu}(n'+l'-1)(n'-l')\delta_n^{\,n'-3} + (n'-l'+2)\delta_n^{\,n'-1}\right\} + (n'-l')\frac{l'+1}{2l'+1}\delta_l^{\,l'+1}$$
$$\left.\times\left\{\frac{k_BT}{\mu}(n'+l'+1)(n'-l'-2)\delta_n^{\,n'-3} + (n'+l'+3)\delta_n^{\,n'-1}\right\}\right]A^{n,l,m}$$
$$+\left[\left(\frac{\partial}{\partial t} + \mathbf{c}_0\cdot\frac{\partial}{\partial \mathbf{x}}\right)\mathbf{c}_0 - (\mathbf{\Psi} + \mathbf{\Omega}\times\mathbf{c}_0)\right]^M \frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\delta_n^{\,n'-1}\left[\frac{l'}{2l'+1}(n'+l'+1)\delta_l^{\,l'-1}\right.$$
$$\left.+\frac{l'+1}{2l'+1}(n'-l')\delta_l^{\,l'+1}\right]A^{n,l,m} + i[\mathbf{\Omega} - \tfrac{1}{2}\,\mathrm{curl}\mathbf{c}_0]^M\delta_n^{\,n'}\delta_l^{\,l'}\frac{\langle l',m'|1,M;l',m'\rangle}{\langle l',0|1,0;l',0\rangle}A^{n,l,m}$$
$$+\left[\frac{1}{2}\left(\frac{\partial \ln T}{(t} + \mathbf{c}_0\cdot\frac{\partial \ln T}{\partial \mathbf{x}}\right) + \frac{1}{3}\,\mathrm{div}\mathbf{c}_0\right]\frac{k_BT}{\mu}(n'+l'+1)(n'-l')A^{n'-2,l',m'} + \left[\frac{\partial}{\partial \mathbf{x}}\mathbf{c}_0\right]^{2,M}$$
$$\times\frac{\langle l',m'|2,M;l,m\rangle}{\langle l',0|2,0;l,0\rangle}\left[(n'+l'+1)\frac{l'(l'-1)}{(2l'-1)(2l'+1)}\delta_l^{\,l'-2}\left\{\frac{k_BT}{\mu}(n'+l'-1)\delta_n^{\,n'-2} + \delta_n^{\,n'}\right\}\right.$$
$$+\frac{2}{3}\frac{l'(l'+1)}{(2l'-1)(2l'+3)}\delta_l^{\,l'}\left\{\frac{k_BT}{\mu}(n'+l'+1)(n'-l')\delta_n^{\,n'-2} + (n'+\tfrac{3}{2})\delta_n^{\,n'}\right\}$$
$$\left.+(n'-l')\frac{(l'+2)(l'+1)}{(2l'+1)(2l'+3)}\delta_l^{\,l'+2}\left\{\frac{k_BT}{\mu}(n'-l'+2)\delta_n^{\,n'-2} + \delta_n^{\,n'}\right\}\right]A^{n,l,m}. \tag{3.28}$$

IV. DISCUSSION

The general result (3.28) includes all kinds of applications of the kinetic theory. The discussion therefore requires a restriction to a special field. So let us consider a system that has already reached a steady state. Moreover let $\mathbf{c}_0 \equiv 0$. Then Eq. (3.28) is reduced to

$$D_{n,l,m}^{n',l',m'}A^{n,l,m} = \frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\left[\frac{l'}{2l'+1}\delta_l^{\,l'-1}\left\{(n'+l'+1)\frac{k_BT}{\mu}\delta_n^{\,n'-1} + \delta_n^{\,n'+1}\right\} + \frac{l'+1}{2l'+1}\delta_l^{\,l'+1}\right.$$
$$\left.\times\left\{(n'-l')\frac{k_BT}{\mu}\delta_n^{\,n'-1} + \delta_n^{\,n'+1}\right\}\right]\left(\frac{\partial}{\partial \mathbf{x}}\right)^M A^{n,l,m} + \frac{1}{2}\frac{k_BT}{\mu}\left(\frac{\partial \ln T}{\partial \mathbf{x}}\right)^M\frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}$$
$$\times\left[(n'+l'+1)\frac{l'}{2l'+1}\delta_l^{\,l'-1}\left\{(n'+l'-1)(n'-l')\frac{k_BT}{\mu}\delta_n^{\,n'-3} + (n'-l'+2)\delta_n^{\,n'-1}\right\}\right.$$
$$\left.+(n'-l')\frac{l'+1}{2l'+1}\delta_l^{\,l'+1}\left\{(n'+l'+1)(n'-l'-2)\frac{k_BT}{\mu}\delta_n^{\,n'-3} + (n'+l'+3)\delta_n^{\,n'-1}\right\}\right]A^{n,l,m}$$
$$-\Psi^M\frac{\langle l',m'|1,M;l,m\rangle}{\langle l',0|1,0;l,0\rangle}\delta_n^{\,n'-1}\left[(n'+l'+1)\frac{l'}{2l'+1}\delta_l^{\,l'-1} + (n'-l')\frac{l'+1}{2l'+1}\delta_l^{\,l'+1}\right]A^{n,l,m}$$
$$+i\Omega^M\frac{\langle l',m'|1,M;l',m'\rangle}{\langle l',1|1,0;l',1\rangle}A^{n',l',m'}. \tag{4.1}$$

If the system is completely homogeneous, we are left with

$$D_{n,l,m}^{n',l',m'}A^{n,l,m} = -\Psi^M\,{}^+C_{M,m}^{\,m'}(l')(n'-l')\frac{l'+1}{2l'+1}A^{n'-1,l'+1,m}$$
$$-\Psi^M\,{}^-C_{M,m}^{\,m'}(l')(n'+l'+1)\frac{l'}{2l'+1}A^{n'-1,l'-1,m} + i\Omega^M\,{}^0C_{M,m}^{\,m'}(l')A^{n',l',m}, \tag{4.2}$$

where the ratios of Clebsch Gordan coefficients, which can be taken from standard literature,[18] have been abbreviated by the coefficients C, which are listed in the following table.

TABLE I. Ratios of the Clebsch Gordan coefficients for $l = l' \pm 1$ (${}^{\pm}C$) and $l = l'$ (${}^0C$)

M	${}^-C_{M,m}^{\,m'}(l')$	${}^0C_{M,m}^{\,m'}(l')$	${}^+C_{M,m}^{\,m'}(l')$
-1	$\frac{1}{l'}\left(\frac{(l'-m')(l'-m'-1)}{2}\right)^{1/2}$	$\left(\frac{(l'-m')(l'+m'+1)}{2}\right)^{1/2}$	$-\frac{1}{l'+1}\left(\frac{(l'+m'+1)(l'+m'+2)}{2}\right)^{1/2}$
0	$\frac{1}{l'}(l'^2-m'^2)^{1/2}$	m'	$\frac{1}{l'+1}((l'+1)^2-m'^2)^{1/2}$
$+1$	$\frac{1}{l'}\left(\frac{(l'+m')(l'+m'-1)}{2}\right)^{1/2}$	$-\left(\frac{(l'+m')(l'-m'+1)}{2}\right)^{1/2}$	$\frac{-1}{l'+1}\left(\frac{(l'-m'+1)(l'-m'+2)}{2}\right)^{1/2}$

Taking the 3-axis along the field $\boldsymbol{\Psi}$,

$$\Psi^{M} = \Psi\delta_0{}^{M}, \tag{4.3}$$

omitting the gyrofield $\boldsymbol{\Omega}$, and restricting ourselves to the homogeneous moments

$$A^{n,l}_{0} := A^{n,l,0}, \tag{4.4}$$

we obtain the simple differential term

$$D^{n',l',m'}_{n,l,m}A^{n,l,m} = -\delta^{m'}_{0}\Psi\left[\frac{l'}{2l'+1}(n'+l'+1)A^{n'-1,l'-1}_{0} + \frac{l'+1}{2l'+1}(n'-l')A^{n'-1,l'+1}_{0}\right], \tag{4.5}$$

which has been used already.[8]

We can take into account an inhomogeneity caused by the particle density $n(\mathbf{x})$ as a function of the spatial coordinate. The ansatz

$$A^{n,l,m}(\mathbf{x}) = n(\mathbf{x})a^{n,l,m} \tag{4.6}$$

enables us to rewrite the expression (4.1) with a vanishing temperature gradient and without the gyrofield as

$$\begin{aligned} d^{n',l',m'}_{n,l,m}a^{n,l,m} = & \left[\frac{l'+1}{2l'+1}\;{}^{+}C^{m'}_{M,m}(l')\left\{\frac{k_BT}{\mu}(n'-l')a^{n'-1,l'+1,m} + a^{n'+1,l'+1,m}\right\}\right. \\ & \left. + \frac{l'}{2l'+1}\;{}^{-}C^{m'}_{M,m}(l')\left\{\frac{k_BT}{\mu}(n'+l'+1)a^{n'-1,l'-1,m} + a^{n'+1,l'-1,m}\right\}\right]\left(\frac{\partial \ln n}{\partial \mathbf{x}}\right)^{M} \\ & - \Psi^{M}\left(\frac{l'+1}{2l'+1}\;{}^{+}C^{m'}_{M,m}(l')(n'-l')a^{n'-1,l'+1,m} + \frac{l'}{2l'+1}\;{}^{-}C^{m'}_{M,m}(l')(n'+l'+1)a^{n'-1,l'-1,m}\right) \end{aligned} \tag{4.7}$$

with the abbreviation

$$d^{n',l',m'}_{n,l,m} := n^{-1}(\mathbf{x})D^{n',l',m'}_{n,l,m}n(\mathbf{x}). \tag{4.8}$$

Finally we discuss the normalization of the differential term (4.7) to dimensionless quantities. By means of the ansatz

$$a^{n,l,m} =: (k_BT/\mu)^{n/2}\check{a}^{n,l,m} \tag{4.9}$$

and the abbreviation

$$\check{d}^{n',l',m'}_{n,l,m} = (k_BT/\mu)^{-n'/2}n^{-1}(\mathbf{x})D^{n',l',m'}_{n,l,m}(k_BT/\mu)^{n/2}n(\mathbf{x}), \tag{4.10}$$

the expression (4.7) can be written as

$$\begin{aligned} \check{d}^{n',l',m'}_{n,l,m}\check{a}^{n,l,m} = & \left(\frac{k_BT}{\mu}\right)^{1/2}\left(\frac{\partial \ln n(\mathbf{x})}{\partial \mathbf{x}}\right)^{M}\left[\frac{l'+1}{2l'+1}\;{}^{+}C^{m'}_{M,m}(l')\{(n'-l')\check{a}^{n'-1,l'+1,m} + \check{a}^{n'+1,l'+1,m}\}\right. \\ & \left. + \frac{l'}{2l'+1}\;{}^{-}C^{m'}_{M,m}(l')\{(n'+l'+1)\check{a}^{n'-1,l'-1,m} + \check{a}^{n'+1,l'-1,m}\}\right] - \left(\frac{\mu}{k_BT}\right)^{1/2}\Psi^{M} \\ & \times\left[\frac{l'+1}{2l'+1}\;{}^{+}C^{m'}_{M,m}(l')(n'-l')\check{a}^{n'-1,l'+1,m} + \frac{l'}{2l'+1}\;{}^{-}C^{m'}_{M,m}(l')(n'+l'+1)\check{a}^{n'-1,l'-1,m}\right]. \end{aligned} \tag{4.11}$$

The matrix elements involved in this expression are frequencies, as is the omitted gyrofield Ω^{M}. On the other hand the collision terms [cf., Eq. (1.8b)] balancing the expression (4.11), can be written as

$$\sum_{j;n,N}(k_BT_i/m_i)^{(n-n')/2}(k_BT_j/m_j)^{N/2}n_j(\mathbf{x})B^{n',l',m'}_{n,l,m;N,L,M}(i,j)\check{a}^{n,l,m}_{i}\check{a}^{N,L,M}_{j}, \tag{4.12}$$

where now the particle species indices i,j had to be taken into account again. The collision elements B in (4.12) are expressible as "reduced collision elements" multiplied by Clebsch Gordan coefficients, which determine the dependence on m, m' and M.[19,20] So the expression (4.12) can be written in terms of collision frequencies

$$\omega^{n',l'}_{n,l;N,L}(i,j) := (k_BT_i/m_i)^{(n-n')/2}(k_BT_j/m_j)^{N/2}n_j(\mathbf{x})B^{n',l',0}_{n,l,0;N,L,0} \tag{4.13}$$

as

$$\sum_{j;l,L}\omega^{n',l'}_{n,l;N,L}(i,j)\frac{\langle l',m'|l,m;L,M\rangle}{\langle l',0|l,0;L,0\rangle}\check{a}^{n,l,m}_{i}\check{a}^{N,L,M}_{j}. \tag{4.14}$$

In the special case of the drift-tube situation the small ratio $n_i/n_0 \ll 1$ of the ion and neutral densities causes the ion–ion interaction term in the expression (4.14) to be neglegible, and we are left with the ion–neutral collision term. Moreover, looking at the Maxwell interaction as a guiding case, we have[5] $n' \geqslant n + N$ and hence for $n' = 1$ the expression (4.14) degenerates to

$$\omega^{1,1}_{1,1;0,0}(i,0)\check{a}^{1,1,0}_{i} + \omega^{1,1}_{0,0;1,1}(i,0)\check{a}^{1,1,0}_{0}, \tag{4.15}$$

where we may neglect the second term, as the neutrals are nearly in perfect equilibrium. So in the case under consideration the collision frequency

$$\omega := \omega^{1,1}_{1,1;0,0}(i,0) \tag{4.16}$$

can serve as a normalizing quantity. Dividing the expression (4.11) by ω, we obtain the dimensionless differential term

$$\left(\frac{\partial \ln n}{\partial \xi}\right)^M \Bigg[\frac{l'+1}{2l'+1} \, {}^{+}C^{\,m'}_{M,m}(l')\{(n'-l')\check{a}^{n'-1,l'+1,m} + \check{a}^{n'+1,l'+1,m}\} + \frac{l'}{2l'+1} \, {}^{-}C^{\,m'}_{M,m}(l')$$
$$\times\{(n'+l'+1)\check{a}^{n'-1,l'-1,m} + \check{a}^{n'+1,l'-1,m}\}\Bigg] - \Phi^M \frac{l'+1}{2l'+1} \, {}^{+}C^{\,m'}_{M,m}(l')(n'-l')\check{a}^{n'-1,l'+1,m}$$
$$- \Phi^M \frac{l'}{2l'+1} \, {}^{-}C^{\,m'}_{M,m}(l')(n'+l'+1)\check{a}^{n'-1,l'-1,m} \tag{4.17}$$

with the spatial coordinate

$$\xi := \omega\left(\frac{\mu}{k_B T}\right)^{1/2} \mathbf{x} \tag{4.18}$$

and the field

$$\mathbf{\Phi} := \left(\frac{\mu}{k_B T}\right)^{1/2} \mathbf{\Psi}/\omega. \tag{4.19}$$

Both are dimensionless quantities and have been used in the context of drift-tube calculations. Obviously the normalizing collision frequency ω can be used to scale the omitted gyrofrequency as well as the collision frequencies, which appear for $n' \geqslant 2$ and for interactions different from the Maxwell case.

CONCLUSION

In Sec. IV, it is shown how the general result of this paper, Eq. (3.28), can be simplified to obtain familiar expressions, used for drift-tube calculations. As mentioned, however, this is only a small field of application. For example, if a reference velocity **c** is taken into account, we are able to evaluate the fourth-order viscosity tensor. Moreover, the convergence of the basic series expansions (1.3) of the distribution functions is improved by inclusion of **c**. On the other hand, the results of Sec. III should be extremely useful to handle time-dependent problems as for instance the application of ac electromagnetic fields or the approach to the steady state, and provide basic partial differential equations in the case of spatially inhomogeneous moments. Both features point to an application in the field of plasma physics.

In the context of this paper, we have restricted ourselves to the general expressions for the differential matrix. Only in Sec. IV do we briefly derive formulas for the calculation of mobility and diffusion, special "transport coefficients." Some more of them will be given subsequently as special applications of our formalism.

ACKNOWLEDGMENTS

I would like to thank the Max Kade Foundation, who enabled the completion of this paper by sponsoring my stay at Brown University, as well as Professor E.A. Mason, who made it possible for me to work with him during this period of time.

[1] J.L. Delcroix, *Plasma Physics* (Wiley, New York, 1970), Chap. 7.
[2] U. Weinert, thesis (University of Düsseldorf, 1975), Chap. 1.
[3] K. Kumar, Australian J. Phys. **20**, 205 (1967).
[4] J. Aisbett, J.M. Blatt, and A.H. Opie, J. Stat. Phys. **11**, 441 (1974).
[5] U. Weinert and K. Suchy, Z. Naturforschung **32a**, 390 (1977).
[6] U. Weinert, Z. Naturforschung **33a**, 480 (1978).
[7] K. Kumar, Australian J. Phys. **23**, 505 (1970).
[8] J.H. Whealton and E.A. Mason, Ann. Phys. (N.Y.) **84**, 8 (1974), Eq. (30).
[9] H. Grad, Commun. Pure Appl. Math. **2**, 325 (1949).
[10] E.A. Desloge, Am. J. Phys. **32**, 733 (1964).
[11] Ref. 2, Chaps. 2, 3.
[12] S. Chapman and T.G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge U. P., Cambridge, 1952), 2nd ed., Sec. 2.2.
[13] A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton U.P. Princeton, 1960), 2nd ed., Sec. 2.5.
[14] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970), 7th ed., Eqs. (22.2.12) and (22.3.9).
[15] Ref. 6, Appendix A1.
[16] U. Weinert, Report 77-01 (University of Düsseldorf, 1977), Appendix A1.
[17] J.A.R. Coope, R.F. Snider, and F.R. McCourt, J. Chem. Phys. **43**, 2269 (1965).
[18] Ref. 13, Sec. 3.6.
[19] Ref. 7, Eq. (6).
[20] Ref. 16, Sec. 2.6.

A new approach to inhomogeneous cosmologies: Intrinsic symmetries. I.

C. B. Collins[a)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
D. A. Szafron[b)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1 and Department of Mathematics, Portland State University, P. O. Box 751, Portland, Oregon 97207 [c)]

(Received 10 October 1978)

After briefly discussing systematic ways of studying inhomogeneous cosmologies, we propose the technique of "intrinsic symmetries," in which restrictions are placed on *sub*manifolds of space–time. This leads to a broad classification of inhomogeneous cosmologies, and to an associated specialization diagram. By enforcing additional restrictions we obtain a further useful specialization.

1. INTRODUCTION

The universe is inhomogeneous. Is spite of this obvious fact, nearly all theoretical investigations in general relativistic cosmology to date have involved models that are spatially homogeneous (and possibly isotropic): Universe *models* are *homogeneous.*

There is considerable justification for this situation. On a sufficiently large scale ($\gtrsim$ 50–100 Mpc.) the universe appears to be roughly homogeneous, for, if it were not, then large clumps of matter would produce anisotropies in the microwave background radiation in excess of those observed today ($\lesssim 0.1\%$; Sachs and Wolfe[1], Rees and Sciama[2]).

From a mathematical point of view, the study of spatially homogeneous universes ("Bianchi" and "Kantowski–Sachs" models) has enormous advantages. This is because the field equations become much more tractable (reducing from partial to ordinary differential equations), and because a "Bianchi–Behr" classification scheme has been devised that indicates which (of the Bianchi models) are the most general (see, e.g., Ellis and MacCallum[3]).

On the other hand, spatially homogeneous cosmological models have several defects. For instance, they may not be sufficiently general for problems where generality considerations are of prime importance (e.g., in the study of singularities). They seem incapable of explaining phenomena such as the homogenization and isotropization of the universe, as required in the chaotic cosmology program of Misner[4] (see, e.g., MacCallum,[5] Collins and Hawking,[6] and references cited). Moreover, it seems that spatially homogeneous models do not provide a suitable background for the formation of galaxies from small inhomogeneous perturbations (see, e.g., Ryan and Shepley,[7] and references cited). In order to study questions such as these, it will probably be necessary to consider inhomogeneous cosmologies, with all their associated mathematical problems.

By a "cosmological model," we shall mean

(i) a solution of Einstein's field equations,

$$R_{ij} - \tfrac{1}{2}Rg_{ij} + \Lambda g_{ij} = T_{ij},$$

in which the distribution of matter is approximated by a smeared-out fluid whose total energy density is nonzero. We do not therefore regard the "Gowdy universe"[8,9] as being a cosmological model, although it might naturally represent one in some approximation. Possibly the best-known inhomogeneous cosmologies are the spherically symmetric "Bondi–Tolman" models[10,11] and the plane-symmetric models of Taub.[12,13] More recently, Szekeres[14] examined a class of inhomogeneous models by starting from the somewhat unsatisfactory premise that the metrics have a particular diagonal form; he then discovered that this led in a remarkable way to a very tractable system of field equations, which he solved to obtain a useful and fairly wide set of cosmologies.

The main stumbling block to the development of inhomogeneous cosmologies has been the need to impose, in a covariant way, symmetries which are sufficiently strong to render the field equations tractable, while not being *so* strong that they require spatial homogeneity. Spatial homogeneity is specified mathematically by requiring that the space–time admit an r-parameter group of isometries, G_r, acting transitively on spacelike hypersurfaces (thus $r \geqslant 3$; for details of terminology, see, e.g., Ellis and MacCallum[3]). The simplest spatially homogeneous cosmologies are the isotropic models of Friedmann, Robertson, and Walker (FRW), while the simplest spatially homogeneous anisotropic models are the (perfect fluid) Bianchi I models, which have flat ("Euclidean") sections, and a metric of the form

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)dy^2 + Z^2(t)dz^2. \quad (1.1)$$

This allows the possibility of varying rates of expansion in different directions, yet, because of the homogeneity, denies any variation of conditions from point to point on each (flat) hypersurface (t = constant).

Perhaps the most obvious way of introducing inhomogeneity is to require that the space–time still admit a group of isometries, but that the (spacelike) orbits of the group not be three-dimensional. The simplest situations occur when a G_3 acts on two-dimensional spacelike orbits. In these cases, the orbits of the group are spaces of constant curvature, which

a)Supported in part by Operating Grant #A3978 from the National Research Council of Canada, and by a University of Waterloo Research Grant.
b)Supported by a Postgraduate Scholarship from the National Research Council of Canada.
c)Permanent address.

can be positive, negative, or zero. When the curvature is positive, the models are spherically symmetric, and when the curvature is zero, they are plane symmetric. All such models are fairly well known, and form a particular subset of the "locally rotationally symmetric" (LRS) models[15,16], in which the space–time is invariant under a spatial rotation about a spacelike axis of symmetry at each point. Another fairly symmetric case occurs when the space–time admits a 2-parameter group (e.g., the Gowdy universe). However, without further symmetries, even this case appears to be very intractable (for a perfect fluid). In view of this impasse, some investigators have sought alternative means of imposing symmetries. The most popular way of doing this appears to be by using conformal motions ("consometries"), particularly homothetic motions("homothesies"), and other "curvature collineations" (see, e.g., Katzin, Levine, and Davis[17] and Yano[18]), in place of the isometries employed in the homogeneous models. The use of homothetic motions leads to a classification similar to the Bianchi–Behr classification of spatially homogeneous models.[19,20] However, in order for a homothetic vector not to degenerate into a Killing vector, either the fluid must obey a Zeldovich ("stiff matter") equation of state ($p = \mu$), or the flow is necessarily *not* orthogonal to the group orbits,[21] i.e., these models are *necessarily* "tilted."[22] Little work has been done on the examination of these "self-similar" cosmologies, presumably because of the complicated features of tilted models.

