Matrix elements for infinitesimal operators of the groups U($p+q$) and U(p,q) in a U(p)$\times$U(q) basis. I

Anatoli U. Klimyk
Institute for Theoretical Physics, Academy of Sciences of the Ukrainian SSR, Kiev-130, USSR

Bruno Gruber
Department of Physics, Southern Illinois University at Carbondale, Carbondale, Illinois 62901

(Received 10 August 1978)

In this article explicit expressions are obtained for the action of the infinitesimal operators of the principal nonunitary series representations of the groups U(p,q) in a U(p)$\times$U(q) basis. It is moreover shown how the finite dimensional irreducible representations of the group U(p,q) and the group U($p+q$) with respect to a U(p)$\times$U(q) basis are obtained from the principal nonunitary series representations of the group U(p,q).

1. INTRODUCTION

The significance of semisimple Lie groups and Lie algebras in physics is well established,[1] especially in particle physics,[2] nuclear physics,[3] and atomic physics.[4] More recently, the range of their application has been extended into molecular physics,[5] solid state physics,[6] and quantum chemistry.[7] Moreover, their range of applications has been extended from originally purely kinematical aspects of physical systems to dynamical aspects, thus adding significantly to their overall importance in physics.[8]

The Lie algebras and Lie groups find their way into physics in the form of representations. Their elements act as linear operators on the states of the physical systems. Thus it is representation theory of semisimple Lie algebras and Lie groups which plays a dominating role in their application.

In this article we will be concerned with the Lie groups U($p+q$) and U(p,q), $p,q\geqslant 1$, integer, which have found a wide area of applications in physics. As it is representation theory which characterizes the use of these groups in physics, the most important aspects to be dealt with are the calculation of the matrix elements (ME) of the infinitesimal operators and finite transformations, and the calculation of the Clebsch–Gordon coefficients (CGC). For the compact Lie groups U(n) [and SO(n)] the ME's and CGC's are normally evaluated in the so-called canonical basis, corresponding to the reduction U($n-1$)$\supset$U($n-2$)$\supset\cdots\supset$U(1) [SO($n-1$)$\supset$SO($n-2$)$\supset\cdots\supset$SO(2) for SO(n)]. In fact, for the groups U(n) the ME's have been obtained in explicit form in this basis.[9] However, it turns out that in many physical applications it is not this canonical basis that describes the actual physical states most conveniently. In fact it is often necessary to go over, by means of a transformation, from these "canonical states" to the "physical states." That is, to those states which are well defined with respect to the quantum numbers of the physical system which is under consideration. Of the many examples the nuclear Elliott model is possibly the best known. The simple Lie group utilized in this model is U(3), with the canonical chain of subgroups U(2)$\supset$U(1). The states defined by this chain are, however, not the physically relevant states. Indeed, the physically relevant chain is U(3)$\supset$O(3). Thus given the states and ME's with respect to the canonical basis, these states have to be transformed into the basis U(3)$\supset$O(3) by means of the Moshinsky transformation brackets.[10] Another special case is the conformal group U(2,2)$\sim$O(4,2) with its maximal compact subgroup U(2)$\times$U(2), which is of considerable interest to physics.[11]

While the ME's are known for the groups U(n) with respect to their canonical basis, this is in general not the case for the noncanonical basis of U(n), even though the noncanonical basis may be the one which the physical problem requires. The reason for this situation is that the calculation of the ME's and CGC's in a noncanonical basis presents greater problems than in the canonical basis. In this article we will address ourselves to the calculation of the ME's and CGC's in certain noncanonical bases.

One of the methods to study representations of compact Lie groups in a noncanonical basis is by means of the theory of the principal nonunitary series representations of an appropriately chosen noncompact semisimple Lie group.[12] This method appears to be the simplest for the calculation of the ME's and CGC's of the groups U($p+q$) and U(p,q) in a noncanonical basis. The idea behind this method is the following.

Let G denote a connected linear (i.e., matrix) semisimple or reductive (a direct product of semisimple and commutative groups) real noncompact Lie group with maximal compact subgroup K. Let $[G]$ denote the complexification of G and let G_k denote a *compact* form of $[G]$. The groups G and G_k have the same irreducible finite dimensional representations. Moreover, every *complex analytic* irreducible finite dimensional representation (i.e., representations whose ME's are complex analytic functions of the complex group parameters) of $[G]$ is also an irreducible representation of its subgroups G and G_k.[13] Every irreducible finite dimensional representation is obtained in this manner through restriction from $[G]$ to its subgroups G and G_k. Thus, the CGC's for the finite dimensional representations of the groups G and G_k are also the same. The ME's of *finite transformations* which correspond to finite dimensional representations of G_k can be obtained from the matrix elements of these representations of G by means of analytic continuation of the group

 0022-2488/79/101995-16$01.00

TABLE I.

Type of group	noncompact real Lie group G	corresponding compact Lie group G_k	maximal compact subgroup K in G	remarks
A I	$SL(n,R)$	$SU(n)$	$SO(n)$	$n>1$
A II	$SU^*(2n)$	$SU(2n)$	$Sp(n)$	$n>1$
A III	$SU(p,q)$	$SU(p+q)$	$S(U(p)\times U(q))$	
	$U(p,q)$	$U(p+q)$	$U(p)\times U(q)$	
B I,D I	$SO_0(p,q)$	$SO(p+q)$	$SO(p)\times SO(q)$	
D III	$SO^*(2n)$	$SO(2n)$	$U(n)$	$n>2$
C I	$Sp(n,R)$	$Sp(n)$	$U(n)$	
C II	$Sp(p,q)$	$Sp(p+q)$	$Sp(p)\times Sp(q)$	

parameters in the frame of the group $[G]$. Thus, an investigation of all the ME's and CGC's for the group G_k is equivalent to an investigation of all the ME's and CGC's of the group G (for the case of the finite dimensional irreducible representations).

Every irreducible *finite* dimensional representation of G is, however, contained as a subrepresentation in an appropriately chosen principal nonunitary series representation of G (the principal nonunitary series representations are obtained from the principal unitary series representations by means of analytic continuation of the continuous representation parameters onto the entire complex space of these parameters). Thus, if the ME's are known of the infinitesimal operators and the finite transformations for the principal nonunitary series representations of G, then we know (in principle) the finite dimensional irreducible representations of G. It turns out that it is possible to obtain the ME's of the infinitesimal operators for the principal nonunitary series representations of the group G in the basis of its maximal compact subgroup K (K basis). From the argument given above it follows that the finite dimensional irreducible representations of the infinitesimal operators of G (and consequently of G_k) can be obtained in the K basis. By utilizing the methods developed in Refs. 12 and 14 it is then possible to obtain the ME's of the *finite transformations*, once the ME's of the infinitesimal operators have been obtained.

The finite dimensional representations of the groups G_k obtained in this manner are not yet in unitary form. In order to obtain unitary representations, the similarity transformations have to be found which unitarize these representations. The problem of finding the matrices corresponding to the unitarizing similarity transformations is indeed the most difficult problem of the entire method for the construction of the finite dimensional representations of the group $U(p+q)$ in a $U(p)\times U(q)$ basis as described in this article. For the case that every irreducible representation of K is contained in an irreducible representation of G_k at most once, the unitarizing matrix is easily found. The unitarizing matrix for an irreducible representation of G_k with this property is simply a diagonal matrix.

From the discussion given above it follows that it is possible to find the infinitesimal operators for any finite dimensional irreducible representation of a compact semisimple (or reductive) Lie group, in the basis of one of its subgroups K, if there exists a real form G of the complexification $[G_k]$ of G_k for which K forms a *maximal compact* subgroup. The possibilities permitted by this method for the classical Lie groups are enumerated in Table I. The notation for the groups is chosen to coincide with the notation of Ref. 15. Table I shows that, for the case of the groups $SU(n)$ the method outlined above permits us to obtain the finite dimensional irreducible representations with respect to the subgroup $SO(n)$, the subgroup $Sp(n)$ (if n is even), and the subgroups $S(U(p)\times U(q))$, $p+q=n$.

In this article explicit formulas are given for the action of the infinitesimal operators of the principal nonunitary series representations of $U(p,q)$ in a $U(p)\times U(q)$ basis. Then, given a finite dimensional representation of $U(p,q)$, the principal nonunitary series representation is defined which contains this finite dimensional representation. It is explained how to obtain the infinitesimal operators for the finite dimensional representations of $U(p+q)$ in a $U(p)\times U(q)$ basis in *unitarized form*. The unitarization can be carried out for each individual case separately. In a second article to follow it will be shown, for the case of the degenerate series representations of $U(p,q)$, how to obtain the explicit form for the infinitesimal operators of the finite dimensional representations of $U(p,q)$ with highest weights $(\lambda_1,0\ldots,0,\lambda_2)$ in a $U(p)\times U(q)$ basis in a *general* manner (i.e., not merely on a case to case basis). In concluding it should be mentioned that the maximally degenerate series representations of $U(p,q)$ in a $U(p)\times U(q)$ basis were discussed in Ref. 16.

2. PRINCIPAL NONUNITARY SERIES REPRESENTATIONS OF SEMISIMPLE (REDUCTIVE) LIE GROUPS

This section of the article serves the purpose to introduce those concepts and definitions of the theory of semisimple and reductive Lie algebras which will be needed in the following. Particular attention will be given the definition of the principal nonunitary series representations, and a basic lemma will be quoted. For more detailed information on the theory of semisimple (reductive) Lie algebras and Lie groups, as well as their representations, the reader is referred to Ref. 17.

Let G denote a connected linear (i.e., matrix) semisimple or reductive Lie group, and let $\mathfrak{g}$ denote its Lie algebra. Let K be a maximal compact subgroup of G, and let $\mathfrak{k}$ denote its Lie subalgebra in $\mathfrak{g}$. Let $B(\cdot,\cdot)$ denote the Killing–Cartan form on the space $\mathfrak{g}$. Moreover, let $\mathfrak{p}$ denote the orthogonal complement of $\mathfrak{k}$ in $\mathfrak{g}$ with respect to $B(\cdot,\cdot)$ Then $\mathfrak{g} = \mathfrak{p} + \mathfrak{k}$ (direct sum). This decomposition of $\mathfrak{g}$ defines the Cartan involution θ; the involution θ leaves every element of $\mathfrak{k}$ invariant and multiplies every element of $\mathfrak{p}$ by the factor (-1). The involution θ is extended to $\mathfrak{g}$ by linearity. The form $\langle \mathfrak{x},\mathfrak{y}\rangle = -cB(\mathfrak{x},\theta\mathfrak{y})$, $c>0$, $\mathfrak{x},\mathfrak{y}\in\mathfrak{g}$, defines a positive definite scalar product on $\mathfrak{g}$. This scalar product can be extended to the complexification $\mathfrak{g}_c$ of $\mathfrak{g}$ by means of linearity.

Let $\mathfrak{a}$ be a maximal commutative subalgebra in $\mathfrak{p}$. The dimensionality of $\mathfrak{a}$ is called the real rank of $\mathfrak{g}$. Let $\mathfrak{m}$ denote the centralizer of $\mathfrak{a}$ in $\mathfrak{k}$. If $\mathfrak{b}$ is a Cartan subalgebra of $\mathfrak{m}$, then $\mathfrak{b} + \mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$. Consider the set of operators $\mathrm{ad}\mathfrak{H}$, $\mathfrak{H}\in\mathfrak{a}$, which act on the space $\mathfrak{g}$. With respect to these operators, the space $\mathfrak{g}$ can be decomposed into a direct sum of subspaces corresponding to eigenvalues λ,

$$\mathfrak{g} = \mathfrak{g}_0 + \sum_\lambda \mathfrak{g}_\lambda.$$

The sum extends over all nonzero linear forms (eigenvalues) λ on the subspace $\mathfrak{a}$ of $\mathfrak{g}$. The subspace $\mathfrak{g}_0$ is the eigenspace which corresponds to zero eigenvalue. The linear forms λ on $\mathfrak{a}$ are called restricted roots of the pair $(\mathfrak{g}, \mathfrak{a})$. The roots may have more then unit multiplicity (this implies that the subspaces $\mathfrak{g}_\lambda$ may have dimensionality greater than 1). The restricted roots can be obtained from the roots of the complexification $\mathfrak{g}_c$ of $\mathfrak{g}$ by restriction to $\mathfrak{a}$. The nonzero restricted roots can be separated into two sets; namely, positive roots $(\lambda>0)$ and negative roots $(\lambda<0)$. Define $\mathfrak{n} = \Sigma_{\lambda>0}\, g_\lambda$. Then $\mathfrak{n}$ is a maximal nilpotent subalgebra in $\mathfrak{g}$. The Iwasawa decomposition of $\mathfrak{g}$ is then given by the direct sum $\mathfrak{g} = \mathfrak{k} + \mathfrak{a} + \mathfrak{n}$. If N, A denote the analytic subgroups of G which correspond to the Lie algebras $\mathfrak{n}$ and $\mathfrak{a}$, respectively, then the Iwasawa decomposition of G is given by $G = ANK$. Moreover, every element $g\in G$ can be decomposed uniquely into the product $g = hnk$, $h\in A$, $n\in N$, and $k\in K$. If M is the centralizer of A in K, then $\mathfrak{m}$ is the Lie algebra of M.

Let δ denote an irreducible finite dimensional unitary representation of M on a space V and let Λ be a complex linear form on $\mathfrak{a}$. The map $h\rightarrow\exp[\Lambda(\log h)]$, $h\in A$, defines a representation of A. Let $L^2_\delta(K,V)$ denote the Hilbert space of measurable vector functions f from K into V, such that

$$\int_K \|f(k)\|^2_V\, dk < \infty, \quad f(mk) = \delta(m)f(k), \quad m\in M,$$

(dk is an invariant measure on K, $\|\ \|_V$ means norm in V), with scalar product

$$(f_1,f_2) = \int_K (f_1(k),f_2(k))_V dk.$$

Then the operators $\pi_{\delta,\Lambda}(g)$, $g\in G$, which act in $L^2_\delta(K,V)$ according to the formula

$$\pi_{\delta,\Lambda}(g)f(k) = \exp[\Lambda(\log\ h)]\,f(k_g), \quad kg = hnk_g, \tag{1}$$

where $h\in A$, $n\in N$, $k_g\in K$, define a representation of G on the space $L^2_\delta(K,V)$. The representations $\pi_{\delta,\Lambda}$ form the *principal nonunitary series* representations of G.

It has been shown in Refs. 18 and 19 that every completely (or infinitesimally) irreducible representation of G which decomposes into a sum of finite dimensional unitary representations of K, each of which occurs a finite number of times only, is infinitesimally equivalent to a subquotient of some principal nonunitary series representation (for the definitions of infinitesimal equivalence and infinitesimal irreducibility see Ref. 17). In particular, every finite dimensional irreducible representation of G is contained in this manner in some representation $\pi_{\delta,\Lambda}$ (more details follow below).

The space $L^2_\delta(K,V)$ can be decomposed into an (infinite) orthogonal sum of finite dimensional subspaces H_ν, on which $\pi_{\delta,\Lambda}$ realizes representations of K which are multiples of irreducible representations of K with highest weight ν. The space of finite linear combinations of vectors of different H_ν, i.e., the space of K finite vectors of $L^2_\delta(K,V)$, will be denoted by $dL^2_\delta(K,V)$. A representation of the algebra $\mathfrak{g}$ on $dL^2_\delta(K,V)$, is obtained through differentiation of the representation $\pi_{\delta,\Lambda}$ of G. It will be denoted by $d\pi_{\delta,\Lambda}$.

Apart from the compact subgroups K and M of G, other compact subgroups of G will be needed in what follows. In order to introduce these subgroups additional information is needed concerning the restricted roots of the pair $(\mathfrak{g}, \mathfrak{a})$. Let Δ denote the set of all positive restricted roots (with multiplicities), and F a set of simple roots in Δ (i.e., a smallest subset of roots of Δ such that any root of Δ can be represented as a linear combination of F with nonnegative integer coefficients). The number of simple restricted roots equal the real rank of G. If $\alpha_1,\alpha_2,\ldots,\alpha_l$ denote such a set of simple roots, then let $F_i = \{\alpha_i\ \alpha_{i+1},\ldots,\alpha_l\}$, $i = 1,2,\ldots,l$. Let Δ_i denote the subset of roots of Δ which can be expressed as linear combinations of the roots of F_i. Let

$$\mathfrak{n}_i = \sum_{\lambda\in\Delta_i}\mathfrak{g}_\lambda \quad\text{and}\quad \mathfrak{n}_i^- = \sum_{\lambda\in\Delta_i}\mathfrak{g}_{-\lambda}.$$

Let G_i be that subgroup of G which is generated by the subgroups $N_i = \exp\mathfrak{n}_i$, $N_i^- = \exp\mathfrak{n}_i^-$, and M. Let $K_i = G_i\cap K$. Then the following sequence of subgroups is obtained,

$$K\equiv K_1\supset K_2\supset\cdots\supset K_l\supset M\equiv K_{l+1}. \tag{2}$$

This chain of subgroups can be further added to. Let $\mathfrak{h}_1$, $\mathfrak{h}_2,\ldots,\mathfrak{h}_l$ denote a set of elements of the subalgebra $\mathfrak{a}$ of $\mathfrak{g}$ for which $\alpha_i(\mathfrak{h}_j) = \delta_{ij}$ holds, where α_i, $i = 1,2,\ldots,l$ are the simple roots. Let Δ^j_{i+1} (j a positive integer) denote that subset of roots of Δ_i for which either $\alpha(\mathfrak{h}_i)\geqslant j$ or $\alpha(\mathfrak{h}_i) = 0$. $\Delta^7_{i+1} = \Delta_{i+1}$ always holds, and for the classical Lie algebras $\Delta^3_{i+1} = \Delta_{i+1}$ holds. Let $\mathfrak{g}^j_{i+1}$ be that subalgebra of $\mathfrak{g}$ which is generated by the root spaces $\mathfrak{g}_\lambda$ and $\mathfrak{g}_{-\lambda}$, $\lambda\in\Delta^j_{i+1}$, and by the subalgebra $\mathfrak{m}$; let G^j_{i+1} be an analytic subgroup of G with Lie algebra $\mathfrak{g}^j_{i+1}$ and let $K^j_{i+1} = G^j_{i+1}\cap K$. Thus, to every pair of subgroups K_i and K_{i+1} the following sequence of subgroups has been obtained,

$$K_i\equiv K^1_{i+1}\supseteq K^2_{i+1}\supseteq\cdots\supseteq K_{i+1}. \tag{3}$$

Let $\mathfrak{k}^j_i$ denote the Lie algebra of K^j_i, considered as subalgebra of $\mathfrak{g}$. Then the scalar product $<\cdot,\cdot>$ defined on $\mathfrak{g}$ defines a scalar product on the subspace $\mathfrak{k}^j_i$. Let $\mathfrak{X}_1$, $\mathfrak{X}_2,\ldots,\mathfrak{X}_t$ be an orthonormal basis for $\mathfrak{k}^j_i$. The left action of the elements $\mathfrak{X}_1$, $\mathfrak{X}_2,\ldots,\mathfrak{X}_t$ on the space $dL^2_\delta(K,V)$ is given by

$$\sigma(X_i)f(k) = \frac{d}{dt}f((\exp t\mathfrak{X}_i)k)\big|_{t=0}. \qquad (4)$$

The operator

$$\omega_i^j = -\sum_{s=1}^{t}\sigma(\mathfrak{X}_s^2)$$

is the Casimir operator of $\mathfrak{k}_i^j$. In addition to the Casimir operators, the following operators Q_i will be required.

$$Q_i = \omega_i^1 - \sum_{j=1}^{5}\frac{1}{j(j+1)}\omega_i^{j+1}, \qquad (5)$$

where $\omega_i^{j+1} = 0$ if $K_i^{j+1} = K_i$.

Let $\mathfrak{H}_i$, $i = 1,2,\ldots,l$ denote an orthonormal basis in $\mathfrak{a}$. Let R denote half the sum of the roots of Δ (with multiplicities taken into account), and let ρ be an element of $\mathfrak{a}$ such that $R(\mathfrak{H}) = \langle\rho,\mathfrak{H}\rangle$, $\mathfrak{H}\in\mathfrak{a}$.

In Ref. 20 Thieleker proved a lemma (not published), which reformulated, can be stated as follows (the proof of this lemma can be found in Ref. 21.)

Basic Lemma: Let $f\in dL_\delta^2(K,V)$. Then for an element $\mathfrak{Y}\in\mathfrak{p}_c$, $\mathfrak{p}_c$ the complexification of $\mathfrak{p}$, the following holds,

$$d\pi_{\delta,\Lambda}(\mathfrak{Y})f(k)$$

$$= \sum_{i=1}^{l}\Lambda(\mathfrak{H}_i)\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{H}_i\rangle f(k) - \langle(\mathrm{ad}k)\mathfrak{Y},\rho\rangle f(k)$$

$$+ \tfrac{1}{2}\sum_{i=1}^{l}[Q_i,\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle]f(k), \qquad (6)$$

where $\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{H}\rangle$, $\mathfrak{H}\in\mathfrak{a}$, is a function on K defined by the scalar product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}_c$, $[\cdot,\cdot]$ is a commutator [the pointwise multiplication of $f(k)\in dL_\delta^2(K,V)$ by $\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle$ is to be considered as action of $\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle$ taken as operator, on the function $f(k)$], $\mathfrak{h}_i$, $i = 1,2,\ldots,l$ are the elements of $\mathfrak{a}$ which were introduced earlier and which satisfy $\alpha_j(\mathfrak{h}_i) = \delta_{ij}$. The operators Q_i act on the vector functions from the left [see Eqs. (4) and (5)].

Using the lemma stated above and the CGC's for the group $K\cong U(p)\times U(q)$, explicit formulas are obtained for the action of the infinitesimal operators of the group $U(p,q)$ for the representations $\pi_{\delta,\Lambda}$ with respect to a $U(p)\times U(q)$ basis.

3. THE GROUP *U(p,q)* AND ITS SUBGROUPS

The group $U(p,q)$ is the group of complex nonsingular $(p+q)\times(p+q)$ matrices which leave invariant the form

$$|z_1|^2 + |z_2|^2 + \cdots + |z_p|^2 - |z_{p+1}|^2 - \cdots - |z_{p+q}|^2.$$

The maximal compact subgroup K is isomorphic to the group $U(p)\times U(q)$ and can be chosen as the set of matrices

$$\begin{bmatrix} U(p) & 0 \\ 0 & U(q)\end{bmatrix}.$$

The Cartan involution θ for the Lie algebra $u(p,q)$ of the group $U(p,q)$ is given by $\theta(\mathfrak{X}) = -(\bar{\mathfrak{X}})^T$, $\mathfrak{X}\in u(p,q)$. The bar denotes complex conjugation and the subscript T stands for transposition. The scalar product defined by

$$\langle\mathfrak{X},\mathfrak{Y}\rangle = \mathrm{Tr}\,\mathfrak{X}\bar{\mathfrak{Y}}^T, \quad \mathfrak{X},\mathfrak{Y}\in u(p,q), \qquad (7)$$

coincides, up to a scalar constant, with the scalar product introduced in the previous section.

The subspace $\mathfrak{p}$ of $u(p,q)$ consists of the linear combinations of the matrices

$$\sqrt{-1}(E_{ij} - E_{ji}), \quad E_{ij} + E_{ji}, \quad 1\leqslant i\leqslant p,$$
$$p+1\leqslant j\leqslant p+q, \qquad (8)$$

where the E_{ij} are defined by the relation $(E_{ij})_{ks} = \delta_{ik}\delta_{js}$.

The subalgebra $\mathfrak{a}$ can be selected as the set of matrices

$$\begin{array}{c} p-q\,\{ \\ q\,\{ \\ q\,\{ \end{array}\left(\begin{array}{c|ccc|ccc} 0 & & 0 & & & 0 & \\ \hline 0 & & 0 & & & & a_q \\ & & & & & \iddots & \\ & & & & & a_2 & \\ & & & & a_1 & & \\ \hline & & & a_1 & & & \\ & & a_2 & & & & \\ 0 & \iddots & & & & 0 & \\ & a_q & & & & & \end{array}\right), \qquad (9)$$

$a_i\in R$.

The subgroup M is isomorphic to the group $U(p-q)\times[U(1)\times\cdots\times U(1)]$ (there are q terms in the square brackets) and consists of the matrices

$$\left(\begin{array}{c|ccc|ccc} U(p-q) & & 0 & & & 0 & \\ \hline & u_q & & & & & \\ & & u_{q-1} & & & & \\ 0 & & & \ddots & & 0 & \\ & & & u_1 & & & \\ \hline & & & & u_1 & & \\ & & & & & u_2 & \\ 0 & & 0 & & & \ddots & \\ & & & & & & u_q \end{array}\right), \qquad (10)$$

$u_i\in U(1)$.

In order to obtain the subgroups K_i^j explicitly, the following transformation φ is used, $\varphi: g\to dgd^{-1}$, $g\in U(p,q)$, where

$$d = \left[\begin{array}{c|c} E_{p-q} & 0 \\ \hline 0 & D\end{array}\right],$$

with E_{p-q} the $(p-q)$-dimensional unit matrix and

$$D=\frac{1}{\sqrt{2}}\left(\begin{array}{cccc|cccc} 1 & & & & & & & 1 \\ & 1 & & & & & 1 & \\ & & \ddots & & & \cdot^{\cdot^{\cdot}} & & \\ & & & 1 & 1 & & & \\ \hline & & & 1 & -1 & & & \\ & & 1 & & & -1 & & \\ & \cdot^{\cdot^{\cdot}} & & & & & \ddots & \\ 1 & & & & & & & -1 \end{array}\right)\begin{array}{l} \}\, q \\ \\ \}\, q \end{array}.$$

Under φ the group $U(p,q)$ transforms into a group $U'(p,q)$ which is isomorphic to $U(p,q)$. As the two groups are isomorphic, the prime will be deleted in the following—there being no danger of confusion. It is clear that φ is also a transformation for the corresponding algebras $u(p,q) \rightarrow u'(p,q)$. The transformation φ carries the algebra $\mathfrak{a}$ into the algebra $\mathfrak{a}'$, consisting of the matrices

$$\begin{array}{l} p-q \\ \\ \\ \\ \\ \end{array}\left(\begin{array}{c|cccc|cccc} 0 & & 0 & & & & 0 & & \\ \hline & a_q & & & & & & & \\ & & a_{q-1} & & & & & & \\ 0 & & & \ddots & & & 0 & & \\ & & & & a_1 & & & & \\ \hline & & & & & -a_1 & & & \\ & & & & & & -a_2 & & \\ 0 & & 0 & & & & & \ddots & \\ & & & & & & & & -a_q \end{array}\right), \quad a_i \in R. \qquad (11)$$

The transformation φ leaves the subgroup M elementwise invariant.

It will turn out to be convenient to define the restricted roots and the root spaces with respect to the subalgebra $\mathfrak{a}'$. It is easy to go from the subalgebra $\mathfrak{a}'$ to the algebra $\mathfrak{a}$ by means of the transformation φ whenever it turns out to be necessary.

The matrices

$\tilde{\mathfrak{h}}_1 = E_{pp} - E_{p+1,p+1}$, $\tilde{\mathfrak{h}}_2 = E_{p-1,p-1} - E_{p+2,p+2}, \ldots,$
$\tilde{\mathfrak{h}}_q = E_{p-q+1,p-q+1} - E_{p+q,p+q}$,

form a basis for $\mathfrak{a}'$. Thus an arbitrary element $\mathfrak{H} \in \mathfrak{a}'$ can be written as $\mathfrak{H} = \Sigma_{i=1}^{q} \omega_i \tilde{\mathfrak{h}}_i$, $\omega_i \in R$. Following Jacobson[22] the restricted roots will be characterized by the meaning they have on the elements $\mathfrak{H} \in \mathfrak{a}'$. Thus, the linear forms

$$\omega_i - \omega_j,\ \ i \neq j; \quad \omega_i + \omega_j,\ \ i \neq j; \ -\omega_i - \omega_j,\ \ i \neq j;$$
$$\omega_i;\ -\omega_i;\quad 2\omega_i;\ -2\omega_i;$$
$$i,j = 1,2,\ldots,q$$

on the elements $\mathfrak{H} = \Sigma_{i=1}^{q} \omega_i \tilde{\mathfrak{h}}_i$ are the restricted roots of the pair $(u'(p,q),\mathfrak{a}')$. The roots

$$\omega_i - \omega_j,\ \ i<j; \quad \omega_i + \omega_j,\ \ i<j; \ \omega_i; \ \ 2\omega_i$$

can be selected as the positive roots. Then the roots

$$\alpha_1 = \omega_1 - \omega_2,\quad \alpha_2 = \omega_2 - \omega_3,\ \ldots,$$
$$\alpha_{q-1} = \omega_{q-1} - \omega_q,\quad \alpha_q = \omega_q \qquad (12)$$

are the simple roots. The algebra $su(p,q)$ [the semisimple part of $u(p,q)$]is a Lie algebra of type A III (see Ref. 23). Hence the roots $\omega_i - \omega_j$, $\omega_i + \omega_j$, $-\omega_i - \omega_j$, $i \neq j$, have the multiplicity 2. The roots ω_i and $-\omega_i$ have the multiplicity $2(p-q)$. The multiplicity of the roots $2\omega_i$ and $-2\omega_i$ is 1. The root system of the restricted roots of $u(p,q)$ is similar to the system of roots of the complex Lie algebra B_n (see, Sec. 6,

TABLE II.

root	multiplicity	basis of root subspace
$\omega_i - \omega_j,\ i \neq j$	2	$E_{p-i+1,p-j+1} - E_{p+j,p+i},\ \sqrt{-1}(E_{p-i+1,p-j+1} + E_{p+j,p+i})$
$\omega_i + \omega_j,\ i \neq j$	2	$E_{p-i+1,p+j} - E_{p-j+1,p+i},\ \sqrt{-1}(E_{p-i+1,p+j} + E_{p-j+1,p+i})$
$-\omega_i - \omega_j,\ i \neq j$	2	$E_{p+j,p-i+1} - E_{p+i,p-j+1},\ \sqrt{-1}(E_{p+j,p-i+1} + E_{p+i,p-j+1})$
ω_i	$2(p-q)$	$E_{p-i+1,k} - E_{k,p+i},\ \sqrt{-1}(E_{p-i+1,k} + E_{k,p+i}), k = 1,2,\ldots,p-q$
$-\omega_i$	$2(p-q)$	$E_{k,p-i+1} - E_{p+i,k},\ \sqrt{-1}(E_{k,p-i+1} + E_{p+i,k}), k = 1,2,\ldots,p-q$
$2\omega_i$	1	$\sqrt{-1}E_{p-i+1,p+i}$
$-2\omega_i$	1	$\sqrt{-1}E_{p+i,p-i+1}$

Chap. IV of Ref. 22). Due to this similarity it turns out to be easy to find the root spaces of u(p,q).

In Table II the basis elements for the root spaces $\mathfrak{g}_\lambda$ are listed, with respect to the subalgebra $\mathfrak{a}'$. The particular choice of the basis elements as linear combinations of elements E_{ij}, instead of the most obvious choice of the elements E_{ij} themselves, is motivated by the fact that the matrices $\varphi^{-1}E_{ij}$ do not belong to u(p,q), while $\varphi^{-1} = \varphi$ acting on the basis elements of Table II transforms them into basis elements (root vectors) of the subalgebra $\mathfrak{a}$ of u(p,q).

The basis elements of $\mathfrak{a}$ are then obtained by means of the transformation φ from the basis elements of Table II as

$$\begin{aligned}
e_{\omega_i-\omega_j} &= E_{p-i+1,p-j+1} + E_{p-i+1,p+j} + E_{p+i,p-j+1} \\
&\quad + E_{p+i,p+j} - E_{p-j+1,p-i+1} + E_{p-j+1,p+i} \\
&\quad + E_{p+j,p-i+1} - E_{p+j,p+i}, \\
e_{\omega_i-\omega_j} &= \sqrt{-1}(E_{p-i+1,p-j+1} + E_{p-i+1,p+j} \\
&\quad + E_{p+i,p-j+1} + E_{p+i,p+j} + E_{p-j+1,p-i+1} \\
&\quad - E_{p-j+1,p+i} - E_{p+j,p-i+1} + E_{p+j,p+i}), \\
e_{\omega_i+\omega_j} &= E_{p-i+1,p-j+1} - E_{p-i+1,p+j} + E_{p+i,p-j+1} \\
&\quad - E_{p+i,p+j} - E_{p-j+1,p-i+1} + E_{p-j+1,p+i} \\
&\quad - E_{p+j,p-i+1} + E_{p-j,p-i}, \\
e_{\omega_i+\omega_j} &= \sqrt{-1}(E_{p-i+1,p-j+1} - E_{p-i+1,p+j} \\
&\quad + E_{p+i,p-j+1} - E_{p+i,p+j} + E_{p-j+1,p-i+1} \\
&\quad - E_{p-j+1,p+i} + E_{p+j,p-i+1} - E_{p+j,p+i}), \\
e_{-\omega_i-\omega_j} &= E_{p-j+1,p-i+1} + E_{p-j+1,p+i} - E_{p+j,p-i+1} \\
&\quad - E_{p+j,p+i} - E_{p-i+1,p-j+1} - E_{p-i+1,p+j} \\
&\quad + E_{p+i,p-j+1} + E_{p+i,p+j}, \\
e_{-\omega_i-\omega_j} &= \sqrt{-1}(E_{p-j+1,p-i+1} + E_{p-j+1,p+i} \\
&\quad - E_{p+j,p-i+1} - E_{p+j,p+i} + E_{p-i+1,p-j+1} \\
&\quad + E_{p-i+1,p-j} - E_{p+i,p-j+1} - E_{p+i,p+j}), \\
e_{\omega_i} &= E_{p-i+1,k} + E_{p+i,k} - E_{k,p-i+1} + E_{k,p+i}, \\
k &= 1,2,\ldots,q, \\
e_{\omega_i} &= \sqrt{-1}(E_{p-i+1,k} + E_{p+i,k} + E_{k,p-i+1} - E_{k,p+i}), \\
k &= 1,2,\ldots,q, \\
e_{-\omega_i} &= E_{p-i+1,k} - E_{p+i,k} - E_{k,p-i+1} - E_{k,p+i}, \\
k &= 1,2,\ldots,q, \\
e_{-\omega_i} &= \sqrt{-1}(E_{p-i+1,k} - E_{p+i,k} + E_{k,p-i+1} + E_{k,p+i}), \\
k &= 1,2,\ldots,q, \\
e_{2\omega_i} &= \sqrt{-1}(E_{p-i+1,p-i+1} - E_{p+i,p+i} + E_{p+i,p-i+1} \\
&\quad - E_{p-i+1,p+i}), \\
e_{-2\omega_i} &= \sqrt{-1}(E_{p-i+1,p-i+1} - E_{p+i,p+i} \\
&\quad + E_{p-i+1,p+i} - E_{p+i,p-i+1}).
\end{aligned}$$

The elements $\mathfrak{h}_i' = \Sigma_{j=1}^{i}\tilde{\mathfrak{h}}_j, i = 1,2,\ldots,q$, have the property $\alpha_j(\mathfrak{h}_i') = \delta_{ij}$. The elements $\mathfrak{H}_i' = (1/\sqrt{2})\tilde{\mathfrak{h}}_i$, $i = 1,2,\ldots,q$, form an orthonormal basis for $\mathfrak{a}'$. The transformation φ^{-1} transforms the elements $\mathfrak{h}_i'$ into the elements $\mathfrak{h}_i$ of $\mathfrak{a}$, and the elements $\mathfrak{H}_i'$ into the elements $\mathfrak{H}_i$ of $\mathfrak{a}$, such that

$$\mathfrak{h}_i = \sum_{j=1}^{i} \tilde{\mathfrak{h}}_j', \quad \mathfrak{H}_i = \frac{1}{\sqrt{2}}\tilde{\mathfrak{h}}_i', \tag{13}$$

with $\tilde{\mathfrak{h}}_i' = E_{p-i+1,p+i} + E_{p+i,p-i+1}$.

So far the roots $\alpha_i, \alpha_{i+1},\ldots,\alpha_q$ for the sets F_i (defined in Sec. 2) have not yet been specified. They are now chosen to be the simple roots of Eq. (12).

All the information needed for the evaluation of the subgroups K_i^j is now available. It follows from the description of the root spaces given above that the sequence, Eq. (2), of subgroups K_i of the group U(p,q) consists of the subgroups $K_i \cong \mathrm{U}(p-i+1)\times\mathrm{U}(q-i+1)\times[\mathrm{U}(1)\times\cdots\times\mathrm{U}(1)]$ i.e., (there are $i-1$ terms in the square brackets). Moreover, for the sequence of subgroups given by Eq. (3), the following holds,

$$K_i \equiv K_{i+1}^1 \supset K_{i+1}^2 \supset K_{i+1} \equiv K_{i+2}^1,$$

i.e., there exists only one subgoups between K_i and K_{i+1}. This subgroup is given as

$$K_{i+1}^2 \cong \mathrm{U}(p-i)\times\mathrm{U}(q-i)\times[\underbrace{\mathrm{U}(1)\times\cdots\times\mathrm{U}(1)}_{(i+1)}].$$

The subgroup $K_i \equiv K_{i+1}^1$ consists of the matrices

$$\left(\begin{array}{c|c|c|c}
\mathrm{U}(p-i+1) & 0 & 0 & 0 \\ \hline
0 & \begin{matrix} u_{i-1} & \\ & \ddots \\ & & u_1 \end{matrix} & 0 & 0 \\ \hline
0 & 0 & \begin{matrix} u_1 & \\ & \ddots \\ & & u_{i-1} \end{matrix} & 0 \\ \hline
0 & 0 & 0 & U(q-i+1)
\end{array}\right), \quad u_j \in \mathrm{U}(1). \tag{14}$$

The subgroup K^2_{i+1} consists of the matrices

$$\left(\begin{array}{c|ccc|ccc|c}
U(p-i) & & 0 & & & 0 & & 0 \\
\hline
 & u_i & & & & & & \\
 & & u_{i-1} & & & & & \\
 & & & \ddots & & & & \\
0 & & & u_1 & & 0 & & 0 \\
\hline
 & & & & u_1 & & & \\
 & & & & & \ddots & & \\
 & & & & & & u_{i-1} & \\
0 & & 0 & & & & u_i' & 0 \\
\hline
0 & & 0 & & & 0 & & U(q-i)
\end{array}\right), \quad \begin{array}{l} u_j \in U(1), \\ u_i' \in U(1) \end{array} \tag{15}$$

Thus the following sequence of subgroups has been obtained,

$$K \equiv K_1 \equiv K^1_2 \supset K^2_2 \supset K_2 \equiv K^1_3 \supset K^2_3 \supset \cdots \supset K_q$$
$$\equiv K^1_{q+1} \supset K^2_{q+1} \supset M. \tag{16}$$

This sequence of subgroups will be of great importance in what follows.

4. BASIS FOR THE SPACE OF PRINCIPAL NONUNITARY SERIES REPRESENTATIONS

In this section a convenient basis will be constructed for the space $L^2_\delta(K,V)$ of the principal nonunitary series representations of U(p,q).

The restriction of the representation $\pi_{\delta,\Lambda}$ to the subgroup K acts on the space $L^2_\delta(K,V)$ as the right regular representation. The representation of K on the space $L^2_\delta(K,V)$, which is obtained by restricting $\pi_{\delta,\Lambda}$ of the subgroup K, divides the space $L^2_\delta(K,V)$ into an orthogonal sum of subspaces. Each of the subspaces carries a representation of K which is a direct sum of identical irreducible representation of K. These subspaces are denoted by $L^\lambda_\delta(K,V)$, where λ is the index of the irreducible representation of K. Hence, $L^2_\delta(K,V)$ consists of all measurable square integrable vector functions f which transfor according to the representation λ of K, if f is acted upon by elements of K from the right, and for which holds $f(mk) = \delta(m) f(k)$. It is clear that the functions of $L^\lambda_\delta(K,V)$, if acted upon from the left by the elements of K, transform also according to the representation λ of K. Under action from the left by the elements of K^j_i, *some* of the functions of $L^\lambda_\delta(K,V)$ will transform according to a given irreducible representation of K^j_i. The space $L^\lambda_\delta(K,V)$ can thus be represented as an orthogonal sum of subspaces, each of which consists of functions which transform under action from the left by the elements of a subgroup of Eq. (16) according to a definite irreducible representation of this subgroup [to each subspace of $L^\lambda_\delta(K,V)$ corresponds a fixed sequence of irreducible representations]. If a function f of $L^\lambda_\delta(K,V)$ transforms according to the irreducible representation μ of K^j_i, if acted upon from the left by the elements of K^j_i, then the representation μ is contained in the representation $\lambda\,|K^j_i$ (restriction of λ to K^j_i) and $\mu|M$ contains the representation δ of M. If, in addition, the function transforms according to the irreducible representation μ' of $K^{j'}_{i'}$, if acted upon from the left by the elements of $K^{j'}_{i'}$, and if $K^{j'}_{i'} \subset K^j_i$, then the representation $\mu|K^{j'}_{i'}$ contains the representation μ'.

Corresponding to the sequence of subgroups, Eq. (16), the sequence Γ of their irreducible representations is then given as

$$\lambda \equiv \lambda^1_2, \lambda^2_2, \lambda^1_3, \lambda^2_3, \ldots, \lambda^1_{q+1}, \lambda^2_{q+1}, \delta, \tag{17}$$

such that $\lambda^j_i\,|K^{j'}_{i'}$ contains $\lambda^{j'}_{i'}$, if $K^j_i \supset K^{j'}_{i'}$.

Consider the subset of all vector functions f of the space $L^\lambda_\delta(K,V)$ which under left action by the elements of the groups of sequence, Eq. (16), transform according to a *given* sequence of irreducible representations, Eq. (17). From the definition of the space $L^\lambda_\delta(K,V)$ follows that this subspace is not empty. This subspace is denoted by

$$L(\Gamma) \equiv L(\lambda, \lambda^2_2, \lambda^1_3, \lambda^2_3, \ldots, \lambda^1_{q+1}, \lambda^q_{2+1}, \delta).$$

Functions f of different subspaces are orthogonal. Moreover, the space $L^\lambda_\delta(K,V)$ is the orthogonal sum of the subspaces $L(\Gamma)$. Therefore,

$$L^2_\delta(K,V) = \sum_\Gamma L(\Gamma) \quad \text{(orthogonal sum)}. \tag{18}$$

The representation $\pi_{\delta,\Lambda}\,|K$ realizes on each of the subspaces $L(\Gamma)$ the *irreducible* representation λ of K if $\Gamma = (\lambda, \cdots)$. The irreducibility of $\pi_{\delta,\Lambda}\,|K$ on $L(\Gamma)$ follows by comparing the multiplicity of a given irreducible representation λ of K in $\pi_{\delta,\Lambda}\,|K$ with the number of subspaces $L(\Gamma)$ with $\Gamma = (\lambda, \cdots)$.

The subgroups of Eq. (16) have the property that any irreducible representation of a group of this chain is contained in the irreducible representation preceding it with not more than unit multiplicity. It follows that the number of subspaces $L(\Gamma)$ for given λ must be equal to the number of irreducible representations λ in $\pi_{\delta,\Lambda}|K$.

The decomposition Eq. (18) is therefore a decomposition of the space $L^2_\delta(K,H)$ into an orthogonal sum of K-irreducible subspaces. A basis for the subspaces $L(\Gamma)$ is constructed as follows.

Let $|\delta,s\rangle, s = 1,2,...,\dim\delta$, denote an orthonormal basis for the space of the representation δ of M. Let $|\lambda,t\rangle$ $t = 1,2,..., \dim\lambda$, denote any orthonormal basis for the space of the representation λ of K. Let $|\lambda,\Gamma,\delta,s,\rangle$, $\Gamma = (\lambda, \lambda^2_2,...,\delta)$, denote an orthonormal basis of the space of the representation λ of K, constructed according to the reduction of the sugroups of Eq. (16). Let

$$D^\lambda_{(\lambda,\Gamma,\delta,s),t}(k) = \sqrt{\dim\lambda}\ \langle\lambda,\Gamma,\delta,s|\lambda(k)|\lambda,t\rangle, \quad k\in K.$$

Then the vector functions with vector index s, $s = 1,2,...,\dim\delta$,

$$\{D^\lambda_{(\lambda,\Gamma,\delta,s),t}(k)\}^{\dim\delta}_{s=1} \equiv D^\lambda_{(\lambda,\Gamma,\delta),t}(k), \quad t = 1,2,...,\dim\lambda, \tag{19}$$

for different t, form an orthonormal basis of $L(\Gamma)$ (for fixed λ,Γ,δ). The union of all basis, Eq. (19), for all $L(\Gamma)$, forms an orthonormal basis for the space $L^2_\delta(K,V)$ and for the space $dL^2_\delta(K,V)$.

In the following, the sequence of representations, Eq. (17), which correspond to the sequence of groups, Eq. (16), are considered. This sequence of representations has the property that each representation is contained in the representation preceding it. The representation δ of the subgroup $M \cong U(p-q)\times[U(1)\times\cdots\times U(1)]$ (for q terms in [ ]) is given by the product of the representation of $U(p-q)$ with highest weight

$$(m_{1,p-q}, m_{2,p-q},...,m_{p-q,p-q}),$$
$$m_{1,p-q} \geqslant m_{2,p-q} \geqslant \cdots \geqslant m_{p-q,p-q}, \quad m_{ik} \text{ integers},$$

and the representation (character) of the group $U(1)\times\cdots\times U(1)$, given by the set of integers $\Lambda_1,\Lambda_2,...,\Lambda_q$ [Λ_i is the character of that group $U(1)$ of M which corresponds in the matrix equation (10) to the parameter u_i].

The irreducible representation λ^1_{i+1} of the subgroup

$$K_i \equiv K^1_{i+1} \cong U(p-i+1)\times U(q-i+1) \times [U(1)\underset{(i-1)}{\times\cdots\times} U(1)],$$

Eq. (14), is given by the product of the representation of $U(p-i+1)$ with highest weight

$$(m_{1,p-i+1}, m_{2,p-i+1},...,m_{p-i+1,p-i+1}),$$
$$m_{1,p-i+1} \geqslant m_{2,p-i+1} \geqslant \cdots \geqslant m_{p-i+1,p-i+1},$$
m_{ik} integers,

the representation of $U(q-i+1)$ with the highest weight

$$(m_{1,q-i+1}, m_{2,q-i+1},...,m_{q-i+1,q-i+1}),$$
$$m_{1,q-i+1} \geqslant m_{2,q-i+1} \geqslant \cdots \geqslant m_{q-i+1,q-i+1},$$
m_{ik} integers,

and the representation of $U(1)\times\cdots\times U(1)$ with weight

$$\Lambda'_1, \Lambda'_2,...,\Lambda'_{i-1},$$

whereby the Λ' are integers. If this representation is to contain the representation δ of the subgroup M, then $\Lambda'_j = \Lambda_j$, $j = 1,2,...,i-1$, must hold.

The representation λ^2_{i+1} of the subgroup

$$K^2_{i+1} \cong U(p-i)\times U(q-i)\times[U(1)\underset{(i-1)}{\times\cdots\times} U(1)] \times U(1)\times U(1)$$

[the last two subgroups U(1) correspond to the parameters u_i and u'_i in the matrix of Eq. (15)] is given by the product of the representation of the group $U(p-i)$ with highest weight

$$(m_{1,p-i}, m_{2,p-i},...,m_{p-i,p-i}),$$
$$m_{1,p-i} \geqslant m_{2,p-i} \geqslant \cdots \geqslant m_{p-i,p-i}, \quad m_{ik} \text{ integers},$$

the representation of the group $U(q-i)$ with highest weight

$$(m_{1,q-i}, m_{2,q-i},...,m_{q-i,q-i}),$$
$$m_{1,q-i} \geqslant m_{2,q-i} \geqslant \cdots \geqslant m_{q-i,q-i}, \quad m_{ik} \text{ integers},$$

the representation of the group $U(1)\times\cdots\times U(1)$, given by the integers

$$\Lambda'_1, \Lambda'_2,...,\Lambda'_{i-1}$$

and by the representation of the group $U(1)\times U(1)$ which is given by the two integers ω^i_1, ω^i_2. If this representation of K^2_{i+1} is to contain the representation δ of the group M, then it must hold that $\Lambda'_j = \Lambda_j$, $j = 1,2,...,i-1$ and $\omega^i_1 + \omega^i_2 = \Lambda_i$.

If, moreover the representation λ^1_{i+1} of K^1_{i+1}, is to contain the representation λ^2_{i+1} of K^2_{i+1}, then the following conditions must hold:

$$\omega^i_1 = \sum_{j=1}^{p-i+1} m_{j,p-i+1} - \sum_{j=1}^{p-i} m_{j,p-i},$$
$$\omega^i_2 = \sum_{j=1}^{q-i+1} m_{j,q-i+1} - \sum_{j=1}^{q-i} m_{j,q-i},$$

i.e., the integers ω_1 and ω_2 are defined by the integers m_{js} which define the repersentations of the groups $U(s)$.

Therefore, the sequence of representations Eq. (17), which corresponds to the sequence of groups Eq. (16), is completely defined by the representations of the subgroups $U(s)$ and the representation δ of M. Thus, the scheme

$$\Omega = \begin{pmatrix} m_{1p} & m_{2p} & \cdots\cdots & \cdots\cdots & \cdots\cdots & m_{pp} & m'_{1q} & m'_{2q} & \cdots\cdots & \cdots\cdots & \cdots\cdots & m'_{qq} \\ & m_{1,p-1} & \cdots\cdots & \cdots\cdots & \cdots\cdots & m_{p-1,p-1} & & m'_{1,q-1} & \cdots & \cdots\cdots & m'_{q-1,q-1} & \\ & & \cdots\cdots & \cdots\cdots & \cdots\cdots & & & & & \cdots\cdots & & \\ & m_{1,p-q+1} & \cdots\cdots & \cdots\cdots & \cdots\cdots & m_{p-q+1,p-q+1} & & & & m'_{1,1} & & \\ & & m_{1,p-q} & \cdots\cdots & m_{p-q,p-q} & & & & & \cdot & & \end{pmatrix} \tag{20}$$

corresponds to the sequence of representations given by Eq. (17). The integers m_{ik} in the scheme Eq. (20) satisfy the conditions

$$m_{ik} \geqslant m_{i,k-1} \geqslant m_{i+1,k}, \quad m'_{ik} \geqslant m'_{i,k-1} \geqslant m'_{i+1,k} \tag{21}$$

and the conditions

$$\omega_1^i + \omega_2^i = \sum_{j=1}^{p-i+1} m_{j,p-i+1} - \sum_{j=1}^{p-i} m_{j,p-i} + \sum_{j=1}^{q-i+1} m_{j,q-i+1} - \sum_{j=1}^{q-i} m_{j,q-i} = \Lambda_i. \tag{22}$$

The scheme Ω, together with the representation δ of M, defines completely the representations of Eq. (17).

We can therefore draw the conclusion that all possible schemes Ω, for which the conditions Eq. (21) and Eq. (22) are satisfied for fixed δ, label uniquely the subspaces $L(\Gamma)$ of the space $L^2_\delta(K,V)$ on which irreducible representations of K are realized. The double Gel'fand–Zetlin schemes

$$\Sigma = \begin{pmatrix} m_{1p} & & \cdots & & m_{pp} & m'_{1q} & & \cdots & \cdots & \cdots & m'_{qq} \\ & \bar{m}_{1,p-1} & \cdots & \bar{m}_{p-1,p-1} & & & \bar{m}'_{1,q-1} & \cdots & \bar{m}'_{q-1,q-1} & & \\ & & \cdots & & & & & \cdots & & & \\ & & \bar{m}_{1,1} & & & & & \bar{m}'_{1,1} & & & \end{pmatrix} \tag{23}$$

label the basis elements in the K-irreducible subspaces $L(\Gamma)$. Therefore, the orthonormal basis elements of the space $L^2_\delta(K,V)$, given by Eq. (19), are uniquely labeled by the two schemes Ω and Σ. The label t corresponds to the scheme Σ, while the label (λ,Γ,δ) corresponds to the scheme Ω. The basis elements of the space $L^2_\delta(K,V)$ will, in the following, be denoted by the schemes

$$\begin{bmatrix} \Omega \\ \Sigma \end{bmatrix} \equiv [\Omega\,|\Sigma\,].$$

It should be noted that the first line of Ω and Σ are identical. This line gives the highest weight of the irreducible representation λ of K.

In the following the scheme $[\Omega\,|\Sigma\,]$ will be used to represent the *concrete vector functions* given by Eq. (19).

A remark should be made at this point. The basis used above for the K-irreducible subspaces $L(\Gamma)$ of $L^2_\delta(K,V)$ was the Gel'fand–Zetlin basis. That implies that in the following the infinitesimal generators of U(p,q) and U$((p+q)$ will be obtained in this basis.

One could, however, choose any other basis for the K-irreducible subspaces $L(\Gamma)$ of $L^2_\delta(K,V)$ and thus obtain the infinitesimal operators in another basis. In what follows the Gel'fand–Zetlin basis will be used for the K-irreducible subspaces. It will be shown, however, what changes have to be made if any other basis for the K-invariant subspaces $L(\Gamma)$ is to be used.

5. INFINITESIMAL GENERATORS OF THE PRINCIPAL NONUNITARY SERIES REPRESENTATIONS OF U(p,q) IN THE $[\Omega|\Sigma]$ BASIS

In order to obtain the matrix elements of the infinitesimal generators of the group U(p,q) in the $\pi_{\delta,\Lambda}$ representation with respect to the $[\Omega\,|\Sigma\,]$ basis in explicit form, the basic lemma introduced in Sec. 2 will be utilized. Before this can be done, however, the formula Eq. (6) has to be brought into a more suitable form.

Let Λ be a linear form defined on the subalgebra $\mathfrak{a}'$ (or $\mathfrak{a}$). Let α_i denote the simple restricted roots. Then the formula $(\alpha_i, \Lambda) = \Lambda(\mathfrak{h}_{\alpha_i})$ represents the roots α_i as elements $\mathfrak{h}_{\alpha_i}$ of $\mathfrak{a}'$ (or $\mathfrak{a}$). A direct evaluation shows that the simple restricted roots α_i correspond to the following elements of $\mathfrak{a}$,

$$\alpha_i \to \mathfrak{h}_{\alpha_i} = \tfrac{1}{2}(\tilde{\mathfrak{h}}'_i - \tilde{\mathfrak{h}}'_{i+1}), \quad i = 1,2,\ldots,q-1, \quad \alpha_q \to \mathfrak{h}_{\alpha_q} = \tfrac{1}{2}\tilde{\mathfrak{h}}'_q, \tag{24}$$

where, as above, $\tilde{\mathfrak{h}}'_i = E_{p-i+1,p+i} + E_{p+i,p-i+1}$.

Since $\mathfrak{H}_i = (1/\sqrt{2})\tilde{\mathfrak{h}}'_i$, Eq. (13), the first term on the right-hand side of Eq. (6) can be rewritten, as

$$\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{H}_i\rangle\Lambda(\mathfrak{H}_i) = \tfrac{1}{2}\langle(\mathrm{ad}k)\mathfrak{Y},\tilde{\mathfrak{h}}'_i\rangle\Lambda(\tilde{\mathfrak{h}}'_i). \tag{25}$$

With the $\mathfrak{h}_i$ defined in Eq. (13) it moreover follows that

$$\begin{aligned} &\sum_{i=1}^{q} \langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle\,\langle\Lambda,\alpha_i\rangle \\ &= \tfrac{1}{2}\sum_{i=1}^{q-1}\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle\Lambda(\tilde{\mathfrak{h}}'_i - \tilde{\mathfrak{h}}'_{i+1}) + \tfrac{1}{2}\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_q\rangle\Lambda(\tilde{\mathfrak{h}}'_q) \\ &= \tfrac{1}{2}\sum_{i=1}^{q}\sum_{j=1}^{i}\langle(\mathrm{ad}k)\mathfrak{Y},\tilde{\mathfrak{h}}'_j\rangle\Lambda(\tilde{\mathfrak{h}}'_i) - \tfrac{1}{2}\sum_{i=1}^{q-1}\sum_{j=1}^{i}\langle(\mathrm{ad}k)\mathfrak{Y},\tilde{\mathfrak{h}}'_j\rangle \\ &\quad\times\Lambda(\tilde{\mathfrak{h}}'_{i+1}) = \tfrac{1}{2}\sum_{i=1}^{q}\langle(\mathrm{ad}k)\mathfrak{Y},\tilde{\mathfrak{h}}'_i\rangle\Lambda(\tilde{\mathfrak{h}}'_i). \end{aligned} \tag{26}$$

Now, the term $\langle(\mathrm{ad}k)\mathfrak{Y},\rho\rangle$ of Eq. (6) is considered. The element ρ of $\mathfrak{a}$ can be expressed as (Ref. 20).

$$\rho = \tfrac{1}{2}\sum_{i=1}^{q}(p_i + 2q_i)\langle\alpha_i,\alpha_i\rangle\mathfrak{h}_i, \tag{27}$$

where p_i is the multiplicity of the simple restricted root α_i, q_i the multiplicity for the restricted root $2\alpha_i$ and $\langle\alpha_i,\alpha_i\rangle = \alpha_i(\mathfrak{h}_{\alpha_i}) = \langle\mathfrak{h}_{\alpha_i},\mathfrak{h}_{\alpha_i}\rangle$. Equation (27) is obtained as follows: A reflection S_{α_i} of the Weyl group of the system of restricted roots of the pair $(u(p,q),\mathfrak{a})$ maps the set of roots Δ, without the roots α_i and $2\alpha_i$, onto itself. The roots α_i and $2\alpha_i$ are mapped by S_{α_i} onto the roots $-\alpha_i$ and $-2\alpha_i$. Thus

$$\begin{aligned}\langle 2R,\alpha_i\rangle &= \langle R_1,\alpha_i\rangle - \langle R,S_{\alpha_i}\rangle \\ &= \langle R - S_{\alpha_i}R,\alpha_i\rangle = (p_i+2q_i)\langle\alpha_i,\alpha_i\rangle,\end{aligned} \tag{28}$$

where R is defined by the relation $R(\mathfrak{H}) = \langle\rho,\mathfrak{H}\rangle, \mathfrak{H}\in\mathfrak{a}$. Equation (28) however implies

$$\alpha_i(\rho) = \tfrac{1}{2}(p_i+2q_i)\langle\alpha_i,\alpha_i\rangle.$$

This proves Eq. (27), since $\alpha_i(\mathfrak{h}_j) = \delta_{ij}$.

Equations (25), (26), and (27) show that the first two terms of the right-hand side of Eq. (6) can be rewrittten as

$$\begin{aligned}&\sum_{i=1}^{q}\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{H}_i\rangle\Lambda(\mathfrak{H}_i)f(k) - \langle(\mathrm{ad}k)\mathfrak{Y},\rho\rangle f(k) \\ &\quad = \sum_{i=1}^{q}[\langle\Lambda,\alpha_i\rangle - \tfrac{1}{2}(p_i+2q_i)\langle\alpha_i,\alpha_i\rangle]\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle f(k),\end{aligned} \tag{29}$$

where $\langle\Lambda,\alpha_i\rangle = \Lambda(\mathfrak{h}_{\alpha_i})$.

In order to bring the third term of Eq. (6) into a more suitable form, this term is evaluated by letting it act on the *basis elements* $[\Omega\,|\Sigma\,]$ of the space $L^2_\delta(K,V)$. Since with respect to action from the left by elements of the subgroups K^j_i the vector functions $[\Omega\,|\Sigma\,]$ transforms according to irreducible representations of these groups, it holds

$$Q_i[\Omega\,|\Sigma\,] = q_i[\Omega\,|\Sigma\,], \tag{30}$$

where q_i is a number. The eigenvalue q_i will be evaluated later on. Due to Eq. (30) it follows that

$$\begin{aligned}&\sum_{i=1}^{q}[Q_i,\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle][\Omega\,|\Sigma\,] \\ &\quad = \sum_{i=1}^{q}(Q_i-q_i)\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle[\Omega\,|\Sigma\,].\end{aligned} \tag{31}$$

Hence, for the case that $f(k) = [\Omega\,|\Sigma]$ the formula given by Eq. (6) can be rewritten as

$$\begin{aligned}&d\pi_{\delta,\Lambda}(\mathfrak{Y})[\Omega\,|\Sigma\,] \\ &\quad = \sum_{i=1}^{q}[\langle\Lambda,\alpha_i\rangle - \tfrac{1}{2}(p_i+2q_i)\langle\alpha_i,\alpha_i\rangle + \tfrac{1}{2}(Q_i-q_i)] \\ &\qquad \times\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle[\Omega\,|\Sigma\,].\end{aligned} \tag{32}$$

In order to obtain an explicit expression for the action of the infinitesimal generators of the representation $\pi_{\delta,\Lambda}$ onto the basis elements $[\Omega\,|\Sigma]$, the functions $\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle[\Omega\,|\Sigma\,]$ will be expanded in terms of the basis elements $[\Omega'|\Sigma']$. In order to do this, the functions $\langle(\mathrm{ad}k)\mathfrak{Y},\mathfrak{h}_i\rangle$ is first studied.

It is not necessary to evaluate Eq. (32) for arbitrary elements $\mathfrak{Y}\in\mathfrak{p}_c$. It is sufficient to evaluate Eq. (32) for only those elements which form an orthonormal basis for $\mathfrak{p}_c$ with respect to the scalar product $\langle\cdot,\cdot\rangle$. The matrices E_{ij}, $i\leqslant p, j>p$ and $E_{ij}, i>p, j\leqslant p$ form an orthonormal basis for $\mathfrak{p}_c$. The operators $d\pi_{,\delta\Lambda}(E_{ij})$, $i\leqslant p, j>p$, form a tensor operator which transforms according to the adjoint representation $(\mathrm{ad}k)$ of the group K. More precisely, this tensor operator transforms according to the tensor product of the vector representation, denoted by $\{\bar{1}\}_p$, of the subgroup U(p), and the representation contragredient to the vector representation, denoted by $\{\bar{1}\}_q$, of the subgroup U(q) of K. The operators $d\pi_{\delta,\Lambda}(E_{ij}), i>p, j\leqslant p$, form a tensor operator with respect to the adjoint representation $(\mathrm{ad}k)$ of K which transforms like the tensor product of the representation $\{\bar{1}\}_p$ of U(p) and the representation $\{1\}_q$ of U(q). As these tensor operators transform like the states of the representations $\{1\}_p\times\{\bar{1}\}_q$ and $\{\bar{1}\}_p\times\{1\}_q$ of the group $K\cong$ U$(p)\times$U(q), they can be uniquely labeled by means of the Gel'fand–Zetlin basis of these representations. The one to one correspondence between the operators $d\pi_{\delta,\Lambda}(E_{ij}), i\leqslant p, j>p$, and the Gel'fand–Zetlin basis for the representation $\{1\}_p\times\{\bar{1}\}_q$ is given by

$$E_{ij} \longleftrightarrow \left.\begin{matrix}p-i+1\\ \text{rows}\end{matrix}\right\{ \left[\begin{array}{cccccc} 1 \quad 0 \quad 0 & & 0 \quad 0 & & 0 \quad 0 \quad -1 \\ 1 \quad 0 & & 0 \quad 0 & & 0 \quad -1 \\ \cdots\cdots & & & & \cdots\cdots\cdots & \\ 1 \quad 0 \;.\;.\;.\; 0 & & & & 0 \;.\;.\;.\; 0 \quad -1 & \\ 0 \;.\;.\;.\; 0 & & & & 0 \;.\;.\;.\; 0 & \\ \cdots & & & & \cdots & \\ 0 & & & & 0 & \end{array}\right] \left.\right\} \begin{matrix}j-p\\ \text{rows}\end{matrix} \tag{33}$$

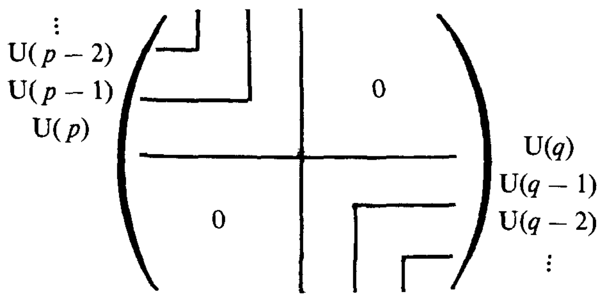
E_{ij}

For the $d\pi_{\delta,\Lambda}(E_{ij})$, $i>p, j\leqslant$p, the correspondence is

$$E_{ij} \longleftrightarrow \left.\begin{matrix}p-j+1\\ \text{rows}\end{matrix}\right\{ \left[\begin{array}{cc} 0 \;.\;.\;.\;.\;.\;.\;.\;.\; 0 \quad 0 \quad -1 & 1 \quad 0 \quad 0 \;.\;.\;.\;.\;.\;.\;.\;.\; 0 \\ 0 \;.\;.\;.\;.\;.\;.\;.\;.\; 0 \;.\; -1 & 1 \quad 0 \;.\;.\;.\;.\;.\;.\;.\; 0 \\ \cdots\cdots\cdots & \cdots\cdots\cdots \\ 0 \;.\;.\;.\; 0 \quad -1 & 1 \quad 0 \;.\;.\;.\; 0 \\ 0 \;.\;.\;.\; 0 & 0 \;.\;.\;.\; 0 \\ \cdots & \cdots \\ 0 & 0 \end{array}\right] \left.\right\} \begin{matrix}i-p\\ \text{rows}\end{matrix} \tag{34}$$

For reasons of simplicity the schemes (33) and (34) will be denoted by (E_{ij}). The schemes (33) and (34) correspond to the chain of subgroups $U(p)\times U(q)\supset U(p-1)\times U(q-1)$

$\supset U(p-2)\times U(q-2)\cdots$ such that



In the following the expression $\langle(\mathrm{ad}k)E_{ij},\mathfrak{h}_k\rangle$ will be evaluated for $i\leqslant p, j>p$. Utilizing Eq. (13) it follows

$$\langle(\mathrm{ad}k)E_{ij},\mathfrak{h}_k\rangle=\sum_{s=1}^{k}\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle. \qquad (35)$$

Since $\langle(\mathrm{ad}k)E_{ij},\mathfrak{h}'_s=\langle(\mathrm{ad}k)E_{ij},E_{p-s+1,p+s}\rangle, i\leqslant p, p>j$, the functions

$$\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle,\quad i\leqslant p,\ j>p, \qquad (36)$$

are the matrix elements of the irreducible representation $\{1\}_p\times\{\bar{1}\}_q$ of the group $K\cong U(p)\times U(q)$ with respect to the orthonormal basis (E_{tr}), i.e.,

$$\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle=D^{\{1\}_p\times\{\bar{1}\}_q}_{(E_{p-s+1,p+s}),(E_{ij})}(k). \qquad (37)$$

It should be remembered that (E_{ij}) denotes the scheme Eq. (33).

Similarly one obtains for the case $i>p, j\leqslant p$,

$$\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle=D^{\{\bar{1}\}_p\times\{1\}_q}_{(E_{p+s,p-s+1}),(E_{ij})}(k). \qquad (38)$$

Thus the functions $\langle(\mathrm{ad}k)E_{ij},\mathfrak{h}_k\rangle$ are defined by Eqs. (35), (37), and (38).

The functions of Eq. (37) and Eq. (38) are invariant under left action by the elements of M, since the $\tilde{\mathfrak{h}}'_s$ are elements of the subalgebra $\mathfrak{a}$ of $u(p,q)$.

Now, the expression

$$\langle(\mathrm{ad}k)E_{ij},\mathfrak{h}_k\rangle[\Omega\,|\Sigma\,]$$
$$\equiv\sum_{s=1}^{q}\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle[\Omega\,/\Sigma\,] \qquad (39)$$

will be evaluated. According to Eq. (19)

$$[\Omega\,|\Sigma\,]\equiv\{D^{\lambda}_{(\Gamma,\delta,s),t}(k)\}^{\dim\delta}_{s=1}. \qquad (40)$$

Thus, the sum of the right-hand side of Eq. (39) contains as terms the product of the functions

$$D^{\{1\}_p\times\{\bar{1}\}_q}_{(E_{p-s+1,p+s}),(E_{ij})}(k)D^{\lambda}_{(\Gamma,\delta,s),t}(k) \qquad (41)$$

for $i\leqslant p, j>p$. The function given by Eq. (41) is an element of the space of the tensor product of the representation $\{1\}_q\times\{\bar{1}\}_q$ and the irreducible representation λ of the group K. Thus

Eq. (41)

$$=\sum_{\lambda',\Gamma',t'}\left(\frac{\dim\lambda}{\dim\lambda'}\right)^{1/2}\langle\{1\}_p\times\{\bar{1}\}_q,(E_{p-s+1,p+s});$$
$$\lambda,\Gamma,\delta,s|\lambda',\Gamma',\delta,s\rangle\,|D^{\lambda'}_{(\Gamma',\delta,s),t'}(k)$$
$$\times\ \langle\lambda',t'\,|\{1\}_p\{\bar{1}\}_q(E_{ij});\lambda,t\,\rangle|, \qquad (42)$$

where $\langle;|\rangle$ and $\langle|;\rangle$ are the matrix elements of the operator U and its inverse which decomposes the tensor product of the two representations $\{1\}_p\times\{\bar{1}\}_q$ and λ of K into a direct sum of irreducible representations λ' of K, i.e., the Clebsch–Gordan coefficients (CGC's). The CGC's $\langle\{1\}_p\times\{\bar{1}\}_q,(E_{p-s+1,p+s});\lambda,\Gamma,\delta,s|\lambda',\Gamma',\delta,s\rangle$ do not depend on s. This is seen as follows. The CGC's of the group K can be factored into a product (or into a sum of products) of CGC's of the subgroup M of K and M scalar factors. These M scalar factors depend only on the representation of M, but not on the indices labeling the individual states of the represenation. The vectors $|\{1\}_p\times\{\bar{1}\}_q,(E_{p-s+1,p+s})\rangle$, on the other hand, are invariant with respect to M. Thus, in this factorization of the CGC's of the group K into CGC's of the subgroup M of K and M scalar factors, the CGC of the subgroup M has to be equal to 1. This implies that the CGC's of K are independent of the index s. Therefore, Eq. (42) can be written as

$$D^{\{1\}_p\times\{\bar{1}\}_q}_{(E_{p-s+1,p+s}),(E_{ij})}(k)D^{\lambda}_{(\Gamma,\delta),t}(k)$$
$$=\sum_{\lambda',\Gamma',t'}\left(\frac{\dim\lambda}{\dim\lambda'}\right)^{1/2}\langle;|\rangle D^{\lambda'}_{(\Gamma',\delta),t'}(k)\langle\ |;\rangle. \qquad (43)$$

Equation (43) thus is a relation for the basis vector functions $[\Omega\,|\Sigma]$, while Eq. (42) is a relation for the functions (matrix elements) $D^{\lambda}_{(\Gamma,\delta,s),t}(k)$. With respect to the notation $[\Omega\,|\Sigma]$, Eq. (43) can be reexpressed as

$$\langle(\mathrm{ad}k)E_{ij},\tilde{\mathfrak{h}}'_s\rangle[\Omega\,|\Sigma\,]$$
$$=\sum_{\Omega'\Sigma'}\left(\frac{\dim\lambda}{\dim\lambda'}\right)^{1/2}\langle\{1\}_p\times\{\bar{1}\}_q,(E_{p-s+1,p+s});\Omega\,|\Omega'\rangle$$
$$\times[\Omega'|\Sigma\,]'\langle\Sigma'|\{1\}_p\times\{\bar{1}\}_q,(E_{ij});\Sigma\,\rangle. \qquad (44)$$

It should be noted that in Eq. (43) the index t in the basis element $D^{\lambda}_{(\Gamma,\delta),t}(k)$ corresponds to an *arbitrary* basis for the space of the representation λ of K (i.e., to an arbitrary sequence of subgroups). In Eq. (44), however, the index t labels the states of the representation λ of K with respect to the subgroup chain $K\supseteq U(p)\times U(q)\supset U(p-1)\times U(q-1)\supset U(p-2)\times U(q-2)\supset\cdots$. In the discussion to follow, this particular chain of subgroups is used to label the states of the representation λ of K. The same discussion can be carried out on the basis of Eq. (43), instead of Eq. (44), leading to results of greater generality.

The CGC's of Eq. (44) factor into the product of the CGC's of the subgroups $U(p)$ and $U(q)$. The CGC's needed for these two groups are the CGC's for the direct product of the vector representation $\{1\}$ (and of the contragredient representation $\{\bar{1}\}$) with an arbitrary irreducible representation of these groups. These CGC's are however known, and thus the right-hand side of Eq. (44) can be evaluated in explicit form.[9] In order to do so, some notational conventions have to be made.

For given schemes Ω and Σ the schemes

$$\Omega^{+i_p, i_{p-1}, \dots, i_r}_{-j_q, j_{q-1}, \dots, j_{r'}}, \quad \Sigma^{+i_p, i_{p-1}, \dots, i_r}_{-j_q, j_{q-1}, \dots, j_{r'}}$$

are obtained from the schemes Ω and Σ by adding 1 to each element $m_{i_p p}, m_{i_{p-1}, p-1}, \dots, m_{i_r r}$ and by subtracting 1 from each element $m'_{j_q q}, m'_{j_{q-1}, q-1}, \dots, m'_{j_{r'} r'}$. Similarly, from given schemes Ω and Σ the new schemes

$$\Omega^{-i_p i_{p-1}, \dots, i_r}_{+j_q j_{q-1}, \dots, j_{r'}}, \quad \Sigma^{-i_p i_{p-1}, \dots, i_r}_{+j_q j_{q-1}, \dots, j_{r'}},$$

are obtained by subtracting 1 from each element $m_{i_p p}, m_{i_{p-1}, p-1}, \dots, m_{i_r r}$ and by adding 1 to each element $m'_{j_q q}, m'_{j_{q-1}, q-1}, \dots, m'_{j_{r'} r'}$.

Taking into account the properties of the CGC's, Eq. (44) can be written as

$$\langle (\mathrm{ad}k) E_{ij}, \tilde{\mathfrak{h}}'_s \rangle [\Omega | \Sigma] = \sum_{\substack{i_p, \dots, i_{p-s+1} \\ j_q, \dots, j_{q-s+1} \\ i'_p, \dots, i'_i \\ j'_q, \dots, j'_{p+q-j+1} \\ i_p = i'_p, j_q = j'_q}} \left(\frac{\dim\lambda}{\dim\lambda'} \right)^{1/2} \left[\Omega^{+i_p, \dots, i_{p-s+1}}_{-j_q, \dots, j_{q-s+1}} \middle| \Sigma^{+i'_p, \dots, i'_i}_{-j'_q, \dots, j'_{p+q-j+1}} \right] \langle \{1\}_p \times \{\bar{1}\}_q, (E_{p-s+1, p+s}); \Omega | \Omega' \rangle \langle \Sigma' | \{1\}_p \times \{\bar{1}\}_q (E_{ij}); \Sigma \rangle, \tag{45}$$

where the sum runs from $1, 2, \dots, r$ for each of the indices i_r, i'_r, j_r, j'_r. The symbols Ω' and Σ' in Eq. (45) have been introduced for reasons of simplification. In each term of the sum, Ω' and Σ' are to be set identical to the $\Omega^{\dots}_{\dots}$ and $\Sigma^{\dots}_{\dots}$ of the basis element $[\Omega^{\dots}_{\dots} | \Sigma^{\dots}_{\dots}]$.