As an alternative approach to space–time symmetries, we suggest in the present article a general technique which may be much more valuable. Instead of placing conditions (such as isometries) on the full space–time manifold, we shall impose restrictions on certain *sub*manifolds. For instance, one could postulate that the matter flow be irrotational, and that the hypersurfaces orthogonal to the fluid flow, *when considered as 3-manifolds*, admit symmetries which are not necessarily symmetries of the full space–time (we shall henceforth call these "*intrinsic symmetries*"). The simplest situation would occur when these hypersurfaces are flat. In this case, each hypersurface orthogonal to the fluid flow is a Euclidean 3-space, admitting the maximal number (6) of associated Killing vectors, which generate the familiar translations and rotations of rigid body motion. However, none of these Killing vectors is necessarily a Killing vector of the full space–time, so the model is not necessarily spatially homogeneous. For example, the particular Taub plane-symmetric model with metric,

$$ds^2 = -dt^2 + t^{-2/3}[t + C(x)]^2 dx^2 + t^{4/3}(dy^2 + dz^2),$$

where $C(x)$ is arbitrary, describes incoherent matter ("dust") flowing orthogonally to *flat* hypersurfaces ($t =$ constant), but the model is spatially homogeneous if and only if $C(x)$ is constant [Ellis[15]; cf. Eq. (1.1) above].

Just as in the full space–time case, there is a variety of ways of imposing intrinsic symmetries, such as by "intrinsic isometries," by "intrinsic consometries," or by requiring the Ricci tensor of the submanifolds to have certain properties. In the present article we shall exhibit a general classification scheme for *all* cosmological models (in fact, our classification is even independent of Einstein's field equations being valid, although we shall always assume this to be the case). We then consider a subclass of models satisfying the additional conditions

(ii) the cosmological constant, Λ, is zero,

(iii) the matter content is a perfect fluid, whose flowlines form a geodesic congruence orthogonal to a family, $\mathscr{F}$, of spacelike hypersurfaces, $\mathscr{S}$.

In subsequent articles, we shall impose further conditions, such as

(iv) each spacelike hypersurface, $\mathscr{S}$, in the family $\mathscr{F}$ is a conformally flat 3-manifold.

As an alternative to condition (iv), we could instead invoke other symmetries, such as

(va) each hypersurface $\mathscr{S}$ is a 3-space of constant curvature,

(vb) the Ricci tensor, R^{*}_{ij}, of each hypersurface $\mathscr{S}$ is isotropic, i.e., $R^{*}_{ij} = \frac{1}{3}R^{*}h_{ij}$, where h_{ij} is the metric induced on $\mathscr{S}$ by the space–time metric,

(vc) each hypersurface $\mathscr{S}$ admits a maximal group of isometries, i.e., a 6-parameter intrinsic isometry group, or we could impose the condition

(vi) each hypersurface $\mathscr{S}$ is flat.

Clearly condition (vi) is a particular case of (va), and, in fact, conditions (va, vb, vc) are all equivalent, as follows from well-known theorems in differential geometry.[23] However, conditions (v) imply (iv) (Ref. 23), and so (vi)$\Rightarrow$(va)$\Leftrightarrow$(vb)$\Leftrightarrow$(vc)$\Rightarrow$(iv). Another symmetry which we shall invoke from time to time is

(vii) both the second fundamental form and the Ricci tensor of the hypersurfaces $\mathscr{S}$ possess two equal eigenvalues. This condition imposes further symmetry on the intrinsic and extrinsic properties of the hypersurfaces in the family $\mathscr{F}$. When it holds, we shall take a basis vector $\mathbf{e}_1$ to be the "preferred" spacelike eigenvector.

In general, the validity of our results will be independent of any energy conditions. Some or our later results will require either

(viiia) $\mu + p > 0$ and $\mu + 3p > 0$,

or

(viiib) an equation of state, $p = p(\mu)$, where $dp/d\mu \geqslant 0$,

conditions which relate the energy density, μ, to the fluid pressure, p.

Cosmological models satisfying conditions (i)–(iv) generalize almost all of the well-known inhomogeneous models mentioned above (i.e., Bondi–Tolman, Taub, Szekeres, and many LRS models), and, as we shall show in a subsequent article, they also generalize some of the spatially homogeneous models (viz. the isotropic FRW models, the anisotropic Bianchi I models, some of the Bianchi–Behr type VI_h models with $h = -1$ (Bianchi type III), and all of the Kantowski–Sachs models).

Models in which condition (vi) holds are particularly interesting from a physical point of view. Much of the analysis of conditions near the cosmological singularity centers around the "velocity-dominated" case, in which the "3-Ricci" curvature terms are unimportant (the term "velocity-dominated" was coined by Eardley, Liang, and Sachs,[24] who had in mind the case when the characteristic length scale of inhomogeneities was very much greater than the particle horizon; this case was in fact treated earlier, though less rigorously, by Lifschitz and Khalatnikov[25]. The velocity dominated case contrasts with the Newtonian approximation, in which wavelengths of inhomogeneous perturbations are insignificant compared with the particle horizon; for this reason, it is sometimes known as the "anti-Newtonian" approximation, and as such it is used in problems relating to the formation of galaxies out of chaotic initial conditions (see, e.g., Tomita[26]). Inhomogeneous models satisfying condition (vi) are therefore ones for which the velocity dominated approximation is exact to all orders (cf. comments in Ref. 27).

In Sec. 2 we introduce the general formalism involving orthonormal tetrads with the timelike vector aligned along the fluid flow. We then obtain a classification and associated specialization diagram for all (fluid-filled) space–times. In Secs. 3 and 4 we impose the restrictions (i)–(iii) and obtain a number of results which lead to a further specialization diagram.

2. FORMALISM

We shall begin by considering space–times containing a fluid with 4-velocity **u**, where **u** is normalized so that $\mathbf{u}\cdot\mathbf{u} = -1$. We shall use the notation of Ellis[28] in which the metric tensor g_{ij} has signature $(-,+,+,+)$; Latin indices range from 0–3 (with i,j,k... denoting coordinate indices and a,b,c... denoting tetrad indices), Greek indices range from 1–3, partial differentiation is denoted by a comma, and covariant differentiation by a semicolon.

The tensor h_{ij} is defined as the projection tensor into the instantaneous rest space of the fluid by

$$h_{ij} = g_{ij} + u_i u_j.$$

We now introduce the standard decomposition (cf. Ref. 28) of $u_{i;j}$ into

$$u_{i;j} = \theta_{ij} + \omega_{ij} - \dot{u}_i u_j,$$

where the "expansion tensor," θ_{ij}, is the symmetric part orthogonal to **u**, the "vorticity tensor," ω_{ij}, is the antisymmetric part orthogonal to **u**, and $\dot{\mathbf{u}}$ is the "acceleration vector." In the special case when **u** is hypersurface orthogonal, the vorticity vanishes and the tensor **h** is the induced metric on the hypersurfaces orthogonal to **u**.

We now introduce the orthonormal tetrad formalism (a more detailed description is given in MacCallum[29]). We choose an orthonormal tetrad $\{\mathbf{e}_a\}$ with the timelike basis vector $\mathbf{e}_0$ aligned along the fluid flow, i.e., $\mathbf{e}_0 = \mathbf{u}$. Since the tetrad $\{\mathbf{e}_a\}$ is orthonormal, the tetrad components of the metric tensor are

$$\mathbf{e}_a\cdot\mathbf{e}_b = g_{ab} = \mathrm{diag}(-1,+1,+1,+1),$$

with

$$g_{ij} = e_i{}^a e_j{}^b g_{ab}. \tag{2.1}$$

Given any function ϕ, the directional derivative of ϕ along $\mathbf{e}_a$ is defined to be

$$\partial_a\phi = \phi_{,j} e^j{}_a .$$

The "Ricci rotation coefficients," Γ_{abc}, are defined by

$$\Gamma_{abc} = \mathbf{e}_a\cdot\nabla_b\mathbf{e}_c = e_a{}^i e_{ci;j} e_b{}^j .$$

In general, the directional derivatives do not commute, so we obtain the "commutation functions," $\gamma^a{}_{bc}$, defined by the Lie bracket

$$[\mathbf{e}_a,\mathbf{e}_b] = \gamma^c{}_{ab}\mathbf{e}_c, \quad \gamma^a{}_{bc} = \gamma^a{}_{[bc]} . \tag{2.2}$$

The commutation functions are related to the Ricci rotation coefficients by

$$\gamma^c{}_{ab} = \Gamma^c{}_{ab} - \Gamma^c{}_{ba} .$$

The quantity $\gamma^\alpha{}_{\beta\lambda}$ is decomposed into its symmetric and antisymmetric parts

$$\gamma^\alpha{}_{\beta\lambda} = \epsilon_{\beta\lambda\mu} n^{\mu\alpha} + \delta^\alpha_\lambda a_\beta - \delta^\alpha_\beta a_\lambda, n^{\mu\alpha} = n^{(\mu\alpha)}. \tag{2.3}$$

Since $\mathbf{e}_0 = \mathbf{u}$, the remaining $\gamma^a{}_{bc}$'s can be expressed in terms of kinematical quantities

$$\gamma_{o o \nu} = -\dot{u}_\nu, \quad \gamma_{\mu o \nu} = -\epsilon_{\mu\nu\sigma}(\Omega^\sigma + \omega^\sigma) - \theta_{\mu\nu},$$
$$\gamma_{o\mu\nu} = 2\epsilon_{\mu\nu\sigma}\omega^\sigma. \tag{2.4}$$

The "vorticity vector," ω^a, is defined by

$$\omega^a \equiv \tfrac{1}{2}\eta^{abcd}u_b\omega_{cd} \Leftrightarrow \omega_{ab} = \eta_{abcd}u^c\omega^d$$

is orthogonal to **u** and contains all the information that ω_{cd} does. The vector

$$\Omega^a \equiv \tfrac{1}{2}\eta^{abcd}u_b\dot{\mathbf{e}}_c\cdot\mathbf{e}_d$$

is the angular velocity of the triad $\{\mathbf{e}_\alpha\}$ with respect to a set of Fermi-propagated axes along $\mathbf{e}_0 = \mathbf{u}$. In Appendix A, we write out Eqs. (2.2) in full, using Eqs. (2.3) and (2.4). Follow-

TABLE I. The relationship between the invariants θ,σ,τ and the canonical form of $\theta_{\alpha\beta}$. From Eq. (2.5), we see that $\sigma = 0$ implies $\tau = 0$, so there are no canonical forms for $\sigma = 0$ and $\tau\neq 0$. Where suitable, the $\mathbf{e}_1$ vector is taken to be preferred.

	$\theta\sigma\tau\neq 0$	$\theta\sigma\neq 0$, $\tau = 0$	$\theta\neq 0$, $\sigma=\tau=0$	$\theta = 0$, $\sigma\tau\neq 0$	$\theta=\tau=0$, $\sigma\neq 0$	$\theta=\sigma=\tau=0$
$\theta_{\alpha\beta}$	$\mathrm{diag}(\theta_1,\theta_2,\theta_3)$ $\theta_1,\theta_2,\theta_3$ distinct $\theta_1+\theta_2+\theta_3\neq 0$	$\mathrm{diag}(\theta_1,\theta_2,\theta_2)$ $\theta_1\neq\theta_2$	$\mathrm{diag}(\theta_1,\theta_1,\theta_1)$ $\theta_1\neq 0$	$\mathrm{diag}(\theta_1,\theta_2,-(\theta_1+\theta_2))$ $(\theta_1-\theta_2)(2\theta_1+\theta_2)(\theta_1+2\theta_2)\neq 0$	$\mathrm{diag}(-2\theta_1,\theta_1,\theta_1)$ $\theta_1\neq 0$	$\mathrm{diag}(0,0,0)$

TABLE II. Algebraic classification of fluid-filled space–times, according to the vanishing of the quantities $\theta,\sigma,\tau,R^*,S,T$. Note that in each case there is the possibility of having the acceleration, $\dot{\mathbf{u}}$, and the vorticity, $\boldsymbol{\omega}$, zero or nonzero.

		I. $\theta\sigma\tau R^*ST\neq 0$			
IIa. $R^*=0$	IIb. $T=0$		IIc. $\tau=0$		IId. $\theta=0$
IIIa. $R^*=T=0$	IIIb. $R^*=\tau=0$	IIIc. $T=\tau=0$	IIId. $\theta=R^*=0$	IIIe. $\theta=T=0$	IIIf. $\theta=\tau=0$
IVa. $R^*=\tau=T=0$	IVb. $\theta=R^*=T=0$		IVc. $\theta=R^*=\tau=0$		IVd. $\theta=\tau=T=0$
		V. $\theta=R^*=\tau=T=0$			
	VIa. $S=0$		VIb. $\sigma=0$		
VIIa. $S=R^*=0$	VIIb. $S=\tau=0$	VIIc. $S=\theta=0$	VIId. $\sigma=R^*=0$	VIIe. $\sigma=T=0$	VIIf. $\sigma=\theta=0$
VIIIa. $S=R^*=\tau=0$	VIIIb. $S=\tau=\theta=0$	VIIIc. $S=R^*=\theta=0$	VIIId. $\sigma=\theta=R^*=0$	VIIIe. $\sigma=T=R^*=0$	VIIIf. $\sigma=\theta=T=0$
	IXa. $S=R^*=\theta=\tau=0$		IXb. $\sigma=\theta=R^*=T=0$		
		X. $\sigma=S=0$			
	XIa. $S=\sigma=R^*=0$		XIb. $\sigma=S=\theta=0$		
		XII. $\theta=\sigma=\tau=R^*=S=T=0$			

ing MacCallum,[29] we define a quantity $R^*_{\alpha\beta}$ by

$$R^*_{\alpha\beta}\equiv\partial_{(\alpha}a_{\beta)}-\epsilon_{\delta\gamma(\alpha}\partial^\delta n_{\beta)}{}^\gamma-2\epsilon_{\gamma\delta(\alpha}n_{\beta)}{}^\gamma a^\delta+2n^\gamma{}_{(\alpha}n_{\beta)\gamma}$$

$$-nn_{\alpha\beta}-\delta_{\alpha\beta}(2a_\gamma a^\gamma+n^{\gamma\delta}n_{\gamma\delta}-\tfrac{1}{2}n^2-\partial_\gamma a^\gamma),$$

where $n\equiv n^\alpha{}_\alpha$.

This quantity is defined so that when the vorticity is zero, it is the 3-Ricci tensor of the hypersurfaces orthogonal to $\mathbf{u}$. Notice, however, that $R^*_{\alpha\beta}$ depends only on the $n^{\alpha\beta},a^\lambda$ and their spatial derivatives, but not on ω^α, so it is well defined, regardless of whether or not $\boldsymbol{\omega}=\mathbf{0}$. We define the "shear tensor" σ_{ij}, the "shear scalar" $\sigma\geqslant 0$, the trace free "3-Ricci" quantity $S_{\alpha\beta}$ and the scalar $S\geqslant 0$ by

$$\sigma_{ij}=\theta_{ij}-\tfrac{1}{3}\theta h_{ij},\quad S_{\alpha\beta}=R^*_{\alpha\beta}-\tfrac{1}{3}R^{*\delta}_{\alpha\beta},$$

$$2\sigma^2=\sigma_{ij}\sigma^{ij},\quad 2S^2=S_{\alpha\beta}S^{\alpha\beta},$$

where the "expansion," θ, and the "3-Ricci" scalar R^* are

$$\theta=\theta^i{}_i,\quad R^*=R^*{}_\alpha{}^\alpha.$$

We define two additional scalars τ and T by

$$\begin{aligned}\tau&=4\sigma^6-3(\sigma_{ij}\,\sigma^{jk}\sigma_k{}^i)^2,\\ T&=4S^6-3(S_{\alpha\beta}\,S^{\beta\gamma}S_\gamma{}^\alpha)^2.\end{aligned}\tag{2.5}$$

Note that

$$\sigma=0\Leftrightarrow\sigma_{ij}=0\Rightarrow\tau=0,\quad S=0\Leftrightarrow S_{\alpha\beta}=0\Rightarrow T=0.$$

In Table I, we express the relationship between the algebraically independent quantities θ, σ, τ and the canonical forms of $\theta_{\alpha\beta}$ (i.e., with respect to an eigenframe of θ_{ij}). A similar relationship exists between the invariants R^*, S, T and the canonical forms of $R^*_{\alpha\beta}$.

In a general space–time we would expect the six algebraically independent quantities θ, σ, τ, R^*, S, and T to be nonzero on an open set. However, there would also be "special" space–times in which at least one of these six quantities vanished on an open set. We therefore obtain an algebraic classification, based on the vanishing of all possible combinations of θ, σ, τ, R^*, S, and T (see Table II; the labeling of the various types has been done in such a way as to obtain as much symmetry as possible between θ and R^* and between τ and T).

In each case, there is the possibility of having the fluid acceleration, $\dot{\mathbf{u}}$, and vorticity, $\boldsymbol{\omega}$, zero or nonzero. Note, however, that we do not at present require the field equations of general relativity to be satisfied, nor the matter to be a perfect fluid. Restrictions such as these could conceivably require that certain subcases in our classification must be void in any *open* region of space–time.

3. PHYSICAL RESTRICTIONS

In this section we shall investigate the consequences of imposing conditions (i)–(iii), (vii), and (viii) onto the classification given in Table II. The Einstein field equations for a perfect fluid are

$$R_{ij}-\tfrac{1}{2}Rg_{ij}+\Lambda g_{ij}=\mu u_iu_j+ph_{ij}.\tag{3.1}$$

The tetrad form of these equations, in the special case $\boldsymbol{\omega}=\dot{\mathbf{u}}=\mathbf{0}$, is given in Appendix B, along with the Jacobi identities

$$[[\mathbf{e}_a,\mathbf{e}_b],\mathbf{e}_c]+[[\mathbf{e}_c,\mathbf{e}_a],\mathbf{e}_b]+[[\mathbf{e}_b,\mathbf{e}_c],\mathbf{e}_a]=0$$

(cf. Ref. 29). (3.2)

The vector $\mathbf{u}$ is a timelike eigenvector of σ_{ab} and R^*_{ab}, with zero eigenvalue. Henceforth when we refer to eigenvectors and eigenvalues of σ_{ab} and R^*_{ab}, we shall be considering *spacelike* eigenvectors and their associated eigenvalues.

Proposition 3.1: In any space-time satisfying conditions (i), (iii), and (vii), any shear eigenframe is necessarily an eigenframe of R^*_{ab}. If $\sigma\neq 0$, then the Fermi propagation vector, $\boldsymbol{\Omega}$, is parallel to the preferred eigenvector of σ_{ab} and R^*_{ab}.

Proof: If $\sigma\equiv 0$, then Eq. (A3) shows that $R^*_{\alpha\beta}$ is isotropic, and so any frame in which $\mathbf{e}_0=\mathbf{u}$ is an eigenframe of $R^*_{\alpha\beta}$. Henceforth, we assume $\sigma\not\equiv 0$. We consider the Jacobi and field Eqs. (A1)–(A6) in a shear eigenframe (because of the equality of two eigenvalues of $\sigma_{\alpha\beta}$, this eigenframe is not of course unique, but, at each point in the hypersurface, there is the freedom to rotate the "2" and "3" axes about the "1" direction).

It follows immediately from Eqs. (A3) that in this frame $R^*_{23}=0$ and $R^*_{22}=R^*_{33}$, i.e., with respect to the shear eigenframe, $R^*_{\alpha\beta}$ is of the form

$$\begin{pmatrix}R^*_{11}&R^*_{12}&R^*_{13}\\R^*_{12}&R^*_{22}&0\\R^*_{13}&0&R^*_{22}\end{pmatrix}.\tag{3.3}$$

TABLE III. Algebraic classification of space–times satisfying conditions (i)–(iii) and (viiia).

	I. $\theta\sigma\tau R^*ST\neq 0$	
IIa. $R^*=0$	IIb. $T=0$	IIc. $\tau=0$
IIIa. $R^*=T=0$	IIIb. $R^*=\tau=0$	IIIc. $T=\tau=0$
	VIa. $S=0$	
	VIIa. $S=R^*=0$ VIIb. $S=\tau=0$	
	VIIIa. $S=R^*=\tau=0$	
	X. $\sigma=S=0$	
	XIa. $S=\sigma=R^*=0$	

Our condition (vii) requires that $R^{\,*}_{\alpha\beta}$ possesses two equal eigenvalues. The characteristic equation for the matrix (3.3) is

$$(\lambda - R^{\,*}_{22})Q(\lambda)=0,$$

where

$$Q(\lambda)\equiv\lambda^2-\lambda\,(R^{\,*}_{11}+R^{\,*}_{22})+(R^{\,*}_{11}R^{\,*}_{22}-R^{\,*2}_{12}-R^{\,*2}_{13}).$$

Hence either $R^{\,*}_{22}$ is a repeated eigenvalue, or $Q(\lambda)$ has two equal roots. In the former case $Q(R^{\,*}_{22})=0\Leftrightarrow R^{\,*}_{12}=R^{\,*}_{13}=0$; on the other hand, $Q(\lambda)$ has two equal roots if and only if $(R^{\,*}_{11}-R^{\,*}_{22})^2+4(R^{\,*2}_{12}+R^{\,*2}_{13})=0\Leftrightarrow R^{\,*}_{12}=R^{\,*}_{13}=R^{\,*}_{11}-R^{\,*}_{22}=0$. Thus, in either case, $R^{\,*}_{12}=R^{\,*}_{13}=0$, and the orthonormal tetrad is an eigenframe of both σ_{ab} and $R^{\,*}_{ab}$. Equations (A3) now show that $\Omega_2=\Omega_3=0$.□

Corollary: Under the conditions of the lemma, $S_{\alpha\beta}\propto\sigma_{\alpha\beta}$.

Proof: If $\sigma\not\equiv 0$, the result follows immediately. If $\sigma\equiv 0$, the space–time is FRW, so $S_{\alpha\beta}\equiv 0$.□

Proposition 3.2: In any space–time satisfying conditions (i), (iii), and (vii), if the hypersurfaces $\mathscr{S}$ have 3-Ricci scalar $R^*\equiv 0$, then these hypersurfaces are flat.

Proof: The contraction of Eq. (A3) together with Eq. (A1) yields the two algebraically independent equations

$$R^*-2\sigma^2=2(\mu+\Lambda)-\tfrac{2}{3}\theta^2 \tag{3.4}$$

and

$$\dot\theta=\tfrac{1}{2}(R^*-6\sigma^2)-\tfrac{3}{2}(\mu+p). \tag{3.5}$$

Using Eqs. (3.4), (3.5), and (A7) we obtain

$$\partial_0(R^*-2\sigma^2)=-\tfrac{2}{3}\theta\,(R^*-6\sigma^2), \tag{3.6}$$

which is due to Raychaudhuri.[30] The trace-free part of (A3) contracted with $\sigma^{\alpha\beta}$ yields

$$\partial_0\sigma^2+2\theta\sigma^2=-\sigma^{\alpha\beta}S_{\alpha\beta} \tag{3.7}$$

(this is most readily seen using either a Fermi-propagated tetrad or a shear eigenframe).

With $R^*\equiv 0$, Eqs. (3.6) and (3.7) require

$$\sigma^{\alpha\beta}S_{\alpha\beta}=0.$$

By the Corollary to Proposition 3.1 we obtain $S_{\alpha\beta}=0$. Thus, $R^{\,*}_{\alpha\beta}=S_{\alpha\beta}+\tfrac{1}{3}R^*\delta_{\alpha\beta}=0$, and the hypersurfaces are flat.

Proposition 3.3: In any space–time satisfying conditions (i)–(iii) and (viiia), $\theta\not\equiv 0$.

Proof: If conditions (i), (ii), and (iii) are satisfied, along with $\theta\equiv 0$, then Eq. (A1) becomes

$$2\sigma^2+\tfrac{1}{2}(\mu+3p)=0.$$

But this requires $(\mu+3p)\leqslant 0$, which contradicts condition (viiia) Thus we must have $\theta\not\equiv 0$.□

Using Propositions 3.1–3.3, the algebraic classification of space–times satisfying conditions (i)–(iii) and (viiia) is given in Table III. Those cases arising from Table II which are void are omitted.

4. FURTHER INVARIANTS

Some classes of space–times from Table II or Table III may contain additional invariants. When this situation arises a subclassification scheme can be introduced. We illustrate how this can be done by examining those space–times from Table III which are in class IIIc, i.e., $T=\tau=0$. We first determine the additional invariants.

Lemma 4.1: If four quantities x, y, z, and w undergo a transformation T:

$$x'=x\cos\psi+y\sin\psi,$$
$$y'=-x\sin\psi+y\cos\psi,$$
$$z'=z\cos2\psi+w\sin2\psi,$$
$$w'=-z\sin2\psi+w\cos2\psi,$$

there are *in general* exactly three independent algebraic invariants, viz. x^2+y^2, z^2+w^2 and $[z(x^2-y^2)+2xyw]/[w(x^2-y^2)-2xyz]$.

Proof: Clearly x^2+y^2 and z^2+w^2 are independent invariants. Consider the effect of the transformation T on the points $P_1=(x,y)$ and $P_2=(z,w)$ in a plane. If 0 is the origin and $0X$ is the horizontal axis, the line $0P_1$ is rotated through an angle ψ, while $0P_2$ is rotated through an angle 2ψ. Since $(P_2\hat0X)'=P_2\hat0X+2\psi$ and $(P_1\hat0X)'=P_1\hat0X+\psi$, the angle $P_2\hat0X-2P_1\hat0X$ is invariant. But,

$$P_2\hat0X=\tan^{-1}(z/w)\quad\text{and}$$

$$2P_1\hat0X=\tan^{-1}\left(\frac{2(y/x)}{(1-(y/x)^2)}\right),$$

whence $\tan(P_2\hat0X-2P_1\hat0X)=[z(x^2-y^2)+2xyw]/[w(x^2-y^2)-2xyz]$ is invariant.□

Remarks: Since this proof depends on geometrical intuition, we present an alternative proof which is more systematic, and so is more easily generalized to other situations. The technique is to change from variables x, y, z and w to a set of four variables, two of which are the invariants $J_1=x^2+y^2$ and $J_2=z^2+w^2$, already found, the other two being $u=y/x$ and $v=z/w$. Then J_1, J_2, u and v are four independent variables which transform under the transformation T as follows:

$$J'_1 = J_1,$$
$$J'_2 = J_2,$$
$$u' = \frac{u - \tan\psi}{1 + u\tan\psi},$$
$$v' = \frac{v - \tan 2\psi}{1 + v\tan 2\psi},$$

where a prime denotes new values. The problem of determining further invariants reduces to one of finding all independent functions of u and v invariant under the transformation T. This is equivalent to eliminating ψ in the above equations. Writing $\tan\psi = (u - u')/(1 + uu')$ and $\tan 2\psi = (v - v')/(1 + vv')$, we obtain

$$\frac{v - v'}{1 + vv'} = \frac{2(u - u')(1 + uu')}{1 + u^2u'^2 + 4uu' - u^2 - u'^2}.$$

It follows after straightforward computation that

$$[v(1 - u^2) - 2u][1 - u'^2 + 2u'v'] = [v'(1 - u'^2) - 2u'][1 - u^2 + 2uv],$$

from which

$$J_3 = \frac{1 + 2uv/(1 - u^2)}{v - 2u/(1 - u^2)}$$

is invariant. Clearly, there can be no further independent invariants.

Note also that if $z \equiv w \equiv 0$ the only independent invariant is $x^2 + y^2$.

Theorem 4.2: In any space–time satisfying conditions (i), (iii), and (vii), the following quantities are invariant under a rotation of the triad $\{\mathbf{e}_\alpha\}$ about the "1" direction:

$$I_1 = a_1,$$
$$I_2 = (n_{31} - a_2)^2 + (n_{12} + a_3)^2,$$
$$I_3 = n_{11},$$
$$I_4 = (n_{22} - n_{33})^2 + 4n_{23}^2,$$
$$I_5 = \{(n_{22} - n_{33})[(n_{12} + a_3)^2 - (n_{31} - a_2)^2] + 4n_{23}(n_{12} + a_3) \times (n_{31} - a_2)\}\{2n_{23}[(n_{12} + a_3)^2 - (n_{31} - a_2)^2] - 2(n_{22} - n_{33})(n_{12} + a_3)(n_{31} - a_2)\}^{-1},$$

together with θ_1 and θ_2. If $\sigma \neq 0$, there are no further independent algebraic invariants (cf. Ref. 15).

Proof: With the change of basis $\mathbf{e}'_0 = \mathbf{e}_0$, $\mathbf{e}'_1 = \mathbf{e}_1$, $\mathbf{e}'_2 = \mathbf{e}_2\cos\psi + \mathbf{e}_3\sin\psi$, $\mathbf{e}'_3 = -\mathbf{e}_2\sin\psi + \mathbf{e}_3\cos\psi$, we obtain from the commutation relations of Appendix A:

$$\Omega'_1 = \Omega_1 + \partial_0\psi,$$
$$\theta'_1 = \theta_1,$$
$$\theta'_2 = \theta_2,$$
$$n'_{11} = n_{11},$$
$$n'_{22} = n_{22}\cos^2\psi + n_{33}\sin^2\psi + n_{23}\sin 2\psi + \partial_1\psi,$$
$$n'_{33} = n_{22}\sin^2\psi + n_{33}\cos^2\psi - n_{23}\sin 2\psi + \partial_1\psi,$$
$$n'_{23} = n_{23}\cos 2\psi - \tfrac{1}{2}(n_{22} - n_{33})\sin 2\psi,$$
$$n'_{31} = n_{31}\cos\psi - n_{12}\sin\psi + \tfrac{1}{2}(\sin\psi\partial_2\psi - \cos\psi\partial_3\psi),$$
$$n'_{12} = n_{12}\cos\psi + n_{31}\sin\psi - \tfrac{1}{2}(\cos\psi\partial_2\psi + \sin\psi\partial_3\psi),$$
$$a'_1 = a_1,$$
$$a'_2 = a_2\cos\psi + a_3\sin\psi + \tfrac{1}{2}(\sin\psi\partial_2\psi - \cos\psi\partial_3\psi),$$

and

$$a'_3 = a_3\cos\psi - a_2\sin\psi + \tfrac{1}{2}(\cos\psi\partial_2\psi + \sin\psi\partial_3\psi),$$

We have chosen the shear eigenframe of Proposition 3.1. The five quantities ψ, $\partial_0\psi$, $\partial_1\psi$, $\partial_2\psi$, and $\partial_3\psi$ are regarded as being arbitrary and independent, since we are concerned here only with an *algebraic* classification, in which we consider conditions at any chosen point of the space–time, and ignore (for the time being) the variation of quantities in the neighborhood of the point.

It follows immediately that I_1, I_3, θ_1, and θ_2 are invariants. The only other independent algebraic invariant is I_5. This can be seen from the way in which the above transformations "decouple." Out of the eight quantities Ω_1, n_{22}, n_{33}, n_{23}, n_{31}, n_{12}, a_2, and a_3, the quantities Ω_1 and $(n_{22} + n_{33})$ are "translated" through arbitrary amounts $\partial_0\psi$ and $\partial_1\psi$, respectively (i.e., no invariants). The quantities $2n_{23}$ and $(n_{22} - n_{33})$ transform as under a rotation through 2ψ, while the quantities $(n_{31} - a_2)$ and $(n_{12} + a_3)$ transform as under a rotation ψ; it follows from Lemma 4.1 that associated with these transformations are exactly three independent algebraic invariants, viz., $I_4 = 4n_{23}^2 + (n_{22} - n_{33})^2$, $I_2 = (n_{31} - a_2)^2 + (n_{12} + a_3)^2$, and $I_5 = [z(x^2 - y^2) + 2xyw]/[w(x^2 - y^2) - 2xyz]$, where $x = (n_{12} + a_3)$, $y = (n_{31} - a_2)$, $z = (n_{22} - n_{33})$, and $w = 2n_{23}$. Finally, the quantities $n_{31} + a_2$ and $n_{12} - a_3$ transform as under a combination of a rotation through ψ, together with translations governed by arbitrary amounts $\partial_2\psi$ and $\partial_3\psi$ (no invariants).□

In Proposition 3.1 we found that if $\sigma \not\equiv 0$, $\Omega_2 = \Omega_3 = 0$. Naturally, we cannot obtain any similar restriction on Ω_1, this being a reflection of the fact that we have not yet used up the freedom of rotation of the shear eigenframe about the "1" axis. In fact, we have

Proposition 4.3: In any space–time satisfying conditions (i), (iii), and (vii), the shear eigenframe may be chosen to be Fermi-propagated. In general ($\sigma \not\equiv 0$), the remaining freedom of choice is an arbitrary rotation about the "1" direction at each point in an initial hypersurface. In $\sigma \equiv 0$, the frame may be chosen arbitrarily in an initial hypersurface.

Proof: Choose a shear eigenframe everywhere in an initial hypersurface $\mathscr{S}$. Fermi-propagation of this eigenframe off $\mathscr{S}$ requires, from Proposition 3.1 and Eqs. (A3), that

$$R^{\,*}_{\alpha\beta} = 0 \Leftrightarrow \partial_0\sigma_{\alpha\beta} + \theta\sigma_{\alpha\beta} = 0, \quad \alpha \neq \beta,$$

and hence $\sigma_{\alpha\beta} = 0\,(\alpha \neq \beta)$ on $\mathscr{S}$ implies $\sigma_{\alpha\beta} \equiv 0\,(\alpha \neq \beta)$ everywhere in the space–time.

One still has the full freedom to choose the shear eigenframe everywhere in $\mathscr{S}$, which $\sigma \not\equiv 0$, is a rotation about the "1" direction, and when $\sigma \equiv 0$ is an arbitrary rotation.□

Remark: Equivalently, we can prove Proposition 4.3 by using a result obtained in the proof of Theorem 4.2, namely that one is free to rotate the shear eigenframe through an angle ψ about the "1" direction, in which case Ω_1 changes to $\Omega'_1 = \Omega_1 + \partial_0\psi$. By choosing the scalar ψ to satisfy $\Omega_1 + \partial_0\psi = 0$ we obtain a Fermi-propagated shear eigenframe. The rotation angle would henceforth be constrained to satisfy $\partial_0\psi = 0$, in order to maintain the condition $\Omega_1 = 0$.

In subsequent papers, we will examine the consequences of imposing the condition (iv) of conformally flat comoving slices, in addition to some of the constraints considered in the present article.

ACKNOWLEDGMENT

We thank C.B.G. McIntosh for illuminating discussions on the role of homothetic motions in inhomogeneous cosmologies.

APPENDIX A: COMMUTATION FUNCTIONS IN TERMS OF KINEMATICAL QUANTITIES

The commutators (2.2), using Eqs. (2.3) and (2.4), are:

$$[\mathbf{e}_0,\mathbf{e}_1] = \dot{u}^1\mathbf{e}_0 - \theta_1\mathbf{e}_1 - (\sigma_{12} - \omega_3 - \Omega_3)\mathbf{e}_2 - (\sigma_{31} + \omega_2 + \Omega_2)\mathbf{e}_3, \tag{A1}$$

$$[\mathbf{e}_0,\mathbf{e}_2] = \dot{u}^2\mathbf{e}_0 - (\sigma_{12} + \omega_3 + \Omega_3)\mathbf{e}_1 - \theta_2\mathbf{e}_2 - (\sigma_{23} - \omega_1 - \Omega_1)\mathbf{e}_3, \tag{A2}$$

$$[\mathbf{e}_0,\mathbf{e}_3] = \dot{u}^3\mathbf{e}_0 - (\sigma_{31} - \omega_2 - \Omega_2)\mathbf{e}_1 - (\sigma_{23} + \omega_1 + \Omega_1)\mathbf{e}_2 - \theta_3\mathbf{e}_3, \tag{A3}$$

$$[\mathbf{e}_2,\mathbf{e}_3] = -2\omega_1\mathbf{e}_0 + n_{11}\mathbf{e}_1 + (n_{12} - a_3)\mathbf{e}_2 + (n_{31} + a_2)\mathbf{e}_3, \tag{A4}$$

$$[\mathbf{e}_3,\mathbf{e}_1] = -2\omega_2\mathbf{e}_0 + (n_{12} + a_3)\mathbf{e}_1 + n_{22}\mathbf{e}_2 + (n_{23} - a_1)\mathbf{e}_3, \tag{A5}$$

$$[\mathbf{e}_1,\mathbf{e}_2] = -2\omega_3\mathbf{e}_0 + (n_{31} - a_2)\mathbf{e}_1 + (n_{23} + a_1)\mathbf{e}_2 + n_{33}\mathbf{e}_3. \tag{A6}$$

APPENDIX B: FIELD EQUATIONS AND JACOBI IDENTITIES IN AN ORTHONORMAL TETRAD

The field equations (3.1) in tetrad form for the special case $\dot{\mathbf{u}} = \boldsymbol{\omega} = \mathbf{0}$ are:

$$\dot{\theta} + \theta^{\alpha\beta}\theta_{\alpha\beta} + \tfrac{1}{2}(\mu + 3p) = \Lambda, \tag{B1}$$

$$\tfrac{2}{3}\partial_\alpha\theta - \partial_\beta\,\sigma_\alpha{}^\beta + 3\sigma_\alpha{}^\beta a_\beta - \epsilon_{\alpha\beta\gamma}n^{\gamma\delta}\sigma_\delta{}^\beta = 0, \tag{B2}$$

$$-R^{*}_{\alpha\beta} = \partial_0\theta_{\alpha\beta} + \theta\theta_{\alpha\beta} + 2\theta^{\gamma}{}_{(\alpha}\epsilon_{\beta)\delta\gamma}\Omega^{\delta} - \tfrac{1}{2}\delta_{\alpha\beta}(\mu - p + 2\Lambda). \tag{B3}$$

Jacobi identities: The Jacobi identities (3.2) for the special case $\dot{\mathbf{u}} = \boldsymbol{\omega} = \mathbf{0}$ are:

$$\partial_\alpha n^{\alpha\delta} + \epsilon^{\delta\alpha\beta}\partial_\alpha a_\beta - 2n^\delta{}_\beta a^\beta = 0, \tag{B4}$$

$$2\partial_0 a_\alpha - \partial_\delta\theta^{\delta}_{\alpha} + \partial_\alpha\theta + \epsilon_\alpha{}^{\delta\epsilon}\partial_\delta\Omega_\epsilon + 2\theta_\alpha{}^\beta a_\beta - 2\epsilon_{\alpha\beta\gamma}a^\beta\Omega^\gamma = 0, \tag{B5}$$

$$\partial_0 n^{\alpha\beta} - \epsilon^{\delta\gamma(\alpha}\partial_\gamma\theta_\delta{}^{\beta)} + \partial^{(\alpha}\Omega^{\beta)} - 2n^{\gamma(\alpha}\epsilon^{\beta)}{}_{\gamma\delta}\Omega^{\delta} - 2n_\gamma{}^{(\alpha}\theta^{\beta)\gamma} + n^{\alpha\beta}\theta - \delta^{\alpha\beta}\partial_\gamma\Omega^\gamma = 0. \tag{B6}$$

The contracted Bianchi identities are:

$$\dot{\mu} + (\mu + p)\theta = 0, \tag{B7}$$

$$\partial_\alpha p = 0. \tag{B8}$$

[1] R.K. Sachs and A.M. Wolfe, Astrophys. J. **147**, 43 (1967).
[2] M.J. Rees and D.W. Sciama, Nature **217**, 511 (1968).
[3] G.F.R. Ellis and M.A.H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).
[4] C.W. Misner, "Relativistic Fluids in Cosmology" in *Battelle Rencontres, 1967 Lectures in Mathematics and Physics*, edited by C.M. deWitt and J.A. Wheeler (Benjamin, New York and Amsterdam, 1968), p. 117.
[5] M.A.H. MacCallum, Nature Phys. Sci. **230**, 112 (1971).
[6] C.B. Collins and S.W. Hawking, Astrophys. J. **180**, 317 (1973).
[7] M.P. Ryan and L.C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton U.P., Princeton, N.J., 1975).
[8] R.H. Gowdy, Ann. Phys. (N.Y.) **83**, 203 (1974).
[9] R.H. Gowdy, J. Math. Phys. **16**, 224 (1975).
[10] R.C. Tolman, Proc. Nat. Acad. Sci. **20**, 169 (1934).
[11] H. Bondi, Mon. Not. R. Astr. Soc. **107**, 410 (1947).
[12] A.H. Taub, Ann. Math. **53**, 472 (1951).
[13] A.H. Taub, Phys. Rev. **103**, 454 (1956).
[14] P. Szekeres, Commun. Math. Phys. **41**, 55 (1975).
[15] G.F.R. Ellis, J. Math. Phys. **8**, 1171 (1967).
[16] J.M. Stewart and G.F.R. Ellis, J. Math. Phys. **9**, 1072 (1968).
[17] G.H. Katzin, J. Levine, and W.R. Davis, J. Math. Phys. **10**, 617 (1969).
[18] K. Yano, *The Theory of Lie Derivatives and its Applications* (North-Holland, Amsterdam, Noordhoff, Groningen, Interscience, New York, 1957).
[19] D.M. Eardley, Commun. Math. Phys. **37**, 287 (1974).
[20] J.-P. Luminet, Gen. Rel. Grav. **9**, 673 (1978).
[21] C.B.G. McIntosh, Gen. Rel. Grav. **7**, 199 (1976).
[22] A.R. King and G.F.R. Ellis, Commun. Math. Phys. **31**, 209 (1973).
[23] L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1926).
[24] D.M. Eardley, E.P.T. Liang, and R.K. Sachs, J. Math. Phys. **13**, 99 (1972).
[25] E.M. Lifschitz and I.M. Khalatnikov, Adv. Phys. (Philos. Mag. Suppl.) **12**, 185 (1963).
[26] K. Tomita, "Relativistic Theory of Gravitational Instability and Inhomogeneous Cosmological Models," in *Bianchi Universes and Relativistic Cosmology*, edited by R. Ruffini (in preparation).
[27] B.K. Berger, D.M. Eardley, and D.W. Olson, Phys. Rev. D. **16**, 3086 (1977).
[28] G.F.R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology, Proceedings of the International School of Physics "Enrico Fermi" Course* XLVII, edited by R.K. Sachs (Academic, London and New York, 1969), p. 104.
[29] M.A.H. MacCallum, "Cosmological Models from a Geometric Point of View," in *Cargèse Lectures in Physics* **6**, *Lectures at the International Summer School of Physics, Cargèse, Corsica*, 1971, edited by E. Schatzman (Gordon and Breach, New York, 1973), p. 61.
[30] A. Raychaudhuri, Phys. Rev. **98**, 1123 (1955).

A new approach to inhomogeneous cosmologies: Intrinsic symmetries. II. Conformally flat slices and an invariant classification

D. A. Szafron[a)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1 and Department of Mathematics, Portland State University, P. O. Box 751, Portland, Oregon 97207 [b)]
C. B. Collins[c)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

(Received 13 October 1978)

We continue our analysis of space–times with intrinsic symmetries, by specializing to the case where the matter moves orthogonally to a family of conformally flat hypersurfaces. In so doing, we obtain a characterization of the Szekeres inhomogeneous cosmological models as being those (perfect fluid) solutions of Einstein's equations with conformally flat comoving slices whose second fundamental form and Ricci tensor possess two equal eigenvalues. An invariant characterization of these models is thereby obtained, and they are then classified in an invariant way.

I. INTRODUCTION

Having established a broad classification of inhomogenous cosmologies in a previous article[1] (hereafter referred to as I), we now specialize to the case where the fluid flow is orthogonal to a family of conformally flat hypersurfaces.

Szekeres[2] recently examined a class of inhomogeneous dust models in which it was assumed that the metric could be written in the form

$$ds^2 = -dt^2 + e^{2A}dx^2 + e^{2B}(dy^2+dz^2), \tag{1.1}$$

where $A = A(t,x,y,z)$, $B = B(t,x,y,z)$, and where the fluid flow vector $\mathbf{u} = \partial/\partial t$. Szafron[3] later extended this study to the perfect fluid case, in which the pressure, p, necessarily satisfies $p = p(t)$ (cf. the generalization by Bondi[4] of the spherically symmetric dust solutions of Tolman[5] and Bondi[4]. We shall henceforth call all space–times with metric (1.1) "Szekeres models." All Szekeres models have the property that the matter is flowing orthogonally to the $\{t = \text{constant}\}$ hypersurfaces, which are conformally flat (see Theorem 2.4 of Sec. 2). This result leads to an invariant characterization of the Szekeres models, in terms of the fluid congruence. Moreover, we can subclassify these models according to the vanishing of various invariants, and from this we obtain a detailed hierachy of all Szekeres space–times.

In I, we considered the progressive imposition of various symmetries, which, for convenience, we repeat here. We use the notation and conventions of I throughout. A "cosmological model" is a model satisfying condition

(i) a solution of Einstein's field equations

$$R_{ij} - \tfrac{1}{2}Rg_{ij} + \Lambda g_{ij} = T_{ij},$$

in which the distribution of matter is approximated by a smeared out fluid whose total energy density is nonzero.

We shall consider in the present article cosmological models satisfying the additional conditions

(ii) the cosmological constant, Λ, is zero,

(iii) the matter content is a perfect fluid, whose flowlines form a geodesic congruence orthogonal to a family, $\mathscr{F}$, of spacelike hypersurfaces,

and

(iv) each spacelike hypersurface, $\mathscr{S}$, in the family $\mathscr{F}$ is a conformally flat 3-manifold.

Later, we shall specialize condition (iv) to either of conditions (v) or (vi) below:

(va) each hypersurface $\mathscr{S}$ is a 3-space of constant curvature,

(vb) the Ricci tensor, R^{*}_{ij}, of each hypersurface $\mathscr{S}$ is isotropic, i.e., $R^{*}_{ij} = \frac{1}{3}R^{*}h_{ij}$, where h_{ij} is the metric induced on $\mathscr{S}$ by the space–time metric,

(vc) each hypersurface $\mathscr{S}$ admits a maximal group of isometries, (i.e., a 6-parameter intrinsic isometry group,

and

(vi) each hypersurface $\mathscr{S}$ is flat. It was pointed out in I that condition (vi)$\Longrightarrow$(va)$\Longleftrightarrow$(vb)$\Longleftrightarrow$ (vc)$\Longrightarrow$(iv) (see, e.g., Ref. 6).

Many of our results will be subject to the further condition

(vii) both the second fundamental form and the Ricci tensor of the hypersurfaces $\mathscr{S}$ possess two equal eigenvalues, and for some of our results to hold we require either

(viiia) $\mu + p > 0, \quad \mu + 3p > 0,$

or

(viiib) an equation of state, $p = p(\mu)$, with $dp/d\mu \geqslant 0$, relating the energy density, μ, to the fluid pressure, p.

[a)]Supported by a Postgraduate Fellowship from the National Research Council of Canada.
[b)]Permanent address.
[c)]Supported in part by Operating Grant #A3978 from the National Research Council of Canada, and by a University of Waterloo Research Grant.

In Sec. 2 we show that cosmological models satisfy conditions (i)–(iv), and (vii) if and only if they belong to the Szekeres class. These models are then subdivided into invariant subclasses.