For the case E_{ij}, $i > p, j \leqslant p$ one obtains the relation

$$\langle (\mathrm{ad}k) E_{ij}, \tilde{\mathfrak{h}}'_s \rangle [\Omega | \Sigma] = \sum_{\substack{i_p, \dots, i_{p-s+1} \\ j_q, \dots, j_{q-s+1} \\ i'_p, \dots, i'_j \\ j'_q, \dots, j'_{p+q-i+1} \\ i_p = i'_p, j_q = j'_q}} \left(\frac{\dim\lambda}{\dim\lambda'} \right)^{1/2} [\Omega^{-i_p, \dots, i_{p-s+1}}_{+j_q, \dots, j_{q-s+1}} | \Sigma^{-i'_p, \dots, i'_j}_{+j'_q, \dots, j'_{p+q-i+1}}]$$

$$\times \langle \{\bar{1}\}_p \times \{1\}_q, (E_{p+s, p-s+1}); \Omega | \Omega' \rangle \langle \Sigma' | \{\bar{1}\}_p \times \{1\}_q, (E_{ij}); \Sigma \rangle \tag{46}$$

[the symbols Ω' and Σ' replace the symbols $\Omega^{\dots}_{\dots}$ and $\Sigma^{\dots}_{\dots}$ of the basis vector preceding them, as in Eq. (45).] where the sum runs from 1 to r for each of the indices i_r, i'_r, j_r, j'_r.

Thus, according to Eq. (35) and Eq. (45), the formula Eq. (32) can be written for $r \leqslant p, t > p$, as

$$d\pi_{\delta, \Lambda}(E_{rt})[\Omega | \Sigma] = \sum_{i=1}^{q} (\langle \Lambda, \alpha_i \rangle - \tfrac{1}{2}(p_i + 2q_i) \langle \alpha_i, \alpha_i \rangle + \tfrac{1}{2} Q_i - \tfrac{1}{2} q_i)$$

$$\times \sum_{s=1}^{i} \sum_{\substack{i_p, \dots, i_{p-s+1} \\ j_q, \dots, j_{q-s+1} \\ i'_p, \dots, i'_r \\ j'_q, \dots, j'_{p+q-t+1} \\ i_p = i'_p, j_q = j'_q}} \left(\frac{\dim\lambda}{\dim\lambda'} \right)^{1/2} [\Omega^{+i_p, \dots, i_{p-s+1}}_{-j_q, \dots, j_{q-s+1}} | \Sigma^{+i'_p, \dots, i'_r}_{-j'_q, \dots, j'_{p+q-t+1}}]$$

$$\times \langle \{1\}_p \times \{\bar{1}\}_q, (E_{p-s+1, p+s}); \Omega | \Omega' \rangle \langle \Sigma' | \{1\}_p \times \{\bar{1}\}_q, (E_{rt}); \Sigma \rangle. \tag{47}$$

In a similar manner an expression is obtained for $d\pi_{\delta, \Lambda}(E_{r,t})$ for $r > p, t \leqslant p$, utilizing Eq. (46).

The basis functions $[\Omega | \Sigma]$ are eigenfunctions of the operators Q_i with eigenvalues q_i [Eq. (30)]. Hence, Eq. (47) can be rewritten in more explicit form. The value of q_i depends on the state $[\Omega | \Sigma]$ upon which the operator Q_i acts, i.e., $q_i = q_i([\Omega | \Sigma])$. The first factor on the right-hand side of Eq. (47) can thus be written as

$$\langle \Lambda, \alpha_i \rangle - \tfrac{1}{2}(p_i + 2q_i) \langle \alpha_i, \alpha_i \rangle + \tfrac{1}{2} \{ q_i([\Omega' | \Sigma']) - q_i([\Omega | \Sigma]) \}, \tag{48}$$

where $[\Omega'|\Sigma']$ represents the basis vectors of the right-hand side of Eq (47).

From Eq. (24), and utilizing the explicit form of the scalar product in $\mathfrak{a}$, it follows that $\frac{1}{2}(p_i + 2q_i)\langle\alpha_i,\alpha_i\rangle$ equals 1, for $i = 1,2,\ldots,q-1$, and equals $(p-q+1)/2$, for $i = q$.

The term $q_i([\Omega'|\Sigma']) - q_i([\Omega|\Sigma])$ is evaluated as follows. According to Eq. (5), Q_i is a linear combination of the Casimir operators of the subgroups K_i^1 and K_i^2, which act upon the functions $[\Omega|\Sigma]$ from the left side. Acting from the left implies that the q_i depend on Ω, but not on Σ. For an irreducible finite dimensional representation of a semisimple compact Lie group with highest weight Λ', the quadratic Casimir operator, Eq. (4), is given by $\langle\Lambda'+R,\Lambda'+R\rangle - \langle R,R\rangle$, where R is half the sum of positive roots of the group. For the commutative part of the subgroup K_i^j the eigenvalue of the Casimir operator is given by $\langle\Lambda''|\Lambda''\rangle = \Lambda''(h_{\Lambda''}) = \langle h_{\Lambda''},h_{\Lambda''}\rangle$, where Λ'' is that linear form on the Lie algebra of the commutative subgroup which defines the representation (that is, $\langle\Lambda''|\Lambda''\rangle$ is equal to a sum of squares of integers, with the integers defining the representation). Thus the eigenvalue of the Casimir operator ω_i^j of the group K_i^j is given by the number $\langle\Lambda+R,\Lambda+R\rangle - \langle R,R\rangle = \langle\Lambda,\Lambda+2R\rangle$, with $\Lambda = \Lambda' + \Lambda''$.

Assuming now that the function $[\Omega|\Sigma]$ transforms, with respect to left action by the elements of K_i^j, according to the irreducible representation with highest weight Λ_i^j, and that the function $[\Omega'|\Sigma']$ transforms, with respect to left action by the elements of K_i^j, according to the irreducible representation with highest weight $\Lambda_i^j + \tau_i^j$ (thus τ_i^j represents the effect of the action of $d\pi_{\delta,\Lambda}(E_{rt})$ on the representation with highest weight Λ_i^j), then

$$\omega_i^j([\Omega'|\Sigma']) - \omega_i^j([\Omega|\Sigma]) = \langle\Lambda_i^j+\tau_i^j,\Lambda_i^j+\tau_i^j+2R_i^j\rangle - \langle\Lambda_i^j,\Lambda_i^j+2R_i^j\rangle = 2\langle\Lambda_i^j+R_i^j,\tau_i^j\rangle + \langle\tau_i^j,\tau_i^j\rangle. \tag{49}$$

Thus, according to Eq. (5),

$$q_i([\Omega'|\Sigma']) - q_i([\Omega|\Sigma]) = 2\langle\Lambda_i^1+R_i^1,\tau_i^1\rangle + \langle\tau_i^1,\tau_i^1\rangle - \langle\Lambda_i^2+R_i^2,\tau_i^2\rangle - \tfrac{1}{2}\langle\tau_i^2,\tau_i^2\rangle. \tag{50}$$

By means of Eq. (50) and Eq. (48), Eq. (47) can be brought into the following form {note that in Eq. (47) there are $q-s+1$ terms for which $[\Omega^{+i_p,\ldots,i_{p-s+1}}_{-j_q,\ldots,j_{q-s+1}}|\Sigma']$ is fixed}

$$d\pi_{\delta,\Lambda}(E_{rt})[\Omega|\Sigma] = \sum_{s=1}^{q} \sum_{\substack{i_p,\ldots,i_{p-s+1}\\ j_q,\ldots,j_{q-s+1}\\ i'_p,\ldots,i'_r\\ j'_q,\ldots,j'_{p+q-t+1}\\ i_p=i'_p,j_q=j'_q}} \left(\frac{\dim\lambda}{\dim\lambda'}\right)^{1/2} [\Omega^{+i_p,\ldots,i_{p-s+1}}_{-j_q,\ldots,j_{q-s+1}}|\Sigma^{+i'_p,\ldots,i'_r}_{-j'_q,\ldots,j'_{p+q-t+1}}]$$
$$\times\left(\sum_{i=s}^{q}\langle\Lambda,\alpha_i\rangle + l_{i_{p-s+1},p-s+1} - l_{j_{q-s+1},q-s+1} - \tfrac{1}{2}(\Lambda_s - 2k_{q-s+1}) - \frac{p+q}{2} + s\right)$$
$$\times\langle\{1\}_p\times\{\bar{1}\}_q,(E_{p-s+1,p+s});\Omega|\Omega'\rangle\,\langle\Sigma'|\{1\}_p\times\{\bar{1}\}_q,(E_{rt});\Sigma\rangle, \tag{51}$$

where $l_{ik} = m_{ik} + k/2 - i + \frac{1}{2}$, and $k_s = \Sigma_{i=1}^{s} m'_{is} - \Sigma_{i=1}^{s-1} m_{i,s-1}$ (with $k_1 = m'_{11}$, $r\leqslant p, t>p$).

For the case $r>p$, $t\leqslant p$ a similar expression is obtained,

$$d\pi_{\delta,\Lambda}(E_{rt})[\Omega|\Sigma] = \sum_{s=1}^{q} \sum_{\substack{i_p,\ldots,i_{p-s+1}\\ j_q,\ldots,j_{q-s+1}\\ i'_p,\ldots,i'_t\\ j'_q,\ldots,j'_{p+q-r+1}\\ i_p=i'_p,j_q=j'_q}} \left(\frac{\dim\lambda}{\dim\lambda'}\right)^{1/2} [\Omega^{-i_p,\ldots,i_{p-s+1}}_{+j_q,\ldots,j_{q-s+1}}|\Sigma^{-i'_p,\ldots,i'_t}_{+j'_q,\ldots,j'_{p+q-r+1}}]$$

$$\times\left(\sum_{i=s}^{q}\langle\Lambda,\alpha_i\rangle - l_{i_{p-s+1},p-s+1} + l'_{j_{q-s+1},q-s+1} + \tfrac{1}{2}(\Lambda_s - 2k_{q-s+1}) - \frac{p+q}{2} + s\right)$$

$$\times\langle\{\bar{1}\}_p\times\{1\}_q,(E_{p+s,p-s+1});\Omega|\Omega'\rangle\langle\Sigma'|\{\bar{1}\}_p\times\{1\}_q(E_{rt});\Sigma\rangle. \tag{52}$$

The formulas Eq. (51) and Eq. (52) give in *explicit* form the action of the infinitesimal operators $d\pi_{\delta,\Lambda}(E_{rt})$ of the representation $d\pi_{\delta,\Lambda}$ on the basis elements $[\Omega|\Sigma]$.

If, in place of the basis $|\Sigma\rangle$, another basis were chosen for the space of the irreducible representation λ of the subgroup K, then the explicit form for the action of the infinitesimal operators $d\pi_{\delta,\Lambda}(E_{rt})$ would be obtained in a similar manner. It is easy to recognize that this would lead to a replacement of the CGC's $\langle \Sigma' | \{1\}_p \times \{\bar{1}\}_q, (E_{rt}), \Sigma \rangle$ and $\langle \Sigma' | \{\bar{1}\}_p \times \{1\}_q, (E_{rt}), \Sigma \rangle$ by the CGC's for the new basis.

As was pointed out before, the CGC's entering Eq. (51) and Eq. (52) are products of CGC's of the groups U(p) and U(q). In fact, the CGC's of these two groups that enter these equations are the ones for a tensor product of an (arbitrary) finite dimensional irreducible representation with the vector representation or the contragredient to the vector representation. These have been obtained by Baird and Biedenharn in Ref. 9, and are given below.

The formulas (51) and (52) for the conformal group SU(2,2) were obtained in Refs. 24 and 25.

6. CLEBSCH–GORDAN COEFFICIENTS FOR THE TENSOR PRODUCTS $[m_n]\times\{1\}$ AND $[m_n]\times\{\bar{1}\}$ OF U(n) REPRESENTATIONS IN THE GEL'FAND–ZETLIN BASIS

In the following the CGC's are given for the tensor product of any finite dimensional representation $[m_n]\equiv[m_{1n},m_{2n},\ldots,m_{nn}]$ of U(n) with the representation $\{1\}\equiv[1,0,\ldots,0]$ and $\{\bar{1}\}\equiv[0,0,\ldots,0,-1]$ of U(n), In these tensor products each irreducible representation of U(n) occurs at most once, i.e., the multiplicity is 1. The CGC's of U(n) can be factored into a product of U(k)-scalar factors of the groups U(k), $k=3,4,\ldots,n$ and a CGC of U(2). Thus it is sufficient to list only the U(k)-scalar factors.

The U(n)-scalar factors for the group U($n-1$) are given by the formulas ($\{0\}$ denotes the identity representation)

$$\left(\begin{matrix} m_n & \{1\} \\ m_{n-1} & \{0\} \end{matrix}\middle| \begin{matrix} m_n^{+i} \\ m_{n-1} \end{matrix}\right) = \left|\frac{\Pi_{j=1}^{n-1}(m_{j,n-1}-m_{in}-j+i-1)}{\Pi_{j\neq i}(m_{jn}-m_{in}-j+i)}\right|^{1/2}, \quad i=1,2,\ldots,n,$$

$$\left(\begin{matrix} m_n & \{\bar{1}\} \\ m_{n-1} & \{0\} \end{matrix}\middle| \begin{matrix} m_n^{-i} \\ m_{n-1} \end{matrix}\right) = \left|\frac{\Pi_{j=1}^{n-1}(m_{j,n-1}-m_{in}-j+i)}{\Pi_{j\neq i}(m_{jn}-m_{in}-j+i)}\right|^{1/2}, \quad i=1,2,\ldots,n,$$

$$\left(\begin{matrix} m_n & \{1\} \\ m_{n-1} & \{1\} \end{matrix}\middle| \begin{matrix} m_n^{+i} \\ m_{n-1}^{+j} \end{matrix}\right) = S(i,j)\left|\frac{\Pi_{k\neq j}(m_{k,n-2}-m_{in}-k+i-1)\Pi_{k\neq i}(m_{kn}-m_{j,n-1}-k+j)}{\Pi_{k\neq i}(m_{kn}-m_{in}-k+i)\Pi_{k\neq j}(m_{k,n-1}-m_{j,n-1}-k+j)}\right|^{1/2},$$

$$\left(\begin{matrix} m_n & \{1\} \\ m_{n-1} & \{1\} \end{matrix}\middle| \begin{matrix} m_n^{-i} \\ m_{n-1}^{-j} \end{matrix}\right) = S(i,j)\left|\frac{\Pi_{k\neq j}(m_{k,n-1}-m_{in}-k+i)\Pi_{k\neq i}(m_{kn}-m_{j,n-1}-k+j+1)}{\Pi_{k\neq j}(m_{kn}-m_{in}-k+i)\Pi_{k\neq j}(m_{k,n-1}-m_{j,n-1}-k+j+1)}\right|^{1/2},$$

where $S(i,j)=+1$, if $i\leqslant j$, and $S(i,j)=-1$, if $i>j$. The proof of these formulas can be found in Ref. 9.

7. INFINITESIMAL OPERATORS OF FINITE DIMENSIONAL REPRESENTATIONS OF THE GROUPS U(p,q) AND U($p+q$) IN A U(p)$\times$U(q) BASIS

It now becomes essential to know which principal nonunitary series representation of U(p,q) a given finite dimensional representation contains. In order to gain this information, Proposition 3.2 of Ref. 26 is utilized. In this reference the principal nonunitary series representations are defined by left multiplication, while in this article they are defined by right multiplication [Eq. (1)]. It is however known that the representations induced for the same δ and Λ are equivalent for the two definitions. Moreover, Proposition 3.2 of Ref. 26 can be proved for the principal nonunitary series representations as defined by Eq. (1) in exactly the same manner as for the case of left multiplication. Thus, the proposition can be applied.

It was pointed out in the Introduction that the finite dimensional representations of U(p,q), U($p+q$), and the finite dimensional complex analytic representations of GL($p+q,C$), can be obtained through analytic continuation and restriction from the representations of anyone of these groups. Thus the discussion to follow can be restricted to the group U(p,q) alone.

The weights of the finite dimensional representations of U(p,q) [or $GL(p+q,C]$ will be considered with respect to the Cartan subalgebra $\mathfrak{b}+\mathfrak{a}'$, where $\mathfrak{b}$ is a Cartan subalgebra of $\mathfrak{m}$ (see Sec. 2). Let D_λ denote a finite dimensional irreducible representation of U(p,q) with highest weight λ. The weights ν of D_λ will be written in terms of coordinates $\nu_1,\nu_2,\ldots,\nu_{p+q}$ with respect to the basis $E_{1,1}\,E_{2,2},\ldots,E_{p+q,p+q}$ in $(\mathfrak{b}+\mathfrak{a}')_c$, the complexification of the subalgebra $\mathfrak{b}+\mathfrak{a}'$. What is needed is the restriction of the weights ν to the direct sum $\mathfrak{a}'+\mathfrak{b}'$ of the subalgebras $\mathfrak{a}'$ and $\mathfrak{b}'$, with $\mathfrak{b}'$ the Lie algebra of the subgroup U(1)$\times\cdots\times$U(1) (q times) of M. It can be verified directly that the restriction of ν to $\mathfrak{a}'+\mathfrak{b}'$ leads to the numbers $\nu_p+\nu_{p+1},\nu_{p-1}+\nu_{p+2},\ldots,\nu_{p-q+1}+\nu_{p+q}$ [the representation of the subgroups U(1) in $\mathfrak{m}$] and to the numbers $\nu_p-\nu_{p+1},\nu_{p-1}-\nu_{p+2},\ldots,\nu_{p-q+1}-\nu_{p+q}$ (the representation of the vector subgroups R of $A'=\exp\mathfrak{a}'$.) Each number $\nu_{p-k}+\nu_{p+k+1}$, $k=0,1,2,\ldots,q-1$, defines the representation of the subgroup U(1) which is given by u_{k+1} in Eq. (10). The number $\nu_{p-k}-\nu_{p+k+1}$, $k=0,1,2,\ldots,q-1$, defines the representation of the subgroup R in A' which corresponds to a_{k+1} in Eq. (11).

The weight of D_λ has to be found which, if restricted, leads to the lowest restricted weight (Proposition 3.2 in Ref. 26). Taking advantage of the similarity of the restricted root system with respect to the pair $(u'(p,q),\mathfrak{a}')$ and the root system of the classical Lie algebra B_n, it is easy to observe that a weight ν goes over into the lowest restricted weight if its coordinates satisfy

$$\nu_p - \nu_{p+1} \leqslant \nu_{p-1} - \nu_{p+2} \leqslant \cdots \leqslant \nu_{p-q+1} - \nu_{p+q} \leqslant 0 \tag{53}$$

and if there exists no other weight ν' which also satisfies Eq. (53), and for which $\nu' < \nu$ holds after restriction to $\mathfrak{a}'$. A computation shows that the weight λ' with coordinates

$$\lambda_{q+1}, \lambda_{q+2}, \ldots, \lambda_p, \lambda_{p+1}, \ldots, \lambda_{p+q}, \lambda_1, \lambda_2, \ldots, \lambda_q, \tag{54}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_{p+q}$ are the coordinates of the highest weight λ, leads to the lowest restricted weight. Moreover, every weight whose last $2q$ coordinates are identical with the last coordinates of the weight of Eq. (54), also leads to the lowest restricted weight. The weight vectors which belong to the lowest restricted weight form the basis for a subspace on which the irreducible representation of $U(p-q)$ (subgroup of M) with highest weight $\lambda_{q+1}, \lambda_{q+2}, \ldots, \lambda_p$ is realized.

Comparing these discussions with Proposition 3.2 of Ref. 26 it follows that:

The representation D_λ with highest weight $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{p+q})$ is contained as subrepresentation in that representation $\pi_{\delta,\Lambda}$ of the principal nonunitary series for which the representation δ is given by the numbers $\lambda_{q+1}, \lambda_{q+2}, \ldots, \lambda_p$ [the highest weight of the representation of $U(p-q)$] *and the numbers* $\Lambda_1 = \lambda_1 + \lambda_{p+q}, \Lambda_2 = \lambda_2 + \lambda_{p+q-1}, \ldots, \Lambda_q = \lambda_q + \lambda_{p+1}$ [the representations of the subgroups $U(1)$ in M], *and for which the linear form Λ is given by* $\lambda_{p+q} - \lambda_1, \lambda_{p+q-1} - \lambda_2, \ldots, \lambda_{p+1} - \lambda_q$ (the representation of the vector subgroups R in $A = \exp \mathfrak{a}$, i.e.,

$$\langle \lambda, \alpha_i \rangle = \tfrac{1}{2}(\lambda_{p+q-i+1} - \lambda_{p+q-i} - \lambda_i + \lambda_{i+1}),$$

$i = 1,2,\ldots,q-1,$

$$\langle \lambda, \alpha_q \rangle = \tfrac{1}{2}(\lambda_{p+1} - \lambda_q).$$

Let D_λ be a finite dimensional irreducible representation of $U(p,q)$ and $\pi_{\delta,\Lambda}$ the principal nonunitary series representation which contains D_λ as a subrepresentation. The representation $\pi_{\delta,\Lambda}$ is defined on the space $L^2_\delta(K,V)$. Thus it becomes necessary to extract from the space $L^2_\delta(K,V)$ the subspace V_λ which is the carrier space for the irreducible representation D_λ. In order to do this it is sufficient to find a basis of V_λ in terms of *linear combinations* of the vector functions $[\Omega \,|\Sigma]$. (In most cases it is not possible to select a basis for V_λ which consists simply of a subset of the $[\Omega\,|\Sigma]$). The desired basis for V_λ should correspond to the reduction $U(p,q) \supset U(p) \times U(q)$. Since each basis element for V_λ is to transform, with respect to right multiplication, according to a given irreducible representation of $U(p)\times U(q)$, it follows that each vector $[\Omega\,|\Sigma]$ in the linear combination has to transform according to the same irreducible representation. Thus, it follows that the basis elements for V_λ are of the form

$$\sum_\Omega \alpha(\Omega)[\Omega\,|\Sigma], \tag{55}$$

where the coefficients $\alpha(\Omega)$ do not depend on Σ (i.e., for all Σ the linear combinations are the same), and all terms of the sum correspond to the same irreducible representation of $U(p)\times U(q)$.

An orthonormal basis for the space V_λ can then be found in the following manner. First the irreducible representation η of $K \cong U(p)\times U(q)$ is extracted which is contained in D_λ and in $\pi_{\delta,\Lambda}$ with unit multiplicity. For this representation η the basis elements form a subset of functions $[\Omega\,|\Sigma]$, consisting of those $[\Omega\,|\Sigma]$ which transform under right multiplication according to the representation η of K, i.e., for this case the sum in Eq. (55) reduces to one term). These functions $[\Omega\,|\Sigma]$ can thus be included into an orthonormal basis for V_λ. In order to obtain the other basis elements for V_λ, it is now necessary to act upon the basis elements obtained so far by means of the operators $d\pi_{\delta,\Lambda}(E_{rt})$, $r>p, t\leqslant p$, and $r \leqslant p, t>p$. This leads to functions which belong to the space V_λ [since V_λ is invariant with respect to $d\pi_{\delta,\Lambda}(E_{rt})$]. Acting with the operators $d\pi_{\delta,\Lambda}(E_{rt})$ on the functions $[\Omega\,|\Sigma]$ leads to a sum of sums of functions, such that the functions of a given sum transform according to the same irreducible representation of $U(p)\times U(q)$, while functions of different sums transform according to nonequivalent representations of $U(p)\times U(q)$. Each fixed sum belongs to V_λ, as it is a linear combination of functions $[\Omega'\,|\Sigma']$ which transform according to a *given* irreducible representation of K. Thus functions of the space V_λ are obtained which transform according to irreducible representations of K. These functions can be orthonormalized and added to the set of basis elements for V_λ. Acting with the operators $d\pi_{\delta,\Lambda}(E_{rt})$ on the basis functions obtained in this manner, new orthonormal basis elements are obtained. After a finite number of steps an orthonormal basis is obtained for the space V_λ, which corresponds to the reduction $U(p,q)\supset U(p)\times U(q)$. Since the action of the operators $d\pi_{\delta,\Lambda}(E_{rt})$ on the function $[\Omega\,|\Sigma]$ is known, it is easy to define the action of the $d\pi_{\delta,\Lambda}(E_{rt})$ on the orthonormal basis for V_λ. The basis elements of V_λ will be denoted by $[\omega|\Sigma]$, where ω labels the subspaces of V_λ which are carrier spaces of equivalent irreducible representations of $K\cong U(p)\times U(q)$.

The finite dimensional irreducible representation of $U(p,q)$ obtained in this manner can be considered as a representation of $U(p+q)$. However, this representation may be nonunitary. In order to bring it into unitary form, a different basis may have to be chosen for V_λ. Let B denote the operator which transforms the basis constructed according to the procedure outlined above into a basis for V_λ for which the representation D_λ of $U(p+q)$ is unitary. The operator B must commute with $D_\lambda\,|K$. Thus, the basis functions of the new basis can be written as

$$\sum_\omega \alpha(\omega)[\omega|\Sigma], \tag{56}$$

where the numerical coefficients $\alpha(\omega)$ do not depend on Σ. If, in particular, the multiplicities of the irreducible representations of the subgroup K in $D_\lambda\,|K$ do not exceed 1, then the operator B is diagonal and is a multiple of the unit operator for every irreducible subspace of K.

For finite dimensional representations of $U(p+q)$ the unitarity condition is $E^*_{ij} = -E_{ji}$, where E_{ij} denotes the operators in the representation D_λ of $U(p+q)$. If this condition is written down in terms of the matrix elements of these operators in the basis Eq. (56), then a system of linear equations is obtained for the coefficients $\alpha(\omega)$. A solution of this system of equations leads to the explicit form of the basis for which the representation D_λ of $U(p+q)$ is unitary. This system of equations does however, in general, not lead to a unique solution. The various possible solutions are related to each other by unitary transformations.

ACKNOWLEDGMENTS

One of the authors (B.G.) wishes to thank the Academy of Sciences of the USSR and the Institute of Theoretical Physics of the Academy of Sciences of the Ukrainian SSR for the hospitality extended to him during a visit.

[1] *Group Theory and its Applications*, edited by M. W. Loebl (Academic, New York, 1968).
[2] L. O'Raifeartaigh, "Broken Symmetry," in Ref. 1, Vol. I; R. E. Behrends, "Broken SU(3) as a Particle Symmetry," in Ref. 1, Vol. I.
[3] P. Kramer and M. Moshinsky, "Harmonic Oscillators and Nuclear Structure," in Ref. 1, Vol. I.
[4] B. R. Judd, "Group Theory in Atomic Spectroscopy," in Ref. 1, Vol. I.
[5] B. G. Wybourne, "Lie Algebras in Quantum Chemistry: Symmetrized Orbitals," Int. J. Quant. Chem. **7**, 1117 (1973).
[6] B. R. Judd, "Lie groups and the Jahn–Teller effect," Canadian J. Phys. **52**, 999 (1974); B. R. Judd, "The Jahn–Teller effect for degenerate systems," Colloques internationaux C.N.R.S. No. 255-Spectroscopic des Elements lourds dans les Solides, Lyon (1976).
[7] F. A. Matsen, Int. J. Quant. Chem. **10**, 525 (1976).
[8] A. O. Barut, "Dynamical Groups and Generalized Symmetries in Quantum Theory," University of Canterbury, New Zealand (1972); L. Armstrong, J. Math. Phys. **12**, 553 (1971); E. Chaćon, D. Levi, and M. Moshinsky, J. Math. Phys. **17**, 1919 (1976).
[9] G. E. Baird and L. C. Biedenharn, J. Math. Phys. **4**, 1449 (1963); **5**, 1723, 1730 (1964).
[10] M. Moshinsky, Rev. Mod. Phys. **34**, 813 (1962).
[11] A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (PWN-Polish Scientific Publishers, Warsaw, 1977); A. O. Barut, "Introduction to the de Sitter and conformal groups and their physical applications," Lect. Theor. Phys. **13**, 3 University of Colorado (1971).
[12] A. U. Klimyk and A. M. Gavrilik, "Matrix Elements and Clebsch Gordan Coefficients of Group Representations," to appear in J. Math. Phys.
[13] M. A. Naimark, "Theory of Group Representations" (in Russian) (NAUKA, Moscow, 1976).
[14] S. J. Alisauskas and A. P. Jucys, Rep. Math. Phys. **5**, 7 (1974).
[15] S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962).
[16] R. Raczka and J. Fischer, Commun. Math. Phys. **3**, 233 (1966); **4**, 8 (1967).
[17] G. Warner, *Harmonic Analysis on Semisimple Lie Groups* (Springer Verlag, Berlin, 1972), Vol. I, II.
[18] Harish-Chandra, Trans. Amer. Math. Soc. **76**, 26 (1954).
[19] J. Lepowsky, Trans. Amer. Math. Soc. **176**, (1973).
[20] E. Thieleker, "Quasisimple and unitary representations of semisimple Lie groups," preprint.
[21] A. U. Klimyk, "Matrix Elements and Clebsch–Gordan Coefficients of Group Representations" (Naukova Dumka, Kiev 1979) (in Russian).
[22] N. Jacobson, *Lie Algebras* (Interscience, New York, 1965).
[23] S. Araki, J. Math., Asaka City University **13**, 1 (1962).
[24] A. Kilberg, W. Miller, and Halbwachs, Commun. Math. Phys. **3**, 194 (1966).
[25] P. I. Holod, preprint, ITP-74-100P, Kiev (1974).
[26] J. Lepowsky and N. R. Wallach, Trans. Amer. Math. Soc. **184**, 223 (1973).

Matrix elements for infinitesimal operators of the groups U($p+q$) and U(p,q) in a U(p)$\times$U(q) basis. II

A. U. Klimyk
Institute for Theoretical Physics, Academy of Sciences of the Ukrainian SSR, Kiev—130, USSR
Bruno Gruber
Department of Physics, Southern Illinois University, Carbondale, Illinois, 62901, USA

(Received 18 January 1979; accepted for publication 13 April 1979)

The general expressions derived in Ref. 1 for the matrix elements of the infinitesimal operators of the groups U($p+q$) and U(p,q) in a U(p)$\times$U(q) basis are utilized in this article to obtain *explicit* expressions for the matrix elements of the infinitesimal operators of (a) the degenerate series representations of U(p,q) in a U(p)$\times U$(q) basis, and (b) the representations of U($p+q$) with highest weight $(m_1, 0, ..., 0, m_2)$ in a U(p)$\times$U(q) basis. The operator which unitarizes the U($p+q$) representations considered is given in *explicit* form.

1. MATRIX ELEMENTS FOR THE DEGENERATE SERIES REPRESENTATIONS OF THE GROUP U(p,q) IN A U(p)$\times$U(q) BASIS

The following principal nonunitary series representations $\pi_{\delta\Lambda}$ of the group U(p,q) are considered:

(a) the representation δ of the subgroup $M \cong \mathrm{U}(p-q)\times\mathrm{U}(1)\times\mathrm{U}(1)\times\cdots\times\mathrm{U}(1)$ [U(1) — q times] is such that it differs from the identity representation only on the *first* subgroup U(1), while the representation δ is the identity representation for the subgroup $\mathrm{U}(p-q)\times\mathrm{U}(1)\times\cdots\times\mathrm{U}(1)$ of M [U(1) — ($q-1$) times; the *first* U(1) is deleted].

(b) the linear form Λ satisfies the conditions $\Sigma_{i=s}^{q}\langle\Lambda,\alpha_i\rangle = 0$, $s = 2,3,...,q$.

Thus, the representations $\pi_{\delta\Lambda}$ of U(p,q) which will be considered in the following are specified by two numbers. The first number Λ_1 defines the representation of (the first) U(1), while the second number $\langle\Lambda,\alpha_1\rangle$ defines the linear form Λ.

Consider the subspace of the space $L_\delta^2(K,V)$ for the representation $\pi_{\delta\Lambda}$ of U(p,q), spanned by the basis elements $[\Omega/\Sigma]$ for which holds

$$\Omega = \left(\begin{array}{cccccccccccccccccc} m_{1p} & & 0 & \cdot & \cdot & \cdot & \cdot & 0 & m_{pp} & m'_{1q} & & 0 & \cdot & \cdot & \cdot & 0 & & m'_{qq} \\ & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & & & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \\ & & \cdot & \cdot & \cdot & \cdot & \cdot & & & & & \cdot & \cdot & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & & & & & & & 0 & & 0 & & & \\ & & & 0 & & 0 & & & & & & & & 0 & & & & \\ & & & & 0 & & & & & & & & & & & & & \end{array}\right). \qquad (1)$$

This subspace is invariant with respect to $\pi_{\delta\Lambda}$. This can be seen by means of Eqs. (51) and (52) of Ref. 1, which give the action of the infinitesimal operators $d\pi_{\delta\Lambda}(E_{rt})$ on the states $[\Omega/\Sigma]$. In fact, the coefficients

$$\sum_{i=s}^{q}\langle\Lambda,\alpha_i\rangle + l_{i_{p-s+1},p-s+1} - l'_{j_{q-s+1},q-s+1} - \tfrac{1}{2}(\Lambda_s - 2K'_{q-s+1}) - (p+q)/2 + s,$$

$$\sum_{i=s}^{q}\langle\Lambda,\alpha_i\rangle - l_{i_{p-s+1},p-s+1} + l'_{j_{q-s+1},q-s+1} + \tfrac{1}{2}(\Lambda_s - 2K'_{q-s+1}) - (p+q)/2 + s,$$

of Eqs. (51) and (52) of Ref. 1 become zero for the values of Λ specified above and for the values $l_{1j}, l_{jj}, l'_{1i}, l'_{ii}$, $j = 1,2,..., p-1$, $i = 1,2,...,q-1$, as defined by means of the pattern equation (1). This however, leads to invariance.

The representation on the invariant subspace obtained in this manner is denoted by $\pi_{\lambda_1\lambda_2}$, where

$$\lambda_1 = \tfrac{1}{2}\Lambda_1 - \langle\Lambda,\alpha_1\rangle, \quad \lambda_2 = \tfrac{1}{2}\Lambda_1 + \langle\Lambda,\alpha_1\rangle. \qquad (2)$$

The form of the pattern equation (1) shows that the irreducible representations of U(p)$\times$U(q) occur in $\pi_{\lambda_1\lambda_2}\,|\,K$ with multiplicity $\leqslant 1$. Moreover, $\pi_{\lambda_1\lambda_2}\,|\,K$ contains only irreducible representations with highest weights

$$(m_{1p},0,...,0,m_{pp})_{\mathrm{U}(p)}\times(m'_{1q},0,...,0,m'_{qq})_{\mathrm{U}(q)}, \quad m_{1p}\geqslant 0, \quad m_{pp}\leqslant 0, \quad m'_{1q}\geqslant 0, \quad m'_{qq}\leqslant 0.$$

These irreducible representations of U(p)$\times$U(q) will be denoted by $[m_p]\times[m_q]$, where $m_p = (m_{1p},0,...,0,m_{pp})$ and $m_q = (m'_{1q},0,...,0,m'_{qq})$. The basis vectors for the space $[m_p]\times[m_q]$ will be denoted by $|\,m_p,\alpha,m_q,\beta\rangle$, where α and β denote Gelfand–Zetlin schemes. Since the multiplicity of $[m_p]\times[m_q]$ in $\pi_{\lambda_1\lambda_2}|K$ is $\leqslant 1$, the symbol Ω can be deleted in the basis states $[\Omega/\Sigma]$. Thus, the basis vectors $|\,m_p,\alpha,m_q,\beta\rangle$ correspond to the schemes Σ.

The formulas (51) and (52) of Ref. 1 lead to an explicit form for the infinitesimal operators $d\pi_{\lambda_1\lambda_2}(E_{rs})$ with respect to the

 0022-2488/79/102011-03$01.00

basis $|m_p,\alpha,m_q,\beta\rangle$. In particular, the operators $d\pi_{\lambda_1\lambda_2}(E_{p,p+1})$ and $d\pi_{\lambda_1\lambda_2}(E_{p+1,p})$ have the form ($m'_{qq}=\lambda_1+\lambda_2-m_{1p}-m_{pp}-m'_{1q}$)

$$\begin{aligned}d\pi_{\lambda_1\lambda_2}(E_{p,p+1})|m_p,\alpha,m_q,\beta\rangle &= (\lambda_2-m_{pp}-m'_{1q}-q+1)K^{+1}_{-1}K^{+1}_{-1}(\alpha,\beta)|m_p^{+1},\alpha,m_q^{-1},\beta\rangle+(-\lambda_1+m_{1p}+m'_{1q})\\ &\times K^{+1}_{-q}K^{+1}_{-q}(\alpha,\beta)|m_p^{+1},\alpha,m_q^{-q},\beta\rangle+(\lambda_2-m_{1p}-m'_{1q}-p-q+2)K^{+p}_{-1}K^{+p}_{-1}(\alpha,\beta)\\ &\times|m_p^{+p},\alpha,m_q^{-1},\beta\rangle+(-\lambda_1+m_{pp}+m'_{1q}-p+1)K^{+p}_{-q}K^{+p}_{-q}(\alpha,\beta)|m_p^{+p},\alpha,m_q^{-q},\beta\rangle,\end{aligned}\tag{3}$$

$$\begin{aligned}d\pi_{\lambda_1\lambda_2}(E_{p+1,p})|m_p,\alpha,m_q,\beta\rangle &= (-\lambda_1+m_{pp}+m'_{1q}-p+1)K^{-1}_{+1}K^{-1}_{+1}(\alpha,\beta)|m_p^{-1},\alpha,m_p^{+1},\beta\rangle\\ &+(\lambda_2-m_{1p}-m'_{1q}-p-q+2)K^{-1}_{+q}K^{-1}_{+q}(\alpha,\beta)|m_p^{-1},\alpha,m_q^{+q},\beta\rangle+(-\lambda_1+m_{1p}\\ &+m'_{1q})K^{-p}_{+1}K^{-p}_{+1}(\alpha,\beta)|m_p^{-p},\alpha,m_q^{+1},\beta\rangle+(\lambda_2-m_{pp}-m'_{1q}\\ &-q+1)K^{-p}_{+q}K^{-p}_{+q}(\alpha,\beta)|m_p^{-p},\alpha,m_q^{+q},\beta\rangle.\end{aligned}\tag{4}$$

The symbols $K^{\pm1(\pm p)}_{\mp1(\mp q)}$ and $K^{\pm1(\pm p)}_{\mp1(\mp q)}(\alpha,\beta)$ represent the following CGC's of the group $U(p)\times U(q)$:

$$K^{+1(+p)}_{-1(-q)}=\langle\{1\}_p\times\{\bar 1\}_q,(E_{p,p+1});\Omega\,|\,\Omega^{+1(+p)}_{-1(-q)}\rangle,$$

$$K^{-1(-p)}_{+1(+q)}=\langle\{\bar 1\}_p\times\{1\}_q,(E_{p+1,p});\Omega\,|\,\Omega^{-1(-p)}_{+1(+q)}\rangle,$$

$$\begin{aligned}K^{+1(+p)}_{-1(-q)}(\alpha,\beta)&=\left(\frac{\dim[m_p]\times[m_q]}{\dim[m_p^{+1(+p)}]\times[m_q^{-1(-q)}]}\right)^{1/2}\\ &\times\langle m_p^{+1(+p)},\alpha,m_q^{-1(-q)},\beta\,|\,\{1\}_p\times\{\bar 1\}_q,\\ &(E_{p,p+1});m_p,\alpha,m_q,\beta\rangle,\end{aligned}$$

$$\begin{aligned}K^{-1(-p)}_{+1(+q)}(\alpha,\beta)&=\left(\frac{\dim[m_p]\times[m_q]}{\dim[m_p^{-1(-p)}]\times[m_q^{+1(+q)}]}\right)^{1/2}\\ &\times\langle m_p^{-1(-p)},\alpha,m_q^{+1(+q)},\beta\,|\,\{\bar 1\}_p\times\{1\}_q,\\ &(E_{p+1,q});m_p,\alpha,m_q,\beta\rangle.\end{aligned}$$

The symbols $(E_{p,p+1})$ and $(E_{p+1,p})$ denote Gelfand–Zetlin schemes, given by Eqs. (33) and (34) of Ref. 1, for an appropriate choice of the indices i and j. The other infinitesimal operators $d\pi(E_{st})$ can either be obtained from the formulas given by Eqs. (51) and (52) of Ref. 1 or by means of the commutation relations of the operators $d\pi(E_{p,p+1})$ and $d\pi(E_{p+1,p})$ with the compact infinitesimal operators.

The set of irreducible representations of $U(p)\times U(q)$ in $\pi_{\lambda_1\lambda_2}|K$ is determined by means of the condition $m_{1p}+m_{pp}+m'_{1q}\leqslant\lambda_1+\lambda_2$.

2. REPRESENTATIONS OF U($p+q$) WITH HIGHEST WEIGHT (m_1,0,...,0,m_2) IN A U(p)×U(q) BASIS

The finite dimensional representations of $U(p,q)$ with highest weight $(m_1,0,\dots,0,m_2)$, $m_1\geqslant0$, $m_2\leqslant0$, are contained in the principal nonunitary series representation $\pi_{\delta\Lambda}$ of $U(p,q)$ for which it holds

$$\langle\Lambda,\alpha_1\rangle=\tfrac{1}{2}(m_2-m_1),\quad \Lambda_1=m_1+m_2,$$
$$\langle\Lambda,\alpha_2\rangle=\cdots=\langle\Lambda,\alpha_q\rangle=0,$$
$$\Lambda_2=\Lambda_3=\cdots=\Lambda_q=0,$$
$$m_{1,p-q}=m_{2,p-q}=\cdots=m_{p-q,p-q}=0.$$

[The numbers $m_{1,p-q}\dots m_{p-q,p-q}$ define the representations of $U(p-q)$ which form a component of the representation δ of M. See Sec. 7 of Ref. 1.]

The representation $\pi_{\delta\Lambda}$ which satisfies the conditions given above contains $\pi_{\lambda_1\lambda_2}$ as a subrepresentation. On the other side, the finite dimensional representation of $U(p,q)$ with highest weights $(m_1,0,\dots,0,m_2)$ contains the irreducible representation of $U(p)\times U(q)$ with highest weights $(m_1,0,\dots,0)_{U(p)}$ $(0,\dots,0,m_2)_{U(q)}$. This irreducible representation of $U(p)\times U(q)$ is contained in the subrepresentation $\pi_{\lambda_1\lambda_2}$ of $\pi_{\delta\Lambda}$. Therefore, all finite dimensional representations of $U(p,q)$ with highest weight $(m_1,0,\dots,0,m_2)$ are subrepresentations of $\pi_{\lambda_1\lambda_2}$. It is easy to show that $\lambda_1=m_1$ and $\lambda_2=m_2$ must hold. Thus, the finite dimensional representation of $U(p,q)$ with highest weight $(m_1,0,\dots,0,m_2)$ is contained as subrepresentation in the representation $\pi_{m_1m_2}$ of $U(p,q)$. It follows in particular that the irreducible representations of $U(p)\times U(q)$, found in the restriction of these finite dimensional representations of $U(p,q)$ [and $U(p+q)$] to the subgroup $U(p)\times U(q)$, occur with unit multiplicity.

Therefore, the finite dimensional representation with highest weight $(m_1,0,\dots,0,m_2)$ is given for the infinitesimal operators $E_{p,p+1}$ and $E_{p+1,p}$ of $U(p,q)$ by Eqs. (3) and (4), if $\lambda_1\to m_1$ and $\lambda_2\to m_2$.

The representations obtained in this manner through Eqs. (3) and (4) do not, however, satisfy the unitarity condition for the group $U(p+q)$:

$$E^*_{p,p+1}=E_{p+1,p}.$$

In order to obtain the irreducible representations of $U(p+q)$ in unitary form, a similarity transformation has to be performed. A new basis is introduced by means of a diagonal operator A such that in this new basis the representation becomes unitary. The similarity transformation is given by[2]

$$|m_p,\alpha,m_q,\beta\rangle=a(m_{1p},m_{pp},m'_{1q},m'_{qq})^{1/2}|m_p,\alpha,m_q,\beta\rangle',$$

where the matrix elements $a(m_{1p},m_{pp},m'_{1q},m'_{qq})$ of A are given by the formulas

$$\begin{aligned}a(m^0_{1p}+i,m^0_{pp},m^0_{1q}-i,m^0_{qq})&=a(m^0_{1p},m^0_{pp},m^0_{1q},m^0_{qq})\\ &\times\frac{\Pi^{i-1}_{j=0}(m_1-m^0_{pp}-m^0_{1q}+p+j)}{\Pi^{i}_{j=1}(m_2-m^0_{pp}-m^0_{1q}-q+j)},\end{aligned}\tag{5}$$

$$a(m_{1p}+i, m^0_{pp}, m_{1q}, m^0_{qq}-i) = a(m_{1p}, m^0_{pp}, m_{1q}, m^0_{qq}) \times \frac{\Pi^{i-1}_{j=0}(-m_2+m_{1p}+m_{1q}+p+q+j-1)}{\Pi^{i-1}_{j=0}(-m_1+m_{1p}+m_{1q}+j)}, \tag{6}$$

$$a(m_{1p}, m^0_{pp}+i, m_{1q}-i, m_{qq}) = a(m_{1p}, m^0_{pp}, m_{1q}, m_{qq}) \times \frac{\Pi^{i}_{j=1}(m_1-m_{1p}-m_{1q}+j)}{\Pi^{i}_{j=1}(m_2-m_{1p}-m_{1q}-p-q+j+1)}, \tag{7}$$

$$a(m_{1p}, m_{pp}+i, m_{1q}, m_{qq}-i) = a(m_{1p}, m_{pp}, m_{1q}, m_{qq}) \times \frac{\Pi^{i-1}_{j=0}(-m_2+m_{pp}+m_{1q}+q+j)}{\Pi^{i}_{j=1}(-m_1+m_{pp}+m_{1q}-p+j)} \tag{8}$$

Here m^0_{1p}, m^0_{pp}, m^0_{1q}, m^0_{qq} are appropriately chosen values of m_{1p}, m_{pp}, m_{1q}, and m_{qq}, respectively, $a(m^0_{1p}, m^0_{pp}, m^0_{1q}, m^0_{qq})$ is an arbitrary constant.

In the new basis $|m_p, \alpha, m_q, \beta\rangle'$, the infinitesimal operators $E_{p,p+1}$ and $E_{p+1,p}$ take on the form

$$\begin{aligned}
E_{p,p+1}|m_p,\alpha,m_q,\beta\rangle'
&= [(m_1-m_{pp}-m'_{1q}+p)(m_2-m_{pp}-m'_{1q}-q+1)]^{1/2}\times K^{+1}_{-1}K^{+1}_{-1}(\alpha,\beta)|m_p^{+1},\alpha,m_q^{-1},\beta\rangle' \\
&+ [(-m_1+m_{1p}+m'_{1q})(-m_2+m_{1p}+m'_{1q}+p+q-1)]^{1/2}K^{+1}_{-q}K^{+1}_{-q}(\alpha,\beta)|m_p^{+1},\alpha,m_q^{-q},\beta\rangle' \\
&+ [(m_1-m_{1p}-m'_{1q}+1)(m_2-m_{1p}-m'_{1q}-p-q+2)]^{1/2}K^{+p}_{-1}K^{+p}_{-1}(\alpha,\beta)|m_p^{+p},\alpha,m_q^{-1},\beta\rangle' \\
&+ [(-m_1+m_{pp}+m'_{1q}-p+1)(-m_2+m_{pp}+m'_{1q}+q)]^{1/2}K^{+p}_{-q}K^{+p}_{-q}(\alpha,\beta)|m_p^{+p},\alpha,m_q^{-q},\beta\rangle',
\end{aligned} \tag{9}$$

$$\begin{aligned}
E_{p+1,p}|m_p,\alpha,m_q,\beta\rangle'
&= -[(m_1-m_{pp}-m'_{1q}+p+1)(m_2-m_{pp}-m'_{1q}-q)]^{1/2}K^{-1}_{+1}K^{-1}_{+1}(\alpha,\beta)|m_p^{-1},\alpha,m_q^{+1},\beta\rangle' \\
&- [(-m_1+m_{1p}+m'_{1q}-1)(-m_2+m_{1p}+m'_{1q}+p+q+2)]^{1/2}K^{-1}_{+q}K^{-1}_{+q}(\alpha,\beta)|m_p^{-1},\alpha,m_q^{+q},\beta\rangle' \\
&- [(m_1-m_{1p}-m'_{1q})(-m_2-m_{1p}-m'_{1q}-p-q+1)]^{1/2}K^{-p}_{+1}K^{-p}_{+1}(\alpha,\beta)|m_p^{-p},\alpha,m_q^{+1},\beta\rangle' \\
&- [(-m_1+m_{pp}+m'_{1q}-p)(-m_2+m_{pp}+m'_{1q}+q-1)]^{1/2}K^{-p}_{+q}K^{-p}_{+q}(\alpha,\beta)|m_p^{-p},\alpha,m_q^{+q},\beta\rangle'.
\end{aligned} \tag{10}$$

The above equations give the infinitesimal generators $E_{p,p+1}$ and $E_{p+1,p}$ of $U(p+q)$ in unitary form. Again, as was mentioned before, the other infinitesimal operators E_{rs} can be obtained by utilizing the commutation relation of these operators with the compact infinitesimal operators which are given by the Gelfand–Zetlin formulas.

Degenerate unitary representations of the group $U(p,q)$ were studied previously by Raczka and Fischer.[3]

If in Eq. (5)–(8), which define the operator A, the symbols m_1 and m_2 are replaced by λ_1 and λ_2, then A can also be used to obtain (in unitary form) the representations of the supplementary degenerate unitary series of $U(p,q)$ from the representations $\pi_{\lambda_1\lambda_2}$ for special values of the parameters λ_1, λ_2.

ACKNOWLEDGMENT

One of the authors (B.G.) gratefully acknowledges the hospitality extended to him by the Institute for Theoretical Physics of the Academy of Sciences of the Ukrainian SSR.

[1]A.U. Klimyk and B. Gruber, "Matrix elements for infinitesimal operators of the groups $U(p+q)$ and $U(p,q)$ in a $U(p)\times U(q)$ basis. I," J. Math. Phys., see preceding paper.

[2]A.U. Klimyk, "Matrix elements and Clebsch–Gordan coefficients for group representations," Academy of Sciences of Ukrainian SSR, Naukova Dumka, Kiev, 1979.

[3]R. Raczka and J. Fisher, Commun. Math. Phys. **3**, 233 (1966); J. Fisher and R. Raczka, Commun. Math. Phys. **4**, 8 (1967); A.O. Barut and R. Raczka, *Theory of Group Representations and Applications* (PWN, Warzawa, 1977).

Sum rules for the matrices of the generators of SU(3) in an SO(3) basis

A. Partensky
Institut de Physique Nucléaire (and I N2 P3), Université Claude Bernard Lyon — 1, 69621 Villeurbanne, France
C. Quesne[a)]
Service de Physique Théorique et Mathématique CP 229, Université Libre de Bruxelles, 1050 Brussels, Belgium

(Received 8 March 1979; accepted for publication 27 April 1979)

We consider various sum rules for the semireduced [i.e., reduced with respect to SO(3)] matrix elements of the generators of SU(3) in a basis of an irreducible representation $[pq]$ corresponding to the group reduction SU(3)⊃SO(3)⊃SO(2). We use basis states which diagonalize an additional labeling operator K, but avoid their explicit construction. We build all the needed operators from the two independent SU(3) vector operators X and V, where $X = (L, Q)$ is made of the SU(3) generators and $V = (V^L, V^Q)$ is defined in terms of them. First we obtain an analytical formula for the linear sum rule satisfied by the diagonal semireduced matrix elements of Q. Then, from the set of quadratic equations fulfilled by the semireduced matrix elements of Q and V^Q, we obtain explicit expressions for the quadratic sum rules satisfied by these quantities. All the above-mentioned sum rules are independent of the selection made for K. When K is defined as the third order operator $L.V^L$, we show that a relation between some nondiagonal matrix elements of Q and V^Q exists enabling the determination of a k-weighted quadratic sum rule for the semireduced matrix elements of Q. As by-products of the preceding results we obtain general formulas for the eigenvalues k of K for all the values of L whose multiplicity does not exceed 3, and we show that we are able to compute analytically the individual matrix elements of Q for not too high dimensionalities by working out the case of the irreducible representation [10,5].

1. INTRODUCTION

It is well known from the work of Gel'fand and Tseitlin and others[1-3] that the chain of subgroups of SU(3)

$$\mathrm{SU}(3)\supset S[\mathrm{U}(2)\times \mathrm{U}(1)]\supset S[\mathrm{U}(1)\times \mathrm{U}(1)\times \mathrm{U}(1)] \quad (1.1)$$

is canonical, i.e., gives rise to a complete set of commuting Hermitian operators (the Casimir operators of the group and of all the subgroups of the chain), providing us with a complete labeling for the states transforming under an irreducible representation (IR) of SU(3). The matrices of the generators of SU(3) in the corresponding canonical orthonormal basis are also easily constructed.[1,3,4]

However in the many-body problem, the physically interesting chain of subgroups is[5]

$$\mathrm{SU}(3)\supset \mathrm{SO}(3)\supset \mathrm{SO}(2). \quad (1.2)$$

The Casimir operators of SU(3), I_2 and I_3, and the angular momentum operators, L^2 and L_0, are not sufficient to characterize the basis of an IR of SU(3) completely. Several methods have been proposed to define the missing label.[6] They can be divided into two classes.

The first type of methods gives rise to a simple labeling of the states by integers but to nonorthogonal basis.[7-10] The second type of solution uses a labeling Hermitian operator K, which belongs to the enveloping algebra of SU(3) and commutes with all the Casimir operators of the group and subgroup.[8,11] The common eigenstates of the whole set of commuting operators are orthogonal but the corresponding eigenvalues of K are in general irrational numbers.[11]

During the last few years, the latter type of method has been extensively used, and new interesting results obtained. It has been shown that the functionally independent missing label operators, whose number is equal to twice that of missing labels,[12] can be chosen as operators of third and fourth order in the generators of SU(3) respectively.[13] Moreover, the eigenvalues and eigenstates of these operators have been calculated analytically or numerically for most IR's of SU(3) likely to appear in physical applications.[13,14]

The method used in Refs. 13 and 14 is based on the Gel'fand and Tseitlin canonical basis. Starting from the well known matrices of the generators of SU(3) in this basis, those of K are constructed and diagonalized. In this way, the eigenvalues of K and the corresponding states of the physical basis are obtained. As a byproduct, the matrices of the generators of SU(3) in this physical basis could then be calculated by a standard basis transformation.

However in many physical applications (diagonalization of a many-body Hamiltonian, calculation of transition strengths, etc.) we are mainly interested by those matrices and much less by the eigenstates themselves. Moreover, the most important physical information about those matrices is often contained in a small number of sum rules for their elements.[15,16] We might therefore prefer to reverse the operation sequence with respect to the one of Refs. 13 and 14 and try to directly compute the interesting sum rules without explicitly constructing the eigenstates of K.

The main purpose of the present paper is to show that such a procedure is practicable for various sum rules and to establish for them some analytical formulas from which their numerical values could then be computed either directly or through a computer code. Our procedure is similar to that one proposed by Partensky and Maguin for the chain SU(4)⊃SU(2)×SU(2).[17] It uses extensively the Lie algebra

[a)]Maître de recherches F.N.R.S.

 0022-2488/79/102014-14$01.00

of SU(3) in the physical chain (1.2), as well as standard Wigner–Racah SU(2) algebraic techniques. It is therefore well suited to analytical calculations, which generally represent a heavy burden when one does not take advantage of the simplifications implied by the Wigner–Racah SU(2) calculus.[18] It should be emphasized that the applicability of our method is not restricted by the dimensionality of the IR of SU(3), nor by the multiplicities of the IR's of SO(3) contained in that IR. However, whenever the latter multiplicities are low enough (generally speaking not greater than three), we can go one step further and obtain the eigenvalues of K, as well as the individual matrix elements of the SU(3) generators. This point will be illustrated by an example.

In Sec. 2 we define the notations used in this paper for the generators $X = (L,Q)$ of SU(3). Section 3 deals with the basis of the IR's of SU(3), both the Gel'fand and Tseitlin basis and the physical one obtained by diagonalizing an additional labeling operator K. The latter is defined in terms of an SU(3) vector operator $V = (V^L, V^Q)$ which will play an important part in the remainder of this paper.

In Sec. 4 we introduce the semireduced matrix elements of Q and V^Q and study some of their symmetry properties. The next sections are then devoted to the determination of sum rules for those matrix elements in a given IR of SU(3). In Sec. 5 we obtain an analytical expression for the linear sum rule satisfied by the semireduced matrix elements of Q, i.e., the sum of its diagonal semireduced matrix elements over the additional quantum number. In Sec. 6 we write the set of quadratic equations fulfilled by the semireduced matrix elements of Q and V^Q. Finally, in Sec. 7, we obtain from these equations explicit expressions for the quadratic sum rules (squares or products summed over the additional quantum numbers) satisfied by these quantities.

In Sec. 8 we particularize our results by choosing for K a third order labeling operator. We then show that the eigenvalues of the latter are proportional to the diagonal semireduced matrix elements of Q. Moreover, the operator satisfies an important commutation relation which implies a link between the nondiagonal semireduced matrix elements of Q and V^Q. As a consequence, we are able to find explicit expressions for some k-weighted quadratic sum rules.

In Sec. 9, using the results established in the preceding sections for the diagonal sum rules, we obtain general formulas for the eigenvalues of the third order labeling operator when the multiplicity is equal to two or three. Then we solve the equations established in the preceding sections for the semireduced matrix elements of Q and the eigenvalues of K in the case of the IR [10,5] of SU(3), for which the multiplicity of the IR's of SO(3) is not greater than three.

2. THE SU(3) ALGEBRA

As is well known,[1-3] the Lie algebra of SU(3) is generated by the operators E_{ij}, $i,j = 1,2,3$, which obey the following relations:

(1) The commutation rule

$$[E_{ij}, E_{kl}] = \delta_{jk} E_{il} - \delta_{il} E_{kj}, \tag{2.1}$$

(2) the Hermitian conjugate relation

$$E_{ij}^{+} = E_{ji}, \tag{2.2}$$

(3) the unimodular condition

$$\sum_{i=1}^{3} E_{ii} = N, \tag{2.3}$$

where N is the number operator of U(3). This basis of the SU(3) algebra is adapted to the canonical chain of subgroups (1.1).

When considering the physical chain of subgroups (1.2), we define new generators L_α ($\alpha = \pm 1,0$), and Q_μ ($\mu = \pm 2, \pm 1,0$), which are to be interpreted as the angular momentum and the quadrupole operators respectively.[5] They are defined in terms of the canonical generators E_{ij} by

$$L_0 = E_{11} - E_{22},$$
$$L_1 = -E_{13} - E_{32}, \quad L_{-1} = E_{31} + E_{23},$$

and

$$Q_0 = 3^{-1/2}[E_{11} + E_{22} - 2E_{33}], \quad Q_1 = E_{32} - E_{13},$$
$$Q_{-1} = E_{31} - E_{23}, Q_2 = \sqrt{2} E_{12}, \quad Q_{-2} = \sqrt{2} E_{21}. \tag{2.4}$$

They obey the following commutation rules:

$$[L,L]^\lambda_\kappa = -\sqrt{2}\, \delta_{\lambda,1} L_\kappa, \tag{2.5a}$$
$$[L,Q]^\lambda_\kappa = -\sqrt{6}\, \delta_{\lambda,2} Q_\kappa, \tag{2.5b}$$
$$[Q,Q]^\lambda_\kappa = \sqrt{10}\, \delta_{\lambda,1} L_\kappa, \tag{2.5c}$$

where we define the coupled commutator $[S^{\lambda_1}, T^{\lambda_2}]^\lambda_\kappa$ of two irreducible tensor operators S^{λ_1} and T^{λ_2}, of rank λ_1 and λ_2 respectively, by

$$[S^{\lambda_1}, T^{\lambda_2}]^\lambda_\kappa = \sum_{\kappa_1\kappa_2} \langle \lambda_1\kappa_1\lambda_2\kappa_2 | \lambda\kappa\rangle [S^{\lambda_1}_{\kappa_1}, T^{\lambda_2}_{\kappa_2}]$$
$$= [S^{\lambda_1} \times T^{\lambda_2}]^\lambda_\kappa - (-1)^{\lambda_1+\lambda_2-\lambda} \times [T^{\lambda_2} \times S^{\lambda_1}]^\lambda_\kappa, \tag{2.6}$$

and $\langle \lambda_1\kappa_1\lambda_2\kappa_2 | \lambda\kappa\rangle$ is an SU(2) Wigner coefficient. Finally, these generators obey the Hermitian conjugate relations

$$L_\alpha^{+} = (-1)^\alpha L_{-\alpha}, \tag{2.7a}$$
$$Q_\mu^{+} = (-1)^\mu Q_{-\mu}. \tag{2.7b}$$

3. BASIS OF IR'S OF SU(3)

In the Gel'fand and Tseitlin scheme, the basis states of an IR$[pq]$ of SU(3) are represented by the patterns[1-3]

$$\left|\begin{matrix} p & & q & & 0 \\ & m_{12} & & m_{22} & \\ & & m_{11} & & \end{matrix}\right\rangle, \tag{3.1}$$

where p, q, and m_{ij} are integers such that $p \geqslant m_{12} \geqslant q \geqslant m_{22} \geqslant 0$, and $m_{12} \geqslant m_{11} \geqslant m_{22}$.

In the physical scheme, the basis states of the IR $[pq]$ are defined as the common eigenstates of the Casimir operators I_2 and I_3 of SU(3), the angular momentum operators $L^2 = \Sigma_\alpha(-1)^\alpha L_\alpha L_{-\alpha}$ and L_0, and an additional labeling operator K.[5] These operators are most easily constructed in terms of the independent SU(3) vector operators. Following Biedenharn,[3] there are two such operators which we shall call $X = (X^L, X^Q)$ and $V = (V^L, V^Q)$ respectively. The first one is made of the generators $L_\alpha (= X^L_\alpha)$, and $Q_\mu (= X^Q_\mu)$, and the second one is obtained from them by symmetrical coupling. It can be easily shown that

$$V^L_\alpha = -2\sqrt{\tfrac{10}{3}}[L\times Q]^1_\alpha,$$
$$V^Q_\mu = \sqrt{\tfrac{14}{3}}[Q\times Q]^2_\mu + \sqrt{2}[L\times L]^2_\mu, \tag{3.2}$$

and that these operators satisfy the commutation relations

$$[L,V^L]^\lambda_\kappa = -\sqrt{2}\,\delta_{\lambda,1}V^L_\kappa, \tag{3.3a}$$
$$[Q,V^L]^\lambda_\kappa = [L,V^Q]^\lambda_\kappa = -\sqrt{6}\,\delta_{\lambda,2}V^Q_\kappa, \tag{3.3b}$$
$$[Q,V^Q]^\lambda_\kappa = \sqrt{10}\,\delta_{\lambda,1}V^L_\kappa, \tag{3.3c}$$

as it should be. Let us note that in Eq. (3.2) we have arbitrarily fixed the overall multiplicative factor. Under Hermitian conjugation, the operators V^L_α and V^Q_μ behave as L_α and Q_μ respectively. Some nontrivial commutation relations can be deduced from the definitions (3.2) and the commutation relations (3.3). For instance, it can be shown that

$$[V^Q,V^Q]^\lambda_\kappa = [1-(-1)^\lambda]\{a_\lambda[V^L\times Q]^\lambda_\kappa + b_\lambda[L\times V^Q]^\lambda_\kappa\}, \tag{3.4}$$

where

$$a_1 = -7/\sqrt{3},\quad b_1 = -\sqrt{3},\quad a_3 = -b_3 = 2\sqrt{2}. \tag{3.5}$$

In terms of the vector operators X and V, the two independent Casimir operators of SU(3) can be written as[3]

$$I_2 = \sum_\alpha(-1)^\alpha L_\alpha L_{-\alpha} + \sum_\mu(-1)^\mu Q_\mu Q_{-\mu}, \tag{3.6}$$

and

$$I_3 = \sum_\alpha(-1)^\alpha L_\alpha V^L_{-\alpha} + \sum_\mu(-1)^\mu Q_\mu V^Q_{-\mu}. \tag{3.7}$$

One further invariant operator can be constructed in this way, namely the fourth-order Casimir operator

$$I_4 = \sum_\alpha(-1)^\alpha V^L_\alpha V^L_{-\alpha} + \sum_\mu(-1)^\mu V^Q_\mu V^Q_{-\mu}, \tag{3.8}$$

which can be rewritten in terms of I_2 as

$$I_4 = \tfrac{4}{3}I_2(I_2+3). \tag{3.9}$$

It is straightforward to show that we can choose the third-order operator $\Sigma_\alpha(-1)^\alpha L_\alpha V^L_{-\alpha}$ or the fourth-order one $\Sigma_\alpha(-1)^\alpha V^L_\alpha V^L_{-\alpha}$ for a labeling operator. All the formulas derived in Secs. 4–7 are valid for either choice for K. However in Secs. 8 and 9, we set

$$K = \sum_\alpha(-1)^\alpha L_\alpha V^L_{-\alpha}, \tag{3.10}$$

and the results obtained there depend on this choice. In Tables I and II we give the connections between the labeling operators chosen by other authors and those used in the present paper.

The eigenvalues of the two fundamental Casimir invariants uniquely label an IR of SU(3). They are given in terms of p and q by

$$\langle I_2\rangle = \tfrac{1}{3}(p^2-pq+q^2+3p),$$
$$\langle I_3\rangle = \tfrac{1}{9}(p-2q)(2p-q+3)(p+q+3). \tag{3.11}$$

Let $L(L+1)$ and M be the eigenvalues of L^2 and L_0 respectively, and k denote that of K. Then the basis states of the IR $[pq]$ in the physical chain (1.2) are written as

$$|[pq]kLM\rangle \quad\text{or}\quad |kLM\rangle, \tag{3.12}$$

when there is no ambiguity. The values of L and M are restricted to the range

$$0\leqslant L\leqslant p,\quad -L\leqslant M\leqslant L. \tag{3.13}$$

The multiplicity of the IR L of SO(3) in the IR $[pq]$ of SU(3) is given by Racah's formula[19]

$$N_L(pq) = [\tfrac{1}{2}P(p-L+2)] - [\tfrac{1}{2}P(q-L+1)] - [\tfrac{1}{2}P(p-q-L+1)], \tag{3.14}$$

where $[x]$ denotes the greatest integer contained in x, and

$$P(y) = \tfrac{1}{2}(y+|y|). \tag{3.15}$$

TABLE II. Connection between fourth-order labeling operators used by other authors and the operator $K=\Sigma_\alpha(-1)^\alpha V^L_\alpha V^L_{-\alpha}$ considered in this paper. Here $L^2=\Sigma_\alpha(-1)^\alpha L_\alpha L_{-\alpha}$, and $Q^2=\Sigma_\mu(-1)^\mu Q_\mu Q_{-\mu}$.

Ref.	Name	Connection with K
11	y	$\frac{1}{4}K$
18	Q^0_l	$-\frac{9}{2}K+12L^2Q^2-36L^2$
13 Eq. (17)	$X^{(4)}$	$\frac{1}{16}K$
13 Eq. (26′)	$X^{(4)}$	$\frac{9}{4}K-6L^2Q^2-63Q^2+\frac{165}{2}L^2$
6	$X^{(4)}$	$\frac{9}{4}K-6L^2Q^2-\frac{45}{2}L^2$

TABLE I. Connection between third-order labeling operators used by other authors and the operator $K=\Sigma_\alpha(-1)^\alpha L_\alpha V^L_{-\alpha}$ considered in this paper.

Ref.	Name	Connection with K
8	Ω	$-\frac{1}{4}K$
11	x	$\frac{1}{2}K$
18	0^0_l	$3\sqrt{\frac{3}{2}}K$
13 Eq. (17)	$X^{(3)}$	$\frac{1}{4}K$
13 Eq. (26)	$X^{(3)}$	$\frac{3}{4}K$
6	$X^{(3)}$	$\frac{3}{4}K$

From now on, we are going to work directly in the physical basis (3.12) with the purpose of deriving sum rules for the matrix elements of the generators in this basis. Use of the Gel'fand and Tseitlin basis (3.1) will be made only in one case as an intermediate step.

4. MATRICES OF THE GENERATORS IN THE PHYSICAL BASIS

In a given IR$[pq]$ of SU(3), the matrix elements of the angular momentum operators L_α are given by the Wigner–Eckart theorem applied to the SO(3) subgroup:

$$\langle k'L'M'|L_\alpha|kLM\rangle = \delta_{k'k}\delta_{L'L}\sqrt{L(L+1)} \times \langle LM\,1\alpha|L'M'\rangle. \quad (4.1)$$

Those of the quadrupole operators Q_μ can be expressed in terms of the so-called semireduced matrix elements $\langle k'L'\|Q\|kL\rangle$, i.e., reduced with respect to the SO(3) subgroup:

$$\langle k'L'M'|Q_\mu|kLM\rangle = \langle LM\,2\mu|L'M'\rangle \times \langle k'L'\|Q\|kL\rangle. \quad (4.2)$$

The calculation of the matrices of the generators is thus equivalent to that of the semireduced matrix elements of Q.

The semireduced matrix elements of V^L and V^Q can all be expressed in terms of those of Q:

$$\langle k'L'\|V^L\|kL\rangle = 2(-1)^{L+L'}[10L'(L'+1)(2L'+1)]^{1/2} \times \begin{Bmatrix} 1 & 2 & 1 \\ L & L' & L' \end{Bmatrix} \langle k'L'\|Q\|kL\rangle, \quad (4.3)$$

$$\langle k'L'\|V^Q\|kL\rangle = (-1)^{L+L'}\sqrt{\tfrac{70}{3}} \sum_{k''L''} \Big[(2L''+1)^{1/2} \times \begin{Bmatrix} 2 & 2 & 2 \\ L & L' & L'' \end{Bmatrix} \langle k'L'\|Q\|k''L''\rangle \times \langle k''L''\|Q\|kL\rangle \Big] + \delta_{k'k}\,\delta_{L'L} \times [\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}. \quad (4.4)$$

The Hermitian conjugate relation (2.7b) leads to the following symmetry relation for the semireduced matrix elements of Q:

$$\langle k'L'\|Q\|kL\rangle = (-1)^{L-L'}[(2L+1)(2L'+1)^{-1}]^{1/2} \times \langle kL\|Q\|k'L'\rangle. \quad (4.5)$$

A similar relation is satisfied by those of V^Q.

The semireduced matrix elements of Q in two contragredient IR's of SU(3), $[pq]$ and $[p, p-q]$, can be related to one another in the following way. Let

$$T\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix}$$

be an irreducible tensor transforming under the IR$[pq]$ of SU(3). The adjoint tensor is characterized by

$$T\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix}^+ = h\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix} \times T\begin{pmatrix} p & & p-q \\ & k^+ & \\ L & & -M \end{pmatrix} \quad (4.6)$$

where

$$h\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix}$$

is a phase factor, and k^+ denotes the eigenvalue of K in the contragredient IR $[p, p-q]$ corresponding to k in the IR $[pq]$. By definition of

$$T\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix},$$

the following commutation relation is satisfied:

$$\left[X_\kappa, T\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix}\right] = \sum_{k'L'M'} \langle [pq]k'L'M'|X_\kappa|[pq]kLM\rangle \times T\begin{pmatrix} p & & q \\ & k' & \\ L' & & M' \end{pmatrix}, \quad (4.7)$$

where X_κ is any generator of SU(3) (L_α or Q_μ). The Hermitian conjugate of this equation leads to the following results:

$$h\begin{pmatrix} p & & q \\ & k & \\ L & & M \end{pmatrix} = (-1)^{L-M} h\begin{pmatrix} p & q \\ k & L \end{pmatrix}, \quad (4.8)$$

where

$$h\begin{pmatrix} p & q \\ k & L \end{pmatrix}$$

is a phase factor independent of M, and

$$h\begin{pmatrix} p & q \\ k' & L' \end{pmatrix} \langle [pq]k'L'\|Q\|[pq]kL\rangle + h\begin{pmatrix} p & q \\ k & L \end{pmatrix} \times \langle [p,p-q]k'^+L'\|Q\|[p,p-q]k^+L\rangle = 0, \quad (4.9)$$

which is the desired relation between semireduced matrix elements of Q in two contragredient IR's of SU(3). As a special case, the diagonal semireduced matrix elements satisfy the relation

$$\langle [pq]kL\|Q\|[pq]kL\rangle = -\langle [p,p-q]k^+L\|Q\|[p,p-q]k^+L\rangle. \quad (4.10)$$

5. LINEAR SUM RULE SATISFIED BY THE SEMIREDUCED MATRIX ELEMENTS OF Q

In this section, we are going to derive an expression for the sum

$$\langle L|Q|L\rangle \equiv \sum_k \langle kL\|Q\|kL\rangle \quad (5.1)$$

in a given IR $[pq]$ of SU(3). The method used can be applied in principle to the trace

$$\langle L|T^\lambda|L\rangle \equiv \sum_k \langle kL\|T^\lambda\|kL\rangle \quad (5.2)$$

of any irreducible tensor operator $T^{\lambda}_{\mu}\,(\mu=-\lambda,\ldots,\lambda)$ of rank λ with respect to SO(3). We are going, therefore, to formulate it in the general case.

Let us first assume that we are able to calculate the trace of T^{λ}_{0} in the basis states of the IR $[pq]$ of SU(3) characterized by a given angular momentum projection M:

$$S_M \equiv \sum_{L\geqslant|M|}\sum_{k}\langle kLM\,|\,T^{\lambda}_{0}|kLM\rangle. \tag{5.3}$$

Then from the Wigner–Eckart theorem applied to the SO(3) subgroup, we get

$$S_M \equiv \sum_{L\geqslant|M|}\langle LM\lambda\,0\,|\,LM\rangle\langle L\,|T^{\lambda}|L\rangle. \tag{5.4}$$

When we consider Eq. (5.4) for all nonnegative values of M $(M=0,1,\ldots,p)$, we get a set of $p+1$ simultaneous linear equations in the $p+1$ unknowns $\langle L\,|T^{\lambda}|L\rangle, L=0,1,\ldots,p$, whose solution leads to the expression of the trace (5.2) in terms of the S_M's.