2. CONFORMALLY FLAT SLICES

We shall be concerned with a subclass of those models in Table II of I, namely, those which belong to type IIIc in our classification, for which the invariants $\tau = T = 0$ and for which the hypersurfaces $\mathscr{S}$ are conformally flat. The condition $\tau = 0$ is equivalent to requiring that the shear tensor, σ_{ij}, have two equal eigenvalues, whereas the condition $T = 0$ is equivalent to requiring that the 3-Ricci tensor R^*_{ij} have two equal eigenvalues, where in each case the eigenvalues are associated with spacelike eigenvectors (the timelike vector $\mathbf{u}$ is an eigenvector of σ_{ij} and R^*_{ij}, with zero eigenvalues; when in the future we refer to eigenvalues of tensors orthogonal to $\mathbf{u}$, we shall have in mind those associated with eigenvectors orthogonal to $\mathbf{u}$). The subclass of models we consider therefore consists precisely of *those models satisfying conditions* (*i*)–(*iv*) *and* (*vii*) *of Sec.* 1. In the particular case where $\sigma \equiv 0$, these conditions imply that the models are isotropic,[7] and we shall frequently ignore this case. Thus we shall often require $\sigma \not\equiv 0$; then the expansion tensor $\theta_{\alpha\beta}$ will have the canonical form $(\theta_1, \theta_2, \theta_2)$ of Table I in I, where $\theta_1 \neq \theta_2 \Longleftrightarrow \sigma \neq 0$.

The Jacobi and field equation (I,B1)–(I,B8) are now specialized to the situation where conditions (i)–(iii) and (vii) (i.e., $\tau = T = 0$) hold, and where the orthonormal tetrad is a Fermi-propagated shear eigenframe [i.e., $\mathbf{\Omega} = \mathbf{0}$ and $\theta_{\alpha\beta} = \operatorname{diag}(\theta_1, \theta_2, \theta_2)$; cf. I, Proposition 4.3]. From I, Proposition 3.1 it follows that, in this frame, $R^*_{\alpha\beta} = \operatorname{diag}(R^*_{11}, R^*_{22}, R^*_{22})$. The equations are written out explicitly in Appendix A.

We shall find it convenient to introduce an algebraic classification of our models, based on the vanishing on an open set not only of the invariants R^* and S (we shall have $\theta\sigma \not\equiv 0$ and $\tau = T = 0$), but also of the other algebraically independent quantities that are invariant under a spacelike rotation about the "1" direction. In fact, from I, Theorem 4.2, we find that there are five algebraic invariants I_1–I_5 associated with the intrinsic curvature of the hypersurfaces $\mathscr{S}$, together with two invariants, θ_1 and θ_2, associated with the extrinsic curvature. We shall base the classification on the vanishing on an open set of the invariants R^*, S, I_1, and I_2. This is because the field equations of general relativity will require $I_3 = I_4 = 0$ (see Theorem 2.3 below), in which case I_5 is not defined and the only algebraically independent invariants associated with the intrinsic curvature are I_1 and I_2 (see remark after I, Lemma 4.1). The classification is simplified somewhat by the results of I, Proposition 3.2, which shows that the case $R^* = 0, S \neq 0$ is inadmissible, and by the Corollary of Lemma 2.1, which shows that $\sigma = 0$, if $S = 0$ and $R^* \neq 0$. We also choose to classify not according to the vanishing of the invariants θ_1 and θ_2, but according to the vanishing of the algebraically dependent quantities $\theta = \theta_1 + 2\theta_2$ and $\sigma = (1/\sqrt{3})|\theta_1 - \theta_2|$. Since for reasonable matter content we have $\theta \not\equiv 0$ from I, Proposition 3.3, and since we assume $\sigma \not\equiv 0$, no further specializations occur.

Lemma 2.1: In any cosmological model satisfying conditions (i)–(iv) and (vii), either the invariants $I_3 = I_4 = 0$, or the spacelike hypersurfaces $\mathscr{S}$ are flat, or the model is FRW.

Proof: The hypersurfaces $\mathscr{S}$ satisfy condition (iv) (i.e., are conformally flat) if and only if the tensor

$$C_{ijk} \equiv R^*_{ij|k} - R^*_{ik|j} - \tfrac14 (h_{ij} R^*_{|k} - h_{ik} R^*_{|j}) = 0, \tag{2.1}$$

where | denotes covariant derivative with respect to the metric induced on $\mathscr{S}$ (Ref. 6). Using the terminology of I, Eq. (2.1) becomes

$$n_{11}(R^*_{11} - R^*_{22}) = 0, \tag{2.2}$$

$$(n_{33} - n_{22} - n_{11})(R^*_{11} - R^*_{22}) = 0, \tag{2.3}$$

$$\partial_2(R^*_{11} - \tfrac14 R^*) + (n_{13} - a_2)(R^*_{11} - R^*_{22}) = 0, \tag{2.4}$$

$$\partial_3(R^*_{22} - \tfrac14 R^*) = 0, \tag{2.5}$$

$$\partial_1(R^*_{22} - \tfrac14 R^*) + (a_1 - n_{23})(R^*_{11} - R^*_{22}) = 0, \tag{2.6}$$

$$\partial_3(R^*_{11} - \tfrac14 R^*) - (n_{12} + a_3)(R^*_{11} - R^*_{22}) = 0, \tag{2.7}$$

$$\partial_2(R^*_{22} - \tfrac14 R^*) = 0, \tag{2.8}$$

and

$$\partial_1(R^*_{22} - \tfrac14 R^*) + (a_1 + n_{23})(R^*_{11} - R^*_{22}) = 0. \tag{2.9}$$

Ostensibly $C_{ijk} = C_{i[jk]}$ has nine independent components. However $\eta^{ijk} C_{ijk} = 0$, and $C^i_{\ ij} = 0$ by the contracted Bianchi identities, and so there are essentially only five independent equations in (2.1). York[8] (cf. Ref. 9) has elegantly expressed the five independent components of C_{ijk} in terms of of a tracefree symmetric tensor, Y_{ij}; this is done by employing a decomposition analogous to that of $\gamma^\alpha_{\beta\lambda}$ in [I, Eq. (2.3)], where we now have that the analog of $a_\alpha \equiv \frac12 \gamma^\beta_{\alpha\beta}$ is zero, and that the analog of $n^{\alpha\beta} \equiv \frac12 \gamma^{(\alpha}{}_{\nu\lambda}\epsilon^{\beta)\nu\lambda}$ is tracefree. However, we shall not use this result, and instead we find it more convenient to use Eqs. (2.2)–(2.9).

The proof now divides into two parts, depending on whether or not R^*_{ij} is isotropic.

Case 1. $R^*_{11} \neq R^*_{22}$: By Eqs. (2.2) and (2.3), it follows that $n_{11} = n_{22} - n_{33} = 0$. Equations (2.6) and (2.9) require $n_{23} = 0$. Thus we have $I_3 = I_4 = 0$.

Case 2. $R^*_{11} = R^*_{22}$: In this case $R^*_{\alpha\beta} = \frac13 R^* \delta_{\alpha\beta}$, i.e., $R^*_{\alpha\beta}$ is isotropic. If $R^* = 0$, the hypersurfaces are flat (since their Riemann tensor vanishes). We shall therefore assume $R^* \neq 0$, and we will show that the model is FRW. Equations (2.2)–(2.9) are equivalent to

$$\partial_1 R^* = \partial_2 R^* = \partial_3 R^* = 0. \tag{2.10}$$

A straightforward computation gives an evolution equation for R^*,

$$\dot{R}^* = -\tfrac23 R^*(\theta_1 + 2\theta_2). \tag{2.11}$$

Equation (2.11) is obtained using (A22) in the form

$$\mu = \theta_2(2\theta_1 + \theta_2) + \tfrac12 R^*. \tag{2.12}$$

Substituting (2.12) into Eq. (A13), we obtain

$$p = -\tfrac{1}{3}[2(\theta_1 + 2\theta_2)] + 2\theta_1^2 + 2\theta_1\theta_2 + 5\theta_2^2 + \tfrac{1}{2}R^*. \tag{2.13}$$

By differentiating Eq. (2.12) along $\mathbf{e}_0$, and using the energy conservation equation (A24) in conjunction with (2.13) and the shear propagation equation (A23), we finally obtain (2.11).

The $[\mathbf{e}_0, \mathbf{e}_\alpha]$ commutation relation (I, A1–A3), together with (2.10), show that $\partial_\alpha R^* = 0$, so from (2.11), $\partial_\alpha(\theta_1 + 2\theta_2) = 0$. The same commutation relations now show that $\partial_\alpha(\dot\theta_1 + 2\dot\theta_2) = 0$. Differentiating (2.13) along $\mathbf{e}_\alpha$, and recalling the momentum conservation equations (A25), we obtain $(\theta_1 - \theta_2)\,\partial_\alpha\theta_2 = 0$. The case $\theta_1 = \theta_2$ corresponds to a FRW solution.[7] On the other hand, if $\theta_1 \neq \theta_2$, then $\partial_\alpha\theta_1 = \partial_\alpha\theta_2 = 0$. The field equations (A15) and (A16) show that $n_{31} = a_2$ and $n_{12} = -a_3$, and consequently, using (A22),

$$-R^* = -4(\partial_1 a_1 + \partial_2 a_2 + \partial_3 a_3 + 6a_1^2 + 8(a_2^2 + a_3^2) + \tfrac{1}{2}(n_{11}^2 + n_{22}^2 + n_{33}^2) - n_{11}(n_{22} + n_{33}) - n_{22}n_{33} + 2n_{23}^2. \tag{2.14}$$

Differentiation of this along $\mathbf{e}_0$, using (2.11), the evolution equations (A4)–(A10) for $n_{\alpha\beta}$ and a_α, and substituting (2.14), gives, on division by $\theta_1 - \theta_2$,

$$-\tfrac{2}{3}R^* + 8\partial_1 a_1 - 12a_1^2 + n_{11}^2 - (n_{22} - n_{33})^2 - 4n_{23}^2 = 0. \tag{2.15}$$

Repeating this process of differentiation, substitution of (2.15), and division by $(\theta_1 - \theta_2)$, we obtain

$$R^* = -\tfrac{9}{2}n_{11}^2. \tag{2.16}$$

Differentiation of (2.16) yields $R^* = 0$, which contradicts our assumption. Consequently the only cases to arise when $\overset{*}{R}_{\alpha\beta}$ is isotropic occur when the hypersurfaces $\mathscr{S}$ are flat or when the model is FRW. □

Corollary: Any cosmological model satisfying conditions (i)–(iii), (v), and (vii) either admits flat hypersurfaces $\mathscr{S}$, or is FRW. □

Proof: If condition (v) holds, then (iv) is necessarily true. The proof of the Lemma shows that in the case where condition (v) is satisfied (i.e., $\overset{*}{R}_{\alpha\beta}$ is isotropic), either $\theta_1 = \theta_2$ or $\overset{*}{R}_{\alpha\beta} = 0$. □

Remarks: Note that Lemma 2.1 and its Corollary hold regardless of there being energy restrictions or an equation of state on the matter.

Lemma 2.2: In any cosmological model satisfying conditions (i)–(iii), (vi), and (vii), either $\sigma \equiv 0$, and the model is FRW, or $\sigma \not\equiv 0$, and the invariant $I_3 = 0$.

Proof: Either $\theta_1 = \theta_2$, and the model is FRW,[7] or $\theta_1 \neq \theta_2$, which we shall henceforth assume. We start with the equation obtained by eliminating $\partial_1 a_3$ between Eqs. (A2) and (A20), in the form

$$\begin{aligned}\partial_2 n_{11} = &-\partial_2(n_{22} - n_{33}) - 2\partial_3(a_1 + n_{23})\\ &- 2n_{11}(n_{31} - a_2) + 2(n_{22} - n_{33})(n_{31} + a_2)\\ &- 4n_{23}(n_{12} - a_3).\end{aligned} \tag{2.17}$$

The technique we shall use is an extension of that used in the proof of Lemma 2.1 We repeatedly differentiate (2.17) along $\mathbf{e}_0$. Second derivative terms are eliminated by applying the commutation relations (I,A1–A3) and the Jacobi identities (A1)–(A12), and as many first derivative terms as possible are eliminated, again using the Jacobi identities. Then, substituting the original equation (2.17), we eliminate one of the terms. Thus, differentiating (2.17), we obtain

$$\begin{aligned}&2(\theta_1 - \theta_2)\partial_2(n_{22} - n_{33}) + (n_{22} - n_{33})\partial_2(\theta_1 - 2\theta_2)\\ &\quad + 2(\theta_1 - \theta_2)\partial_3(a_1 + 2n_{23}) + 2(a_1 - 2n_{23})\partial_3\theta_2\\ &\quad + 2n_{23}\partial_3\theta_1 - 4(\theta_1 - \theta_2)(n_{22} - n_{33})(n_{31} + a_2)\\ &\quad + 8(\theta_1 - \theta_2)n_{23}(n_{12} - a_3) - n_{11}\partial_2(3\theta_1 - 2\theta_2) = 0.\end{aligned} \tag{2.18}$$

Eliminating $\partial_2(n_{22} - n_{33})$ between (2.17) and (2.18), and dividing by $(\theta_1 - \theta_2)$ yields

$$\begin{aligned}&2(\theta_1 - \theta_2)\partial_3 a_1 + 2a_1\partial_3\theta_2 - 6n_{23}\partial_3\theta_1 + 2n_{23}\partial_3\theta_2 + (n_{22} - n_{33})\\ &\quad \times\partial_2\theta_2 - 3(n_{22} - n_{33})\partial_2\theta_1 - 3n_{11}\partial_2\theta_1 + 9n_{11}\partial_2\theta_2 = 0.\end{aligned} \tag{2.19}$$

Differentiation of (2.19) along $\mathbf{e}_0$, elimination of $\partial_3 a_1$, and division by $(\theta_1 - \theta_2)$ gives

$$\begin{aligned}&4n_{23}\partial_3\theta_1 + 2n_{23}\partial_3\theta_2 + (n_{22} - n_{33})\partial_2\theta_2 + 2(n_{22} - n_{33})\partial_2\theta_1\\ &\quad + 7n_{11}\partial_2\theta_2 = 0.\end{aligned} \tag{2.20}$$

Finally, differentiation of (2.20) along $\mathbf{e}_0$, elimination of $\partial_3\theta_1$, and division by $(\theta_1 - \theta_2)$ gives

$$n_{11}\partial_2\theta_2 = 0. \tag{2.21a}$$

In a similar manner, repeated differentiation of the equation obtained by eliminating $\partial_1 a_2$ between Eqs. (A3) and (A21) gives

$$n_{11}\partial_3\theta_2 = 0. \tag{2.21b}$$

This can be recognized immediately by rotating the frame at each point about the $(\mathbf{e}_2 + \mathbf{e}_3)$-direction: $\mathbf{e}_1 \to -\mathbf{e}_1$, $\mathbf{e}_2 \to \mathbf{e}_3, \mathbf{e}_3 \to \mathbf{e}_2$ and observing the transformation laws of the commutation functions, γ^a_{bc}. Thus, each Eq. (2.7)–(2.21b) will have a counterpart under this rotation.

We now show that Eqs. (2.21a) and (2.21b) imply $n_{11} = 0$. If this were false, i.e., if $n_{11} \neq 0$, we have $\partial_2\theta_2 = \partial_3\theta_2 = 0$. Then $[\mathbf{e}_2, \mathbf{e}_3]\theta_2 = 0 \Rightarrow n_{11}\partial_1\theta_2 = 0 \Rightarrow \partial_1\theta_2 = 0$. Now Eq. (A14) implies $a_1 = 0$, in which case Eqs. (2.19) and (2.20) require $\partial_2\theta_1 = 0$. Similarly, the equations resulting from differentiation of Eqs. (A3) and (A21) require $\partial_3\theta_1 = 0$, and then $[\mathbf{e}_2, \mathbf{e}_3]\theta_1 = 0 \Rightarrow \partial_1\theta_1 = 0$. By Eq. (2.12) in the case $R^* = 0$ it follows that $\partial_\alpha\mu = 0$ $(\alpha = 1,2,3)$, and we already have $\partial_\alpha p = 0$ $(\alpha = 1,2,3)$, by Eq. (A25). It follows from Proposition 2.2 in a subsequent article in the present series that this model is spatially homogeneous of Bianchi type I, and is necessarily LRS.[10] Since for a Bianchi I model there is a Fermi-propagated shear eigenframe in which $n_{11} = 0$ (Ref. 10) and since n_{11} is invariant, it follows that $n_{11} = 0$ in our tetrad, which contradicts our assumption that $n_{11} \neq 0$. Consequently either the model is FRW or $I_3 = n_{11} = 0$. □

Remark: The repeated differentiations leading from (2.17) to (2.21b) were checked using a REDUCE program, written by J. Wainwright.

Theorem 2.3: In any cosmological model satisfying conditions (i)–(iv) and (vii), either $\sigma\equiv 0$, and the model is FRW, or $\sigma\not\equiv 0$, and $I_3 = I_4 = 0$.

Proof: By Lemmas 2.1 and 2.2, either the model is FRW, or $I_3 = I_4 = 0$, or the hypersurfaces $\mathscr{S}$ are flat and $I_3 = 0 \neq I_4$. We show that the last possibility is untenable. We first prove that with $R^{*}_{\alpha\beta} = 0$, $I_3 = 0 \neq I_4 \Rightarrow I_2 = 0$, and deduce that $I_2 = I_3 = 0 \neq I_4$ implies a contradiction.

(i) $R^{*}_{\alpha\beta} = 0, I_3 = 0, I_4 \neq 0 \Rightarrow I_2 = 0$: The sum of Eqs. (2.18), (2.20), and $2(\theta_1 - \theta_2)\times$ Eq. (2.17) gives, on comparison with Eq. (2.19)

$$a_1\partial_3\theta_2 = 0, \tag{2.22}$$

and, by symmetry,

$$a_1\partial_2\theta_2 = 0. \tag{2.23}$$

By a suitable rotation, we can choose a shear eigenframe on an initial hypersurface with the property that $n_{23} = 0$ (see I, Proposition 4.3 and the proof of I, Theorem 4.2). The Jacobi Eq. (A10) shows that $n_{23}\equiv 0$, in which case we obtain from Eq. (2.20) and its counterpart

$$\partial_2(2\theta_1 + \theta_2) = \partial_3(2\theta_1 + \theta_2) = 0, \tag{2.24}$$

recalling that now $I_4 \neq 0 \Longleftrightarrow n_{22} - n_{33} \neq 0$. If $a_1 \neq 0$, then (2.22)–(2.24) show that $\partial_2\theta_1 = \partial_3\theta_1 = \partial_2\theta_2 = \partial_3\theta_2 = 0$. If, on the other hand, $a_1 = 0$, we obtain from Eq. (2.19) and its counterpart that

$$\partial_2(3\theta_1 - \theta_2) = \partial_3(3\theta_1 - \theta_2) = 0. \tag{2.25}$$

Now Eq. (2.24) and (2.25) require $\partial_2\theta_1 = \partial_3\theta_1 = \partial_2\theta_2 = \partial_3\theta_2 = 0$.

Consequently, in either case

$$\partial_2\theta_1 = \partial_3\theta_1 = \partial_2\theta_2 = \partial_3\theta_2 = 0, \tag{2.26}$$

and now Eqs. (A15) and (A16) imply $n_{31} - a_2 = n_{12} + a_3 = 0$, i.e., $I_2 = 0$.

(ii) $R^{*}_{\alpha\beta} = 0$, $I_2 = I_3 = 0 \Rightarrow I_4 = 0$: Assume that $I_4 \neq 0$. Choosing a frame in which $n_{23}\equiv 0$, we have from part (i) of the proof that (2.26) holds. The commutators $[\mathbf{e}_1, \mathbf{e}_2]$ and $[\mathbf{e}_3, \mathbf{e}_1]$ acting on θ_2, together with Eq. (A14), require

$$\partial_2 a_1 = \partial_3 a_1 = 0; \tag{2.27}$$

application of the same commutators on a_1 gives, on using (2.27),

$$\partial_2(\partial_1 a_1) = \partial_3(\partial_1 a_1) = 0. \tag{2.28}$$

From Eq. (A17) we obtain

$$R^{*}_{11} = 0 \Longleftrightarrow \partial_1 a_1 = a_1^2 + \tfrac{1}{4}(n_{22} - n_{33})^2. \tag{2.29}$$

The derivatives of (2.29) along $\mathbf{e}_2$ and $\mathbf{e}_3$ give, using (2.27) and (2.28),

$$(n_{22} - n_{33})\,\partial_2(n_{22} - n_{33}) = (n_{22} - n_{33})\,\partial_3(n_{22} - n_{33}) = 0. \tag{2.30}$$

Moreover, subtracting Eqs. (A18a) and (A18b), and recalling $I_2 = 0$, we find that

$$(n_{22} - n_{33})\,(n_{22} + n_{33}) = 0. \tag{2.31}$$

Recall $I_4 \neq 0 \Longleftrightarrow (n_{22} - n_{33}) \neq 0$, Then Eq. (2.31) requires $n_{33} = -n_{22} \neq 0$, and so, from (2.30)

$$\partial_2 n_{22} = \partial_3 n_{22} = 0. \tag{2.32}$$

Equation (2.17) shows that $a_2 = 0$, and by symmetry we obtain the counterpart, $a_3 = 0$. Substituting into Eqs. (A17) and (A18a), and eliminating $\partial_1 a_1$, we obtain

$$a_1 = \pm\, n_{22} \neq 0. \tag{2.33}$$

In order for Eq. (2.33) to be valid for all time, Eqs. (A4) and (A8) require $\partial_1\theta_2 = 0$. But then Eq. (A14) requires $a_1 = 0$, which contradicts (2.33). It follows that $I_4 = 0$. □

We can extend the result of Theorem 2.3, and so obtain a characterization of the Szekeres models.

Theorem 2.4 (*Characterization of the Szekeres models*: cf. Refs. 3,11,12): For models satisfying conditions (i)–(iii) and (vii), with $\sigma \neq 0$, the following are equivalent:

(a) Condition (iv), i.e., conformally flat hypersurfaces $\mathscr{S}$;

(b) $I_3 = I_4 = 0$;

(c) The model is of the Szekeres class.

Proof: Theorem 2.3 has shown that (a)⇒(b). We complete the proof by showing that (b)⇒(c) and that (c)⇒(a).

(i) (b)⇒(c): We choose a Fermi-propagated shear eigenframe for which $n_{11} = n_{22} - n_{33} = n_{23} = 0$. By I, Proposition 4.3, and the proof of I, Theorem 4.2, we can arrange for $n_{22} = 0$ everywhere on an initial hypersurface $\mathscr{S}$, and by (A8) it follows that $n_{22}\equiv 0$, and hence $n_{33}\equiv 0$. From the commutation relations (I,A1–A6), we find that the integral curves of $\mathbf{e}_0$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are each orthogonal to hypersurfaces, which we label with coordinates t,x,y, and z respectively. It follows that $\mathbf{e}_a(x^i) \propto \delta^i_a$, and hence

$$\mathbf{e}_0 = \frac{1}{T}\frac{\partial}{\partial t},\quad \mathbf{e}_1 = \frac{1}{X}\frac{\partial}{\partial x},\quad \mathbf{e}_2 = \frac{1}{Y}\frac{\partial}{\partial y},$$

and

$$\mathbf{e}_3 = \frac{1}{Z}\frac{\partial}{\partial z},$$

where T,X,Y, and Z are, *a priori*, functions of all four coordinates. From the fact that $\mathbf{e}_0 = \mathbf{u}$ is geodesic, we obtain $\mathbf{e}_0 \cdot [\mathbf{e}_0, \mathbf{e}_\alpha] = 0 \Longleftrightarrow T = T(t)$; by suitably redefining the coordinate t we obtain $T = 1$. The metric in these coordinates is obtained from (I, 2.1) and is thus of the form

$$ds^2 = -dt^2 + X^2\,dx^2 + Y^2\,dy^2 + Z^2 dz^2. \tag{2.34}$$

Using the commutation relations (I, A1)–(I, A6) again, we see that: $n_{23} = 0 \Longleftrightarrow \mathbf{e}_2 \cdot [\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_3 \cdot [\mathbf{e}_1, \mathbf{e}_3] \Longleftrightarrow Y_x/Y = Z_x/Z$ (where a subscript denotes differentiation). The condition that the shear tensor have two equal eigenvalues becomes: $\theta_{22} = \theta_{33} \Longleftrightarrow \mathbf{e}_2 \cdot [\mathbf{e}_0, \mathbf{e}_2] = \mathbf{e}_3 \cdot [\mathbf{e}_0, \mathbf{e}_3] \Longleftrightarrow Y_t/Y = Z_t/Z$. Thus $Z = Yf(y,z)$, for some function $f(y,z)$, and (2.34) becomes

$$ds^2 = -dt^2 + X^2 dx^2 + Y^2(dy^2 + f^2(y,z)\,dz^2). \tag{2.35}$$

Since any 2-metric is conformally flat, there is a transformation of coordinates $y' = y'(y,z)$, $z' = z'(y,z)$ such that $dy^2 + f^2(y,z)\,dz^2 = F^2(y',z')\,(dy'^2 + dz'^2)$; absorbing $F(y',z')$ into Y in (2.35), and dropping the primes, we see that the metric is of the Szekeres form

$$ds^2 = -dt^2 + e^{2A}\,dx^2 + e^{2B}\,(dy^2 + dz^2), \tag{2.36}$$

where $X \equiv e^A$, $Y \equiv e^B$, $A = A(t,x,y,z)$, and $B = B(t,x,y,z)$.