Next we calculate the trace S_M by using its invariance under the transformation from the physical basis to the Gel'fand one. As the eigenvalue of L_0 corresponding to the Gel'fand state (3.1) is equal to $2m_{11}-m_{12}-m_{22}$, the trace S_M becomes

$$S_M=\sum_{\alpha\beta}\left\langle\begin{matrix}\beta & 2\alpha-\beta-M\\ & \alpha\end{matrix}\right|T^{\lambda}_{0}\left|\begin{matrix}\beta & 2\alpha-\beta-M\\ & \alpha\end{matrix}\right\rangle, \tag{5.5}$$

where $\alpha=m_{11}$ and $\beta=m_{12}$ are restricted to those values satisfying the following inequalities:

$$\max(q,|M|)\leqslant\beta\leqslant p,$$
$$\tfrac{1}{2}(\beta+M)\leqslant\alpha\leqslant\min[\beta,\beta+M,\tfrac{1}{2}(q+\beta+M)]. \tag{5.6}$$

To complete the derivation, we then only need to compute the diagonal matrix elements of T^{λ}_{0} in the Gel'fand basis and to perform the summations over α and β in Eq. (5.5) under the conditions (5.6).

The method we have just explained can be trivially applied to the unit operator for which $\lambda=0$. We then get Racah's formula (3.14) for the multiplicity $N_L(pq)$ again.

Let us turn now to the trace of Q for which $\lambda=2$. The set of linear equations (5.4) becomes

$$S_M=\sum_{L\geqslant|M|}[L(L+1)(2L-1)(2L+3)]^{-1/2}\times[3M^2-L(L+1)]\langle L\,|Q\,|L\rangle. \tag{5.7}$$

Its solution can be shown to be

$$\langle L\,|Q\,|L\rangle=[L(2L-1)]^{-1/2}[(L+1)(2L+3)]^{1/2}\times\Big\{S_L-S_{L+1}+3[(L+1)(2L+3)]^{-1}\times\sum_{M\geqslant L+1}(2M+1)[S_M-S_{M+1}]\Big\}. \tag{5.8}$$

It remains now to perform the summations over α and β in Eq. (5.5) taking into account that the eigenvalue of Q_0 corresponding to the Gel'fand state (3.1) is equal to $3^{-1/2}[3(m_{12}+m_{22})-2(p+q)]=3^{-1/2}[3(2\alpha-M)-2(p+q)]$, and to substitute the result for S_M into Eq. (5.8).

When $p\geqslant 2q$, we get the following expression for $\langle L\,|Q\,|L\rangle$:

$$\langle L\,|Q\,|L\rangle=\tfrac{1}{2}[3L(L+1)(2L-1)(2L+3)]^{-1/2}\times\begin{cases}(-1)^L L(L+1)[(-1)^p(p-2q)-(-1)^q(2p-q+3)+(-1)^{p+q}(p+q+3)], & \text{if } 0\leqslant L<q,\\ \{L(L+1)[(-1)^{p+q+L}(p+q+3)+(-1)^{p+L}(p-2q)\\ \quad+(q+1)(2p-q+3)]-q(q+1)(q+2)(2p-q+3)\}, & \text{if } q\leqslant L<p-q,\\ (p-2q)\{(p+1)(p+2)(p+3)-L(L+1)[p+2-(-1)^{p+L}]\}, & \text{if } p-q\leqslant L\leqslant p.\end{cases} \tag{5.9}$$

When $p<2q$, we use the symmetry relation (4.10) to show that

$$\langle[pq]L\,|\,Q\,|\,[pq]L\rangle=-\langle[p,p-q]L\,|\,Q\,|\,[p,p-q]L\rangle, \tag{5.10}$$

and we apply Eq. (5.9) to the IR $[p,p-q]$.

In this section,we derived a linear sum rule satisfied by the semireduced matrix elements of Q. We will now get some quadratic relations between the semireduced matrix elements of Q and V^Q.

6. QUADRATIC RELATIONS SATISFIED BY THE SEMIREDUCED MATRIX ELEMENTS OF Q AND V^Q

A first class of relations is obtained by taking the semireduced matrix elements of the second, third, and fourth-order Casimir operators of SU(3), defined in Eqs. (3.6), (3.7), and (3.8) respectively. Using standard SU(2) Wigner–Racah calculus and the symmetry relation (4.5), we get straightforwardly

$$\sum_{k'L'}\langle kL\,\|Q\,\|k'L'\rangle\langle\bar{k}L\,\|Q\,\|k'L'\rangle=[\langle I_2\rangle-L(L+1)]\,\delta_{\bar{k}k}, \tag{6.1}$$

$$\sum_{k'L'}\langle kL\,\|Q\,\|k'L'\rangle\langle\bar{k}L\,\|V^Q\|k'L'\rangle=\langle I_3\rangle\delta_{\bar{k}k}-[L(L+1)]^{1/2}\langle\bar{k}L\,\|V^L\|kL\rangle, \tag{6.2}$$

$$\sum_{k'L'}\langle kL\,\|V^Q\|k'L'\rangle\langle\bar{k}L\,\|V^Q\|k'L'\rangle=\langle I_4\rangle\delta_{\bar{k}k}-\sum_{k'L'}\langle kL\,\|V^L\|k'L'\rangle\langle\bar{k}L\,\|V^L\|k'L'\rangle, \tag{6.3}$$

where $\langle I_2\rangle$, $\langle I_3\rangle$, and $\langle I_4\rangle$ are well-known functions of p and q, given in Eqs. (3.11) and (3.9), and $\langle kL \| V^L \| k'L'\rangle$ is proportional to $\langle kL \| Q \| k'L'\rangle$ according to Eq. (4.3).

A second class of relations is deduced from the commutation relations (2.5c), (3.3c), and (3.4) of the operators Q_μ and V^Q_μ with themselves. Taking the semireduced matrix elements of both sides of these equations, we get

$$[1-(-1)^\lambda][(2\lambda+1)(2L+1)]^{1/2}\sum_{L'}(-1)^{\bar{L}+L'}\begin{Bmatrix} L & \bar{L} & \lambda \\ 2 & 2 & L' \end{Bmatrix}\sum_{k'}\langle kL\|Q\|k'L'\rangle\langle \bar{k}\bar{L}\|Q\|k'L'\rangle = -[10L(L+1)]^{1/2}\delta_{\lambda 1}\delta_{\bar{k}k}\delta_{\bar{L}L}, \quad (6.4)$$

$$[(2\lambda+1)(2L+1)]^{1/2}\sum_{L'}(-1)^{\bar{L}+L'}\begin{Bmatrix} L & \bar{L} & \lambda \\ 2 & 2 & L' \end{Bmatrix}\sum_{k'}[\langle kL\|V^Q\|k'L'\rangle\times\langle \bar{k}\bar{L}\|Q\|k'L'\rangle - (-1)^\lambda\langle kL\|Q\|k'L'\rangle\langle \bar{k}\bar{L}\|V^Q\|k'L'\rangle] = -\sqrt{10}\,\delta_{\lambda 1}\langle \bar{k}\bar{L}\|V^L\|kL\rangle, \quad (6.5)$$

and

$$[1-(-1)^\lambda]\sum_{L'}(-1)^{\bar{L}+L'}\begin{Bmatrix} L & \bar{L} & \lambda \\ 2 & 2 & L' \end{Bmatrix}\sum_{k'}\langle kL\|V^Q\|k'L'\rangle\langle \bar{k}\bar{L}\|V^Q\|k'L'\rangle = [1-(-1)^\lambda]\sum_{L'}(-1)^{\bar{L}+L'}\begin{Bmatrix} L & \bar{L} & \lambda \\ 1 & 2 & L' \end{Bmatrix}\Big\{a_\lambda\sum_{k'}\langle kL\|Q\|k'L'\rangle\langle \bar{k}\bar{L}\|V^L\|k'L'\rangle + b_\lambda\,\delta_{L'\bar{L}}[\bar{L}(\bar{L}+1)]^{1/2}\langle kL\|V^Q\|\bar{k}\bar{L}\rangle\Big\}, \quad (6.6)$$

where the semireduced matrix elements of V^L and V^Q are given by Eqs. (4.3) and (4.4) respectively, and the coefficients a_λ and b_λ by Eq. (3.5).

Equations (6.1)–(6.6) can now be used as the initiating stage to determine some quadratic sum rules satisfied by the semireduced matrix elements of Q and V_Q. We now proceed to show how these sum rules can be defined and calculated explicitly.

7. QUADRATIC SUM RULES SATISFIED BY THE SEMIREDUCED MATRIX ELEMENTS OF Q AND V^Q

Let us consider the following quantities

$$\langle L'|Q^2|L\rangle \equiv \sum_{kk'}\langle k'L'\|Q\|kL\rangle^2, \quad (7.1)$$

$$\langle L'|QV^Q|L\rangle \equiv \sum_{kk'}\langle k'L'\|Q\|kL\rangle\langle k'L'\|V^Q\|kL\rangle, \quad (7.2)$$

and

$$\langle L'|(V^Q)^2|L\rangle \equiv \sum_{kk'}\langle k'L'\|V^Q\|kL\rangle^2, \quad (7.3)$$

summed over the additional quantum numbers k and k'. By definition their values are of course independent of the way these quantum numbers have been defined. By making $\bar{L}=L$, $\bar{k}=k$, and summing over k in Eqs. (6.1)–(6.6), we get a system of three equations for each of the three sum rules (7.1), (7.2), and (7.3). As each of these systems can be handled in the same way, we hereafter write only the detailed derivation of the first sum rule $\langle L'|Q^2|L\rangle$.

Let us note first that from the symmetry relation (4.5), it follows that

$$\langle L'|Q^2|L\rangle = (2L'+1)^{-1}(2L+1)\langle L|Q^2|L'\rangle. \quad (7.4)$$

Therefore, the five sum rules $\langle L'|Q^2|L\rangle$, corresponding to a given value of L and $L'=L\pm 2, L\pm 1, L$, belong to three classes according to the value of $|\Delta L| = |L'-L| = 2, 1$ or 0. Taking Eq. (7.4) into account, Eq. (6.1) and Eq. (6.4) with $\lambda=1$ and 3 lead to the following relations:

$$(2L-3)\langle L-2|Q^2|L\rangle + (2L-1)\langle L-1|Q^2|L\rangle + (2L+1)\langle L|Q^2|L\rangle + (2L+1)\langle L|Q^2|L+1\rangle + (2L+1)\times\langle L|Q^2|L+2\rangle = (2L+1)[\langle I_2\rangle - L(L+1)]N_L(pq), \quad (7.5)$$

$$-2(L+1)(2L-3)\langle L-2|Q^2|L\rangle - (L+3)(2L-1)\langle L-1|Q^2|L\rangle - 3(2L+1)\langle L|Q^2|L\rangle + (L-2)(2L+1)\times\langle L|Q^2|L+1\rangle + 2L(2L+1)\langle L|Q^2|L+2\rangle = -5L(L+1)(2L+1)N_L(pq), \quad (7.6)$$

and

$$-(L+1)(L+2)(2L-3)(2L+3)\langle L-2|Q^2|L\rangle + 2(L-2)(L+2)(2L-1)(2L+3)\langle L-1|Q^2|L\rangle + 12(L-1)$$

$$\times(L+2)(2L+1)\langle L\,|Q^2|L\rangle-2(L-1)(L+3)(2L-1)(2L+1)\langle L\,|Q^2|L+1\rangle+(L-1)L(2L-1)$$
$$\times(2L+1)\langle L\,|Q^2|L+2\rangle=0. \tag{7.7}$$

These relations are valid for any value of L such that $0\leqslant L\leqslant p$.

With the purpose of getting a recursion relation for the sum rules with $|\Delta L|=2$, we first eliminate $\langle L\,|Q^2|L\rangle$ between Eqs. (7.5) and (7.6), and Eqs. (7.6) and (7.7). The first resulting equation is then summed over L going from $\bar{L}$ to p, while the second one (for $L\geqslant 2$) is first multiplied by $(-1)^{L-\bar{L}}[(L-1)L(L+1)(L+2)]^{-1}$, and then summed over L between the same limits. In this way they become

$$(2L-3)(2L-1)\langle L-2|Q^2|L\rangle+(2L-1)(2L+1)\langle L-1|Q^2|L+1\rangle+L(2L-1)\langle L-1|Q^2|L\rangle$$
$$=A_L^{(1)}(pq), \tag{7.8}$$

and

$$-(L+1)(2L-3)(2L-1)\langle L-2|Q^2|L\rangle+(L-1)(2L-1)(2L+1)\langle L-1|Q^2|L+1\rangle$$
$$-2L(2L-1)\langle L-1|Q^2|L\rangle=-(L-1)L(L+1)B_L^{(1)}(pq),\quad L\geqslant 2, \tag{7.9}$$

where we have changed $\bar{L}$ into L, and we have defined

$$A_L^{(1)}(pq)=\sum_{L'=L}^{p}\{(2L'+1)[8L'(L'+1)-3\langle I_2\rangle]N_{L'}(pq)\},\quad L\geqslant 0, \tag{7.10}$$

$$B_L^{(1)}(pq)=4\sum_{L'=L}^{p}[(-1)^{L'-L}(2L'+1)N_{L'}(pq)],\quad L\geqslant 0. \tag{7.11}$$

It is now easy to eliminate $\langle L-1|Q^2|L\rangle$ between Eqs. (7.8) and (7.9) and to get the desired recursion relation

$$-(L-1)(2L-3)(2L-1)\langle L-2|Q^2|L\rangle+(L+1)(2L-1)(2L+1)\langle L-1|Q^2|L+1\rangle$$
$$=2A_L^{(1)}(pq)-(L-1)L(L+1)B_L^{(1)}(pq),\quad L\geqslant 2. \tag{7.12}$$

Its solution is obtained by changing L into L', multiplying the equation by L', and summing over L' going from L to p:

$$\langle L-2|Q^2|L\rangle=[(L-1)L(2L-3)(2L-1)]^{-1}\sum_{L'=L}^{p}\{L'[(L'-1)L'(L'+1)B_{L'}^{(1)}(pq)-2A_{L'}^{(1)}(pq)]\},\ L\geqslant 2. \tag{7.13}$$

When all the sum rules with $|\Delta L|=2$ have been determined, the remaining ones, with $|\Delta L|=1$ or 0, can be calculated in terms of them. We have indeed

$$\langle L-1|Q^2|L\rangle=-[2(2L+1)]^{-1}(L+5)(2L-3)\langle L-2|Q^2|L\rangle$$
$$+[2L(2L-1)]^{-1}(L-1)(L+2)(2L+3)\langle L\,|Q^2|L+2\rangle+C_L^{(1)}(pq),\quad L\geqslant 2, \tag{7.14}$$

and

$$\langle L\,|Q^2|L\rangle=[2(L+1)(2L+1)]^{-1}(L-1)(2L-3)(2L+5)\langle L-2|Q^2|L\rangle$$
$$-[2L]^{-1}(L+2)(2L-3)\langle L\,|Q^2|L+2\rangle+D_L^{(1)}(pq),\quad L\geqslant 1, \tag{7.15}$$

where

$$C_L^{(1)}(pq)=[L(2L-1)]^{-1}(L-1)[2L(L+1)(L+6)-3\langle I_2\rangle]N_L(pq),\quad L\geqslant 2, \tag{7.16}$$

and

$$D_L^{(1)}(pq)=-[L(L+1)]^{-1}[3L(L+1)(L^2+L-4)-(L^2+L-3)\langle I_2\rangle]N_L(pq),\quad L\geqslant 1. \tag{7.17}$$

The sum rules $\langle L'|QV^Q|L\rangle$ and $\langle L'|(V^Q)^2|L\rangle$ are given by relations similar to Eqs. (7.13), (7.14), and (7.15), with $A_L^{(1)}(pq)$, $B_L^{(1)}(pq)$, $C_L^{(1)}(pq)$, and $D_L^{(1)}(pq)$, replaced by $A_L^{(i)}(pq)$, $B_L^{(i)}(pq)$, $C_L^{(i)}(pq)$, $D_L^{(i)}(pq)$, where $i=2$ and 3 respectively. The explicit expressions of these new quantities can be found in Appendix A.

The summations over L' in Eqs. (7.10), (7.11), and (7.13), as well as in the corresponding equations for the other sum rules, extend from L to p. They are most easily calculated for high values of L. When one is interested in low values of L, it is more convenient to have summations over L' going from zero to some upper limit determined by L. In Appendix B we show how to convert the summations in Eqs. (7.10), (7.11), and (7.13) into summations from zero to $L-1$.

Up to now all the relations obtained for the semireduced matrix elements of Q do not depend on the way the additional quantum number k is defined. From now on we choose to use for k the eigenvalue of the third-order labeling operator K, defined in Eq. (3.10). We will hereafter examine the consequences of this assumption for the semireduced matrix elements of Q and show that this choice enables us to determine some k-weighted quadratic sum rules for those quantities.

8. THIRD ORDER LABELING OPERATOR AND *k*-WEIGHTED QUADRATIC SUM RULES FOR THE SEMIREDUCED MATRIX ELEMENTS OF *Q*

When using the operator K defined in Eq. (3.10) as a labeling operator, the additional quantum number k can be related very simply to the corresponding diagonal semireduced matrix element of Q. If we take, indeed, the semireduced matrix element of both sides of Eq. (3.10) and take into account that the semireduced matrix elements of V^L are proportional to those of Q as shown in Eq. (4.3), we get

$$k\delta_{k'k} = 2[\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}\langle k'L\|Q\|kL\rangle. \tag{8.1}$$

We conclude therefore that with this choice of labeling operator the matrix elements of Q which are diagonal in L are also diagonal in k. Moreover, we see clearly the importance of the calculation of the semireduced matrix elements of Q for the solution of the state labeling problem. Introducing Eq. (8.1) into the symmetry relation (4.10), we obtain that

$$k^+ = -k, \tag{8.2}$$

i.e., the eigenvalue of K in the contragredient IR$[p, p-q]$, corresponding to k in the IR$[pq]$, is equal to $-k$. Equations (8.1) and (8.2) are well known relations, derived previously by other authors.[20,13]

We now show that the operator K gives rise to an additional relation, which is new to our knowledge, and leads to a simple connection between the nondiagonal matrix elements of Q and V^Q. This relation is a special case of a more general one, which writes

$$\Big[\sum_\alpha(-1)^\alpha L_\alpha L_{-\alpha}, T^Q_\mu\Big] = \Big[\sum_\alpha(-1)^\alpha L_\alpha T^L_{-\alpha}, Q_\mu\Big], \tag{8.3}$$

and is valid for any vector operator $T = (T^L, T^Q)$.

To prove Eq. (8.3), let us first transform it into

$$\sum_\alpha(-1)^\alpha\{L_\alpha[L_{-\alpha}, T^Q_\mu] + [L_\alpha, T^Q_\mu]L_{-\alpha}\} = \sum_\alpha(-1)^\alpha\{L_\alpha[T^L_{-\alpha}, Q_\mu] + [L_\alpha, Q_\mu]\,T^L_{-\alpha}\}, \tag{8.4}$$

and use the equalities

$$[L_\alpha, T^Q_\mu] = [T^L_\alpha, Q_\mu], \tag{8.5}$$

and

$$\sum_\alpha(-1)^\alpha L_\alpha T^L_{-\alpha} = \sum_\alpha(-1)^\alpha T^L_\alpha L_{-\alpha}, \tag{8.6}$$

which derive directly from Eqs. (3.3a) and (3.3b), written for T instead of V. We get the following relation

$$\sum_\alpha(-1)^\alpha T^L_\alpha Q_\mu L_{-\alpha} = \sum_\alpha(-1)^\alpha L_\alpha Q_\mu T^L_{-\alpha}, \tag{8.7}$$

which remains now to be proved. For this purpose, let us consider the identity

$$[I_2, T^Q_\mu] = 0, \tag{8.8}$$

valid for any operator T^Q_μ built from the generators of SU(3). By introducing the definition (3.6) of I_2 into this equation, and using Eq. (8.5) as well as the relation

$$\sum_\nu(-1)^\nu Q_\nu[Q_{-\nu}, T^Q_\mu] + \sum_\alpha(-1)^\alpha[L_\alpha, Q_\mu]T^L_{-\alpha} = \sum_\nu(-1)^\nu[Q_\nu, T^Q_\mu]Q_{-\nu} + \sum_\alpha(-1)^\alpha T^L_\alpha \times[L_{-\alpha}, Q_\mu] = 0, \tag{8.9}$$

which is a direct consequence of the commutation relations (3.3), we get

$$\sum_\alpha(-1)^\alpha\{L_\alpha[T^L_{-\alpha}, Q_\mu] + [T^L_\alpha, Q_\mu]L_{-\alpha} - [L_\alpha, Q_\mu]T^L_{-\alpha} - T^L_\alpha[L_{-\alpha}, Q_\mu]\} = 0. \tag{8.10}$$

It is now straightforward to show that this equation reduces to Eq. (8.7) when we write each commutator as $[A,B] = AB - BA$, and take Eq. (8.6) into account. This completes the proof of Eq. (8.3).

In particular, when T is replaced by V, we reach the following important relation

$$[L^2, V^Q_\mu] = [K, Q_\mu]. \tag{8.11}$$

Taking the semireduced matrix element of both sides of this equation, we get

$$(k'-k)\langle k'L'\|Q\|kL\rangle = [L'(L'+1) - L(L+1)] \times\langle k'L'\|V^Q\|kL\rangle. \tag{8.12}$$

We conclude that for $L' \neq L$, the semireduced matrix elements of V^Q are given in terms of those of Q by the following relation:

$$\langle k'L'\|V^Q\|kL\rangle = [L'(L'+1) - L(L+1)]^{-1} \times(k'-k)\langle k'L'\|Q\|kL\rangle, \quad L' \neq L. \tag{8.13}$$

As a consequence of Eq. (8.13), the sum rules (7.2) and (7.3) with $L' \neq L$ can be rewritten in terms of semireduced matrix elements of Q only:

$$\langle L'|QV^Q|L\rangle = [L'(L'+1) - L(L+1)]^{-1} \times\sum_{kk'}(k'-k)\langle k'L'\|Q\|kL\rangle^2, \quad L' \neq L, \tag{8.14}$$

$$\langle L'|(V^Q)^2|L\rangle = [L'(L'+1) - L(L+1)]^{-2} \times\sum_{kk'}(k'-k)^2\langle k'L'\|Q\|kL\rangle^2, \quad L' \neq L. \tag{8.15}$$

We can now proceed one step further and establish analytical expressions for the following k-weighted sum rules

$$\langle L'|Q^2\cdot|L\rangle \equiv \sum_{kk'} k\,\langle k'L'\|Q\|kL\rangle^2, \tag{8.16a}$$

and

$$\langle L'|\cdot Q^2|L\rangle \equiv \sum_{kk'} k'\langle k'L'\|Q\|kL\rangle^2, \tag{8.16b}$$

where the point on the right (left) of Q^2 on the left-hand side means that the weight k (k') refers to the ket (bra). From Eq. (4.5) those sum rules satisfy the symmetry relation

$$\langle L'|Q^2\cdot|L\rangle = (2L+1)(2L'+1)^{-1}\langle L|\cdot Q^2|L'\rangle. \tag{8.17}$$

It is straightforward to see that Eq. (8.14) can be rewritten as

$$\langle L'|QV^Q|L\rangle = [L'(L'+1)-L(L+1)]^{-1} \times[\langle L'|\cdot Q^2|L\rangle - \langle L'|Q^2\cdot|L\rangle], \quad L'\neq L. \tag{8.18}$$

It is shown in Appendix C that the sum rule $\langle L-1\mid Q^2\cdot|L\rangle$ satisfies a recursion relation, whose solution is

$$\begin{aligned}\langle L-1|Q^2\cdot|L\rangle &= [(L-1)L(L+1)(2L-1)]^{-1} \\ &\times\sum_{L'=L}^{P}\{L'(L'+1)(2L'+1)\times[-2(\tfrac{1}{3}L'(L'+1) \\ &\times(2L'-1)(2L'+3))^{1/2}(\langle I_2\rangle+3)\langle L'|Q|L'\rangle \\ &+(2L'-1)(2L'+3)\langle L'|QV^Q|L'\rangle - 2(L'+1) \\ &\times(L'+2)\langle L'|QV^Q|L'+1\rangle]\}.\end{aligned} \tag{8.19}$$

From it the other sum rules can be calculated by successively applying the following relations:

$$\langle L+1|Q^2\cdot|L\rangle = (2L+3)^{-1}(2L+1)\langle L|Q^2\cdot|L+1\rangle - 2(L+1)\langle L+1|QV^Q|L\rangle, \tag{8.20}$$

$$\begin{aligned}\langle L|Q^2\cdot|L\rangle &= (2L+1)^{-1}[-(L-1)(L+1) \\ &\times\langle L-1|Q^2\cdot|L\rangle + L(L+2)\langle L+1|Q^2\cdot|L\rangle \\ &+L(L+1)(2L+1)\langle L|QV^Q|L\rangle],\end{aligned} \tag{8.21}$$

$$\begin{aligned}\langle L+2|Q^2\cdot|L\rangle &= [2L(2L+1)(2L+5)]^{-1}\{(L-1) \\ &\times(L+2)(2L-1)(2L+3)\langle L-1|Q^2\cdot|L\rangle \\ &+3(L+2)(2L-1)(2L+1)\langle L|Q^2\cdot|L\rangle \\ &-L(2L+3)(2L^2+3L-8)\langle L+1|Q^2\cdot|L\rangle \\ &-2L(L+1)(L+2)(2L+1)(2L+3)[\tfrac{1}{3}L(L+1) \\ &\times(2L-1)(2L+3)]^{1/2}\langle L|Q|L\rangle\},\end{aligned} \tag{8.22}$$

and

$$\begin{aligned}\langle L-2|Q^2\cdot|L\rangle &= [2(L+1)(2L-3)(2L+1)]^{-1} \\ &\times\{(L+1)(2L-1)(2L^2+L-9) \\ &\times\langle L-1|Q^2\cdot|L\rangle \\ &+3(L-1)(2L+1)(2L+3)\langle L|Q^2\cdot|L\rangle \\ &-(L-1)(L+2)(2L-1)(2L+3) \\ &\times\langle L+1|Q^2\cdot|L\rangle \\ &-2(L-1)L(L+1)(2L-1)(2L+1) \\ &\times[\tfrac{1}{3}L(L+1)(2L-1) \\ &\times(2L+3)]^{1/2}\langle L|Q|L\rangle\}.\end{aligned} \tag{8.23}$$

Eq. (8.20) is a special case of Eq. (8.18), while the other equations are proved in Appendix C.

9. CALCULATION OF THE EIGENVALUES OF K AND THE MATRIX ELEMENTS OF Q IN THE PHYSICAL BASIS

All the sum rules which we have considered up to now can be evaluated in principle for any IR[pq] of SU(3) and any multiplicities $N_L(pq)$. It would be indeed quite straightforward to develop a computer code in order to compute them numerically whenever the dimensionality of the IR of SU(3) is too large to allow us to calculate them by hand.

It is, however, much more difficult, even for not too high dimensionalities, to get detailed information about the individual matrix elements of Q in a given IR of SU(3), i.e., to calculate the semireduced matrix elements $\langle k'L'\|Q\|kL\rangle^2$ for all the values of k, L, k', and L' contained in this IR. In this case the usefulness of our procedure is restricted by the magnitude of the multiplicities of the IR's of SO(3) contained in the IR of SU(3) considered. In this section we are going to show through an example how far we can go in that field. It should be remembered, however, that the computation of the individual matrix elements of Q is only a byproduct of our study, whose main goal was to calculate sum rules.

To begin with, let us turn to the determination of the eigenvalues of the third-order labeling operator K, or equivalently, to that of the diagonal matrix elements of Q owing to Eq. (8.1). The diagonal sum rules $\langle L|Q|L\rangle$, $\langle L|Q^2|L\rangle$, and $\langle L|Q^2\cdot|L\rangle$ are simply proportional to $S=\Sigma_{j=1}^{N_L}k_j$, $C=\Sigma_{j=1}^{N_L}k_j^2$, and $D=\Sigma_{j=1}^{N_L}k_j^3$, respectively, where the summations are restricted to those values of k corresponding to L. We are thus able to calculate the eigenvalues of K in all the cases where the multiplicity $N_L(pq)$ does not exceed three, with the results:

(i) If $N_L(pq)=2$,

$$k_{1,2}=\tfrac{1}{2}(S\pm\sqrt{2C-S^2}). \tag{9.1}$$

(ii) If $N_L(pq)=3$,

$$k_j=\tfrac{1}{3}\left(S+\sqrt{2(3C-S^2)}\cos\frac{a+2j\pi}{3}\right), \quad j=1,2,3, \tag{9.2}$$

where

$$\cos a=\sqrt{2}(3C-S^2)^{-3/2}(9D+2S^3-9SC). \tag{9.3}$$

In principle it is now possible to obtain also the nondiagonal matrix elements $\langle k'L'\|Q\|kL\rangle^2$ for all the values of k, L, k', and L' contained in the IR[pq] if all the multiplicities are less than or equal to three. However, when N_L and $N_{L'}$ are equal to two or three, the sum rule equations are not sufficient for that purpose. It is then necessary to use in addition Eqs. (6.1)–(6.6) for $\bar{k}=k$ and to solve them step by step for all the values of L and k. As an example we have worked out completely the case of the IR[10,5], for which the possible L values are (with their corresponding multiplicity written as an exponent) 10, 9, 8^2, 7^2, 6^3, 5^3, 4^2, 3^2, 2, and 1. We show the results in Table III. The IR[10,5] is self-contragredient so that we could use Eqs. (4.9) and (4.10) to simplify the calculations. In spite of this, the computation was rather tedious. Moreover, a systematic procedure for handling the system of equations does not seem to emerge, so that it might be difficult to write a computer program to calculate the matrix elements of Q for general IR's of SU(3) for which $N_L(pq)\leqslant 3$. Therefore, our procedure might not be suited to the calculation of matrix elements of Q for IR's of SU(3) whose dimensionality is much higher than that of [10,5] (equal to 216). For IR's which are not self-contragredient, the limit on the dimensionality should be still lower.

TABLE III. Eigenvalues k of $K=\Sigma_\alpha(-1)^\alpha L_\alpha V^L_{-\alpha}$ and semireduced matrix elements $\langle k'L'\|Q\|kL\rangle^2$ for the IR[10,5] of SU(3). Only the upper triangular matrix is shown; the lower one can be deduced from it by use of Eq. (4.5).

L'	L / k' \ k	10 0	9 0	8 $12\sqrt{209}$	8 $-12\sqrt{209}$	7 $12\sqrt{561}$	7 $-12\sqrt{561}$
10	0	0	$\frac{110}{9}$	$\frac{80}{9}$	$\frac{80}{9}$	0	0
9	0		0	$\frac{220}{19}$	$\frac{220}{19}$	$\frac{20}{3}$	$\frac{20}{3}$
8	$12\sqrt{209}$			$\frac{11}{10}$	0	$\frac{16685}{1428}+\frac{165}{28}\sqrt{\frac{57}{17}}$	$\frac{16685}{1428}-\frac{165}{28}\sqrt{\frac{57}{17}}$
8	$-12\sqrt{209}$				$\frac{11}{10}$	$\frac{16685}{1428}-\frac{165}{28}\sqrt{\frac{57}{17}}$	$\frac{16685}{1428}+\frac{165}{28}\sqrt{\frac{57}{17}}$
7	$12\sqrt{561}$					$\frac{891}{182}$	0
7	$-12\sqrt{561}$						$\frac{891}{182}$

L'	L / k' \ k	6 $8\sqrt{2769}$	6 $-8\sqrt{2769}$	6 0	5 $8\sqrt{4665}$	5 $-8\sqrt{4665}$	5 0
10	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
8	$12\sqrt{209}$	$\frac{108296}{32305}+\frac{136}{35}\sqrt{\frac{627}{923}}$	$\frac{108296}{32305}-\frac{136}{35}\sqrt{\frac{627}{923}}$	$\frac{202500}{15691}$	0	0	0
8	$-12\sqrt{209}$	$\frac{108296}{32305}-\frac{136}{35}\sqrt{\frac{627}{923}}$	$\frac{108296}{32305}+\frac{136}{35}\sqrt{\frac{627}{923}}$	$\frac{202500}{15691}$	0	0	0
7	$12\sqrt{561}$	$\frac{90184}{6461}+\frac{216}{7}\sqrt{\frac{187}{923}}$	$\frac{90184}{6461}-\frac{216}{7}\sqrt{\frac{187}{923}}$	$\frac{10780}{2769}$	$\frac{155872}{84903}+\frac{1632}{91}\sqrt{\frac{55}{5287}}$	$\frac{155872}{84903}-\frac{1632}{91}\sqrt{\frac{55}{5287}}$	$\frac{2704}{311}$
7	$-12\sqrt{561}$	$\frac{90184}{6461}-\frac{216}{7}\sqrt{\frac{187}{923}}$	$\frac{90184}{6461}+\frac{216}{7}\sqrt{\frac{187}{923}}$	$\frac{10780}{2769}$	$\frac{155872}{84903}-\frac{1632}{91}\sqrt{\frac{55}{5287}}$	$\frac{155872}{84903}+\frac{1632}{91}\sqrt{\frac{55}{5287}}$	$\frac{2704}{311}$

L'	L / k' \ k	6 $8\sqrt{2769}$	6 $-8\sqrt{2769}$	6 0	5 $8\sqrt{4665}$	5 $-8\sqrt{4665}$	5 0
6	$8\sqrt{2769}$	$\frac{7384}{385}$	0	0	$\frac{353746035}{26121823}+\frac{295245}{91}\sqrt{\frac{5}{287053}}$	$\frac{353746035}{26121823}-\frac{295245}{91}\sqrt{\frac{5}{287053}}$	$\frac{450560}{287053}$
6	$-8\sqrt{2769}$		$\frac{7384}{385}$	0	$\frac{353746035}{26121823}-\frac{295245}{91}\sqrt{\frac{5}{287053}}$	$\frac{353746035}{26121823}+\frac{295245}{91}\sqrt{\frac{5}{287053}}$	$\frac{450560}{287053}$
6	0			0	$\frac{1531904}{18658445}$	$\frac{1531904}{18658445}$	$\frac{6997914}{287053}$
5	$8\sqrt{4665}$				$\frac{2488}{39}$	0	0
5	$-8\sqrt{4665}$					$\frac{2488}{39}$	0

TABLE III (cont.)

5	0						0

L'	k' \ L, k	4, $12\sqrt{105}$	4, $-12\sqrt{105}$	3, $4\sqrt{2145}$	3, $-4\sqrt{2145}$	2, 0	1, 0
6	$8\sqrt{2769}$	$\frac{606360}{131989} + \frac{3240}{143}\sqrt{\frac{35}{923}}$	$\frac{606360}{131989} - \frac{3240}{143}\sqrt{\frac{35}{923}}$	0	0	0	0
6	$-8\sqrt{2769}$	$\frac{606360}{131989} - \frac{3240}{143}\sqrt{\frac{35}{923}}$	$\frac{606360}{131989} + \frac{3240}{143}\sqrt{\frac{35}{923}}$	0	0	0	0
6	0	$\frac{921536}{59995}$	$\frac{921536}{59995}$	0	0	0	0
5	$8\sqrt{4665}$	$\frac{1756}{3421} + \frac{36}{11}\sqrt{\frac{7}{311}}$	$\frac{1756}{3421} - \frac{36}{11}\sqrt{\frac{7}{311}}$	$\frac{206948}{51315} + \frac{196}{3}\sqrt{\frac{13}{3421}}$	$\frac{206948}{51315} - \frac{196}{3}\sqrt{\frac{13}{3421}}$	0	0
5	$-8\sqrt{4665}$	$\frac{1756}{3421} - \frac{36}{11}\sqrt{\frac{7}{311}}$	$\frac{1756}{3421} + \frac{36}{11}\sqrt{\frac{7}{311}}$	$\frac{206948}{51315} - \frac{196}{3}\sqrt{\frac{13}{3421}}$	$\frac{206948}{51315} + \frac{196}{3}\sqrt{\frac{13}{3421}}$	0	0
5	0	$\frac{6240}{311}$	$\frac{6240}{311}$	$\frac{23328}{3421}$	$\frac{23328}{3421}$	0	0

L'	k' \ L, k	4, $12\sqrt{105}$	4, $-12\sqrt{105}$	3, $4\sqrt{2145}$	3, $-4\sqrt{2145}$	2, 0	1, 0
4	$12\sqrt{105}$	$\frac{81}{11}$	0	$\frac{1513}{90} + \frac{1}{2}\sqrt{1001}$	$\frac{1513}{90} - \frac{1}{2}\sqrt{1001}$	$\frac{160}{9}$	0
4	$-12\sqrt{105}$		$\frac{81}{11}$	$\frac{1513}{90} - \frac{1}{2}\sqrt{1001}$	$\frac{1513}{90} + \frac{1}{2}\sqrt{1001}$	$\frac{160}{9}$	0
3	$4\sqrt{2145}$			$\frac{143}{3}$	0	$\frac{32}{7}$	$\frac{64}{7}$
3	$-4\sqrt{2145}$				$\frac{143}{3}$	$\frac{32}{7}$	$\frac{64}{7}$
2	0					0	$\frac{286}{5}$
1	0						0

10. CONCLUSION

In this paper we have shown that it is possible to obtain analytical formulas for various sum rules satisfied by the matrix elements of the generators of SU(3) in an SO(3) basis. For that purpose we have chosen to use an orthogonal basis which diagonalizes an additional labeling operator K, and we have established our relations without explicitly constructing the basis states. As a byproduct of our study, we have been able to deduce general formulas for the eigenvalues of K when the multiplicity of the SO(3) IR's does not exceed three and to calculate the individual matrix elements of the SU(3) generators for the IR[10,5] of SU(3).

Were we concerned with the determination of the eigenvalues of K or the individual matrix elements of Q for general IR's of SU(3), it would be clear that our procedure is not the best one available. It is indeed well known that the eigenvalues of K can be calculated straightforwardly either in a nonorthogonal basis[8] or in the canonical orthogonal one.[13,14] Concerning the matrix elements of Q, the use of either basis would lead to a clear-cut and systematic procedure that could be easily programmed. Some works are in process along these lines using nonorthogonal basis.[21]

However the usefulness of our procedure is obvious as compared with other ones when considering sum rules for the matrix elements of Q. With a nonorthogonal basis, we should have to combine the overlap matrix of the basis states with both diagonal and nondiagonal matrix elements of Q. With the canonical orthogonal basis we should have to restore the SO(3) symmetry by using the transformation matrix from this basis to a physical orthogonal one. In both cases it is clear that no direct procedure would be available to calculate sum rules and that no analytical results could be obtained except in extremely simple cases. On the other hand, by working directly with a physical orthogonal basis as we do in the present paper, one is straightforwardly led to general analytical formulas for various sum rules, which we consider as the achievement of the present paper.

APPENDIX A: EXPLICIT EXPRESSIONS OF $A_L^{(i)}(pq)$, $B_L^{(i)}(pq)$, $C_L^{(i)}(pq)$, AND $D_L^{(i)}(pq)$ FOR $i = 2$ AND 3

In terms of

$$\langle L \mid V^L \mid L \rangle \equiv \sum_k \langle kL \| V^L \| kL \rangle = 2[\tfrac{1}{3}(2L-1)(2L+3)]^{1/2} \langle L \mid Q \mid L \rangle, \tag{A1}$$

and

$$\begin{aligned}\langle L \mid V^Q \mid L \rangle \equiv \sum_k \langle kL \| V^Q \| kL \rangle = [\tfrac{70}{3}(2L+1)]^{1/2} \sum_{L_1} (-1)^{L-L_1} \begin{Bmatrix} 2 & 2 & 2 \\ L & L & L_1 \end{Bmatrix} \langle L \mid Q^2 \mid L_1 \rangle \\ + [\tfrac{1}{3} L(L+1)(2L-1)(2L+3)]^{1/2} N_L(pq),\end{aligned} \tag{A2}$$

respectively, we get

$$A_L^{(2)}(pq) = \sum_{L'=L}^{p} \left\{(2L'+1)\left[8\sqrt{L'(L'+1)}\langle L' \mid V^L \mid L' \rangle - 3\langle I_3 \rangle N_{L'}(pq)\right]\right\}, \tag{A3}$$

$$B_L^{(2)}(pq) = 4 \sum_{L'=L}^{p} \{(-1)^{L'-L} [L'(L'+1)]^{-1/2} (2L'+1) \langle L' \mid V^L \mid L' \rangle\}, \tag{A4}$$

$$C_L^{(2)}(pq) = [L(2L-1)]^{-1}(L-1)\left[2(L+6)\sqrt{L(L+1)}\langle L \mid V^L \mid L \rangle - 3\langle I_3 \rangle N_L(pq)\right], \tag{A5}$$

$$D_L^{(2)}(pq) = -[L(L+1)]^{-1}\left[3(L^2+L-4)\sqrt{L(L+1)}\langle L \mid V^L \mid L \rangle - (L^2+L-3)\langle I_3 \rangle N_L(pq)\right], \tag{A6}$$

and

$$\begin{aligned}A_L^{(3)}(pq) = \sum_{L'=L}^{p} \{(2L'+1)[(3L'(L'+1)(2L'-1)(2L'+3))^{1/2} \langle L' \mid V^Q \mid L' \rangle \\ + \tfrac{26}{3}(3(2L'+1)^{-1}(L'-1)(L'+1)(2L'-1)\langle L'-1 \mid Q^2 \mid L' \rangle + (2L'-1)(2L'+3)\langle L' \mid Q^2 \mid L' \rangle \\ + 3L'(L'+2)\langle L' \mid Q^2 \mid L'+1 \rangle) - 3\langle I_4 \rangle N_{L'}(pq)]\},\end{aligned} \tag{A7}$$

$$\begin{aligned}B_L^{(3)}(pq) = \tfrac{8}{3} \sum_{L'=L}^{p} \{(-1)^{L'-L} [L'(L'+1)]^{-1} [3(L'-2)(L'+1)(2L'-1)\langle L'-1 \mid Q^2 \mid L' \rangle \\ + 2(2L'-1)(2L'+1)(2L'+3)\langle L' \mid Q^2 \mid L' \rangle + 3L'(L'+3)(2L'+1)\langle L' \mid Q^2 \mid L'+1 \rangle]\},\end{aligned} \tag{A8}$$

$$\begin{aligned}C_L^{(3)}(pq) = [L(2L-1)]^{-1}(L-1)\{[3L(L+1)(2L-1)(2L+3)]^{1/2} \langle L \mid V^Q \mid L \rangle + \tfrac{2}{3}[3(2L+1)^{-1} \\ \times (L+1)(2L-1)(2L^2+13L-21)\langle L-1 \mid Q^2 \mid L \rangle + (2L-1)(2L+3)(4L+21) \\ \times \langle L \mid Q^2 \mid L \rangle + 3L(L+2)(2L+19)\langle L \mid Q^2 \mid L+1 \rangle] - 3\langle I_4 \rangle N_L(pq)\},\end{aligned} \tag{A9}$$

$$\begin{aligned}D_L^{(3)}(pq) = [L(L+1)]^{-1}\{[3L(L+1)(2L-1)(2L+3)]^{1/2} \langle L \mid V^Q \mid L \rangle - 2[(2L+1)^{-1} \\ \times (L-1)(L+1)(2L-1)(4L^2+2L-21)\langle L-1 \mid Q^2 \mid L \rangle + (2L-1)(2L+3)(2L^2+2L-7) \\ \times \langle L \mid Q^2 \mid L \rangle + L(L+2)(4L^2+6L-19)\langle L \mid Q^2 \mid L+1 \rangle] + (L^2+L-3)\langle I_4 \rangle N_L(pq)\}.\end{aligned} \tag{A10}$$

APPENDIX B: ALTERNATIVE EXPRESSIONS FOR $A_L^{(1)}(pq)$, $B_L^{(1)}(pq)$, AND $\langle L-2 \mid Q^2 \mid L \rangle$

In this appendix we show how to convert the summations over L' going from L to p in Eqs. (7.10), (7.11), and (7.13) into summations from zero to $L-1$.

Let us begin with Eq. (7.10). It is obvious that it can be rewritten as

$$A_L^{(1)}(pq) = A_0^{(1)}(pq) - \sum_{L'=0}^{L-1} \{(2L'+1)[8L'(L'+1) - 3\langle I_2 \rangle] N_{L'}(pq)\}, \tag{B1}$$

where the value of $A_0^{(1)}(pq)$ can be found from Eq. (7.8) for $L = 0$. We get

$$A_0^{(1)}(pq) = 0, \tag{B2}$$

so that

$$A_L^{(1)}(pq) = -\sum_{L'=0}^{L-1} \{(2L'+1)[8L'(L'+1) - 3\langle I_2 \rangle] N_{L'}(pq)\}, \tag{B3}$$

which is the desired expression for $A_L^{(1)}(pq)$.

Eq. (7.11) can be transformed in the same way into

$$B_L^{(1)}(pq) = (-1)^L B_0^{(1)}(pq) - 4 \sum_{L'=0}^{L-1} (-1)^{L-L'} (2L'+1) N_{L'}(pq). \tag{B4}$$

However in this case, $B_0^{(1)}(pq)$ is not equal to zero because Eq. (7.9) is not valid for $L = 0$. We have therefore to calculate

$B_0^{(1)}(pq)$ from the definition (7.11). We obtain

$$B_0^{(1)}(pq) = 4 \sum_{L'=0}^{p} (-1)^{L'}(2L'+1)N_{L'}(pq), \tag{B5}$$

where $N_{L'}(pq)$ is determined by Racah's formula (3.14). It is straightforward to show that

$$\begin{aligned} B_0^{(1)}(pq) &= 2(p^2 - pq + q^2 + 3p + 2), && \text{when } p \text{ and } q \text{ are even,} \\ &= 2(p - q + 1)(q + 1), && \text{when } p \text{ is even and } q \text{ odd,} \\ &= -2(p - q + 1)(p + 2), && \text{when } p \text{ is odd and } q \text{ even,} \\ &= -2(p + 2)(q + 1), && \text{when } p \text{ and } q \text{ are odd.} \end{aligned} \tag{B6}$$

It now remains to transform the summation in Eq. (7.13). For this purpose, let us go back to Eq. (7.12) and consider only those values of L which are greater than 2. Let us change L into L', multiply the equation by L' and sum over L' going from 2 to $L-1$. We get

$$(L-1)L(2L-3)(2L-1)\langle L-2|Q^2|L\rangle - 6\langle 0|Q^2|2\rangle = -\sum_{L'=2}^{L-1}\{L'[(L'-1)L'(L'+1)B_{L'}^{(1)}(pq) - 2A_{L'}^{(1)}(pq)]\},$$
$$L \geqslant 3. \tag{B7}$$

The quantity $\langle 0|Q^2|2\rangle$ can now be evaluated directly from Eq. (7.8) for $L = 1$:

$$3\langle 0|Q^2|2\rangle = A_1^{(1)}(pq). \tag{B8}$$

Introducing this expression into Eq. (B7), we see that $\langle L-2|Q^2|L\rangle$ is given by a summation over L' going from 1 to $L-1$ (or equivalently from 0 to $L-1$ as the term corresponding to $L' = 0$ is equal to zero), and that the resulting formula is valid not only for $L \geqslant 3$ but also for $L = 2$:

$$\langle L-2|Q^2|L\rangle = -[(L-1)L(2L-3)(2L-1)]^{-1}\sum_{L'=0}^{L-1}\{L'[(L'-1)L'(L'+1)B_{L'}^{(1)}(pq) - 2A_{L'}^{(1)}(pq)]\}, \quad L \geqslant 2. \tag{B9}$$

Eqs. (B3), (B4) and (B9) are the alternative expressions we were looking for to replace Eqs. (7.10), (7.11), and (7.13) respectively.

APPENDIX C: CALCULATION OF THE SUM RULES $\langle L'|Q^2\cdot|L\rangle$

To calculate the sum rules $\langle L'|Q^2\cdot|L\rangle$, defined in Eq. (8.16a), we start from Eqs. (6.1) and (6.4). By equating L' to L and k' to k, multiplying by k, summing over k, and using Eqs. (8.1) and (8.17), we obtain a system of three equations for those sum rules, which writes

$$\begin{aligned} &(2L-3)\langle L-2|Q^2\cdot|L\rangle + (2L-1)\langle L-1|Q^2\cdot|L\rangle + (2L+1)\langle L|Q^2\cdot|L\rangle + (2L+3)\langle L+1|Q^2\cdot|L\rangle + (2L+5) \\ &\quad\times\langle L+2|Q^2\cdot|L\rangle = 2(2L+1)[\langle I_2\rangle - L(L+1)][\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}\langle L|Q|L\rangle, \end{aligned} \tag{C1}$$

$$\begin{aligned} &-2(L+1)(2L-3)\langle L-2|Q^2\cdot|L\rangle - (L+3)(2L-1)\langle L-1|Q^2\cdot|L\rangle - 3(2L+1)\langle L|Q^2\cdot|L\rangle \\ &\quad + (L-2)(2L+3)\langle L+1|Q^2\cdot|L\rangle + 2L(2L+5)\langle L+2|Q^2\cdot|L\rangle \\ &\qquad = -10L(L+1)(2L+1)[\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}\langle L|Q|L\rangle, \end{aligned} \tag{C2}$$

and

$$\begin{aligned} &-(L+1)(L+2)(2L-3)(2L+3)\langle L-2|Q^2\cdot|L\rangle + 2(L-2)(L+2)(2L-1)(2L+3)\langle L-1|Q^2\cdot|L\rangle \\ &\quad + 12(L-1)(L+2)(2L+1)\langle L|Q^2\cdot|L\rangle - 2(L-1)(L+3)(2L-1)(2L+3)\langle L+1|Q^2\cdot|L\rangle \\ &\quad + (L-1)L(2L-1)(2L+5)\langle L+2|Q^2\cdot|L\rangle = 0. \end{aligned} \tag{C3}$$

From them we can express $\langle L\pm 2|Q^2\cdot|L\rangle$ in terms of the other sum rules. The resulting expressions are respectively contained in Eqs. (8.22) and (8.23). By eliminating $\langle L\pm 2|Q^2\cdot|L\rangle$, we are left with a single relation between the remaining sum rules:

$$\begin{aligned} &(L-1)(L+1)(L+3)(2L-1)\langle L-1|Q^2\cdot|L\rangle + (2L-1)(2L+1)(2L+3)\langle L|Q^2\cdot|L\rangle - (L-2)L(L+2)(2L+3) \\ &\quad\times\langle L+1|Q^2\cdot|L\rangle = 2L(L+1)(2L+1)[\langle I_2\rangle + 3][\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}\langle L|Q|L\rangle. \end{aligned} \tag{C4}$$

We can get two additional relations by making use of the known expressions of $\langle L|QV^Q|L\rangle$ and $\langle L+1|QV^Q|L\rangle$. On the one hand from Eqs. (8.1), (8.17), (8.22), and (8.23), we obtain successively

$$\langle L|QV^Q|L\rangle = \tfrac{1}{2}[\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{-1/2}\sum_k k\langle kL\|V^Q\|kL\rangle$$

$$= [L(L+1)(2L+1)]^{-1}[(L-1)(L+1)\langle L-1|Q^{2\cdot}|L\rangle + (2L+1)\langle L|Q^{2\cdot}|L\rangle - L(L+2)\langle L+1|Q^{2\cdot}|L\rangle]. \quad \text{(C5)}$$

This equation leads to the expression of $\langle L|Q^{2\cdot}|L\rangle$ in terms of $\langle L\pm 1|Q^{2\cdot}|L\rangle$ and $\langle L|QV^Q|L\rangle$, contained in Eq. (8.21). On the other hand, when applied to the case $L'=L+1$, Eqs. (8.17) and (8.18) give rise to the expression of $\langle L+1|Q^{2\cdot}|L\rangle$ in terms of $\langle L|Q^{2\cdot}|L+1\rangle$ and $\langle L+1|QV^Q|L\rangle$, contained in Eq. (8.20).

It now remains to introduce both expressions into Eq. (C5) such as to obtain a recursion relation for $\langle L-1|Q^{2\cdot}|L\rangle$ writing

$$(L-1)L(L+1)(2L-1)\langle L-1|Q^{2\cdot}|L\rangle - L(L+1)(L+2)(2L+1)\langle L|Q^{2\cdot}|L+1\rangle$$
$$= L(L+1)(2L+1)\{-2[\tfrac{1}{3}L(L+1)(2L-1)(2L+3)]^{1/2}[\langle I_2\rangle+3]\langle L|Q|L\rangle + (2L-1)(2L+3)\langle L|QV^Q|L\rangle - 2(L+1)(L+2)\langle L|QV^Q|L+1\rangle\}. \quad \text{(C6)}$$

Its solution is easily found by replacing L by L' and summing over L' going from L to p. It is given in Eq. (8.19). This completes the derivation of the sum rules $\langle L'|Q^{2\cdot}|L\rangle$.

[1]I.M. Gel'fand and M.L. Tseitlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).
[2]M. Moshinsky, J. Math. Phys. **4**, 1128 (1963).
[3]L.C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G. Baird and L.C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[4]J. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
[5]M. Moshinsky, *Group Theory and the Many Body Problem* (Gordon and Breach, New York, 1968).
[6]M. Moshinsky, J. Patera, R.T. Sharp, and P. Winternitz, Ann. Phys. (N.Y.) **95**, 139 (1975).
[7]J.P. Elliott, Proc. Roy. Soc. A **245**, 128, 562 (1958).
[8]V. Bargmann and M. Moshinsky, Nucl. Phys. **18**, 967 (1960); **23**, 177 (1961).
[9]M. Moshinsky and V. Syamala Devi, J. Math. Phys. **10**, 455 (1969).
[10]R.T. Sharp and C.S. Lam, J. Math. Phys. **10**, 2033 (1969).
[11]G. Racah, in *Group Theoretical Concepts and Methods in Elementary Particle Physics*, edited by F. Gürsey (Gordon and Breach, New York, 1964); Y.L. Ilamed, in *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch, S.G. Cohen, A. de-Shalit, S. Sambursky, and I. Talmi (North-Holland, Amsterdam, 1968).
[12]A. Peccia and R.T. Sharp, J. Math. Phys. **17**, 1313 (1976).
[13]B.R. Judd, W. Miller Jr., J. Patera, and P. Winternitz, J. Math. Phys. **15**, 1787 (1974).
[14]W. Mc Kay, J. Patera, and R.T. Sharp, Comp. Phys. Comm. **10**, 1 (1975).
[15]J.B. French, in *Nuclear Structure*, edited by A. Hossain, Harun-ar-Rachid, and M. Islam (North-Holland, Amsterdam, 1967), p. 124.
[16]Application of the method of Ref. 20 to the problem studied by A. Pais, Phys. Rev. D **10**, 2147 (1974), gives rise to the need of sum rules.
[17]A. Partensky and C. Maguin, J. Math. Phys. **19**, 511 (1978).
[18]J.W.B. Hughes, J. Phys. A **6**, 48, 281 (1973).
[19]G. Racah, Rev. Mod. Phys. **21**, 494 (1949).
[20]M. Moshinsky, J. Patera, R.T. Sharp, and P. Winternitz, Phys. Rev. D **10**, 1587 (1974).
[21]R.T. Sharp, private communication.

Clebsch–Gordan coefficients of finite magnetic groups

P. M. van den Broek
Institute for Theoretical Physics, University of Nijmegen, Nijmegen, The Netherlands

(Received 24 January 1979; accepted for publication 13 April 1979)

A detailed method is given for the calculation of Clebsch–Gordan coefficients of finite magnetic groups. This method is a generalization of a new method for the calculation of Clebsch–Gordan coefficients of finite nonmagnetic groups which makes use of the fact that the Clebsch–Gordan coefficients may be arranged into vectors which are eigenvectors of certain projection matrices.

1. INTRODUCTION

According to the well-known results of Bargmann [1] and Wigner [2] the Hilbert space of state vectors of a physical system with a symmetry group G carriers a projective unitary–antiunitary (PUA) representation of G. All relevant physical information, such as transition probabilities, is contained in matrix elements of irreducible tensor operators which transform according to irreducible PUA representations of G. An important tool for the calculation of these matrix elements is the Wigner–Eckart theorem, which has been generalized for finite magnetic groups by Aviran and Zak. [3] To apply the Wigner–Eckart theorem one needs the Clebsch–Gordan coefficients of G.

Recently a step forward in the theory of Clebsch–Gordan coefficients of nonmagnetic groups has been made by the observation that columns of Clebsch–Gordan coefficients are eigenvectors of some projection matrices. This observation was made first by Schindler and Mirman [4] for the symmetric groups. A detailed procedure for the construction of the Clebsch–Gordan coefficients was then given by van den Broek and Cornwell [5] and Dirl. [6] The advantage of this method over previous methods is that the calculations are straightforward, problems related to multiplicity (overlooked for instance by Cornwell [7] and Berenson and Birman [8]) are properly taken care of and tedious ad hoc methods (Sakata [9]) and orthogonalization procedures (Koster [10]) are avoided.

The present paper extends the method of Ref. 5 to magnetic groups. It will turn out that as in the case of nonmagnetic groups the columns of Clebsch–Gordan coefficients are eigenvectors of projection matrices but that some additional constraints have to be imposed on these eigenvectors which depend on the type (I, II, or III) of UA representations in the decomposition. I will show that these three types of constraints can easily be dealt with.

Relations for the Clebsch–Gordan coefficients of finite magnetic groups have been given by Aviran and Litvin, [11] Rudra [12] and Kotsev. [13] However, none of these authors proves that these relations are sufficient, i.e., that each solution indeed provides a set of Clebsch–Gordan coefficients. Also, a systematic way of solving these relations has not been given.

Different approaches to the problem of calculating Clebsch–Gordan coefficients for finite magnetic groups have been given by Gard and Backhouse [14] and Sakata. [15] Gard and Backhouse [14] assume that the Clebsch–Gordan coefficients of the nonmagnetic subgroup are known and give a method to construct from them the Clebsch–Gordan coefficients of the whole group. Sakata [15] obtains a collection of matrices which fulfill the requirements to be matrices of Clebsch–Gordan coefficients except for being unitary (or even nonsingular). The problem then remains to find a unitary matrix among this collection of matrices.

In Sec. 3 we will discuss the relations for the Clebsch–Gordan coefficients of finite magnetic groups of Ref. 11 and Ref. 13 and show that these relations are indeed sufficient. Moreover, we will show that the solutions of these relations automatically satisfy the orthogonality relations, which means that each solution gives a matrix of Clebsch–Gordan coefficients which is unitary. The main part of this paper will be Sec. 4, where we will derive a straightforward method to solve the equations of Sec. 3. A summary of this method will be given in Sec. 5, especially for those readers who wish to use the method without bothering for the proofs and the derivations. In Sec. 6 we will give a derivation of the Wigner–Eckart theorem for finite magnetic groups which is different from the derivation of Aviran and Zak [3] and which shows the close connection between the form of the Wigner–Eckart theorem and the lemma of Schur for finite magnetic groups.

Although one has in general to consider PUA representation of G, we will restrict ourselves ot UA representations in this paper. The generalization to PUA representations, however, is trivial; all results hold for PUA representations as well.

2. DEFINITIONS AND PRELIMINARIES

Let G be a finite group and G_0 a subgroup of G of index 2. A unitary–antiunitary (UA) representation of G with respect to G_0 is a homomorphic mapping T from G into the group of unitary and antiunitary operators on some Hilbert space $\mathscr{H}$ such that $T(g)$ is unitary if $g\in G_0$ and $T(g)$ is antiunitary if $g\notin G_0$. In the following we shall drop the phrase "with respect to G_0." Suppose $\mathscr{H}$ is finite dimensional and let $\{\phi_1,...,\phi_n\}$ be an orthonormal basis of $\mathscr{H}$. For each operator $T(g)$ a unitary $n\times n$ matrix $D(g)$ is defined by $D_{ij}(g)=(\phi_i,T(g)\phi_j)$. The matrices $D(g)$ satisfy

$$D(g)D^{g}(g')=D(gg'),\quad \forall g,g'\in G, \tag{2.1}$$

where D^{g} is defined by

 0022-2488/79/102028-08$01.00

$$D^g = \begin{cases} D & \text{if } g \in G_0 \\ D^* & \text{if } g \notin G_0 \end{cases}. \tag{2.2}$$

A UA representation of G can also be defined to be a mapping from G into the unitary matrices of some dimension n such that Eq. (2.1) holds. The connection with the previous definition lies in the choice of the basis $\{\phi_1,...,\phi_n\}$ of $\mathscr{H}$. In the sequel T and D will always denote an operator and a matrix, respectively.

A UA representation T of G is reducible if these exists a proper subspace of $\mathscr{H}$ which is invariant under $T(G)$; otherwise T is irreducible. Two UA representations T_1 and T_2 of G in the Hilbert spaces H_1 and H_2 respectively are equivalent if these exists a unitary mapping $U{:}H_2 \rightarrow H_1$ such that

$$T_1(g)U = UT_2(g), \quad \forall g \in G, \tag{2.3a}$$

or, equivalently, in terms of matrices, two UA representations D_1 and D_2 of G are equivalent if there exists a unitary matrix U such that

$$D_1(g)U^g = UD_2(g), \quad \forall g \in G. \tag{2.3b}$$

According to Wigner [16] the irreducible UA representations are divided into three types: the restriction $D \downarrow G_0$ of an irreducible UA representation D of G to G_0, which is a unitary representation of G_0, is irreducible (Type I), is reducible into two equivalent irreducible components (Type II), or is reducible into two inequivalent irreducible components (Type III).

Let a be a fixed element of $G \setminus G_0$. An irreducible UA representation of G of type II is equivalent with a UA representation which has the form

$$D(g) = \begin{pmatrix} \Delta(g) & 0 \\ 0 & \Delta(g) \end{pmatrix} \quad \forall g \in G_0; \quad D(a) = \begin{pmatrix} 0 & U \\ -U & 0 \end{pmatrix}, \tag{2.4}$$

where Δ is an irreducible unitary representation of G_0 and U satisfies

$$UU^* = -\Delta(a^2) \tag{2.5}$$

and

$$U\Delta^*(a^{-1}ga)U^{-1} = \Delta(g) \quad \forall g \in G_0. \tag{2.6}$$

An irreducible UA representation of G of type III is equivalent with a UA representation which has the form

$$D(g) = \begin{pmatrix} \Delta(g) & 0 \\ 0 & \Delta^*(a^{-1}ga) \end{pmatrix} \quad \forall g \in G_0;$$

$$D(a) = \begin{pmatrix} 0 & \Delta(a^2) \\ \mathbb{1} & 0 \end{pmatrix}, \tag{2.7}$$

where Δ is an irreducible unitary representation of G_0 which is not equivalent with $\bar{\Delta}$ defined by $\bar{\Delta}(g) = \Delta^*(a^{-1}ga)$.

An irreducible UA representation is said to be in standard form if it is of type I, if it is of type II and satisfies Eq. (2.4) or if it is of type III and satisfies Eq. (2.7). So each irreducible UA representation is equivalent with an irreducible UA representation which is in standard form. From now on we will only consider one fixed chosen representative which is in standard form of each class of equivalent irreducible UA represenatations. These representatives will be denoted by D^α, D^β,..., and their dimensions by $d_\alpha, d_\beta, \cdots$.

The lemma of Schur may be generalized for irreducible UA representations as follows. Let D be an irreducible UA representation of G which is in standard form and let U be a matrix which satisfies

$$D(g)U^g = UD(g) \quad \forall g \in G. \tag{2.8}$$

Then

$$U = \lambda \mathbb{1}; \lambda \in \mathbb{R} \quad \text{if } D \text{ is of type I}, \tag{2.9a}$$

$$U = \begin{pmatrix} \lambda \mathbb{1} & \mu \mathbb{1} \\ -\mu^* \mathbb{1} & \lambda^* \mathbb{1} \end{pmatrix}; \lambda, \mu \in \mathbb{C} \text{ if } D \text{ is of type II}, \tag{2.9b}$$

and

$$U = \begin{pmatrix} \lambda \mathbb{1} & 0 \\ 0 & \lambda^* \mathbb{1} \end{pmatrix}; \lambda \in \mathbb{C} \text{ if } D \text{ is of type III} \tag{2.9c}$$

($\mathbb{1}$ always denotes a unit matrix).

From this the following orthogonality relations may be derived [13,17]: If D^γ is of type I,

$$\frac{d_\gamma}{|G|} \sum_{g \in G_0} D^{\gamma *}(g)_{pq} D^{\eta}(g)_{kl} = \tfrac{1}{2} \delta_{pk} \delta_{ql} \delta_{\gamma\eta}, \tag{2.10a}$$

$$\frac{d_\gamma}{|G|} \sum_{g \notin G_0} D^{\gamma *}(g)_{pq} D^{\eta}(g)_{kl} = \tfrac{1}{2} \delta_{pk} \delta_{ql} \delta_{\gamma\eta}. \tag{2.10b}$$

If D^γ is of type II,

$$\frac{d_\gamma}{|G|} \sum_{g \in G_0} D^{\gamma *}(g)_{pq} D_{\eta}(g)_{kl} = \delta_{\gamma\eta} (\delta_{pk} \delta_{ql} + \delta_{p,k+(d_\gamma/2)} \delta_{q,l+(d_\gamma/2)}), \quad (p \sim q), \tag{2.11a}$$

$$\frac{d_\gamma}{|G|} \sum_{g \notin G_0} D^{\gamma *}(g)_{pq} D^{\eta}(g)_{kl} = \delta_{\gamma\eta} (\delta^{pk} \delta_{ql} - \delta_{p,k+(d_\gamma/2)} \delta_{q,l+(d_\gamma/2)}), \quad (p \sim q). \tag{2.11b}$$

If D^γ is of type III

$$\frac{d_\gamma}{|G|} \sum_{g \in G_0} D^{\gamma *}(g)_{pq} D^{\eta}(g)_{kl} = \delta_{pk} \delta_{ql} \delta_{\gamma\eta}, \quad (p \sim q), \tag{2.12a}$$

$$\frac{d_\gamma}{|G|} \sum_{g \notin G_0} D^{\gamma *}(g)_{pq} D^{\eta}(g)_{kl} = \delta_{pk} \delta_{ql} \delta_{\gamma\eta}, \quad (p \not\sim q). \tag{2.12b}$$

Here $|G|$ is the order of G; $p \sim q$ means $1 \leqslant p,q \leqslant (d_\gamma/2)$ or $(d_\gamma/2) + 1 \leqslant p,q \leqslant d_\gamma$. In Eq.(2.11) and in the remainder of this paper indices of the type $k + d_\gamma/2$ should be taken modulo d_γ. It should be noted that the orthogonality relations in this form do not hold in general if the UA representations are not in standard form.

The direct product $D^\alpha \otimes D^\beta$ of two irreducible UA representations D^α and D^β of G is defined by

$$(D^\alpha \otimes D^\beta)(g)_{ij,kl} = D^\alpha(g)_{ik} D^\beta(g)_{jl}, \quad \forall g \in G. \tag{2.13}$$

In general $D^\alpha \otimes D^\beta$ is a reducible UA representation of dimension $d_\alpha d_\beta$; suppose it is equivalent with the direct sum $\Sigma_\gamma \oplus m_\gamma D^\gamma$. Then there exists a unitary matrix U with the property

$$(D^\alpha \otimes D^\beta)(g)U^g = U \sum_\gamma \oplus m_\gamma D^\gamma(g), \quad \forall g \in G. \tag{2.14}$$

The elements of the matrix U are the Clebsch–Gordan coefficients. We label the row of U by the pairs (i,j);

$i = 1,2,...,d_\alpha; j = 1,2,...,d_\beta$ and the columns of U by the triples (γ,τ,k) with $m_\gamma \neq 0$; $\tau = 1,2,...,m_\gamma$ and $k = 1,2,\cdots d_\gamma$. The Clebsch–Gordan coefficients may then be denoted by

$$\begin{pmatrix} \alpha & \beta & \big| & \gamma,\tau \\ i & j & \big| & k \end{pmatrix},$$

Another definition of the Clebsch–Gordan coefficients, which is easily shown to be completely equivalent with the preceding definition is the following. Let T be a homomorphic mapping from G into the operators on a $d_\alpha d_\beta$-dimensional Hilbert space $\mathscr{H}$ such that $T(g)$ is unitary if $g \in G_0$ and $T(g)$ is antiunitary if $g \notin G_0$; let $\{\phi_{ij}^{\alpha\beta}\}$ $(i = 1,2,...,d_\alpha;$ $j = 1,2,...,d_\beta)$ be an orthonormal basis of $\mathscr{H}$ which transforms under T according to

$$T(g)\phi_{ij}^{\alpha\beta} = \sum_{kl}(D^\alpha \otimes D^\beta)(g)_{kl,ij}\phi_{kl}^{\alpha\beta}, \quad \forall g \in G. \tag{2.15}$$

Then there exists in $\mathscr{H}$ an orthonormal basis $\{\psi_k^{\gamma,\tau}\}$ $(m_\gamma \neq 0;\ \tau = 1,2,...,m_\gamma;\ k = 1,2,...,d_\gamma)$ such that

$$T(g)\psi_k^{\gamma,\tau} = \sum_l D^\gamma(g)_{lk}\psi_l^{\gamma,\tau}, \quad \forall g \in G. \tag{2.16}$$

The Clebsch–Gordan coefficients are then defined by

$$\psi_k^{\gamma,\tau} = \sum_{i,j}\begin{pmatrix} \alpha & \beta & \big| & \gamma,\tau \\ i & j & \big| & k \end{pmatrix}\phi_{ij}^{\alpha\beta}. \tag{2.17}$$

Since both bases are orthonormal it follows

$$\phi_{ij}^{\alpha\beta} = \sum_{\gamma,\tau,k}\begin{pmatrix} \alpha & \beta & \big| & \gamma,\tau \\ i & j & \big| & k \end{pmatrix}^*\psi_k^{\gamma,\tau} \tag{2.18}$$

Now we define for each γ with $m_\gamma > 0$ and for $p,q = 1,2,...d_\gamma$ the operators P_{pq}^γ and Q_{pq}^γ by

$$P_{pq}^\gamma = \frac{d_\gamma}{|G|}\sum_{g \in G_0} D^\gamma(g)_{pq}^* T(g) \tag{2.19}$$

and

$$Q_{pq}^\gamma = \frac{d_\gamma}{|G|}\sum_{g \notin G_0} D^\gamma(g)_{pq}^* T(g). \tag{2.20}$$

If D^γ is of type II or type III, P_{pq}^γ is equal to zero if $p \not\sim q$ and Q_{pq}^γ is equal to zero if $p \sim q$.

From the definitions the following expressions are derived immediately:

$$T(g)P_{pq}^\gamma = \sum_k D_{kp}^\gamma(g)P_{kq}^\gamma, \quad \forall g \in G_0, \tag{2.21a}$$

$$T(g)P_{pq}^\gamma = \sum_k D_{kp}^\gamma(g)Q_{kq}^\gamma, \quad \forall g \in G \setminus G_0, \tag{2.21b}$$

$$T(g)Q_{pq}^\gamma = \sum_k D_{kp}^\gamma(g)Q_{kq}^\gamma, \quad \forall g \in G_0, \tag{2.21c}$$

$$T(g)Q_{pq}^\gamma = \sum_k D_{kp}^\gamma(g)P_{kq}^\gamma, \quad \forall g \in G \setminus G_0. \tag{2.21d}$$

Using the orthogonality relations one can derive in a straightforward way the following expressions. If D^γ is of type I,

$$P_{pq}^\gamma P_{rs}^\eta = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qr}P_{ps}^\gamma, \tag{2.22a}$$

$$Q_{pq}^\gamma Q_{rs}^\eta = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qr}P_{ps}^\gamma, \tag{2.22b}$$

$$P_{pq}^\gamma Q_{rs}^\eta = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qr}Q_{ps}^\gamma, \tag{2.22c}$$

$$Q_{pq}^\gamma P_{rs}^\eta = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qr}Q_{ps}^\gamma, \tag{2.22d}$$

if D^γ is of type II,

$$P_{pq}^\gamma P_{rs}^\eta = \delta_{\gamma\eta}(\delta_{qr}P_{ps}^\gamma + \delta_{q,r+(d_\gamma/2)}P_{p+(d_\gamma/2),s}^\gamma), \quad (p \sim q), \tag{2.23a}$$

$$Q_{pq}^\gamma Q_{rs}^\eta = \delta_{\gamma\eta}(\delta_{qr}P_{ps}^\gamma - \delta_{q,r+(d_\gamma/2)}P_{p+(d_\gamma/2),s}^\gamma), \quad (p \not\sim q), \tag{2.23b}$$

$$P_{pq}^\gamma Q_{rs}^\eta = \delta_{\gamma\eta}(\delta_{qr}Q_{ps}^\gamma + \delta_{q,r+(d_\gamma/2)}Q_{p+(d_\gamma/2),s}^\gamma), \quad (p \sim q), \tag{2.23c}$$

$$Q_{pq}^\gamma P_{rs}^\eta = \delta_{\gamma\eta}(\delta_{qr}Q_{ps}^\gamma - \delta_{q,r+(d_\gamma/2)}Q_{p+(d_\gamma/2),s}^\gamma), \quad (p \not\sim q), \tag{2.23d}$$

and finally if D^γ is of type III

$$P_{pq}^\gamma P_{rs}^\eta = \delta_{\gamma\eta}\delta_{qr}P_{ps}^\gamma, \quad (p \sim q), \tag{2.24a}$$

$$Q_{pq}^\gamma Q_{rs}^\eta = \delta_{\gamma\eta}\delta_{qr}P_{ps}^\gamma, \quad (p \not\sim q), \tag{2.24b}$$

$$P_{pq}^\gamma Q_{rs}^\eta = \delta_{\gamma\eta}\delta_{qr}Q_{ps}^\gamma, \quad (p \sim q), \tag{2.24c}$$

$$Q_{pq}^\gamma P_{rs}^\eta = \delta_{\gamma\eta}\delta_{qr}Q_{ps}^\gamma, \quad (p \not\sim q). \tag{2.24d}$$

From the orthogonality relations and Eq. (2.16) it follows that the operators P_{rs}^γ and Q_{rs}^γ act as follows on the basis elements $\psi_k^{\gamma,\tau}$. If D^γ is of type I

$$P_{pq}^\gamma\psi_k^{\eta,\tau} = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qk}\psi_p^{\gamma,\tau}, \tag{2.25a}$$

$$Q_{pq}^\gamma\psi_k^{\eta,\tau} = \tfrac{1}{2}\delta_{\gamma\eta}\delta_{qk}\psi_p^{\gamma,\tau}, \tag{2.25b}$$

if D_γ is of type II

$$P_{pq}^\gamma\psi_k^{\eta,\tau} = \delta_{\gamma\eta}(\delta_{qk}\psi_p^{\gamma,\tau} + \delta_{q,k+(d_\gamma/2)}\psi_{p+(d_\gamma/2)}^{\gamma,\tau}), \quad (p \sim q), \tag{2.26a}$$

$$Q_{pq}^\gamma\psi_k^{\eta,\tau} = \delta_{\gamma\eta}(\delta_{qk}\psi_p^{\gamma,\tau} - \delta_{q,k+(d_\gamma/2)}\psi_{p+(d_\gamma/2)}^{\gamma,\tau}), \quad (p \not\sim q), \tag{2.26b}$$

and if D^γ is of type III

$$P_{pq}^\gamma\psi_k^{\eta,\tau} = \delta_{\gamma\eta}\delta_{qk}\psi_p^{\gamma,\tau}, \quad (p \sim q), \tag{2.27a}$$

$$Q_{pq}^\gamma\psi_k^{\eta,\tau} = \delta_{\gamma\eta}\delta_{qk}\psi_p^{\gamma,\tau}, \quad (p \not\sim q). \tag{2.27b}$$

3. EQUATIONS FOR THE CLEBSCH-GORDAN COEFFICIENTS

Let us consider the matrix elements of P_{pq}^γ and Q_{pq}^γ with respect to the basis $\{\phi_{ij}^{\alpha\beta}\}$. From Eq. (2.15) it follows

$$(\phi_{mn}^{\alpha\beta}, P_{pq}^\gamma\phi_{ij}^{\alpha\beta}) = \frac{d_\gamma}{|G|}\sum_{g \in G_0} D_{pq}^{\gamma *}(g)D_{mi}^\alpha(g)D_{nj}^\beta(g) \tag{3.1a}$$

and

$$(\phi_{mn}^{\alpha\beta}, Q_{pq}^\gamma\phi_{ij}^{\alpha\beta}) = \frac{d_\gamma}{|G|}\sum_{g \notin G_0} D_{pq}^{\gamma *}(g)D_{mi}^\alpha(g)D_{nj}^\beta(g). \tag{3.1b}$$

On the other hand, it follows from Eq. (2.18) that

$$(\phi_{mn}^{\alpha\beta}, P_{pq}^\gamma\phi_{ij}^{ab}) = \sum_{\eta,\tau,k}\sum_{\eta',\tau',k'}\begin{pmatrix} \alpha & \beta & \big| & \eta & \tau \\ m & n & \big| & k & \end{pmatrix} \times\begin{pmatrix} \alpha & \beta & \big| & \eta' & \tau' \\ i & j & \big| & k' & \end{pmatrix}^*(\psi_k^{\eta,\tau}, P_{pq}^\gamma\psi_{k'}^{\eta',\tau'}) \tag{3.2a}$$

and

$$(\phi_{mn}^{\alpha\beta}, Q_{pq}^\gamma\phi_{ij}^{\alpha\beta}) = \sum_{\eta,\tau,k}\sum_{\eta',\tau',k'}\begin{pmatrix} \alpha & \beta & \big| & \eta & \tau \\ m & n & \big| & k & \end{pmatrix}$$

$$\times\begin{pmatrix}\alpha & \beta & \Big| & \eta' & \tau' \\ i & j & \Big| & k' & \end{pmatrix}(\psi_k^{\eta,\tau},Q_{pq}^{\gamma}\psi_{k'}^{\eta',\tau'}) \tag{3.2b}$$

Using Eq. (2.25), (2.26), and (2.27) this becomes, if D^γ is of type I,

$$(\phi_{mn}^{\alpha\beta},P_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) = \tfrac{1}{2}\sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix}^* \tag{3.3a}$$

and

$$(\phi_{mn}^{\alpha\beta},Q_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) = \tfrac{1}{2}\sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix}; \tag{3.3b}$$

if D^γ is of type II,

$$\begin{aligned}(\phi_{mn}^{\alpha\beta},P_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) &= \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix}^* \\ &+ \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p+\frac{d_\gamma}{2} & \end{pmatrix} \\ &\times\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q+\frac{d_\gamma}{2} & \end{pmatrix}^* \quad (p\sim q),\end{aligned} \tag{3.4a}$$

and

$$\begin{aligned}(\phi_{mn}^{\alpha\beta},Q_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) &= \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix} \\ &- \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p+\frac{d_\gamma}{2} & \end{pmatrix} \\ &\times\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q+\frac{d_\gamma}{2} & \end{pmatrix} \quad (p\not\sim q),\end{aligned} \tag{3.4b}$$

and if D^γ is of type III,

$$(\phi_{mn}^{\alpha\beta},P_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) = \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix}^* \quad (p\sim q), \tag{3.5a}$$

and

$$(\phi_{mn}^{\alpha\beta},Q_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}) = \sum_\tau\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ m & n & \Big| & p & \end{pmatrix}\begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & q & \end{pmatrix} \quad (p\not\sim q) \tag{3.5b}$$

Let $A(\gamma,p,q)$ and $B(\gamma,p,q)$ be the matrices of P_{pq}^γ and Q_{pq}^γ respectively with respect to the basis $\{\phi_{ij}^{\alpha\beta}\}$

$$A(\gamma,p,q)_{mn,ij} = (\phi_{mn}^{\alpha\beta},P_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}), \tag{3.6a}$$

$$B(\gamma,p,q)_{mn,ij} = (\phi_{mn}^{\alpha\beta},Q_{pq}^{\gamma}\phi_{ij}^{\alpha\beta}). \tag{3.6b}$$

Let $c(\gamma,\tau,k)$ be column matrix of dimension $d_\alpha d_\beta$ given by

$$c(\gamma,\tau,k)_{ij} = \begin{pmatrix}\alpha & \beta & \Big| & \gamma & \tau \\ i & j & \Big| & k & \end{pmatrix}. \tag{3.7}$$

Then Eqs. (3.3), (3.4), and (3.5) can be written as Eqs. (3.8), (3.9), and (3.10), respectively:

$$A(\gamma,p,q) = \tfrac{1}{2}\sum_\tau c(\gamma,\tau,p)\tilde{c}^*(\gamma,\tau,q), \tag{3.8a}$$

$$B(\gamma,p,q) = \tfrac{1}{2}\sum_\tau c(\gamma,\tau,p)\tilde{c}(\gamma,\tau,q), \tag{3.8b}$$

$$\begin{aligned}A(\gamma,p,q) = &\sum_\tau c(\gamma,\tau,p)\tilde{c}^*(\gamma,\tau,q) \\ &+ \sum_\tau c\Big(\gamma,\tau,p+\frac{d_\gamma}{2}\Big)\tilde{c}^*\Big(\gamma,\tau,q+\frac{d_\gamma}{2}\Big), \quad (p\sim q),\end{aligned} \tag{3.9a}$$

$$\begin{aligned}B(\gamma,p,q) = &\sum_\tau c(\gamma,\tau,p)\tilde{c}(\gamma,\tau,q) \\ &- \sum_\tau c\Big(\gamma,\tau,p+\frac{d_\gamma}{2}\Big)\tilde{c}\Big(\gamma,\tau,q+\frac{d_\gamma}{2}\Big), \quad (p\not\sim q),\end{aligned} \tag{3.9b}$$

$$A(\gamma,p,q) = \sum_\tau c(\gamma,\tau,p)\tilde{c}^*(\gamma,\tau,q), \quad (p\sim q), \tag{3.10a}$$

$$B(\gamma,p,q) = \sum_\tau c(\gamma,\tau,p)\tilde{c}(\gamma,\tau,q), \quad (p\not\sim q), \tag{3.10b}$$

Here $\tilde{c}$ denotes the row matrix which is the transpose of c.

Now we will show that a solution of these equations is indeed a unitary matrix of Clebsch–Gordan coefficients. The orthogonality relations of the Clebsch–Gordan coefficients take the form

$$\sum_{\gamma,\tau,k} c(\gamma,\tau,k)\tilde{c}^*(\gamma,\tau,k) = 1. \tag{3.11}$$

But from Eqs. (3.8a), (3.9a) and (3.10a) it follows that $\Sigma_{\gamma,\tau,k}\, c(\gamma,\tau,k)\,\tilde{c}^*(\gamma,\tau,k)$ is the same for each solution, and therefore this sum must be equal to the unit matrix, since we know that a unitary matrix of Clebsch–Gordan coefficients exists. So each solution indeed gives a unitary matrix. To show that this unitary matrix is a matrix of Clebsch–Gordan coefficients we have to prove Eq. (2.16) if Eq. (2.15) holds and if the basis $\{\psi_k^{\gamma,\tau}\}$ is defined by Eq. (2.17) where the coefficients are some solution of Eq. (3.8) (for the D^γ of type I), (3.9) (for the D^γ of type II), and (3.10) (for the D^γ of type III).

From Eqs. (2.17) and (2.15) it follows that

$$P_{pq}^{\gamma}\psi_k^{\eta,\tau} = \sum_{i,j,m,n}\begin{pmatrix}\alpha & \beta & \Big| & \eta & \tau \\ i & j & \Big| & k & \end{pmatrix}A(\gamma,p,q)_{mn,ij}\,\phi_{ij}^{\alpha\beta}. \tag{3.12}$$

If we substitute for $A(\gamma,p,q)_{mn,ij}$ the right-hand side of Eq. (3.3a), (3.4a), or (3.5a), take into account the orthogonality relations for the coefficients and use Eq. (2.17) again, then we obtain Eqs. (2.25a), (2.26a), and (2.27a), respectively. Equations (2.25b), (2.26b) and (2.27b) are derived similarly. From Eqs. (2.25a), (2.26a), and (2.27a) it follows that

$$P_{pq}^{\gamma}\psi_q^{\gamma,\tau} = \alpha\psi_p^{\gamma,\tau}, \tag{3.13}$$

where $\alpha = \frac{1}{2}$ if D^γ is of type I and $\alpha = 1$ if D^γ is of type II or of type III. If we apply $T(g)$ on both sides of this equation for $g\in G_0$, it follows from Eq. (2.21a) that

$$\sum_k D^{\gamma}_{kp}(g) P^{\gamma}_{kq} \psi^{\gamma,\tau}_q = \alpha T(g) \psi^{\gamma,\tau}_p, \tag{3.14}$$

and using Eq. (3.13) we obtain Eq. (2.16) for $g\in G_0$. From Eqs. (2.25b), (2.26b), and (2.27b) it follows that

$$Q^{\gamma}_{pq} \psi^{\gamma,\tau}_q = \alpha \psi^{\gamma,\tau}_p. \tag{3.15}$$

Applying $T(g)$ on both sides of this Eq. for $g\in G \setminus G_0$ we again obtain Eq. (3.14), and thus Eq. (2.16), by using Eq. (2.21d).

4. SOLUTION OF THE EQUATIONS

In this section we will give a method for solving Eqs. (3.8), (3.9), and (3.10). Let us start with Eq. (3.8). Since here D^{γ} is of type I it follows from the Eqs. (2.22) and (3.6) that

$$A(\gamma,p,q)A(\gamma,r,s) = \tfrac{1}{2}\delta_{qr} A(\gamma,p,s), \tag{4.1a}$$

$$B(\gamma,p,q)B^*(\gamma,r,s) = \tfrac{1}{2}\delta_{qr} A(\gamma,p,s), \tag{4.1b}$$

$$A(\gamma,p,q)B(\gamma,r,s) = \tfrac{1}{2}\delta_{qr} B(\gamma,p,s), \tag{4.1c}$$

$$B(\gamma,p,q)A^*(\gamma,r,s) = \tfrac{1}{2}\delta_{qr} B(\gamma,p,s). \tag{4.1d}$$

Moreover, since $P^{\gamma\dagger}_{pq} = P^{\gamma}_{qp}$ and $Q^{\gamma\dagger}_{pq} = Q_{qp}$ it follows from Eq. (3.6) that

$$\widetilde{A}^*(\gamma,p,q) = A(\gamma,q,p), \tag{4.2a}$$

$$\widetilde{B}(\gamma,p,q) = B(\gamma,q,p). \tag{4.2b}$$

Since we know that the solutions of Eq. (3.8) obey the orthogonality relations, it follows that these solutions also satisfy

$$A(\gamma,p,q)c(\gamma,\tau,k) = \tfrac{1}{2}\delta_{qk} c(\gamma,\tau,p), \tag{4.3a}$$

$$B(\gamma,p,q)c^*(\gamma,\tau,k) = \tfrac{1}{2}\delta_{qk} c(\gamma,\tau,p), \tag{4.3b}$$

and, with $p = q = k$,

$$A(\gamma,k,k)c(\gamma,\tau,k) = \tfrac{1}{2}c(\gamma,\tau,k), \tag{4.4a}$$

$$B(\gamma,k,k)c^*(\gamma,\tau,k) = \tfrac{1}{2}c(\gamma,\tau,k). \tag{4.4b}$$

Suppose we have found m_γ orthonormal solutions $c(\gamma,\tau,k)$ $(\tau = 1,2,\cdots m_\gamma)$ of Eq. (4.4) for some fixed value of k, then it follows from Eq. (4.3a) that for $p\neq k$ the vectors $c(\gamma,\tau,p)$ are given by

$$c(\gamma,\tau,p) = 2A(\gamma,p,k)c(\gamma,\tau,k) \tag{4.5}$$

We will show that Eq. (3.8) is then satisfied. Since $2P^{\gamma}_{kk}$ is a projection operator which projects on a subspace of dimension m_γ [see Eq. (2.25a)] the matrix $A(\gamma,k,k)$ has exactly m_γ independent eigenvectors with eigenvalue $\frac{1}{2}$ and the other independent eigenvectors have all eigenvalue 0. Since we supposed we found m_γ orthogonal eigenvectors $c(\gamma,\tau,k)$ $(\tau = 1,2,\ldots,m_\gamma)$ with eigenvalue $\frac{1}{2}$, this set may be extended with eigenvectors with eigenvalue 0 to an orthonormal basis of eigenvectors of $A(\gamma,k,k)$. The relation

$$A(\gamma,k,k) = \tfrac{1}{2}\sum_\tau c(\gamma,\tau,k)\tilde{c}^*(\gamma,\tau,k) \tag{4.6a}$$

is now proved by verifying that both sides give the same result if they are applied to this orthonormal set of eigenvectors. In the same way the relation

$$B(\gamma,k,k) = \tfrac{1}{2}\sum_\tau c(\gamma,\tau,k)\tilde{c}(\gamma,\tau,k) \tag{4.6b}$$

is proved, using the complex conjugates of the orthonormal set of eigenvectors. Equation (3.8) may now be derived immediately from Eqs. (4.5), (4.6), (4.1), and (4.2).

So the problem of solving Eq. (3.8) is reduced to the problem of finding m_γ orthonormal vectors $c(\gamma,\tau,k)$ $(\tau = 1,2,\cdots m_\gamma)$ which satisfy Eq. (4.4).

Let ψ be a nonzero column of $A(\gamma,k,k)$, say the mnth column:

$$\psi_{ij} = A(\gamma,k,k)_{ij,mn}. \tag{4.7}$$

Note that $\tilde{\psi}^*\psi = \frac{1}{2}A(\gamma,k,k)_{mn,mn}$, which means that a column of $A(\gamma,k,k)$ is nonzero if and only if the diagonal term of this column in nonzero. Let ϕ be the mnth column of $B(\gamma,k,k)$:

$$\phi_{ij} = B(\gamma,k,k)_{ij,mn}. \tag{4.8}$$

It follows immediately from Eq. (4.1) that

$$A(\gamma,k,k)\psi = \tfrac{1}{2}\psi, \tag{4.9a}$$

$$A(\gamma,k,k)\phi = \tfrac{1}{2}\phi, \tag{4.9b}$$

$$B(\gamma,k,k)\psi^* = \tfrac{1}{2}\phi, \tag{4.9c}$$

$$B(\gamma,k,k)\phi^* = \tfrac{1}{2}\psi. \tag{4.9d}$$

If $A(\gamma,k,k)_{mn,mn}$ is not equal to $-B(\gamma,k,k)_{mn,mn}$ then $\psi + \phi$ is nonzero and satisfies Eq. (4.4). So after normalizing this vector we may define it to be $c(\gamma,1,k)$:

$$c(\gamma,1,k) = (\psi+\phi)\{A(\gamma,k,k)_{mn,mn} + \tfrac{1}{2}B^*(\gamma,k,k)_{mn,mn} + \tfrac{1}{2}B(\gamma,k,k)_{mn,mn}\}^{-1/2}. \tag{4.10}$$

If it happens that $A(\gamma,k,k)_{mn,mn} = -B(\gamma,k,k)_{mn,mn}$ then the vector $i(\psi-\phi)$ is nonzero and satisfies Eq. (4.4); so we may then take $c(\gamma,1,k)$ equal to this vector after normalizing it:

$$c(\gamma,1,k) = \frac{i(\psi-\phi)}{\{2A(\gamma,k,k)_{mn,mn}\}^{1/2}}. \tag{4.11}$$

If $m_\gamma \geqslant 2$ we define $A'(\gamma,k,k)$ and $B'(\gamma,k,k)$ by

$$A'(\gamma,k,k) = A(\gamma,k,k) - \tfrac{1}{2}c(\gamma,1,k)\tilde{c}^*(\gamma,1,k) \tag{4.12a}$$

$$B'(\gamma,k,k) = B(\gamma,k,k) - \tfrac{1}{2}c(\gamma,1,k)\tilde{c}(\gamma,1,k) \tag{4.12b}$$

It is easy to check that the Eqs. (4.1) and (4.2) (with p,q,r, and s all equal to k) hold as well for $A'(\gamma,k,k)$ and $B'(\gamma,k,k)$. Therefore, we obtain in the same way as above a normalized vector $c(\gamma,2,k)$ which satisfies

$$A'(\gamma,k,k)c(\gamma,2,k) = \tfrac{1}{2}c(\gamma,2,k), \tag{4.13a}$$

$$B'(\gamma,k,k)c^*(\gamma,2,k) = \tfrac{1}{2}c(\gamma,2,k). \tag{4.13b}$$

Now since

$$A(\gamma,k,k)A'(\gamma,k,k) = \tfrac{1}{2}A'(\gamma,k,k) \tag{4.14}$$

and

$$B(\gamma,k,k)B'^*(\gamma,k,k) = \tfrac{1}{2}A'(\gamma,k,k), \tag{4.15}$$

we obtain from Eq. (4.13a)

$$A(\gamma,k,k)c(\gamma,2,k) = \tfrac{1}{2}c(\gamma,2,k), \tag{4.16a}$$

and from Eq. (4.13b)

$$B(\gamma,k,k)c^*(\gamma,2,k) = \tfrac{1}{2}c(\gamma,2,k). \tag{4.16b}$$

If we apply both sides of Eq. (4.12a) to the vector $c(\gamma,2,k)$ we find $\tilde{c}^*(\gamma,1,k)c(\gamma,2,k) = 0$; so $c(\gamma,1,k)$ and $c(\gamma,2,k)$ are orthogonal. Thus, we have shown that the vector $c(\gamma,2,k)$ satisfies all requirements. It is obvious that if $m_\gamma \geqslant 3$ we may proceed in this way until m_γ orthonormal vectors satisfying Eq. (4.4) are obtained.

Let us now consider Eq. (3.9). Note that we only have to consider Eq. (3.9a) for $p,q \leqslant d_\gamma/2$ and Eq. (3.9b) for $p \leqslant d_\gamma/2$ and $q > d_\gamma/2$. Let us define the vectors $d(\gamma,\tau,p)$ for $\tau = 1,2,...,2m_\gamma$ and $p = 1,2,...,d_\gamma/2$ by

$$d(\gamma,\tau,p) = c(\gamma,\tau,p) \quad 1 \leqslant \tau \leqslant m_\gamma, \tag{4.17a}$$

$$d(\gamma,\tau,p) = c\left(\gamma,\tau - m_\gamma,p + \frac{d_\gamma}{2}\right) \quad m_\gamma < \tau \leqslant 2m_\gamma. \tag{4.17b}$$

This definition only implies a different labeling of the vectors $c(\gamma,\tau,p)$. Equation (3.9a), for $p,q \leqslant d_\gamma/2$ now becomes

$$A(\gamma,p,q) = \sum_{\tau=1}^{2m_\gamma} d(\gamma,\tau,p)\tilde{d}^*(\gamma,\tau,q), \tag{4.18a}$$

and Eq. (3.9b), for $p \leqslant d_\gamma/2$ and $q > d_\gamma/2$ becomes

$$B(\gamma,p,q) = \sum_{\tau=1}^{m_\gamma} d(\gamma,\tau,p)\tilde{d}\left(\gamma,\tau + m_\gamma,q - \frac{d_\gamma}{2}\right) - \sum_{\tau=1}^{m_\gamma} d(\gamma,\tau + m_\gamma,p)\tilde{d}\left(\gamma,\tau,q - \frac{d_\gamma}{2}\right),$$

which may be written as

$$C(\gamma,p,q) = \sum_{\tau=1}^{2m_\gamma} \epsilon(\tau) d(\gamma,\tau,p)\tilde{d}(\gamma,\tau + m_\gamma,q), \tag{4.18b}$$

where $p,q \leqslant d_\gamma/2$, $C(\gamma,p,q)$ is defined by

$$C(\gamma,p,q) = B\left(\gamma,p,q + \frac{d_\gamma}{2}\right), \tag{4.19}$$

$\epsilon(\tau)$ is defined by

$$\epsilon(\tau) = \begin{cases} 1 & \text{if } 1 \leqslant \tau \leqslant m_\gamma, \\ -1 & \text{if } m_\gamma < \tau \leqslant 2m_\gamma, \end{cases} \tag{4.20}$$

and $\tau + m_\gamma$ should be read as $\tau + m_\gamma \pmod{2m_\gamma}$. So Eq. (3.9) has been rewritten as Eq. (4.18), with the definitions of Eqs. (4.17), (4.19), and (4.20), and the parameters p,q now run from 1–$d_\gamma/2$.

From Eqs. (2.23), (3.6), and (4.19) and the fact that

$$Q^\gamma_{pq} = -Q^\gamma_{p+(d_\gamma/2),q+(d_\gamma/2)}$$

and

$$P^\gamma_{pq} = P^\gamma_{p+(d_\gamma/2),q+(d_\gamma/2)}$$

if D^γ is of type II, it follows that for $p,q,r,s \leqslant d_\gamma/2$

$$A(\gamma,p,q)A(\gamma,r,s) = \delta_{qr}A(\gamma,p,s), \tag{4.21a}$$

$$C(\gamma,p,q)C^*(\gamma,r,s) = -\delta_{qr}A(\gamma,p,s), \tag{4.21b}$$

$$A(\gamma,p,q)C(\gamma,r,s) = \delta_{qr}C(\gamma,p,s), \tag{4.21c}$$

$$C(\gamma,p,q)A^*(\gamma,r,s) = \delta_{qr}C(\gamma,p,s). \tag{4.21d}$$

Moreover, since $P^{\gamma\dagger}_{pq} = P^\gamma_{qp}$ and $Q^{\gamma\dagger}_{pq} = Q^\gamma_{qp}$ it follows from Eqs. (3.6) and (4.19) that

$$\tilde{A}^*(\gamma,p,q) = A(\gamma,q,p), \tag{4.22a}$$

$$\tilde{C}(\gamma,p,q) = -C(\gamma,q,p). \tag{4.22b}$$

Since we know that the solutions of Eq. (4.18) satisfy the orthogonality relations, it follows that these solutions also satisfy

$$A(\gamma,p,q)d(\gamma,\tau,k) = \delta_{qk}d(\gamma,\tau,p), \tag{4.23a}$$

$$C(\gamma,p,q)d^*(\gamma,\tau,k) = \delta_{qk}\epsilon(\tau + m_\gamma)d(\gamma,\tau + m_\gamma,p), \tag{4.23b}$$

and, with $p = q = k$,

$$A(\gamma,k,k)d(\gamma,\tau,k) = d(\gamma,\tau,k), \tag{4.24a}$$

$$C(\gamma,k,k)d^*(\gamma,\tau,k) = \epsilon(\tau + m_\gamma)d(\gamma,\tau + m_\gamma,k). \tag{4.24b}$$

Suppose we found $2m_\gamma$ orthonormal solutions $d(\gamma,\tau,k)$ $(\tau = 1,2,\cdots 2m_\gamma)$ of Eq. (4.24) for some fixed value of k; it then follows from Eq. (4.23a) that for $p \neq k$ the vectors $d(\gamma,\tau,p)$ are given by

$$d(\gamma,\tau,p) = A(\gamma,p,k)d(\gamma,\tau,k). \tag{4.25}$$

These vectors satisfy Eq. (4.18) as may be verified immediately, noting that P^γ_{kk} is a projection operator which projects on a subspace of dimension $2m_\gamma$ [see Eq. (2.26a)] and reasoning along the same lines as above, where we derived Eq. (3.8) from Eqs. (4.4) and (4.5).

So the problem of solving Eq. (4.18) is reduced to the problem of finding $2m_\gamma$ orthonormal vectors $d(\gamma,\tau,k)$ $(\tau = 1,2,...,2m_\gamma)$ which satisfy Eq. (4.24). This can be done as follows. Let ψ be a nonzero column of $A(\gamma,k,k)$, say the mnth column:

$$\psi_{ij} = A(\gamma,k,k)_{ij,mn}, \tag{4.26}$$

and let ϕ be the mnth column of $C(\gamma,k,k)$:

$$\phi_{ij} = C(\gamma,k,k)_{ij,mn}. \tag{4.27}$$

Then the following properties may be derived immediately from Eqs. (4.21) and (4.22):

$$A(\gamma,k,k)\psi = \psi, \tag{4.28a}$$

$$A(\gamma,k,k)\phi = \phi, \tag{4.28b}$$

$$C(\gamma,k,k)\psi^* = \phi, \tag{4.28c}$$

$$C(\gamma,k,k)\phi^* = -\psi, \tag{4.28d}$$

$$\tilde{\psi}^*\psi = \tilde{\phi}^*\phi = A(\gamma,k,k)_{mn,mn}, \tag{4.28e}$$

$$\tilde{\psi}^*\phi = 0. \tag{4.28f}$$

Therefore, we may define

$$d(\gamma,1,k) = \frac{\psi}{\{A(\gamma,k,k)_{mn,mn}\}^{1/2}} \tag{4.29a}$$

and

$$d(\gamma,m_\gamma + 1,k) = \frac{-\phi}{\{A(\gamma,k,k)_{mn,mn}\}^{1/2}}; \tag{4.29b}$$

these vectors are orthonormal and satisfy Eq. (4.24). If $m_\gamma \geqslant 2$ we define $A'(\gamma,k,k)$ and $C'(\gamma,k,k)$ by

$$A'(\gamma,k,k) = A(\gamma,k,k) - d(\gamma,1,k)\tilde{d}^*(\gamma,1,k) - d(\gamma,m_\gamma + 1,k)\tilde{d}^*(\gamma,m_\gamma + 1,k), \tag{4.30a}$$

$$C'(\gamma,k,k) = C(\gamma,k,k) - d(\gamma,1,k)\tilde{d}(\gamma,m_\gamma + 1,k) + d(\gamma,m_\gamma + 1,k)\tilde{d}(\gamma,1,k). \tag{4.30b}$$

It is easy to check that Eqs. (4.21) and (4.22) (with p,q,r and s

all equal to k) hold as well for $A'(\gamma,k,k)$ and $C'(\gamma,k,k)$. Therefore, we obtain in the same way as above two orthonormal vectors $d(\gamma,2,k)$ and $d(\gamma,m_\gamma+2,k)$ which satisfy

$$A'(\gamma,k,k)d(\gamma,2,k)=d(\gamma,2,k), \tag{4.31a}$$

$$A'(\gamma,k,k)d(\gamma,m_\gamma+2,k)=d(\gamma,m_\gamma+2,k), \tag{4.31b}$$

$$C'(\gamma,k,k)d^*(\gamma,2,k)=-d(\gamma,m_\gamma+2,k), \tag{4.31c}$$

$$C'(\gamma,k,k)d^*(\gamma,m_\gamma+2,k)=d(\gamma,2,k). \tag{4.31d}$$

Now since

$$A(\gamma,k,k)A'(\gamma,k,k)=A'(\gamma,k,k) \tag{4.32a}$$

and

$$C(\gamma,k,k)C'^*(\gamma,k,k)=-A'(\gamma,k,k), \tag{4.32b}$$

we obtain from Eq. (4.31)

$$A(\gamma,k,k)d(\gamma,2,k)=d(\gamma,2,k), \tag{4.33a}$$

$$A(\gamma,k,k)d(\gamma,m_\gamma+2,k)=d(\gamma,m_\gamma+2,k), \tag{4.33b}$$

$$C(\gamma,k,k)d^*(\gamma,m_\gamma+2,k)=d(\gamma,2,k), \tag{4.33c}$$

$$C(\gamma,k,k)d^*(\gamma,2,k)=d(\gamma,m_\gamma+2,k). \tag{4.33d}$$

Since $A'(\gamma,k,k)d(\gamma,2,k)=A(\gamma,k,k)\,d(\gamma,2,k)$ Eq. (4.30a) gives $\{d(\gamma,1,k)\tilde{d}^*(\gamma,1,k)-d(\gamma,m_\gamma+1,k)\tilde{d}^*(\gamma,m_\gamma+1,k)\}$ $d(\gamma,2,k)=0$. By multiplying on the left with $\tilde{d}^*(\gamma,1,k)$ we find, since $\tilde{d}^*(\gamma,1,k)\,d(\gamma,1,k)=1$ and $\tilde{d}^*(\gamma,1,k)$ $d(\gamma,m_\gamma+1,k)=0$, that $\tilde{d}^*(\gamma,1,k)\,d(\gamma,2,k)=0$; so $d(\gamma,1,k)$ and $d(\gamma,2,k)$ are orthogonal. In the same way we may show that $d(\gamma,1,k)$, $d(\gamma,2,k)$, $d(\gamma,m_\gamma+1,k)$ and $d(\gamma,m_\gamma+2,k)$ are orthonormal vectors. Thus, we indeed obtained correct solutions $d(\gamma,2,k)$ and $d(\gamma,m_\gamma+2,k)$. It is obvious that if $m_\gamma \geqslant 3$ we may proceed in this way until $2m_\gamma$ orthonormal vectors satisfying Eq. (4.24) are obtained. The solutions $c(\gamma,\tau,p)$ of Eq. (3.9) are then given by Eqs. (4.25) and (4.17).

Let us finally consider Eq. (3.10). Since here D^γ is of type III, it follows from Eqs. (2.24) and (3.6) that

$$A(\gamma,p,q)A(\gamma,r,s)=\delta_{qr}A(\gamma,p,s) \quad (p\sim q), \tag{4.34a}$$

$$B(\gamma,p,q)B^*(\gamma,r,s)=\delta_{qr}A(\gamma,p,s) \quad (p\not\sim q), \tag{4.34b}$$

$$A(\gamma,p,q)B(\gamma,r,s)=\delta_{qr}B(\gamma,p,s) \quad (p\sim q), \tag{4.34c}$$

$$B(\gamma,p,q)A^*(\gamma,r,s)=\delta_{qr}B(\gamma,p,s) \quad (p\not\sim q). \tag{4.34d}$$

Moreover, since $P^{\gamma\dagger}_{pq}=P^{\gamma}_{qp}$ and $Q^{\gamma\dagger}_{pq}=Q^{\gamma}_{qp}$ it follows from Eq. (3.6) that

$$\tilde{A}^*(\gamma,p,q)=A(\gamma,q,p), \tag{4.35a}$$

$$\tilde{B}(\gamma,p,q)=B(\gamma,q,p). \tag{4.35b}$$

Since we know that the solutions of Eq. (3.10) satisfy the orthogonality relations, it follows that these solutions also satisfy

$$A(\gamma,p,q)c(\gamma,\tau,k)=\delta_{qk}c(\gamma,\tau,p) \quad (p\sim q), \tag{4.36a}$$

$$B(\gamma,p,q)c^*(\gamma,\tau,k)=\delta_{qk}c(\gamma,\tau,p) \quad (p\not\sim q), \tag{4.36b}$$

and, with $p=q=k$

$$A(\gamma,k,k)c(\gamma,\tau,k)=c(\gamma,\tau,k). \tag{4.37}$$

Suppose we have found m_γ orthonormal solutions $c(\gamma,\tau,k)$ $(\tau=1,2,..,m_\gamma)$ of Eq. (4.37) for some fixed value of k; it then follows from Eq. (4.36) that for $p\neq k$ the vectors $c(\gamma,\tau,p)$ are given by

$$c(\gamma,\tau,p)=A(\gamma,p,k)c(\gamma,\tau,k) \tag{4.38a}$$

if $p\sim k$, and by

$$c(\gamma,\tau,p)=B(\gamma,p,k)c^*(\gamma,\tau,k) \tag{4.38b}$$

if $p\not\sim k$. As in the preceding cases we can show that Eq. (3.10) is then satisfied. So the problem of solving Eq. (3.10) is reduced to the problem of finding m_γ orthonormal vectors $c(\gamma,\tau,k)$ $(\tau=1,2,...,m_\gamma)$ which satisfy Eq. (4.37).

But this problem is exactly the same as the corresponding problem for unitary representations of nonmagnetic groups,[5] the solution being as follows: let ψ be a nonzero column of $A(\gamma,k,k)$, say the mnth column. Then

$$c(\gamma,1,k)=\frac{\psi}{\{A(\gamma,k,k)_{mn,mn}\}^{1/2}}. \tag{4.39}$$

Define (if $m_\gamma \geqslant 2$)

$$A'(\gamma,k,k)=A(\gamma,k,k)-c(\gamma,1,k)\tilde{c}^*(\gamma,1,k), \tag{4.40}$$

and take $c(\gamma,2,k)$ equal to a normalized column of $A'(\gamma,k,k)$. In this way one may proceed until m_γ orthonormal solutions $c(\gamma,\tau,k)$ $(\tau=1,2,...,m_\gamma)$ of Eq. (4.37) are obtained.

5. SUMMARY OF THE METHOD

The decomposition of the Kronecker product and the notation of the Clebsch–Gordan coefficients is given in Eq. (2.14) and the lines below. For each D^γ appearing in the decomposition the Clebsch–Gordan coefficients are calculated separately, the method depending on the type of D^γ.

Suppose D^γ is of type I. We may choose $k=1$ and calculate the matrices $A(\gamma,p,k)$ and $B(\gamma,p,k)$ for $1\leqslant p\leqslant d_\gamma$ from Eqs. (3.1) and (3.6). The Clebsch–Gordan coefficients are written as column matrices with Eq. (3.7). We choose a diagonal element $A(\gamma,k,k)_{mn,mn}$ which is not equal to zero. If $A(\gamma,k,k)_{mn,mn}$ is not equal $-B(\gamma,k,k)_{mn}$ then $c(\gamma,1,k)$ is given by Eqs. (4.10), (4.7), and (4.8), and otherwise by Eqs. (4.11), (4.7), and (4.8). If $m_\gamma\geqslant 2$ then $A'(\gamma,k,k)$ and $B'(\gamma,k,k)$ are obtained with Eq. (4.12) and from these matrices $c(\gamma,2,k)$ is obtained in the same way as $c(\gamma,1,k)$ was obtained from $A(\gamma,k,k)$ and $B(\gamma,k,k)$. This procedure is continued until $c(\gamma,\tau,k)$ is known for $1\leqslant\tau\leqslant m_\gamma$. The remaining Clebsch–Gordan coefficients are given by Eq. (4.5).

Suppose D^γ is of type II. We may choose $k=1$ and calculate the matrices $A(\gamma,p,k)$ and $C(\gamma,p,k)$ for $1\leqslant p\leqslant d_\gamma/2$ from Eqs. (3.1), (3.6), and (4.19). The Clebsch–Gordan coefficients are written as column matrices with Eqs. (3.7) and (4.17). We choose a diagonal element $A(\gamma,k,k)_{mn,mn}$ which is not equal to zero. Then $d(\gamma,1,k)$ is given by Eqs. (4.29a) and (4.26); $d(\gamma,m_\gamma+1,k)$ is given by Eqs. (4.29b) and (4.27). If $m_\gamma\geqslant 2$, then $A'(\gamma,k,k)$ and $C'(\gamma,k,k)$ are obtained with Eq. (4.30), and from these matrices we obtain $d(\gamma,2,k)$ and $d(\gamma,m_\gamma+2,k)$ in the same way as we obtained $d(\gamma,1,k)$ and $d(\gamma,m_\gamma+1,k)$ from $A(\gamma,k,k)$ and $C(\gamma,k,k)$. This procedure is continued until $d(\gamma,\tau,k)$ is known for $1\leqslant\tau\leqslant 2m_\gamma$. The remaining Clebsch–Gordan coefficients are given by Eq. (4.25).

Suppose D^γ is of type III. We may choose $k=1$ and calculate the matrices $A(\gamma,p,k)$ and $B(\gamma,p,k)$ for $1\leqslant p\leqslant d_\gamma$ from Eqs. (3.1) and (3.6). The Clebsch–Gordan coefficients are written as column matrices with Eq. (3.7). We choose a diagonal element $A(\gamma,k,k)_{mn,mn}$ which is not equal to zero. Then $c(\gamma,1,k)$ is given by Eqs. (4.39) and (4.26). If $m_\gamma \geqslant 2$ then $A'(\gamma,k,k)$ is obtained with Eq. (4.40) and we obtain $c(\gamma,2,k)$ from $A'(\gamma,k,k)$ in the same way as we obtained $c(\gamma,1,k)$ from $A(\gamma,k,k)$. This procedure is continued until $c(\gamma,\tau,k)$ is known for $1 \leqslant \tau \leqslant m_\gamma$. The remaining Clebsch–Gordan coefficients are given by Eq. (4.38a) for $2 \leqslant p \leqslant d_\gamma/2$ and by Eq. (4.39b) for $d_\gamma/2 < p \leqslant d_\gamma$.

6. THE WIGNER ECKART THEOREM FOR FINITE MAGNETIC GROUPS

The Wigner–Eckart theorem has been generalized for UA representations of finite magnetic groups by Aviran and Zak.[3] Here we will give a derivation of the Wigner–Eckart theorem which brings to light the close connection with the lemma of Schur.

Let T be a UA representation in a Hilbert space H and let $\{\phi_i^\beta\}$ $(i = 1,\ldots,d_\beta)$ and $\{\phi_j^\gamma\}$ $(j = 1,\ldots,d_\gamma)$ be orthonormal sets of elements of H which transform under T according to the irreducible UA representations D^β and D^γ of G:

$$T(g)\phi_i^\beta = \sum_j D_{ji}^\beta(g)\phi_j^\beta, \quad \forall g \in G, \tag{5.1}$$

and similarly for $\{\phi_j^\gamma\}$. Let O^α be a irreducible tensor operator transforming under T according to the irreducible UA representation D^α of G:

$$T(g)O_k^\alpha T(g)^{-1} = \sum_l D_{lk}^\alpha(g)O_l^\alpha, \quad \forall g \in G. \tag{5.2}$$

We assume that the sets $\{\phi_i^\alpha\}$, $\{\phi_j^\beta\}$ and $\{O_k^\alpha\}$ have been chosen in such a way that D^α, D^β and D^γ are the fixed representatives of their equivalence classes (which is always possible). The elements $O_i^\alpha\phi_j^\beta$ transform according to $D^\alpha \otimes D^\beta$:

$$T(g)O_i^\alpha\phi_j^\beta = \sum_{kl}(D^\alpha \otimes D^\beta)_{kl,ij}O_k^\alpha\phi_l^\beta \quad \forall g \in G. \tag{5.3}$$

If $D^\alpha \otimes D^\beta$ is equivalent with the direct sum $\Sigma_\oplus m_\eta D^\eta$ then the elements $\psi_p^{\eta,\tau}(m_\eta \neq 0; \tau = 1,2,\cdots m_\eta, p = 1,\cdots d_\eta)$, defined by

$$\psi_p^{\eta,\tau} = \sum_{i,j}\begin{pmatrix}\alpha & \beta & \big| & \eta & \tau \\ i & j & \big| & p & \end{pmatrix} O_i^\alpha\phi_j^\beta, \tag{5.4}$$

where the $\left(\begin{smallmatrix}\alpha & \beta & | & \eta & \tau\\ i & j & | & p & \end{smallmatrix}\right)$ are the Clebsch–Gordan coefficients, transform according to

$$T(g)\psi_p^{\eta,\tau} = \sum_q D_{qp}^\eta(g)\psi_q^{\eta,\tau}, \quad \forall g \in G. \tag{5.5}$$

We are interested in the matrix elements $(\phi_k^\gamma, O_i^\beta\phi_j^\alpha)$. We have

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_{\eta,\tau,p}\begin{pmatrix}\alpha & \beta & \big| & \eta & \tau \\ i & j & \big| & p & \end{pmatrix}^* (\phi_k^\gamma, \psi_p^{\eta,\tau}). \tag{5.6}$$

Define the matrix U by $U_{kp} = (\phi_k^\gamma, \psi_p^{\eta,\tau})$. One verifies immediately that

$$D^\gamma(g)U^g = UD^\eta(g) \quad \forall g \in G. \tag{5.7}$$

Therefore, if $\gamma \neq \eta$, U is equal to the zero matrix[17]; if $\gamma = \eta$, then from the generalized lemma of Schur it follows that U has the form of Eq. (2.9). From Eq. (5.6) it follows that the Wigner–Eckart theorem now reads as follows: If D^γ is of type I, then

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k & \end{pmatrix}^* \lambda_\gamma(\tau), \tag{5.8}$$

where $\lambda_\gamma(\tau)$ are real numbers which do not depend on i, j and k; if D^γ is of type II, then

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k & \end{pmatrix}^* \lambda_\gamma(\tau) + \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k + \frac{d_\gamma}{2} & \end{pmatrix}^* \mu_\gamma(\tau), \tag{5.9a}$$

if $k \leqslant d_\gamma/2$, and

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k & \end{pmatrix}^* \lambda_\gamma^*(\tau) - \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k - \frac{d_\gamma}{2} & \end{pmatrix}^* \mu_\gamma^*(\tau), \tag{5.9b}$$

if $d_\gamma/2 < k \leqslant d_\gamma$, where $\lambda_\gamma(\tau)$ and $\mu_\gamma(\tau)$ are complex numbers which do not depend on i, j and k; if D^γ is of type III, then

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k & \end{pmatrix}^* \lambda_\gamma(\tau), \tag{5.10a}$$

if $k \leqslant d_\gamma/2$, and

$$(\phi_k^\gamma, O_i^\beta\phi_j^\alpha) = \sum_\tau \begin{pmatrix}\alpha & \beta & \big| & \gamma & \tau \\ i & j & \big| & k & \end{pmatrix}^* \lambda_\gamma^*(\tau), \tag{5.10b}$$

if $d_\gamma/2 < k \leqslant d_\gamma$, where $\lambda_\gamma(\tau)$ are complex numbers which do not depend on i, j and k. Since the generalized lemma of Schur does not hold in general if the irreducible UA representations are not in standard form, the Wigner–Eckart theorem as formulated above will also not be valid in general for irreducible UA representations which are not in standard form.

ACKNOWLEDGEMENT

I thank Dr. T. Janssen for reading the manuscript.

[1] V. Bargmann, Ann. Math. **59**, 1, (1954).
[2] E.P. Wigner, Ann. Math. **40**, 149 (1939).
[3] A. Aviran and J. Zak, J. Math. Phys. **9**, 2138 (1968).
[4] S. Schindler and R. Mirman, J. Math. Phys. **18**, 1678 (1977).
[5] P.M. van den Broek and J.F. Cornwell, Phys. Stat. Sol. B **90**, 211 (1978).
[6] R. Dirl, J. Math. Phys. **20**, 659 (1979).
[7] J.F. Cornwell, Phys. Stat. Sol. **37**, 225 (1970).
[8] R. Berenson and J.L. Birman, J. Math. Phys., **16**, 227 (1975).
[9] I. Sakata, J. Math. Phys. **15**, 1702 (1974).
[10] G.F. Koster, Phys. Rev. **109**, 227 (1957).
[11] A. Aviran and D.B. Litvin, J. Math. Phys. **14**, 1491 (1973).
[12] P. Rudra, J. Math. Phys. **15**, 2031 (1974).
[13] I.N. Kotsev, Sov. Phys. Crystallogr. **19**, 286 (1974).
[14] P. Gard and N.B. Backhouse, J. Phys. A.: Math. Gen. **8**, 450 (1975).
[15] I. Sakata, J. Math. Phys. **15**, 1710 (1974).
[16] E.P. Wigner, Group Theory (Academic, New York, 1959).
[17] J.O. Dimmock, J. Math. Phys. **4**, 1307 (1963).

Stokes multipliers for a class of ordinary differential equations

F. L. Hinton
Fusion Research Center, The University of Texas at Austin, Austin, Texas 78712

(Received 22 January 1979; accepted for publication 1 June 1979)

A new method is presented for calculating the Stokes multipliers for a class of linear second-order ordinary differential equations. The Stokes multipliers allow the asymptotic solutions of these equations to be continued across the Stokes lines on which they are dominant. The differential equations, of the class considered here, have an irregular singular point at infinity and a singular point at the origin, which may be either regular or irregular. The Stokes multipliers, as functions of the coefficients in the differential equation, are obtained in the form of convergent infinite series, whose terms must be obtained from the solution of recursion relations, which are derived. In the case of Whittaker's equation (when the origin is a regular singular point), the known results are obtained analytically. When the origin is an irregular singular point, numerical evaluation of the series is necessary, but the method seems to be quite efficient for use with digital computers. In the special case of an equation, with two irregular singular points, which can be transformed to Mathieu's equation, the numerical results for the Stokes multiplier show good agreement with available known results for the characteristic exponents of Mathieu's equation.

I. INTRODUCTION

The second-order ordinary differential equations considered here have the form

$$\frac{d^2\psi}{d\zeta^2} + Q(\zeta)\psi = 0,$$

where

$$Q(\zeta) = -\tfrac{1}{4} + \sum_{n=1}^{\infty} Q_n \zeta^{-n}$$

which converges for $|\zeta| > R$ for some $R > 0$. These equations have an irregular singular point at infinity; the asymptotic forms of the solutions are

$$\psi \sim e^{\pm \zeta/2}$$

which have an essential singularity at infinity. The equations may have singular points in the finite ζ plane also, depending on the behavior of $Q(\zeta)$ for finite values of ζ. For example, if the series for $Q(\zeta) + \frac{1}{4}$ terminates with M terms, the origin is a regular singular point when M is 1 or 2, and is an irregular singular point when $M > 2$. If the series for $Q(\zeta) + \frac{1}{4}$ does not terminate, it may include the series expansions of a finite number of terms like $(\zeta - \zeta_j)^{-m_j}$, which corresponds to a singular point at $\zeta = \zeta_j$. However, we shall not explicitly consider finite singular points other than the origin $\zeta = 0$.

In this paper, a new method for determining the Stokes multipliers, is presented. We first demonstrate, in Sec. II, the relation between the Stokes multipliers and the reflection and transmission coefficients for wave problems which arise in physical applications. The terminology used in the paper is explained in this section also, and a brief introduction to the concept of Stokes multipliers is included. In Sec. III, a pair of coupled integral equations is derived, which is the starting point of the method used here. An asymptotic solution of these equations is obtained in Sec. IV, including a convergent series expression for the Stokes multiplier, in terms of the solution of certain recursion relations. In Sec. V, we show that the Stokes multipliers for Bessel's equation and for Whittaker's equation are obtained analytically and quite simply by the present method. In Sec. VI, the results of numerical solutions of the recursion relations are compared with the known results for Whittaker's equation (the standard equation with two singular points, one regular and one irregular) and for Mathieu's equation (which can be transformed to an equation with two irregular singular points).

II. WAVE REFLECTION AND TRANSMISSION

Second-order ordinary differential equations of the form

$$\frac{d^2\phi}{dx^2} + q(x)\phi = 0 \tag{1}$$

for $-\infty < x < \infty$, occur frequently in physics. In many problems, this is a time-independent version of a wave equation, and the function q also depends on a parameter ω, the wave frequency. In some problems, a wave of unit amplitude is assumed to be incident from $-\infty$, and the amplitudes of the reflected and transmitted waves are desired, as functions of ω. In other problems, those values of ω are sought for which a solution exists with no incident wave: These are normal modes of the system.

The asymptotic boundary conditions in these problems can be stated in terms of the Liouville–Green (WKBJ) approximation.[1] Assuming that

$$q = \bar{q} + \tilde{q}, \tag{2}$$

where

$$\left|\frac{\tilde{q}}{\bar{q}}\right| \to 0, \quad \bar{q}^{-3/4}\frac{d^2}{dx^2}\bar{q}^{-1/4} \to 0$$

for $|x| \to \infty$, then the following is an asymptotic form of the solution

$$\phi \sim (-\bar{q})^{-1/4}\left\{C^+ \exp\left[\int dx(-\bar{q})^{1/2}\right] + C^- \exp\left[-\int dx(-\bar{q})^{1/2}\right]\right\} \tag{3}$$

(where the values of C^+ and C^- are different, in general, for $x \to +\infty$ and $x \to -\infty$). If $C^-(-\infty)$ is the amplitude of the incident wave, then $C^+(-\infty)$ is the amplitude of the reflected wave, $C^-(+\infty)$ is the amplitude of the transmitted wave, and $C^+(+\infty)$ is assumed to be zero. For normal modes, the values of ω are sought for which $C^+(-\infty)$ and $C^-(+\infty)$ are finite when $C^-(-\infty) = 0$.

In this paper, we address the problem of determining the relation which connects the asymptotic solutions for $x \to -\infty$ and $x \to +\infty$. The connection formulas, which give $C^-(-\infty)$ and $C^+(-\infty)$ in terms of $C^-(+\infty)$, are related to the Stokes multipliers, which we shall define shortly.

A. Singular points of the wave equation

We shall consider the solution of Eq. (1) to be a function of a complex variable Z, and that the function $q(Z)$ has the asymptotic form

$$q \sim q_{N-2} Z^{N-2}, \quad \text{for } |Z| \to \infty \tag{4}$$

with $N > 0$. Then the solution, according to Eq. (3), behaves as

$$\phi \sim \exp[\pm (2/N)(-q_{N-2})^{1/2} Z^{N/2}]$$

which has an essential singularity at $Z = \infty$. Equation (1) is said to have an irregular singular point at $Z = \infty$, in this case. We shall assume here that N is an integer greater than one. The *rank* of the irregular singular point is the least integer k for which $N/2 \leqslant k$; we assume that k is finite.

The solution behaves differently as $Z \to \infty$ along different *Stokes lines*, which will be defined by

$$\text{Im}[(-q_{N-2})^{1/2} Z^{N/2}] = 0. \tag{5}$$

As $Z \to \infty$ along one of these N rays, a solution ϕ which increases exponentially is called *dominant*, while one which decreases exponentially is called *recessive*.

The following change of variables is found to be convenient. Let

$$\phi(Z) = Z^{1/2 - N/4} \psi(\zeta), \quad \zeta = (4/N)(-q_{N-2})^{1/2} Z^{N/2}. \tag{6}$$

Then the equation for $\psi(\zeta)$ is

$$d^2\psi/d\zeta^2 + Q(\zeta)\psi = 0, \tag{7}$$

where

$$Q = -\frac{1}{4} - \frac{[q(Z) - q_{N-2} Z^{N-2}]}{4 q_{N-2} Z^{N-2}} - \frac{(N^2 - 4)}{64 q_{N-2}} Z^{-N}.$$

This equation has an irregular singular point of rank one at infinity. The asymptotic form of the solutions which are dominant and recessive on the positive real ζ axis are

$$\psi \sim e^{\pm \zeta/2}$$

and the Stokes lines in the ζ plane are given by

$$\arg\zeta = 0, \pi, 2\pi, \cdots.$$

We shall consider, in this paper, equations for which the function $Q(\zeta)$ is *single valued*. For example, Eq. (1) with either of the functions

$$q(Z) = AZ^2 + B + CZ^{-2} + DZ^{-4} + EZ^{-6}$$

or

$$q(Z) = AZ + BZ^{-2} + CZ^{-5}$$

transforms into Eq. (7) with $Q(\zeta)$ having the form

$$Q(\zeta) = -\frac{1}{4} + \frac{\nu}{2\zeta} + \frac{(\frac{1}{4} - \beta^2)}{\zeta^2} + \frac{4p}{\zeta^3} - \frac{4q^2}{\zeta^4}, \tag{8}$$

where ν, β, p, and q are parameters related to A, B, C, D, and E. Besides the irregular singular point at $\zeta = \infty$, we allow a singular point at $\zeta = 0$, which may be regular or irregular. Referring to Eq. (8), if $p = q^2 = 0$, then the singular point at $\zeta = 0$ is regular, and we have Whittaker's equation. If $p \neq 0$ or $q \neq 0$, then the singular point at $\zeta = 0$ is irregular; in the special case $\nu = 0, p = 0$, the equation can be transformed into Mathieu's equation, as discussed in Sec. VI. The Whittaker and Mathieu special cases are useful for comparing the results of the present paper with known results.

In general, we shall assume that $Q(\zeta)$ is defined by an infinite series

$$Q(\zeta) = -\frac{1}{4} + \sum_{n=1}^{\infty} Q_n \zeta^{-n} \tag{9}$$

which converges for sufficiently large $|\zeta|$.

Since the function $q(Z)$ is allowed to have singularities in the finite Z plane, the original wave problem will be assumed to have been defined along a contour which avoids these singularities. We assume that this contour passes *above* all of these singularities. For definiteness, we assume that it lies just above the real Z axis, so that *all singularities of $q(Z)$ lie on or below the real axis.*

An example is given by the Budden equation[2]

$$\frac{d^2\phi}{dZ^2} + \left(\beta_1 + \frac{\beta_2}{Z}\right)\phi = 0$$

which describes the propagation of waves in a medium whose index of refraction has a zero and an infinity. The direction in which we are to continue the solution around the singularity at $Z = 0$ is determined by including the effect of collisions in the index of refraction.

Since we have assumed that $q(Z)$ has no singularities in the upper half of the finite Z plane, the solution of Eq. (1) can be analytically continued to a path which passes from $+\infty$ to $-\infty$, always at a large distance from the origin in the upper half plane. This path maps into a path in the ζ plane which passes, at a large distance from the origin, from $\infty e^{i\alpha}$ [where $\alpha = \arg(-q_{N-2})^{1/2}$] counterclockwise to $\infty e^{i\beta}$ (where $\beta = \alpha + N\pi/2$). We are thus led to the problem of continuing the asymptotic solution $\psi(\zeta)$ which gives the appropriate asymptotic behavior to $\phi(Z)$, for $Z \to +\infty$, around counterclockwise in the ζ plane at a large distance from the origin. At least one Stokes line will be crossed, before arriving at the point which maps into $Z = -\infty$.

Figure 1 shows an example of these paths for an equation in which $q \sim q_2 Z^2$ (when $N = 4$) where q_2 is real and positive. Asymptotically, for large $|Z|$, the equation is the same as Weber's equation. The corresponding Stokes lines in the Z plane and the ζ plane are shown, and since they are defined only in terms of the asymptotic form of $q(Z)$, they are not shown in regions near the origin.

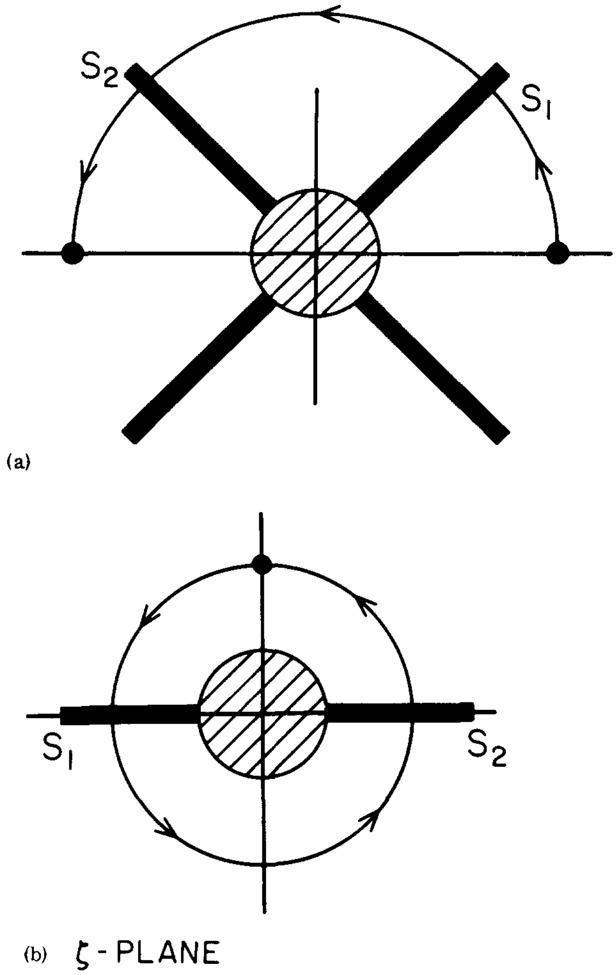


FIG. 1. Analytic continuation paths in the Z plane (a) and the ζ plane (b), for an example with $N = 4$.

B. The Stokes Phenomenon

It is natural to attempt a solution of Eq. (7), with Q given by Eq. (9), by substituting

$$\psi = \zeta^{-\rho} e^{\sigma(\zeta/2)} \sum_{n=0}^{\infty} c_n \zeta^{-n}. \tag{10}$$

One finds[3]

$$\rho = Q_1/\sigma, \quad \sigma = \pm 1,$$

and, for $n \geqslant 1$,

$$\sigma n c_n = (\rho + n)(\rho + n - 1)c_{n-1} + \sum_{m=0}^{n-1} Q_{n+1-m} c_m \tag{11}$$

with c_0 arbitrary. Such a formal solution, called a Thomés normal solution, is an asymptotic expansion of a solution in a sector bounded by Stokes lines, assuming[4] that the infinite series in Eq. (9) converges for sufficiently large $|\zeta|$. It cannot be continued across a Stokes line on which it is dominant, however.

Stokes[5] discovered that the coefficients in such asymptotic expansions must change discontinuously as a Stokes line is crossed. This is called the Stokes Phenomenon, and we shall represent it in the following form:

$$\psi \sim \zeta^{Q_1} e^{-\zeta/2} \sum_{n=0}^{\infty} c_n^{(1)} \zeta^{-n} + T_1 \theta_1(\zeta) \zeta^{-Q_1} e^{\zeta/2} \times \sum_{n=0}^{\infty} c_n^{(2)} \zeta^{-n}, \tag{12}$$

where θ_1 is a step function defined by

$$\theta_1(\zeta) = \begin{cases} 0, & \arg\zeta \leqslant \pi, \\ 1, & \arg\zeta > \pi, \end{cases}$$

and $c_0^{(2)} = c_0^{(1)}$. This gives a complete (in the sense of Watson[6]) asymptotic expansion, for $-\pi < \arg\zeta < 2\pi$, of that solution which is recessive on $\arg\zeta = 0$.

The *Stokes multiplier* T_n on the Stokes line $\arg\zeta = n\pi$ (to be denoted in what follows by S_n) may be defined as follows[7]:

(coefficient of recessive term *after* crossing S_n)
= (coefficient of recessive term *before* crossing S_n)
+ T_n x(coefficient of dominant term on S_n).

Following this definition, the asymptotic expansion for $-\pi < \arg\zeta < 5\pi$, for example, (neglecting the negative powers of ζ) is

$$\psi \sim c_0^{(1)}\{1 + T_1T_2\theta_2 + [T_1 + (1 + T_1T_2)T_3]T_4\theta_4\}\zeta^{Q_1}e^{-\zeta/2} + c_0^{(1)}[T_1\theta_1 + (1 + T_1T_2)T_3\theta_3]\zeta^{-Q_1}e^{\zeta/2}, \tag{13}$$

where θ_n is the appropriate step function on the Stokes line S_n, and T_n is the corresponding Stokes multiplier.

The connection formulas for the original wave problem are determined by the values of the Stokes multipliers. When only one Stokes line is crossed, we have, by comparing Eq. (12) with Eq. (3),

$$C^-(+\infty) = c_0^{(1)}$$

for the amplitude of the transmitted wave, and

$$C^-(-\infty) = C^-(+\infty), \quad C^+(-\infty) = T_1 C^-(+\infty)$$

for the amplitudes of the incident and reflected waves. When two Stokes lines are crossed, the incident and reflected wave amplitudes are

$$C^-(-\infty) = (1 + T_1T_2)C^-(+\infty),$$
$$C^+(-\infty) = T_1 C^-(+\infty),$$

and so on.

The Stokes multipliers $T_2, T_3, T_4, \ldots$ are simply related to the Stokes multiplier T_1, considered as a function of the parameters in the function Q in Eq. (7). We denote a solution of Eq. (7) which is recessive on $\arg\zeta = 0$ by

$$\psi_1(\zeta) = W_1(\zeta; \{Q_{2n+1}\}, \{Q_{2n}\}),$$

where the dependence on the coefficients of the odd and even powers of ζ^{-1} in Eq. (9) is indicated explicitly. Because Eq. (7) is unchanged when ζ is replaced by $\zeta e^{-i\pi}$, and the coefficients Q_{2n+1} of the odd powers of ζ^{-1} are replaced by their negatives, a second solution is

$$\psi_2(\zeta) = W_1(\zeta e^{-i\pi}; \{-Q_{2n+1}\}, \{Q_{2n}\}).$$

This solution is recessive on $\arg\zeta = \pi$, and is dominant on $\arg\zeta = 0$, so it is linearly independent of the first solution. By comparison with Eq. (13), it is easy to show that the analytic continuation of the solution ψ_1 is determined by

$$\psi_1 \sim \zeta^{Q_1} e^{-\zeta/2} + T_1 \theta_1 e^{-i\pi Q_1} \psi_2 \tag{14}$$

for $\arg\zeta < 3\pi$. In particular, it follows that the Stokes multiplier T_2 is given in terms of the function defined by

$$T_1 = T_1(\{Q_{2n+1}\}, \{Q_{2n}\})$$

by

$$T_2 = e^{-2i\pi Q_1} T_1(\{-Q_{2n+1}\}, \{Q_{2n}\}).$$

Similarly, it can be shown that

$$T_3 = e^{4i\pi Q_1} T_1, \quad T_4 = e^{-4i\pi Q_1} T_2.$$

Thus, it is sufficient to determine T_1.

In this paper, we shall obtain a convergent series expression for the Stokes multiplier T_1, as a function of the parameters in the differential equation.

III. COUPLED WAVE INTEGRAL EQUATIONS

In this section we shall derive a pair of integral equations for the amplitudes of right-going and left-going waves. The dependence of the function Q on position is responsible for coupling these waves. For large values of $|\zeta|$, the wave amplitudes approach constants, consistent with the WKBJ approximation. These integral equations are exact, and represent a reformulation of the problem of solving Eq. (7).

A. Simplifying transformation

We begin with Eq. (7) with Q given by Eq. (9) and write the solution as the product of two functions

$$\psi(\zeta) = X(\zeta)U(\zeta). \tag{15}$$

We require $X(\zeta)$ to be a solution of

$$\frac{d^2X}{d\zeta^2} + (Q + \mu^2)X = 0, \tag{16}$$

where

$$\mu = \tfrac{1}{2} - Q_1/\zeta. \tag{17}$$

Then

$$Q + \mu^2 = \frac{Q_1^2 + Q_2}{\zeta^2} + \sum_{n=3}^{\infty} Q_n \zeta^{-n} \tag{18}$$

so that Eq. (16) has a *regular* singular point at $\zeta = \infty$. Then at least one series solution of the form

$$X = \zeta^c\left(1 + \sum_{n=1}^{\infty} a_n \zeta^{-n}\right) \tag{19}$$

can be obtained by standard methods. It converges for sufficiently large $|\zeta|$, uniformly with respect to $\arg\zeta$. When Q is given by Eq. (8), for example, a solution of Eq. (16) is

$$X(\zeta) = \zeta^c e^{-2q/\zeta} M(a,b,4q/\zeta),$$

where M is the regular Kummer function,[8] which is defined by an infinite series with an infinite radius of convergence. The parameters a, b, and c are related to the parameters in Eq. (8),

$$a = \tfrac{1}{2} + (\beta^2 - \nu^2/4)^{1/2} - p/q,$$

$$b = 1 + 2(\beta^2 - \nu^2/4)^{1/2},$$

$$c = \tfrac{1}{2} - (\beta^2 - \nu^2/4)^{1/2}.$$

B. Coupled wave equations

For $|\zeta|$ sufficiently large that $X \neq 0$, the equation for U is

$$X^{-2}\frac{d}{d\zeta}X^2\frac{dU}{d\zeta} - \mu^2 U = 0. \tag{20}$$

This equation can be written in matrix form,

$$\frac{d\Phi}{d\zeta} = M\Phi, \tag{21}$$

where

$$\Phi = \begin{pmatrix} U \\ X^2 \dfrac{dU}{d\zeta} \end{pmatrix}, \quad M = \begin{pmatrix} 0 & X^{-2} \\ X^2\mu^2 & 0 \end{pmatrix}.$$

We use the method of Keller and Keller[9] to derive the coupled wave differential equations.

The eigenvalues of the matrix M are $\pm\mu$. For $|\zeta|$ sufficiently large that neither μ nor X vanishes, the matrix

$$S = \begin{pmatrix} 1 & 1 \\ \mu X^2 & -\mu X^2 \end{pmatrix}$$

can be used to transform M to diagonal form

$$S^{-1}MS = \mathrm{diag}(\mu, -\mu) \equiv \Lambda.$$

If we let

$$\Phi = S\Psi,$$

then Eq. (21) transforms to

$$\frac{d\Psi}{d\zeta} = \left(\Lambda - S^{-1}\frac{dS}{d\zeta}\right)\Psi,$$

where

$$S^{-1}\frac{dS}{d\zeta} = \frac{d}{d\zeta}\log(\mu^{1/2}X)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Finally, by defining the amplitudes A_1, A_2 by

$$\Psi_1 = A_1\mu^{-1/2}X^{-1}\exp\left(\int \mu\, d\zeta\right),$$

$$\Psi_2 = A_2\mu^{-1/2}X^{-1}\exp\left(-\int \mu\, d\zeta\right),$$

we obtain the coupled wave differential equations

$$\frac{d}{d\zeta}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \left(\frac{1}{2}\frac{\mu'}{\mu} + \frac{X'}{X}\right) \times \begin{pmatrix} 0 & \exp\left(-2\int \mu\, d\zeta\right) \\ \exp\left(2\int \mu\, d\zeta\right) & 0 \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{22}$$

In terms of the solution of Eq. (7), the amplitudes A_1 and A_2 are defined by

$$\psi = \mu^{-1/2}\left[A_1(\zeta)\exp\left(\int \mu d\zeta\right) + A_2(\zeta)\exp\left(-\int \mu d\zeta\right)\right]. \tag{23}$$

Since an arbitrary additive constant is implied by the indefinite phase integral, we may define it, using Eq. (17), by

$$\int \mu d\zeta = \tfrac{1}{2}\zeta - Q_1 \ln\zeta.$$

The lower limit of integration (the "phase reference point") has no importance in the present method, in contrast to the (approximate) phase integral method,[7] since it can always be eliminated by redefining A_1 and A_2.

With the assumption that ψ is recessive on the Stokes line $\arg\zeta = 0$, we must have

$$A_1(\zeta) \to 0, \quad \text{for}\ \zeta \to +\infty$$

while A_2 approaches a constant; we choose this constant to be unity

$$A_2(\zeta) \to 1, \quad \text{for}\ \zeta \to +\infty.$$

By integrating Eq. (22) from $+\infty$ to ζ, using the above intitial conditions, we obtain the coupled wave integral equations

$$A_1(\zeta) = \int_{+\infty}^{\zeta} d\xi\, B(\xi)\xi^{\nu} e^{-\xi} A_2(\xi), \tag{24}$$

$$A_2(\zeta) = 1 + \int_{+\infty}^{\zeta} d\xi\, B(\xi)\xi^{-\nu} e^{\xi} A_1(\xi), \tag{25}$$

where

$$\nu = 2Q_1$$

and

$$B = \frac{1}{2}\frac{\mu'}{\mu} + \frac{X'}{X}. \tag{26}$$

We emphasize that Eqs. (23)–(25) are *exact* and merely reformulate the problem of solving Eq. (7)

In some problems, it may be more convenient to begin with the more general equation

$$\frac{1}{P(\zeta)}\frac{d}{d\zeta}\left(P(\zeta)\frac{d\psi}{d\zeta}\right) + Q(\zeta)\psi = 0$$

rather than Eq. (7). In this case Eq. (16) is replaced by

$$\frac{1}{P}\frac{d}{d\zeta}\left(P\frac{dX}{d\zeta}\right) + (Q+\mu^2)X = 0$$

while, in Eqs. (20)–(22), X^2 is replaced by PX^2. In Eq. (23), $\mu^{-1/2}$ must be replaced by $\mu^{-1/2}P^{-1/2}$, but the phase integral is unchanged. We still obtain Eqs. (24) and (25), but with B defined by

$$B = \frac{1}{2}\frac{\mu'}{\mu} + \frac{1}{2}\frac{P'}{P} + \frac{X'}{X}.$$

C. Determination of the function $B(\zeta)$

One problem which must be addressed, before considering the solution of Eqs. (24) and (25), is the determination of the function $B(\zeta)$ defined by Eq. (26). Using the definition of μ, Eq. (17), we have, for $|\zeta| > |\nu|$,

$$\frac{1}{2}\frac{\mu'}{\mu} = \frac{1}{2\nu}\sum_{n=2}^{\infty}\left(\frac{\nu}{\zeta}\right)^n. \tag{27}$$

Thus, it suffices to find a similar convergent series for the function

$$v(\zeta) = X'/X. \tag{28}$$

This function satisfies the Riccati equation

$$\frac{dv}{d\zeta} + v^2 + Q + \mu^2 = 0,$$

where $Q + \mu^2$ is given by Eq. (18). By substituting

$$v(\zeta) = \sum_{n=1}^{\infty} v_n \zeta^{-n} \tag{29}$$

we find

$$-v_1 + v_1^2 + Q_1^2 + Q_2 = 0, \tag{30}$$

$$v_2 = Q_3/2(1-v_1), \tag{31}$$

and, for $n \geqslant 3$,

$$v_n = \left[\sum_{m=2}^{n-1} v_m v_{n+1-m} + Q_{n+1}\right]\Big/(n-2v_1). \tag{32}$$

By choosing v_1 to be the root of Eq. (30) with the smaller real part, we obtain one series solution which can be shown[10] to converge for sufficiently large $|\zeta|$. The recursion relations, Eqs. (31) and (32), may be used to compute the coefficients v_n; in cases where the series for $v(\zeta)$ terminates at a small number of terms, an analytical solution may be obtained. We thus obtain a convergent series (for sufficiently large $|\zeta|$) for the function $B(\zeta)$ which appears in Eqs. (24) and (25),

$$B(\zeta) = \sum_{n=1}^{\infty} B_n \zeta^{-n}, \tag{33}$$

where $B_1 = v_1$, and, for $n \geqslant 2$,

$$B_n = \tfrac{1}{2}\nu^{n-1} + v_n, \tag{34}$$

where v_n is obtained from Eqs. (30)–(32).

IV. SOLUTION OF THE INTEGRAL EQUATIONS

A. Asymptotic expansion of the solution

We shall look for solutions of Eqs. (24) and (25) of the following form, for $-\pi < \arg\zeta < 2\pi$:

$$A_1(\zeta) = \zeta^{\nu} e^{-\zeta}\left[\sum_{n=1}^{N} \alpha_n^{(1)}\zeta^{-n} + R_{11}^{(N)}(\zeta)\right] + T_1\theta_1(\zeta)\left[\sum_{n=0}^{N} \beta_n^{(2)}\zeta^{-n} + R_{12}^{(N)}(\zeta)\right], \tag{35}$$

$$A_2(\zeta) = \sum_{n=0}^{N-1} \beta_n^{(1)}\zeta^{-n} + R_{21}^{(N)}(\zeta) + T_1\theta_1(\zeta)\zeta^{-\nu} e^{\zeta}\left[\sum_{n=1}^{N} \alpha_n^{(2)}\zeta^{-n} + R_{22}^{(N)}(\zeta)\right]. \tag{36}$$

Here θ_1 is the step function on $\arg\zeta = \pi$, $\beta_0^{(1)} = \beta_0^{(2)} = 1$, the other coefficients $\alpha_n^{(1)}$, $\beta_n^{(1)}$, $\alpha_n^{(2)}$, $\beta_n^{(2)}$, and T_1 are to be determined, and for $|\zeta| \to \infty$,

$$R_{11}^{(N)} \sim \alpha_{N+1}^{(1)}\zeta^{-(N+1)}, \quad R_{12}^{(N)} \sim \beta_{N+1}^{(2)}\zeta^{-(N+1)},$$
$$R_{21}^{(N)} \sim \beta_N^{(1)}\zeta^{-N}, \quad \text{and}\ R_{22}^{(N)} \sim \alpha_{N+1}^{(2)}\zeta^{-(N+1)},$$

with N an arbitrary positive integer. These expressions are thus assumed to be asymptotic expansions of $A_1(\zeta)$ and $A_2(\zeta)$. In the limit $N \to \infty$, Eqs. (35) and (36), combined with Eq. (23), give an expression for ψ of the form given in Eq. (12).

In evaluating the integral in Eq. (24), we use the contours shown in Fig. 2. Term-by-term integration of the series in Eq. (36) will lead us to evaluate the integrals in terms of the complementary incomplete gamma function, defined by

$$\Gamma(\alpha,\zeta) = -\int_{\infty}^{\zeta} d\xi\, \xi^{\alpha-1} e^{-\xi} \tag{37}$$

We must digress briefly to discuss the asymptotic expansion of this function.

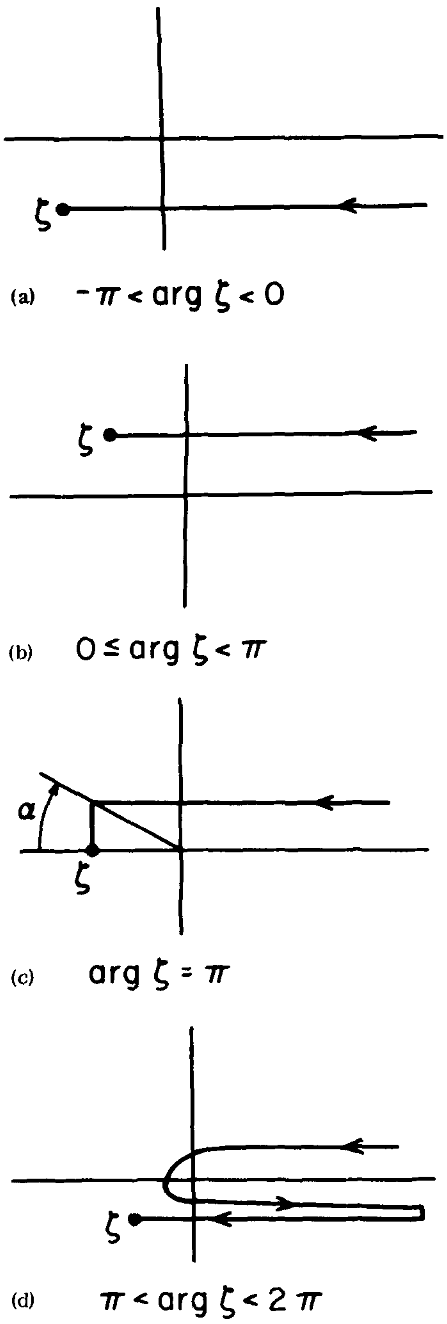


FIG. 2. Integration contours used in Eq. (24), in the complex ξ plane, for different ranges of $\arg\zeta$. In (c), α is an arbitrary positive acute angle: $0 < \alpha < \pi/2$.