(ii)(c)$\Rightarrow$(a): We use the orthonormal tetrad $\mathbf{e}_0 = \partial/\partial t$, $\mathbf{e}_1 = e^{-A}\partial/\partial x$, $\mathbf{e}_2 = e^{-B}\partial/\partial y$, and $\mathbf{e}_3 = e^{-B}\partial/\partial z$. In this tetrad, $n_{11} = n_{22} = n_{23} = n_{33} = 0$. The necessary and sufficient conditions for the slices $\{t = \text{constant}\}$ to be conformally flat are given by Eqs. (2.2)–(2.9) With our tetrad specialization, Eqs. (2.2) and (2.3) are identically satisfied, and Eqs. (2.6) and (2.9) are equivalent. We must therefore verify the remaining equations (2.4)–(2.8). In doing so, we shall use a result of Szekeres[2] and of Szafron[3] that (in our notation) with conditions (i)–(iii) and (vii) in force,

$$\dot{B}_y = \dot{B}_z = 0 \Longleftrightarrow \partial_2\theta_2 = \partial_3\theta_2 = 0. \qquad (.2.37)$$

Moreover, from Eqs. (A15) and (A16)

$$\partial_2\theta_1 = (\theta_1 - \theta_2)(a_2 - n_{31}) \qquad (2.38)$$

and

$$\partial_3\theta_1 = (\theta_1 - \theta_2)(a_3 + n_{12}). \qquad (2.39)$$

It follows from the field equations that

$$R^{*}_{11} - \tfrac{1}{4}R^{*} = -\dot{\theta}_1 - \theta_1^2 - \theta_1\theta_2 + \tfrac{1}{2}\theta_2^2 - \tfrac{1}{2}p, \qquad (2.40)$$

and

$$R^{*}_{22} - \tfrac{1}{4}R^{*} = -\dot{\theta}_2 - \tfrac{3}{2}\theta_2^2 - \tfrac{1}{2}p. \qquad (2.41)$$

The derivatives of Eq. (2.41) along $\mathbf{e}_2$ and $\mathbf{e}_3$ give (2.5) and (2.8), on using (2.37), (A25), and the commutation relations, (I,A1)–(I,A6). The derivative of (2.40) along $\mathbf{e}_2$ gives Eq. (2.4), on using Eqs. (2.37), (2.38), (A5), (A11), (A23), (A25), and the commutation relations (I,A1.1–A1.6). Equation (2.7) is similarly verified [it is the counterpart of (2.4) under the rotation $\mathbf{e}_1 \to -\mathbf{e}_1$, $\mathbf{e}_2 \to \mathbf{e}_3$, $\mathbf{e}_3 \to \mathbf{e}_2$], by differentiating (2.40) along $\mathbf{e}_3$ and using Eqs. (2.37) (2.39), (A6), (A12), (A23), (A25) and the commutation relations (I,A1)–(I,A6). Finally, the derivative of Eq. (2.41) along $\mathbf{e}_1$ gives Eq. (2.6), on using Eqs. (A4), (A14), (A23), (A25), and the commutation relations (I,A1)–(I,A6). □

Corollary 1: In the case where $\sigma \not\equiv 0$, the freedom of choice in the Szekeres coordinates (t,x,y,z) of (2.36), and in the associated orthonormal shear eigenframe

$$\mathbf{e}_0 = \frac{\partial}{\partial t}, \quad \mathbf{e}_1 = e^{-A}\frac{\partial}{\partial x},$$

$$\mathbf{e}_2 = e^{-B}\frac{\partial}{\partial y}, \quad \mathbf{e}_3 = e^{-B}\frac{\partial}{\partial z},$$

is as follows:

$$\mathbf{e}_0 \to \mathbf{e}_0, \quad \mathbf{e}_1 \to \mathbf{e}_1, \quad \mathbf{e}_2 \to \mathbf{e}_2\cos\psi + \mathbf{e}_3\sin\psi,$$

$$\mathbf{e}_3 \to -\mathbf{e}_2\sin\psi + \mathbf{e}_3\cos\psi,$$
$$t \to t + t_0, \quad X \to C(x), \quad y \to g(y+iz) + \bar{g}(y-iz),$$
$$z \to -i[g(y+iz) - \bar{g}(y-iz)],$$

where C and g are arbitrary functions satisfying $C'(x)g'(y+iz) \not\equiv 0$, t_0 is an arbitrary constant, and $\psi = \psi(y,z) = n\pi - \arg\{g'(y+iz)\}$, n being an integer. The metric functions A and B transform as

$$A \to A - \ln|C'(x)| \quad \text{and } B \to B - \ln|2g'(y+iz)|.$$

Proof: Using the results in the proof of I, Theorem 4.2 together with the preservation of $\sigma_{23} = \Omega_1 = n_{22} = 0$, we find that the allowable freedom in the tetrad is a rotation through ψ about the 1-direction and that $\partial_0\psi = \partial_1\psi = 0$. Consider now a general simultaneous change of coordinates, $x^{i'} = x^{i'}(x^j)$, and of tetrad, $\mathbf{e}_{i'} = L_{i'}{}^j\mathbf{e}_j$, subject to the conditions that $e_0{}^0 = 1$, $e_1{}^1 \neq 0$, $e_2{}^2 = e_3{}^3 \neq 0$, all other $e_i{}^a \equiv e_a(x^i) = 0$, both before and after the transformations. Then $e_{0'}{}^{0'} = 1$, $e_{\alpha'}{}^{0'} = 0 \Rightarrow x^{0'} = x^0 + \text{constant}$, while $e_{a'}{}^{1'} = 0$ $(a = 0,2,3) \Rightarrow x^{1'} = x^{1'}(x')$ (and $e_1{}^1 \neq 0 \Longleftrightarrow dx^{1'}/dx^1 \neq 0$). Also $e_{0'}{}^{a'} = e_{1'}{}^{a'} = 0$ $(a = 2,3)$ requires $y' = y'(y,z)$ and $z' = z'(y,z)$, where we have written $x^2 = y$ and $x^3 = z$ and the Jacobian $\partial(y',z')/\partial(y,z) \neq 0$. The constraints $e_{3'}{}^{2'} = e_{2'}{}^{3'} = 0$ require

and

$$-\sin\psi\frac{\partial y'}{\partial y} + \cos\psi\frac{\partial y'}{\partial z} = 0. \qquad (2.43)$$

Finally, $e_{2'}{}^{2'} = e_{3'}{}^{3'}$ requires

$$\cos\psi\left(\frac{\partial y'}{\partial y} - \frac{\partial z'}{\partial z}\right) + \sin\psi\left(\frac{\partial y'}{\partial z} + \frac{\partial z'}{\partial y}\right) = 0. \qquad (2.44)$$

Adding Eqs. (2.42) and (2.43), and eliminating ψ with (2.44), we obtain

$$\frac{\partial y'}{\partial y} = \frac{\partial z'}{\partial z} \quad \text{and} \quad \frac{\partial y'}{\partial z} = -\frac{\partial z'}{\partial y}, \qquad (2.45)$$

which are the Cauchy–Riemann equations obtained in a general (analytic) transformation of the complex variable $y + iz$ into $y' + iz' = g(y+iz)$, where g is arbitrary. Thus $y' = g(y+iz) + \bar{g}(y-iz)$ and $z' = -i[g(y+iz) - g(y-iz)]$ {Alternatively we can obtain from Eqs. (2.45) that

$$\frac{\partial^2 y'}{\partial y^2} + \frac{\partial^2 y'}{\partial z^2} = \frac{\partial^2 z'}{\partial y^2} + \frac{\partial^2 z'}{\partial z^2} = 0,$$

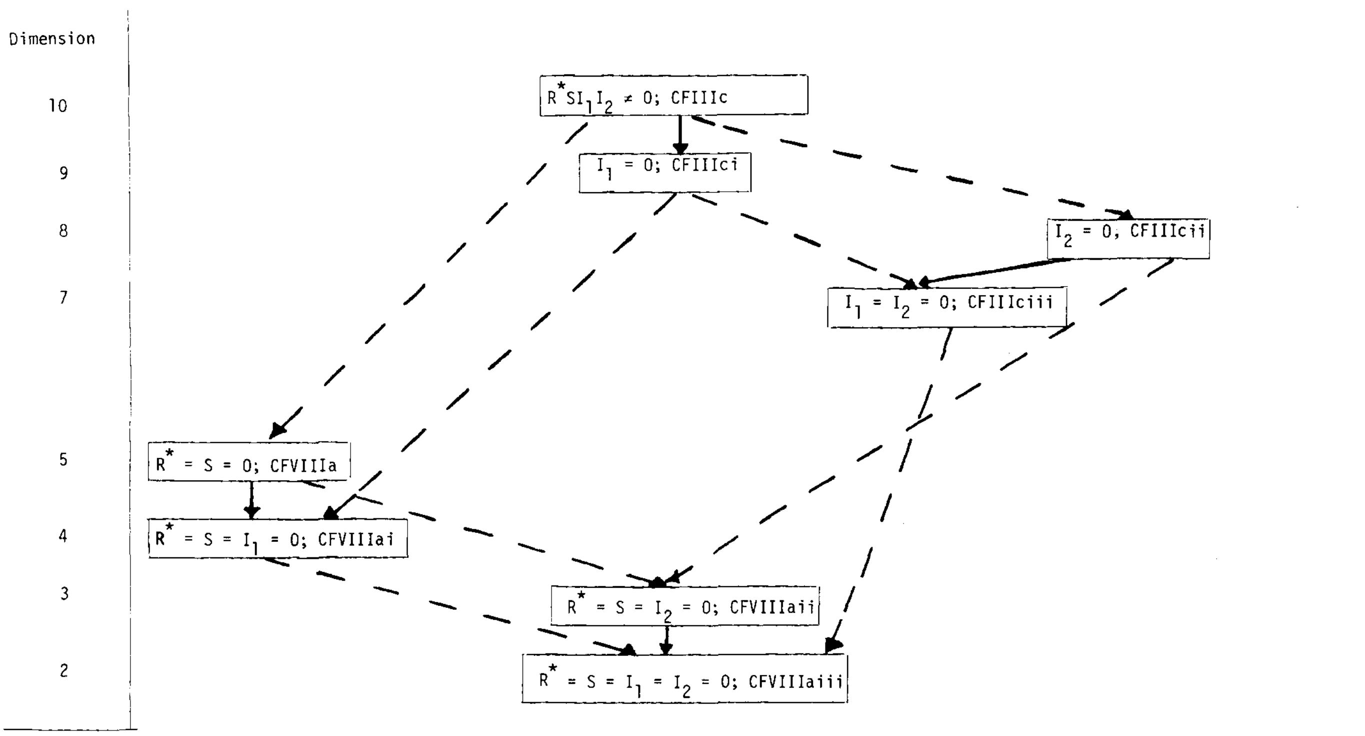
hence $y' = g(y+iz) + h(y-iz)$ and $z' = -i[g(y+iz) - h(y-iz)] + c$, where the functions g and h and the constant c are arbitrary. For real coordinates y' and z', h is the complex conjugate of g, and c is real. The constant c may be made zero by means of the transformations $g(\zeta) \to g(\zeta) + ic/2$, $h(\bar{\zeta}) = \bar{g}(\bar{\zeta}) \to \bar{g}(\bar{\zeta}) - ic/2$.} The Jacobian $\partial(y',z')/\partial(y,z)$ is nonzero if an only if $g'(\zeta) \not\equiv 0$. The only equation that remains to be satisfied is (2.42), which becomes $|g'(y+iz)|\sin(\psi + \arg[g'(y+iz)]) = 0$, i.e., $\psi(y,z) = n\pi - \arg[g'(y+iz)]$, where n is an integer.

The transformations of the metric functions are most readily calculated directly from the metric (2.36) □

Corollary 2: The subclassification of the Szekeres metrics, employed by Szekeres[2] and Szafron,[3] into Class I ($B_x \neq 0$) and Class II ($B_x = 0$) is an invariant classification in the case $\sigma \not\equiv 0$. It is not invariant in the FRW case $\sigma \equiv 0$.

Proof: In the orthonormal tetrad of Corollary 1, we find, using the commutation relations (I, A1)–(I,A6), that $I_1 = a_1 = -B_x e^{-A}$. The result follows in the case $\sigma \not\equiv 0$, using I, Theorem 4.2 and I, Proposition 4.3. The result for $\sigma \equiv 0$ follows by observing that both Class I and Class II admit the same FRW metrics. [In the notation of Szafron,[3] Class I metrics are FRW if $k = \text{constant}$ and $\phi = \phi(t)$ and Class II metrics are FRW if $\lambda = 0$; in each case $R^{*} = 6k/\phi^2$, and this can be positive, negative, or zero.] □

TABLE I. Specialization diagram for anisotropic cosmological models with conditions (i)–(iv), (vii), and (viiia) in force. Algebraic dimension is assigned on the left. Arrows denote specialization to models of lower dimension, a solid arrow denoting a reduction of one dimension, and a broken arrow denoting a reduction of more than one dimension.



Remarks: An invariant classification of the Szekeres line element (2.36) was first given for dust by Wainwright,[12] in terms of the principal null congruences. Szafron[3] generalized this result to the case of a perfect fluid. Berger, Eardley, and Olson[11] proved that the Szekeres dust models of Class I admit conformally flat comoving slices, and that the "marginally bound quasispherical" subset (Class CFVIIIaii in Table I above) have *flat* comoving slices. Wainwright and Szafron later proved that in *all* perfect Szekeres metrics, the comoving slices are conformally flat. Their proof essentially consisted of checking the validity of Eqs. (2.2)–(2.9) using a REDUCE program, written by Wainwright; it preceded our analytic proof.

The Petrov classification of the Szekeres models follows immediately upon using Eqs. (2.37) in conjunction with the tetrad components of the "electric" and "magnetic" parts of the Weyl tensor, $E_{\alpha\beta}$ and $H_{\alpha\beta}$, respectively (cf. Theorem 2.4 of a subsequent article in this series). Following Ellis,[7] the Weyl tensor C_{abcd} may be decomposed into the "electric" part, E_{ab}, with

$$E_{ac} \equiv C_{abcd} u^b u^d, \quad E_{ab} = E_{(ab)}, \quad E_a{}^a = 0, \text{ and}$$

$$E_{ab} u^b = 0,$$

and the "magnetic" part, H_{ab}, with

$$H_{ac} \equiv \tfrac{1}{2}\, \eta_{ab}{}^{gh} C_{ghcd}\, u^b\, u^d, \quad H_{ab} = H_{(ab)}, \quad H_a{}^a = 0,$$

and

$$H_{ab}\, u^b = 0.$$

MacCallum[13] has expressed the components of E_{ab} and H_{ab} in a general orthonormal tetrad. In the Fermi-propagated shear eigenframe of Corollary 1, we obtain $E_{\alpha\beta} = \operatorname{diag}(-2E,E,E)$ and $H_{\alpha\beta} = 0$. By evaluating the eigenvalues and eigenvectors of $E_{\alpha\beta} + iH_{\alpha\beta}$ (cf. Ref. 14), we find that the models are of Petrov type D if $E \neq 0$, and of type O (i.e., conformally flat, and hence FRW[7]) if $E = 0$. This result is in agreement with those of Wainwright[12] in the dust case, and of Szafron[3] in the perfect fluid case.

In Appendix B, we give the commutation functions, Jacobi identities, and field equations for the Szekeres class in the orthonormal tetrad of Corollary 1.

We now classify those space–times satisfying conditions (i)–(iv), (vii), and (viiia), with $\sigma \not\equiv 0$. This classification is based on the vanishing of the invariant quantities R^*, S, I_1, and I_2, bearing in mind the results of the Corollary to Lemma 2.1 and of I, Proposition 3.2. The results of this are given in a "specialization diagram," Table I.

The tensor $R^*_{\alpha\beta}$ possesses 5 algebraically independent components, since it has two equal eigenvalues. Since $I_3 = I_4 = 0$, and since we can make $n_{22} \equiv 0$ (see I, Proposition 4.3 and the proof of I, Theorem 4.2), there are five algebraically independent quantities amongst $n_{\alpha\beta}$ and a_β. Consequently we can ascribe a dimension of 10 to the most general case. The vanishing of R^* or I_1 alone reduces the dimension by 1, the vanishing of I_2 alone reduces the dimension by 2, and the vanishing of S alone reduces the dimension by 4. We have used the notation of I, Tables II and III to describe each type of model. This is prefixed with the letters "CF" (to denote conformally flat comoving slices), and suffixed with the numbers i, ii, and iii to denote the vanishing of I_1, I_2, and both I_1 and I_2, respectively.

In a subsequent article, we shall carefully investigate the members of each subclass in Table I, and relate them to known space–times, in the case where the matter either is dust ($p=0$), or it obeys an equation of state $p=p(\mu)$ with $dp/d\mu\neq 0$.

ACKNOWLEDGMENTS

We are grateful to A. Spero for helpful discussions, and to B. K. Berger, D. M. Eardley, and D. W. Olson for useful correspondence and for informing us of details of their work prior to publication. Above all, we thank J. Wainwright for discussions and constructive criticisms.

APPENDIX A: SPECIALIZATION TO THE CASE WHERE CONDITIONS (i)–(iii) AND (vii) PERTAIN, IN A FERMI-PROPAGATED ORTHONORMAL SHEAR EIGENFRAME

$$\theta_{\alpha\beta}=\operatorname{diag}(\theta_1,\theta_2,\theta_2)$$

Jacobi identities:

$$\partial_1 n_{11}+\partial_2 n_{12}+\partial_3 n_{31}+\partial_2 a_3-\partial_3 a_2-2(n_{11}a_1+n_{12}a_2+n_{31}a_3)=0, \tag{A1}$$

$$\partial_2 n_{22}+\partial_3 n_{23}+\partial_1 n_{12}+\partial_3 a_1-\partial_1 a_3-2(n_{22}a_2+n_{23}a_3+n_{12}a_1)=0, \tag{A2}$$

$$\partial_3 n_{33}+\partial_1 n_{31}+\partial_2 n_{23}+\partial_1 a_2-\partial_2 a_1-2(n_{33}a_3+n_{31}a_1+n_{23}a_2)=0, \tag{A3}$$

$$\partial_0 a_1+\partial_1\theta_2+a_1\theta_1=0, \tag{A4}$$

$$2\partial_0 a_2+\partial_2(\theta_1+\theta_2)+2a_2\theta_2=0, \tag{A5}$$

$$2\partial_0 a_3+\partial_3(\theta_1+\theta_2)+2a_3\theta_2=0, \tag{A6}$$

$$\partial_0 n_{11}-n_{11}(\theta_1-2\theta_2)=0, \tag{A7}$$

$$\partial_0 n_{22}+n_{22}\theta_1=0, \tag{A8}$$

$$\partial_0 n_{33}+n_{33}\theta_1=0, \tag{A9}$$

$$\partial_0 n_{23}+n_{23}\theta_1=0, \tag{A10}$$

$$2\partial_0 n_{31}-\partial_2(\theta_1-\theta_2)+2n_{31}\theta_2=0, \tag{A11}$$

and

$$2\partial_0 n_{12}+\partial_3(\theta_1-\theta_2)+2n_{12}\theta_2=0, \tag{A12}$$

Field equations:

$$\dot\theta_1+2\dot\theta_2+\theta_1^2+2\theta_2^2+\tfrac12(\mu+3p)=0, \tag{A13}$$

$$\partial_1\theta_2+a_1(\theta_1-\theta_2)=0, \tag{A14}$$

$$\partial_2(\theta_1+\theta_2)+(n_{31}-a_2)(\theta_1-\theta_2)=0, \tag{A15}$$

$$\partial_3(\theta_1+\theta_2)-(n_{12}+a_3)(\theta_1-\theta_2)=0, \tag{A16}$$

$$\begin{aligned}-R^{*}_{11}\equiv{}&-\partial_1 a_1+\partial_2 n_{31}-\partial_3 n_{12}+2(n_{12}a_3-n_{31}a_2)\\&-2(n_{11}^2+n_{12}^2+n_{31}^2)+nn_{11}\\&+\left(2a_\gamma a^\gamma+n^{\gamma\delta}n_{\gamma\delta}-\frac{n^2}{2}-\partial_\gamma a^\gamma\right)\\={}&\dot\theta_1+\theta_1(\theta_1+2\theta_2)-\tfrac12(\mu-p),\end{aligned} \tag{A17}$$

$$\begin{aligned}-R^{*}_{22}\equiv{}&-\partial_2 a_2+\partial_3 n_{12}-\partial_1 n_{23}+2(n_{23}a_1-n_{12}a_3)\\&-2(n_{22}^2+n_{23}^2+n_{12}^2)+nn_{22}\\&+\left(2a_\gamma a^\gamma+n^{\gamma\delta}n_{\gamma\delta}-\frac{n^2}{2}-\partial_\gamma a^\gamma\right)\end{aligned} \tag{A18a}$$

$$\begin{aligned}&=\dot\theta_2+\theta_2(\theta_1+2\theta_2)-\tfrac12(\mu-p)\\&=-R^{*}_{33}\\&=-\partial_3 a_3+\partial_1 n_{23}-\partial_2 n_{31}+2(n_{31}a_2-n_{23}a_1)\\&\quad-2(n_{33}^2+n_{31}^2+n_{23}^2)+nn_{33}\\&\quad+\left(2a_\gamma a^\gamma+n^{\gamma\delta}n_{\gamma\delta}-\frac{n^2}{2}-\partial_\gamma a^\gamma\right),\end{aligned} \tag{A18b}$$

$$\begin{aligned}-R^{*}_{23}\equiv{}&\tfrac12(\partial_2 a_3+\partial_3 a_2)+\tfrac12(\partial_3 n_{31}+\partial_1(n_{22}-n_{33})-\partial_2 n_{12})\\&+n_{33}a_1-n_{31}a_3+n_{12}a_2-n_{22}a_1\\&-2(n_{12}n_{31}+n_{22}n_{23}+n_{23}n_{33})+nn_{23}\\={}&0,\end{aligned} \tag{A19}$$

$$\begin{aligned}-R^{*}_{31}\equiv{}&-\tfrac12(\partial_3 a_1+\partial_1 a_3)+\tfrac12(\partial_1 n_{12}+\partial_2(n_{33}-n_{11})-\partial_3 n_{23})\\&+n_{11}a_2-n_{12}a_1+n_{23}a_3-n_{33}a_2\\&-2(n_{23}n_{12}+n_{33}n_{31}+n_{31}n_{11})\\&+nn_{31}=0,\end{aligned} \tag{A20}$$

and

$$\begin{aligned}-R^{*}_{12}\equiv{}&-\tfrac12(\partial_1 a_2+\partial_2 a_1)+\tfrac12(\partial_2 n_{23}+\partial_3(n_{11}-n_{22})-\partial_1 n_{31})\\&+n_{22}a_3-n_{23}a_2+n_{13}a_1-n_{11}a_3\\&-2(n_{31}n_{23}+n_{11}n_{12}+n_{12}n_{22})+nn_{12}=0,\end{aligned} \tag{A21}$$

when $n\equiv n^{\alpha}_{\alpha}$.