Integration in Eq. (37) by parts yields the recursion relation

$$\Gamma(\alpha,\zeta) = \zeta^{\alpha-1}e^{-\zeta} - (1-\alpha)\Gamma(\alpha-1,\zeta)$$

and iteration then gives

$$e^{\zeta}\Gamma(\alpha,\zeta) = \zeta^{\alpha-1}\sum_{m=0}^{M-1}(-1)^m\frac{\Gamma(1-\alpha+m)}{\Gamma(1-\alpha)}\zeta^{-m} + (-1)^M\frac{\Gamma(1-\alpha+M)}{\Gamma(1-\alpha)}e^{\zeta}\Gamma(\alpha-M,\zeta). \tag{38}$$

The asymptotic expansion of $\Gamma(\alpha,\zeta)$ is obtained by evaluation or estimation of the last term in Eq. (38), called the remainder. The contribution to the remainder, from any of the horizontal line segments in Fig. 2, is easily estimated to be $O(|\zeta^{-(M+1-\alpha)}|)$ for $|\zeta| \to \infty$, which is of the order of the first term omitted from the series. The contribution from the vertical line segment in Fig. 2(c) can be evaluated by making a change of variable to $y = (\xi - \zeta)/i|\zeta|$, integrating by parts and using the Riemann–Lebesque lemma. It is found to be of the same order as the horizontal line segment contributions. Finally, the loop contribution in Fig. 2(d) is essentially Hankel's loop integral for the gamma function,[11] so that we obtain the asymptotic expansion (for $-\pi < \arg\zeta < 2\pi$)

$$e^{\zeta}\Gamma(\alpha,\zeta) \sim \sum_{m=0}^{\infty}(-1)^m\frac{\Gamma(1-\alpha+m)}{\Gamma(1-\alpha)}\zeta^{-m+\alpha-1} + \frac{2\pi i\theta_1(\zeta)}{\Gamma(1-\alpha)}e^{i\pi(\alpha-1)}e^{\zeta}, \tag{39}$$

where θ_1 is the step function on $\arg\zeta = \pi$.

Of course, the term which is exponentially small, for $|\zeta| \to \infty$ with $\pi/2 < \arg\zeta < 3\pi/2$, has a coefficient which is not uniquely determined by this analysis, since an exponentially small term can always be included in the remainder. In fact, Eq. (39), without the discontinuous term, is the asymptotic expansion usually given for $|\arg\zeta| < 3\pi/2$. The rays $\arg\zeta = \pm\pi/2, \pm 3\pi/2, \cdots$ across which recessive terms become dominant, and conversely, are known as anti-Stokes lines. The recessive term must be included if $\Gamma(\alpha,\zeta)$ is to be continued across the anti-Stokes line $\arg\zeta = 3\pi/2$, since it then becomes the dominant term.

The above analysis only shows that it is not inconsistent to choose the coefficient of the recessive term to be discontinuous on the Stokes line $\arg\zeta = \pi$. But there is a good reason for this choice. Note that the leading terms in Eq. (39) may be written as

$$\zeta^{-\nu}e^{\zeta/2}\Gamma(\alpha,\zeta) \sim \zeta^{\nu}e^{-\zeta/2} + t_1\theta_1(\zeta)\zeta^{-\nu}e^{\zeta/2},$$

where $\nu = (\alpha-1)/2$ and

$$t_1 = 2\pi i e^{i\pi(\alpha-1)}/\Gamma(1-\alpha)$$

is the Stokes multiplier for this function. Now a Stokes line is defined, according to Eq. (5), as a ray on which the modulus of the ratio of the recessive exponential to the dominant one, $|e^{\pm\zeta}|$, has a minimum, for a given $|\zeta|$. The asymptotic approximation which is *least discontinuous* (and therefore most accurate) is obtained by choosing the discontinuity to be on the Stokes line. Stokes[5] showed explicitly, for the Airy function, that the discontinuity in the recessive term is *not* larger than the error made in truncating the asymptotic series multiplying the dominant exponential, at its smallest term, *only* if this discontinuity occurs on the Stokes line. By using the Stokes phenomenon for $\Gamma(\alpha,\zeta)$, as given by Eq. (39), we will be led to a self-consistent determination of the Stokes multiplier T_1 which we are seeking.

When Eq. (36) is substituted into Eq. (24), we obtain terms which don't contain the factor T_1 plus terms which do contain it; we begin by considering the former terms. Using

Eq. (33) for B, and the definition

$$I_{11}(\zeta)\equiv\int_{+\infty}^{\zeta}d\xi\,B(\xi)\xi^{\nu}e^{-\xi}\Big[\sum_{n=0}^{N-1}{}'\beta_n^{(1)}\xi^{-n}+R_{21}^{(N)}(\xi)\Big]$$

we have

$$I_{11}(\zeta)=\int_{+\infty}^{\zeta}d\xi\,\xi^{\nu}e^{-\xi}\Big[\sum_{s=1}^{N}\xi^{-s}\sum_{n=0}^{s-1}\beta_n^{(1)}B_{s-n}+O\big(\beta_N^{(1)}\xi^{-(N+1)}\big)\Big].$$

Using the definition of the complementary incomplete gamma function, the sum can be integrated term by term. Since asymptotic order relations may be integrated[12] we find

$$\begin{aligned}\zeta^{-\nu}e^{\zeta}I_{11}(\zeta)&=-\zeta^{-\nu}e^{\zeta}\sum_{s=1}^{N}\Big(\sum_{n=0}^{s-1}\beta_n^{(1)}B_{s-n}\Big)\Gamma(1-s+\nu,\zeta)\\&\quad+O[\zeta^{-(N+1)}]+O[\beta_N^{(1)}\zeta^{-\nu}e^{\zeta}\Gamma(\nu-N,\zeta)].\end{aligned}$$

Using the asymptotic expansion Eq. (39), and rearranging the summation gives

$$\begin{aligned}\zeta^{-\nu}e^{\zeta}I_{11}(\zeta)&=-\sum_{p=1}^{N}(-1)^p\Gamma(p-\nu)\zeta^{-p}\sum_{s=1}^{p}\frac{(-1)^s}{\Gamma(s-\nu)}\\&\quad\times\sum_{n=0}^{s-1}{}'\beta_n^{(1)}B_{s-n}-2\pi i\theta_1(\zeta)e^{i\pi\nu}\zeta^{-\nu}e^{\zeta}\\&\quad\times\Big[\sum_{s=1}^{N}\frac{(-1)^s}{\Gamma(s-\nu)}\sum_{n=0}^{s-1}{}'\beta_n^{(1)}B_{s-n}\\&\quad+O\Big(\frac{\beta_N^{(1)}}{\Gamma(N+1-\nu)}\Big)\Big]+O(\zeta^{-(N+1)}).\end{aligned}\tag{40}$$

The integration of the terms in Eq. (36) which are proportional to T_1, is straightforward and does not involve the incomplete gamma function. The path of integration in the ξ plane is taken to be the limit, as $R\to\infty$, of a path consisting of a circular arc of radius R from $\arg\xi=0$ to $\arg\xi=\arg\zeta$, plus a straight-line segment radially inward to the point $\xi=\zeta$. The result is

$$\begin{aligned}&T_1\int_{+\infty}^{\zeta}d\xi\,B(\xi)\theta_1(\xi)\Big(\sum_{n=1}^{N}\alpha_n^{(2)}\xi^{-n}+R_{22}^{(N)}(\xi)\Big)\\&\quad=-T_1\theta_1(\zeta)\Big[\sum_{n=1}^{N}\Big(\sum_{r=1}^{n}\alpha_r^{(2)}B_{n+1-r}\Big)\frac{\zeta^{-n}}{n}\\&\qquad+O(\zeta^{-(N+1)})\Big].\end{aligned}\tag{41}$$

We now collect results from Eqs. (40) and (41), compare with Eq. (35), and obtain

$$\alpha_n^{(1)}=(-1)^{n+1}\Gamma(n-\nu)\sum_{s=1}^{n}\frac{(-1)^s}{\Gamma(s-\nu)}\times\sum_{m=0}^{s-1}{}'\beta_m^{(1)}B_{s-m},\tag{42}$$

$$\beta_n^{(2)}=-\frac{1}{n}\sum_{r=1}^{n}\alpha_r^{(2)}B_{n+1-r},\qquad n\geqslant1\tag{43}$$

and

$$T_1=-2\pi ie^{i\pi\nu}\Big[\sum_{n=1}^{N}\frac{(-1)^n}{\Gamma(n-\nu)}\sum_{m=0}^{n-1}{}'\beta_m^{(1)}B_{n-m}+O\Big(\frac{\beta_N^{(1)}}{\Gamma(N+1-\nu)}\Big)\Big].\tag{44}$$

We next substitute Eq. (35) into Eq. (25). The path of integration is taken to be the limit, for $R\to\infty$, of a path consisting of a circular arc of radius R from $\arg\xi=0$ to its intersection with a horizontal line segment through $\xi=\zeta$, plus this horizontal segment. The integration of the terms in Eq. (35) which don't contain the factor T_1, is straightforward. The terms containing the factor T_1 can be integrated in terms of the complementary incomplete gamma function by using

$$\int_{-\infty}^{\zeta}d\xi\,\xi^{\alpha-1}e^{\xi}=e^{i\pi(\alpha-1)}\Gamma(\alpha,\zeta e^{-i\pi}),$$

and we obtain

$$\begin{aligned}&\int_{\infty}^{\zeta}d\xi\,B(\xi)\xi^{-\nu}e^{\xi}A_1(\xi)\\&\quad=-\sum_{p=1}^{N-1}\frac{\zeta^{-p}}{p}\sum_{m=1}^{p}\alpha_m^{(1)}B_{p-m+1}+O(\zeta^{-N})\\&\qquad+T_1\theta_1(\zeta)\zeta^{-\nu}e^{\zeta}\Big[\sum_{s=1}^{N}\zeta^{-s}\sum_{p=1}^{s}\frac{\Gamma(s-\nu)}{\Gamma(p-\nu)}\\&\qquad\times\sum_{m=0}^{p-1}{}'\beta_m^{(2)}B_{s-m}+O(\zeta^{-(N+1)})\Big].\end{aligned}\tag{45}$$

By comparing with Eq. (36), we obtain

$$\beta_n^{(1)}=-\frac{1}{n}\sum_{m=1}^{n}\alpha_m^{(1)}B_{n-m+1},\qquad n\geqslant1\tag{46}$$

$$\alpha_n^{(2)}=\sum_{p=1}^{n}\frac{\Gamma(n-\nu)}{\Gamma(p-\nu)}\sum_{m=0}^{p-1}{}'\beta_m^{(2)}B_{p-m}.\tag{47}$$

We have thus obtained recursion relations for the determination of the coefficients in Eqs. (35) and (36), and an expression for the Stokes multiplier T_1, in terms of the solution of these recursion relations. It is convenient to define the coefficients

$$\gamma_n=\frac{(-1)^n\alpha_n^{(1)}}{\Gamma(n-\nu)},\qquad\text{for }n\geqslant1\tag{48}$$

and

$$\gamma_0=0.$$

Then from Eq. (42) we obtain

$$\gamma_n-\gamma_{n-1}=a_n,\qquad\text{for }n\geqslant1\tag{49}$$

where

$$a_n=\frac{(-1)^{n-1}}{\Gamma(n-\nu)}\sum_{m=0}^{n-1}{}'\beta_m^{(1)}B_{n-m}\tag{50}$$

and from Eq. (46) we have

$$\beta_m^{(1)}=-\frac{1}{m}\sum_{p=1}^{m}(-1)^p\Gamma(p-\nu)\gamma_pB_{m-p+1},\qquad\text{for }m\geqslant1\tag{51}$$

(Recall that $\beta_0^{(1)}=1$.)

B. The Stokes multiplier

The expression for the Stokes multiplier T_1 given in Eq. (44) can be expressed quite simply in terms of the coefficients a_n. By taking the limit $N\to\infty$, as is consistent with regarding Eqs. (35) and (36) as infinite asymptotic series, we find

$$T_1 = 2\pi i e^{i\pi\nu} \sum_{n=1}^{\infty} a_n. \tag{52}$$

Alternatively, this result may be expressed, using Eq. (49), as

$$T_1 = 2\pi i e^{i\pi\nu} \lim_{n\to\infty} \gamma_n. \tag{53}$$

The principal *result* of this paper is Eq. (52), which gives an infinite series representation for the Stokes multiplier. The terms in the infinite series are to be obtained by solving the recursion relations, Eqs. (49)–(51). The convergence of this infinite series is easily demonstrated by showing that the limit in Eq. (53) exists. Assuming that, for $p \to \infty$, γ_p either increases in absolute magnitude or approaches a nonzero limit, the series in Eq. (51) is dominated by the $p = m$ term

$$\beta_m^{(1)} \sim \frac{(-1)^{m+1}}{m} \Gamma(m-\nu)\gamma_m B_1, \quad \text{for } m \to \infty. \tag{54}$$

Then it is clear that the sum in Eq. (50) is dominated by the $m = n - 1$ term, so that

$$a_n \sim \frac{-B_1^2 \gamma_{n-1}}{(n-1)(n-1-\nu)}, \quad \text{for } n \to \infty. \tag{55}$$

Using Eq. (49), we have

$$\frac{\gamma_n}{\gamma_{n-1}} \sim 1 - \frac{B_1^2}{(n-1)(n-1-\nu)}, \quad \text{for } n \to \infty \tag{56}$$

so that, for sufficiently large M and N,

$$\frac{\gamma_N}{\gamma_M} \sim \prod_{n=M+1}^{N} \left[1 - \frac{B_1^2}{(n-1)(n-1-\nu)}\right]. \tag{57}$$

In the limit $N \to \infty$, we obtain a convergent infinite product, showing that the limit in Eq. (53) exists. If $B_1 = 0$, this argument must be modified. The convergence is even faster in this case, because

$$\frac{\gamma_n}{\gamma_{n-1}} \sim 1 - \frac{B_2^2}{(n-2)(n-1-\nu)(n-2-\nu)(n-3-\nu)}.$$

The recursion relations for $\alpha_n^{(2)}$ and $\beta_n^{(2)}$, given by Eqs. (43) and (47), are similar in form to those for $\alpha_n^{(1)}$ and $\beta_n^{(1)}$. By comparing Eqs. (35), (36), and (23) with Eq. (14), we find that $\alpha_n^{(2)}$ and $\beta_n^{(2)}$ are the coefficients in the second linearly independent solution ψ_2. These coefficients will not concern us further, however.

C. Relation to Thomé normal solutions

We now derive another useful expression for the Stokes multiplier, in terms of the coefficients c_n in the Thomé normal solutions, Eq. (10). From Eqs. (35), (36), and (23), we have

$$\psi \sim \mu^{-1/2}\Big\{\zeta^{\nu/2} e^{-\zeta/2}\Big[1 + \sum_{n=1}^{\infty} (\alpha_n^{(1)} + \beta_n^{(1)})\zeta^{-n}\Big] + T_1\theta_1(\zeta)\zeta^{-\nu/2}e^{\zeta/2}\Big[1 + \sum_{n=1}^{\infty} (\alpha_n^{(2)} + \beta_n^{(2)})\zeta^{-n}\Big]\Big\} \tag{58}$$

for $-\pi < \arg\zeta < 2\pi$, where μ is defined by Eq. (17). Using the notation

$$\mu^{1/2} = \sum_{m=0}^{\infty} \mu_m \zeta^{-m},$$

we multiply ψ, as given by Eq. (12), by $\mu^{1/2}$ and compare with Eq. (58) to obtain

$$\alpha_n^{(1)} + \beta_n^{(1)} = \sum_{m=0}^{n} c_m^{(1)} \mu_{n-m}/c_0^{(1)}. \tag{59}$$

From Eq. (11) it follows that

$$c_m/c_{m-1} = O(m), \quad \text{for } m \to \infty$$

so that the sum in Eq. (59) is dominated by the $m = n$ term

$$\alpha_n^{(1)} + \beta_n^{(1)} \sim \mu_0 c_n^{(1)}/c_0^{(1)} = c_n^{(1)}, \quad \text{for } n \to \infty$$

(by choosing $c_0^{(1)} = \mu_0 = 2^{-1/2}$). Also, since

$$\beta_n^{(1)}/\alpha_n^{(1)} \sim -B_1/n, \quad \text{for } n \to \infty,$$

we have

$$\alpha_n^{(1)} \sim c_n^{(1)}$$

and so we obtain, from Eqs. (48) and (53),

$$T_1 = 2\pi i e^{i\pi\nu} \lim_{n\to\infty} \frac{(-1)^n c_n^{(1)}}{\Gamma(n-\nu)}. \tag{60}$$

This expression is useful whenever the solution of the Thome' recursion relations, Eq. (11), is easier than the solution of Eqs. (49)–(51). An example of such a case will be found in the next section.

V. ANALYTICALLY SOLVABLE CASES

The simplest example, to which the method of the present paper can be applied, is Bessel's equation, the normal form of which may be written as

$$\frac{d^2\psi}{d\zeta^2} + \left(-\frac{1}{4} + \frac{\frac{1}{4} - \beta^2}{\zeta^2}\right)\psi = 0. \tag{61}$$

This equation results when the transformation of Eq. (6) is applied to the equation

$$\frac{d^2\phi}{dz^2} - z^p\phi = 0. \tag{62}$$

This is an equation with *one turning point* [where $q(z) = 0$] of order p; then $\beta = 1/(p+2)$ in Eq. (61). The function $B(\zeta)$, which is contained in the integral equations, Eqs. (24) and (25), is given in this case by

$$B(\zeta) = \frac{\frac{1}{2} - \beta}{\zeta}.$$

Thus, the coefficients B_n, which appear in Eqs. (49)–(51), are zero for $n > 1$, while $B_1 = \frac{1}{2} - \beta$. Also, $\nu = 0$, so that Eq. (51) gives

$$\beta_m^{(1)} = (-1)^{m+1}\frac{B_1}{m}\Gamma(m)\gamma_m$$

while Eq. (50) gives

$$a_n = \frac{(-1)^{n-1}B_1\beta_{n-1}^{(1)}}{\Gamma(n)} = -\frac{B_1^2\gamma_{n-1}}{(n-1)^2}.$$

Thus, Eq. (49) becomes the two-term recursion relation

$$\gamma_n = \gamma_{n-1}[1 - B_1^2/(n-1)^2]$$

which can be solved immediately. An expression for the Stokes multiplier then follows from Eq. (53):

$$T_1 = 2\pi i \gamma_1 \prod_{n=1}^{\infty} [1 - (\tfrac{1}{2} - \beta)^2/n^2]$$
$$= 2i\cos(\pi/(p+2)) \tag{63}$$

using $\gamma_1 = B_1$, $\beta = 1/(p+2)$, and the infinite-product representation of the trigonometric functions.[13] This expression agrees with the known result, which was obtained by a different method.[14] The most famous special case of this result is the case $p = 1$, for which Eq. (62) becomes Airy's equation, and Eq. (63) gives $T_1 = i$. This special case forms the basis of the approximation WKBJ connection formula, and the phase integral method.[7]

An analytical result for the Stokes multiplier can also be obtained in the case of Whittaker's equation

$$\frac{d^2\psi}{d\zeta^2} + \left(-\frac{1}{4} + \frac{\nu}{2\zeta} + \frac{(\frac{1}{4} - \beta^2)}{\zeta^2} \right) \psi = 0. \tag{64}$$

In this case, the Thome' recursion relation, Eq. (11), reduces to a two-term recursion relation

$$c_n = \frac{[(n - \nu/2)(n - 1 - \nu/2) + \frac{1}{4} - \beta^2]}{(-n)} c_{n-1}. \tag{65}$$

The solution can again be obtained immediately

$$\frac{c_n}{c_0} = \frac{(-1)^n}{n!} \frac{\Gamma(n + 1 - p_1)\Gamma(n + 1 - p_2)}{\Gamma(1 - p_1)\Gamma(1 - p_2)}, \tag{66}$$

where

$$p_{1,2} = \tfrac{1}{2} + \nu/2 \pm \beta. \tag{67}$$

By using Eq. (60) and the asymptotic form[15]

$$\Gamma(n + a)/\Gamma(n + b) \sim n^{a-b}, \quad \text{for } n \to \infty,$$

we obtain

$$T_1 = \frac{2\pi i e^{i\pi\nu}}{\Gamma(\frac{1}{2} - \frac{\nu}{2} - \beta)\Gamma(\frac{1}{2} - \frac{\nu}{2} + \beta)}. \tag{68}$$

This result, and the corresponding results for T_2, T_3, and T_4 obtained from the discussion in Sec. II, agree with the known results[16] obtained by methods unrelated to the present method.

VI. NUMERICAL CALCULATIONS

A. Method

The approximate evaluation of the series expression for the Stokes multiplier, as given by Eq. (52), is straightforward using a digital computer. The terms are to be obtained by solving the recursion relations, Eqs. (49)–(51). When β_m is eliminated, we obtain the following, after rearrangement:

$$\gamma_n = \gamma_{n-1} + a_n, \tag{69}$$

$$a_n = \sum_{m=1}^{n-1} (-1)^m \frac{\Gamma(n - m - \nu)}{\Gamma(n - \nu)} S_m^{(n)} \gamma_{n-m} - \frac{(-1)^n B_n}{\Gamma(n - \nu)}, \tag{70}$$

where

$$S_m^{(n)} = \sum_{r=1}^{m} \frac{B_r B_{m+1-r}}{(n - r)}. \tag{71}$$

When n is large, only a few terms are needed in the m summation, since

$$\Gamma(n - m - \nu)/\Gamma(n - \nu) \sim n^{-m}, \quad \text{for } n \to \infty.$$

Numerical evaluation of the coefficients B_n is straightforward using Eqs. (30)–(32) and Eq. (34).

The result obtained, by truncation of the infinite series in Eq. (52) after N terms, can easily be improved. Using Eq. (55) with

$$\gamma_{n-1} \simeq \gamma_\infty \equiv \lim_{n \to \infty} \gamma_n \tag{72}$$

the remaining terms can be summed approximately. The use of the trapezoidal rule, to replace the sum by an integral, leads to

$$\gamma_\infty - \gamma_N = \sum_{n=N+1}^{\infty} a_n \sim \frac{-B_1^2 \gamma_\infty}{N} + O(N^{-2}), \tag{73}$$

where

$$\gamma_N = \sum_{n=1}^{N} a_n$$

from which we obtain γ_∞, and hence T_1, with an error $O(N^{-2})$. By keeping both the $m = 1$ and $m = 2$ terms in Eq. (70), a little more effort yields an expression with an error $O(N^{-3})$. By keeping the $m = 1, 2$ and 3 terms in Eq. (70), and using the Euler–Maclaurin summation formula, we find

$$\gamma_\infty = \gamma_N / F(N) + O(N^{-4}), \tag{74}$$

where

$$F(N) = 1 + B_1^2/N + \left[\tfrac{1}{2}(\nu + 1)B_1^2 + \tfrac{1}{2}B_1^4 - B_1 B_2\right]/N^2 + \left[(\tfrac{1}{3}\nu^2 + \tfrac{1}{2}\nu + \tfrac{1}{6})B_1^2 + (\tfrac{1}{2}\nu + \tfrac{2}{3})B_1^4 + \tfrac{1}{6}B_1^6 - B_1^3 B_2 - (\tfrac{4}{3}\nu + 2)B_1 B_2 + \tfrac{2}{3}B_1 B_3 + \tfrac{1}{3}B_2^2\right]/N^3. \tag{75}$$

In the case of Whittaker's equation, comparison with the exact result verifies the order relation for the error, as given by Eq. (74). Sample calculations indicate that at least four decimal place accuracy is achieved, with N between 50 and 100.

B. An equation with two irregular singular points

For differential equations with two irregular singular points, few results are available for comparison with the results obtained by the present method. When $\nu = p = 0$ in Eq. (8), Eq. (7) can be transformed into Mathieu's equation. Comparison of known results for the characteristic exponents for Mathieu's equation, with the Stokes multiplier obtained by the method of this paper, is made possible by the following reasoning:

We consider the equation

$$\frac{d^2\psi}{d\zeta^2} + \left(-\frac{1}{4} + \frac{(\frac{1}{4} - \beta^2)}{\zeta^2} - \frac{4q^2}{\zeta^4} \right) \psi = 0. \tag{76}$$

Since $Q(\zeta)$ is a single-valued analytic function for $|\zeta| > \epsilon > 0$, it is known[3] that a solution of Eq. (76) exists in the form

$$\psi_1(\zeta) = \zeta^{-\rho} \Psi_1(\zeta),$$

where $\Psi_1(\zeta)$ is a single-valued analytic function for $|\zeta| > \epsilon$. Since, however, $Q(\zeta)$ contains only even powers of ζ, the function

$$\psi_2(\zeta) = (\zeta e^{i\pi})^{-\rho} \Psi_1(\zeta e^{i\pi})$$

is also a solution, and so is the function

$$\psi(\zeta) = \psi_1(\zeta) + e^{i\pi\rho}\psi_2(\zeta).$$

This latter solution has the form

$$\psi(\zeta) = \zeta^{-\rho}\Psi(\zeta), \tag{77}$$

where Ψ is an analytic function such that

$$\Psi(\zeta e^{i\pi}) = \Psi(\zeta). \tag{78}$$

We now make the change of variables

$$\psi = \zeta^{1/2}y(\theta), \quad \zeta = \zeta_0 e^{i\theta}, \tag{79}$$

where

$$\zeta_0 = 2iq^{1/2} \tag{80}$$

and θ is real, corresponding to a value of ζ which lies on the circle $|\zeta| = |\zeta_0|$. Then Eq. (76) is transformed to

$$\frac{d^2y}{d\theta^2} + (\beta^2 - 2q\cos 2\theta)y = 0 \tag{81}$$

which is Mathieu's equation. It follows, from the results of the preceding paragraph, that a solution of Mathieu's equation exists in the form

$$y(\theta) = e^{i\nu\theta}Y(\theta), \tag{82}$$

where

$$Y(\theta) = \Psi(\zeta_0 e^{i\theta})$$

is periodic with period π,

$$Y(\theta + \pi) = Y(\theta). \tag{83}$$

This is Floquet's theorem; the characteristic exponent ν is related to the value of ρ, in Eq. (77), by

$$\nu = -\rho - \tfrac{1}{2}. \tag{84}$$

On the other hand, the value of ρ is related to the Stokes multiplier T_1 for Eq. (76), as follows: The solution of Eq. (76) which is recessive on $\arg\zeta = 0$ is

$$\psi^{(1)}(\zeta) = A_1^{(1)}e^{\zeta/2} + A_2^{(1)}e^{-\zeta/2}.$$

The A's satisfy Eqs. (24) and (25) where B is given by Eq. (26) with $\mu = \frac{1}{2}$ and

$$X = \zeta^{1/2}J_\beta(2iq/\zeta),$$

where J_β is a Bessel function. A second, linearly independent, solution is

$$\psi^{(2)}(\zeta) = \psi^{(1)}(\zeta e^{-i\pi}).$$

The result of analytically continuing these solutions, along a path from $\zeta = \zeta_1$ to $\zeta = \zeta_1 e^{i\pi}$, where $\arg\zeta_1 = \pi/2$, is given by the circuit relation

$$\begin{pmatrix}\psi^{(1)}(\zeta_1 e^{i\pi})\\ \psi^{(2)}(\zeta_1 e^{i\pi})\end{pmatrix} = \begin{pmatrix}T_1 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi^{(1)}(\zeta_1)\\ \psi^{(2)}(\zeta_1)\end{pmatrix}, \tag{85}$$

where T_1 is the Stokes multiplier on $\arg\zeta = \pi$. A linear combination of $\psi^{(1)}$ and $\psi^{(2)}$

$$\psi = a\psi^{(1)} + b\psi^{(2)}$$

is of the form given by Eqs. (77) and (78) provided that a and b are the elements of an eigenvector of the matrix in Eq. (85), with $e^{-i\pi\rho}$ the corresponding eigenvalue. Since the product of the two eigenvalues is -1, and their sum is T_1, we have

$$T_1 = e^{-i\pi\rho} - e^{i\pi\rho}.$$

Finally, using Eq. (84), we have the relation between the Stokes multiplier T_1 and the Floquet characteristic exponent ν

$$T_1 = 2i\cos\pi\nu. \tag{86}$$

The following known results were used to compare Eq. (86) with results which were numerically calculated using the method of this paper:

(a) Periodic solutions

When $\nu = n$, an integer, relations between the coefficients in Mathieu's equation exist

$$\beta^2 = 1 + q - q^2/8 + \cdots \quad (n = 1)$$
$$\beta^2 = 4 - q^2/12 + \cdots \quad (n = 2)$$

(b) Stable nonperiodic solutions

When $\nu \neq n$, β, ν, and q are related by

$$\beta^2 = \nu^2 + q^2/2(\nu^2 - 1) + \cdots$$

for small q.

(c) Stable and unstable nonperiodic solutions for small β^2

The relation between β, ν, and q is

$$\cos\pi\nu = (1 - \pi^2\beta^2/2 + \cdots) - \pi^2 q^2/4 + \cdots$$

for small q. The full expressions given by Eqs. (20.2.25), (20.3.15), and (20.3.18) of Ref. 8 were used in all cases. Good agreement was obtained, within the regions of validity of these approximate formulas. It seems very likely that the present method gives results which are reasonably accurate over a wider range of parameters than these approximate formulas.

C. Discussion

The recursion relations derived in this paper provide an efficient means for calculating numerically the Stokes multipliers for a class of ordinary differential equations. The main limitation seems to be that the numerical values of the parameters in the differential equation must not be much larger than unity. The rate of convergence seems to be quite slow, when these parameters are large. On the other hand, that is precisely when the more sophisticated analytical approximate methods[17] work best. Thus, the method of this paper and the analytical phase integral methods are complementary. The present method may be expected to give more accurate results for the smallest eigenvalue in wave problems, for which the phase integral method is least accurate.

Although the eigenvalues for a wave problem are related to the Stokes multiplier, the structure of the eigenfunctions is not obtained by calculating only the Stokes multiplier. However, the eigenfunctions may be obtained by solving an initial-value problem, once the eigenvalues are known, and this is a straightforward numerical problem which does not require iteration.

The computer time required for the calculation of Stokes multipliers by this method seems to be much less than is required by another, recently published, method.[18]

ACKNOWLEDGMENT

This work was supported in part by the U.S. Department of Energy Contract EY-76-C-05-4478.

[1]F.W.J. Olver, *Asymptotics and Special Functions* (Academic, New York, 1974), p. 190.
[2]K.G. Budden, *Radio Waves in the Ionosphere* (Cambridge U.P., London, 1961), p. 476.
[3]A. Erdelyi, *Asymptotic Expansions* (Dover, New York, 1956), pp. 58–63.
[4]Ref. 1, p. 232.
[5]G.G. Stokes, Trans. Cambridge, Philos. Soc. **10**, 105 (1864).
[6]G.N. Watson, Philos. Trans. Roy. Soc. London A **211**, 279 (1911); see also R.B. Dingle, Proc. Roy. Soc. London, Ser. A **244**, 456 (1958).
[7]J. Heading, *An Introduction to Phase-Integral Methods* (Methuen, London, 1962), p. 57.
[8]M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965), p. 504.
[9]H.B. Keller and J.B. Keller, J. Soc. Indust. Appl. Math. **10**, 246 (1962).
[10]E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956), pp. 295–296; the Riccati equation is easily transformed into the equation of Briot and Bouquet, discussed in this reference.
[11]E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis* (Cambridge U.P., London, 1927), p. 245.
[12]Ref. 1, p. 8.
[13]I.S. Gradshteyn and I.W. Ryzhik, *Tables of Integrals Series and Products* (Academic, New York, 1965), p. 37.
[14]Ref. 7, p. 63.
[15]Ref. 8, p. 257.
[16]J. Heading, J. London Math. Soc. **37**, 195 (1962).
[17]F.W.J. Olver, Philos, Trans. Roy Soc. London, Ser. A **289**, 501 (1978).
[18]A.M. Emamzadeh, J. Inst. Math. Appl. **19**, 77, 149 (1977).

Gel'fand–Levitan equations with comparison measures and comparison potentials

H. E. Moses[a)]
University of Lowell, College of Pure and Applied Science, Center for Atmospheric Research, 450 Aiken Street, Lowell, Massachusetts 01854

(Received 22 September 1978)

Using an abstract form of the Gel'fand–Levitan equation, it is shown how a solution of the equation corresponding to a given weight operator can be found in terms of a solution for the equation with a different weight operator. The resulting Gel'fand–Levitan equation is a generalization of the original one. To achieve our result, an analog of a canonical transformation for direct scattering is used. The effect of the use of the transformation is to include part of the scattering potential (the comparison potential) in the unperturbed Hamiltonian. The generalized Gel'fand–Levitan equation has the advantage that if the weight operator for a given Gel'fand–Levitan equation is close to that for an already solved Gel'fand–Levitan equation, the solution of the first can be obtained from the second by using the solution of the second as a first approximation in an iteration procedure or as a trial function in a variational procedure. The method is illustrated by considering the inverse problem for the one-dimensional Schrödinger equation, a generalized radial Schrödinger equation, and the Marchenko equation for the zero angular momentum radial Schrödinger equation. Though the use of a comparison weight function for some of the cases above has been given by others, the work of the present paper represents a systematic approach to the problem. The role of a variational principle will also be discussed.

1. INTRODUCTION

In Refs. 1 and 2 an abstract formulation of the Gel'fand–Levitan algorithm was given in which a Hamiltonian H could be obtained from the weight operator (essentially the spectral measure function of H) and from the knowledge of the spectrum of the unperturbed operator H_0 (in scattering problems identified with the kinetic energy). It was shown that this abstract formalism contained as special cases the Gel'fand–Levitan equations for the one-dimensional Schrödinger equation,[3] the radial Schrödinger equation,[4] the generalized radial Schrödinger equation,[5] the Marchenko version of the inverse problem for the radial Schrödinger equation,[4] and versions of the inverse problem for the three-dimensional Schrödinger equation.[6,7] It is the object of the present paper to offer a generalization of the abstract Gel'fand–Levitan algorithm and then apply the generalization to some of the above cases.

In this generalization it is assumed that for a particular weight operator (which we call the "comparison" operator) the corresponding inverse problem has been solved, and thus the solution of the Gel'fand–Levitan equation and the scattering potential (called the "comparison potential") have been obtained. Now, instead of using the weight operator of the kinetic energy and the weight operator associated with a general Hamiltonian as the given quantities for the inverse problem for the general Hamiltonian, we use the weight operators for the solved inverse problem and the general Hamiltonian. The Gel'fand–Levitan kernel obtained from the generalized algorithm is now used to construct an operator which maps the eigenfunctions of the Hamiltonian obtained from the comparison weight operator to those of the general Hamiltonian. Moreover, the total scattering potential is the sum of the comparison potential and an increment obtained from the new kernel.

The method used in the inverse problem of the present paper is an analog of the method of using canonical transformations in the direct scattering problem to introduce the interaction picture. One recollects that in the direct problem one transforms away the kinetic energy (or the sum of the kinetic energy and a portion of the scattering potential) and thereby expresses the scattering problem in terms of the potential (or remainder of the potential). The advantage of the use of the generalized method of this paper is analagous to the advantage of using the interaction picture in direct scattering problems. If the comparison weight operator is close to the weight operator of the Hamiltonian which we wish to find, the Gel'fand–Levitan kernel of the generalized algorithm will be small, and it might be possible to obtain it in terms of a perturbation expansion in terms of the difference of the two weight operators. Moreover, since (as will be shown) a variational principle given in Ref. 8 is valid for the kernel of the generalized algorithm, it could give more accurate results for a given amount of work than the use of the variational principle for the original version of the Gel'fand–Levitan algorithm. Still another advantage is that the difference between the potential which we seek and the comparison potential can be interpreted as the error in the potential due to the error in the weight operator, which is taken as the

a)Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-77-3169.

 0022-2488/79/102047-07$01.00

difference between the comparison operator and the given one.

Finally we shall show how the generalized algorithm looks in cases mentioned above, namely for the generalized radial equation, the one-dimensional equation, and the Marchenko equation. Though the use of a comparison potential and comparison weight operator have been given before for the usual radial Schrödinger equation,[9,10] the one-dimensional equation,[11-13] and the Marchenko equation,[13] they have not been obtained as particular applications of a general formalism, as in the present paper. The present exposition thus unifies previous treatments and can be adapted to other inverse problems. Moreover, as will be seen in later papers, the general formalism of the present paper lends itself to applications other than for the inverse problem.

2. THE GENERALIZED GEL'FAND–LEVITAN ALGORITHM IN AN ABSTRACT FORMULATION

For the sake of brevity we shall write this paper as a direct continuation of Refs. 1 and 2 and shall modify the notation only slightly to conform to the problem of the present paper. As in the earlier papers H_0 is the unperturbed Hamiltonian which for scattering problems is usually identified as the kinetic energy. It will be convenient, for the sake of a later application and for other generalizations, to assume H_0 has point eigenvalues in addition to the continuous spectrum, these point eigenvalues being denoted by E_{0i} and the corresponding eigenfunctions $|H_0;E_{0i}\rangle$. (For point eigenvalues generally we shall not need the degeneracy operator A_0 and therefore will not indicate it). Then, instead of the completeness relation (2.3) of Ref. 2, we shall use

$$\iint |H_0,A_0;E,a\rangle\, dE\, da\langle H_0,A_0;E,a| + \sum_i (C_{0i})^{-1}|H_0;E_{0i}\rangle\langle H_0;E_{0i}| = \eta(H_0), \tag{1}$$

where C_{0i} are the normalizations of $|H_0;E_{0i}\rangle$ and where the integration over E is over the continuous spectrum of H_0 $\eta(H_0)$ is the identity operator in Hilbert space (as opposed to the q-extended space) as in Ref. 2. The generalization of the present paper could have been accomplished by including the comparison potential in H_0 and using a comparison weight operator for the continuous spectrum corresponding to appropriate boundary conditions. Indeed this was the procedure which the author used for introducing comparison weight operators and potentials for the one-dimensional problem in Ref. 11. However, the procedure of the present paper is more satisfying.

We now introduce a set of scattering potentials V_j, which are labelled by the subscript $j = 1,2,3,\cdots$. We reserve $j = 0$ for the unperturbed Hamiltonian and thus write $V_0 \equiv 0$. The total Hamiltonians associated with these scattering potentials are written H_j, where

$$H_j = H_0 + V_j, \tag{2}$$

in which here and later we set ϵ of Refs. 1 and 2 equal to unity for simplicity. In addition to a continuous spectrum which coincides with that of H_0, the operators may have discrete spectra with point eigenvalues E_{ji}. Since for each j the point eigenvalues of H_j will differ in position and number, the index i, which labels their number, will vary in range with j.

The continuous spectrum has a weight operator $\langle a|\omega_{cj}(E)|a'\rangle$ while the eigenfunctions corresponding to the point eigenvalues E_{ji} will have the normalizations C_{ji}. The completeness relations for the continuous spectrum eigenfunctions $|H_j, A_j;E,a\rangle$ and discrete spectrum eigenfunctions $|H_j;E_{ji}\rangle$ are

$$\int dE \iint |H_j, A_j;E,a\rangle\, da\, \langle a|\omega_{cj}(E)|a'\rangle\, da'\langle H_j, A_j;E,a'| + \sum_i (C_{ji})^{-1}|H_j;E_{ji}\rangle\langle H_j;E_{ji}| = \eta(H_0). \tag{3}$$

As in Refs. 1 and 2 we note that wave operators U_j can be introduced such that

$$|H_j,A_j;E,a\rangle = U_j|H_0,A_0;E,a\rangle, \quad |H_j;E_{ji}\rangle = U_j|H_0;E_{ji}\rangle, \tag{4}$$

where $|H_0;E_{ji}\rangle$ formally satisfy the eigenfunction equation $H_0|H_0;E_{ji}\rangle = E_{ji}|H_0;E_{ji}\rangle$. They are not true eigenfunctions of H_0 unless E_{ji} equals an eigenvalue $E_{0i'}$ of H_0 for some i and i' in which case it is an eigenket of H_0 belonging to the point eigenvalue $E_{0i'}$. As in Ref. 2 they are needed to complete the q-space.

The completeness relation (3) is equivalent to the relation

$$U_jW_jU_j^* = \eta(H_0). \tag{5}$$

In Eq. (5) W_j is the weight operator (a generalization of the notion of measure) given explicitly by

$$W_j = \int dE \iint |H_0,A_0;E,a\rangle \times da\langle a|\omega_{cj}(E)|a'\rangle\, da'\langle H_0,A_0;E,a'| + \sum_i (C_{ji})^{-1}|H_0;E_{ji}\rangle\langle H_0;E_{ji}|. \tag{6}$$

We also introduce $U_{0,j} = U_j^{-1}$.

As shown in Refs. 1 and 2, Eq. (5) is equivalent to the Gel'fand–Levitan equation. For on writing

$$U_j = I + K_j \quad \text{and} \quad U_{0,j} = I + K_{0,j} \tag{7}$$

and requiring K_j and $K_{0,j}$ to satisfy the triangularity conditions

$$\langle q|K_j|q'\rangle = \langle q|K_{0,j}|q'\rangle = 0 \quad \text{for } q' > q, \tag{8}$$

we obtain the Gel'fand–Levitan equation for K_j from Eq. (5) (see Refs. 1 and 2):

$$\langle q|K_j|q'\rangle = -\langle q|\Omega_j|q'\rangle - \int_{q_0}^{q}\langle q|K_j|q''\rangle\, dq''\, \langle q''|\Omega_j|q'\rangle, \tag{9}$$

for $q \geqslant q'$, where

$$\Omega_j = W_j - \eta(H_0). \tag{10}$$

Equation (4) is then

$$\langle q|H_j,A_j;E,a\rangle = \langle q|H_0,A_0;E,a\rangle + \int_{q_0}^{q}\langle q|K_j|q'\rangle\, dq'\, \langle q'|H_0, A_0;E,a\rangle,$$

$$\langle q|H_j;E_{ji}\rangle = \langle q|H_0;E_{ji}\rangle + \int_{q_0}^{q}\langle q|K_j|q'\rangle\, dq'\, \langle q'|H_0;E_{ji}\rangle. \tag{11}$$

Equations (11) imply the following boundary conditions on the eigenfunctions of H_j:

$$\lim_{q\to q_0} [\langle q|H_j,A_j;E,a\rangle - \langle q|H_0,A_0;E,a\rangle] = 0,$$

$$\lim_{q\to q_0} [\langle q|H_j;E_{ji}\rangle - \langle q|H_0;E_{ji}\rangle] = 0. \quad (12)$$

It is convenient to define U_{j0}, $U_{0,j0}$, W_{j0}, K_{j0}, $K_{0,j0}$ in the following way:

$$U_{j0} = U_j, \quad U_{0,j0} = U_{0,j}, \quad K_{j0} = K_j, \quad K_{0,j0} = K_{0,j}. \quad (13)$$

We now introduce our analog of a canonical transformation. We define

$$U_{jk} = U_{j0}U_{0,k0} = U_{j0}U_{k0}^{-1},$$
$$U_{0,jk} = U_{jk}^{-1} = U_{k0}U_{0,j0} = U_{kj},$$
$$W_{jk} = U_{k0}W_{j0}U_{k0}^{*}, \quad K_{jk} = U_{jk} - I,$$
$$K_{0,jk} = U_{0,jk} - I, \quad \Omega_{jk} = U_{k0}[W_{j0} - W_{k0}]\, U_{k0}^{*}. \quad (14)$$

In Eq. (14), the asterisk means Hermitian adjoint.

The following theorem is the principal result of the present paper.

Theorem: (a) The operators K_{jk} and $K_{0,jk}$ have the same triangularity properties (8) as K_{j0}, $K_{0,j0}$.

(b) The kernels $\langle q|K_{jk}|q'\rangle$ satisfy the generalized Gel'fand–Levitan equation (for $q\geqslant q'$)

$$\langle q|K_{jk}|q'\rangle = -\langle q|\Omega_{jk}|q'\rangle - \int_{q_0}^{q}\langle q|K_{jk}|q''\rangle\langle q''|\Omega_{jk}|q'\rangle. \quad (15)$$

(c) The operators U_{jk} transform the eigenfunctions of H_k to those of H_j, i.e.,

$$|H_j, A_j;E,a\rangle = U_{jk}|H_k, A_k;E,a\rangle, \quad |H_j;E_{ji}\rangle = U_{jk}|H_k;E_{ji}\rangle. \quad (16)$$

or, equivalently,

$$\langle q|H_j,A_j;E,a\rangle = \langle q|U_{jk}|H_k,A_k;E,a\rangle$$
$$= \langle q|H_k,A_k;E,a\rangle + \int_{q_0}^{q}\langle q|K_{jk}|q''\rangle\, dq''\, \langle q''|H_k,A_k;E,a\rangle,$$
$$\langle q|H_j;E_{ji}\rangle = \langle q|U_{ij}|H_k;E_{ji}\rangle$$
$$= \langle q|H_k,E_{ji}\rangle + \int_{q_0}^{q}\langle q|K_{jk}|q''\rangle\, dq''\, \langle q''|H_k;E_{ji}\rangle. \quad (16')$$

In Eqs. (16) and (16') $|H_k;E_{ji}\rangle = U_{k0}|H_0;E_{ji}\rangle$. Hence $|H_k;E_{ji}\rangle$ is a formal eigenket of H_k (i.e., satisfies $H_k|H_k;E_{ji}\rangle = E_{ji}|H_k;E_{ji}\rangle$) having the boundary condition $\lim_{q\to q_0}[\langle q|H_k;E_{ji}\rangle - \langle q|H_0;E_{ji}\rangle] = 0$. The ket $|H_k;E_{ji}\rangle$ is a true eigenket of H_k only if $E_{ji} = E_{ki'}$ for some i' (i.e., if E_{ji} is a point eigenvalue of H_k).

(d) $U_{kp}U_{pm} = U_{km}$. (17)

(e) $\langle q|K_{kp}|q\rangle + \langle q|K_{pm}|q\rangle = \langle q|K_{km}|q\rangle$, (18)

$$\langle q|K_{jk}|q\rangle = -\langle q|K_{kj}|q\rangle. \quad (19)$$

(f) $W_{jk} = U_{km}W_{jm}W_{km}^{*}$,
$\Omega_{jk} = U_{km}[W_{jm} - W_{km}]\, U_{km}^{*}$ for any m. (20)

(g) $W_{jj} = \eta(H_0)$, $U_{jj} = I$,

$$U_{0,jj} = I, \quad K_{jj} = K_{0,jj} = 0. \quad (21)$$

Before proving the theorem we shall discuss it briefly. The generalized Gel'fand–Levitan equation (15) is seen to depend on the difference between the weight operators W_{j0} and W_{k0} which are the weight operators for the inverse problem which we wish to solve and for the inverse problem which has been solved. If the difference is small, we expect K_{jk} to be small. In this sense the use of Eq. (15) is analagous to the use of the interaction picture for direct scattering problems.

The driving kernel $\langle q|\Omega_{jk}|q'\rangle$ of Eq. (15) can be written in terms of the eigenfunctions of H_k (including the formal eigenkets $|H_k;E_{ji}\rangle$) as follows:

$$\langle q|\Omega_{jk}|q'\rangle$$
$$= \int dE \iint \langle q|H_k,A_k;E,a\rangle\, da[\langle a|\omega_{cj}(E)|a'\rangle - \langle a|\omega_{ck}(E)|a'\rangle]\, da'\langle H_k,A_k;E,a'|q'\rangle$$
$$+ \sum_i (C_{ji})^{-1}\langle q|H_k;E_{ji}\rangle\langle H_k;E_{ji}|q'\rangle$$
$$- \sum_i (C_{ki})^{-1}\langle q|H_k;E_{ki}\rangle\langle H_k;E_{ki}|q'\rangle. \quad (22)$$

This result follows from Eqs. (6) and (16)

A second point is that the theorem is given in a "covariant" formulation. That is, the Hamiltonian H_o plays a favored role only in giving the boundary conditions on the eigenfunctions of H_j as in Eq. (12), and even here one could have given the boundary condition with respect to any fixed Hamiltonian H_k.

Equations (18) and (19) are of importance in relating the scattering potentials. In the applications discussed in the present paper the scattering potentials V_j are given in terms of the Gel'fand–Levitan kernel $\langle q|K_j|q\rangle \equiv \langle q|K_{j0}|q\rangle$ either as

$$V_j(q) = 2\frac{d}{dq}\langle q|K_{j0}|q\rangle \quad (23)$$

or as

$$V_j(q) = -2\frac{d}{dq}\langle q|K_{j0}|q\rangle. \quad (24)$$

Let us define V_{j0} and V_{jk} by

$$V_{j0}(q) \equiv V_j(q), \quad V_{jk}(q) = \pm\frac{d}{dq}\langle q|K_{jk}|q\rangle, \quad (25)$$

where the plus or minus sign holds according as Eq. (23) or (24) is true.

Then from Eqs. (18) and (19)

$$V_{kp}(q) + V_{pm}(q) = V_{km}(q), \quad V_{jk}(q) = -V_{kj}(q). \quad (26)$$

In particular, let p be zero in the first of Eq. (26). Then, on changing indices in an obvious manner,

$$V_{jk}(q) = V_j(q) - V_k(q). \quad (27)$$

Thus

$$H_j = H_k + V_{jk}. \quad (27')$$

Thus the generalized Gel'fand–Levitan equation shows how the difference in the spectra of H_j and H_k leads to a difference in the potentials V_j and V_k. In Ref. 11 we discuss in more detail for the example considered (but which applies more generally) how changes in the spectrum of H_k—for example, the addition of point eigenvalues—affect the potential of V_k.

Equation (17) is reminiscent of conditions satisfied by the conditional probabilities of a Markov chain. One can introduce the notion of "paths" as in Ref. 11. This notion suggests that there might be alternative ways of solving the inverse problem. Also, though we have treated the subscripts as though they came from a denumerable set, the subscripts may also be regarded as being values of a continuous variable. Perhaps this idea will also prove to be useful in getting other relations associated with the inverse spectral problem.

We shall now prove portions of the theorem which are not self-evident. For simplicity we shall assume that all the operators H_i have a continuous spectrum only. Then $\eta(H_0) = I$, since the Hilbert space and the q-extended space are identical.

To prove part (a) of the theorem, we note that from the first of Eqs. (14)

$$\langle q|K_{jk}|q'\rangle = \langle q|K_{j0}|q'\rangle + \langle q|K_{0,k0}|q'\rangle + \eta(q-q')\int_{q'}^{q}\langle q|K_{j0}|q''\rangle \times dq''\langle q''|K_{0,k0}|q'\rangle, \tag{28}$$

where $\eta(x)$ is the Heaviside function, $\eta(x) = 1$ if $x>0$, $=0$ if $x<0$. It follows immediately the $\langle q|K_{jk}|q'\rangle = 0$ if $q'>q$.

To prove part (b), we note from Eq. (14) that

$$U_{jk}[U_{k0}W_{j0}U^*_{k0}]\,U^*_{jk} = \eta(H_0), \tag{29}$$

or

$$U_{jk}[U_{k0}W_{j0}U^*_{k0} - U_{k0}W_{k0}U^*_{k0} + \eta(H_0)] = \eta(H_0)\,U^*_{0,jk}, \tag{30}$$

or

$$U_{jk}\Omega_{jk} + U_{jk}\eta(H_0) = \eta(H_0)\,U^*_{0,jk}. \tag{31}$$

We now write $U_{jk} = \eta(H_0) + K_{jk}$, $U^*_{0,jk} = \eta(H_0) + K^*_{0,jk}$, and use $\langle q|\eta(H_0)|q'\rangle = \delta(q-q')$ and the triangularity properties of $\langle q|K_{jk}|q'\rangle$, $\langle q|K_{0,jk}|q'\rangle$ to obtain the generalized Gel'fand–Levitan equation (15) in a manner identical to the obtaining of the Gel'fand–Levitan equation of Ref. 1.

The proof of the remaining parts of the theorem are obvious, though the obviousness is due to the choice of notation.

Finally, we note that the variational principle of Ref. 8 applies to Eq. (15). This result follows from the fact that $\langle q|\eta(H_0) + \Omega_{jk}|q'\rangle \equiv \delta(q-q') + \langle q|\Omega_{jk}|q'\rangle$ is the kernel of a positive-definite operator in q-space precisely in the sense of Ref. 8.

We shall now apply the above formalism to the one-dimensional Schrödinger equation, the generalized radial Schrödinger equation, and the Marchenko equation. For the sake of brevity we shall assume acquaintance with the results of earlier references.

3. APPLICATION TO THE INVERSE PROBLEM FOR THE ONE-DIMENSIONAL SCHRÖDINGER EQUATION

We shall follow the notation and results of Ref. 3. In the present case we identify the variable q with x, where $-\infty < x < \infty$. The triangularity condition on the kernels $\langle x|K_{jk}|x'\rangle$ is

$$\langle x|K_{jk}|x'\rangle = 0, \quad x'>x. \tag{32}$$

The variable a is the sign of the momentum and takes on only two values $a = \pm 1$. From Ref. 3

$$\langle 1|\omega_{cj}(E)|1\rangle = \langle -1|\omega_{cj}(E)|-1\rangle = 1,$$
$$\langle 1|\omega_{cj}(E)|1\rangle = b_j(p), \quad \langle -1|\omega_{cj}(E)|1\rangle = b^*_j(p), \tag{33}$$

where $p = E^{1/2}$ and $b_j(p)$ is the reflection coefficient (from the left) for the potential V_j. Through $b_j(p)$ is defined for $p>0$, it can be defined for $p<0$ using analytic continuation and the fact that from analytic continuation $b_j(-p) = b^*_j(p)$.

We define

$$\psi_j(x|p) = (2|p|)^{1/2}\langle x|H_j,A_j;E,a\rangle \quad (p = aE^{1/2}), \tag{34}$$
$$\psi_{ki}(x) = \langle x|H_k;E_{ki}\rangle, \psi_{jki}(x) = \langle x|H_k;E_{ji}\rangle. \tag{35}$$
$$\psi_{ki}(x) \equiv \psi_{kki}(x). \tag{35'}$$

The Gel'fand–Levitan equation (15) becomes

$$\langle x|K_{jk}|x'\rangle = -\langle x|\Omega_{jk}|x'\rangle - \int_{-\infty}^{x}\langle x|K_{jk}|x''\rangle\,dx''\,\langle x''|\Omega_{jk}|x'\rangle, \tag{36}$$

with

$$\langle x|\Omega_{jk}|x'\rangle = \int_{-\infty}^{\infty}\psi^*_k(x|p)\psi^*_k(x'|p)\{b_j(p) - b_k(p)\}\,dp + \sum_i\frac{\psi_{jki}(x)\psi_{jki}(x')}{C_{ji}} - \sum_i\frac{\psi_{ki}(x)\psi_{ki}(x')}{C_{ki}}. \tag{37}$$

In Eq. (37) we require the eigenfunctions $\psi_k(x|p)$ to satisfy $\lim_{x\to-\infty}\psi_k(x|p) = \psi_0(x|p) = (2\pi)^{-1/2}e^{ipx}$ because of the triangularity conditions on the kernels $\langle x|K_{jk}|x'\rangle$. The amplitude of $\psi_0(x|p)$ is determined by the completeness relation Eq. (1), noting that $H_0 = -d^2/dx^2$ has no point eigenvalue.

It is convenient to require $\psi_{k0i}(x) = \exp[(-E_{ki})^{1/2}x]$ so that
$$\lim_{x\to-\infty}\psi_{jki}(x) = \psi_{j0i}(x) = \exp[(-E_{ji})^{1/2}x].$$
With these boundary conditions $\psi_k(x|-p) = \psi^*_k(x|p)$; $\psi_{jki}(x)$, $\langle x|K_{jk}|x'\rangle$ are real.

Because $V_k(x) = 2(d/dx)\langle x|K_{k0}|x\rangle$ it follows that

$$V_{jk}(x) = 2\frac{d}{dx}\langle x|K_{jk}|x\rangle = V_j(x) - V_k(x), \tag{38}$$

which is what Eq. (25) becomes in the present case.

Finally, Eq. (16') becomes

$$\psi_j(x|p) = \psi_k(x|p) + \int_{-\infty}^{x}\langle x|K_{jk}|x'\rangle\,dx'\,\psi_k(x'|p),$$
$$\psi_{ji}(x|p) = \psi_{jki}(x) + \int_{-\infty}^{x}\langle x|K_{jk}|x'\rangle\,dx'\,\psi_{jki}(x'). \tag{39}$$

It should be mentioned that Eqs. (36), (37), and (38) are given in Ref. 12 without proof and in Ref. 13 as a limit of a discretized formulation. The thrusts of Refs. 12 and 13 are quite different from those of Ref. 11 and the present paper, however.

4. APPLICATION TO THE GENERALIZED RADIAL SCHRÖDINGER EQUATION

The generalized radial Schrödinger equation is discussed in Ref. 5 for both the direct and inverse problems. The variable q is replaced by the variable r ($0 < r < \infty$). The operators H_k are given by

$$H_k = -\frac{d^2}{dr^2} + V_k(r), \quad V_0(r) \equiv 0. \tag{40}$$

Instead of requiring the functions of the domain of H_k to satisfy the condition $f(0) = 0$, our generalization consists of generalizing the domain to require

$$f'(0) = \alpha_k f(0), \tag{41}$$

where the α_k are real numbers which generally differ from each other for different k. The usual radial equation case is recovered by requiring $\alpha_k \to \infty$, all k. There is no degeneracy variable a.

Let us define

$$\psi_k(r|p) = (2p)^{1/2} \langle r| H_k; E\rangle, \quad (p = E^{1/2}, E > 0), \tag{42}$$

$$\psi_{ki}(r) = \langle r| H_k; E_{ki}\rangle. \tag{43}$$

In the direct problem we require $\psi_k(r|p)$ and $\psi_{ki}(r)$ to satisfy the differential equations

$$\left[-\frac{d^2}{dr^2} + V_k(r)\right]\psi_k(r|p) = p^2\psi_k(r|p),$$

$$\left[-\frac{d^2}{dr^2} + V_k(r)\right] E_{ki}\psi_{ki}(r), \tag{44}$$

with the boundary conditions

$$\psi_k(0|p) = (2/\pi)^{1/2}\frac{p}{[p^2 + (\alpha_k)^2]^{1/2}}, \quad \psi_{ki}(0) = 1, \tag{45}$$

$$\frac{d}{dr}\psi_k(0|p) = \alpha_k\psi_k(0|p), \quad \frac{d}{dr}\psi_{ki}(0) = \alpha_k\psi_{ki}(0) = \alpha_k. \tag{45'}$$

The set of boundary conditions (45') assure us that the operators H_k operate on functions in its domain [see Eq. (41)].

The one can show that eigenfunctions satisfy the completeness relation [we use the fact that all of the eigenfunctions are real because of the boundary conditions (45) and (45')]

$$\int_0^\infty \psi_k(r|p)W_k(p)\psi_k(r'|p)\,dp + \sum_i \psi_{ki}(r)[C_{ki}]^{-1}\psi_{ki}(r') = \delta(r - r'). \tag{46}$$

In can be shown that the continuous part of the weight operator $\omega_{ck}(E)$ is given by

$$\omega_{ck}(E) = W_k(p). \tag{47}$$

In the direct problem, V_k and α_k are given and the weight $W_k(p)$ and normalizations C_{ki} are to be found. In the inverse problem it can be shown that if $W_k(p)$ and C_{ki} are given, V_k and α_k can be found.

In particular, let us consider the eigenfunctions of H_0. There are two cases: (a) $\alpha_0 > 0$, (b) $\alpha_0 < 0$.

In both cases the eigenfunctions of the continuous spectrum $\psi_0(x|p)$ are given by

$$\psi_0(x|p) = (2/\pi)^{1/2}\sin[pr + \gamma(p)], \tag{48}$$

where $\gamma(p)$ is given by

$$\sin[\gamma(p)] = \frac{p}{[p^2 + (\alpha_0)^2]^{1/2}},$$

$$\cos[\gamma(p)] = \frac{\alpha_0}{[p^2 + (\alpha_0)^2]^{1/2}}. \tag{48'}$$

If $\alpha_0 < 0$, then H_0 has a point eigenvalue and corresponding eigenfunction

$$E_{01} = -(\alpha_0)^2, \quad \psi_{01}(r) = \exp[\alpha_0 r]. \tag{49}$$

However, if $\alpha_0 > 0$, the continuous spectrum comprises the entire spectrum of H_0 and there is no point eigenvalue.

The weight function $W_0(p) = 1$ for all values of α_0. The normalization of the eigenfunction $\psi_{01}(r)$ for $\alpha_0 < 0$ is $C_{01} = (-2\alpha_0)^{-1}$. The completeness relation (46) for eigenfunctions of H_0 is now of the form Eq. (1).

The Gel'fand–Levitan equation (15) becomes in the present case

$$\langle r|K_{jk}|r'\rangle = 0 \quad \text{for } r' > r,$$

$$\langle r|K_{jk}|r'\rangle = -\langle r|\Omega_{jk}|r'\rangle - \int_0^r \langle r|K_{jk}|r''\rangle\,dr''\,\langle r''|\Omega_{jk}|r'\rangle \quad \text{for } r \geqslant r'. \tag{50}$$

On introducing the pseudo-eigenfunctions $\psi_{jki}(r)$ of H_k defined by

$$\psi_{jki}(r) = \langle r|H_k; E_{ji}\rangle, \quad \psi_{jki}(0) = 1, \quad \frac{d}{dr}\psi_{jki}(0) = \alpha_k \tag{51}$$

and on using Eqs. (42) and (43), $\langle r|\Omega_{jk}|r'\rangle$ is given by

$$\langle r|\Omega_{jk}|r'\rangle = \int_0^r \psi_k(r|p)\psi_k(r'|p)[W_j(p) - W_k(p)]\,dp + \sum_i \frac{\psi_{jki}(r)\psi_{jki}(r')}{C_{ji}} - \sum_i \frac{\psi_{ki}(r)\psi_{ki}(r')}{C_{ki}}. \tag{52}$$

Because $V_k(r) = 2(d.dr)\langle r|K_{k0}|r\rangle$, it follows that

$$V_{jk}(r) = 2\frac{d}{dr}\langle r|K_{jk}|r\rangle = V_j(r) - V_k(r). \tag{53}$$

As usual we find V_j from Eq. (53). To find α_j, we use

$$\alpha_j = \alpha_k + \langle 0|\Omega_{jk}|0\rangle. \tag{54}$$

Finally, Eqs. (16') become

$$\psi_j(r|p) = \psi_k(r|p) = \int_0^r \langle r|K_{jk}|r'\rangle\,dr'\psi_k(r'|p),$$

$$\psi_{ji}(r) = \psi_{jki}(r|p) = \int_0^r \langle r|K_{jk}|r'\rangle\,dr'\psi_{jki}(r'). \tag{55}$$

For the special case $\alpha_k \to \infty$, i.e., where the boundary condition (41) is replaced by $f(0) = 0$, our results go over into the earlier results of Refs. 9 and 10.

5. APPLICATION TO THE MARCHENKO EQUATION

The Marchenko equation is concerned with the problem of obtaining the scattering potential for a given angular momentum (zero, say) from the scattering phase and bound state information. In a certain sense it uses information complementary to that needed for the Gel'fand–Levitan equation for the radial equation of the last section for $\alpha_j \to \infty$. Specifically let us consider the Hamiltonian H_j.

$$H_j = -\frac{d^2}{dr^2} + V_j(r) \quad (0 \leqslant r < \infty) \tag{56}$$

acting on the subspace of Hilbert space of functions $f(r)$, which are twice differentiable and satisfy the condition

$$f(0) = 0. \tag{57}$$

The operators H_j have a continuous spectrum which ranges from zero to infinity. On calling the spectral variable p^2 with $p > 0$, we require the corresponding eigenfunctions $\psi_j(r|p)$ to satisfy

$$H_j \psi_j(r|p) = p^2 \psi_j(r|p) \tag{58}$$

with the boundary conditions

$$\psi_j(0|p) = 0, \quad \frac{d}{dr}\psi_j(0|p) = \left(\frac{2}{\pi}\right)^{1/2} p. \tag{59}$$

It is convenient to introduce the Jost wavefunctions $f_j(r|p)$ and $f_j^*(r|p)$, which satisfy the differential equations (58) with the boundary condition

$$\lim_{r\to\infty} f_j(r|p) = e^{-ipr}. \tag{60}$$

These solutions are, of course, the well-known Jost wavefunctions (see, e.g., Ref. 4, which can be consulted for much of the development of the present section).

The one can show

$$\psi_j(r|p) = (2/\pi)^{1/2}(1/2i)[f_j(p) f_j^*(r|p) - f_j^*(p) f_j(r|p)]. \tag{61}$$

In Eq. (61) $f_j(p)$ are the Jost functions defined by

$$f_j(p) = f_j(0|p). \tag{62}$$

The scattering operator $S_j(p)$ is defined by

$$f_j(p)/f_j^*(p) = S_j(p). \tag{63}$$

In addition to the continuous spectrum H_j may have a discrete spectrum which we shall take to be negative. The points of the discrete spectrum will be denoted by $E_{ji} = -\kappa_{ji}^2$ ($\kappa_{ji} > 0$). The corresponding eigenfunctions $\psi_{ji}(r)$ satisfy

$$H_j \psi_{ji}(r) = -\kappa_{ji}^2(r), \quad \psi_{ji}(0) = 0. \tag{64}$$

In addition to the boundary condition at $r = 0$ of Eq. (64), we must specify a condition at $r \to \infty$. It is simplest to choose

$$\lim_{r\to\infty} \psi_{ji}(r) = \exp(-\kappa_{ji} r). \tag{65}$$

Having fixed the boundary conditions, the normalizations C_{ji} are given by

$$C_{ji} = \int_0^\infty [\psi_{ji}(r)]^2 \, dr. \tag{66}$$

Furthermore, $\psi_j(r|p)$ and $\psi_{ji}(r)$ are real.

From the asymptotic condition (60) and that of Eq. (65) it is seen that

$$\psi_{ji}(r) = f_j(r| - i\kappa_{ji}). \tag{67}$$

[In Eq. (67) the i in front of κ_{ji} is the unit imaginary and *not* the counting index i which appears in the subscript.] The right-hand side of Eq. (67) is defined by analytic continuation in the complex p plane.

To introduce the Marchenko equation, we follow Ref. 14, which has a brief description of the standard method of derivation.

It can be shown that a triangular kernel $\langle r|K_j|r'\rangle$ eixsts such that

$$\langle r|K_j|r'\rangle = 0 \quad \text{for } r > r,$$

$$\langle r|K_j|r'\rangle = -\langle r|\Omega_j|r'\rangle - \int_r^\infty \langle r|K_j|r''\rangle \, dr'' \, \langle r''|\Omega_j r'\rangle, \quad \text{for } r \leqslant r', \tag{68}$$

where $\langle r|\Omega_j|r'\rangle$ is given by

$$\langle r|\Omega_j|r'\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[ip(r+r')][1 - S_j(p)] \, dp + \sum_i \frac{\exp[-\kappa_{ji}(r+r')]}{C_{ji}}. \tag{68a}$$

In Eq. (68a) the values of $S_j(p)$ appear for $p < 0$, though originally $S_j(p)$ was defined only for $p > 0$. To define $S_j(p)$ for negative values of p, we use analytic continuation:

$$S_j(-p) = S^*_{(p)}. \tag{68b}$$

We shall now show that the Marchenko equation (68) can be written in the form of Eq. (5). We shall first define an operator $K_{0,j}$ by defining the kernel $\langle r|K_{0,j}|r'\rangle$ when $K_{0,j}$ is written as an integral operator:

$$\langle r|K_{0,j}|r'\rangle = 0 \quad \textit{for } r > r',$$

$$\langle r|K_{0,j}|r'\rangle = \langle r|\Omega_j|r'\rangle + \int_{r'}^\infty \langle r|\Omega_j|r''\rangle \, dr'' \, \langle r''|K_j^*|r'\rangle \quad \text{for } r \leqslant r', \tag{69}$$

where $\langle r|K_j^* r'\rangle$ is the kernel of the operator K_j^*, i.e. (remembering that the kernels are real)

$$\langle r|K_j^*|r'\rangle = \langle r'|K_j|r\rangle. \tag{69'}$$

Now define U_j, $U_{0,j}$, W_j in a manner analogous to Eqs. (7) and (10). One obtains Eq. (5) after the manner of Ref. 1. Hence the Marchenko equation falls into the formalism of Refs. 1, 2, and the present paper.

The functions $f_j(r|p)$ are only formal eigenfunctions of H_j since they do not satisfy the boundary condition (57). Nevertheless, the formalism of the present paper goes through. [It should be noted that $f_j(r| - i\kappa_{ji})$ are eigenfunctions of H_j because of Eq. (67). A treatment of the problem in which the Jost wavefunctions are true eigenfunctions can be achieved by embedding the problem of the present section into a one-dimensional problem ($-\infty < x < +\infty$) with a suitable potential. Nevertheless, it is not necessary to do so for the purposes of the present paper.] Note that the "pseudo-eigenfunctions" corresponding to $\langle x|H_k; E_{ji}\rangle$ are $f_k(r|i\kappa_{ji})$, the analytic continuation of $f_k(r|p)$, and that they are real.

Then our Gel'fand–Levitan equation is

$$\langle r|K_{jk}|r'\rangle = 0 \quad \text{for } r > r',$$

$$\langle r|K_{jk}r'\rangle = -\langle r|\Omega_{jk}|r'\rangle \cdot \int_r^\infty \langle r|K_{jk}|r''\rangle\, dr''\, \langle r''|\Omega_{jk}|r'\rangle \quad \text{for} \quad r\leqslant r', \tag{70}$$

where

$$\langle r|\Omega_{jk}|r'\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} [S_k(p) - S_j(p)]\, f_k^*(r/p)\, f_k^*(r'|p)\, dp + \sum_i \frac{f_k(r|-i\kappa_{ji}) f_k(r'|-i\kappa_{ji})}{C_{ji}} + \sum_i \frac{f_k(r|-i\kappa_{ki}) f_k(r'|-i\kappa_{ki})}{C_{ki}}. \tag{71}$$

Furthermore, in defining $V_{jk}(r)$ by

$$V_{jk}(r) = -2\frac{d}{dr}\langle r|K_{jk}|r\rangle, \tag{72}$$

we have, as expected from the general theory

$$V_{jk}(r) = V_j(r) - V_k(r). \tag{73}$$

Finally the Jost wavefunctions and the eigenfunctions corresponding to point eigenvalues are

$$f_j(r|p) = f_k(r|p) + \int_r^\infty \langle r|K_{jk}|r'\rangle\, dr'\, f_k(r'|p),$$

$$f_j(r|-i\kappa_{ji}) = f_k(r|-i\kappa_{ji}) + \int_r^\infty \langle r|K_{jk}|r'\rangle dr'\, f_k(r'|-i\kappa_{ji}). \tag{74}$$

It should be mentioned that these results were obtained in Ref. 13 as the limit of a discretized version of the problem.

[1]I. Kay and H.E. Moses, Nuovo Cimento **2**, 917 (1955).
[2]I. Kay and H.E. Moses, Nuovo Cimento **3**, 66 (1956).
[3]I. Kay and H.E. Moses, Nuovo Cimento **3**, 276 (1956); **5**, 230 (1957).
[4]L.D. Faddeev, J. Math. Phys. **4**, 72 (1963) [transl. by B. Seckler].
[5]H.E. Moses, Stud. Appl. Math. **58**, 187 (1978).
[6]I. Kay and H.E. Moses, Nuovo Cimento **22**, 689 (1961); also Comm. Pure Appl. Math. **14**, 435 (1961).
[7]L.D. Faddeev, J. Sov. Math. **5**, 334 (1976).
[8]M. Kanal and H.E. Moses, J. Math. Phys. **18**, 2445 (1977).
[9]R. Jost and W. Kohn, Det. Kg. Dan. Vidensk. Selskab., Math. Fys. Medd. **27**, No. 9 (1953).
[10]N. Levinson, Phys. Rev. **89**, 755 (1953).
[11]H.E. Moses, J. Math. Phys. **18**, 2243 (1977).
[12]V.E. Zakharov and L.D. Faddeev, Functional Anal. Appl. **5**, 280 (1971).
[13]K.M. Case and S.C. Chiu, J. Math. Phys. **14**, 1643 (1973).
[14]V. de Alfaro and T. Regge, *Potential Scattering* (North-Holland, Amsterdam, 1965).

Noether's theorem, time-dependent invariants and nonlinear equations of motion

John R. Ray and James L. Reid
Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29631

(Received 17 April 1979; accepted for publication 18 May 1979)

Noether's theorem is applied to a Lagrangian for a system with nonlinear equations of motion. Noether's theorem leads to a time-dependent constant of the motion along with an auxiliary equation of motion. Special cases of this invariant have been used to quantize the time-dependent harmonic oscillator. We also discuss the solution of the original equations of motion in terms of the solutions to the auxiliary equation.

I. INTRODUCTION

The time-dependent harmonic oscillator equation

$$\ddot{x} + \omega^2(t)x = 0 \tag{1.1}$$

occurs in many physical problems. Lewis [1,2] proved that a conserved quantity for the harmonic oscillator is the time-dependent invariant

$$I = \tfrac{1}{2}[(x/\rho)^2 + (x\dot{\rho} - \dot{x}\rho)^2], \tag{1.2}$$

where $x(t)$ satisfies (1.1) and $\rho(t)$ satisfies the auxiliary equation

$$\ddot{\rho} + \omega^2(t)\rho = 1/\rho^3. \tag{1.3}$$

Lewis obtained this result by showing that an adiabatic invariant for (1.1) was, in fact, an exact invariant. As a point of interest, we mention that the invariant I was derived by Ermakov in 1880. [3] In his treatment, Ermakov assumed (1.1) and (1.3) and eliminated the frequency $\omega^2(t)$ between these two equations. After a few simple manipulations, he arrived at the invariant (1.2). Ermakov's derivation is, thus, not as general as that of Lewis, since Lewis obtained both (1.2) and (1.3) from (1.1). However, Ermakov's procedure admits immediate generalization.

Recently we discussed and generalized Ermakov's method of deriving the invariant (1.2). [4] Our generalization leads to the invariant

$$I = \tfrac{1}{2}[\phi((x/\rho) + \theta(\rho/x) + (x\dot{\rho} - \rho\dot{x})^2)], \tag{1.4}$$

for the coupled system of differential equations

$$\ddot{x} + \omega^2(t)x = g(\rho/x)/(x^2\rho), \tag{1.5}$$

$$\ddot{\rho} + \omega^2(t)\rho = f(x/\rho)/(\rho^2 x), \tag{1.6}$$

where $g(\rho/x)$ and $f(x/\rho)$ are arbitrary functions. Functions ϕ and θ are related to f and g, respectively, by the following equations:

$$\phi(x/\rho) = 2\int^{x/\rho} f(u)du, \tag{1.7}$$

$$\theta(\rho/x) = 2\int^{\rho/x} g(u)du. \tag{1.8}$$

The invariant (1.2) played a central role in the exact quantum treatment, by Lewis and Riesenfeld, [5] of the time-dependent harmonic oscillator. It is possible that more of the invariants (1.4) will prove useful in solving physical problems.

Lutzky, [6] in a recent paper derived the invariant (1.2) and the auxiliary equation (1.3) by a straightforward applications of Noether's theorem to the Lagrangian

$$L = \tfrac{1}{2}[\dot{x}^2 - \omega^2(t)x^2] \tag{1.9}$$

for the time-dependent harmonic oscillator. Noether's theorem yields an invariant for each symmetry transformation that leaves the action invariant. For the Lagrangian (1.9) Lutzky showed that Noether's theorem leads to the invariant (1.2) and the auxiliary equation (1.3). Thus, Lutzky also derived both (1.2) and (1.3) from (1.1), as did Lewis. However, the derivation by Lutzky is simpler and is also easier to generalize.

In this paper we apply Noether's theorem to the Lagrangian

$$L = \tfrac{1}{2}[\dot{\rho}^2 - \omega^2(t)\rho^2 + 2G(t)F(\rho)], \tag{1.10}$$

which is associated with the equation of motion

$$\ddot{\rho} + \omega^2(t)\rho = G(t)F'(\rho), \quad F'(\rho) = \frac{dF}{d\rho}, \tag{1.11}$$

where $G(t)$ and $F(\rho)$ are, initially, arbitrary functions of their arguments. We shall show that Noether's theorem and Lagrangian (1.10) lead to invariants which are special cases of our results (1.4), (1.5), and (1.6).

The invariant (1.2) can be interpreted as a link between the two differential equations (1.1) and (1.3). This link is most clearly illustrated by the theorem that if x_1 and x_2 are linearly independent solutions of (1.1), with Wronskian $W = x_1\dot{x}_2 - x_2\dot{x}_1$, then the general solution to (1.3) can be written

$$\rho = (Ax_1^2 + Bx_2^2 + 2Cx_1x_2)^{1/2}, \tag{1.12}$$

where A, B, and C are constants related by

$$AB - C^2 = 1/W^2. \tag{1.13}$$

In their discussions, both Lewis and Lutzky mention that this theorem can be proven by writing the invariant (1.2) for the two solutions x_1 and x_2, keeping ρ the same, and eliminating $\dot{\rho}$ between the two resulting invariants. The invariant (1.2) not only links $x(t)$ and $\rho(t)$ but in addition provides a path to solve the nonlinear equation (1.3) in terms of solutions to the linear equation (1.1). Our generalization (1.4), (1.5), and (1.6) shows that the invariant (1.4) links solutions to (1.5) and (1.6); however, since Eqs. (1.5) and (1.6) are

 0022-2488/79/102054-04$01.00

coupled we cannot use the technique of eliminating $\dot\rho$ between two invariants to solve (1.5) in terms of solutions of (1.6), or vice versa. Since use of the Lagrangian (1.10) with Lutzky's procedure yields special cases of (1.4), (1.5), and (1.6) the technique of eliminating $\dot\rho$ to generate a solution of either (1.5) or (1.6) also does not work. Nonetheless, we find, somewhat surprisingly, that Lagrangian (1.10) leads to certain nonlinear equations of motion in ρ which can be solved in terms of solutions to the related auxiliary equation. It is not clear whether such solutions are associated with the invariant linking these two differential equations. It is, however, still possible that this points to a deeper connection between being able to solve a nonlinear equation in terms of solutions to an auxiliary equation and the existence of the associated invariant.

In Sec. II we apply Noether's theorem to the Lagrangian (1.10) and derive the form of the invariant and the equations linked by the invariant. In Sec. III we discuss some properties of special cases and present some further examples of solving nonlinear equations in terms of solutions to linear equations. Finally, in Sec. IV we present our conclusions along with suggestions for further work.

II. NOETHER's THEOREM

We shall use the formulation of Noether's theorem given by Lutzky. [6] The symmetry transformation is described by the group operator

$$X = \xi(\rho,t)\frac{\partial}{\partial t} + \eta(\rho,t)\frac{\partial}{\partial\rho}. \tag{2.1}$$

If the symmetry transformation defined by (2.1) leaves the action A,

$$A = \int L(\rho,\dot\rho,t)\,dt,$$

invariant, then the combination of terms

$$\xi\frac{\partial L}{\partial t} + \eta\frac{\partial L}{\partial\rho} + (\dot\eta - \dot\rho\dot\xi)\frac{\partial L}{\partial\dot\rho} + \dot\xi L$$

is a total time derivative of a function $f(\rho,t)$, i.e.,

$$\xi\frac{\partial L}{\partial t} + \eta\frac{\partial L}{\partial\rho} + (\dot\eta - \dot\rho\dot\xi)\frac{\partial L}{\partial\dot\rho} + \dot\xi L = \dot f. \tag{2.2}$$

It follows from this that a constant of the motion for the system is

$$I = (\xi\dot\rho - \eta)\frac{\partial L}{\partial\dot\rho} - \xi L + f. \tag{2.3}$$

The Lagrangian to be investigated is

$$L = \tfrac12[\dot\rho^2 - \omega^2(t)\rho^2 + 2G(t)F(\rho)], \tag{2.4}$$

which yields the equations of motion

$$\ddot\rho + \omega^2(t)\rho = G(t)F'(\rho). \tag{2.5}$$

We assume $F(\rho)$ is not one of the following functions, all of which lead to trivial cases: const, const$\cdot\rho$ or const$\cdot\rho^2$. $F(\rho)$ can be any function except one of these three. Using the Lagrangian (2.4) in (2.2) and noting that this equation must hold for all values of ρ and $\dot\rho$ we can equate the coefficients of powers of $\dot\rho$. The $\dot\rho^3$ terms give

$$\frac{\partial\xi}{\partial\rho} = 0 \Rightarrow \xi = \xi(t). \tag{2.6}$$

The $\dot\rho^2$ terms yield

$$\eta = \tfrac12\dot\xi\rho + \psi(t), \tag{2.7}$$

where $\psi(t)$ is an arbitrary function of time. The $\dot\rho$ terms give

$$f = \tfrac14\ddot\xi\rho^2 + \dot\psi\rho + \chi(t), \tag{2.8}$$

where $\chi(t)$ is an arbitrary function of time. If we use (2.6), (2.7), and (2.8), the remaining terms in (2.2) are

$$(-\xi\omega\dot\omega - \dot\xi\omega^2 - \tfrac14\dddot\xi)\rho^2 - (\ddot\psi + \omega^2\psi)\rho + (\xi\dot G + \dot\xi G)F + \tfrac12 G\rho F'\dot\xi + GF'\psi - \dot\chi = 0. \tag{2.9}$$

Due to our assumptions concerning the form of $F(\rho)$ the coefficients of ρ^2 and ρ must vanish, requiring

$$\ddot\psi + \omega^2\psi = 0, \tag{2.10}$$

$$\dddot\xi + 4\xi\omega\dot\omega + 4\dot\xi\omega^2 = 0. \tag{2.11}$$

Multiplication of (2.11) by ξ and integration result in the first integral

$$\xi\ddot\xi - \tfrac12\dot\xi^2 + 2\xi^2\omega^2 = 2k = \text{const}, \tag{2.12}$$

which becomes

$$\ddot x + \omega^2(t)x = k/x^3 \tag{2.13}$$

through the transformation $\xi = x^2$. This is the time-dependent harmonic oscillator equation if we choose the arbitrary constant k to vanish. This is the equation linked to (2.5) by Noether's theorem. The remaining terms in (2.9) are

$$(\xi\dot G + \dot\xi G)F + \tfrac12 G\rho F'\dot\xi + GF'\psi - \dot\chi = 0. \tag{2.14}$$

We see that $\dot\chi = 0$, and, since a constant would only add a constant to f by (2.8), we can drop χ. We next suppose that F, $\rho F'$, and F' are all linearly independent. Then from (2.14) we obtain $\dot\xi = 0$. For $\dot\xi = 0$, Eq. (2.11) yields $\dot\omega = 0$. We exclude this case as we are interested in discussing the time-dependent harmonic oscillator. Thus, some of functions $F, \rho F', F'$ must be linearly dependent. The possibilities are

$$\text{(a)}\ \rho F' = -2m F, \tag{2.15}$$

where m is an arbitrary constant, and,

$$\text{(b)}\ F' = cF, \tag{2.16}$$

where c is an arbitrary constant. Case (b) is easily seen to lead $\dot\omega = 0$ and, hence, we have only case (a) to consider. Case (a) implies

$$F = F_0\rho^{-2m}, \quad F_0 = \text{const}. \tag{2.17}$$

Using this form for F in (2.14) we obtain the equations

$$\xi\dot G + \dot\xi G + \tfrac12 G(-2m)\dot\xi = 0, \tag{2.18}$$

$$\psi = 0. \tag{2.19}$$

The solution to (2.18) is

$$G = G_0\xi^{m-1} = G_0 x^{2m-2}, \quad G_0 = \text{const}. \tag{2.20}$$

Substituting the form of F (2.17) and the form of G (2.20) back into the equation of motion for ρ, we obtain the equation of motion

$$\ddot\rho + \omega^2(t)\rho = c\frac{x^{2m-2}}{\rho^{2m+1}} = \frac{c}{x\rho^2}\left(\frac{x}{\rho}\right)^{2m-1}, \tag{2.21}$$

where $c = -2mG_0 F_0$ is an arbitrary constant. Recall that $x(t)$ satisfies the following equation of motion

$$\ddot{x} + \omega^2(t)x = k/x^3. \tag{2.22}$$

The Noether theorem invariant is given by Eq. (2.3) and takes the form

$$I = \tfrac{1}{2}\left[\frac{c}{m}\left(\frac{x}{\rho}\right)^{2m} + k\left(\frac{\rho}{x}\right)^2 + (\rho\dot{x} - x\dot{\rho})^2\right]. \tag{2.23}$$

From our previous assumptions on $F(\rho)$, m cannot be equal to 0, $-1/2$, or -1. It is clear that these results are a special case of (1.4), (1.5), and (1.6) with $g = k\rho/x$, $f = c(x/\rho)^{2m-1}$.

Equations (2.21), (2.22), and (2.23) represent our results of applying Noether's theorem to the Lagrangian (1.10). Notice that, whereas we started from the Lagrangian with $F(\rho)$ and $G(t)$ arbitrary functions, in order for L to admit a symmetry transformation (2.1), F and G are restricted to the forms (2.17) and (2.20) respectively. The two cases treated by Lutzky can be recovered from our general results by (1) $G = 0$ and (2) $m = 1$. Note that in the case $m = 1$ Eq. (2.21) is independent of x. This is important since, in general, uncoupled equations are much easier to handle.

Further reflection on the form of (2.14) shows that one can generalize the equation of motion to

$$\ddot{\rho} + \omega^2(t)\rho = \frac{1}{x\rho^2}\sum_i c_i\left(\frac{x}{\rho}\right)^{2m_i - 1} = \sum_i G_i(t)F_i'(\rho), \tag{2.24}$$

where $i = 1,2,\cdots$ and c_i and m_i are arbitrary constants, $m_i \neq 0, -1/2, -1$. The Lagrangian changes to the form

$$L = \tfrac{1}{2}(\dot{\rho}^2 - \omega^2(t)\rho^2 + 2\textstyle\sum G_iF_i) \tag{2.25}$$

with a corresponding change in the invariant I. Equation (2.24) is the general equation of motion linked to the time-dependent harmonic oscillator equation through the Noether invariant associated with the Lagrangian (2.25). This generalizes Lutzky's results and shows that the Noether theorem invariants for this Lagrangian are special cases of our generalization of Ermakov's results (1.4).

III. EXAMPLE

Lewis and Lutzky indicated how solutions of the nonlinear equation (1.3) could be obtained in terms of solutions to the time-dependent harmonic oscillator equation (1.1) by using the invariant (1.2). The form of this solution was given by (1.12). The method of eliminating $\dot{\rho}$ between two invariants will generally not work for equations of the form (2.22) and (2.24) because of the coupling of x in the ρ equation. Nevertheless, we present an example that shows we can obtain solutions of a special case of (2.24) in terms of solutions to (2.22). The solution has the same general form as (1.12).

Consider the following special case of (2.24),

$$\ddot{\rho} + \omega^2\rho = \frac{c_1}{x\rho^2}\left(\frac{x}{\rho}\right)^{2m_1 - 1} + \frac{c_2}{x\rho^2}\left(\frac{x}{\rho}\right)^{2m_2 - 1}. \tag{3.1}$$

For $m_1 = (m-2)/2$, $m_2 = (2m-2)/2$, we obtain

$$\ddot{\rho} + \omega^2\rho = c_1 x^{m-4}\rho^{1-m} + c_2 x^{2m-4}\rho^{1-2m}. \tag{3.2}$$

A special solution to the auxiliary equation (2.22) is

$$x = (-4k/W^2)^{1/4}(uv)^{1/2}, \tag{3.3}$$

where u and v are linearly independent solutions of the time-dependent harmonic oscillator equation (1.1) with Wronskian W. Equation (3.3) is of the form (1.12) with A and B both zero. We define the constants c_1 and c_2 as

$$c_1 = (b/4)(m-2)\,W^2\left(\frac{-4k}{W^2}\right)^{-(m-4)/4}, \tag{3.4}$$

$$c_2 = (m-1)(ac - b^2/4)\,W^2\left(\frac{-4k}{W^2}\right)^{-(m-2)/2}, \tag{3.5}$$

where a, b, and c are new arbitrary constants. Using the solution (3.3) and the constants (3.4) and (3.5), we have for the ρ equation of motion

$$\ddot{\rho} + \omega^2(t)\rho = (b/4)(m-2)\,W^2(uv)^{(m-2)/2}\rho^{1-m} + (m-1)(ac - b^2/4)(uv)^{m-2}W^2\rho^{1-2m}. \tag{3.6}$$

It can be proven by direct substitution that a solution to (3.6) is

$$\rho = [au^m + b(uv)^{m/2} + cv^m]^{1/m}. \tag{3.7}$$

This solution was first found by Reid[7] in his investigation of homogeneous solutions to nonlinear equations. Although we see no way to obtain this solution by using the invariant linking equations (2.22) and (3.6), the similarity between the solution (1.12) and (3.7) is obvious. That Eq. (3.6) can be derived directly from Eqs. (2.22) and (2.24), linked by the Noether invariant, suggests a deeper connection among these results. As a final point we note that, whereas (1.12) is the general solution to (1.3), (3.7) is only a particular solution to (3.6).

IV. CONCLUSIONS

Noether's theorem applied to the Lagrangian

$$L = \tfrac{1}{2}\left[\dot{\rho}^2 - \omega^2(t)\rho^2 + 2\sum_i G_i(t)F_i(\rho)\right], \tag{4.1}$$

leads to the equations of motion

$$\ddot{\rho} + \omega^2(t)\rho = \frac{1}{x\rho^2}\sum_i c_i\left(\frac{x}{\rho}\right)^{2m_i - 1} = \sum G_i(t)F_i'(\rho), \quad m_i \neq 0, -\tfrac{1}{2}, -1, \tag{4.2}$$

where x satisfies the auxiliary equation

$$\ddot{x} + \omega^2(t)x = k/x^3. \tag{4.3}$$

The Noether invariant linking these equations is

$$I = \frac{1}{2}\left[\sum_i \frac{c_i}{m_i}\left(\frac{x}{\rho}\right)^{2m_i} + k\left(\frac{\rho}{x}\right)^2 + (\rho\dot{x} - x\dot{\rho})^2\right]. \tag{4.4}$$

For $k = 0$ the invariant I links an infinite number of Eqs. (4.2) to the time-dependent harmonic oscillator equation of motion. These results are a special case of our general results derived in Ref. 4 using Ermakov's technique. In Ermakov's technique one does not make use of Noether's theorem. The results in this paper furnish a further example of how Noether's theorem applied to a Lagrangian leads to an invariant and an auxiliary equation. Lutzky[6] considered the case

$G(t) = 0$ and found the results which were earlier derived and used by Lewis.[1,2] In this case the time-dependent invariant linking the equations was used to solve physical problems and to obtain the solution of a nonlinear equation in terms of solutions to a linear equation. It is possible some of the Noether invariants (4.4) could be used to solve physical problems. We have shown in Sec. III how solutions of a special case of (4.2) can be written in terms of solutions of the auxiliary equation (4.3).

It seems to us that many important properties of nonlinear equations of motion are associated with Noether identities as derived in Ref. 6 and in this paper. The working out of more explicit examples would help clarify this point.

[1]H.R. Lewis, Jr., Phys. Rev. Lett. **18**, 510 (1967).
[2]H.R. Lewis, Jr. J. Math. Phys. **9**, 1976 (1968).
[3]V.P. Ermakov, Univ. Izv. Kiev **20**, No. 9, 1 (1880).
[4]J.R. Ray and J.L. Reid, Phys. Lett. A **71**, 317 (1979).
[5]H.R. Lewis, Jr. and W.B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969).
[6]M. Lutzky, Phys. Lett. A **68**, 3 (1978).
[7]J.L. Reid, Proc. Am. Math. Soc. **38**, 532 (1973); P.B. Burt and J.L. Reid, J. Math. Anal. Appl. **55**, 43 (1976).

Gauge theory in Hamiltonian classical mechanics: The electromagnetic and gravitational fields

B. Boisseau and C. Barrabès
Départment de Physique, Faculté des Sciences de Tours, Parc de Grandmont, Tours 37, France

(Received 6 March 1979; accepted for publication 18 May 1979)

Gauge potentials are directly defined from Hamiltonian classical mechanics. Gauge transformations belong to canonical transformations and are determined by a first order development of generating functions. The electromagnetic and gravitational fields, and they only, are obtained.

I. INTRODUCTION

We are indebted to Weyl[1] for the prototype of gauge theories. He introduced the electromagnetic interaction as a gauge field associated with the group of phase transformations in charged fields. On the same model, the Yang and Mills[2] field is associated with the rotations in isospace. Then the concept is drawn from internal space and carried into space–time. So the gravitational field is related to the Lorentz group by Utiyama[3] and to the Poincaré group by Kibble.[4] These early works provide a good insight into the process for determining the interaction fields.

Let us consider a free field whose equation of evolution is covariant under a group of global transformations (the parameters of group elements do not depend on space–time). This equation will be no longer covariant if we make these transformations local (the parameters of group elements depend on space–time). The covariance will be restored by introducing gauge connections (or gauge potentials) according to the minimal coupling principle. The variance of these potentials is determined by requiring that the equation of evolution should be covariant.

Thus, gauge theories appear to be a powerful tool of research in interacting models of matter; Weinberg's[5] and Salam's[6] works give a proof of this. Abers and Lee[7] gave a whole review on the subject.

The purpose of the present paper is to build up a gauge theory in Hamiltonian classical mechanics. The intended transformations will obviously belong to canonical ones in order to preserve the frame of the Hamiltonian formalism. Gauge fields will be generated according to the very similar principle expressed above.

Formal invariance of the Hamiltonian, under some canonical transformations, implies the introduction of gauge potentials, by a "minimal coupling principle." The variance of these potentials is determined by requiring that the Hamiltonian should be form invariant. One difference, however, is that gauge transformations will not derive, contrarily to usual theories, from a group of global transformations.

In Sec. II we point out, as an example, that it is possible to introduce the electromagnetic field in Hamiltonian mechanics for a particular canonical transformation. The existing analogy with the local phase transformation in quantum mechanics is used as an introduction to the gauge theory in Hamiltonian classical mechanics that we develop in the two next sections. Section III determines the canonical transformations that we may consider as gauge transformations. These are defined by their generating functions. In Sec. IV, after a transposition in a relativistic picture of Hamiltonian mechanics and a physical justification of approximations considered in Sec. III, the electromagnetic and gravitational fields are obtained only as gauge fields.

II. GAUGE THEORY OF THE ELECTROMAGNETIC FIELD IN QUANTUM MECHANICS AND IN HAMILTONIAN CLASSICAL MECHANICS

By taking the electromagnetic field as an example and by analogy with quantum mechanics, this section points out how it is possible to introduce a gauge field in Hamiltonian mechanics. The unusual formulation of the electromagnetic gauge theory in quantum mechanics (Sec. II.A), allows an easy transposition in Hamiltonian classical mechanics. (Sec. II.B).

A. Point of view of quantum mechanics ($\hbar = c = 1$)

Let us consider the Schrödinger equation of a free particle

$$i\,\partial_t\psi = H_0\psi, \tag{2.1}$$

where

$$H_0 = \frac{1}{2m}\sum_j(-i\,\partial_j)^2. \tag{2.2}$$

This equation is invariant under the global phase transformation

$$\psi' = \exp(ie\alpha)\psi.$$

However, if we work out the local phase transformation

$$\psi'(x,t) = \exp[ie\alpha(x,t)]\psi(x,t), \tag{2.3}$$

Eq. (2.1) is transformed in the equivalent equation

$$i\,\partial_t\psi' = H'_0\psi', \tag{2.4}$$

where

$$H'_0 = \frac{1}{2m}\sum_j(-i\,\partial_j - e\partial_j\alpha)^2 - e\partial_t\alpha. \tag{2.5}$$

Comparison between Eqs. (2.2) and (2.5) shows obviously the noninvariance of the Hamiltonian under the local gauge transformation (2.3). In order to require this invariance, we are induced to substitute in H'_0, for the derivatives $\partial_j\alpha$ and $\partial_t\alpha$, a four-component field $A_\mu(x,t)$ according to

 0022-2488/79/102058-05$01.00

the correspondence

$$\partial_i \alpha \to A_i, \quad \partial_t \alpha \to A_0. \tag{2.6}$$

In this way one obtains, by the minimal coupling principle, the Hamiltonian of the particle in the presence of the gauge potentials A_μ:

$$H = \frac{1}{2m}\sum_j (-i\partial_j - eA_j)^2 - eA_0 \tag{2.7}$$

and the associated Schrödinger equation

$$i\partial_t \psi = H\psi. \tag{2.8}$$

Then, let us perform again the gauge transformation (2.3): Eq. (2.8) is transformed into the equivalent equation

$$i\partial_t \psi' = H'\psi', \tag{2.9}$$

where

$$H' = \frac{1}{2m}\sum_j (-i\partial_j - e\partial_j \alpha - eA_j)^2 - e\partial_t \alpha - eA_0. \tag{2.10}$$

Let us define

$$A'_j = A_j + \partial_j \alpha, \quad A'_0 = A_0 + \partial_t \alpha, \tag{2.11}$$

Then, the new Hamiltonian can be written as follows:

$$H' = \frac{1}{2m}\sum_j (-i\partial_j - eA'_j)^2 - eA'_0. \tag{2.12}$$

Therefore, the formal invariance of the Hamiltonian is obtained on condition that the fields A_μ should be jointly transformed according to Eq. (2.11). Now Eq. (2.11) is the gauge transformation of the electromagnetic potentials; moreover, the coupling in the Hamiltonian (2.7) between the particle and the gauge potentials is just the usual one of a charged particle and the electromagnetic field. We can interpret this result in the following manner[8]: the formal invariance of the Hamiltonian under a local phase transformation implies the existence of the electromagnetic field. In this example we see the essential part of gauge theories emerge: the requirement of a symmetry (for reasons sometimes abstruse) makes the introduction of an external field necessary.

B. Point of view of the Hamiltonian classical mechanics

Let us consider a free classical particle. The equations of motion are

$$\dot{p}_i = -\frac{\partial H_0}{\partial x^i}, \quad \dot{x}^i = \frac{\partial H_0}{\partial p_i}, \tag{2.13}$$

where

$$H_0 = \frac{1}{2m}\sum_i p_i^2. \tag{2.14}$$

Let us work out the canonical transformation

$$x'^i = x'^i(x^j, p_j, t), \quad p'_i = p'_i(x^j, p_j, t), \tag{2.15}$$

generated by

$$F_3(p, x', t) = -e\alpha(x', t) - x'^j p_j. \tag{2.16}$$

The transformation (2.15) and the Hamiltonian H'_0 of the new system are given by

$$p'_i = -\frac{\partial F_3}{\partial x'^i} = e\frac{\partial \alpha}{\partial x'^i} + p_i, \tag{2.17}$$

$$x^i = -\frac{\partial F_3}{\partial p_i} = x'^i, \tag{2.18}$$

$$H'_0 - H_0 = \frac{\partial F_3}{\partial t} = -e\frac{\partial \alpha}{\partial t}. \tag{2.19}$$

This provides us with the Hamiltonian of the new system

$$H'_0 = \frac{1}{2m}\sum_i \left(p'_i - e\frac{\partial \alpha}{\partial x'^i}\right)^2 - e\frac{\partial \alpha}{\partial t}, \tag{2.20}$$

to which one associates the equations equivalent to Eq. (2.13):

$$\dot{p}'_i = -\frac{\partial H'_0}{\partial x'^i}, \quad \dot{x}'^i = \frac{\partial H'_0}{\partial p'_i}. \tag{2.21}$$

Comparison between Eqs. (2.14) and (2.20) shows obviously the noninvariance of the Hamiltonian under the canonical transformation generated by Eq. (2.16). Let us consider the latter as a gauge transformation: We are induced, as it was done at the beginning of this section, to substitute in H'_0 a four-component field $A'_\mu(x', t)$ for the derivatives $\partial\alpha/\partial x'^i$ and $\partial\alpha/\partial t$.

In this way one obtains, by the minimal coupling principle, the Hamiltonian of the particle in the presence of the gauge potentials A'_μ:

$$H' = \frac{1}{2m}\sum_i (p'_i - eA'_i)^2 - eA'_0. \tag{2.22}$$

In order not to multiply notations we suppress the "primed" variables and simply write the new Hamiltonian (2.22) in the following manner:

$$H = \frac{1}{2m}\sum_i (p_i - eA_i)^2 - eA_0. \tag{2.23}$$

The canonical transformation generated by Eq. (2.16) yields

$$H' = \frac{1}{2m}\sum_i \left(p'_i - e\frac{\partial \alpha}{\partial x'^i} - eA_i\right)^2 - e\frac{\partial \alpha}{\partial t} - eA_0;$$

then, using again Eq. (2.11),

$$H' = \frac{1}{2m}\sum_i (p'_i - eA'_i)^2 - eA'_0.$$

Hence, we obtain the formal invariance of the Hamiltonian (2.23), on the condition that the fields A_μ should be jointly transformed according to Eq. (2.11).

Henceforth, we can have similar conclusions to the ones drawn in Sec. IIA. The Hamiltonian (2.23) describes a charged classical particle in the presence of the electromagnetic field.

Thus, the Hamiltonian theory (Sec. IIB) is perfectly similar to the quantum theory (Sec. IIA). One simply substitutes a canonical transformation for a phase transformation. This will not surprise us, since Van Hove[9] showed the close links between canonical and unitary transformations. However, in the two next sections, we shall completely ignore the quantum point of view and construct a gauge theory entirely based on Hamiltonian classical mechanics.

III. CANONICAL AND GAUGE TRANSFORMATIONS

The canonical transformations are changes of phase space coordinates

$$q'^i = q'^i(q^j, p_j), \quad p'_i = p'_i(q^j, p_j), \tag{3.1}$$

so that the 1 form $pdq - p'dq'$ is an exact form

$$p_i dq^i - p'_i dq'^i = dS(p,q). \tag{3.2}$$

These transformations preserve the canonical form of Hamilton's equations.

In a $2n$ dimensional phase space, Arnold[10] shows that all the canonical transformations can be generated by a minimal system $\mathscr{S}_1$ of 2^n types of generating functions $F(q, p'_i, q'^j)$ in which

$$q = (q^1 \ldots q^n), \quad p'_i = (p'_{i_1} \ldots p'_{i_k}),$$

$$q'^j = (q'^{j_1} \ldots q'^{j_{n-k}}). \tag{3.3}$$

Here, $(i_1 \ldots i_k)(j_1 \ldots j_{n-k})$ designates one partition of the set $(1 \ldots n)$ in two disconnected subsets. The number of these partitions is 2^n.

The generating functions $F(q, p'_i, q'^j)$ have to satisfy

$$\det\left(\frac{\partial^2 F}{\partial q\, \partial r'}\right) \neq 0, \quad (r' = p'_i, q'^j). \tag{3.4}$$

Except for this restriction, the generating functions are arbitrary. The functions $F_1(q,q')$ and $F_2(q, p')$ (Leech's[11] or Goldstein's[12] notation) belong to this minimal system $\mathscr{S}_1$. The identity is generated by a function of type F_2:

$$F_2 = q^i p'_i. \tag{3.5}$$

However, no generating function of other type belonging to the minimal system $\mathscr{S}_1$ can generate the identity. Indeed, the independence of the variables involved in the other types of generating functions belonging to $\mathscr{S}_1$ is inconsistent with the identity. Then we deduce that all the canonical transformations continuously connected with the identity admit a generating function of the type $F_2(q, p')$.

Up till now, we have only used this minimal system $\mathscr{S}_1$ associated with $(q^1 \ldots q^n)$ but it is obvious that we can choose among 2^n minimal systems $\mathscr{S}_S$ $(s = 1 \ldots 2^n)$ associated with the 2^n systems of variables:

$$p_l = (p_{l_1} \ldots p_{l_h}), \quad q^m = (q^{m_1} \ldots q^{m_{n-h}}), \tag{3.6}$$

where $(l_1 \ldots l_h)(m_1 \ldots m_{n-h})$, as previously, designates one partition of $(1 \ldots n)$ in two disconnected subsets. In this way we obtain $2^n \cdot 2^n = 4^n$ types of generating functions $F(p_l, q^m, p'_i, q'^j)$ constituting a nonminimal system.

It will be convenient, for reasons that will appear in the next section, to use the minimal system $\mathscr{S}_{2^n}$ composed of generating functions of the type $F(p, p'_i, q'^j)$. The function $F_3(p,q')$ and $F_4(p,p')$ (Leech's or Goldstein's notation) belong to the minimal system $\mathscr{S}_{2^n}$. The identity is generated in this system only by a function of the type F_3:

$$F_3 = -p_i q'^i. \tag{3.7}$$

As previously, we deduce that all the canonical transformations continuously connected to the identity can be generated by functions of the type $F_3(p,q')$.

In the usual gauge theories, transformations are continuously connected to the identity. This leads us to consider as gauge transformations only the canonical transformations continuously connected to the identity. Let us emphasize that they will be all generated by functions of the type $F_3(p,q')$. Let us expand the functions $F_3(p,q')$ in power of p:

$$F_3(p,q') = F_3(0,q') + \frac{\partial F_3}{\partial p_i}(0,q')\, p_i + \cdots. \tag{3.8}$$

Suppose, for some reason, that we may neglect the terms of second order; the equation (3.8) can be written

$$F_3(p,q') = -\alpha(q') - g^j(q')\, p_j. \tag{3.9}$$

The functions $\alpha(q')$ and $g^j(q')$ may be chosen arbitrarily except for the restrictions analogous to Eq. (3.4) assuring that Eq. (3.9) generates a change of coordinates in phase space. Indeed, the equation (3.9) gives

$$F_3(0,q') = -\alpha(q'), \quad \frac{\partial F_3}{\partial p_i}(0,q') = -g^i(q'). \tag{3.10}$$

Let us remark that the first order is the lowest one from which the development may be stopped; the zero order cannot follow the conditions analogous to Eq. (3.4). Now let us put together the results obtained in this section; we shall get the following conclusion: To the first order approximation in momentum, all the canonical transformations continuously connected with the identity can be generated by Eq. (3.9).

In the following section we are to justify physically the rightness of the first order approximation. The transformations continuously connected with the identity generated by Eq. (3.9) will be considered as gauge transformations.

IV. GAUGE THEORY IN HAMILTONIAN CLASSICAL MECHANICS: THE ELECTROMAGNETIC AND GRAVITATIONAL FIELDS

Let us consider a Minkowski space (x^α) to which one associates an eight dimension phase space (x^α, p_α). The equations of motion of a free particle are Hamilton's equations

$$\frac{d}{d\tau} p_\alpha = -\frac{\partial H_0}{\partial x^\alpha}, \quad \frac{d}{d\tau} x^\alpha = \frac{\partial H_0}{\partial p_\alpha}, \tag{4.1}$$

where the Hamiltonian[13]

$$H_0 = (1/2m)\, p_\alpha p_\beta\, \eta^{\alpha\beta}, \quad (\eta^{\alpha\beta} = -1,1,1,1) \tag{4.2}$$

is a scalar quantity and τ designates the proper time

$$d\tau = (-dx^\alpha dx^\beta \eta_{\alpha\beta})^{1/2}. \tag{4.3}$$

In order to consider cannonical transformations in that eight dimension space, we suppose τ is an evolution parameter external to the space time. We shall give it back its meaning of proper time at the last moment to solve the equations. So the relation between the external parameter τ and space–time is similar to the relation between Newtonian time and space. Therefore, the theory of canonical transformations is identical with the usual theory. It is even simplified, since the parameter τ will never occur in the transformations.

Let us consider the canonical transformations in this eight dimension phase space

$$x'^{\mu} = x'^{\mu}(x^{\alpha}, p_{\alpha}),$$
$$p'^{\mu} = p'^{\mu}(x^{\alpha}, p_{\alpha}). \tag{4.4}$$

In the new system, the equations of motion become

$$\frac{d}{d\tau} p'_{\mu} = -\frac{\partial H'_0}{\partial x^{\mu}}, \quad \frac{d}{d\tau} x'^{\mu} = \frac{\partial H'_0}{\partial p'_{\mu}}, \tag{4.5}$$

where

$$H'_0(x'^{\mu}, p'_{\mu}) = H_0(x^{\alpha}, p_{\alpha}), \tag{4.6}$$

since τ does not occur in Eq. (4.4).

Let us choose now a system of units in which the mass unit is very large (the mass M of the universe, if it exists), the unit of velocity being naturally c, the velocity of light in empty space. Conjugate momenta, which in the case of the free particle become identical with linear momenta and energy, turn out to be very small quantities in usual physical experiments. Then one may consider the first order approximation in p, exposed in Sec. III. All the canonical transformations, continuously connected with the identity, can be generated, to this approximation, by

$$F_3(p_{\alpha}, x'^{\mu}) = -\alpha(x'^{\mu}) - g^{\alpha}(x'^{\mu}) p_{\alpha}. \tag{4.7}$$

The canonical transformations (4.4) generated by Eq. (4.7) are given by

$$p'_{\mu} = -\frac{\partial F_3}{\partial x'^{\mu}} = \frac{\partial \alpha}{\partial x'^{\mu}} + \frac{\partial g^{\alpha}}{\partial x'^{\mu}} p_{\alpha},$$
$$x^{\alpha} = -\frac{\partial F_3}{\partial p_{\alpha}} = g^{\alpha}(x'^{\mu}). \tag{4.8}$$

Let us put

$$x^{\alpha} = g^{\alpha}(x'^{\mu}) = x^{\alpha}(x'^{\mu}), \tag{4.9}$$

the equations (4.8) and their inverses are written

$$p'_{\mu} = \frac{\partial \alpha}{\partial x'^{\mu}} + \frac{\partial x^{\alpha}}{\partial x'^{\mu}} p_{\alpha}, \quad x^{\alpha} = x^{\alpha}(x'^{\mu}),$$
$$p_{\alpha} = \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \left(p'_{\mu} - \frac{\partial \alpha}{\partial x'^{\mu}} \right), \quad x'^{\mu} = x'^{\mu}(x^{\alpha}), \tag{4.10}$$

respectively. By using Eqs. (4.6) and (4.10), we may now write the Hamiltonian of the new system as follows:

$$H'_0 = \frac{1}{2m} \left(p'_{\mu} - \frac{\partial \alpha}{\partial x'^{\mu}} \right) \left(p'_{\nu} - \frac{\partial \alpha}{\partial x'^{\nu}} \right) \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \times \frac{\partial x'^{\nu}}{\partial x^{\beta}} \eta^{\alpha\beta}. \tag{4.11}$$

In this expression the derivatives $\partial x'^{\mu}/\partial x^{\alpha}$ must naturally be functions of x'.

By comparing Eqs. (4.2) and (4.11), it is obvious that the Hamiltonian is not form invariant under the transformations generated by Eq. (4.7). If one considers these transformations as gauge transformations, this invariance will be obtained by a minimal coupling principle: In the Hamiltonian H'_0, one substitutes the fields $G'_{\nu}(x')$ for the derivatives $\partial\varphi(x')/\partial x'^{\nu}$ coming from the gauge transformation. These are gauge potentials whose variance will enable us to obtain the form invariance of the Hamiltonian. The latter will be named the substitution Hamiltonian.

In the Hamiltonian (4.11) it is therefore possible to perform the substitution

$$\frac{\partial \alpha(x')}{\partial x'^{\mu}} \to B'_{\mu}(x'). \tag{4.12}$$

On the other hand, the derivatives $\partial x'^{\mu}/\partial x^{\alpha}$ cannot be replaced immediately. Their definitions are not consistent with those we gave in the minimal coupling principle, but algebraic expressions of these derivatives are.

As matter of fact we have

$$\frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\alpha}}{\partial x'^{\nu}} = \delta^{\mu}_{\nu}. \tag{4.13}$$

It is therefore possible to introduce $V'^{\alpha}{}_{\nu}(x')$ by the substitution

$$\frac{\partial x^{\alpha}}{\partial x'^{\nu}} \to V'^{\alpha}{}_{\nu}(x'). \tag{4.14}$$

One can define $V'_{\alpha}{}^{\mu}(x')$ such as

$$V'_{\alpha}{}^{\mu} V'^{\alpha}{}_{\nu} = \delta^{\mu}_{\nu}. \tag{4.15}$$

The relations (4.13), (4.14), and (4.15) then give the substitution

$$\frac{\partial x'^{\mu}}{\partial x^{\alpha}} \to V'_{\alpha}{}^{\mu}(x'). \tag{4.16}$$

We may write now the substitution Hamiltonian, that is to say, after suppressing primed variables:

$$H = (1/2m)(p_{\mu} - B_{\mu})(p_{\nu} - B_{\nu}) V_{\alpha}{}^{\mu} V_{\beta}{}^{\nu} \eta^{\alpha\beta}. \tag{4.17}$$

Let us examine the variance of this Hamiltonian under a canonical transformation generated by a generating function (4.7). By using again the equations (4.10), we obtain

$$H' = \left(\frac{1}{2m}\right) \left[\frac{\partial x'^{\lambda}}{\partial x^{\mu}} \left(p'_{\lambda} - \frac{\partial \alpha}{\partial x'^{\lambda}} \right) - B_{\mu} \right] \times \left[\frac{\partial x'^{\rho}}{\partial x^{\nu}} \left(p'_{\rho} - \frac{\partial \alpha}{\partial x'^{\rho}} \right) - B_{\nu} \right] V_{\alpha}{}^{\mu} V_{\beta}{}^{\nu} \eta^{\alpha\beta} \tag{4.18}$$

or equivalently

$$H' = \left(\frac{1}{2m}\right) \left[p'_{\lambda} - \frac{\partial \alpha}{\partial x'^{\lambda}} - B_{\epsilon} \frac{\partial x^{\epsilon}}{\partial x'^{\lambda}} \right] \times \left[p'_{\rho} - \frac{\partial \alpha}{\partial x'^{\rho}} - B_{\kappa} \frac{\partial x^{\kappa}}{\partial x'^{\rho}} \right] \times \cdots \times \frac{\partial x'^{\lambda}}{\partial x^{\mu}} V_{\alpha}{}^{\mu} \frac{\partial x'^{\rho}}{\partial x^{\nu}} V_{\beta}{}^{\nu} \eta^{\alpha\beta}. \tag{4.19}$$

If we define

$$B'_{\lambda}(x') = \frac{\partial x^{\epsilon}}{\partial x'^{\lambda}} B_{\epsilon} + \frac{\partial \alpha}{\partial x'^{\lambda}}, \tag{4.20}$$

$$V'_{\alpha}{}^{\lambda}(x') = \frac{\partial x'^{\lambda}}{\partial x^{\mu}} V_{\alpha}{}^{\lambda}, \tag{4.21}$$

then the Hamiltonian is written

$$H' = (1/2m)(p'_{\lambda} - B'_{\lambda})(p'_{\rho} - B'_{\rho}) V'_{\alpha}{}^{\lambda} V'_{\beta}{}^{\rho} \eta^{\alpha\beta}. \tag{4.22}$$

The comparison of Eq. (4.22) with (4.17) shows clearly the formal invariance of the substitution Hamiltonian on the condition that the gauge potentials should be transformed jointly according to Eqs. (4.20) and (4.21).

Let us interpret now the gauge potentials. The equation (4.20) shows that the gauge B_μ is transformed, except for a gradient, as a covariant vector under a general transformation of space–time, which agrees exactly with the properties of the electromagnetic potentials. The equations (4.21) show that the $V_\alpha{}^\mu$ constitute four vector fields which can easily be interpreted. In Minkowski space (our starting point), the element of length is

$$ds^2 = -\eta_{\alpha\beta}\, dx^\alpha\, dx^\beta. \tag{4.23}$$

The transformations generated by Eq. (4.7), changing a system of rectilinear coordinates (x^α) into a curvilinear system (x'^μ), allow us to write

$$dx^\alpha = \frac{\partial x^\alpha}{\partial x'^\mu}\, dx'^\mu. \tag{4.24}$$

Then the element of length (4.23) is written

$$\begin{aligned} ds^2 &= -\eta_{\alpha\beta}\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\, dx'^\mu\, dx'^\nu \\ &= -g'_{0\mu\nu}\, dx'^\mu\, dx'^\nu. \end{aligned} \tag{4.25}$$

where $g'_{0\mu\nu}$ designates Minkowski's metric expressed in curvilinear coordinates. The substitutions (4.14) and (4.16) translate the metric $g'_{0\mu\nu}$ expressed in Eq. (4.25) into

$$g'_{\mu\nu} = \eta_{\alpha\beta} V'^\alpha{}_\mu\, V'^\beta{}_\nu\,. \tag{4.26}$$

This allows us to interpret $g_{\mu\nu}$ as a Riemannian metric associated with tetrad[14] fields $V^\alpha{}_\mu$ or ($V_\alpha{}^\mu$). The substitution Hamiltonian is given the form

$$H = (1/2m)(p_\mu - B_\mu)(p_\nu - B_\nu)g^{\mu\nu}. \tag{4.27}$$

If we define A_μ as

$$B_\mu = eA_\mu, \tag{4.28}$$

where e designates an electrical charge, which gives to A_μ the dimension of an electromagnetic potential, we obtain the Hamiltonian of a charged particle in the presence of electromagnetic and gravitational field

$$H = (1/2m)(p_\mu - eA_\mu)(p_\nu - eA_\nu)g^{\mu\nu}. \tag{4.29}$$

The equations of motion are

$$\frac{d}{d\tau}p_\mu = -\frac{\partial H}{\partial x^\mu}, \quad \frac{d}{d\tau}x^\mu = \frac{\partial H}{\partial p_\mu}, \tag{4.30}$$

where H is given by Eq. (4.29) and τ now recovers its significance of proper time

$$d\tau^2 = -g_{\mu\nu}\, dx^\mu\, dx^\nu. \tag{4.31}$$

From the equations (4.30) and (4.31), it is easy to obtain the well known equations of a charged particle in a gravitational field $g^{\mu\nu}$ and, in the presence of an electromagnetic field $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$,

$$m\left(\frac{dU^\lambda}{d\tau} + \{{}^{\lambda}_{\mu\nu}\}U^\mu U^\nu\right) = g^{\lambda\rho}eF_{\rho\sigma}U^\sigma; \tag{4.32}$$

$U^\lambda = dx^\lambda/d\tau$ is the four-velocity and $\{{}^{\lambda}_{\mu\nu}\}$ designates Christoffel's symbols.

V. DISCUSSION

Thus, we have defined a gauge theory in Hamiltonian classical mechanics in a perfectly autonomous way. In fact, this theory is not a complete one since in this work we never dealt with the equations of evolution of gauge fields. However, it is well known that they are governed by the equations of Einstein and Maxwell. In our opinion, the essential result of this paper is to point out, in the perspective of gauge theories, the necessary existence of long range interactions: the electromagnetic and gravitational fields. The result turns out to be even more precise: We've only found those two fields, which is satisfying in classical mechanics.

Let us note, however, that we used a first order approximation in p. This leads us to suppose that, if we had not assumed its validity, we could have found some new physical features (whether some small effects in the existing fields, or other fields). However, this raises difficult problems which we shall not evoke here and that we set aside for a further work.

[1] H. Weyl. *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950).
[2] C.N. Yang and R.L. Mills, Phys. Rev. **96**, 191 (1954).
[3] R. Utiyama, Phys. Rev. **101**, 1597 (1956).
[4] T.W.B. Kibble, J. Math. Phys. **2**, 212 (1961).
[5] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).
[6] A. Salam, in *Elementary Particle Theory Nobel Symposium* 1968, edited by Nils Svartholm (Wiley, New York, 1968).
[7] E.S. Abers and B.W. Lee, Phys. Rep. C **9**, 1 (1973).
[8] L. Fonda and G.C. Ghirardi, *Symmetry Principles in Quantum Physics* (Dekker, New York, 1970), p. 226.
[9] L. Van Hove, Acad. R. Belg. Bull. Sem Sér. Sci. **37**, 610 (1951).
[10] V. Arnold, *Mathematical Methods of Classical Mechanics* (Mir, Moscow, 1975); *Springer Graduate Texts in Math.* N°60 (Springer, New York, 1978).
[11] J.W. Leach, *Classical Mechanics* (Methuen, London, 1957).
[12] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1950).
[13] See Ref. 11, Chap. 10.
[14] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), p. 365–373.

Quasiprobability distributions and the analysis of the linear quantum channel with thermal noise

Carl W. Helstrom[a)]
Department of Applied Physics & Information Science, University of California, San Diego, La Jolla, California 92093

(Received 7 November 1978)

The quasiprobability distributions (QPD) introduced by Cahill and Glauber much simplify calculating the density operator of the output of a linear quantum channel from the density operator of the input when the channel attenuates and is corrupted by thermal noise. This channel models, for instance, an attenuator viewed as a simple harmonic oscillator in contact with a heat bath, or the relation between the radiating field mode in the aperture of a transmitter and the field mode it excites at the aperture of a distant receiver. It is shown how the attenuation and the noise in the channel modify the dependence of the QPD on the ordering parameter s. The form of the QPD at the output indicates that the greater the attenuation, the more nearly the state of the output resembles a classical state. Indeed, the addition of thermal noise contributing on the average only one photon suffices to convert an arbitrary unimodal state into a classical state. The principal result of the paper is applied to determining the output of a quantum channel whose input mode is in a generalized coherent state.

1. THE QUASIPROBABILITY DISTRIBUTION

The quasiprobability distribution (QPD) was introduced by Cahill and Glauber [1] as a comprehensive representation of the density operator ρ of a one-dimensional quantum system, and they showed how it could be used to evaluate expected values of variously ordered functions of the photon annihilation and creation operators a and a^+. The QPD $W(\alpha,s)$ is defined by

$$W(\alpha,s) = \mathrm{Tr}[\,\rho\, T(\alpha,s)], \tag{1.1}$$

where $\alpha = \alpha_x + i\alpha_y$ is a complex variable and the operator $T(\alpha,s)$ is the two-dimensional Fourier transform of the s-ordered displacement operator

$$D(\xi,s) = \exp(\tfrac{1}{2}s|\xi|^2)\, D(\xi) = \exp(\xi a^+ - \xi^* a + \tfrac{1}{2}s|\xi|^2\mathbf{1}), \tag{1.2}$$

that is,

$$T(\alpha,s) = \int D(\xi,s)\exp(\alpha\xi^* - \alpha^*\xi)\, d^2\xi/\pi = \int \exp[\xi(a^+ - \alpha^*\mathbf{1}) - \xi^*(a - \alpha\mathbf{1}) + \tfrac{1}{2}s|\xi|^2\mathbf{1}] \times d^2\xi/\pi, \tag{1.3}$$

in which, as in all integrals here, the integration is carried over the entire plane of the complex integration variable; $d^2\xi = d\xi_x\, d\xi_y$ and $\mathbf{1}$ is the identity operator. The density operator ρ can inversely be expressed as

$$\rho = \int W(\alpha,s)\, T(\alpha,-s)\, d^2\alpha/\pi. \tag{1.4}$$

The parameter s is called the *ordering parameter*; $s = 1$ corresponds to normal ordering, $s = 0$ to symmetrical ordering, and $s = -1$ to antinormal ordering of the operator functions of a and a^+ being averaged. The mathematical properties of the QPD $W(\alpha,s)$ as a function of the continuous parameter s were thoroughly explored by Cahill and Glauber. For $s = -1$, in terms of the familiar coherent states $|\alpha\rangle$, the QPD is

$$W(\alpha,-1) = \langle\alpha|\,\rho\,|\alpha\rangle, \tag{1.5}$$

which always exists and is nonnegative; $W(\alpha,0)$ is the Wigner distribution; [2] and if a P representation of the density operator exists, its P representative is

$$P(\alpha) = W(\alpha,1)/\pi. \tag{1.6}$$

It is the behavior of the QPD $W(\alpha,s)$ as s approaches 1 through real values in $-\infty < s < 1$ that is of primary interest. Here we shall show how the QPD can be used to simplify calculating the density operator ρ_1 at the output of a linear quantum channel subject to thermal noise when the density operator ρ_0 of the input is given. The attenuation and the noise in the channel significantly alter the dependence of the QPD on the ordering parameter s.

2. THE LINEAR QUANTUM CHANNEL

The linear quantum channel was introduced by Takahasi [3] in extending the calculations of photon statistics for attenuators and masers initiated by Shimoda, Takahasi, and Townes. [4] At the input to the channel are a principal mode and a parasitic mode, which can be considered as coupled quantum harmonic oscillators whose states are situated in Hilbert spaces $\mathscr{H}_0$ and $\mathscr{H}_0'$, respectively. Operating on states in $\mathscr{H}_0$ are the photon annihilation and creation operators a and a^+, and the density operator of the principal input mode is ρ_0. Operating on states in $\mathscr{H}_0'$ are the annihilation and creation operators a' and a'^+, and the density operator ρ_0' of the parasitic input mode is that of a harmonic oscillator in thermal equilibrium with a heat bath,

[a)]This research was carried out under Research Grant ENG 77-04500 from the National Science Foundation.

$$\rho_0' = (\pi N)^{-1} \int \exp(-|\alpha|^2/N)|\alpha\rangle\langle\alpha|\, d^2\alpha, \quad (2.1)$$

where N, the mean number of thermal photons, is related to the absolute temperature $\mathscr{T}$ through the Planck formula,

$$N = [\exp(h\nu/k\mathscr{T}) - 1]^{-1},$$

$h\nu$ being the photon energy and k Boltzmann's constant. The density operator of the pair of input modes is $\rho_0 \otimes \rho_0'$ in the product $\mathscr{H}_0 \otimes \mathscr{H}_0'$ of the Hilbert spaces.

At the output of the quantum channel are a principal mode, for which the annihilation and creation operators are b and b^+, and a parasitic mode, for which the annihilation and creation operators are b' and b'^+. These operators, which act in the product space $\mathscr{H}_0 \otimes \mathscr{H}_0'$, are related to those for the input through the unitary transformation

$$b = \kappa a \otimes \mathbf{1}' + \kappa' \mathbf{1} \otimes a', \quad b = -\kappa' a \otimes \mathbf{1}' + \kappa^* \mathbf{1} \otimes a', \quad (2.2)$$

where $\mathbf{1}$ and $\mathbf{1}'$ are the identity operators in $\mathscr{H}_0$ and $\mathscr{H}_0'$, respectively. We call κ the *transmittance*, and we define $\kappa' = (1 - |\kappa|^2)^{1/2}$ as real. Because the pairs a, a^+ and a', a'^+ obey the usual commutation rules

$$[a,a^+] = aa^+ - a^+a = \mathbf{1},$$

$$[a',a'^+] = a'a'^+ - a'^+a' = \mathbf{1}',$$

so do the pairs b, b^+ and b', b'^+,

$$[b,b^+] = [b',b'^+] = \mathbf{1} \otimes \mathbf{1}',$$

and b and b' commute

$$[b, b'] = [b,b'^+] = [b^+,b'] = [b^+,b'^+] = 0,$$

all by virtue of (2.2).

The eigenstates of the number operators a^+a and a'^+a' form orthogonal bases in $\mathscr{H}_0$ and $\mathscr{H}_0'$, respectively, their products constituting an orthogonal basis in $\mathscr{H}_0 \otimes \mathscr{H}_0'$. We can construct an alternative basis for this product space by operating on its ground state $|\mathbf{0}\rangle$ with various powers of b^+ and b',

$$(m!\, n!)^{-1/2}\, b^{+m}\, b'^{+n}\, |\mathbf{0}\rangle = |m\rangle_b |n\rangle_{b'},$$

where $|m\rangle_b$ and $|n\rangle_{b'}$ are the eigenstates of the commuting number operators b^+b and b'^+b', respectively. The product space $\mathscr{H}_0 \otimes \mathscr{H}_0'$ can thus alternatively be represented by the product $\mathscr{H}_1 \otimes \mathscr{H}_1'$ of Hilbert spaces whose orthogonal bases are the sets $|m\rangle_b$ and $|n\rangle_{b'}$ of number eigenstates, respectively. The space $\mathscr{H}_1$ bears the states of the principal output mode; $\mathscr{H}_1'$ bears those of the parasitic output mode. In these spaces the identity operators are $\mathbf{1}_b$ and $\mathbf{1}_b'$, respectively. We can then replace b and b' in (2.2) by $b \otimes \mathbf{1}_b'$, and $\mathbf{1}_b \otimes b'$, with b and b' photon annihilation operators in $\mathscr{H}_1$ and $\mathscr{H}_1'$, respectively. Inversely, the original annihilation operators are given by

$$a \otimes \mathbf{1}' = \kappa^* b \otimes \mathbf{1}_b', - \kappa' \mathbf{1}_b \otimes b',$$

$$\mathbf{1} \otimes a' = \kappa' b \otimes \mathbf{1}_b', + \kappa \mathbf{1}_b \otimes b'. \quad (2.3)$$

The principal input mode might represent a harmonic oscillator at the moment it is placed in thermal contact with a heat bath. The effect of the heat bath is accounted for by the coupling with the parasitic mode, whose density operator is ρ_0' of (2.1). The principal output mode then represents the oscillator at a later time t, and the transmittance κ equals $\exp(-\gamma t)$, where γ characterizes the rate of conversion of coherent energy into heat. In this way the linear quantum channel models the behavior of an attenuator.[3,5]

Alternatively, the principal input mode might be the radiating mode of the field in the aperture of a transmitter, and the principal output mode might be the field mode it excites in the aperture of a distant receiver. The parasitic input mode then represents the environment as it contributes thermal noise, and the parasitic output mode represents the sink of transmitted energy that misses the receiving aperture.[3] In both examples, the density operator of the parasitic output mode must be calculated by forming the trace over the states in the space $\mathscr{H}_1'$ of the parasite.

3. THE OUTPUT QUASIPROBABILITY DISTRIBUTION

The state of the output of the quantum channel can be calculated from the input state $\rho_0 \otimes \rho_0'$ by the method described previously[6] for handling the R-representation of the density operators of coupled harmonic oscillators. It is easier to employ their quasiprobability distributions, and we shall derive a simple formula for the QPD of the output density operator ρ_1 in terms of that of the input density operator ρ_0.

The state of the combined input modes is

$$\rho_0 \otimes \rho_0' = \int\int W_0(\alpha,s)\, W_0'(\alpha',s) T(\alpha, -s) \otimes T'(\alpha', -s) \times d^2\alpha\, d^2\alpha'/\pi^2, \quad (3.1)$$

where $T'(\alpha',s)$ is defined in terms of the operators a', a'^+ by an equation like (1.3). By a formula of Cahill and Glauber's [Ref. 7, Eq. (6.25)], we can write

$$T(\alpha, -s) \otimes T'(\alpha', -s) = 4(1+s)^{-2} \exp\left[\ln\left(\frac{s-1}{s+1}\right) [(a^+ - \alpha^*\mathbf{1})(a - \alpha\mathbf{1}) \otimes \mathbf{1}' + \mathbf{1} \otimes (a'^+ - \alpha'^*\mathbf{1}')(a' - \alpha'\mathbf{1}')]\right]. \quad (3.2)$$

Defining the complex variables β, β' by the transformation like that in (2.2),

$$\beta = \kappa\alpha + \kappa'\alpha', \quad \beta' = -\kappa'\alpha + \kappa^*\alpha', \quad (3.3)$$

we find by (2.3)

$$(a^+ - \alpha^*\mathbf{1})(a - \alpha\mathbf{1}) \otimes \mathbf{1}' + \mathbf{1} \otimes (a'^+ - \alpha'^*\mathbf{1}')(a' - \alpha'\mathbf{1}') = (b^+ - \beta^*\mathbf{1}_b)(b - \beta\mathbf{1}_b) \otimes \mathbf{1}_b' + \mathbf{1}_b \otimes (b'^+ - \beta^*\mathbf{1}_b')(b' - \beta\mathbf{1}_b'). \quad (3.4)$$

In the transition from input to output of the quantum channel, the pair of operators a, a' is transformed into the pair b, b' given by (2.2); and the operator $T(\alpha, -s) \otimes T'(\alpha', -s)$ goes into

$$T(\beta, -s) \otimes T'(\beta', -s) = 4(1+s)^{-2} \exp\left[\ln\left(\frac{s-1}{s+1}\right)(b^+ - \beta^*\mathbf{1}_b) \times (b - \beta\mathbf{1}_b) \otimes \mathbf{1}_b' + \mathbf{1}_b \otimes (b'^+ - \beta^*\mathbf{1}_b') \times (b' - \beta\mathbf{1}_b')\right]. \quad (3.5)$$

The density operator of the pair of modes at the output of the channel is then

$$\rho_f = \int\int W_0(\alpha,s)\, W_0'(\alpha',s) T(\beta,-s) \otimes T'(\beta',-s) \times d^2\alpha\, d^2\alpha'/\pi^2, \tag{3.6}$$

with β, β' given by (3.3) in terms of α, α'.

The parasitic mode described by states in the space $\mathscr{H}_1'$ being unobserved, the density operator ρ_1 of the principal output mode is found by taking the trace of ρ_f over those states. Since Tr $T'(\beta',-s) = 1$, we find

$$\rho_1 = \int\int W_0(\alpha,s)\, W_0'(\alpha',s) T(\kappa\alpha + \kappa'\alpha', -s) \times d^2\alpha\, d^2\alpha'/\pi^2. \tag{3.7}$$

The QPD of the density operator ρ_1 is therefore, by (1.1)

$$\begin{aligned} W_1(\gamma,s) &= \mathrm{Tr}[\rho_1\, T(\gamma,s)] \\ &= \int\int W_0(\alpha,s)\, W_0'(\alpha',s) \times \delta^{(2)}(\gamma - \kappa\alpha - \kappa'\alpha')\, d^2\alpha\, d^2\alpha'/\pi \\ &= \int W_0(\alpha,s) W_0'(\kappa'^{-1}(\gamma - \kappa\alpha), s) \times d^2\alpha/\pi\kappa'^2, \end{aligned} \tag{3.8}$$

since by [Ref. 7, Eq. (6.40)],

$$\mathrm{Tr}[T(\beta,-s)\, T(\gamma,s)] = \pi\delta^{(2)}(\beta - \gamma)$$

in terms of the two-dimensional delta function.

The QPD of the density operator ρ_0' of the parasitic input mode, given in (2.1), is

$$W_0'(\alpha',s) = 2(2N+1-s)^{-1} \exp[-2|\alpha'|^2/(2N+1-s)] \tag{3.9}$$

by [Ref. 1, Eq. (7.8)], and we can write (3.8) as

$$\begin{aligned} W_1(\gamma,s) &= 2(2N+1-s)^{-1} \int W_0(\alpha,s) \times \exp\left(-\frac{2|\kappa|^2|\alpha - \kappa^{-1}\gamma|^2}{\kappa'^2(2N+1-s)}\right)\frac{d^2\alpha}{\pi\kappa'^2} \\ &= 2|\kappa|^{-2}(s-\tau)^{-1}\int W_0(\alpha,s) \times \exp[-2|\alpha - \kappa^{-1}\gamma|^2(s-\tau)^{-1}]\, d^2\alpha/\pi, \end{aligned}$$

where τ is defined by

$$|\kappa|^2(s-\tau) = \kappa'^2(2N+1-s)$$

or

$$\tau = 1 - |\kappa|^{-2}(1 - s + 2N\kappa'^2).$$

By Cahill and Glauber's displacement formula

$$W_0(\beta,\tau) = 2(s-\tau)^{-1}\int W_0(\alpha,s) \exp[-2|\alpha-\beta|^2 \times (s-\tau)^{-1}]\, d^2\alpha/\pi, \quad \mathrm{Re}\,\tau < \mathrm{Re}s \tag{3.10}$$

[Ref. 7, Eq. (6.41)], we finally obtain for the QPD of the output density operator ρ_1

$$W_1(\gamma,s) = |\kappa|^{-2}\, W_0(\gamma/\kappa, 1 - |\kappa|^{-2}(1 - s + 2N\kappa'^2)). \tag{3.11}$$

The term $2N\kappa'^2$ on the right-hand side represents the addition of Gaussian random noise from the heat bath or the environment. A model for the addition of such noise to an oscillator with initial density operator ρ determines the final density operator ρ' by

$$\rho' = \int P_n(\alpha)\, D(\alpha)\rho D^+(\alpha)\, d^2\alpha, \tag{3.12}$$

where

$$P_n(\alpha) = (\pi N_0)^{-1}\exp(-|\alpha|^2/N_0),$$

N_0 is the mean number of noise photons added, and $D(\alpha)$ is the displacement operator

$$D(\alpha) = \exp(\alpha a^+ - \alpha^* a). \tag{3.13}$$

By using (1.1), (3.9), and the displacement formula (3.10), it is not hard to show that the QPD's of ρ and ρ' are related by

$$W'(\alpha,s) = W(\alpha, s - 2N_0). \tag{3.14}$$

In (3.11), $N_0 = N\kappa'^2$, as is to be expected from (2.3).

The QPD $W_0(\alpha,-1) = \langle\alpha|\rho_0|\alpha\rangle$ is always nonnegative and free of singularities.[1] If we rewrite (3.11) as

$$W_1(\gamma, 1 + 2\kappa'^2 N - |\kappa|^2(1-\tau)) = |\kappa|^{-2}\, W_0(\gamma/\kappa, \tau) \tag{3.15}$$

and put $\tau = -1$, we see that for any transmittance κ, $W_1(\gamma,1) = P_1(\gamma)/\pi$ will be nonnegative and free of singularities for a sufficiently large noise level, $N > |\kappa|^2/\kappa'^2$. [If $W_1(\gamma,s) \geqslant 0$ for any real value of s, the displacement formula (3.10) shows that $W_1(\gamma,\tau) \geqslant 0$ for any real value of τ less than s.] If we call a state with a nonnegative and well-behaved P representative a "classical state," we conclude that the addition of sufficient thermal noise will turn an arbitrary initial state into a classical state.

In interpreting (3.11) further, we take $N = 0$, writing it as

$$W_1(\gamma,s) = |\kappa|^{-2}\, W_0(\gamma/\kappa, 1 - |\kappa|^{-2}(1-s)). \tag{3.16}$$

Putting $s = 1$, we see that if the input density operator ρ_0 possesses a P representation, so does the output density operator ρ_1, and by (1.6) their P representatives are related by

$$P_1(\gamma) = |\kappa|^{-2} P_0(\gamma/\kappa).$$

The complex amplitudes of the mixture of coherent states at the input,

$$\rho_0 = \int P_0(\alpha)|\alpha\rangle\langle\alpha|\, d^2\alpha,$$

contract by the transmittance κ, as is to be expected. If, on the other hand, $W_0(\alpha,s)$ is singular at $s = 1$, so that no P representation exists for the input density operator ρ_0, none will exist for the output density operator ρ_1 when $N = 0$.

Putting $N = 0$, $\tau = -1$ into (3.15), we obtain

$$\begin{aligned} W_1(\gamma, 1 - 2|\kappa|^2) &= |\kappa|^{-2}\, W_0(\gamma/\kappa, -1) \\ &= |\kappa|^{-2}\langle\gamma/\kappa|\rho_0|\gamma/\kappa\rangle. \end{aligned} \tag{3.17}$$

The right-hand side of this equation is always nonnegative, and by (3.10), therefore, $W_1(\gamma,s) \geqslant 0$ for $s \leqslant 1 - 2|\kappa|^2$. Thus the QPD $W_1(\gamma,s)$ is nonnegative at least over a segment of the Res axis whose right-hand end approaches the point $s = 1$ as $|\kappa|^2$ decreases to zero, the state ρ_1 of the output approaching closer and closer to a classical state.

The nature of this approach to a classical state can be understood by comparing ρ_1 with a density operator ρ_1' whose P representation is

$$\rho_1' = |\kappa|^{-2} \int \langle \gamma/\kappa| \rho_0 | \gamma/\kappa\rangle \, |\gamma\rangle \langle \gamma| d^2\gamma/\pi$$
$$= \int \langle\alpha|\rho_0|\alpha\rangle |\kappa\alpha\rangle \langle\kappa\alpha| d^2\alpha/\pi. \tag{3.18}$$

This density operator represents a classical state whose P representative is

$$P_1'(\alpha) = |\kappa|^{-2}\langle\alpha/\kappa|\rho_0|\alpha/\kappa\rangle/\pi. \tag{3.19}$$

Its QPD, at all values of s for which $W_1(\alpha,s)$ exists, is by (3.10),

$$W_1'(\alpha,s) = W_1(\alpha,s-2|\kappa|^2)$$
$$= \int K(|\alpha-\beta|)\, W_1(\beta,s)\, d^2\beta,$$
$$K(x) = \pi^{-1}|\kappa|^{-2}\exp(-x^2/|\kappa|^2). \tag{3.20}$$

To prove this, we note that (1.6), (3.17), and (3.19) imply (3.20) for $s=1$. Thence the displacement formula (3.10) implies (3.20) for values of s with $\mathrm{Re}\,s<1$, whereupon analytic continuation extends it to all portions of the s plane where $W_1(\beta,s)$ exists. As $|\kappa|^2\to 0$, the kernel $K(|\alpha-\beta|)$ approaches the two-dimensional delta function $\delta^{(2)}(\alpha-\beta)$.

If as $|\kappa|^2\to 0$ we imagine increasing the mean number of photons in the input state, while preserving the form of its density operator and keeping the mean number of photons in the output state fixed, the density operator ρ_1' defined by (3.18) approaches the true output density operator ρ_1 in the sense that

$$\mathrm{Tr}(\rho_1'-\rho_1)^2 = o(|\kappa|^2). \tag{3.21}$$

The proof is given in the Appendix. If the states $|\chi_n\rangle$ form an orthogonal basis,

$$|\langle\chi_n|(\rho_1'-\rho_1)|\chi_m\rangle|^2 = |\mathrm{Tr}\,[(\rho_1'-\rho_1)|\chi_m\rangle\langle\chi_n|]|^2$$
$$\leqslant \mathrm{Tr}(\rho_1'-\rho_1)^2\,\mathrm{Tr}(|\chi_m\rangle$$
$$\times\ \langle\chi_n|\chi_n\rangle\langle\chi_m|)$$
$$= \mathrm{Tr}(\rho_1'-\rho_1)^2 = o(|\kappa|^2)$$

by the Schwarz inequality. Thus in this limit the matrix elements of the operator ρ_1' in any orthogonal representation approach those of the operator ρ_1,

$$|\langle\chi_n|\rho_1'|\chi_m\rangle - \langle\chi_n|\rho_1|\chi_m\rangle| = o(|\kappa|).$$

As the number of photons in the input state becomes larger and larger with the decrease of the transmittance κ to zero, the state it induces at the output of the linear quantum channel becomes ever more difficult to distinguish from a classical state, that is, from a mixture of coherent states with a nonnegative density

$$P_1'(\gamma) = |\kappa|^{-2}\langle\gamma/\kappa|\rho_0|\gamma/\kappa\rangle.$$

If the input state possesses a P representation whose P representative $P_0(\alpha)$ is negative in part of the complex α plane, the P-representative $P_1(\alpha)$ of the output state will also be negative somewhere, and it cannot be approximated by ρ_1', for which $P_1'(\alpha)\geqslant 0$ everywhere. If $P_1(\alpha)<0$ for some values of α, however, as shown by Picinbono and Rousseau,[8] the operator ρ_0 in (3.18) with

$$P_0(\alpha) = |\kappa|^2 P_1(\kappa\alpha)$$

cannot be a density operator for arbitrarily small values of $|\kappa|$, and the passage to the limit $|\kappa|\to 0$ contemplated here cannot be carried out. Eq. (3.20), with $s=1$, indeed furnishes an alternative proof of their result.

If we compare (3.20) with (3.14), we see that the approximate classical state ρ_1' at the output could be created from the actual state ρ_1 by the addition of thermal noise bearing an average number $N_0 = |\kappa|^2$ of photons. Even at the input ($\kappa=1$) the addition of thermal noise contributing an average of only one photon suffices to turn any arbitrary state ρ_0 into a classical state. The relative insignificance of that one photon when the state ρ_0 itself carries a large average number of photons accords with the common belief that fields with a larger number of photons in each mode behave essentially like classical fields.

4. THE ATTENUATION OF A GENERALIZED COHERENT STATE

To illustrate the usefulness of (3.11) we apply it to calculating the density operator at the output of a linear quantum channel when the input mode is a generalized coherent state of the kind described by Stoler[9] and Lu.[10] The properties of such states $|\beta_0;\mu_0,\nu_0\rangle$ were extensively treated by Yuen.[11] They are the right-eigenstates of the operator $\mu_0 a + \nu_0 a^+$, where $\mu_0^2 - |\nu_0|^2 = 1$,

$$(\mu_0 a + \nu_0 a^+)|\beta_0;\mu_0,\nu_0\rangle = \beta_0|\beta_0;\mu_0,\nu_0\rangle, \tag{4.1}$$

and for $\mu_0 = 1$ they reduce to the ordinary coherent states $|\beta_0\rangle$. (We assume μ_0 real throughout.) From [Ref. 11, Eq. (3.20)],

$$\langle\alpha|\beta_0;\mu_0,\nu_0\rangle = \mu_0^{-1/2}\exp\,[-\tfrac12|\alpha|^2 - \tfrac12|\beta_0|^2$$
$$+ (\alpha\beta_0^* + \tfrac12\nu_0^*\beta_0^2 - \tfrac12\nu_0\alpha^{*2})/\mu]. \tag{4.2}$$

Hence by (1.5)

$$W_0(\alpha,-1) = \langle\alpha|\beta_0;\mu_0,\nu_0\rangle\,\langle\beta_0;\mu_0,\nu_0|\alpha\rangle$$
$$= \mu_0^{-1}\exp\Big(-|\alpha-\hat\beta_0|^2 - \frac{\nu_0}{2\mu_0}(\alpha^*-\hat\beta_0^*)^2$$
$$-\frac{\nu_0^*}{2\mu_0}(\alpha-\hat\beta_0)^2\Big) \tag{4.3}$$

with $\hat\beta_0 = \mu_0\beta_0 - \nu_0\beta_0^*$. The QPD $W_0(\gamma,s)$ for arbitrary s is found by substituting (4.3) into the displacement formula (3.10) with β replaced by γ, τ replaced by s, and s replaced by -1. Upon evaluating the Gaussian integral one obtains

$$W_0(\gamma,s) = \{[\mu_0^2 - \tfrac14(1+s)]^2 - \mu_0^2|\nu_0|^2\}^{-1/2}$$
$$\times\exp\{\{[\mu_0^2-\tfrac12(1+s)]^2 - \mu_0^2|\nu_0|^2\}^{-1}$$
$$\times\{[\tfrac12(1+s)-\mu_0^2]|\gamma-\hat\beta_0|^2 - \tfrac12\mu_0\nu_0^*$$
$$\times(\gamma-\hat\beta_0)^2 - \tfrac12\mu_0\nu_0(\gamma^*-\hat\beta_0^*)^2\}\} \tag{4.4}$$

The integration of (3.10) here requires $\mathrm{Re}\,s < -1$, but by analytical continuation the result is valid everywhere in the complex s plane when it is cut along the real axis in $s_2 \leqslant \mathrm{Re}s \leqslant s_1$, with

$$s_1 = (\mu_0 + |\nu_0|)^2 > 1, \quad s_2 = s_1^{-1} = (\mu_0 - |\nu_0|)^2 < 1.$$

For $\beta_0 = 0$, (4.4) reduces to [Ref. 1, Eq. (7.41)], Cahill and Glauber having already noted the interesting analytical properties of the QPD of this type of state. Outside of the cut the QPD is real and positive for all real values of s.

If the principal mode at the input of a linear quantum channel is in such a generalized coherent state, the QPD $W_1(\gamma,s)$ of the density operator ρ_1 at the output is obtained by substituting (4.4) into (3.11). This QPD $W_1(\gamma,s)$ is analytic over the s-plane cut in

$$s_2' \leqslant \mathrm{Re}s \leqslant s_1',$$

where

$$s_1' = \kappa'^2(1+2N) + |\kappa|^2 s_1, \quad s_2' = \kappa'^2(1+2N) + |\kappa|^2 s_2.$$

The cut is now shorter by a factor $|\kappa|^2$ and shifted toward larger values of s by an amount depending on the noise. When $N=0$ the cut still includes the point $s=1$.