It will be convenient to express Eqs. (A17) and (A18) in terms of the trace of $\theta_{\alpha\beta}$ and $R^{*}_{\alpha\beta}$, and their trace free parts, $\sigma_{\alpha\beta}$ and $S_{\alpha\beta}$:

$$\begin{aligned}-R^{*}\equiv-R^{*\alpha}_{\alpha}\equiv{}&-4\partial_\alpha a^\alpha+6a_\alpha a^\alpha+n^{\alpha\beta}n_{\alpha\beta}-\frac{n^2}{2}\\={}&2\theta_2(2\theta_1+\theta_2)-2\mu\end{aligned} \tag{A22}$$

and

$$-(R^{*}_{11}-R^{*}_{22})=(\theta_1-\theta_2)\;+(\theta_1-\theta_2)(\theta_1+2\theta_2). \tag{A23}$$

The contracted Bianchi identities, under which the above equations are compatible, become

$$\dot\mu+(\mu+p)(\theta_1+2\theta_2)=0, \tag{A24}$$

and

$$\partial_\alpha p=0. \tag{A25}$$

APPENDIX 2: COMMUTATION FUNCTIONS, JACOBI IDENTITIES, AND EINSTEIN FIELD EQUATIONS FOR THE PERFECT FLUID SZEKERES METRICS (2.36) IN THE ORTHONORMAL TETRAD $\mathbf{e}_0 = (\partial/\partial t)$, $\mathbf{e}_1 = e^{-A}(\partial/\partial x)$, $\mathbf{e}_2 = e^{-B}(\partial/\partial y)$, and $\mathbf{e}_3 = e^{-B}(\partial/\partial z)$ (with $A = 0$)

Commutation functions:

$$\dot{\mathbf{u}} = \boldsymbol{\omega} = \boldsymbol{\Omega} = 0, \quad \theta_{11} = \dot{A}, \quad \theta_{22} = \theta_{33} = \dot{B},$$

$$\theta_{\alpha\beta} = 0 \quad (\alpha \neq \beta), n_{11} = n_{22} = n_{33} = n_{23} = 0,$$

$$n_{12} = -\tfrac{1}{2}e^{-B}(A_z - B_z), \quad n_{31} = \tfrac{1}{2}e^{-B}(A_y - B_y),$$

$$a_1 = -B_x e^{-A},$$

$$a_2 = -\tfrac{1}{2}e^{-B}(A_y + B_y),$$

and

$$a_3 = -\tfrac{1}{2}e^{-B}(A_z + B_z).$$

Jacobi identities: Equations (A1)–(A12) are identically satisfied.

Field equations:

$$\ddot{A} + 2\ddot{B} + \dot{A}^2 + 2\dot{B}^2 + \tfrac{1}{2}(\mu + 3p) = 0, \tag{B1}$$

$$\dot{B}_x - B_x(\dot{A} - \dot{B}) = 0, \tag{B2}$$

$$\dot{A}_y + \dot{B}_y + A_y(\dot{A} - \dot{B}) = 0, \tag{B3}$$

$$\dot{A}_z + \dot{B}_z + A_z(\dot{A} - \dot{B}) = 0, \tag{B4}$$

$$-R^{\bullet}_{11} = 2e^{-A}(B_x e^{-A})_x + e^{-2B}(A_{yy} + A_{zz} + A_y^2 + A_z^2) + 2B_x^2 e^{-2A} = \ddot{A} + \dot{A}(\dot{A} + 2\dot{B}) - \tfrac{1}{2}(\mu - p), \tag{B5}$$

$$\begin{aligned} -R^{\bullet}_{22} &= e^{-B}[e^{-B}(A_y + B_y)]_y + e^{-B}(B_z e^{-B})_z + 2B_x^2 \\ &\quad \times e^{-2A} + e^{-A}(B_x e^{-A})_x + e^{-2B}(A_y^2 \\ &\quad + A_z B_z + B_y^2 + B_z^2) \\ &= \ddot{B} + \dot{B}(\dot{A} + 2\dot{B}) - \tfrac{1}{2}(\mu - p) \end{aligned} \tag{B6a}$$

$$\begin{aligned} &= -R^{\bullet}_{33} \\ &= e^{-B}[e^{-B}(A_z + B_z)]_z + e^{-B}(B_y e^{-B})_y \\ &\quad + 2B_x^2 e^{-2A} + e^{-A}(B_x e^{-A})_x \\ &\quad + e^{-2B}(A_z^2 + A_y B_y + B_y^2 + B_z^2), \end{aligned} \tag{B6b}$$

$$-R^{\bullet}_{23} = e^{-2B}(A_{yz} - B_y A_z - B_z A_y + A_y A_z) = 0, \tag{B7}$$

$$-R^{\bullet}_{31} = e^{-A-B}(B_{xz} - A_z B_x) = 0, \tag{B8}$$

$$-R^{\bullet}_{12} = e^{-A-B}(B_{xy} - A_y B_x) = 0, \tag{B9}$$

Equations (B1)–(B9) are equivalent to (A13)–(A21).

APPENDIX C: CLASS II SZEKERES METRICS

We use the notation of Szafron.[3] Let L denote the operator

$$L \equiv \phi \frac{\partial^2}{\partial t^2} + \dot{\phi}\frac{\partial}{\partial t} + (\ddot{\phi} + \phi p)I.$$

Then ϕ satisfies

$$L\phi = -k, \tag{C1}$$

and λ satisfies

$$L\lambda = U + kW. \tag{C2}$$

Let μ denote the general solution of

$$L\mu = 0.$$

Then, using (C1), the general solution of (C2) is

$$\lambda = -\frac{\phi}{k}(U + kW) + \mu, \tag{C3}$$

whenever $k \neq 0$. Thus,

$$\alpha = \ln(\lambda + \phi\sigma) \tag{C4a}$$

$$= \ln(\mu + \phi\hat{\sigma}). \tag{C4b}$$

In (C4),

$$\sigma = \frac{2}{1 + k\xi\bar{\xi}}[U\xi\bar{\xi} + V\xi + \overline{V\xi} + W]$$

and

$$\hat{\sigma} = \sigma - \frac{1}{k}(U + kW) = \frac{2}{1 + k\xi\bar{\xi}}[\hat{U}\xi\bar{\xi} + \hat{V}\xi + \overline{\hat{V}\xi} + \hat{W}],$$

where

$$\hat{U} = U - \tfrac{1}{2}(U + kW), \quad \hat{V} = V,$$

and

$$\hat{W} = W - \frac{1}{2k}(U + kW).$$

Thus the term in (C3) that involves ϕ can be absorbed into the term in α involving σ, in such a way that the functions $\hat{U}$ and $\hat{W}$ satisfy $\hat{U} + k\hat{W} = 0$.

[1] C.B. Collins and D.A. Szafron, "A new approach to inhomogeneous cosmologies: intrinsic symmetries, I," J. Math. Phys., see previous paper, this issue.
[2] P. Szekeres, Commun. Math. Phys. **41**, 55 (1975).
[3] D.A. Szafron, J. Math. Phys **18**, 1673 (1977).
[4] H. Bondi, Mon. Not. R. Astr. Soc. **107**, 410 (1947).
[5] R.C. Tolman, Proc. Nat. Acad. Sci. **20**, 169 (1934).
[6] L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, N.J., 1926).
[7] G.F.R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology*, Proceedings of the International School of Physics "Enrico Fermi" Course XLVII, 1969, edited by R.K. Sachs (Academic, London and New York, 1971), p. 104.
[8] J.W. York, Phys. Rev. Lett. **26**, 1656 (1971).
[9] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[10] G.F.R. Ellis and M.A.H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).
[11] B.K. Berger, D.M. Eardley and D.W. Olson, Phys. Rev. D **16**, 3086 (1977).
[12] J. Wainwright, J. Math. Phys. **18**, 672 (1977).
[13] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields* (Pergamon, New York, 1975).
[14] M.A.H. MacCallum, "A Class of Homogeneous Cosmological Models" Ph.D. thesis, Cambridge University (1969).

A new approach to inhomogeneous cosmologies: Intrinsic symmetries. III. Conformally flat slices and their analysis

C. B. Collins[a)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
D. A. Szafron[b)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1 and Department of Mathematics, Portland State University, P. O. Box 751, Portland, Oregon 97207 [c)]

(Received 13 October 1978)

We have already obtained an invariant characterization and classification of the Szekeres inhomogeneous cosmological models. In the present paper, we present a detailed study of the hierarchy of solutions. The approach adopted suggests simple generalizations of the Szekeres models, and we present a number of results and ideas for further study.

1. CONFORMALLY FLAT SLICES WITH AN EQUATION OF STATE

We continue our investigation (see Refs. 1 and 2, hereafter referred to as I and II, respectively) of space–times in which the fluid flows orthogonally to conformally flat spacelike hypersurfaces. We shall be discussing models subject to various conditions. These were listed in I and II, and, for brevity, are not repeated here. We shall investigate exhaustively the contents of each subcase in the classification of II, Table I, when the matter obeys an equation of state of the form $p = p(\mu)$; the discussion naturally divides into two possibilities, depending on whether or not $dp/d\mu \equiv 0$. If $dp/d\mu \equiv 0$, p is constant, and then we consider the only physically realistic case of dust ($p \equiv 0$). We shall also obtain necessary and sufficient conditions on the variables A and B in the Szekeres line element (II,2.26) for our models to be any of the familiar ones (e.g., spatially homogeneous, LRS, Taub, Bondi–Tolman, flat slices, $R^*_{\alpha\beta}$ isotropic, etc.). We shall assume $\sigma \not\equiv 0$ throughout, i.e., we ignore the isotropic (FRW) solutions.

A direct calculation (cf. II, Appendix B) shows that

$$I_1 = a_1 = -B_x e^{-A} \tag{1.1}$$

and

$$I_2 = (n_{13} - a_2)^2 + (n_{12} + a_3)^2 = e^{-2B}(A_y^2 + A_z^2). \tag{1.2}$$

Einstein's field equations impose restrictions on the metric functions A and B [see (II, B1)–(II, B9)]. With these restrictions in force, we now give conditions for the Szekeres metric to satisfy various specializations. The space–times satisfy condition (v) of I and II, i.e., the hypersurfaces $\{t = \text{constant}\}$ have isotropic $R^*_{\alpha\beta}$, if and only if $R^*_{11} = \frac{1}{2}(R^*_{22} + R^*_{33}) \Longleftrightarrow$

$$2e^{-2A}(B_{xx} - A_x B_x) + e^{-2B}(A_{yy} + A_{zz} + A_y^2 + A_z^2 - 2(B_{yy} + B_{zz})) = 0, \tag{1.3}$$

while

$$R^* = 0 \Longleftrightarrow R^*_{11} + R^*_{22} + R^*_{33} = 0$$
$$\Longleftrightarrow e^{-2A}(3B_x^2 - 2A_x B_x + 2B_{xx}) + e^{-2B} \times (A_{yy} + A_{zz} + A_y^2 + A_z^2 + B_{yy} + B_{zz}) = 0. \tag{1.4}$$

Moreover, the hypersurfaces $\{t = \text{constant}\}$ are *flat* if and only if both (1.3) and (1.4) hold, i.e., if and only if

$$B_x^2 e^{-2A} + e^{-2B}(B_{yy} + B_{zz}) = 0 \tag{1.5a}$$

and

$$2e^{-2A}(B_x^2 - A_x B_x + B_{xx}) + e^{-2B}(A_{yy} + A_{zz} + A_y^2 + A_z^2) = 0. \tag{1.5b}$$

The Szekeres space–times are *locally rotationally symmetric* if and only if there are no preferred directions in the 2-plane spanned by $\mathbf{e}_2$ and $\mathbf{e}_3$. This is so if and only if $I_2 = 0 \Longleftrightarrow$

$$A = A(t,x) \tag{1.6}$$

(and since $n_{11} = \omega_1 = 0$, the LRS models necessarily belong to case II of Refs. 3 and 4). Of the inhomogeneous LRS cosmological models, two are historically important: the Bondi–Tolman spherically symmetric models and the Taub plane-symmetric models. In the spherically symmetric models, the surfaces $\{t = \text{constant}, x = \text{constant}\}$ are 2-spheres, while in the plane-symmetric models they are 2-planes. Thus the Szekeres models are *spherically symmetric* if and only if (1.6) holds and

$$e^{-2B}(B_{yy} + B_{zz}) = -r^2(t,x), \quad (r \not\equiv 0), \tag{1.7a}$$

and they belong to the *Bondi–Tolman* class if and only if, *in addition*,

$$B_x \not\equiv 0 \quad \text{and} \quad p = 0. \tag{1.7b}$$

The models are of the *Taub plane-symmetric* class if and only if (1.6) holds and

$$B_{yy} + B_{zz} = 0. \tag{1.8}$$

Finally, the Szekeres metrics are *spatially homogeneous* if and only if

$$\dot{A} = \dot{A}(t) \quad \text{and} \quad \dot{B} = \dot{B}(t). \tag{1.9}$$

[a)] Supported in part by Operating Grant #A3978 from the National Research Council of Canada, and by a University of Waterloo Research Grant.
[b)] Supported by a Postgraduate Fellowship from the National Research council of Canada.
[c)] Permanent address: Department of Mathematics, Portland State University, P.O. Box 751, Portland, Oregon, 97207.

This is seen by employing Eqs. (II, B2)–(II, B4) to show that without loss of generality, $A = A(t)$ and $B = \psi(t) + \nu(y,z)$ (cf. Refs. 5 and 6) where the metric $e^{2\nu(y,z)}(dy^2 + dz^2)$ is that of a 2-space of constant curvature. The line element (II, 2.36) then clearly admits four independent spacelike Killing vectors.

We investigate each subcase occurring in the specialization diagram (II, Table I), determining the associated restrictions on the solutions as given by Szekeres[5] for dust and Szafron[6] for perfect fluid. This is done using Szafron's notation. We prove in II, Appendix A that in the Class II models (defined by the condition $B_x = 0$), without loss of generality, $k\neq 0 \Rightarrow U + kW = 0$, and we take this to be the case henceforth (in the particular form of the dust solutions as given by Szekeres),[5] U and W are necessarily so related.[7] We take this opportunity to point our further typographical errors on p. 61 of Ref. 5 (cf. Refs. 7 and 8): If $k = 0$ and $\phi = t^{2/3}$, then $\mu = \frac{9}{10}u(r)t^{4/3}$, while if $k = 0$ and $\phi = 1$, then $\mu = \frac{1}{2}u(r)t^2$.

1. Type VIIIa models ($S = R^* = 0$, and specializations): In these models, the hypersurfaces $\{t = \text{constant}\}$ are flat. With an equation of state $p = p(\mu)$, with $dp/d\mu\neq 0$, it follows from Proposition 2.2 of Sec. 2 that the models are spatially homogeneous of Bianchi type I, so $I_1 = I_2 = 0$, and the models are necessarily of type VIIIaiii (they admit a G_4 isometry group, multiply transitive on the hypersurfaces $\{t = \text{constant}\}$, i.e., they are LRS, belonging to case IIa of Ref. 4). If the pressure is zero, then there are models in each subclass of type CFVIII. We now discuss each subclass in turn, when the matter content is dust.

(a) Type CFVIIIaiii ($S = R^* = I_1 = I_2 = 0$): The Szekeres metrics are of this type if and only if $A = A(t,x)$, $B_x = 0$, and $B_{yy} + B_{zz} = 0$, as follows from (1.1), (1.5), and (1.6). Since (1.8) then holds, these models belong to the Taub plane-symmetric subclass of the LRS models. The 2-surfaces $\{t = \text{constant}, x = \text{constant}\}$ are therefore flat. This means that the (dust) models belong to case IIbiii of Ref. 3, in which, without loss of generality, the functions A and B are given by

$$e^A = t^{-1/3}[t + C(x)]$$

and

$$e^B = t^{2/3}.$$

This is the generalized Bianchi I metric discussed in I, Sec. 1 (the expression for e^A corrects that given for X by Ellis[3]; it also corrects the expression for X given by Szafron and Wainwright[9]). The space–time is in general invariant under an Abelian G_3 acting multiply transitive on the 2-surfaces $\{t = \text{constant}, x = \text{constant}\}$; in the special case where $C(x)$ is constant, the space–time is invariant under a G_4 multiply transitive in the hypersurfaces $\{t = \text{constant}\}$, i.e., a spatially homogeneous model of Bianchi type I. In the notation of Szafron,[6] all these models are of Class II with $k = U(z) = V(z) = 0$ (without loss of generality, $\sigma = 0$). The models are spatially homogeneous of Bianchi I if and only if in addition $[\ln(\lambda + \phi\sigma)]' = 0$.

(b) Type CFVIIIaii ($S = R^* = I_2 = 0$): The Szekeres metrics are of this type if and only if (1.5) and (1.6) hold, and $B_x \neq 0$. The only models are a subset of those given by Ellis,[3] of case IIaii, in which the hypersurfaces $\{t = \text{constant}\}$ are constrained to be flat. With Ellis's tetrad choice ($n_{12} = a_3 = 0$) this requires $\partial_1 a_1 = a_1^2 = 2\partial_2 a_2 - 4a_2^2$, as follows from (II, A17) and (II, A18a). Consequently $a_0^2(x) = k > 0$ in (4.32a) of Ref. 3 [cf. his Eq. (4.19)]. These models are particular examples of spherically symmetric Bondi–Tolman dust solutions (cf. Ref. 5) and they are the ones described as being "marginally bound" by Berger, Eardley, and Olson.[10] In the notation of Szafron,[6] these models are of Class I with $k(z) = 0$ and $A(z)$, $B(z)$, and $C(z)$ all constant. Equations (1.7) are satisfies with $r \equiv \phi^{-1}$ (cf. Ref. 10).

(c) Type CFVIIIai ($S = R^* = I_1 = 0$): The Szekeres metrics are of this type if and only if $B_x = 0$, $B_{yy} + B_{zz} = 0$, and $A_{yy} + A_{zz} = -(A_y^2 + A_z^2) \neq 0$. The solutions are of Class II with $k = U(z) = 0$ in Szafron's notation.[6]

(d) Type CFVIIIa ($S = R^* = 0$): The necessary and sufficient condition for the metric (II, 2.36) to belong to this subclass is that Eq. (1.5) hold, with $B_x(A_y^2 + A_z^2) \neq 0$. This subclass consists of all remaining Class I (dust) Szekeres metrics which have flat slices and yet are not LRS. In Szafron's notation,[6] the models are of Class I with $k(z) = 0$, and with at least one of $A'(z)$, $B'(z)$, and $C'(z)$ nonzero.

The above discussion of the type VIII models exhausts all our models in which $R^*_{\alpha\beta}$ is isotropic; this is because, in that case, by the Corollary to II, Lemma 2.1, all anisotropic models have $R^*_{\alpha\beta} = 0$. We now turn our attention to those models for which R^*_{ij} is anisotropic, considering first the subset which is LRS.

(2) Type IIIcii models ($I_2 = 0$ and specializations): With an equation of state $p = p(\mu)$, where $dp/d\mu \neq 0$, these models are of case II of Ref. 4, being of case IIa if $I_1 = 0$ (type CFIIIciii) and of case IIb if $I_1 \neq 0$ (type CFIIIcii). However, such case IIb models are necessarily isotropic.[4] Moreover, the case IIa models are spatially homogeneous, and admit a G_4 acting multiply transitively on the hypersurfaces $\{t = \text{constant}\}$; if $\partial_2 a_2 - a_2^2 < 0$, there is a G_3 subgroup of Bianchi type III (case II of Kantowski and Sachs[11]) acting simply transitively on the hypersurfaces $\{t = \text{constant}\}$ [using Eq. (4.19) of Ref. 3], whereas if $\partial_2 a_2 - a_2^2 > 0$ there is no such subgroup, but there is a G_3 subgroup acting multiply transitively on the 2-surfaces $\{t = \text{constant}, x = \text{constant}\}$, and the space–time geometry is that of case I of Kantowski and Sachs.[11] (If $\partial_2 a_2 - a_2^2 = 0$, the space–time is of type CFVIIIaiii.) We now discuss each subclass of type CFIIIcii and CFIIIciii models, when the matter content is dust.

(a) Type CFIIIciii ($I_1 = I_2 = 0$): The Szekeres metrics belong to this subclass if and only if $A = A(t,x)$ and $B_x = 0$ [and $R^* \neq 0$, i.e., (1.4) false]. The (dust) models consist of case IIbi of Ref. 3, and in general admit a multiply transitive G_3 acting on the 2-surfaces $\{t = \text{constant}, x = \text{constant}\}$. If $A = A(t)$, the space–times admit a group, G_4, multiply transitive on the hypersurfaces $\{t = \text{constant}\}$, and belonging to case I of Kantowski and Sachs[11] if $\partial_2 a_2 - a_2^2 > 0$, and to Bianchi type III (case II of Kantowski and Sachs[11]) if $\partial_2 a_2 - a_2^2 < 0$. In the notation of Szafron,[6] all these models are of Class

II with $k \neq 0$ and $U(z) = V(z) = W(z) = 0$ (cf. II, Appendix A). The (generalized) Bianchi III metrics have $k = -1$, and then (1.7a) is satisfied with $r \equiv \phi^{-1}$; the (generalized) Kantowski–Sachs metrics have $k = +1$ (cf. Refs. 5 and 12). The models are spatially homogeneous if and only if in addition $[\ln(\lambda + \phi\sigma)]' = 0$.

(b) Type CFIIIcii ($I_2 = 0$): These models are the remaining LRS models with conformally flat (but not flat) sections. The Szekeres metrics are of this subclass if and only if $A = A(t,x)$ (and $B_x R^* \neq 0$). The solutions are the remaining (dust) models in case IIaii of Ref. 3, excluding the marginally-bound case where $a_0^2(x) = K$ (which is of type CFVIIIaii). These metrics are "generalized Bondi–Tolman" solutions, admitting a multiply transitive G_3 on the 2 surfaces $\{t = \text{constant}, x = \text{constant}\}$. When the (constant) curvature of these 2-surfaces is positive, the metrics are spherically symmetric (Bondi–Tolman), and when the curvature is zero the metrics are plane symmetric (Taub). (In the case when the curvature is negative, Szekeres[5] describes the solutions as "pseudospherically symmetric," and remarks that this class is previously unexplored. However, Ellis[3] has investigated it fairly extensively.) In the notation of Szafron,[6] the metrics are of Class I with $A(z)$, $B(z)$, and $C(z)$ linearly dependent functions and $k(z) \not\equiv 0$. The models are Bondi–Tolman spherically symmetric if $\partial_2 a_2 - a_2^2 > 0$ [$\Longleftrightarrow k(z) > -1$; cf. Ref. 5], and then Eqs. (1.7) are satisfied with $r \equiv [1 + k(z)]^{1/2}\phi^{-1}$. The models are Taub plane symmetric if $\partial_2 a_2 - a_2^2 = 0$ [$\Longleftrightarrow k(z) = -1$; Szekeres[5] ascribes this type of solution to Eardley, Liang, and Sachs,[13] although it occurs as a special case of a general class of solutions studied earlier by, e.g., Taub[14,15] and Ellis[3]].