The output density operator ρ_1, as we shall prove, can be written as

$$\rho_1 = U(\mu,\nu)\rho(\delta)U^+(\mu,\nu), \tag{4.5}$$

where

$$\rho(\delta) = (\pi N')^{-1}\int \exp(-|\alpha-\delta|^2/N')|\alpha\rangle\langle\alpha| d^2\alpha \tag{4.6}$$

is the density operator of a harmonic oscillator containing both a coherent signal of complex amplitude δ and thermal noise contributing an average number N' of photons, and $U(\mu,\nu)$ is the unitary operator transforming the annihilation operator a into $\mu a - \nu a^+$,

$$U^+(\mu,\nu)aU(\mu,\nu) = \mu a - \nu a^+,$$
$$U^+(\mu,\nu)a^+U(\mu,\nu) = \mu a^+ - \nu^* a, \tag{4.7}$$

with $\mu^2 - |\nu|^2 = 1$. The parameters μ, ν, N', and δ of the output state ρ_1 in (4.5) are given by the equations

$$N'\mu^2 + (N'+1)|\nu|^2 = |\kappa|^2|\nu_0|^2 + \kappa'^2 N, \tag{4.8}$$

$$(2N'+1)\mu\nu = \kappa^2\mu_0\nu_0, \tag{4.9}$$

$$\delta = \mu\kappa\hat{\beta}_0 + \nu\kappa^*\hat{\beta}_0^*, \tag{4.10}$$

in terms of the parameters μ_0, ν_0, and $\hat{\beta}_0$ of the input. If for simplicity we take ν_0 and κ real, (4.8) and (4.9) can be solved by determining θ from

$$\tanh 2\theta = \frac{2\kappa^2\mu_0\nu_0}{2\kappa'^2N + 1 + 2\kappa^2\nu_0^2},$$

whereupon

$$\mu = \cosh\theta, \quad \nu = \sinh\theta,$$
$$2N' + 1 = 2\kappa^2\mu_0\nu_0/\sinh 2\theta.$$

Equations (4.5) and (4.6) show that the output state ρ_1 possesses a P representation in terms not of the usual coherent states $|\alpha\rangle$, but of the generalized coherent states

$$|\alpha;\mu,\nu\rangle = U(\mu,\nu)|\alpha\rangle, \tag{4.11}$$

that is,

$$\rho_1 = (\pi N')^{-1}\int \exp(-|\alpha-\delta|^2 N')|\alpha;\mu,\nu\rangle\langle\alpha;\mu,\nu| d^2\alpha. \tag{4.12}$$

To prove (4.5), we use (1.1) to determine the QPD $W_1(\gamma,s)$ and equate the result with (4.4) after its substitution into (3.11). Now

$$\begin{aligned} W_1(\gamma,s) &= \mathrm{Tr}\,\rho_1 T(\gamma,s) \\ &= \mathrm{Tr}[U(\mu,\nu)\rho(\delta)U^+(\mu,\nu)T(\gamma,s)] \\ &= \mathrm{Tr}[\rho(\delta)U^+(\mu,\nu)T(\gamma,s)U(\mu,\nu)] \\ &= \mathrm{Tr}[\rho(\delta)T(\gamma,s;\mu,\nu)], \end{aligned} \tag{4.13}$$

where by (1.3)

$$\begin{aligned} T(\gamma,s;\mu,\nu) &= U^+(\mu,\nu)T(\gamma,s)U(\mu,\nu) \\ &= \int \exp\Big(\gamma\xi^* - \gamma^*\xi + \frac{1}{2}s|\xi|^2\Big) \\ &\quad \times U^+(\mu,\nu)D(\xi)U(\mu,\nu)\,d^2\xi/\pi \\ &= \int \exp\Big(\gamma\xi^* - \gamma^*\xi + \frac{1}{2}s|\xi|^2\Big) \\ &\quad \times D(\mu\xi + \nu\xi^*)\,d^2\xi/\pi \end{aligned} \tag{4.14}$$

since by (4.7) and (3.13)

$$\begin{aligned} &U^+(\mu,\nu)D(\xi)U(\mu,\nu) \\ &\quad = \exp[\xi(\mu a^+ - \nu^* a) - \xi^*(\mu a - \nu a^+)] \\ &\quad = D(\mu\xi + \nu\xi^*) = D(\tau), \end{aligned} \tag{4.15}$$

$$\tau = \mu\xi + \nu\xi^*.$$

After substituting (4.14) into (4.13) we must evaluate

$$\begin{aligned} \mathrm{Tr}[\rho(\delta)D(\tau)] &= (\pi N')^{-1}\int \exp[-|\alpha-\delta|^2/N'] \\ &\quad \times \langle\alpha|D(\tau)|\alpha\rangle\,d^2\alpha \\ &= (\pi N')^{-1}\int \exp(-N'^{-1}|\alpha-\delta|^2 \\ &\quad + \alpha^*\tau - \tau^*\alpha - \tfrac{1}{2}|\tau|^2)\,d^2\alpha \\ &= \exp[\delta^*\tau - \delta\tau^* - \tfrac{1}{2}(2N'+1)|\tau|^2]. \end{aligned} \tag{4.16}$$

Combining (4.13) and (4.14) and using (4.16) we obtain

$$\begin{aligned} W_1(\gamma,s) &= \int \exp[\gamma\xi^* - \gamma^*\xi + \tfrac{1}{2}s|\xi|^2 + \delta^*(\mu\xi + \nu\xi^*) \\ &\quad - \delta(\mu\xi^* + \nu^*\xi) - (N'+\tfrac{1}{2})|\mu\xi + \nu\xi^*|^2]\,d^2\xi/\pi \\ &= \int \exp\{(\gamma-\lambda)\xi^* - (\gamma^*-\lambda^*)\xi \\ &\quad - [(N'+\tfrac{1}{2})(\mu^2+|\nu|^2) - \tfrac{1}{2}s]|\xi|^2 \\ &\quad - (N'+\tfrac{1}{2})\mu\nu\xi^{*2} - (N'+\tfrac{1}{2})\mu\nu^*\xi^2\}\,d^2\xi/\pi \end{aligned} \tag{4.17}$$

after some algebra, where

$$\lambda = \mathrm{Tr}(\rho_1 a) = \mu\delta - \nu\delta^*. \tag{4.18}$$

After evaluating the bivariate Gaussian integral, we find for the QPD

$$W_1(\gamma,s) = M^{-1/2}\exp(-M^{-1}\{[(N'+\tfrac{1}{2})(\mu^2+|\nu|^2) - \tfrac{1}{2}s]|\gamma'|^2 + (N'+\tfrac{1}{2})\mu\nu^*\gamma'^2 + (N'+\tfrac{1}{2})\mu\nu[PW{:}\gamma'^*{:}2]\}), \quad (4.19)$$

where

$$\gamma' = \gamma - \lambda,$$
$$M = [(N'+\tfrac{1}{2})(\mu^2+|\nu|^2) - \tfrac{1}{2}s]^2 - (2N'+1)^2\mu^2|\nu|^2. \quad (4.20)$$

When (4.4) is substituted into (3.11), we find

$$W_1(\gamma,s) = \{[\mu_0^2 - \tfrac{1}{2}(s'+1)]^2 - \mu_0^2|\nu_0|^2\}^{-1/2}|\kappa|^{-2} \times\exp(\{[\mu_0^2 - \tfrac{1}{2}(s'+1)]^2 - \mu_0^2|\nu_0|^2\}^{-1} \times\{[\tfrac{1}{2}(s'+1) - \mu_0^2]|\kappa^{-1}\gamma - \hat{\beta}_0|^2 - \tfrac{1}{2}\mu_0\nu_0 \times(\kappa^{*-1}\gamma^* - \hat{\beta}_0^*)^2 - \tfrac{1}{2}\mu_0\nu_0^* \times(\kappa^{-1}\gamma - \hat{\beta}_0)^2\})$$

with $s' = 1 + |\kappa|^{-2}(s - 1 - 2N\kappa'^2)$. Equating corresponding terms in (4.19) and (4.20) we obtain (4.8), (4.9), and

$$\lambda = \mu\delta - \nu\delta^* = \kappa\hat{\beta}_0.$$

Solving this last equation and its complex conjugate for δ yields (4.10). In Ref. 12 (4.5)–(4.10) were derived for $N=0$ by using the R-representation and the method of Ref. 6; the calculation was far more tedious than this.

When they carry a large average number of photons, these generalized coherent states manifest the similarity to classical states mentioned at the end of Sec. 3. This similarity is illustrated by the behavior of the "bunching ratio"

$$R = [\mathrm{Tr}(\rho_0 a^{+2}a^2) - \langle n\rangle^2]/\langle n\rangle^2,$$
$$\langle n\rangle = \mathrm{Tr}(\rho_0 a^+ a).$$

It is known [11] that R can be negative ("antibunching") for such states if the parameters μ_0, ν_0, and β_0 are appropriately chosen. It can be shown, however, that when the average number

$$\langle n\rangle = |\hat{\beta}_0|^2 + |\nu_0|^2$$

of photons is large, the minimum attainable value of this ratio R is approximately equal to $-\langle n\rangle^{-1}$ and goes to zero as the average number $\langle n\rangle$ increases without bound [13]. If the input to a linear channel is in such a generalized coherent state, the bunching ratio of the state of the principal output mode will be the same as that of the input, but the mean number of photons upon which to try to observe antibunching will be less by a factor of $|\kappa|^2$.

APPENDIX

To prove (3.21), that

$$\mathrm{Tr}(\rho_1' - \rho_1)^2 = \mathrm{Tr}\,\rho_1'^2 - 2\mathrm{Tr}\,\rho_1'\rho_1 + \mathrm{Tr}\rho_1^2 = o(|\kappa|^2) \quad (A1)$$

as $|\kappa|^2\to 0$, we evaluate each term by means of a general formula for the trace of the product of two arbitrary density operators ρ_1 and ρ_2 in terms of their QPD's,

$$\mathrm{Tr}(\rho_1\rho_2) = \int\int W_1(\alpha,s_1)\,W_2(\beta,s_2)\,\mathrm{Tr}[T(\alpha,-s_1) \times T(\beta,-s_2)]\,d^2\alpha\,d^2\beta/\pi^2$$
$$= 2(s_1+s_2)^{-1}\int\int W_1(\alpha,s_1)\,W_2(\beta,s_2) \times\exp\left(-\frac{2|\alpha-\beta|^2}{s_1+s_2}\right)d^2\alpha\,d^2\beta/\pi^2 \quad (A2)$$

by (1.4) and [Ref. 7, Eq. (6.39)], the real part of $s_1 + s_2$ being positive. Thus with $s_1 = s_2 = 1$,

$$\mathrm{Tr}\,\rho_1'^2 = \int\int W_1'(\alpha,1)\,W_1'(\beta,1)\exp(-|\alpha-\beta|^2) \times d^2\alpha\,d^2\beta/\pi^2. \quad (A3)$$

From (3.20) with $s_1 = 1$, $s_2 = 1 - 2|\kappa|^2$, $s_1 + s_2 = 2\kappa'^2$,

$$\mathrm{Tr}\,\rho_1'\rho_1 = 2(s_1+s_2)^{-1}\int\int W_1'(\alpha,s_1)\,W_1'(\beta,s_2+2|\kappa|^2) \times\exp\left(-\frac{2|\alpha-\beta|^2}{s_1+s_2}\right)d^2\alpha\,d^2\beta/\pi^2$$
$$= \kappa'^{-2}\int\int W_1'(\alpha,1)\,W_1'(\beta,1)\exp(-\kappa'^{-2} \times|\alpha-\beta|^{-2})\,d^2\alpha\,d^2\beta/\pi^2. \quad (A4)$$

Finally by (3.20), taking $s_1 = s_2 = 1 - 2|\kappa|^2$ in (A.2),

$$\mathrm{Tr}\,\rho_1^2 = (1-2|\kappa|^2)^{-1}\int\int W_1'(\alpha,1)\,W_1'(\beta,1) \times\exp[-(1-2|\kappa|^2)^{-1}|\alpha-\beta|^2]\,d^2\alpha\,d^2\beta/\pi^2. \quad (A5)$$

Hence, with $P_1'(\alpha) = W_1'(\alpha,1)/\pi$ the P representative of the approximate density operator ρ_1', we find

$$\mathrm{Tr}(\rho_1' - \rho_1)^2 = \int\int P_1'(\alpha)\,P_1'(\beta)\,F(|\alpha-\beta|)\,d^2\alpha\,d^2\beta, \quad (A6)$$

in which

$$F(x) = e^{-x}\{1 - 2\kappa'^{-2}\exp(-|\kappa|^2x^2/\kappa'^2) + (1-2|\kappa|^2)^{-1}\exp[-2|\kappa|^2x^2/(1-2|\kappa|^2)]\}. \quad (A7)$$

Keeping the mean number of photons in the output mode fixed as $|\kappa|^2\to 0$ is equivalent to keeping $P_1'(\alpha)$ constant during the passage to the limit. Expanding the exponential functions in the brackets { } in (A7), we find that the terms of order $|\kappa|^0$ and $|\kappa|^2$ cancel, whereupon $F(x)$ is proportional to $|\kappa|^4$, and $\mathrm{Tr}(\rho_1' - \rho_1)^2$ vanishes faster than $|\kappa|^2$ as $|\kappa|^2\to 0$, as in (3.21).

[1] K.E. Cahill and R.J. Glauber, Phys. Rev. **177**, 1882 (1969).
[2] E.B. Wigner, Phys. Rev. **40**, 749 (1932); J.E. Moyal, Proc. Cambridge Philos. Soc. **45**, 99 (1949).
[3] H. Takahasi, in *Advances in Communication Systems*, edited by A. Balakrishnan (Academic, New York, 1965), Vol. 1, p 227.
[4] K. Shimoda, H. Takahasi, C.H. Townes, J. Phys. Soc. Japan **12**, 686 (1957).
[5] W. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York, 1964). See §7.4, pp. 261–6.
[6] C.W. Helstrom, J. Math. Phys. **8**, 37 (1967).
[7] K.E. Cahill and R.J. Glauber, Phys. Rev. **177**, 1857 (1969).
[8] B. Picinbono and M. Rousseau, Phys. Rev. A **15**, 1648 (1977).
[9] D. Stoler, Phys. Rev D **1**, 3217 (1970).
[10] E.Y.C. Lu, Nuovo Cimento Lett. **2**, 1241 (1971).
[11] H.P. Yuen, Phys. Rev. A **13**, 2226 (1976).
[12] C.W. Helstrom, Trans IEEE **IT-25**, 69 (1979).
[13] C.W. Helstrom, Optics Commun. **28**, 363 (1979).

Gravitational duality and Bäcklund transformations[a)]

N. Papanicolaou
Department of Physics, University of California, Berkeley, California 94720

(Received 11 March 1979; accepted for publication 25 May 1979)

We analyze various forms of duality symmetry transformations occurring in the theory of axially symmetric gravitational fields and their relation to standard Bäcklund transformations. Appropriately interpreted, duality rotations provide genuine Bäcklund transformations for the fundamental SL(2,R) invariants associated with the metric, consisting of elementary algebraic substitutions; thereby the construction of the metric reduces to the solution of linear equations.

I. INTRODUCTION

A. General

Recent work concerning axially symmetric gravitational fields[1–4] revealed an interesting connection with a large variety of two-dimensional nonlinear systems, that were previously integrated by inverse scattering methods. Whereas a direct application of the inverse method to the gravitational problem is halted by a number of technical difficulties, Bäcklund transformations seem to lead to interesting results. A solution-generating procedure, in some respects analogous to a Bäcklund transformation, was already derived through the elaborate work of Kinnersley and Chitre.[3] Therefore, it appeared reasonable to ask whether their program hinges on the existence of a symmetry transformation of the Bäcklund variety. While examining this question, we became aware of a recent paper by Harrison[5] where a Bäcklund transformation is derived. Nevertheless, we have decided to communicate our observations for they address the connection between the duality rotation, on which the construction of the Kinnersley–Chitre hierarchy is based, and standard Bäcklund transformations.

A simple connection does indeed exist. Appropriately interpreted, the duality rotation provides a Bäcklund transformation for the three fundamental SL(2,R) invariants associated with a metric, and consists of simple algebraic substitutions; while the solution obtained this way for the metric itself is only a special solution associated with the same invariants. Of course, knowledge of the invariants reduces the construction of the full metric to a system of first order linear equations. A stronger result is obtained in a special case, the SL(2,R)$\sim$O(2,1) invariant nonlinear σ model. Exploring the duality rotation, the construction of both invariant and variant fields is reduced to elementary algebraic quadratures. Although such an elementary algorithm is not yet available for the full theory, our considerations suggest its existence.

An important link among the gravitational duality rotation, the Bäcklund transformation and previous experience with similar systems is provided by yet another form of duality pointed out by Maison.[4] The reader who feels uncomfortable with the plethora of duality symmetry transformations should find the present article helpful. Otherwise, we do not intent to provide an extensive list of solutions to Einstein's equations. The explicit examples that we work out are chosen to illustrate the main ideas and do not necessarily describe physically interesting situations—no special effort is made at this stage to comply with reasonable boundary conditions.

We furthermore restrict ourselves to the description of nonstationary fields, e.g., collision of gravitational plane waves or cylindrical waves.[6–8] With the identification of variables explained by Maison in,[4] this reduces to the theory of a real chiral field ϕ^a, $a=1,2,3$ satisfying the equations of motion, written here in two equivalent forms:

$$(\tau\phi\times\phi_\eta)_\xi+(\tau\phi\times\phi_\xi)_\eta=0,\quad \tau_{\eta\xi}=0,$$

$$\phi^2\equiv\phi^a\phi_a=-1 \tag{1.1a}$$

$$\phi_{\eta\xi}+\frac{1}{2\tau}(\tau_\eta\phi_\xi+\tau_\xi\phi_\eta)=(\phi_\eta\cdot\phi_\xi)\phi,$$

$$\tau_{\eta\xi}=0,\quad \phi^2=-1. \tag{1.1b}$$

η and ξ are characteristic real coordinates. Scalar contractions are performed with the metric $(g_{ab})=(-1,1,1)$, whereas the cross product is defined with the usual antisymmetric tensor ϵ^{abc}, while $\epsilon_{abc}=-\epsilon^{abc}$. The global SL(2,$R$) symmetry, and invariance under the conformal transformation $\eta\to f_1(\eta)$ and $\xi\to f_2(\xi)$ are manifest.

In the remainder of this section, we state Maison's result and explain the so-called gravitational duality rotation in a way suitable for our subsequent discussion. Section II analyzes in detail the special system obtained from (1.1) by setting the background field τ equal to unity—the O(2,1) nonlinear σ model. Section III deals with the complete problem. Since our conventions differ slightly from those of,[4] an appendix describes the construction of the metric in terms of the fields employed here.

B. Reduction and duality

A moving orthonormal trihedral is defined as

[a)]Research supported by the National Science Foundation under grant number PHY77-23512.

$$e_1=\phi,\quad e_2=\frac{\sqrt{h_2}\,\phi_\eta+\sqrt{h_1}\,\phi_\xi}{2\sqrt{h_1h_2}\,\cos(\alpha/2)},\quad e_3=\frac{\sqrt{h_2}\,\phi_\eta-\sqrt{h_1}\,\phi_\xi}{2\sqrt{h_1h_2}\,\sin(\alpha/2)}$$

$$h_1\equiv\phi_\eta^2,\quad h_2\equiv\phi_\xi^2,\quad \cos\alpha\equiv\frac{(\phi_\eta\cdot\phi_\xi)}{\sqrt{h_1h_2}}=\frac{\sigma}{\sqrt{h_1h_2}}$$

$$(e_i\cdot e_j)=(g_{ij})=(-1,1,1). \tag{1.2}$$

The equations governing its space–time displacement are derived by using the equation of motion (1.1) and elementary completeness arguments. One obtains

$$\begin{bmatrix} e_1\\ e_2\\ e_3\end{bmatrix}_\eta=\begin{bmatrix}0 & \sqrt{h_1}\cos\frac{\alpha}{2}, & \sqrt{h_1}\sin\frac{\alpha}{2}\\ \sqrt{h_1}\cos\frac{\alpha}{2}, & 0, & \frac{\alpha_n}{2}-\left(\frac{h_1}{h_2}\right)^{1/2}\frac{\tau_\xi}{2\tau}\sin\alpha\\ \sqrt{h_1}\sin\frac{\alpha}{2}, & -\frac{\alpha_n}{2}+\left(\frac{h_1}{h_2}\right)^{1/2}\frac{\tau_\xi}{2\tau}\sin\alpha, & 0\end{bmatrix}\begin{bmatrix} e_1\\ e_2\\ e_3\end{bmatrix} \tag{1.3a}$$

$$\begin{bmatrix} e_1\\ e_?\\ e_2\end{bmatrix}_\xi=\begin{bmatrix}0, & \sqrt{h_2}\cos\frac{\alpha}{2}, & -\sqrt{h_2}\sin\frac{\alpha}{2}\\ \sqrt{h_2}\cos\frac{\alpha}{2}, & 0 & -\frac{\alpha_\xi}{2}+\left(\frac{h_2}{h_1}\right)^{1/2}\frac{\tau_\eta}{2\tau}\sin\alpha\\ -\sqrt{h_2}\sin\frac{\alpha}{2}, & \frac{\alpha_\xi}{2}-\left(\frac{h_2}{h_1}\right)^{1/2}\frac{\tau_\eta}{2\tau}\sin\alpha, & 0\end{bmatrix}\begin{bmatrix} e_1\\ e_2\\ e_3\end{bmatrix}. \tag{1.3b}$$

The equations of motion satisfied by the invariants h_1, h_2 and σ (or α) may be obtained as the integrability condition of Eqs. (1.3a) and (1.3b):

$$(\tau h_1)_\xi+\tau_\eta\sigma=0,\quad (\tau h_2)_\eta+\tau_\xi\sigma=0,$$

$$\alpha_{\eta\xi}=\sqrt{h_1h_2}\,\sin\alpha+\left[\frac{\tau_\xi}{2\tau}\left(\frac{h_1}{h_2}\right)^{1/2}\sin\alpha\right]_\xi+\left[\frac{\tau_\eta}{2\tau}\left(\frac{h_2}{h_1}\right)^{1/2}\sin\alpha\right]_\eta,$$

$$\sigma\equiv\sqrt{h_1h_2}\,\cos\alpha. \tag{1.4}$$

This completes the reduction of the original system (1.1). It may be said that (1.3) is the "linear" problem associated with the "reduced" system (1.4). We should emphasize, however, that this is not yet a genuine Lax problem, for an eigenvalue parameter is missing in Eqs. (1.3). To put it crudely, a theory that possesses an invariance group can always be reduced to a system of equations analogous to (1.3,4); which does not in general lead to a Lax problem. This is possible for a distinct class of nonlinear systems that admit a duality symmetry transformation. Such a duality transformation was first discovered by Pohlmeyer, in the O(N) nonlinear σ model.[9] A class of Fermi interactions[10] and the linear ferromagnetic chain[11] were subsequently shown to possess analogous structure. For the theory under present consideration, the duality transformation was derived by Maison.[4] We first outline its content, and describe the consequences throughout the paper.

The first of Eqs. (1.1a) is written in the form of a local conservation law, which is Noether's current conservation associated with the SL(2,R) invariance. We may "trivialize" this equation by introducing a dual vector field Q^a, through the compatible equations

$$\tau\phi\times\phi_\eta=-Q_\eta,\quad \tau\phi\times\phi_\xi=Q_\xi. \tag{1.5}$$

In turn, the dual field Q satisfies the equation of motion

$$Q_{\eta\xi}+(1/\tau)Q_\eta\times Q_\xi=(1/2\tau)(\tau_\eta Q_\xi+\tau_\xi Q_\eta). \tag{1.6}$$

In the special case $\tau=1$, Eq. (1.6) is nothing but a fancy way of expressing the fact that the current densities appearing in (1.1a), with $\tau=1$, are "pure gauge" fields

$$(\phi\times\phi_\xi)_\eta-(\phi\times\phi_\eta)_\xi-2(\phi\times\phi_\eta)\times(\phi\times\phi_\xi)=0. \tag{1.7}$$

This follows merely from the constraint $\phi^2=-1$. What is interesting here is that the presence of the background free field τ in (1.6) does not disturb the essential consequences expected from a pure gauge condition such as (1.7). Namely, a duality transformation may be derived as follows. Given a solution of the equation of motion, the linear system

$$R_\eta=(1-\gamma_s)(\phi\times\phi_\eta)^a\tau_aR=-(1-\gamma_s)\tau^{-1}Q_\eta^a\tau_aR,$$

$$R_\xi=(1-\gamma_s^{-1})(\phi\times\phi_\xi)^a\tau_aR=(1-\gamma_s^{-1})\tau^{-1}Q_\xi^a\tau_aR,$$

$$\gamma_s \equiv \left[\frac{1+2s(\tau-\tilde{\tau})}{1-2s(\tau+\tilde{\tau})}\right]^{1/2}, \tag{1.8}$$

is a system of compatible equations. Here, τ_a are the generators of SL(2,R), $[\tau_a,\tau_b] = \epsilon_{abc}\tau^c$, and s is an arbitrary real parameter; $\tilde{\tau}$ is conjugate to the free field τ, $\tau_\eta = \tilde{\tau}_\eta$ and $\tau_\xi = -\tilde{\tau}_\xi$. Furthermore, we can normalize R such that $\det R = 1 \rightarrow R \in SL(2,R)$.

Most importantly, the one-parameter family of fields defined from

$$\phi^{(s),a}\tau_a \equiv R^{-1}\phi^a\tau_a R \tag{1.9}$$

satisfy the original equation of motion for arbitrary s, in the background field $\tau^{(s)}$:

$$\tau^{(s)} = \frac{1}{2}\left[\frac{\tau-\tilde{\tau}}{1+2s(\tau-\tilde{\tau})} + \frac{\tau+\tilde{\tau}}{1-2s(\tau+\tilde{\tau})}\right],$$
$$\tau^{(s)}_{\eta\xi} = 0. \tag{1.10}$$

This completes the construction of the duality symmetry transformation. Its consequences will be described in later stages.

C. Gravitational duality

We finally turn to a preliminary description of an alternative form of duality that emerged in earlier discussions of axially symmetric gravitational fields, e.g., see Ref. 2 and references contained there. It will be necessary to clearly distinguish it from the duality discussed in Sec. I B. With no exception, we shall refer to it as the *G duality*. In order to reveal its content, we abandon manifest covariance under SL(2,R) transformations. We thus introduce the Lewis–Papapetrou (L–P) parametrization of the chiral field ϕ^a

$$\phi^1+\phi^2 \equiv f, \quad \phi^1-\phi^2 \equiv \omega^2 f + \frac{1}{f}, \quad \phi^3 \equiv \omega f, \tag{1.11}$$

so that the constraint $\phi^a\phi_a = -1$ becomes automatic. The two independent equations satisfied by the real L–P parameters f and ω are easily derived to be

$$(\tau f^2\omega_\eta)_\xi + (\tau f^2\omega_\xi)_\eta = 0$$
$$[\tau(f^2\omega\omega_\eta - f^{-1}f_\eta)]_\xi + [\tau(f^2\omega\omega_\xi - f^{-1}f_\xi)]_\eta = 0, \tag{1.12}$$

to be supplemented by $\tau_{\eta\xi}=0$. Using the dual potential introduced in (1.5), specifically

$$\Omega \equiv Q^1+Q^2; \quad \Omega_\eta = -\tau f^2\omega_\eta, \quad \Omega_\xi = \tau f^2\omega_\xi, \tag{1.13}$$

an equivalent system for f and Ω may be derived

$$(\tau^{-1}f^{-2}\Omega_\eta)_\xi + (\tau^{-1}f^{-2}\Omega_\xi)_\eta = 0$$
$$(\tau^{-1}f^{-2}\Omega\Omega_\eta + \tau f^{-1}f_\eta)_\xi$$
$$+ (\tau^{-1}f^{-2}\Omega\Omega_\xi + \tau f^{-1}f_\xi)_\eta = 0. \tag{1.14}$$

The crucial observation is that (1.14) follows from (1.12) under the *formal* substitution:

$$f \rightarrow \frac{1}{\tau f}, \quad \omega \rightarrow \Omega, \tag{1.15}$$

and, of course, $\tau_{\eta\xi}=0$. This simply means that a new chiral field F^a obtained from (1.11) under the same formal substitution, namely

$$F^1+F^2 = \frac{1}{\tau f}, \quad F^1-F^2 = \frac{\Omega^2}{\tau f} + \tau f,$$
$$F^3 = \frac{\Omega}{\tau f}, \quad F^aF_a = -1, \tag{1.16}$$

is also a solution of the original equation of motion in the *same* background field τ. This symmetry transformation constitutes what we call G duality.

II. CHIRAL O(2,1) THEORY ($\tau = 1$)

We begin the detailed analysis of the concepts described above, and their relation to Bäcklund transformations, in the special system obtained from (1.1) by setting $\tau = 1$. This is a nonlinear σ model with O(2,1) symmetry. The specialization of the equations given in Sec. I to the present problem is obvious. We shall gradually reproduce them as we proceed with this discussion. The theory of the O(2,1) σ model is fairly analogous to the well studied O(3) model [9,10]. The added element here is the G duality outlined in Sec. I C.

The reduced system (1.4) simplifies to

$$h_{1,\xi} = 0 = h_{2,\eta}, \quad \alpha_{\eta\xi} = \sqrt{h_1h_2}\ \sin\alpha. \tag{2.1}$$

It was already mentioned that the equations of motion are invariant under conformal transformations. We may exploit this freedom to impose appropriate gauge conditions. The most convenient choice is $h_1 = 1 = h_2$. We shall refer to it as the *canonical* parametrization or canonical conformal frame. With this choice, the first two equations in (2.1) become trivial, whereas the second is the well-known sine-Gordon theory.

It will be important to examine the transformation of the fundamental invariants h_1, h_2 and $\sigma = (h_1h_2)^{1/2} \times \cos\alpha$ under the duality transformation (1.8), which now reduces to

$$R_\eta = (1-\gamma)(\phi\times\phi_\eta)^a\tau_a R,$$
$$R_\xi = (1-\gamma^{-1})(\phi\times\phi_\xi)^a\tau_a R \tag{2.2}$$

with γ an arbitrary real constant. Starting from a solution ϕ with invariants h_1, h_2, σ the invariants of the associated *dual ray* $\phi^{(\gamma),a}\tau_a \equiv R^{-1}\phi^a\tau_a R$, are easily calculated to be

$$h_1^{(\gamma)} = \gamma^2 h_1, \quad h_2^{(\gamma)} = \gamma^{-2}h_2, \quad \sigma^{(\gamma)} = \sigma. \tag{2.3}$$

We wish to illustrate this situation in a simple example. We start with the solution:

$$\{\phi^a\} = \{\cosh(\eta+\xi), 0, \sinh(\eta+\xi)\} \tag{2.4}$$

in the canonical frame $h_1 = 1 = h_2$; one also calculates $\sigma = 1$, or $\alpha = 0$. Using (2.4) as input in (2.2), the construction of R is trivial. Not surprisingly, the dual ray associated with (2.4) is found to be

$$\{\phi^{\gamma,a}\} = \{\cosh(\gamma\eta + \gamma^{-1}\xi), 0, \sinh(\gamma\eta + \gamma^{-1}\xi)\}, \tag{2.5}$$

with invariants $h_1^{(\gamma)} = \gamma^2$, $h_2^{(\gamma)} = \gamma^{-2}$, $\sigma^{(\gamma)} = 1$ as expected. This example might in fact lead to false impressions. It appears that the effect of a duality transformation is equivalent to an ordinary constant conformal transformation or even a Lorentz transformation. This is not the case, however. In general, a solution depends on a number of parameters ϵ, ϵ', ⋯, aside from the dependence on the duality parameter γ. A duality transformation amounts to redefining γ, but leaves unchanged $\epsilon, \epsilon', \cdots$; which distinguishes it from conformal or Lorentz transformations that uniformly transform all parameters.

We return to the general case and state the derivation of a Lax problem associated with the reduced system (2.1). Observe that the invariants entering the sine-Gordon equation, $h_1 h_2$ and α, remain unchanged under duality transformations or, equivalently, they do not depend on the duality parameter γ. In particular, $h_1 h_2 = 1$ in the canonical frame as well as in the frame $h_1 = \gamma^2, h_2 = \gamma^{-2}$. The transcription of the "linear" problem (1.3) into a genuine Lax problem should now be obvious. We simply set $\tau = 1$ and $h_1 = \gamma^2$, $h_2 = \gamma^{-2}$ in Eqs. (1.3)

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}_\eta = \begin{bmatrix} 0 & \gamma\cos\frac{\alpha}{2} & \gamma\sin\frac{\alpha}{2} \\ \gamma\cos\frac{\alpha}{2} & 0 & \frac{\alpha_\eta}{2} \\ \gamma\sin\frac{\alpha}{2} & -\frac{\alpha_\eta}{2} & 0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}_\xi = \begin{bmatrix} 0 & \gamma^{-1}\cos\frac{\alpha}{2} & -\gamma^{-1}\sin\frac{\alpha}{2} \\ \gamma^{-1}\cos\frac{\alpha}{2} & 0 & -\frac{\alpha_\xi}{2} \\ -\gamma^{-1}\sin\frac{\alpha}{2} & \frac{\alpha_\xi}{2} & 0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}. \tag{2.6}$$

The integrability condition is $\alpha_{\eta\xi} = \sin\alpha$, for arbitrary γ.

Starting from (2.6) a Bäcklund transformation can be derived along the lines of.[10] We only state the result that can be verified by an elementary, but relatively tedious, computation. Given a solution ϕ^a in the conformal frame $h_1 = \gamma^2, h_2 = \gamma^{-2}$ with invariant $(\phi_\eta \cdot \phi_\xi) = \cos\alpha$, the chiral field F^a:

$$F = \frac{1}{\gamma^2 - \epsilon^2}\left\{(\gamma^2 + \epsilon^2)e_1 + 2\epsilon\gamma\left[\cos\frac{\alpha'}{2}e_2 - \sin\frac{\alpha'}{2}e_3\right]\right\}, \quad F^aF_a = -1$$

$$e_1 \equiv \phi, \quad e_2 \equiv \frac{\gamma^{-1}\phi_\eta + \gamma\phi_\xi}{2\cos(\alpha/2)}, \quad e_3 \equiv \frac{\gamma^{-1}\phi_\eta - \gamma\phi_\xi}{2\sin(\alpha/2)} \tag{2.7a}$$

is also a solution, in the same conformal frame and $(F_\eta \cdot F_\xi) = \cos\alpha'$, provided that α and α' are related by the standard Bäcklund transformation:

$$\left(\frac{\alpha' - \alpha}{2}\right)_\eta = \epsilon\sin\left(\frac{\alpha' + \alpha}{2}\right), \quad \left(\frac{\alpha' + \alpha}{2}\right)_\xi = \frac{1}{\epsilon}\sin\left(\frac{\alpha' - \alpha}{2}\right), \tag{2.7b}$$

with ϵ an arbitrary real constant. The Bäcklund parameter ϵ should be clearly distinguished from the duality parameter γ, even though an interesting interplay between the two will be shown to take place in connection with G-duality transformations. Before turning to this question, we apply (2.7) to the simple solution (2.5). The resulting one soliton solution is written in the form:

$$F^1 \pm F^2 = \frac{1}{\gamma^2 - \epsilon^2}\left[(\gamma^2 + \epsilon^2)\cosh\theta - 2\epsilon\gamma\frac{\sinh\chi\sinh\theta \pm 1}{\cosh\chi}\right]$$

$$F^3 = \frac{1}{\gamma^2 - \epsilon^2}\left[(\gamma^2 + \epsilon^2)\sinh\theta - 2\epsilon\gamma\tanh\chi\cosh\theta\right] \tag{2.8}$$

$$h'_1 \equiv (F_\eta \cdot F_\eta) = \gamma^2, \quad h'_2 \equiv (F_\xi \cdot F_\xi) = \gamma^{-2},$$

$$\sigma' \equiv (F_\eta \cdot F_\xi) = 1 - 2\cosh^{-2}\chi$$

$$\theta \equiv \gamma\eta + \gamma^{-1}\xi, \quad \chi \equiv \epsilon\eta + \epsilon^{-1}\xi,$$

modulo an arbitrary constant O(2,1) rotation (trivial position parameters have also been omitted in the above expression).

The Bäcklund transformation (2.7) reduces the construction of arbitrary multi-soliton solutions to quadratures. We now reveal an intricate relation between this Bäcklund transformation and the G duality described in the end of Sec. I. This will in turn provide an elementary algebraic algorithm that can replace the Riccati equations (2.7b), in the construction of higher soliton solutions. The precise content of G duality was already stated in Sec. I C. We shall only have to specialize the appropriate formulas by setting $\tau = 1$. Hence, starting from a solution ϕ^a characterized by the L–P parameters f and ω, see Eq. (1.11), and defining a dual potential Ω from

$$\Omega_\eta = -f^2\omega_\eta, \quad \Omega_\xi = f^2\omega_\xi, \tag{2.9}$$

a new solution is obtained by a G duality transformation, namely

$$F^1 + F^2 = 1/f, \quad F^1 - F^2 = \Omega^2/f + f, \quad F^3 = \Omega/f. \tag{2.10}$$

It is important to examine the transformation of the fundamental invariants h_1, h_2 and σ under G duality. We first express the invariants associated with the solution ϕ^a in terms of the L–P parameters f and ω

$$\begin{aligned} h_1 &\equiv (\phi_\eta \cdot \phi_\eta) = f^2\omega_\eta^2 + f^{-2}f_\eta^2, \\ h_2 &\equiv (\phi_\xi \cdot \phi_\xi) = f^2\omega_\xi^2 + f^{-2}f_\xi^2 \\ \sigma &\equiv (\phi_\eta \cdot \phi_\xi) = f^2\omega_\eta\omega_\xi + f^{-2}f_\eta f_\xi . \end{aligned} \tag{2.11}$$

The invariants associated with the solution F^a are readily calculated from (2.10)

$$\begin{aligned} h_1' &\equiv (F_\eta \cdot F_\eta) = f^2\omega_\eta^2 + f^{-2}f_\eta^2, \\ h_2' &\equiv (F_\xi \cdot F_\xi) = f^2\omega_\xi^2 + f^{-2}f_\xi^2 \\ \sigma' &\equiv (F_\eta \cdot F_\xi) = -f^2\omega_\eta\omega_\xi + f^{-2}f_\eta f_\xi . \end{aligned} \tag{2.12}$$

Comparing with (2.11), we deduce

$$h_1' = h_1, \quad h_2' = h_2, \quad \sigma' + \sigma = 2f^{-2}f_\eta f_\xi . \tag{2.13}$$

The first interesting observation is that G duality preserves the conformal frame ($h_1' = h_1, h_2' = h_2$), but induces a nontrivial transformation for the invariant σ, a property shared by the Bäcklund transformation (2.7). Secondly, the transformation formula (2.13) for the fundamental invariants does not contain the dual potential Ω, contrary to the situation for variant fields, see Eq. (2.10). In order to appreciate the meaning of the above observations, we find it convenient to continue our demonstration with an explicit example.

As we did for the Bäcklund transformation, we apply the G duality transformation to the simple solution (2.5). We immediately identify the associated L–P parameters

$$f = \cosh\theta, \quad \omega = \tanh\theta, \quad \theta \equiv \gamma\eta + \gamma^{-1}\xi \tag{2.14}$$

and thereby calculate the dual potential Ω from (2.9):

$$\Omega = \gamma^{-1}\xi - \gamma\eta, \tag{2.15}$$

up to a trivial additive constant (such position parameters are systematically set equal to zero). Direct application of (2.10) and (2.13) yields:

$$\begin{aligned} &F^1 + F^2 = \frac{1}{\cosh\theta}, \quad F^1 - F^2 = \frac{\Omega^2}{\cosh\theta} + \cosh\theta, \\ &F^3 = \frac{\Omega}{\cosh\theta}, \quad h_1' = h_1 = \gamma^2, \\ &h_2' = h_2 = \gamma^{-2}, \quad \sigma' = 1 - 2\cosh^{-2}\theta. \end{aligned} \tag{2.16}$$

Comparing with Eqs. (2.8), we immediately realize that the effect of G duality on the invariants is identical to a Bäcklund transformation; whereas it is easily seen that the variant fields in (2.16) are not related to the corresponding fields in (2.8) by a constant O(2,1) rotation. The resolution of this apparent paradox is obtained by the following considerations. Notice that, in contrast to (2.16), the variant fields in (2.8) depend on two essential parameters, the duality parameter γ and the Bäcklund parameter ϵ. However, the invariant σ' depends only on the Bäcklund parameter. Hence, despite the fact that we are dealing with a nonlinear system, expanding the equality (equation of motion) $F_{\eta\xi}(\ ;\epsilon,\gamma) = \sigma'(\ ;\epsilon) \times F(\ ;\epsilon,\gamma)$ around $\gamma = \epsilon$, and keeping the leading term, we obtain a *solution* with the same invariants as $F(\ ;\epsilon;\gamma)$. The only provision is that F should be regular at $\gamma = \epsilon$, so that a γ-independent leading term may be identified. Simple inspection of (2.8) reveals the opposite. Recall, however, that an arbitrary constant O(2,1) rotation can be applied to F that may well depend on the parameters γ and ϵ. Indeed, we found that the following O(1,1) rotation in the 12 plane renders the solution regular at $\gamma = \epsilon$

$$R = \frac{1}{2(\gamma-\epsilon)}\begin{bmatrix} 1+(\gamma-\epsilon)^2 & 1-(\gamma-\epsilon)^2 \\ 1-(\gamma-\epsilon)^2 & 1+(\gamma-\epsilon)^2 \end{bmatrix}. \tag{2.17}$$

Applying it to F given in (2.8), and expanding carefully around $\gamma = \epsilon$, one finds for the leading term:

$$\begin{aligned} &F^1 + F^2 = \frac{1}{2\epsilon}\left[\frac{\Omega^2}{\cosh\chi} + \cosh\chi\right], \\ &F^1 - F^2 = \frac{2\epsilon}{\cosh\chi}, \quad F^3 = \frac{\Omega}{\cosh\chi} \\ &h_1' = \epsilon^2, \quad h_2' = \epsilon^{-2}, \quad \sigma' = 1 - 2\cosh^{-2}\chi \\ &\chi \equiv \epsilon\eta + \epsilon^{-1}\xi, \quad \Omega \equiv \epsilon\eta - \epsilon^{-1}\xi. \end{aligned} \tag{2.18}$$

Direct comparison with (2.16) establishes that this solution is equivalent to the result of G duality, modulo a remaining trivial O(2,1) rotation and the identification $\gamma = \epsilon$.

The general pattern should now be clear. We thus summarize our main conclusions for the chiral O(2,1) theory. Combining the Bäcklund transformation given in (2.7a) and the G-duality transformation for the invariant σ (or α) given by (2.13), we have essentially arrived at a completely algebraic procedure for the construction of multi-soliton solutions. It is straighforward and can be explained in words as follows: A solution ϕ^a in the conformal frame $h_1 = \gamma^2$, $h_2 = \gamma^{-2}$ is characterized equivalently by the L–P parameters f_γ, ω_γ; the subscript indicates the dependence on the duality parameter γ. Using (2.13) with $f = f_{\gamma=\epsilon}$, we construct the once-transformed invariant σ'. Inserting $\sigma' = \cos\alpha'$ into (2.7a), we obtain a new solution again parametrized by the duality parameter γ in addition to the Bäcklund parameter ϵ. After a general constant O(2,1) rotation is performed, we identify new L–P parameters f_γ', ω_γ'; inserting $f_{\gamma=\epsilon}'$ into (2.13) we obtain a new potential σ'', etc. We thus obtain a hierarchy of solutions all parametrized by the duality parameter γ, in addition to a sequence of Bäcklund parameters $\epsilon, \epsilon', \epsilon'' \cdots$, and a set of position parameters introduced through the constant O(2,1) rotations, performed at each stage of the iteration. In such a scheme, explicit solution of the Riccati equations (2.7b) is avoided.

It may seem surprising that a manifestly O(2,1) covariant Bäcklund transformation such as (2.7) is intimately related to the "noncovariant" prescription of G duality. We note, however, that alternative formulations of the Bäcklund transformation in terms of a pseudo-potential[5] also abandon manifest covariance. In fact, it appears possible to express the logarithmic derivatives $f^{-1}f_\eta$ and $f^{-1}f_\xi$ in (2.13) in terms of a pseudo-potential satisfying appropriate Riccati equations. We shall not elaborate on this point, as it will not be essential for our subsequent considerations.

We conclude this section with some comments on the important qualitative differences between the duality transformation (2.2,3) and the Bäcklund transformation (2.7a,b). They can be best understood in the context of the inverse problem deriving from Eqs. (2.6), see Ref. 10. Duality interrelates the Jost functions to one another, while it leaves the potential unchanged. In contrast, a Bäcklund transformation induces a nontrivial change on the potential by adding to it one bound state at each stage of the iteration. Such intriguing structure underlies, in fact, most of the surprising properties of a large class of integrable systems.

III. THE GENERAL CASE

The analog of Eq. (2.7a) is not yet available for the complete ($\tau \neq 1$) theory. This prevents the direct generalization of the elementary Bäcklund algorithm derived for the O(2,1) nonlinear σ model. However, the lesson from the preceeding discussion is that a genuine Bäcklund algorithm may be derived for the fundamental SL(2,R) invariants, exploring the G-duality transformation; thereby, the construction of the corresponding variant fields reduces to the solution of linear equations. We anticipate that the analog of (2.7a) will be obtained in the future, in which case some of the subsequent considerations will be substantially simplified.

Aside from certain technical differences, the spirit of the following construction is similar to the considerations of Sec. II. The conformal freedom may be eliminated by restricting to the canonical frame:

$$\tau = \tau_c = \eta + \xi. \tag{3.1}$$

For instance, a simple solution in the canonical frame is given by

$$\{\phi^a\} = \{\cosh(\eta - \xi), \sinh(\eta - \xi), 0\}. \tag{3.2}$$

Most of the following arguments possess an obvious extension to arbitrary solutions. For concreteness, we discuss for the moment the solution (3.2). Along with it, it will be essential to consider the associated dual ray constructed by performing a duality transformation. The function γ_s entering the duality equation (1.8) is now

$$\gamma_s = \left[\frac{1 + 4s\xi}{1 - 4s\eta}\right]^{1/2}, \tag{3.3}$$

since $\tau = \eta + \xi$, $\tilde{\tau} = \eta - \xi$. For the construction of R associated with (3.2), we use the real basis of SL(2,R): $\tau_1 = -(i/2)\sigma_2$, $\tau_2 = \frac{1}{2}\sigma_1$, $\tau_3 = \frac{1}{2}\sigma_3$, $[\tau_a, \tau_b] = \epsilon_{abc}\tau^c$, with $\epsilon_{123} = -1$. Inserting (3.2) into (1.8) and solving the resulting linear equations, one finds

$$R = \exp\left\{-\frac{1}{2}\left[\eta - \xi - \frac{1}{2s} \times \left(1 - \sqrt{(1 + 4s\xi)(1 - 4s\eta)}\right)\right]\sigma_3\right\}$$

$$\det R = 1, \quad R(s = 0) = 1. \tag{3.4}$$

The dual ray associated with (3.2) is readily calculated to be:

$$\{\phi^{(s),a}\} = \{\cosh\theta_s, \sinh\theta_s, 0\}$$
$$\theta_s \equiv \frac{1}{2s}\left[1 - \sqrt{(1 + 4s\xi)(1 - 4s\eta)}\right], \tag{3.5}$$

a solution in the conformal frame

$$\tau^{(s)} = \frac{\xi}{1 + 4s\xi} + \frac{\eta}{1 - 4s\eta}. \tag{3.6}$$

This solution reduces to (3.2) at $s = 0$, but not under the ordinary conformal transformation $\tau^{(s)} \to \eta + \xi$.

With this concrete example in mind, we return to the general problem. Under duality, the invariants transform according to

$$h_1^{(s)} = \gamma_s^2 h_1, \quad h_2^{(s)} = \gamma_s^{-2} h_2, \quad \sigma^{(s)} = \sigma. \tag{3.7}$$

It is also useful to know that

$$\frac{\tau_\eta^{(s)}}{\tau^{(s)}} = \frac{\gamma_s^2}{\eta + \xi}, \quad \frac{\tau_\xi^{(s)}}{\tau^{(s)}} = \frac{\gamma_s^{-2}}{\eta + \xi}. \tag{3.8}$$

The main difference from the discussion of Sec. II is that h_1 and h_2 satisfy nontrivial differential equations, see Eqs. (1.4). In order to arrive at a Lax problem, it will be essential to parametrize these equations in terms of quantities that are invariant under duality transformations (see Sec. V of Ref. 10 for a similar discussion). This task is in fact straightforward. Direct inspection of (3.7) and (3.8) suggests that an appropriate change of variables in (1.4) is obtained by

$$\tau h_1 \equiv \tau_\eta H_1, \quad \tau h_2 \equiv \tau_\xi H_2, \tag{3.9}$$

where H_1 and H_2 are invariant under duality transformations. We also note that $\tau^{-2}\tau_\eta\tau_\xi$ is invariant so that it assumes its canonical value $(\eta + \xi)^{-2}$. With these qualifications, the reduced system (1.4) reads:

$$H_{1,\xi} + \sigma = 0, \quad H_{2,\eta} + \sigma = 0$$
$$\alpha_{\eta\xi} = \frac{(H_1 H_2)^{1/2}}{\eta + \xi}\sin\alpha + \frac{1}{2}\left(\frac{1}{\eta + \xi}\left(\frac{H_1}{H_2}\right)^{1/2}\sin\alpha\right)_\xi + \frac{1}{2}\left(\frac{1}{\eta + \xi}\left(\frac{H_2}{H_1}\right)^{1/2}\sin\alpha\right)_\eta \tag{3.10}$$
$$\sigma \equiv \frac{(H_1 H_2)^{1/2}}{\eta + \xi}\cos\alpha.$$

Further simplification is achieved by noting that an integrating factor β may be introduced through the compatible equations $H_1 = \beta_\eta$, $H_2 = \beta_\xi$, so that the first order equation in (3.10) may be replaced by a second order equation: $\beta_{\eta\xi} + \sigma = 0$. This fact will not be used in our subsequent discussion.

We must transcribe the "linear" problem (1.3) in the above variables. The prescription is standard. The elements of the orthonormal trihedral are constructed in terms of the dual ray, associated with a solution in the canonical frame, and are functions of the duality parameter s. Accordingly, the invariants appearing in the matrices of (1.3) are identically replaced by

$$h_1 = \frac{\tau_\eta^{(s)}}{\tau^{(s)}} H_1 = \frac{\gamma_s^2}{\eta + \xi} H_1,$$

$$h_2 = \frac{\tau^{(s)}_\xi}{\tau^{(s)}} H_2 = \frac{\gamma_s^{-2}}{\eta+\xi} H_2, \quad \sigma^{(s)} = \sigma. \tag{3.11}$$

H_1, H_2 and σ do not depend on the duality parameter s. For future reference, we also note that the matrices in (1.3a) and (1.3b), denoted by C_1 and C_2 respectively, are real superpositions of the generators of SL(2,R) in the vector representation:

$$C_1 = \omega_1^a I_a, \quad C_2 = \omega_2^a I_a, \quad [I_a, I_b] = \epsilon_{abc} I^c,$$
$$I_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad I_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$
$$I_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.12}$$
$$\epsilon^{123} = -\epsilon_{123} = 1.$$

Using (1.3), (3.11) and (3.12) we write explicitly,

$$\{\omega_1^a\} = \Big\{ -\frac{\alpha_\eta}{2} + \frac{1}{2(\eta+\xi)} \Big(\frac{H_1}{H_2}\Big)^{1/2} \sin\alpha,$$
$$\gamma_s \Big(\frac{H_1}{\eta+\xi}\Big)^{1/2} \sin\frac{\alpha}{2}, \gamma_s \Big(\frac{H_1}{\eta+\xi}\Big)^{1/2} \cos\frac{\alpha}{2} \Big\}$$
$$\tag{3.13}$$
$$\{\omega_2^a\} = \Big\{ \frac{\alpha_\xi}{2} - \frac{1}{2(\eta+\xi)} \Big(\frac{H_2}{H_1}\Big)^{1/2} \sin\alpha,$$
$$-\gamma_s^{-1} \Big(\frac{H_2}{\eta+\xi}\Big)^{1/2} \sin\frac{\alpha}{2}, \gamma_s^{-1} \Big(\frac{H_2}{\eta+\xi}\Big)^{1/2} \cos\frac{\alpha}{2} \Big\}.$$

The integrability condition is now written as

$$\omega^c_{1,\xi} - \omega^c_{2,\eta} + \epsilon^c_{ab} \omega_1^a \omega_2^b = 0, \quad a,b,c = 1,2,3 \tag{3.14}$$

and leads to Eqs. (3.10), for arbitrary value of the duality or *eigenvalue* parameter s. This completes the lengthy, but necessary, preliminaries that will enable us to construct soliton solutions to the generalized sine-Gordon system (3.10), by exploring elementary G-duality transformations.

In analogy with Sec. II, our main assertion is that G duality provides Bäcklund transformations for the invariants H_1, H_2 and σ. Starting from a solution characterized by the L–P parameters f and ω and invariants

$$h_1 = f^2 \omega_\eta^2 + f^{-2} f_\eta^2, \quad h_2 = f^2 \omega_\xi^2 + f^{-2} f_\xi^2,$$
$$\sigma = f^2 \omega_\eta \omega_\xi + f^{-2} f_\eta f_\xi, \tag{3.15}$$

the invariants associated with the transformed solution given by (1.16) are calculated to be

$$h_1' \equiv (F_\eta \cdot F_\eta) = f^2 \omega_\eta^2 + \Big(\frac{\tau_\eta}{\tau} + \frac{f_\eta}{f}\Big)^2$$
$$h_2' \equiv (F_\xi \cdot F_\xi) = f^2 \omega_\xi^2 + \Big(\frac{\tau_\xi}{\tau} + \frac{f_\xi}{f}\Big)^2 \tag{3.16}$$
$$\sigma' \equiv (F_\eta \cdot F_\xi) = -f^2 \omega_\eta \omega_\xi + \Big(\frac{\tau_\eta}{\tau} + \frac{f_\eta}{f}\Big)\Big(\frac{\tau_\xi}{\tau} + \frac{f_\xi}{f}\Big).$$

We noticed earlier that G duality transformations do not change the conformal frame. In other words, they preserve the value of the background free field τ. For this reason, we do not have to specify the choice of the conformal frame at this point, as long as $\tau_{\eta\xi} = 0$.

Combining Eqs. (3.15) and (3.16), we obtain

$$h_1' - h_1 = \frac{\tau_\eta}{\tau}\Big(\frac{\tau_\eta}{\tau} + 2\frac{f_\eta}{f}\Big)$$
$$h_2' - h_2 = \frac{\tau_\xi}{\tau}\Big(\frac{\tau_\xi}{\tau} + 2\frac{f_\xi}{f}\Big) \tag{3.17}$$
$$\sigma' + \sigma = \frac{f_\eta f_\xi}{f^2} + \Big(\frac{\tau_\eta}{\tau} + \frac{f_\eta}{f}\Big)\Big(\frac{\tau_\xi}{\tau} + \frac{f_\xi}{f}\Big).$$

In the light of our preceeding discussion, it will be necessary to rewrite these equations in terms of the duality-invariant variables H_1 and H_2. Using the definition (3.9) and Eqs. (3.17), one finds:

$$H_1' - H_1 = \frac{\tau_\eta}{\tau} + \frac{2f_\eta}{f}$$
$$H_2' - H_2 = \frac{\tau_\xi}{\tau} + \frac{2f_\xi}{f} \tag{3.18}$$
$$\sigma' + \sigma = \frac{f_\eta f_\xi}{f^2} + \Big(\frac{\tau_\eta}{\tau} + \frac{f_\eta}{f}\Big)\Big(\frac{\tau_\xi}{\tau} + \frac{f_\xi}{f}\Big).$$

The finer details of the above formulas will be illustrated in a concrete example. The argument explaining the fact that (3.18) may be used in place of a Bäcklund transformation is perhaps laborious. The reader is urged to keep in mind the analysis of Sec. II, which we will not repeat here. In practice, the application of (3.18) is completely elementary.

In what follows we construct what may be called the one-soliton solution of the generalized sine-Gordon system (3.10). The input is provided by the elementary solution given in (3.2), or more accurately by its dual ray constructed in (3.5). Before employing it into the algorithm, we must perform a general constant SL(2,R) rotation. For our demonstration, we only perform a 90° rotation in the 23-plane, thus suppressing a more or less inessential position parameter that would have been introduced in the general case. We write

$$\{\phi^{(s),a}\} = \{\cosh\theta_s, 0, \sinh\theta_s\}$$
$$\theta_s \equiv \frac{1}{2s}\big[1 - \sqrt{(1+4s\xi)(1-4s\eta)}\,\big]$$
$$\tau = \tau^{(s)} = \frac{\xi}{1+4s\xi} + \frac{\eta}{1-4s\eta}. \tag{3.19}$$

The associated invariants are easily calculated to be:

$$H_1 = \eta + \xi = H_2, \quad \sigma = -1 \tag{3.20}$$

and are, of course, s independent. Accordingly, the relevant L–P parameter f is identified as

$$f_s = \cosh\theta_s. \tag{3.21}$$

Following the instructions of Sec. II, Eqs. (3.18) are used with H_1, H_2 and σ given from (3.20) and $\tau = \tau^{(s=\epsilon)}$, $f = f_{s=\epsilon}$ from (3.19) and (3.21). We thus obtain:

$$H_1' = \eta + \xi + \frac{\gamma_\epsilon^2}{\eta+\xi} + 2\gamma_\epsilon \tanh\theta_\epsilon$$
$$H_2' = \eta + \xi + \frac{\gamma_\epsilon^{-2}}{\eta+\xi} - \frac{2}{\gamma_\epsilon}\tanh\theta_\epsilon \tag{3.22a}$$

$$\sigma' = 1 - 2\tanh^2\theta_\epsilon + \frac{\tanh\theta_\epsilon}{\eta+\xi}\left(\frac{1}{\gamma_\epsilon} - \gamma_\epsilon\right) + \frac{1}{(\eta+\xi)^2},$$

where

$$\gamma_\epsilon = \left[\frac{1+4\epsilon\xi}{1-4\epsilon\eta}\right]^{1/2},$$

$$\theta_\epsilon = \frac{1}{2\epsilon}\left[1 - \sqrt{(1+4\epsilon\xi)(1-4\epsilon\eta)}\right]. \qquad (3.22b)$$

This completes the construction of the one-soliton solution of Eqs. (3.10), depending on the Bäcklund parameter ϵ.

We mentioned earlier that knowledge of the fundamental invariants reduces the construction of the associated variant fields to the solution of linear equations, which we now formulate (see Refs. 4, 10). The answer is provided by Eqs. (1.3), incorporating the refinements summarized in Eqs. (3.12) and (3.13), and an appropriate transcription in the spinor representation. For the following discussion, it will be somewhat convenient to use the (complex) basis of the SL(2,R) Lie algebra:

$$\tau_1 = \frac{i}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \quad \tau_2 = \frac{1}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},$$

$$\tau_3 = \frac{i}{2}\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \qquad (3.23)$$

whereas a finite SL(2,R) transformation will be represented by

$$R = \begin{bmatrix} A & B\\ B^* & A^*\end{bmatrix}, \quad AA^* - BB^* = 1. \qquad (3.24)$$

Let ψ be a two-component complex spinor satisfying the equations:

$$\psi_\eta = \omega_1^a \tau_a \psi, \quad \psi_\xi = \omega_2^a\tau_a\psi. \qquad (3.25)$$

Using (3.13) and (3.23), we write explicitly

$$\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}_\eta = \frac{i}{2}\begin{bmatrix} -\dfrac{\alpha_\eta}{2} + \dfrac{1}{2(\eta+\xi)}\left(\dfrac{H_1}{H_2}\right)^{1/2}\sin\alpha, & -\gamma_s\left(\dfrac{H_1}{\eta+\xi}\right)^{1/2} e^{i\alpha/2}\\ \gamma_s\left(\dfrac{H_1}{\eta+\xi}\right)^{1/2} e^{-i\alpha/2}, & \dfrac{\alpha_\eta}{2} - \dfrac{1}{2(\eta+\xi)}\left(\dfrac{H_1}{H_2}\right)^{1/2}\sin\alpha\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}, \qquad (3.26a)$$

$$\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}_\xi = \frac{i}{2}\begin{bmatrix} \dfrac{\alpha_\xi}{2} - \dfrac{1}{2(\eta+\xi)}\left(\dfrac{H_2}{H_1}\right)^{1/2}\sin\alpha, & -\gamma_s^{-1}\left(\dfrac{H_2}{\eta+\xi}\right)^{1/2} e^{-i\alpha/2}\\ \gamma_s^{-1}\left(\dfrac{H_2}{\eta+\xi}\right)^{1/2} e^{i\alpha/2}, & -\dfrac{\alpha_\xi}{2} + \dfrac{1}{2(\eta+\xi)}\left(\dfrac{H_2}{H_1}\right)^{1/2}\sin\alpha\end{bmatrix}\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}, \qquad (3.26b)$$

$$\gamma_s = \left[\frac{1+4s\xi}{1-4s\eta}\right]^{1/2}. \qquad (3.26c)$$

Of course, the integrability condition leads to (3.10) for arbitrary s.

It is easy to see that $(\psi_1^*\psi_1 - \psi_2^*\psi_2)_\eta = 0 = (\psi_1^*\psi_1 - \psi_2^*\psi_2)_\xi$. We may, therefore, normalize ψ such that $\psi_1^*\psi_1 - \psi_2^*\psi_2 = 1$. We further observe that if $\psi = (\psi_1, \psi_2)$ is a solution of (3.26), $\bar\psi = (\psi_2^*, \psi_1^*)$ is also a solution. Hence, the matrix

$$R = \begin{bmatrix}\psi_1 & \psi_2^*\\ \psi_2 & \psi_1^*\end{bmatrix} \in \mathrm{SL}(2,R) \qquad (3.27)$$

satisfies the system of Eqs. (3.26), namely $R_\eta = \omega_1^a\tau_a R$, $R_\xi = \omega_2^a\tau_a R$. The elements of the first row of the three-dimensional matrix representation of R may be identified with the components of the chiral field ϕ^a, satisfying the original equations of motion in the background field

$$\tau = \tau^{(s)} = \frac{\xi}{1+4s\xi} + \frac{\eta}{1-4s\eta}. \qquad (3.28)$$

Explicitly, one finds:

$$\{\phi^a\} = \{(\psi_1^*\psi_1 + \psi_2^*\psi_2), i(\psi_1\psi_2 - \psi_1^*\psi_2^*), (\psi_1\psi_2 + \psi_1^*\psi_2^*)\}. \qquad (3.29)$$

In previous applications of the above procedure, the analog of the "linear" system (3.26) was solved by inverse scattering methods. This is prevented here by a number of technical difficulties. Instead, the system (3.26) is linearized with the aid of the G duality transformation summarized in Eqs. (3.18). For instance, the one-soliton solution of the generalized sine-Gordon system given in (3.22) may be used in Eqs. (3.26) in order to arrive at a linear system for the spinor ψ. Solution of this linear system and application of (3.29) yields a chiral field ϕ parametrized by the duality parameter s and the Bäcklund parameter ϵ. This may in turn be used into (3.18) for the construction of a doublet solution of the generalized sine-Gordon system, etc. We do not carry out this integration here, but note that solving linear (or Riccati) equations can lead to potential complications. This brings us back to an earlier remark, namely that the derivation of the analog of Eq. (2.7a) is most desirable. This would reduce the solution of the linear equations (3.26) to elementary algebraic substitutions, see Sec. II.

A number of concluding remarks are in order. One way or another, an "industry" for the construction of exact solutions to Einstein's equations is being established. Whether this would lead to physically interesting new solutions remains to be seen. One thing is certain. Einstein's equations for axially symmetric fields have revealed a remarkable and worth exploring structure. The present discussion would have achieved its purpose if it contributed to a better understanding of the relative importance of the various forms of duality and Bäcklund transformations encountered in this problem. We may briefly summarize our main conclusion by saying that the gravitational duality rotation is nothing but a special form of a Bäcklund transformation. We have certainly been cavalier concerning boundary conditions; we can hardly claim at this point that the series of soliton solutions indicated here possess special significance. What might be of some general importance is the observation that all solutions seem to develop singularities, as discussed earlier by Szekeres[6] and Kahn and Penrose[7] in the context of special solutions. Technically, the origin of these singularities is here the special analytic form of the function γ_s entering the duality trnsformation.

ACKNOWLEDGMENT

I wish to thank the referee for suggesting the addition of some clarifications to the original text.

APPENDIX: COMPLETE METRIC

We briefly describe the construction of the complete axially symmetric metric, associated with a chiral field ϕ^a solving the nonlinear system (1.1). Its general form is

$$ds^2 = e^{2\Gamma(t,x)}(dt^2 - dx^2) - f_{AB}(t,x)dx^A dx^B$$
$$A,B = 2,3; \quad x^2 = y \quad x^3 = z. \tag{A1}$$

The coordinates t and x are related to the characteristic coordinates employed in the text by

$$\eta = \frac{t+x}{2}, \quad \xi = \frac{t-x}{2}, \tag{A2}$$

whereas the symmetric matrix (f_{AB}) is obtained from the background field τ and the chiral field ϕ^a, or the L–P parameters f and ω, through the relations

$$(f_{AB}) = \tau\begin{bmatrix} -(\phi^1+\phi^2), & \phi_3 \\ \phi^3, & -(\phi^1-\phi^2), \end{bmatrix} = \tau\begin{bmatrix} -f, & \omega f \\ \omega f, & -\omega^2 f - (1/f) \end{bmatrix}. \tag{A3}$$

Finally, given a solution ϕ^a in the canonical frame $\tau = \eta + \xi$, the construction of the function $\Gamma(t,x)$ in (A1) is effected by solving the linear equations

$$\Gamma_{\eta\xi} + \tfrac{1}{4}(\phi_\eta \cdot \phi_\xi) = \frac{1}{4(\eta+\xi)^2}$$
$$\Gamma_\eta = \frac{\eta+\xi}{4}(\phi_\eta\cdot\phi_\eta) - \frac{1}{4(\eta+\xi)}$$
$$\Gamma_\xi = \frac{\eta+\xi}{4}(\phi_\xi\cdot\phi_\xi) - \frac{1}{4(\eta+\xi)}, \tag{A4}$$

which are compatible by virtue of Eqs. (1.1), with $\tau = \eta + \xi$.

As an illustration, we present in this notation the familiar Kasner solution and work out an immediate generalization provided by the duality transformation. With p an arbitrary real constant, the Kasner solution reads ($\tau = \eta + \xi = t$):

$$f = -(\eta+\xi)^p = -t^p, \quad \omega = 0 \tag{A5}$$

or

$$\{\phi^a\} = \{-\tfrac{1}{2}(t^p + t^{-p}), \quad -\tfrac{1}{2}(t^p - t^{-p}), 0\}. \tag{A6}$$

The associated invariants are readily found to be

$$H_1 = H_2 = \frac{p^2}{\eta+\xi}, \quad \sigma = \frac{p^2}{(\eta+\xi)^2}, \quad \cos\alpha = 1. \tag{A7}$$

Furthermore, Eqs. (A4) are trivially solved to yield

$$\Gamma = \frac{p^2-1}{4}\ln(\eta+\xi) = \frac{p^2-1}{4}\ln t. \tag{A8}$$

Assembling the above information, the complete metric reads

$$ds^2 = t^{(p^2-1)/2}(dt^2 - dx^2) - t^{1+p}dy^2 - t^{1-p}dz^2. \tag{A9}$$

It can be brought into its standard form

$$ds^2 = dt^2 - t^{2p_1}dx^2 - t^{2p_2}dy^2 - t^{2p_3}dz^2,$$
$$p_1 = \frac{p^2-1}{p^2+3}, \quad p_2 = \frac{2(1+p)}{p^2+3}, \quad p_3 = \frac{2(1-p)}{p^2+3}, \tag{A10}$$

by a more or less obvious coordinate transformation.

We have worked out the dual ray associated with the Kasner solution, see Sec. III for a similar calculation. In terms of L–P parameters,

$$f_s = -\left[\frac{1}{2s}\frac{\sqrt{1+4s\xi} - \sqrt{1-4s\eta}}{\sqrt{1+4s\xi} + \sqrt{1-4s\eta}}\right]^p, \quad \omega = 0, \tag{A11a}$$

$$\tau = \tau^{(s)} = \frac{\xi}{1+4s\xi} + \frac{\eta}{1-4s\eta}. \tag{A11b}$$

It is found to possess an interesting cyclic property. Under the ordinary conformal transformation $\tau^{(s)} \to \eta + \xi$, required to express the solution in the canonical frame, f_s transforms simply as $f_s \to f_{-s}$. Since the duality parameter s is not required to be positive, it can be said that (A11a) is a solution in both the canonical frame $\tau = \eta + \xi$ *and* the conformal frame $\tau = \tau^{(s)}$.

Further generalizations of the Kasner metric can be achieved by performing Bäcklund (G duality) transformations in the manner described in the main text.

[1]R. Geroch, J. Math. Phys. **13**, 394 (1972).
[2]W. Kinnersley, J. Math. Phys. **18**, 1529 (1977).
[3]W. Kinnersley and D.M. Chitre, Phys. Rev. Lett. **40**, 1608 (1978); additional references may be found here.
[4]D. Maison, Phys. Rev. Lett. **41**, 521 (1978); and Max–Planck preprints (1978).
[5]B.K. Harrison, Phys. Rev. Lett. **41**, 1197 (1978).
[6]P. Szekeres, Nature **228**, 1183 (1970); J. Math. Phys. **13**, 286 (1972).
[7]K. Kahn and R. Penrose,, Nature **229**, 185 (1971).
[8]P. Jordan *et al.*, Akad. Wiss. Lit. Mainz Abh. Math.–Natur. Kl. No. 2 (1960).
[9]K. Pohlmeyer, Commun. Math. Phys. **46**, 207 (1976); related work was independently done by F. Lund and T. Regge, Phys. Rev. D **14**, 1524 (1976); F. Lund, Phys. Rev. D **15**, 1540 (1977). See also M. Lüscher and K. Pohlmeyer, Desy preprint (1977), to be published.
[10]A. Neveu and N. Papanicolaou, Commun. Math. Phys. **58**, 31 (1978).
[11]L.A. Takhtajan, Phys. Lett. A **64**, 235 (1977); Takhtajan's "linear" problem was shown to lead to a duality transformation by A. Jevicki and N. Papanicolaou, Ann. Phys. (N.Y.) (to appear).

Self-graviting fluids with cylindrical symmetry. II

P. S. Letelier
Departamento de Física, Universidade de Brasília 70 910, Brasília, Brazil

(Received 9 April 1979; accepted for publication 4 May 1979)

The Einstein field equations for an irrotational perfect fluid with pressure p, equal to energy density ρ are studied when the space–time has cylindrical symmetry with no reflection symmetry. The coordinate transformation to comoving coordinates is discussed. The energy and the Hawking–Penrose inequalities are studied. Particular classes of solutions are exhibited.

1. INTRODUCTION

In a recent paper[1] we studied the Einstein field equations for an irrotational perfect fluid with pressure p equal to rest energy ρ in a cylindrically symmetric space–time. The metric that we considered had reflection symmetry, i.e., invariance under the transformation $z \to -z$.

The purpose of this paper is to study the Einstein field equations for the same type of fluids, but now the metric is taken as

$$ds^2 = e^{2(\omega-\lambda)}(dt^2 - dr^2) - (\psi^2 e^{2\lambda} + r^2 e^{-2\lambda})d\theta^2 - 2\psi e^{2\lambda}\, d\theta\, dz - e^{2\lambda}\, dz^2, \quad (1.1)$$

with ω, λ, and ψ functions of t and r only, i.e., a cylindrically symmetric metric without reflection symmetry. Note that when $\psi = 0$ the metric (1.1) reduces to a particular case of the metric studied in Ref. 1. The metric (1.1) has been used to study gravitational waves with two degrees of freedom.[2]

The Einstein field equations for an irrotational perfect fluid with equation of state $p = \rho$ are equivalent to the Einstein field equations coupled to a massless scalar field,[3] i.e.,

$$R_{ab} = -2\sigma_{,a}\sigma_{,b}, \quad (1.2a)$$

$$\Box\sigma = [(-g)^{1/2}\sigma_{,a} g^{ab}]_{,b}/(-g)^{1/2} = 0. \quad (1.2b)$$

The units are so chosen that we have for the velocity of light $c = 1$ and Newton's constant of gravitation $G = 1/8\pi$. A comma means partial derivative with respect to the index.

The pressure p, the 4-velocity u_a and the energy–momentum tensor T_{ab} are related to σ by[3]

$$p = \rho = \sigma_{,a}\sigma^{,a}, \quad (1.3)$$

$$u_a = \sigma_{,a}/\sigma_{,b}\sigma^{,b}, \quad (1.4)$$

$$T_{ab} = 2\sigma_{,a}\sigma_{,b} - g_{ab}\sigma_{,c}\sigma^{,c} \quad (1.5)$$

In Sec. 2 we study Eqs. (1.2) when the metric is given by (1.1). In Sec. 3 the coordinate transformation that enables us to write the equations in comoving coordinates is presented. In Sec. 4, the energy and the Hawking–Penrose inequalities[4] are examined. In Sec. 5 we exhibit particular solutions to the field equations.

2. THE FIELD EQUATIONS

The field equations (1.2) and the pressure (1.3) when the metric is (1.1) reduce to,

$$\omega_{00} - \omega_{11} - \omega_1/r - \lambda_{00} + \lambda_{11} + \lambda_1/r + 2\lambda_0^2 + (e^{4\lambda}/2r^2)\psi_0^2 = -2\sigma_0^2, \quad (2.1)$$

$$-\omega_{00} + \omega_{11} - \omega_1/r + \lambda_{00} - \lambda_{11} - \lambda_1/r + 2\lambda_1^2 + (e^{4\lambda}/2r^2)\psi_1^2 = -2\sigma_1^2, \quad (2.2)$$

$$-\omega_0/r + 2\lambda_0\lambda_1 + (e^{4\lambda}/2r^2)\psi_0\psi_1 = -2\sigma_0\sigma_1, \quad (2.3)$$

$$\lambda_{00} - \lambda_{11} - \lambda_1/r - (e^{4\lambda}/2r^2)(\psi_0^2 - \psi_1^2) = 0, \quad (2.4)$$

$$-\psi_{00} + \psi_{11} - \psi_1/r - 4(\lambda_0\psi_0 - \lambda_1\psi_1) = 0, \quad (2.5)$$

$$\sigma_{00} - \sigma_{11} - \sigma_1/r = 0, \quad (2.6)$$

$$p = \rho = e^{-2(\omega-\lambda)}(\sigma_0^2 - \sigma_1^2), \quad (2.7)$$

where the indices 0 and 1 mean derivatives with respect to t and r, respectively, the comma is omitted for brevity.

The system of Eq. (2.1), (2.2), and (2.3) is equivalent to

$$\omega_0 = 2r(\lambda_0\lambda_1 + \sigma_0\sigma_1) + (e^{4\lambda}/2r)\psi_0\psi_1, \quad (2.8a)$$

$$\omega_1 = r(\lambda_0^2 + \lambda_1^2 + \sigma_0^2 + \sigma_1^2) + (e^{4\lambda}/4r)(\psi_0^2 + \psi_1^2), \quad (2.8b)$$

$$\omega_{00} - \omega_{11} - \lambda_{00} + \lambda_{11} + \lambda_1/r + \lambda_0^2 + \sigma_0^2 - \lambda_1^2 - \sigma_1^2 + (e^{4\lambda}/4r^2)(\psi_0^2 - \psi_1^2) = 0. \quad (2.9)$$

Equation (2.9) follows from the other field equations. Equations (2.4) and (2.5) are equivalent to

$$\lambda_{00} - \lambda_{11} - \lambda_1/r = (e^{4\lambda}/2r^2)(\psi_0^2 - \psi_1^2), \quad (2.10)$$

$$\psi_{00} - \psi_{11} - \psi_1/r = -2\psi_1/r - 4(\lambda_0\psi_0 - \lambda_1\psi_1). \quad (2.11)$$

Note that the integrability condition for ω, $\omega_{01} = \omega_{10}$, are exactly Eqs. (2.10), (2.11), and (2.6). Thus, in the present case, the solution to Einstein equations reduces to finding a solution to the coupled nonlinear differential equations (2.10)–(2.11) and the solution of the usual cylindrical wave equation (2.6) (that is well known). We get ω from (2.8) by quadrature.