3. Type IIIc and IIIci models (general case and $I_1 = 0$): The remaining models are those which are not LRS and for which $R^{*}_{\alpha\beta}$ is anisotropic. We show first that the only possible models are those with dust matter content (cf. the argument in Ref. 12). If $p = p(\mu)$ and $dp/d\mu \neq 0$, then Eq. (II, A25) shows that $p = p(t)$ and $\mu = \mu(t)$. By the energy conservation Eq. (II, A24), $\theta_1 + 2\theta_2$ depends on t only. Now Eq. (II,A13) requires that $\theta_1^2 + 2\theta_2^2$ depends on t only, hence $\theta_1 = \theta_1(t)$ and $\theta_2 = \theta_2(t)$. Using Szafron's notation,[6] if the metric is of Class I, this requires $\dot\phi/\phi$ and $(\dot\phi' + \dot\phi v')/(\phi' + \phi v')$ to depend on time only, hence $\phi(t,z)$ is separable and then the model is FRW (which we assume not to be the case). If the metric is of Class II, we find that $(\dot\lambda + \dot\phi\sigma)/(\lambda + \phi\sigma)$ depends on time only, so, differentiating with respect to ξ and $\bar\xi$, we obtain $(\dot\lambda\phi - \lambda\dot\phi)\sigma_\xi = (\dot\lambda\phi - \lambda\dot\phi)\sigma_{\bar\xi} = 0$. Since $\dot\lambda\phi \neq \lambda\dot\phi$ (otherwise the metric is FRW), it follows that $\sigma_\xi = \sigma_{\bar\xi} = 0$, and so Eq. (1.6) is satisfied and the space–times are LRS, which is again a contradiction. We now discuss each subclass of type CFIIIc and CFIIIci models, when the matter content is dust.

(a) Type CFIIIci ($I_1 = 0$): The Szekeres metrics are of

Table I. Contents of each subcase in II. Table I, when the matter either is dust or it obeys an equation of state $p = p(\mu)$, with $dp/d\mu \neq 0$. We assume condition (viiia) of I and II holds.

Type	Dimension	$R^* = 0$	$S = 0$	$I_1 = 0$	$I_2 = 0$	Flat slices	$R^{*}_{\alpha\beta}$ isotropic	LRS	Contains Tolman–Bondi spherically symmetric models	Contains Taub plane-symmetric models	Contains spatially homogeneous models	Contains others	Solutions in notation of Szafron[a]
CFVIIaiii	2	√	√	√	√	√	√	√	×	Bianchi I + Generalized Bianchi I (for $p = 0$ only)	Bianchi I	×	Class II, $k = U(z) = V(z) = 0$ (+ $[\ln(\lambda + \phi\sigma)]' = 0$ for Bianchi I)
CFVIIIaii	3	√	√	×	√	√	√	√	Marginally bound (for $p = 0$ only)	×	×	×	Class I, $k(z) = 0$, $A(z)$, $B(z)$ and $C(z)$ constants
CFVIIIai	4	√	√	√	×	√	√	×	×	×	×	√ (for $p = 0$ only)	Class II, $k = U(z) = 0$, $V(z) \neq 0$
CFVIIIa	5	√	√	×	×	√	√	×	×	×	×	√ (for $p = 0$ only)	Class I, $k(z) = 0$. At least one of $A'(z)$, $B'(z)$ $C'(z)$ nonzero
CFIIIciii	7	×	×	√	√	×	×	√	×	×	Kantowski–Sachs ($k = +1$) and Bianchi III ($k = -1$)	Generalized Kantowski–Sachs ($k = +1$), and Generalized Bianchi III ($k = -1$) (for $p = 0$ only)	Class II, $k \neq 0$, $U(z) = V(z) = W(z) = 0$ (+ $[\ln(\lambda + \phi\sigma)]' = 0$ for Kantowski–Sachs ($k = +1$) and Bianchi III ($k = -1$))
CFIIIcii	8	×	×	×	√	×	×	√	Not marginally bound $k(z) > -1$ (for $p = 0$ only)	$k(z) = -1$ (for $p = 0$ only)	×	$k(z) < -1$ Generalized Bondi–Tolman (for $p = 0$ only)	Class I, $k(z) \neq 0$ $A(z)$, $B(z)$, $C(z)$ linearly dependent
CFIIIci	9	×	×	√	×	×	×	×	×	×	×	Remaining class II solutions (for $p = 0$ only)	Class II $[k^2 + U^2(z)][U^2(z) + V(z)\bar V(z) + W^2(z)] \neq 0$
CFIIIc	10	×	×	×	×	×	×	×	×	×	×	Remaining Class I Solutions (for $p = 0$ only)	Class I, $k(z) \neq 0$ $A(z)$, $B(z)$, $C(z)$ linearly independent

this type if and only if $B_x = 0$ [and $(A_y^2 + A_z^2)R^* \neq 0$]. This means that, in the notation of Szafron,[6] the metrics are of Class II with $[k^2 + U^2(z)][U^2 + V\bar{V} + W^2] \neq 0$. The solutions were discovered by Szekeres.[5]

(b) Type CFIIIc: This is the most general class, in which $B_x(A_y^2 + A_z^2)R^* \neq 0$. In the notation of Szafron,[6] the metrics are of Class I with $k(z) \not\equiv 0$ and with $A(z)$, $B(z)$, and $C(z)$ linearly independent. These solutions were discovered by Szekeres.[5]

Table I summarizes the results of the foregoing analysis.

Clearly, some of these results have been previously obtained by a variety of authors, and we have endeavored to acknowledge this in the text. We should point out that as a consequence of our scrutiny of the Szekeres models, we have rediscovered a result of Wainwright,[16] that the Szekeres models are necessarily LRS if the matter obeys an equation of state $p = p(\mu)$ with $dp/d\mu \neq 0$. This result was also duplicated, and extended, by Spero and Szafron,[12] who found that such models are not only LRS, but also spatially homogeneous, being FRW, Kantowski–Sachs, or Bianchi types I or III. Again this is clear from our table.

Finally, we make some remarks concerning the two solutions of Stephani.[17] These belong to the Szekeres class, and were obtained by Stephani in his search for Petrov type D models of imbedding class I. In general, $I_1 = 0$ and $I_2SR^* \neq 0$. In Stephani's first solution, the matter content cannot be dust; if it satisfies an equation of state of the form $p = p(\mu)$, then $I_1 = I_2 = S = R^* = 0$, i.e., the metric belongs to type CFVIIIaiii and is spatially homogeneous of Bianchi type I [it may also be a FRW($k = 0$) model]. In Stephani's second set of solutions, the only solutions in which $p = p(\mu)$ and $dp/d\mu \neq 0$ are LRS ($I_2 = 0$), and then $SR^* \neq 0$, i.e., they are of Bianchi type III or Kantowski–Sachs [or they may also be FRW($k = \pm 1$) models]. We have found that in the dust solutions, the function $d_1(x)$ given by Stephani[17] is not arbitrary (as he has claimed) but that $d_1(x) \propto \cosh x$. By redefining his variable r, we can arrange for $d_2(r) = 1$, without loss of generality. The models are then LRS if and only if $d_1(x) = 0$. Thus the dust solutions in Stephani's second set are in general of type CFIIIci, but they can also be of type CFIIIciii. Relating these solutions to those obtained by Szekeres,[5] we see that $f^2 = t^2/b$ requires $b > 0$, so $k = -1$, and $\dot\phi/\phi = 1/t$ requires $\phi \propto t$. By Eq. (8) of Ref. 5, $\dot\phi^2 = 1 + l/\phi$ requires $\phi = t$ and $l = 0$. In fact, the Stephani dust solutions form only a small subset of those Class II Szekeres solutions for which $k = -1$ and $l = 0$, since, in the notation of Szafron,[6] the Stephani dust solutions are restricted by the requirement that $V = 0$.

2. FLAT SLICES AND ACCELERATING LRS METRICS

In this section, we obtain a number of results pertaining to space–times with flat comoving slices. We also examine the LRS models in more detail.

Theorem 2.1: Let M be a space–time with a C^3 metric. Suppose that M is foliated by a family of flat spacelike hypersurfaces. Then there exists locally a set of coordinates in which the metric has the form

$$ds^2 = -A^2dt^2 + 2A_\alpha dx^\alpha dt + \delta_{\alpha\beta}dx^\alpha dx^\beta, \tag{2.1}$$

where $A = A(t,x^\alpha)$ and $A_\alpha = A_\alpha(t,x^\beta)$. The flat spacelike hypersurfaces are then specified by $\{t = \text{constant}\}$.

Proof: We may choose comoving coordinates $\{y^a\}$ so that the line element has the form

$$ds^2 = -B^2(dy^0)^2 + g_{\alpha\beta}\,dy^\alpha dy^\beta, \tag{2.2}$$

with $B = B(y^a)$ and $g_{\alpha\beta} = g_{\alpha\beta}(y^a)$. The flat hypersurfaces are given by $\{y^0 = \text{constant}\}$. Consider the curve, Γ, defined by $y^\beta = 0$. For each point p on this curve, let $\mathbf{e}_\alpha(p)$ denote those elements of the tangent space of the hypersurfaces $\{y^0 = \text{constant}\}$ through p, where $\mathbf{e}_\alpha(p)\cdot\mathbf{e}_\beta(p) = \delta_{\alpha\beta}$, i.e., at each point $p \in \Gamma$, the vectors $\mathbf{e}_\alpha(p)$ are orthonormal. We first construct vector fields $\{\mathbf{Y}_\alpha\}$ which are parallel on each flat hypersurface $\{y^0 = \text{constant}\}$, and which agree with the vector fields $\{\mathbf{e}_\alpha\}$ along the curve Γ [i.e., $\mathbf{Y}_\alpha(p) = \mathbf{e}_\alpha(p)$ for all $p \in \Gamma$]. These vector fields $\{\mathbf{Y}_\alpha\}$ are determined by parallel propagation of the fields $\{\mathbf{e}_\alpha\}$ in each submanifold $\{y^0 = \text{constant}\}$, and therefore satisfy the equations

$$\frac{\partial Y_\alpha{}^\beta}{\partial y^\lambda} = -\Gamma^\beta_{\lambda\mu}Y_\alpha{}^\mu, \quad Y_\alpha{}^\beta(p) = e_\alpha{}^\beta(p), \quad Y_\alpha{}^0 = 0. \tag{2.3}$$

For each $\mathbf{Y}_\alpha$, (2.3) is a system of differential equations with independent variables y^λ, and parametrized by y^0. The integrability conditions of this system are clearly satisfied, since the hypersurfaces $\{y^0 = \text{constant}\}$ are flat [i.e., the integrability conditions are $R^{*\beta}_{\ \alpha\lambda\mu}Y_\gamma{}^\lambda = 0$, where $R^{*\beta}_{\ \alpha\lambda\mu}$ is the (zero) Riemann tensor of the hypersurfaces $\{y^0 = \text{constant}\}$; cf. Appendix II of Ref. 18]. Furthermore, since the quantities $\Gamma^\beta_{\lambda\mu}$ are C^2 in the coordinates y^a, so also are the components $Y_\alpha{}^\beta$, as follows from the proofs of the theorem in Appendix II of Ref. 18. Note also that since parallel propagation preserves scalar products, $\mathbf{Y}_\alpha\cdot\mathbf{Y}_\beta = \delta_{\alpha\beta}$ everywhere.

We now seek a new coordinate system (t,x^α), where $t = y^0$ and $x^\alpha = x^\alpha(y^a)$, with respect to which $Y_\alpha = \partial/\partial x^\alpha$ in some open region of space–time through which Γ passes. This will be possible if and only if we can determine suitable solutions of the equations

$$\delta^\beta_\alpha = Y_\alpha{}^\lambda\frac{\partial x^\beta}{\partial y^\lambda}. \tag{2.4}$$

Since the components $Y_\alpha{}^\beta$ are locally C^2 functions of the coordinates y^a, and since the vectors $\mathbf{Y}_\alpha$ are linearly independent on Γ, it follows that the inverse, $Y^\beta{}_\gamma$, of $Y_\alpha{}^\lambda$ exists and is a C^2 function of the coordinates y^a in some open region through which Γ passes. Equations (2.4) are therefore equivalent to the equations

$$\frac{\partial x^\beta}{\partial y^\lambda} = Y^\beta{}_\lambda\,. \tag{2.5}$$

By Eqs. (2.3) it follows that the inverse, $Y^\beta{}_\gamma$, satisfies

$$\frac{\partial Y^\beta{}_\gamma}{\partial y^\lambda} = Y^\beta{}_\gamma\Gamma^\alpha_{\lambda\gamma}\,,$$

from which the integrability conditions for (2.5) are seen to hold (i.e., the integrability conditions are $Y^{\beta}{}_{\alpha}\,\Gamma^{\alpha}_{\lambda\gamma} = Y^{\beta}{}_{\alpha}\Gamma^{\alpha}_{\gamma\lambda}$, which are true since the connection is symmetric). Thus Eqs. (2.5), and hence Eqs. (2.4), possess a solution $x^{\alpha} = x^{\alpha}(y^{a})$ which is C^2 in the coordinates y^{a}. Moreover, with this choice of coordinates (t,x^{α}) the line element (2.2) becomes (2.1), where $A_{\alpha} = g_{\beta\gamma}(\partial y^{\beta}/\partial t)Y_{\alpha}{}^{\gamma}$ and $A^2 = B^2 - A_{\alpha}A_{\beta}\delta^{\alpha\beta}$. □

Corollary 1: Let M be a space–time with a C^3 metric. Suppose that M is foliated by a family, $\mathscr{F}$, of flat spacelike hypersurfaces. Then there exists locally an orthonormal tetrad for which $\mathbf{e}_0$ is orthogonal to the family $\mathscr{F}$, and in which the commutation function $\gamma^{\alpha}_{\beta\gamma} = 0$ ($\Longleftrightarrow n_{\alpha\beta} = 0$ and $a_{\alpha} = 0$).

Proof: Choose the coordinates of Theorem 2.1 with

$$\mathbf{e}_0 = B^{-1}\left(\frac{\partial}{\partial t} - \delta^{\alpha\beta}A_{\beta}\frac{\partial}{\partial x^{\alpha}}\right) \quad \text{and} \quad \mathbf{e}_{\alpha} = \frac{\partial}{\partial x^{\alpha}},$$

where $B^2 = A^2 + A_{\alpha}A_{\beta}\delta^{\alpha\beta} \neq 0$. □

Remark: We shall refer to the tetrad of Corollary 1 as the "*Cartesian tetrad*."

Corollary 2: Let M be a space–time with a C^3 metric. Suppose that M is foliated by a family, $\mathscr{F}$, of flat spacelike hypersurfaces. Suppose further that the normal congruence of $\mathscr{F}$ is geodesic, and that Cartesian tetrad of Corollary 1 is Fermi propagated. Then there exists a function $\phi\,(t,x^{\alpha})$ such that $A_{\beta} = \phi_{,\beta}$ in (2.1), and the line element is expressible in the form

$$ds^2 = -(1 - \phi_{,\alpha}\phi_{,\beta}\delta^{\alpha\beta})dt^2 + 2\phi_{,\alpha}dx^{\alpha}\,dt + \delta_{\alpha\beta}\,dx^{\alpha}\,dx^{\beta}. \tag{2.6}$$

Proof: The condition that the normal congruence of $\mathscr{F}$ is geodesic means that $B = B\,(y^0)$ in Eq. (2.2). By a transformation of the form $y^{0'} = y^{0'}(y^0)$ we can then arrange for $A = 1$, without loss of generality. From the tetrad of Corollary 1 and (I,2.4), we find that

$$\gamma_{\mu 0 \nu} = \frac{\partial A_{\mu}}{\partial x^{\nu}} = -\epsilon_{\mu\nu\sigma}\Omega^{\sigma} - \theta_{\mu\nu}.$$

Consequently $\mathbf{\Omega} = \mathbf{0}$ implies that $A_{[\mu,\nu]} = 0$, where a comma denotes partial derivative. Thus there exists locally a function $\phi\,(t,x^{\alpha})$ such that $A_{\mu} = \phi_{,\mu}$, and the result follows from Theorem 2.1.

Remarks: the coordinates (t,x^{α}) obtained in Theorem 2.1 and Corollary 2 are not in general comoving with the normal congruence of the family $\mathscr{F}$ [in fact they are comoving in Theorem 2.1 if and only if $\mathbf{e}_0(x^{\alpha}) = 0 \Longleftrightarrow A_{\alpha} = 0$, and they are comoving in Corollary 2 if and only if $\phi = \phi\,(t)$, which generates Minkowski space–time]. However, the transformation between these coordinates and comoving ones can be written down explicitly in some special cases. The spatially homogeneous perfect fluid Bianchi type I solutions have a metric of the form

$$ds^2 = -d\hat{t}^2 + X^2(\hat{t})d\hat{x}^2 + Y^2(\hat{t})d\hat{y}^2 + Z^2(\hat{t})d\hat{z}^2$$

[cf. I, Eq. (1.1)], and satisfy the conditions of Corollary 2.[19] If we define functions t, x, y, and z by

$$t = \hat{t}, \quad x = \hat{x}X(\hat{t}), \quad y = \hat{y}Y(\hat{t}), \quad \text{and} \quad z = \hat{z}Z(\hat{t}),$$

and write

$$\phi = -\tfrac{1}{2}[\hat{x}^2X(\hat{t})\dot{X}(\hat{t}) + \hat{y}^2Y(\hat{t})\dot{Y}(\hat{t}) + \hat{z}^2Z(\hat{t})\dot{Z}(\hat{t})],$$

where $\dot{} \equiv d/d\hat{t}$, then the metric has the form of Eq. (2.6) in the coordinates (t,x,y,z).

We shall now show that for certain cosmological models, spatial homogeneity of the fluid's energy density μ requires the model to be spatially homogeneous in the conventional sense of Lie group theory.

Proposition 2.2: Suppose that a cosmological model satisfies conditions (i), (ii), (iii), and (vi) of I and II, that the fluid's expansion tensor possesses two equal eigenvalues ($\tau = 0$), and that the energy density μ is constant on each of the flat hypersurfaces $\mathscr{S}$. Then necessarily the model is spatially homogeneous, and admits a group of Bianchi type I.

Proof: We choose the Fermi-propagated shear eigenframe of I, Proposition 4.3. The contracted Bianchi identities (II, A24) and (II,A25) yield

$$\partial_{\alpha}\theta_1 = \partial_{\alpha}\theta_2 = 0,$$

for all α. Field Eqs. II, (A14) – (A16) now require that either $\theta_1 = \theta_2$, and then the model is FRW[20] or

$$a_1 = n_{31} - a_2 = n_{12} + a_3 = 0.$$

If the model is FRW, then it is spatially homogeneous, and it admits a group of Bianchi type I since $\overset{*}{R}_{\alpha\beta} = 0$.[19] Assume then that the model is not FRW. The definition of $\overset{*}{R}_{11}$, Eq. (II,A17), becomes

$$\tfrac{1}{2}(n_{22}^2 + n_{33}^2 - n_{11}^2) + 2n_{23}^2 - n_{22}n_{33} = 0. \tag{2.7}$$

Differentiation of (2.7) along $\mathbf{e}_0$, using Eqs. (II,A7)–(II,A10), and substituting (2.7) yields

$$n_{11} = 0, \tag{2.8}$$

and then the model is LRS.[3,4] Substitution of Eq. (2.8) into (2.7) implies that

$$n_{22} - n_{33} = n_{23} = 0. \tag{2.9}$$

Equations (2.8) and (2.9) now allow us to choose coordinates (as in the proof of II, Theorem 2.4) so that the line element is of the Szekeres form [II, Eq. (1.1)]. Now from Theorem 2.3 of Ref. 12, or alternatively from the results of Ellis, MacCallum, and Stewart,[3,4,19] the model is spatially homogeneous and admits a group of Bianchi type I, since $\overset{*}{R}_{\alpha\beta} = 0$.

Remark: Instead of requiring $\mu = \mu(t)$ we could alternatively demand an equation of state [i.e., condition (viiib) of I and II] with $dp/d\mu \neq 0$, and appeal to Eq. (II,A25) to deduce that μ satisfies the conditions of the proposition. We note also that the proof of Proposition 2.2 does not require that the cosmological term be zero. We therefore obtain

Corollary: Proposition 4.2 is valid when condition (ii) of I and II is relaxed.

In general the Cartesian tetrad of Corollary 1 of Theorem 2.1, in which $n^{\alpha\beta} = a^{\lambda} = 0$, is not a shear eigenframe. We now determine circumstances in which the two frames agree.

Theorem 2.3: Suppose that a cosmological model satisfies conditions (i), (iii), and (viiia) of I and II. Suppose further that there is a shear eigenframe with respect to which $n^{\alpha\beta}\equiv a^{\lambda}\equiv 0$. Then the model is either spatially homogeneous (and admits a group of Bianchi type I), or it is a spatially inhomogeneous LRS model belonging to case IIa of Ref. 4 if $p\neq 0$, and to case IIbiii of Ref. 3 if $p = 0$. With the additional condition (viiib) of I and II, the only inhomogeneous model is the "generalized Bianchi I" (case IIbiii) dust solution of Ellis[3] (or an unrealistic generalization in which the pressure is a nonzero constant).

Proof: The field Eqs. (I,B3) in a shear eigenframe with $R^{\;*}_{\alpha\beta} = 0$ imply that if θ_1, θ_2, and θ_3 are distinct,

$$\Omega^{\alpha} = 0 \tag{2.10}$$

[if θ_1, θ_2, and θ_3 are not distinct, then a shear eigenframe can be found in which (2.10) is true]. The Jacobi identities (I,B5) and (I,B6) then imply that

$$\partial_\alpha \theta_\beta = 0 \quad (\alpha\neq\beta). \tag{2.11}$$

The commutators (I,A1)–(I,A3) are given by

$$[\mathbf{e}_0,\mathbf{e}_\alpha] = -\theta_\alpha \mathbf{e}_\alpha \quad (\text{no sum}).$$

Applying these commutators to $\theta_\beta - \theta_\lambda$ (α,β,λ all distinct) and using Eqs. (I,B3) and (2.11) yields

$$(\partial_1\theta_1)(\theta_2-\theta_3) = (\partial_2\theta_2)(\theta_1-\theta_3) = (\partial_3\theta_3)(\theta_1-\theta_2) = 0. \tag{2.12}$$

The argument now divides into two cases, depending on whether all $\partial_\alpha\theta_\alpha$ terms vanish in Eq. (2.12) (case 1), or not (case 2).

Case 1: $\partial_1\theta_1 = \partial_2\theta_2 = \partial_3\theta_3 = 0$. then we have $\gamma^a_{bc} = \gamma^a_{bc}(t)$, so Lemma 2.1 from Ref. 19 implies that this model is spatially homogeneous. Since $R^{\;*}_{\alpha\beta} = 0$, it admits a group of Bianchi type I.

Case 2: there exists at least one nonzero $\partial_\alpha\theta_\alpha$ (no sum). Without loss of generality, let $\partial_1\theta_1\neq 0$. Then, Eq. (2.12) implies

$$\theta_2 = \theta_3,$$

from which, using (2.11),

$$\partial_2\theta_2 = \partial_3\theta_3 = 0. \tag{2.13}$$

Equations (2.11) and (II,A22) then imply that

$$\partial_2\mu = \partial_3\mu = 0.$$

We shall show that $\partial_1\mu\neq 0$. Assume that $\partial_1\mu = 0$. The commutator (I,A1) then requires $\partial_1\partial_0\mu = 0$. Using the contracted Bianchi identities (II,A24) and (II,A25) this requires $(\mu+p)\partial_1\theta_1 = 0$ which is impossible, by virtue of our assumptions. Consequently,

$$\partial_1\mu\neq 0, \tag{2.14}$$

so these solutions do not admit a group of Bianchi type I. By Eq. (II,A25), however, p *is* constant on the flat hypersurfaces. These solutions are LRS and belong to case IIa of Ref. 4 if $p\neq 0$, and to case IIbiii of Ref. 3 if $p = 0$. With condition (viiib) of I and II in force, Eq. (2.14) requires $dp/d\mu = 0$, and hence p is constant throughout space–time. The only physically relevant case then occurs for dust ($p = 0$), and the models are the "generalized Bianchi I" (case IIbiii) solutions of Ellis.[3]

We shall now discuss the Petrov classification of some of the cosmological models with flat slices.

Theorem 2.4: Suppose we have a cosmological model in which conditions (i), (ii), (iii), and (vi) of I and II are satisfied. Suppose further that the Cartesian tetrad (see Corollary 1 of Theorem 2.1) is Fermi-propagated. Then this model is in general of Petrov type I. The following special cases occur:

(i) Petrov type D if and only if the expansion tensor has two equal eigenvalues ($\tau = 0, \sigma\neq 0$) or a zero eigenvalue ($\pm[1/3(4\sigma^6-\tau)]^{1/2} - \theta\sigma^2 + \theta^3/9 = 0, \quad \sigma\neq 0$),

(ii) Petrov type O (conformally flat) if and only if the model is FRW.

Proof: Since $\boldsymbol{\omega} = \dot{\mathbf{u}} = \mathbf{0}$, a solution is Petrov type O (conformally flat) if and only if it is FRW (cf. Ref. 20).

We decompose the Weyl tensor into its electric and magnetic parts, as in II, Sec. 2. In the Fermi-propagated Cartesian tetrad, with $\boldsymbol{\omega} = \dot{\mathbf{u}} = \mathbf{0}$, the components of E_{ab} and H_{ab} are[21]

$$E_{\alpha\beta} = -\partial_0\theta_{\alpha\beta} - \tfrac{2}{3}\theta\theta_{\alpha\beta} - \sigma^{\gamma}_{\alpha}\sigma_{\beta\gamma} + \tfrac{1}{3}\delta_{\alpha\beta}$$

$$\times(\tfrac{1}{3}\theta^2 - \tfrac{1}{2}\mu - \tfrac{3}{2}p),$$

$$H_{\alpha\beta} = -\epsilon_{(\beta}{}^{\gamma\delta}\partial_{|\delta|}\sigma_{\alpha)\gamma}.$$

However, in this tetrad, the Jacobi identity (I,B6) becomes

$$\epsilon_{(\beta}{}^{\gamma\delta}\partial_{|\delta|}\sigma_{\alpha)\gamma} = 0,$$

so that

$$H_{\alpha\beta}\equiv 0.$$

The matrix $E_{\alpha B} + iH_{\alpha\beta}$ is now real and symmetric, so it has three real eigenvalues. Thus our solution is of Petrov type I, D, or O (cf. Ref. 22). Since $E^{\;\alpha}_{\alpha} = 0$ we may write $E_{\alpha\beta}$ in the form

$$E_{\alpha\beta} = -\partial_0\sigma_{\alpha\beta} - \tfrac{2}{3}\theta\sigma_{\alpha\beta} - \sigma_{\alpha\gamma}\sigma^{\gamma}_{\beta} + \tfrac{2}{3}\sigma^2\delta_{\alpha\beta}.$$

The trace free part of the field equation (I,B3) is

$$\partial_0\sigma_{\alpha\beta} + \theta\sigma_{\alpha\beta} = 0,$$

so $E_{\alpha\beta}$ becomes

$$E_{\alpha\beta} = \tfrac{1}{3}\theta\sigma_{\alpha\beta} - \sigma_{\alpha\gamma}\sigma^{\gamma}_{\beta} + \tfrac{2}{3}\sigma^2\delta_{\alpha\beta}. \tag{2.15}$$

We now restrict our attention to those solutions which are not FRW, so that $E_{\alpha\beta}\neq 0$ and $\sigma_{\alpha\beta}\neq 0$. A necessary and sufficient condition for the space–times to be Petrov type D is for $E_{\alpha\beta}$ to have two equal nonzero eigenvalues, i.e.,

$$I^3 - 6J^2 = 0,$$

where

$$I\equiv E^{\alpha\beta}E_{\alpha\beta}\neq 0 \quad\text{and}\quad J\equiv E_{\alpha\beta}E^{\beta\gamma}E^{\;\alpha}_{\gamma}\neq 0. \tag{2.16}$$

Substitution of Eq. (2.15) into Eq. (2.16), using Eqs. (I,2.5), yields, after a lengthy calculation,

$$I^3 - 6J^2 = \frac{2}{q}\tau\{\pm[1/3(4\sigma^6-\tau)]^{1/2} - \theta\sigma^2 + \theta^3/9\}^2. \tag{2.17}$$

The vanishing of τ is equivalent to demanding that the expansion tensor θ_{ab} have two equal eigenvalues. The vanishing of the second factor in (2.17) is equivalent to demanding that the expansion tensor possess a zero eigenvalue.

In Sec. 3, we shall consider possible ways of generalizing our results. The class of inhomogeneous cosmological models we consider in the present article includes all the well-known ones (such as those of Bondi and Tolman, of Taub, and of Szekeres, and most LRS models). However, there is a more obscure class of LRS models, case IIc of Ref. 4, in which the acceleration, $\dot{\mathbf{u}}$, does not vanish, and in which the fluid flow is orthogonal to conformally flat slices. What we shall now in effect show is that, if the fluid's expansion, θ, is nonzero, this class of LRS models admits no members in which either the slices are flat or the fluid density is constant on each slice.

Theorem 2.5: Any perfect fluid LRS cosmological model, in which conditions (i) and (viii) of I and II hold, satisfies at least one of the following conditions:

(a) it is not expanding ($\theta = 0$),

(b) it is spatially homogeneous,

(c) the fluid flows orthogonally to conformally flat hypersurfaces.

Proof: Stewart and Ellis[4] (cf. Ref.3) have divided all LRS perfect fluid solutions with $\mu + p \neq 0$ into three mutally exclusive classes. In case I, $\theta = 0$. In case III, the solutions are either spatially homogeneous or they belong to case IIIa of Ref. 4, in which case $\theta = 0$ (since $dp/d\mu \geqslant 0$ by condition viiib of I and II). All solutions in case II have a metric of the form

$$ds^2 = -\frac{-dt^2}{F^2(t,x)} + X^2(t,x)dx^2 + Y^2(t)[dy^2 + T^2(y)dz^2],$$

where $T = T(y)$ is a solution of $T''(y) + KT(y) = 0$ (K is a constant). The $\{t = \text{constant}\}$ sections (which are orthogonal to the fluid flow) are shown to be conformally flat by verifying that, in a suitable orthonormal tetrad, Eqs. (II,2.2)–(II,2.9) are satisfied. □

Theorem 2.6: Any perfect fluid LRS model in which conditions (i), (vi), and (viiia) of I and II hold satisfies at least one of the following conditions:

(a) it is not expanding ($\theta = 0$),

(b) the acceleration, $\dot{\mathbf{u}}$, is necessarily zero.

Proof: The (dust) solutions of Ellis[3] have $\dot{\mathbf{u}} = \mathbf{0}$. Since the fluid flow is hypersurface orthogonal, $\boldsymbol{\omega} = \mathbf{0}$, and so the case I solutions of Stewart and Ellis[4] are excluded. The only remaining solutions with $\dot{\mathbf{u}} \neq \mathbf{0}$ are those of case IIc of Ref. 4. Using the expression for $\overset{*}{R}_{\alpha\beta}$ given in I, Sec. 2, the constraints $\overset{*}{R}_{11} = \overset{*}{R}_{22} = 0$ imply

$$\partial_2 s - s^2 = \partial_1 a = a^2, \tag{2.18}$$

where $s = 2n_{31}$ and we have chosen the tetrad of Stewart and Ellis,[4] in which $n_{\alpha\beta}$ and a_α are of the form $a_\alpha = (a, s/2, 0)$ and

$$n_{\alpha\beta} = \begin{pmatrix} 0 & 0 & s/2 \\ 0 & 0 & 0 \\ s/2 & 0 & 0 \end{pmatrix}.$$

Using the commutator Eq. (I,A2) and the equations of Appendix A in Ref. 4, the derivative of Eq. (2.18) along $\mathbf{e}_0$ results in $a\dot{u}\theta_2 = 0$, where $\dot{u}_\alpha = (\dot{u},0,0)$. Since $a\dot{u} \neq 0$ in case IIc, this requires that $\theta_2 = 0$, and then the (01) field equation yields $\alpha = 0$, in which case $\theta = \alpha + 2\beta = 0$. □

Corollary 1: Any LRS model satisfying conditions (i) and (viii) of I and II with $dp/d\mu > 0$, in which the matter content is a perfect fluid whose congruence has nonzero expansion and flows orthogonally to flat hypersurfaces is spatially homogeneous of Bianchi type I.

Proof: Since the flow is hypersurface orthogonal, the only candidates belong to case II or case III of Ref. 4. However, in case III the solutions are either not expanding (case IIIa) or admit spatially homogeneous hypersurfaces, $\mathscr{S}$, orthogonal to the flow (case IIIb). Since in case IIIb the hypersurfaces $\mathscr{S}$ are not flat, the only allowable LRS models belong to case II. By Theorem 2.6 the matter flow must be geodesic, and the result now follows from our analysis in Sec. 1.

Corollary 2 (cf. Ref.10) Any spherically symmetric model satisfying conditions (i), (ii), and (viii) of I and II with $dp/d\mu > 0$, in which the matter content is a perfect fluid flowing orthogonally to flat hypersurfaces is a FRW ($k = 0$) model.

Proof: The contraction of (I,B3) yields (I,3.4) which shows that $\theta \neq 0$. By Corollary 1, the only possible solutions are spatially homogeneous of Bianchi type I. With the additional requirement of *spherical* symmetry the models must be FRW models, in which $k = 0$, since the spatially homogeneous hypersurfaces are flat. □

Remark: Corollary 2 shows that the result of Berger, Eardley, and Olson[10] can be proved without their assumption of regularity at the center of symmetry.

Theorem 2.7: Suppose that a perfect fluid LRS model satisfies conditions (i)and (viii) of I and II, and that the fluid flow is orthogonal to hypersurfaces of constant energy density. Then the fluid flowlines are geodesics.

Proof: The only LRS models in which the matter is a perfect fluid satisfying conditions (i) and (viiia) of I and II, and in which the fluid flowlines are hypersurface orthogonal but *not* geodesic are those of case IIc of Ref. 4. But the contracted Bianchi identities together with condition (viiib) of I and II then yield $(\mu + p)\dot{u} = -\partial_1 p = (-dp/d\mu)\partial_1 u = 0$, whence $\mu + p = 0$, which violates condition (viiia) of I and II. □

3. CONCLUSION

We have extensively examined space–times satisfying our conditions (i)–(iv) and (vii) of I and II. That is, we have considered space–times in which the matter content is a perfect fluid (satisfying Einstein's field equations with $\Lambda = 0$), whose flowlines are geodesics orthogonal to conformally flat hypersurfaces $\mathscr{S}$, and for which the fluid shear tensor, σ_{ij}, and the Ricci tensor, R_{ij} of the hypersurfaces have two equal eigenvalues. We have shown that these space–times constitute the class of inhomogeneous cosmologies examined by

Szekeres[5] and Szafron.[6] We have specialized to the case where the matter either is dust or it obeys an equation of state $p = p(\mu)$ with $dp/d\mu \neq 0$, and we have carefully investigated every subcase in our classification.

The next step is to generalize our models by relaxing some of the conditions (i)–(iv) and (vii) of I and II. It seems likely that the "next most tractable" subset will be the generalizations of our models either to the accelerating case ($\dot{\mathbf{u}} \neq 0$) or to the case when at least one of σ_{ij} of R_{ij} possesses distinct eigenvalues. We have made several unsuccessful attempts at the latter generalization, imposing the more restrictive condition (vi) of I and II in place of condition (iv). One would perhaps expect that Lorentzian 4-manifolds with flat spacelike hypersurfaces would be "few and far between," and that the imposition of Einstein's field equations for a perfect fluid would create further drastic restrictions. Nevertheless, the solutions appear to be rather elusive. We have attempted to analyze models satisfying conditions (i)–(iii) and (vi) [but not (vii)] of I and II, by examining the propagation of constraints, thereby generalizing the argument of the proof of Lemma 2.2. However, this rapidly becomes very complicated. We have even used a computer to facilitate the differentiation, substitution, and elimination procedure. While this enabled us to achieve a larger number of steps in the argument, we did not find it possible to proceed sufficiently far to reduce the problem to manageable proportions.

The alternative of introducing acceleration [i.e., relaxing condition (iii) of I and II, while retaining (vii)] may also lead to tractable sets of models. We are aware of certain inhomogeneous cosmological models satisfying conditions (i) and (viii) of I and II, in which $\dot{\mathbf{u}} \neq 0$. These are the accelerating perfect fluid locally rotationally symmetric models (case IIc) of Stewart and Ellis.[4] Some exact solutions in this class are known. For instance, the set of nonrotating tilted LRS Bianchi type V models lie in this class, where the hypersurfaces of homogeneity do *not* coincide with the hypersurfaces orthogonal to the fluid flow (in fact, these are the only spatially homogeneous tilted LRS models[23]). These models have been obtained in exact form by Farnsworth[24] in the case of dust ($p = 0$), and by Maartens and Nel[25] for Zel'dovich stiff matter ($p = \mu$). The proof of Theorem 2.5 shows that in *all* case IIc models the comoving slices are conformally flat, i.e., condition (iv) of I and II is necessarily satisfied. On the other hand, Theorem 2.6 shows that there are essentially *no* case IIc models in which the comoving slices are *flat*, i.e., in which condition (vi) of I and II holds.

Another possible way of generalizing our models is to obtain a means of weakening condition (iv) of I and II. Mathematically, this condition is expressed by the vanishing of the "York curvature tensor," Y_{ij} (see e.g. Refs. 26 and 27). Since this tensor is symmetric and tracefree, it seems natural to define scalars $Y_1 \geqslant 0$ and Y_2 by

$$Y_1^2 = Y_{ij} Y^{ij}$$

and

$$Y_2 = 4Y_1^6 - 3(Y_{ij} Y^{jk} Y^i_{\ k})^2,$$

in much the same way as the scalars σ and τ are defined for the symmetric tracefree tensor σ_{ij} [see I, Eq. (2.5)]. Then $Y_1 = 0 \Longleftrightarrow Y_{ij} = 0 \Rightarrow Y_2 = 0$. We can therefore devise an algebraic classification of the York curvature tensor, according as $Y_1 Y_2 \neq 0$, $Y_1 \neq Y_2 = 0$, or $Y_1 = Y_2 = 0$. This classification can then be combined in an obvious way to subclassify the various cases arising in Table II of I.

Wainwright (private communication) has recently considered the intrinsic properties of a class of space–times previously discovered by Oleson[28,29] and Wainwright[30]; we discuss some of his results below. The models satisfy conditions (i), (ii), and (viiia) of I and II, and condition (iii) is relaxed, in the sense that while the matter content is still a perfect fluid whose flowlines form a congruence of curves orthogonal to a family of spacelike hypersurfaces, this congruence is *not* in general geodesic. Moreover, condition (iv) of I and II is no longer necessarily valid: The hypersurfaces $\mathscr{S}$ are not in general conformally flat, and $Y_1 Y_2 \neq 0$. In general, the expansion tensor possesses three distinct eigenvalues, so condition (vii) of I and II no longer holds, and also condition (viiib) is generally violated. Imposing the condition that the flowlines be geodesics reduces the models to either LRS or spatially homogeneous space–times (whilst those of Oleson[29] become FRW). Conditions (v) of I and II are satisfied in the Petrov type II models of Oleson[29] only if they degenerate to FRW ($k = 0$) solutions. Imposing an equation of state, as in condition (viiib) of I and II, requires that it be the stiff matter (Zel'dovich) equation of state $p = \mu$, and that the Ricci tensor of $\mathscr{S}$ has two equal eigenvalues. Finally, solutions in a subclass of the (Petrov type N) models of Oleson[28] have the peculiar property that if they admit at least one Killing vector, then Y_1 and Y_2 are connected by the algebraic constraint $Y = 4Y_1^6$, i.e., the determinant of the York tensor Y_{ij} vanishes.

Despite the various obstacles to generalization that we have so far encountered, we are optimistic that the novel technique of imposing restrictions on the spatial geometry will be of great use in classifying and analyzing inhomogeneous cosmological models in general relativity. In examining inhomogeneous cosmologies, mathematical intractabilities abound, and one is constrained by a lack of technique. We submit that this "constraint" is so restrictive that, without the use of intrinsic symmetries, the "future development" of inhomogenous cosmologies will remain somewhat obscure, in more ways than one!

ACKNOWLEDGMENTS

We are grateful to G.W. Horndeski for helpful discussions, and to B.K. Berger, D.M. Eardley, and D.W. Olson for useful correspondence and for informing us of details of their work prior to publication. We also thank J. Wainwright for discussions relating to his current (unpublished) research, which we refer to in Sec. 3.

[1] C.B. Collins, and D.A. Szafron, "A new approach to inhomogeneous cosmologies: intrinsic symmetries. I," J. Math. Phys., preceding paper.
[2] D.A. Szafron and C.B. Collins, "A new approach to inhomogeneous cosmologies: intrinsic symmetries. II. conformally flat slices and an invariant classification," J. Math. Phys., preceding paper.

[3]G.F.R. Ellis, J. Math. Phys. **8**, 1171 (1967).
[4]J.M. Stewart and G.F.R. Ellis, J. Math. Phys. **9**, 1072 (1968).
[5]P. Szekeres, Commun. Math. Phys. **41**, 55 (1975).
[6]D.A. Szafron, J. Math. Phys. **18**, 1673 (1977).
[7]J. Wainwright, J. Math. Phys. **18**, 672 (1977).
[8]W.B. Bonnor and N. Tomimura, Mon. Not. R. Astron. Soc. **175**, 85 (1976).
[9]D.A. Szafron and J. Wainwright, J. Math. Phys. **18**, 1668 (1977).
[10]B.K. Berger, D.M. Eardley, and D.W. Olson, Phys. Rev. D **16**, 3086 (1977).
[11]R. Kantowski and R.K. Sachs, J. Math. Phys. **7**, 443 (1966).
[12]A. Spero and D.A. Szafron, J. Math. Phys. **19**, 1536 (1978).
[13]D.M. Eardley, E.P.T. Liang, and R.K. Sachs, J. Math. Phys. **13**, 99 (1972).
[14]A.H. Taub, Ann. Math. **53**, 472 (1951).
[15]A.H. Taub, Phys. Rev. **103**, 454 (1956).
[16]J. Wainwright, Gen. Relativ. Grav. **10**, 797 (1977).
[17]H. Stephani, Commun. Math. Phys. **9**, 53 (1968).
[18]D. Laugwitz, *Differential and Riemannian Geometry* (Academic, New York, 1965).
[19]G.F.R. Ellis and M.A.H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).
[20]G.F.R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology, Proceedings of the International School of Physics Enrico Fermi Course* XLVII, 1969, edited by R.K. Sachs (Academic, New York, 1971).
[21]M.A.H. MacCallum, "A Class of Homogeneous Cosmological Models," Ph.D. thesis, Cambridge University (1969).
[22]L.D. Landau and E.M. Lifshitz, The Classical Theory of Fields (Pergamon, New York, 1975).
[23]A.R. King and G.F.R. Ellis, Commun. Math. Phys. **31**, 209 (1973).
[24]D.L. Farnsworth, J. Math. Phys. **8**, 2315 (1967).
[25]R. Maartens and S.P. Nel, "Spatially homogeneous, locally rotationally symmetric cosmological models", B. S. thesis, University of Cape Town (1976).
[26]L.P. Eisenhart, *Riemannian Geometry* (Princeton U.P., Princeton, New Jersey, 1926).
[27]J.W. York, Phys. Rev. Lett. **26**, 1956 (1971).
[28]M. Oleson, J. Math. Phys. **12**, 666 (1971).
[29]M.K. Oleson, "Algebraically special fluid space–times with shearing rays," Ph.D. thesis, University of Waterloo (1972).
[30]J. Wainwright, Int. J. Theor. Phys. **10**, 39 (1974).

Erratum: Generalized *C* metric [J. Math. Phys. 19, 1986 (1978)]

Frederick J. Ernst
Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616

Equations (3) and (4) should both have a factor ρ on the right hand side.

Erratum: Coping with different languages in the null tetrad formulation of general relativity [J. Math. Phys. 19, 489 (1978)]

Frederick J. Ernst
Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616

A number of misprints have been discovered in Appendix B. In the first term on the right hand side of Eq. (1c) the complex conjugate w^* should appear instead of w. In the last term on the right hand side of Eq. (4c) B^*_+ should appear instead of B^*_t. In the last term on the right hand side of both Eqs. (8b) and (8c) $\overline{B}_+$ should appear instead of $\overline{B}_-$. Finally, in the third line of the text the word "frequently" should appear instead of "frequency."

 0022-2488/79/112371-01$00.00

Erratum: Representations of a para-Bose algebra using only a single Bose field
[J. Math. Phys. 20, 390 (1979)]

R. Jagannathan and R. Vasudevan
Department of Electrical Engineering, University of New Hampshire, Durham, New Hampshire 03824

The nonvanishing elements of matrices F_2 and F_3 in Eq. (23) are incorrect. The correct values are given by

$$F^4_{25} = -F^5_{24} = -\tfrac{1}{2}, \quad F^4_{34} = -F^5_{35} = \tfrac{1}{2}.$$

Then the GSO(2,1) considered is seen to be semisimple with

$$\det\|h\|\neq 0, \quad \det\|g\|\neq 0,$$

and hence not satisfying the C-theorem of Pais and Rittenberg,[1] i.e., there does not exist an antisymmetric matrix C relating F and S matrices as $S^m = CF_m, \forall m = 1,2,3$. Hence, some of the statements regarding GSO(2,1) occurring in the penultimate paragraph after Eq. (27) of the text need to be reexamined.

[1]A. Pais and V. Rittenberg, J. Math. Phys. **16**, 2062 (1975).

Erratum: Linearized analysis of inhomogeneous plasma equilibria: General theory
[J. Math. Phys. 20, 413 (1979)]

H. Ralph Lewis and Keith R. Symon
Los Alamos Scientific Laboratory, P. O. Box 1663, M. S. 642, Los Alamos, New Mexico 87545

1) The sentence beginning in the fourth line after Eq. (II.21) should read "See the note added in proof, Ref. 12."

2) Equation (IV.23) should read

$$J^*_{srn'} = \int dQ\, dP\Gamma^{(c)}\mathscr{F}_s(Q,P,p_k)\delta(H_s - E_s) \times z^*_{n'}(Q)u_{sr}(Q,P)\big|_{p_k = p_k^0}, \tag{IV.23}$$

3) Equation (IV.24) should read

$$\langle n'|D|n\rangle = \lambda_n\delta_{nn'} - \delta_{\kappa\kappa'}\sum_{s(r)}\frac{K_{sr}(\omega)J^*_{srn'}H'_{srn}}{\mu_{sr} - \omega}, \tag{IV.24}$$

4) Equation (IV.25) should read

$$J_{srn} = H'_{srn}. \tag{IV.25}$$

5) In the sixth line following Eq. (V.6), "amplitudes and

6) In the third line before Eq. (A1), $\phi(r)$ should read $\phi(\mathbf{r})$.

$$\delta H_s = \sum_i \delta\mathbf{r}\cdot(\nabla\phi_i)\frac{\partial H_s}{\partial\phi_i} + \sum_j\left(\delta q_j\frac{\partial H_s}{\partial q_j} + \delta p_j\frac{\partial H_s}{\partial p_j}\right),$$

7) The equation following Eq. (A12) should read

8) In Ref. 2, "D.C. RD. C. Robinson" should read "D.C. Robinson."

9) In Ref. 10, "K.R." should read "K.R. Symon."

10) In the seventh line of the right-hand column of p. 423, "indeces" should read "indices."

 0022-2488/79/112372-01$00.00