In (2.10) and (2.11) it does not appear σ, so each particular solution λ and ψ of the system of Eq. (2.10)–(2.11) generates a class of functions ω given by

$$\omega[\sigma,\lambda,\psi] = \Omega[\sigma] + \Sigma[\lambda,\psi], \quad (2.12a)$$

where

$$\Omega[\sigma] = \int r[2\sigma_0\sigma_1\, dt + (\sigma_0^2 + \sigma_1^2)\, dr] \quad (2.12b)$$

$$\Sigma[\lambda,\psi] = \Omega[\lambda] + \int (e^{4\lambda}/4r)[2\psi_0\psi_1\, dt + (\psi_0^2 + \psi_1^2)\, dr] \quad (2.12c)$$

and σ is a solution to Eq. (2.6).

3. COMOVING COORDINATES

It can be easily verified that the coordinates R defined by

$$dR(t,r) = r(\sigma_0\, dr + \sigma_1\, dt) \tag{3.1}$$

and $T = \sigma(t,r)$ transform the 4-velocity u^a to $u^{a'} = [(\sigma_{,a}\sigma^{a})^{1/2},0,0,0]$; therefore R and T are comoving coordinates. Equation (2.6) guarantees that the differential that defines R is exact.

The Jacobian of the transformation to comoving coordinates is

$$J\left(\frac{T,R,\theta,z}{t,r,\theta,z}\right) = r(\sigma_0^2 - \sigma_1^2) \tag{3.2}$$

which vanishes where $p = \rho = 0$ in the nonsingular region of the space–time.

The line element in comoving coordinates is

$$ds^2 = [(e^{2(\omega-\lambda)}/(\sigma_0^2-\sigma_1^2)](dT^2 - dR^2/r^2) - (\psi^2 e^{2\lambda} + r^2 e^{-2\lambda})d\theta^2 - 2\psi e^{2\lambda}\, d\theta\, dz - e^{2\lambda}dz^2, \tag{3.3}$$

this line element has a singularity at $r = 0$.

The field equations in comoving coordinates are

$$\dot\omega = r(\sigma_1\dot\lambda^2 + r^2\sigma_1\lambda'^2 + 2r\sigma_0\dot\lambda\lambda' + \sigma_1) + (e^{4\lambda}/4r)(\sigma_1\dot\psi^2 + r^2\sigma_1\psi'^2 + 2r\sigma_0\dot\psi\psi'), \tag{3.4}$$

$$\omega' = \sigma_0\dot\lambda^2 + r^2\sigma_0\lambda'^2 + 2r\sigma_1\dot\lambda\lambda' + \sigma_0 + (e^{4\lambda}/4r^2)(\sigma_0\dot\psi^2 + r^2\sigma_0\psi'^2 + 2r\sigma_1\dot\psi\psi'), \tag{3.5}$$

$$\ddot\lambda - r^2\lambda'' - \frac{2\sigma_0}{\sigma_0^2-\sigma_1^2}\lambda' = \frac{e^{4\lambda}}{2r^2}(\dot\psi^2 - r^2\psi'^2), \tag{3.6}$$

$$\ddot\psi - r^2\psi'' + \frac{2\sigma_1}{r(\sigma_0^2-\sigma_1^2)}\dot\psi = -4(\dot\lambda\dot\psi - r^2\lambda'\psi'), \tag{3.7}$$

where $\dot{(\ )}$ and $(\ )'$ denote derivatives with respect to T and R, respectively. To solve Eqs. (3.6) and (3.7) we must know σ and the transformation $t = t(R,T)$ and $r = r(R,T)$; these transformations, in principle, are known when σ is given.

4. REALITY CONDITIONS

For irrotational perfect fluids with equation of state, $p = \rho$, the energy condition $T_{ab}U^aU^b \geqslant 0$ and the Hawking–Penrose condition[4] $(T_{ab} - 1/2\, T g_{ab})U^aU^b \geqslant 0$ tell us the same, that

$$\rho = e^{-2(\omega-\lambda)}(\sigma_0^2 - \sigma_1^2) \geqslant 0 \tag{4.1}$$

Note that when $\sigma_0^2 < \sigma_1^2$ the metric does not have necessarily a pathological behavior, but in this case it does not represent the metric of a perfect fluid. When $\sigma_0^2 < \sigma_1^2$ the metric (3.3) can be generated by the energy-momentum tensor

$$T^{ab} = -\sigma_{,c}\sigma^{,c}\left[\underset{1}{l}^a\underset{1}{l}^b + \underset{0}{l}^a\underset{0}{l}^b - \underset{2}{l}^a\underset{2}{l}^b - \underset{3}{l}^a\underset{3}{l}^b\right], \tag{4.2}$$

where

$$\underset{0}{l}^a = (\sigma_{,b}\sigma^{,b})^{-1/2}\delta^a_T, \tag{4.3a}$$

$$\underset{1}{l}^a = (\sigma_{,b}\sigma^{,b})^{-1/2}r^{-1}\delta^a_R, \tag{4.3b}$$

$$\underset{2}{l}^a = (\delta^a_\theta + \psi\delta^a_z)e^{\lambda}/r, \tag{4.3c}$$

$$\underset{3}{l}^a = e^{-\lambda}\delta^a_z. \tag{4.3d}$$

The last statement can be easily proved using (1.5) and the fact that when $\sigma_0^2 < \sigma_1^2$, (4.3) is an orthonormal tetrad associated to the metric (3.3). Note that (4.2) is the stress–energy tensor of an anisotropic fluid with positive rest energy density $-\sigma_{,c}\sigma^{,a}$ and vanishing heat flow. In this case both reality conditions are satisfied.

5. PARTICULAR CASES

In this section we shall study different solutions to the system of Eqs. (2.10)–(2.11).

Let us assume

$$\lambda_{00} - \lambda_{11} - \lambda_1/r = 0, \tag{5.1}$$

then (2.10) tells us that ψ is an arbitrary function of either $t - r$ or $t + r$, i.e., $\psi = \psi(t \pm r)$. From (5.1) and (2.11) we get

$$\lambda = (1/4)\ln r + (1/4)\ln 4b, \tag{5.2}$$

where b is an integration constant. Now ω is given by

$$\omega = -(1/16)\ln r + 2b\int(\psi')^2\, d(\mp t + r) + \Omega[\sigma], \tag{5.3}$$

where now the prime means derivative with respect to the argument. This solution, when $\sigma = 0$, is studied in Ref. 2. Note that this solution represents the "superposition" of a gravitational wave, a cylindrical self-gravitating fluid and a wire without reflection symmetry located at $r = 0$.

Let us assume

$$\lambda_1 = \psi_0 = 0, \tag{5.4}$$

thus Eq. (2.11) gives us

$$\psi = (1/2)ar^2, \tag{5.5}$$

where a is an integration constant. We have omitted in (5.5) an additive constant, because it can always be eliminated by a coordinate transformation of the type $z \to z + c\theta$.

From (5.5) and (2.10) we find

$$2\lambda_{00} = -a^2e^{4\lambda}. \tag{5.6a}$$

Equation (5.6a) is equivalent to

$$e^{-2\lambda}(e^{-2\lambda})_{,00} - [(e^{-2\lambda})_{,0}]^2 = a^2. \tag{5.6b}$$

This equation admits the real solution

$$e^{-2\lambda} = b\cosh(ct), \tag{5.7a}$$

where b and c are constants related by

$$b^2c^2 = a^2. \tag{5.7b}$$

From (5.5), (5.7), and (2.12) we obtain

$$\omega(1/8)c^2r^2 + \Omega[\sigma]. \tag{5.8}$$

The pressure in this case is

$$p=\rho=b\,[\cosh(ct)]^{-1}(\sigma_0^2-\sigma_1^2)\exp(-2\Omega-c^2r^2/4). \tag{5.9}$$

If we take the particular solution to (2.6) given by $\sigma=kt$, where k is a constant, the pressure takes the appealing form,

$$p=\rho=bk^2[\cosh(ct)]^{-1}\exp\{-[(k^2+c^2/4)r^2]\}. \tag{5.10}$$

Note that in this case r is related to the comoving coordinate R by $R=(1/2)kr^2$. Thus, the pressure is not singular and goes to zero when either $T=kt$ or R goes to infinity. But the metric (3.3) has a singularity at $r=0$.

Let us assume

$$\psi_0=\lambda_0=0, \tag{5.11}$$

thus Eq. (2.11) gives us

$$\psi_1=ar\,e^{-4\lambda}, \tag{5.12}$$

and Eq. (2.10) reduces to

$$\lambda_{11}+\lambda_1/r=1/2\,a^2e^{-4\lambda}. \tag{5.13a}$$

Equation (5.13a) is equivalent to

$$e^{2\lambda}(e^{2\lambda})_{,11}-[(e^{2\lambda})_{,1}]^2+e^{2\lambda}(e^{2\lambda})_{,1}/r=a^2. \tag{5.13b}$$

This equation admits the real solution

$$e^{2\lambda}=b/2[(cr)^2+1], \tag{5.14}$$

where the constants a, b, and c are related by Eq. (5.7b). From (5.14) and (5.12) we obtain

$$\psi=-\frac{2}{a[1+(cr)^2]}. \tag{5.15}$$

Now ω is given by

$$\omega=1/2\ln[1+(cr)^2]+\Omega[\sigma]. \tag{5.16}$$

The pressure takes the particular simple form

$$p=(1/2)b\,e^{-2\Omega}(\sigma_0^2-\sigma_1^2). \tag{5.17}$$

Note that (5.17) can be obtained from (5.9) doing the following replacements $b\to b/2$ and $c\to 0$. Thus, the remarks that we made about (5.10) also apply in this case, except the one about the limit $T\to\infty$.

[1]P.S. Letelier, J. Math. Phys. **16**, 1488 (1975); see also D. Ray, J. Math. Phys. **17**, 1171 (1976); this paper rectifies an oversight in the first paper. Particular cases of the solution discussed in the first paper can be found in P. S. Letelier and R. Tabensky, J. Math. Phys. **16**, 8 (1975); Nuovo Cimento **28** B, 407 (1975).

[2]J. Stachel, J. Math. Phys. **7**, 1321 (1966) and references therein. See also R.A. d'Inverno and J. Stachel, J. Math. Phys. **19**, 2447 (1978).

[3]R. Tabensky and A.H. Taub, Commun. Math. Phys. **29**, 61 (1973). See also Ref. 1.

[4]S.W. Hawking and R. Penrose, Proc. R. Soc. London, Ser. **314**, 529 (1970).

Gaussian quantum stochastic processes on the CCR algebra

Göran Lindblad
Department of Theoretical Physics, Royal Institute of Technology, S–100 44 Stockholm, Sweden

(Received 22 February 1979; accepted for publication 11 May 1979)

We define a stationary Gaussian quantum stochastic process (GQSP) on the C^*-algebra of the canonical commutation relations over a real symplectic Hilbert space. Physically GQSPs can describe diffusion with non-Markovian memory effects in quantum harmonic oscillators of arbitrary dimension. We find that in analogy with the commutative case a GQSP is completely specified by a operator-valued autocovariance function satisfying certain positive definiteness and reality conditions. The autocovariance function also determines the response of the system to a class of time-dependent generalized forces, and it has a spectral representation in terms of a positive operator-valued measure.

I. INTRODUCTION

In an earlier paper[1] a formalism for (generally) non-Markovian quantum stochastic processes (QSPs) was introduced, providing a general description of irreversible phenomena in quantum systems which included memory effects. Such QSPs were defined by the operationally measurable correlations or, equivalently, in a mathematically more perspicuous fashion by a set of completely positive (CP) maps on tensor products of the algebra of observables [which was taken to be $B(\mathscr{H})$, $\mathscr{H}$ a separable Hilbert space].

In this paper we consider an especially simple class of QSPs defined on the C^*-algebra of the canonical commutation relations (CCR). We call them Gaussian QSPs (GQSPs) in order to stress the analogy with the commutative case. Physically GQSPs can describe diffusion with memory effects in a set of harmonically coupled quantum harmonic oscillators. The formalism defined in Ref. 1 does not immediately apply to the CCR algebra. The simple structure of GQSPs, however, makes it easy to develop the formalism necessary to deal with this case. This is done in the Appendix.

In Sec. 2 we first recall the definition of Gaussian (quasifree) CP (GCP) maps on the CCR algebra over a real Hilbert space provided with a symplectic form. The GQSPs are defined in Sec. 3 via GCP maps on tensor products of CCR algebras, and the corresponding correlation operator form is also derived. It is assumed here that an invariant Gaussian (quasifree) state exists. It is shown in Sec. 4 that a stationary GQSP is uniquely defined by an operator-valued autocovariance function which satisfies certain positive definiteness and reality conditions. A normalized autocorrelation function is also introduced. In Sec. 5 is derived a spectral representation of the autocorrelation function in terms of a positive operator-valued measure which satisfies a reality condition. It is also shown that a large class of positive operator-valued measures satisfying this condition actually exists. The results of Secs. 4 and 5 thus nicely rephrase the classical theorems for Gaussian processes in the noncommutative language. In addition it is shown in Sec. 4 that the autocovariance function defines the response of the system to a certain class of time-dependent generalized forces. This is a property not shared by classical Gaussian SPs without some additional hypothesis.

It should be pointed out that the Markovian case (where the dynamics is given by a semigroup of GCP maps) has been treated elsewhere.[2-4] In that case some greater generality is achieved as the existence of an invariant state is not assumed. On the other hand the autocorrelation function and its spectral representation was not derived.

2. GAUSSIAN COMPLETELY POSITIVE MAPS ON THE CCR ALGEBRA

Let H be a separable real Hilbert space with complexification $\mathscr{H}$, $\sigma(x,y)$ a bounded real symplectic form on H. The CCR algebra $\overline{\Delta(H,\sigma)}$ is the closure in a suitable C^*-norm of the *-algebra $\Delta(H,\sigma)$ which is generated by elements $W(x)$, $x\in H$, satisfying the Weyl relations[5]

$$W(x)W(y)=\exp[-i\sigma(x,y)]W(x+y),$$
$$W(x)^+=W(-x).$$

For simplicity we write

$$\mathscr{A}^0=\Delta(H,\sigma),\quad \mathscr{A}=(\mathscr{A}^0)^-=\overline{\Delta(H,\sigma)}.$$

We can consider $\mathscr{A}$ to be a concrete C^*-algebra of bounded operators in a Hilbert space $\mathscr{K}$. The C^*-algebra

$$\mathscr{A}_n=\overline{\Delta(H_n,\sigma_n)}$$
$$H_n=\overset{n}{\oplus}H,\quad \sigma_n(\mathbf{x},\mathbf{y})=\sum_{i=1}^{n}\sigma(x_i,y_i)$$

is then the n-fold C^* tensor product[6]

$$\mathscr{A}_n=\overset{n}{\otimes}\mathscr{A}\subset B(\mathscr{K}_n),\quad \mathscr{K}_n=\overset{n}{\otimes}\mathscr{K}.$$

We note that in this case the spatial tensor product is the unique C^* tensor product (Ref. 7 Theor. 10.10) and that

$$\mathscr{A}_n=(\mathscr{A}_n^0)^-,\quad \mathscr{A}_n^0=\overset{n}{\otimes}\mathscr{A}^0.$$

We recall that a positive linear functional on $\mathscr{A}^0$ extends uniquely to a positive linear functional on $\mathscr{A}$ and that a completely positive (CP) map of $\mathscr{A}^0$ into itself extends uniquely to a CP map on $\mathscr{A}$.[5,8] We call the state ρ on $\mathscr{A}$ Gaussian (quasifree) if

$$\rho(W(x))=\exp[g(x)],\qquad (2.1)$$

where g is a bounded quadratic form with $g(0)=0$.[9] The

 0022-2488/79/102081-07$01.00

positivity of ρ is equivalent to the positive definiteness of

$$\exp[g(x-y)+i\sigma(x,y)]. \tag{2.2}$$

A map $T\in B(\mathscr{A})$ is Gaussian CP (GCP) if

$$T[W(x)]=W(Ax)\exp[f(x)], \tag{2.3}$$

where $A\in B(H)$ and f is a bounded quadratic form such that

$$\exp[f(x-y)+i\sigma(x,y)-i\sigma(Ax,Ay)] \tag{2.4}$$

is positive definite.[7,8,10] When T is unit preserving, then $f(0)=0$. There is a Gaussian state ρ invariant under T:

$$\rho(T[W(x)])=\rho(W(x))=\exp[g(x)]$$

iff f is of the form

$$f(x)=g(x)-g(Ax). \tag{2.5}$$

We will always assume the existence of such an invariant state from now on. By gauge transformations we can remove the inessential linear terms in f and g, and we will consequently take f and g to be homogenous quadratic forms.

The positive definiteness of (2.2) is equivalent to the conditional positive definiteness of

$$L(x,y)=g(x-y)+i\sigma(x,y)$$

(see Ref. 11) and hence to the positive definiteness of

$$\begin{aligned}K(x,y)&=L(x,y)+L(0,0)-L(x,0)-L(0,y)\\&=g(x-y)-g(x)-g(y)+i\sigma(x,y).\end{aligned} \tag{2.6}$$

We can extend the bilinear forms on H to sesquilinear forms on $\mathscr{H}$ and write

$$\begin{aligned}&\sigma(x,y)=(x,Sy), && g(x-y)-g(x)-g(y)=(x,Gy),\\&K(x,y)=(x,Ky), && g(x)=-\tfrac{1}{2}(x,Gx),\\&x,y\in\mathscr{H}, && G,K,S\in B(\mathscr{H}).\end{aligned}$$

The positive definiteness of (2.6) then reads

$$K\equiv G+iS\geqslant 0.$$

In the same way the positive definiteness of (2.4) gives, using (2.5), the condition for the CP property of T,

$$K-A^{+}KA\geqslant 0.$$

We can introduce the complex conjugation c associated with a real basis spanning H in $\mathscr{H}$. Then

$$G^{c}=G,\quad S^{c}=S,\quad A^{c}=A,\quad K^{c}=G-iS$$

$$G^{+}=G,\quad S^{+}=-S,\quad K^{+}=K.$$

Given a positive K, G and iS may be defined as the symmetric and antisymmetric parts under transposition in this basis.

In Secs. 4 and 5 we will assume that K is nondegenerate and furthermore that K^{-1} is bounded (if this condition is not fulfilled we would have to consider a Gel'fand triplet instead of only H).

3. DEFINITION OF A STATIONARY GAUSSIAN QSP

The definition of a stationary QSP on $B(\mathscr{H})$ ($\mathscr{H}$ a separable Hilbert space) was given in Ref. 1 through what was called a correlation operator representation, and an equivalent tensor product form was derived. From the correlation operators could be calculated the probabilities of events consisting of sets of outcomes of arbitrary sequences of measurements on the system. The measurements were defined by CP map valued measures called instruments.[1,12] In the case of the CCR algebra $\mathscr{A}$, the correlation operators and the probabilities are not so simply related. For this reason we choose to define a stationary GQSP in terms of a tensor product representation. A procedure for defining the correlation operators and the probabilities associated with a certain class of instruments starting from the tensor product form is given in the Appendix. The definitions are chosen to permit a general algebraic treatment not depending on any representation dependent properties. The definitions furthermore coincide with those of Ref. 1 whenever the latter are applicable, e.g., when H is finite dimensional and we consider an irreducible Schrödinger representation in a separable Hilbert space, such that the Gaussian states and the GCP maps are normal.

A GQSP is defined by a set $\{T_n\in\mathrm{GCP}(\mathscr{A}_n)\}$ where T_n is a function of $n+1$ time parameters, i.e., (2.3) reads

$$T_n(t_0,\ldots,t_n)[W(\mathbf{x}_n)]=W(\mathbf{A}_n\mathbf{x}_n)\exp[f_n(\mathbf{x}_n)],$$

where $\mathbf{A}_n\in B(H_n)$ and f_n are bounded homogeneous quadratic forms such that (2.4) is positive definite for $(\mathbf{A}_n,f_n,\sigma_n)$. We must assume some continuity in the time parameters, namely that each $\mathbf{A}_n$ is a strongly continuous function of $(t_0,\ldots,t_n)$ with $\mathbf{A}_n(0)=I$. Furthermore we assume that there is a Gaussian state $\bar\rho$ satisfying the invariance condition (A3)

$$\bar\rho_1(T_n(t_0,\ldots,t_n)[W(\mathbf{x}_n)])=\bar\rho(W(x_1))\times T_{n-1}(t_1,\ldots,t_n)[W(\mathbf{x}_{n-1})],$$

where $n-1$ refers to the $n-1$ last indices. With (2.1) and (2.3) this relation gives

$$f_n(\mathbf{x}_n)+g((\mathbf{A}_n\mathbf{x}_n)_1)=f_{n-1}(\mathbf{x}_{n-1})+g(x_1), \tag{3.1}$$

$$f_1(x)=g(x)-g(A_1x),$$

$$(\mathbf{A}_n\mathbf{x}_n)_{n-1}=\mathbf{A}_{n-1}\mathbf{x}_{n-1}. \tag{3.2}$$

The compatibility condition (A2) reads

$$\begin{aligned}&T_n(t_0,\ldots,t_n)[W(\mathbf{x}_{n-1}\oplus 0)]\\&\quad=T_{n-1}(t_0,\ldots,t_{n-1})[W(\mathbf{x}_{n-1})]\otimes I,\end{aligned}$$

where $n-1$ now refers to the $n-1$ first indices. This gives

$$\mathbf{A}_n(\mathbf{x}_n\oplus 0)=(\mathbf{A}_{n-1}\mathbf{x}_{n-1})\oplus 0, \tag{3.3}$$

$$f_n(\mathbf{x}_{n-1}\oplus 0)=f_{n-1}(\mathbf{x}_{n-1}),$$

where the last relation actually follows from (3.1) and (3.3). Together with the stationarity condition that all quantitites depend only on the time differences

$$\Delta t_k=t_k-t_{k-1},\quad k=1,\ldots,n,$$

(3.2) and (3.3) imply that $\mathbf{A}_n$ is of the form

$$(\mathbf{A}_n)_{ij}=\theta_{ij}A^{(j-i+1)}(\Delta t_i,\ldots,\Delta t_j), \tag{3.4}$$

where i,j refers to the indices in $H_n=\overset{n}{\oplus}H(i)$, $\theta_{ij}=0$ for $j<i$, $=1$ for $j\geqslant i$, and $A^{(k)}\in B(H)$ are strongly continuous functions of k nonegative parameters. From (3.1) follows that

$$f_n(\mathbf{x}_n)=\sum_{i=1}^{n}\left[g(x_i)-g\left(\sum_{j=i}^{n}A^{(j-i+1)}\times(\Delta t_j,\ldots,\Delta t_j)x_j\right)\right]. \tag{3.5}$$

Consequently the continuity of f_n in the time parameters follows from that of the $A^{(k)}$.

In addition to the positive definiteness condition on (2.4) there is one more condition not yet satisfied, and that is the additional compatibility condition (A4) which is most easily expressed in the correlation operator form (A6). Therefore we now turn to the derivation of the correlation operators. We see from the definition of T_n and (3.4) that the conditions for Lemma A-3 are satisfied. Consequently there are $\mathbf{y}_n, \mathbf{y}'_n \in H_n$ such that for all n

$$\begin{aligned} W(0 \oplus \mathbf{y}_{n-1}) T_n [W(-\mathbf{x}_n) W(-\mathbf{y}_{n-1} \oplus 0) \\ \times W(\mathbf{y}'_{n-1} \oplus 0) W(\mathbf{x}'_n)] W(0 \oplus -\mathbf{y}'_{n-1}) \\ = T_n [\{W(-x_k); W(x'_k)\}_1^n] \otimes I \end{aligned} \quad (3.6)$$

gives the correlation operator. From (3.4) follows that

$$\begin{aligned} \mathbf{y}_n &= B_n(\Delta t_1, ..., \Delta t_n)[\mathbf{x}_n] \oplus \mathbf{y}_{n-1} \\ &= \bigoplus_{k=0}^{n-1} B_{n-k}(\Delta t_{k+1}, ..., \Delta t_n)[\mathbf{x}_{n-k}], \end{aligned}$$

where $n-k$ refers to the $n-k$ last indices and the B_n are defined by

$$\begin{aligned} B_n(\Delta t_1, ..., \Delta t_n)[\mathbf{x}_n] = \sum_{k=1}^{n} [A^{(k)}(\Delta t_1, ..., \Delta t_k) x_k] \\ + B_{n-k}(\Delta t_{k+1}, ..., \Delta t_n)[\mathbf{x}_{n-k}], \\ B_0 = 0. \end{aligned}$$

From this relation follows that B_n is of the form

$$B_n(\Delta t_1, ..., \Delta t_n)[\mathbf{x}_n] = \sum_{k=1}^{n} B^{(k)}(\Delta t_1, ..., \Delta t_k) x_k,$$

where $B^{(k)} \in B(H)$ satisfies the recursion formula ($B^{(0)} = I$)

$$\begin{aligned} B^{(n)}(\Delta t_1, ..., \Delta t_n) \\ = \sum_{k=1}^{n} A^{(k)}(\Delta t_1, ..., \Delta t_k) \cdot B^{(n-k)}(\Delta t_{k+1}, ..., \Delta t_n) \end{aligned} \quad (3.7)$$

and are strongly continuous functions of the time parameters. We see from (3.6) that the correlation operators are of the form

$$\begin{aligned} T_n [\{W(-x_k); W(x'_k)\}_1^n] = W(-B_n[\mathbf{x}_n]) W(B_n[\mathbf{x}'_n]) \\ \times \exp[k_n(\mathbf{x}_n, \mathbf{x}'_n)], \end{aligned} \quad (3.8)$$

where the k_n are quadratic forms in $(\mathbf{x}_n, \mathbf{x}'_n)$. Now we can apply the additional compatibility relation (A6) to (3.8) to obtain by recursion

$$\begin{aligned} B^{(n)}(\Delta t_1, ..., \Delta t_n) &= B^{(n-1)}(\Delta t_1, ..., \Delta t_m + \Delta t_{m+1}, ..., \Delta t_n) \\ &= B^{(1)}(t_n - t_0) = A(t_n - t_0), \end{aligned}$$

$$A(t) = A_1(t) = A^{(1)}(t),$$

$$B_n[\mathbf{x}_n] = \sum_1^n A(t_k - t_0) x_k. \quad (3.9)$$

We can invert the relation (3.6) to obtain

$$\begin{aligned} T_n [W(-\mathbf{x}_n) W(\mathbf{x}'_n)] = T_n [\{W(-z_k); W(z'_k)\}_1^n] \\ \otimes W(-\mathbf{y}_{n-1}) W(\mathbf{y}'_{n-1}), \end{aligned}$$

where $\mathbf{z}_n$ is defined by the recursion relation

$$\mathbf{z}_n = \mathbf{x}_n - \mathbf{y}_{n-1}, \quad y_n = B_n[\mathbf{z}_n].$$

This shows the equivalence of the conditions (A1)–(A4) and (A5)–(A7) for GQSPs. The CP condition (A5) is just the positive definiteness of $\exp[k_n(\mathbf{x}, \mathbf{y})]$ which is equivalent to the positive definiteness of

$$K_n(\mathbf{x}, \mathbf{y}) = k_n(\mathbf{x}, \mathbf{y}) + k_n(\mathbf{0}, \mathbf{0}) - k_n(\mathbf{x}, \mathbf{0}) - k_n(\mathbf{0}, \mathbf{y}).$$

The k_n can be determined in terms of the f_n via (3.6), but they can also be derived directly from the invariance condition (A7) which gives with (3.8) the recursion formula

$$\begin{aligned} k_1(x, y) = g(x - y) - g(A(t)(x - y)) + i\sigma(x, y) \\ - i\sigma(A(t)x, A(t)y) \end{aligned}$$

$$\begin{aligned} k_n(\mathbf{x}_n, \mathbf{y}_n) - k_{n-1}(\mathbf{x}_{n-1}, \mathbf{y}_{n-1}) \\ = g(x_1 - y_1 + x_{1,n} - y_{1,n}) - g(x_{0,n} - y_{0,n}) \\ + i\sigma(x_1 + x_{1,n}, y_1 + y_{1,n}) - i\sigma(x_{0,n}, y_{0,n}) \\ - i\sigma(x_1, x_{1,n}) + i\sigma(y_1, y_{1,n}), \end{aligned}$$

where $n-1$ refers to the $n-1$ last indices and where we introduce the notation

$$x_{m,n} = \sum_{k=m+1}^{n} A(t_k - t_m) x_k.$$

The solution is given by

$$\begin{aligned} k_n(\mathbf{x}, \mathbf{y}) = K_n(\mathbf{x}, \mathbf{y}) - \tfrac{1}{2} K_n(\mathbf{x}, \mathbf{x}) - \tfrac{1}{2} K_n(\mathbf{y}, \mathbf{y}) \\ - i \sum_k \{\sigma(x_k, x_{k,n}) - \sigma(y_k, y_{k,n})\}, \end{aligned} \quad (3.10)$$

$$\begin{aligned} K_n(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{n} K(x_k, y_k) + \sum_{k=1}^{n-1} \{K(x_k, y_{k,n}) \\ + K(x_{k,n}, y_k)\} - K(x_{0,n}, y_{0,n}). \end{aligned} \quad (3.11)$$

Obviously k_n is determined by K and $A(t)$. We can sum up this paragraph in the following proposition.

Proposition 3-1: A stationary GQSP is defined by a positive $K \in B(\mathscr{H})$ (giving G and S when the real basis of $\mathscr{H} = H \oplus iH$ is specified) and a real strongly continuous function A: $\mathbf{R} \to B(H)$ with $A(0) = I$ such that the k_n defined by (3.11) are positive definite. The tensor product form is specified by the $A^{(n)}$ defined by the recursion relation (3.7) which now reads

$$\begin{aligned} A^{(n)}(\Delta t_1, ..., \Delta t_n) = A(t_n - t_0) \\ - \sum_{k=1}^{n-1} A^{(k)}(\Delta t_1, ..., \Delta t_k) \cdot A(t_n - t_k) \end{aligned}$$

and the f_n given by (3.5). The correlation operator form is specified by (3.8)–(3.11).

Remark: The GQSP is obviously Markovian iff $A(t)$ forms a semigroup: $A(s + t) = A(s) \cdot A(t)$. This is so if and only if $A^{(2)} = 0$ (and hence $A^{(k)} = 0$ for $k \geqslant 2$).

4. AUTOCOVARIANCE AND AUTOCORRELATION FUNCTIONS

We introduce the operator-valued autocovariance function

$$K(t) = KA(t) \ (t \geqslant 0), \quad = A(-t)^+ K \quad (t < 0).$$

and assume from now on that K^{-1} exists as a bounded operator.

Proposition 4-1: A strongly continuous function $\mathbf{R} \ni t \mapsto K(t) \in B(\mathscr{H})$ with $K(0) = K = G + iS$

uniquely defines a stationary GQSP on $\overline{\Delta(H,\sigma)}$ with an invariant Gaussian state given by G iff

(a) it is positive definite in the sense that

$$\sum_{i,j} X_i^+ K(t_i - t_j) X_j \geqslant 0$$

for all $\{t_i \in \mathbf{R}, X_i \in B(\mathscr{H})\}_1^N$, all N

(b) $A(t) = K^{-1}K(t)$ $(t \geqslant 0)$ is real: $A(t)^c = A(t)$

Proof: We write (3.11) in the form

$$K_n(\mathbf{x},\mathbf{y}) + K(x_{0,n}, y_{0,n}) = \sum_{k,l} (x_k, K(t_l - t_k) y_l)$$

K_n pos. def. for all n and $K \geqslant 0$ obviously implies that $K(t)$ is pos. def..

Conversely, let $K(t)$ be pos. def. and put

$$x_0 = -x_{0,n}.$$

Then we find that

$$\sum_{k,l=0}^{n} (x_k, K(t_l - t_k) y_l) = K_n(\mathbf{x}_n, \mathbf{y}_n)$$

and consequently K_n is pos. def.

We can normalize $K(t)$ and introduce the autocorrelation function

$$R(t) = K^{-1/2} K(t) K^{-1/2}.$$

$R(t)$ is obviously positive definite if $K(t)$ is and $R(0) = I$.

Proposition 4-2: If $R(t)$ is positive definite and $R(0) = I$ then $R(t)^+ R(t) \leqslant I$ and $R(t)$ is strongly continuous in t if it is continuous at $t = 0$.

Proof: In the positive definiteness condition

$$\sum_{k,l} X_k^+ R(t_k - t_l) X_l \geqslant 0$$

we first take $t_1 = 0, t_2 = t, X_1 = I, X_2 = R(t)$ to obtain

$$I - R(t)^+ R(t) \geqslant 0.$$

Secondly we put $t_1 = 0, t_2 = s, t_3 = t, X_1 = R(s)\cdot R(t-s) - R(t), X_2 = -R(t-s), X_3 = I$ to obtain

$$I - X_1^+ X_1 - X_2^+ X_2 \geqslant 0.$$

If $R(t)$ is strongly continuous at the origin, then $\text{s-lim}_{t\to s} X_2 = -I$ and consequently $\text{s-lim}_{t\to s} X_1 = 0$.

We want to justify the names autocovariance and autocorrelation function introduced above. If $\mathscr{A} = \pi(\overline{\Delta(H,\sigma)})$ is a ray continuous *-representation of the CCR in a Hilbert space $\mathscr{K}$, then there is for each $x \in H$ a self-adjoint operator in $\mathscr{K}$ which we denote by (x,X), such that

$$W(x) = \exp[i(x,X)].$$

If we write $x = |x|e$, $X_i = (e_i, X)$, then we find (with some abuse of notation)

$$\bar\rho(X_1 T_1(0,t)[X_2])$$
$$= -\frac{d^2}{d|x_1|\,d|x_2|}\bar\rho(W(x_1)T_1(0,t)[W(x_2)])\big|_{x_1=x_2=0}$$
$$= (e_1, K(t)e_2) = K(t)_{12}.$$

Higher order correlations of the form

$$\bar\rho(T_n[\{X_k; Y_k\}_1^n])$$

are similarly derived from the relation

$$\bar\rho(T_n[\{W(-x_k); W(y_k)\}_1^n]) = \exp[G_n(\mathbf{x},\mathbf{y})],$$

$$G_n(\mathbf{x},\mathbf{y}) = \sum_{k,l=1} (x_k, K(t_l - t_k) y_l) - \sum_{k=1}^{n-1} \{(x_k, Kx_{k,n}) + (y_{k,n}, K y_k)\} - \tfrac{1}{2}\sum_{k=1}^{n} \{(x_k, Kx_k) + (y_k, K y_k)\}.$$

They are obtained as sums of products of matrix elements of $K(t)$ in precisely the same way as in the commutative case, e.g.,

$$\bar\rho(T_4[\{X_k;0\}_1^4]) = K(t_2 - t_1)_{12} K(t_4 - t_3)_{34} + K(t_3 - t_1)_{13} K(t_4 - t_2)_{24} + K(t_4 - t_1)_{14} K(t_3 - t_2)_{23}.$$

In contradistinction to a commutative SP the definition of a QSP includes the response of the system to external forces. In the case of the GQSP we can easily deal with a class of generalized forces defined through the time derivative of the expectation of the generator X introduced above. The corresponding finite transformations are simply gauge transformations

$$U(y)[W(x)] = \exp[i(y,x)]W(x),$$

where $y = f(t)\Delta t \in H$ and $\mathbf{R} \ni t \to f(t) \in H$ is a sufficiently regular function. The calculations are simplified if we assume that S^{-1} exists as a bounded operator (this is not necessary, however). Then we can implement $U(y)$ by a unitary transformation

$$U(y)[W(x)] = W(-\tfrac{1}{2}S^{-1}y)W(x)W(\tfrac{1}{2}S^{-1}y).$$

We let this type of instrument act on the system at the instants $t_k = kt/n$, $k = 1,\ldots,n$. The time development of the system during the time interval $(0,t)$ is then given by

$$T[W(x)] = \exp[i\sigma(y_n, x)]T_n[\{W(-z_k); W(z_k')\}_1^n]$$
$$z_k = \tfrac{1}{2}S^{-1}y_k = \tfrac{1}{2}S^{-1}f(t_k)t/n,$$
$$z_k' = z_k, \quad k < n, \quad z_n' = z_n + x.$$

The expectation of this operator in the state $\bar\rho$ is

$$\exp[i\sigma(z_n,x) + G_n(\mathbf{z},\mathbf{z}')]$$
$$= \exp\Big[\sum_{k=1}^{n} 2i\,\mathrm{Im}(z_k, K(t_n - t_k)x) + i\,\mathrm{Im}(z_n, Kx) - \tfrac{1}{2}(x,Kx) + i(z_n, Sx)\Big]$$
$$= \exp\Big[i\sum_{k=1}^{n} (A(t_n - t_k)^+ f(t_k)t/n, x) - \tfrac{1}{2}(x,Gx)\Big].$$

We take the limit $n \to \infty$ and assume that the integral

$$A^+(f,t) = \int_0^t A(t-\tau)^+ f(\tau)\,d\tau$$

exists in a weak sense. We then obtain

$$\bar\rho(T(t)[W(x)]) = \exp[i(A^+(f,t),x) + g(x)]$$

and the expectation of (X,e) at time t is $(A^+(f,t),e)$. We conclude that the response of the system to this type of force is given by $K(t)$. This is strongly reminiscent of a version of the fluctuation-dissipation theorem in the stochastic description of thermodynamic fluctuations. In the classical case, however, this relation is not a consequence of the definition of a stochastic process, but involves additional information in the form of a linear response hypothesis or a Langevin equation.

5. SPECTRAL REPRESENTATION OF THE AUTOCORRELATION FUNCTION

We want to derive a noncommutative analog of Bochner's theorem for the autocorrelation function $R(t)$. The first step is

Lemma 5-1: $\mathbf{R} \ni t \mapsto R(t) \in B(\mathscr{H})$ is a weakly continuous positive definite function iff there is a positive operator-valued measure (POVM) E on $\mathbf{R}$ such that

$$R(t) = \int_{-\infty}^{\infty} \exp(i\omega t) E(d\omega). \tag{5.1}$$

The definition of a POVM is given, e.g., in Ref. 12. The proof proceeds from Bochner's theorem precisely as the proof of Stone's theorem in Ref. 13 Sec. 138. Note that weak continuity implies strong continuity.

Corollary: Under the same conditions there is a Hilbert space $\mathscr{K}$, a weakly continuous group of unitary operators $U(t)$ in $\mathscr{K}$ and a projection $P: \mathscr{K} \to \mathscr{H}$ such that $R(t) = PU(t)P$. The proof follows from Theorem 8.1 of Ref. 14.

The condition $A(t)^c = A(t)$ in Proposition 4-1 must be transcribed into a condition on $E(\omega)$. Define

$$(2\pi i)^{-1} \int (\omega - z)^{-1} E(d\omega) = E_+(z), \quad \mathrm{Im} z > 0,$$
$$= -E_-(z), \quad \mathrm{Im} z < 0.$$

Obviously

$$E_+(z^*) = E_-(z)^+,$$

$$E_-(z) = (2\pi)^{-1} \int_0^{\infty} R(t) \exp(-izt)\, dt.$$

Introduce

$$V = (K^c)^{1/2} K^{-1/2}, \quad N = V^+ V.$$

Note that $V^c = V^{-1}$. The condition $A(t)^c = A(t)$ then reads

$$E_-(-z^*) = V^+ [NE_-(z)]^c V,$$
$$E_+(-z^*) = V^+ [E_+(z) N]^c V \tag{5.2}$$

(the two equalities are just Hermitian conjugates of each other). From Proposition 4-1 follows

Proposition 5-2: For given K, the strongly continuous function $\mathbf{R} \ni t \mapsto R(t) \in B(\mathscr{H})$ is the autocorrelation function of a GQSP iff it is of the form (5.1) where the POVM E satisfies the reality condition (5.2) and is normalized to unity,

$$E(\mathbf{R}) = I.$$

As the condition (5.2) is quite implicit we want to describe in more detail how to construct a large set of POVMs which satisfy it. We start with POVMs in the commutant of N.

Proposition 5-3: (a) Let ω_0 belong to the point spectrum and put

$$\Delta E = \lim_{\epsilon \to 0} [E(\omega_0) - E(\omega_0 - \epsilon)].$$

Then $\Delta E \in \{N\}'$.

(b) If there is a gap in the spectrum, i.e., an open interval A on $\mathbf{R}$ such that $E(A) = 0$ then $E(\omega) \in \{N\}'$

Proof: (a) It is easily seen that

$$\Delta E = 2 \lim_{\epsilon \to 0} \int_{\gamma(\epsilon)} E_-(z)\, dz,$$

$$\gamma(\epsilon): z = \omega_0 + \epsilon \exp(i\theta), \quad -\pi \leqslant \theta \leqslant 0.$$

The equality (5.2) integrated over $\gamma(\epsilon)$ gives that $-\omega_0$ is a point in the discrete spectrum with a "jump" which must be positive:

$$\Delta E' = V^+ (N \Delta E)^c V \geqslant 0.$$

Hence $N\Delta E = V^+ (\Delta E')^c V$. Hermitian conjugation gives

$$N \Delta E = \Delta E N.$$

(b) If there is a gap in the spectrum, then $E_+(z)$ and $-E_-(z)$ are analytic continuations of each other. From (5.2) we obtain by analytic continuation

$$E_+(-z^*) = V^+ [NE_+(z)]^c V.$$

Consequently $[N, E_+(z)] = 0$ and $E(\omega) \in \{N\}'$. Conversely, given a POVM E_1 in $\{N\}'$ we put

$$E_1'(\omega) = V^+ [NE_1(-\omega)]^c V.$$

E_1' is then a POVM in $\{N\}'$ and

$$E(\omega) = A^{-1/2} [E_1(\omega) + E_1'(\omega)] A^{-1/2},$$
$$A = E_1(\mathbf{R}) + E_1'(\mathbf{R})$$

is a normalized POVM in $\{N\}'$ satisfying (5.2):

$$E(\omega) = V^+ [NE(-\omega)]^c V.$$

This construction gives the most general normalized POVM in $\{N\}'$ satisfying (5.2).

We will now show how to find a POVM which is not in $\{N\}'$ and which consequently covers $\mathbf{R}$. Assume that

$$e(\omega) = dE(\omega)/d\omega$$

exists as a positive weakly measurable function. Define

$$\tau[e](\omega) = \lim_{\epsilon \to 0} V^+ [E_+(-\omega + i\epsilon) N + NE_-(\omega - i\epsilon)]^c V$$

whenever the right-hand side exists as a weakly measurable function. Then (5.2) is satisfied iff

$$e(\omega) = \tau[e](\omega). \tag{5.3}$$

A necessary condition for this to hold is that

$$\lim_{\epsilon \to 0} [E_+(-\omega + i\epsilon) N + NE_-(\omega - i\epsilon)] \tag{5.4}$$

is a positive weakly measurable function. If this is true, then the identity

$$\tau \cdot \tau[e] = e$$

holds. Consequently

$$e(\omega) = \tfrac{1}{2}[e(\omega) + \tau[e](\omega)]$$

is a positive weakly measurable function satisfying (5.3). The problem is then to construct a positive weakly measurable $e(\omega)$ such that (5.4) is positive and weakly measurable. When the limit exists, the expression (5.4) can be rewritten in the form

$$e'(\omega)\equiv\tfrac{1}{2}[e(\omega)N+Ne(\omega)] + (2\pi i)^{-1}\ \mathrm{PV}\int d\lambda\,(\lambda-\omega)^{-1}[e(\lambda),N].$$

In the second term of Hilbert transform type $e(\omega)$ must be sufficiently regular for the integral to represent a measurable function. In order that this function should be integrable over $(-\infty,\infty)$ we must demand that

$$\int d\omega\,[e(\omega),N]=0,\quad \text{i.e.,}\quad \int d\omega\,e(\omega)\in\{N\}'.$$

Note that we cannot nomalize $e(\omega)$ from the start.
Let $a{:}\mathbf{R}\to B(\mathscr{H})$ be a sufficiently regular function of compact support such that

$$a(\omega)^+=a(\omega),\qquad \int d\omega\,a(\omega)=0.$$

It is easy to see that as N and N^{-1} are assumed to be finite, there is a minimal integrable scalar function $c(\omega)$ such that

$$e_1(\omega)\equiv a(\omega)+c(\omega)I\geqslant 0,$$

$$e_1'(\omega)\equiv\tfrac{1}{2}[e_1(\omega)N+Ne_1(\omega)] + (2\pi i)^{-1}\,\mathrm{PV}\int d\lambda\,(\lambda-\omega)^{-1}[a(\lambda),N]\geqslant 0.$$

e_1 and e_1' define POVMs E_1 and E_1'. The set of POVMs E in $\{N\}'$ such that

$$E\leqslant E_1,\quad E\leqslant N^{-1/2}E_1'N^{-1/2}$$

is a partially ordered set. Choose any maximal element E_2. Then

$$E_3=E_1-E_2,\quad E_3'=E_1'-NE_2$$

are POVMs. Note that

$$E_3(\mathbf{R})=N^{-1}E_3'(\mathbf{R})\equiv A\in\{N\}'.$$

Define

$$E=\tfrac{1}{2}[A^{-1/2}E_3A^{-1/2}+V^+(A^{-1/2}E_3'A^{-1/2})^cV].$$

E is then a normalized POVM satisfying (5.2). Furthermore if F is a POVM in $\{N\}'$ satisfying (5.2) such that $\lambda F\leqslant E$ for some $\lambda>0$, then $F=0$.

ACKNOWLEDGMENT

This work was made with financial support from the Swedish Natural Science Research Council.

APPENDIX

A class of instruments relevant to the definition of a GQSP can be specified in the following way.

Let $\Omega=\{\omega\}$ be a countable set of outcomes. An instrument on $\mathscr{A}^0$ is given by a map $\mathscr{E}$ from subsets of Ω to $\mathrm{CP}(\mathscr{A}^0)$ such that

$$\mathscr{E}(\emptyset)=0,\quad \mathscr{E}(\Omega)[I]=I,$$

$$\mathscr{E}(E)[X]=\sum_k\mathscr{E}(E_k)[X],\quad X\in\mathscr{A}^0,$$

for disjoint sets E_k with $E=\cup E_k$, where the sum is assumed to be norm convergent. Of course each $\mathscr{E}(E)$ extends uniquely to a map in $\mathrm{CP}(\mathscr{A})$ and the sum above is in fact uniformly norm convergent on $\mathscr{A}$.[8]

In analogy with the tensor product representation derived in Ref. 1 we introduce the following structure for a stationary QSP on $\mathscr{A}^0$ (extended to $\mathscr{A}$ by continuity).

A QSP consists of a set $\{T_n\in B(\mathscr{A}_n^0)\}_1^\infty$, T_n depending on $n+1$ time parameters (continuously in some topology which we will not specify) through the time differences $\Delta t_k=t_k-t_{k-1}$, and an invariant state $\bar\rho$ on $\mathscr{A}$ satisfying the conditions (a)–(d) below.

(a) *Complete positivity*: $T_n\in\mathrm{CP}(\mathscr{A}_n^0)$ which means that

$$\sum_{i,j}X(i)^+T_n[W(\mathbf{x}(i))^+W(\mathbf{x}(j))]X(j)\geqslant 0 \tag{A1}$$

for all $\{\mathbf{x}(i)\in H_n, X(i)\in\mathscr{A}_n\}_1^N$ and all N. T_n extends to $\mathscr{A}_n$ by continuity.

(b) *Compatibility*:

$$T_n(t_0,\dots,t_n)[X\otimes I]=T_{n-1}(t_0,\dots,t_{n-1})[X]\otimes I \tag{A2}$$

for all $x\in\mathscr{A}_{n-1}$.

(c) *Invariance*:

$$\bar\rho_1(T_n(t_0,\dots,t_n)[X\otimes Y]) =\bar\rho(X)T_{n-1}(t_1,\dots,t_n)[Y] \tag{A3}$$

for all $\{X\in\mathscr{A}, Y\in\mathscr{A}_{n-1}\}$, where $\bar\rho_1$ denotes $\bar\rho$ considered as a partial state on the first factor in $\mathscr{A}_n$.

(d) In addition we must impose a compatibility condition corresponding to that given in Ref. 1 Sec. 3, Remark 1. This condition will be formulated below.

In order to be able to derive the correlation operators and the probabilities associated with instruments of the type defined above we will restrict ourselves to a special class of CP maps on $\mathscr{A}_n^0$.

Lemma A-1: Let $T_n\in B(\mathscr{A}_n^0)$ be of the form (for a given m)

$$T_n[X\otimes W(\mathbf{x}_{n-m}^n)^+W(\mathbf{y}_{n-m}^n)]=T_{m,n}(\mathbf{x}_{n-m}^n,\mathbf{y}_{n-m}^n)[X]\otimes\left[\bigotimes_{k=m}^{n-1}X_{n-k}(\mathbf{x}_{n-k}^n,\mathbf{y}_{n-k}^n)\right],$$

where $X\in\mathscr{A}_m^0$, $\mathbf{x}_{n-k}^n=(x_{n+1},\dots,x_n)$. Let $T_{m,n}(\mathbf{x},\mathbf{y})\in B(\mathscr{A}_m^0)$ be positive definite in the sense that

$$\sum_{i,j}^N X(i)^+T_{m,n}(\mathbf{x}(i),\mathbf{x}(j))[W(\mathbf{y}(i))^+W(\mathbf{y}(j))]X(j)\geqslant 0$$

for all $\{\mathbf{x}(i)\in H_{m-n},\mathbf{y}(i)\in H_m, X(i)\in\mathscr{A}_m\}$ and all N, and the $X_{n-k}(\mathbf{x},\mathbf{y})\in\mathscr{A}^0$ positive definite in the sense that

$$\sum_{i,j}^N X(i)^+X_{n-k}(\mathbf{x}(i),\mathbf{x}(j))X(j)\geqslant 0$$

for all $\{\mathbf{x}(i)\in H_{n-k}, X(i)\in\mathscr{A}\}$, all N and all $m\leqslant k<n$. Then $T_n\in\mathrm{CP}(\mathscr{A}_n^0)$.

Proof: From Ref. 7 Sec. 1,2 we know that $T_{m,n}$ and $\{X_{n-k}\}$ have Kolmogorov decompositions.

$$T_{m,n}(\mathbf{x},\mathbf{y})[X]=V_{m,n}(\mathbf{x})^+\,\pi_{m,n}(X)V_{m,n}(\mathbf{y}),$$

$$X_{n-k}(\mathbf{x},\mathbf{y})=W_{n-k}(\mathbf{x})^+W_{n-k}(\mathbf{y}),$$

where $\pi_{m,n}$ is a *-representation of $\mathscr{A}_m$ in some Hilbert space. It is then obvious that (A1) is satisfied.

Corollary: If T_n is of this form for $m=1$, then it is true for all $1\leqslant m<n$.

Denote by ρ_0 the central state on $\mathscr{A}$ defined by[5]

$$\rho_0(W(x)) = \begin{cases} 1 & \text{for } x=0, \\ 0 & \text{for } x\neq 0, \end{cases}$$

and denote by $\rho_{0,m}$ the partial state on the mth factor in $\mathscr{A}_n$ given by ρ_0.

Lemma A-2: On the set of all $T_n \in \mathrm{CP}(\mathscr{A}_n^0)$ with the properties of Lemma A-1 for a given m there is defined a contraction

$$\Theta_m : T_n \mapsto T'_{n-1} \in \mathrm{CP}(\mathscr{A}^0_{n-1})$$

through

$$T'_{n-1}(X\otimes Y) = \sum_{x\in H} \rho_{0,m}((I_m \otimes W(x) \otimes I_{n-m-1}) \times T_n[X\otimes W(-x)\otimes Y])$$

for all $X\otimes \mathscr{A}^0_{m-1}$, $Y\in \mathscr{A}^0_{n-m}$.

Proof: Note first that the sum is always finite. From the given form of T_n we obtain that T'_{n-1} is of the same form with

$$X'_{n-k} = X_{n-k}, \quad m\leqslant k < n-1,$$

$$\begin{aligned} T'_{m,n-1}(\mathbf{x}^{n-1}_{n-m-1},\mathbf{y}^{n-1}_{n-m-1})[W(\mathbf{x}_m)^+ W(\mathbf{y}_m)] &= T_{m,n}(\mathbf{x}^{n-1}_{n-m},\mathbf{y}^{n-1}_{n-m})[W(\mathbf{x}_{m-1})^+ W(\mathbf{y}_{m-1}) \\ &\quad \otimes X_{n-m}(\mathbf{x}^{n-1}_{n-m},\mathbf{y}^{n-1}_{n-m})] \\ &= V_{m,n}(\mathbf{x}^{n-1}_{n-m})^+ \pi_{m,n}[W(\mathbf{x}_{m-1})^+ W(\mathbf{y}_{m-1}) \\ &\quad \otimes X_{n-m}(\mathbf{x}^{n-1}_{n-m},\mathbf{y}^{n-1}_{n-m})] V_{m,n}(\mathbf{y}^{n-1}_{n-m}), \end{aligned}$$

where $\mathbf{x}_m = (x_1,\ldots,x_m)$. $T'_{m,n-1}$ obviously has the positive definiteness properties defined in Lemma A-1.

Remarks: 1. If T_n belongs to the class of maps of Lemma A-1 for a given m, then so does ($\circ$ denotes composition of CP maps)

$$\Big(\overset{n}{\otimes} T'_{(k)}\Big)\circ T_n \quad \text{for all } T'_{(k)}\in \mathrm{CP}(\mathscr{A}^0),$$

$$T_n \circ (T'_m \otimes I_{n-m}) \quad \text{for all } T'_m \in \mathrm{CP}(\mathscr{A}^0_m).$$

2. If $T'_{(m)} = I_{m-1}\otimes T'\otimes I_{n-m} \in \mathrm{CP}(\mathscr{A}^0_n)$, where $T'\in \mathrm{CP}(\mathscr{A}^0)$, then we have the following relation

$$\Theta_m[T_n\circ T'_{(m)}] = \Theta_m[T'_{(m+1)}\circ T_n].$$

3. If $A,B\in \mathscr{A}^0$ and we put $T'(X) = AXB$, then the preceding relation still holds (T' is a "matrix element" of a CP map on the 2×2 matrix algebra over $\mathscr{A}$).

We can now formulate the additional compatibility relation mentioned above.

(d) *Compatibility*:

$$\Theta_m[T_n(\{t_k\}_1^n)] = T_{n-1}(\{t_k\}_1^n : k\neq m). \tag{A4}$$

The probabilities corresponding to those of Ref. 1 Sec. 2 are defined as follows. Let $\{\mathscr{E}_k\}_1^n$ be a set of instruments of the type defined above. The probability of obtaining a sequence $(\omega_1,\ldots,\omega_n)$ of outcomes when the instruments act at instants $t_1 < t_2 < \ldots < t_n$, given that the state at time $t_0 < t_1$ is ρ, is given by

$$\rho\Big(\Big(\prod_{m=1}^{n-1}\Theta_m\Big)\Big[T_n(t_0,\ldots,t_n)\circ\Big(\overset{n}{\otimes}\mathscr{E}_k(\omega_k)\Big)\Big][I]\Big).$$

This expression is well defined by Remarks 1,2 above. It is obviously positive. The definition coincides with that in Ref. 1 whenever they are both valid.

The correlation operators are defined in the following way

$$T_n(t_0,\ldots,t_n)[\{W(-x_k);W(x'_k)\}_1^n] = \Big(\prod_{m=1}^{n-1}\Theta_m\Big)\Big[T_n(t_0,\ldots,t_n)\circ\Big(\overset{n}{\otimes} T(x_k,x'_k)\Big)\Big][I],$$

where $T(x,y)[X] = W(-x)X\,W(y)$, and they satisfy the following relations which correspond to those of Ref. 1 Sec. 2.

(a) *Complete positivity*:

$$\sum_{i,j}^{N} X(i)^+ T[\{W(-x_k(i));W(x_k(j))\}_1^n]X(j)\geqslant 0 \tag{A5}$$

for all $\{X(i)\in\mathscr{A}_n, \mathbf{x}(i)\in H_n\}$, all N.

(b) *Compatibility*:

$$\begin{aligned} &T_n(\{t_k\}_1^n)[\{W(-x_k);W(x'_k)\}_1^n] \mid x_m = x'_m = 0 \\ &= T_{n-1}(\{t_k\}_1^n : k\neq m)[\{W(-x_k);W(x'_k)\}_1^n : k\neq m] \end{aligned} \tag{A6}$$

which corresponds to (A2) for $m=n$ and to (A4) for $m<n$.

(c) *Invariance*:

$$\begin{aligned} \bar\rho(T_n(t_0,\ldots,t_n)[\{W(-x_k);W(x'_k)\}_1^n]) &= \bar\rho(W(-x_1)T_{n-1}(t_1,\ldots,t_n) \\ &\quad \times[\{W(-x_k);W(x'_k)\}_2^n]W(x'_1)). \end{aligned} \tag{A7}$$

The following lemma helps in the calculation of the correlation operators for GQSPs.

Lemma A-3: If $T_n\in \mathrm{CP}(\mathscr{A}^0_n)$ is of the type defined in Lemma A-1 for $m=1$ and if there are unitary $U_k, U'_k\in \mathscr{A}^0$ such that

$$(I\otimes U)T_n\circ\Big(\overset{n}{\otimes} T(x_k,x'_k)\Big)[U^+U'\otimes I](I\otimes U')^+ \in \mathscr{A}^0\otimes I_{n-1},$$

where $U = \overset{n-1}{\otimes} U_k$, then this expression is equal to

$$T_n[\{W(-x_k);W(x'_k)\}_1^n]\otimes I_{n-1}.$$

Proof: This follows from the definitions of Θ and the correlation operator and Remark 3 above.

[1]G. Lindblad, Commun. Math. Phys. **65**, 281–94 (1979).
[2]G. Lindblad, Rep. Math. Phys. **10**, 393–406 (1976).
[3]B. Demoen, P. Vanheuverzwijn, and A. Verbeure, "Completely positive quasifree maps on the CCR algebra," Preprint KUL-TH-77/008, Leuven 1977.
[4]P. Vanheuverzwijn, Ann. Inst. H. Poincaré **29**, 123–38 (1978).
[5]J. Manuceau, M. Sirugue, D. Testard, and A. Verbeure, Commun. Math. Phys. **32**, 231–43 (1973).
[6]J. Manuceau, Ann. Inst. H. Poincaré **8**, 139–61 (1968).
[7]D.E. Evans and J.T. Lewis, Commun. Dubl. Inst. Adv. Studies Ser. A **24**, (1977).
[8]B. Demoen, P. Vanheuverzwijn, and A. Verbeure, Lett. Math. Phys. **2**, 161–66 (1977).
[9]J. Manuceau, and A. Verbeure, Commun. Math. Phys. **9**, 293–43 (1973).
[10]D.E. Evans, and J.T. Lewis, J. Funct. Anal. **26**, 369–77 (1977).
[11]K.R. Parthasarathy, and K. Schmidt, Springer Lecture Notes in Mathematics **272**, (1972).
[12]E.B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).
[13]F. Riesz, B. Sz-Nagy, *Vorlesungen über Funktionalanalysis* (Deutscher Verlag der Wissenschaften, Berlin, 1956).
[14]R. Beals, *Topics in Operator Theory* (Univ. Chicago Press, Chicago, 1971).

Quantum soliton and classical soliton

H. Matsumoto
Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1
G. Oberlechner
Centre de Recherches Nucleaires, Université Louis Pasteur, 67037 Strasbourg, France
M. Umezawa
Centre de Recherches Nucleaires, Université Louis Pasteur, 67037 Strasbourg, France
H. Umezawa
Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1

(Received 28 March 1979; accepted for publication 18 May 1979)

The interaction between quanta and a static soliton is discussed in the boson theory. As a natural consequence, there appears the quantum coordinate associated with the position of the soliton, which leads to a quantum or classical behavior of the soliton depending on its size and the experimental conditions. It will be shown that the presence of the quantum coordinate (or collective coordinate) is required by the equal-time canonical commutation relation.

1. INTRODUCTION

In recent papers[1,2] we showed how the classical Euler equation is obtained from the Heisenberg equation when the boson transformation is used; this formalism is called the boson theory. According to this method the soliton solutions[3-5] can be regarded as extended objects that are created by the condensation of a certain boson. In Ref. 1 some of us (H.M. and H.U.) applied this method to the (1 + 1)-dimensional $\lambda\phi^4$-model and obtained the static soliton solution from the Heisenberg equation. In Ref. 2, others of us (G.O. and M.U.) made an extensive analysis of the sine-Gordon equation. A remarkable result is the *linear law* for the boson transformation parameter $f(x)$:

$$f^{(N)}(x) = \sum_{i=1}^{N} f_i(x), \tag{1.1}$$

where $f^N(x)$ is the parameter for N solitons while the functions $f_i(x)$ are the single soliton parameters. One merit of this approach lies in the fact that the theory begins with the *usual quantum field theory*, from which creation of extended objects naturally follows. Thus, extended objects (i.e., solitons) appear in a quantum many-body system, and therefore, they interact with the quanta. It is the purpose of this paper to study the interaction between extended objects and the quanta. In this section, we present a brief summary of the boson theory.

Using the same notation as Ref. 1, we consider a Heisenberg equation

$$\Lambda(\partial)\psi = F[\psi], \tag{1.2}$$

where ψ is a scalar Heisenberg field. For the sake of simplicity, we assume that the perturbative method is applicable to Eq. (1.2). Eq. (1.2) leads to the Yang-Feldman equation

$$\psi = \varphi^0 + [\Lambda(\partial)]^{-1}F[\psi], \tag{1.3}$$

where φ^0 is a renormalized free boson field satisfying

$$\Lambda(\partial)\varphi^0 = 0. \tag{1.4}$$

and has the form

$$\varphi^0(x) = Z^{\frac{1}{2}} \int \frac{d^3k}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_k}} \times [\alpha(\mathbf{k})e^{i\mathbf{kx} - i\omega_k t} + \alpha^\dagger(\mathbf{k})e^{-i\mathbf{kx} + i\omega_k t}]. \tag{1.5}$$

Here $\alpha(\mathbf{k})$ is the annihilation operator of the φ^0-quantum with momentum $\mathbf{k}$ and energy ω_k, and Z is the wave function renormalization constant. In (1.3), we used the notation φ^0 instead of φ^{in} in order to make it clear that our choice for the free field is not necessarily limited to the in-field. When φ^0 is different from the in-field, a suitable choice of the Green function $\Lambda^{-1}(\partial)$ should be made accordingly.

When we solve (1.3) by successive iteration we are led to the usual perturbative expansion. The result is an expression of ψ in terms of φ^0:

$$\psi(x) = \psi[x;\varphi^0], \tag{1.6}$$

which is called the dynamical map.

Let us now introduce a c-number function $f(x)$ that satisfies

$$\Lambda(x)f(x) = 0. \tag{1.7}$$

Then we can generalize the Yang–Feldman equation as

$$\psi = \varphi^0 + f + [\Lambda(\partial)]^{-1}F[\psi]. \tag{1.8}$$

Solving (1.8) by successive iteration, we obtain a new solution of (1.2):

$$\psi^f(x) = \psi(x;\varphi^0 + f). \tag{1.9}$$

Note that ψ^f is related to ψ through the boson transformation $\varphi^0 \to \varphi^0 + f$. The fact that both ψ and ψ^f satisfy the same Heisenberg equation (1.2) is the content of the boson transformation theorem[6]:

$$\Lambda(\partial)\psi^f = F[\psi^f]. \tag{1.10}$$

Let us now introduce the classical field $\phi^f(x)$ as

$$\phi^f = \langle 0|\psi^f|0\rangle. \tag{1.11}$$

This describes an extended object created by the condensation

 0022-2488/79/102088-09$01.00

of the φ^0-quanta. It was shown in Ref. 1 that, when we define

$$\phi_0^f = \lim_{\hbar\to 0}\phi^f, \tag{1.12}$$

this satisfies the Euler equation

$$\Lambda(\partial)\phi_0^f = F[\phi_0^f]. \tag{1.13}$$

Equation (1.12) shows that ϕ_0^f *is given by the tree diagrams only*.

In the next section, we derive a kind of Schrödinger equation for the one particle matrix elements in the presence of extended objects; the equation contains a *self-consistent potential* that is created by the soliton. In Sec. 3, the structure of the dynamical map ψ^f is investigated in the tree approximation. The result shows that the introduction of *certain canonical quantum coordinates associated with the position of soliton is required from the canonical commutation relations.* Significance of these coordinates is discussed. A general consideration without any use of the tree approximation for the results of Sec. 3 is developed in Sec. 4. This consideration again confirms the appearance of the quantum coordinates mentioned above. In Sec. 5, we present a brief summary of our results.

Quantization of the classical soliton has been studied by many. It is comforting to us that the boson theory naturally leads to results, several features of which resemble the results obtained by other approaches that began with classical solutions.[4-6] The quantum coordinate (or the collective coordinate) is a kind of concept that is widely used in solid state physics.[7] The study of extended objects by use of the collective coordinate is extensively done in Ref. 6. A significant question from the quantum field theoretical viewpoint is to ask what is the origin of this coordinate. Our conclusion is that the presence of the collective coordinate is required from the canonical relations.

2. THE SELF-CONSISTENT POTENTIAL INDUCED BY SOLITONS

In this section, we study the interaction between the soliton and φ^0-quanta. To do this, we determine the general structure of the dynamical map in (1.9).

The Taylor expansion of the dynamical map in (1.9) leads to

$$\psi^f(x) = \phi^f(x) + \int d^4y\, c(x;y)\varphi^0(y) + \int d^4y \times \int d^4z\, c(x;y,z){:}\varphi^0(y)\varphi_0(z){:} + \cdots, \tag{2.1}$$

where

$$c(x;y) = \frac{\delta\phi^f(x)}{\delta f(y)}, \tag{2.2}$$

$$c(x;y,z) = \frac{1}{2}\frac{\delta}{\delta f(y)}\frac{\delta}{\delta f(z)}\phi^f(x), \text{ etc.,} \tag{2.3}$$

and derivatives mean functional derivatives.

We now introduce

$$J^f(x) = \langle 0|F[\psi^f]|0\rangle. \tag{2.4}$$

Then, the Heisenberg equation (1.2) leads to

$$\Lambda(\partial)\phi^f = J^f(x). \tag{2.5}$$

Let us now change $f(x)$ into $f(x)+\epsilon g(x)$, where $g(x)$ also satifies (1.7), that is,

$$\Lambda(\partial)g(x) = 0. \tag{2.6}$$

Here ϵ is an infinitesimal constant. We have

$$\delta\phi_g^f(x) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}[\phi^{f+\epsilon g}(x) - \phi^f(x)] = \int d^4y\, \frac{\delta\phi^f(x)}{\delta f(y)} g(y). \tag{2.7}$$

Since $\phi^{f+\epsilon g}$ satisfies (2.5),

$$\Lambda(\partial^x)\delta\phi_g^f(x) = \int d^4y\, \frac{\delta J^F(x)}{\delta f(y)} g(y). \tag{2.8}$$

Here ∂^x means the derivative with respect to x. For given $f(x)$, $\phi^f(x)$ is uniquely determined. Furthermore, there is a one-to-one correspondence between $f(x)$ and $\phi^f(x)$, under the given boundary condition for $f(x)$. Then we can regard $f(x)$ as a functional of $\phi^f(x)$. As a result, $J^f(x)$ is also regarded as a functional of $\phi^f(x)$,

$$J^f(x) = J(x;\phi^f). \tag{2.9}$$

Since

$$\frac{\delta J^f(x)}{\delta f(y)} = \int d^4z\, \frac{\delta J(x;\phi^f)}{\delta\phi^f(z)}\frac{\delta\phi^f(z)}{\delta f(y)}, \tag{2.10}$$

we have

$$\Lambda(\partial^x)\,\delta\phi_g^f(x) = \int d^4z\, V^f(x,z)\,\delta\phi_g^f(z), \tag{2.11}$$

with

$$V^f(x,z) = \frac{\delta J(x;\phi^f)}{\delta\phi^f(z)}. \tag{2.12}$$

The function V^f is called the "self-consistent potential." The expression $\delta\phi_g^f(x)$ behaves like a wave function of a quantum in a nonlocal potential $V^f(x,z)$.

The significance of Eq. (2.11) lies in the fact that the linear φ^0-term in the dynamical map is equal to $\delta\phi_{\varphi_0}^f(x)$ and therefore that it satisfies (2.11). This implies that *φ^0-quanta feels the self-consistent potential $V^f(x,z)$ which is induced by soliton.* Choosing

$$g(x) = \frac{Z^{\frac{1}{2}}}{\sqrt{(2\pi)^3 2\omega_k}}\exp[\pm i(\mathbf{k}\cdot\mathbf{x} - \omega_k t)], \tag{2.13}$$

we can determine the linear φ^0-term in the dynamical map (2.1). Here ω_k is the energy of the φ^0-quantum with momentum $\mathbf{k}$. In a similar manner, the higher order terms in (2.1) can be obtained.

When the tree approximation is used [see (1.13)], $V^f(x,z)$ is simply given by

$$V^f(x,z) = F'[\phi^f(x)]\delta^{(4)}(x-z), \tag{2.14}$$

where $F' \equiv \delta F/\delta\phi^f$. Then (2.11) becomes

$$(\Lambda(\partial^x) - F'[\phi^f(x)])\delta\phi_g^f(x) = 0. \tag{2.15}$$

This equation has the same form as the one called the stability equation.[4]

In the following, we restrict our consideration to static cases in which $f(x)$ (and therefore $\phi^f(x)$) is independent of time [i.e., $f(x) = f(\mathbf{x})$]. In these cases the time-translational invariance of the theory leads us to conclude that the expansion coefficients in the dynamical map (2.1) and the self-consistent potential have the following forms;

$$C(x;y) = C(\mathbf{x},\mathbf{y},t_x - t_y), \tag{2.16}$$

$$C(x;y,z) = C(\mathbf{x},\mathbf{y},\mathbf{z},t_x - t_y,t_x - t_z), \quad \text{etc.}, \tag{2.17}$$

$$V^f(x,y) = V^f(\mathbf{x},\mathbf{y};t_x - t_y). \tag{2.18}$$

Equation (2.7) together with (2.13) leads us to the following form of $\delta\phi^f_g$:

$$\delta\phi^f_g(x) = Z^{\frac{1}{2}}u(\mathbf{x},\mathbf{k})\frac{e^{-i\omega_k t_x}}{\sqrt{2\omega_k}} = \int d^4y \frac{\delta\phi^f(x)}{\delta f(y)}\frac{e^{i\mathbf{k}\cdot\mathbf{y} - i\omega_k t_y}}{\sqrt{2(2\pi)^3\omega_k}}Z^{\frac{1}{2}}. \tag{2.19}$$

It should be noted that the normalization of $u(\mathbf{x},\mathbf{k})$ is determined by the asymptotic condition: $u(\mathbf{x},\mathbf{k}) \to \exp(i\mathbf{k}\mathbf{x})/(2\pi)^{3/2}$ for $|\mathbf{x}| \to \infty$.

Now (2.11) leads to the eigenvalue equation

$$\Lambda(\boldsymbol{\partial},\omega_k)u(\mathbf{x},\mathbf{k}) = \int d^3y\, V^f(\mathbf{x},\mathbf{y};\omega_k)u(\mathbf{y},\mathbf{k}), \tag{2.20}$$

where

$$\Lambda(\boldsymbol{\partial},\omega) = \Lambda(\partial)|_{\partial_0 = -i\omega} \tag{2.21}$$

and $V^f(\mathbf{x},\mathbf{y},\omega)$ is defined by

$$V^f(x,y) = \frac{1}{2\pi}\int d\omega \exp[-i\omega(t_x - t_y)]V^f(\mathbf{x},\mathbf{y};\omega). \tag{2.22}$$

In the next section, we study Eq. (2.11) [or (2.20)] by using the tree approximation. A general consideration without any use of the approximation is presented in Sec. 4.

3. THE QUANTUM COORDINATES. I. THE TREE APPROXIMATION

When the tree approximation is made, $\phi^f(x)$ satisfies the Euler equation (1.13):

$$\Lambda(\partial)\phi^f(x) = F[\phi^f(x)]. \tag{3.1}$$

In this case, the self-consistent potential $V^f(x,z)$ becomes

$$V^f(x,z) = F'[\phi^f(x)]\delta(x - z), \tag{3.2}$$

where

$$F' = \frac{\delta F}{\delta\phi^f}. \tag{3.3}$$

Thus, Eq. (2.11) becomes (2.15), i.e.,

$$(\Lambda(\partial) - F'[\phi^f(x)])\delta\phi^f_g(x) = 0. \tag{3.4}$$

Since the renormalization factor Z is equal to 1 when the tree approximation is used, (2.19) reads as

$$\delta\phi^f_g(x) = [2\omega_k]^{-\frac{1}{2}}u(\mathbf{x},\mathbf{k})e^{-i\omega_k t_x} \tag{3.5}$$

$$= [2\omega_k(2\pi)^3]^{-\frac{1}{2}}\int d^4y \frac{\delta\phi^f}{\delta f(y)} \times \exp(i\mathbf{k}\mathbf{y} - i\omega_k t_y). \tag{3.6}$$

Now (3.4) becomes the eigenvalue equation

$$\{\Lambda(\boldsymbol{\partial},\omega_k) - F'[\phi^f(x)]\}u(\mathbf{x},\mathbf{k}) = 0. \tag{3.7}$$

Furthermore, this eigenvalue equation admits solutions with discrete energies:

$$\{\Lambda(\boldsymbol{\partial},\omega_i) - F'[\phi^f(x)]\}u_i(\mathbf{x}) = 0. \tag{3.8}$$

We can choose $u(\mathbf{x},\mathbf{k})$ and $u_i(\mathbf{x})$ to form an orthonormalized complete set of solutions of the eigenvalue equations (3.7) and (3.8):

$$\int d^3x\, u^*(\mathbf{x},\mathbf{k})u(\mathbf{x},\mathbf{l}) = \delta(\mathbf{k} - \mathbf{l}), \tag{3.9}$$

$$\int d^3x\, u_i^*(\mathbf{x})u_j(\mathbf{x}) = \delta_{ij}, \tag{3.10}$$

$$\int d^3x\, u_i^*(\mathbf{x})u(\mathbf{x},\mathbf{k}) = 0. \tag{3.11}$$

We then have the sum rule

$$\int d^3k u^*(\mathbf{x},\mathbf{k})u(\mathbf{y},\mathbf{k}) + \sum_i u_i^*(\mathbf{x})u_i(\mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}). \tag{3.12}$$

Let us now recall that, since Eq. (1.7) is translationally invariant, $f(\mathbf{x} + \mathbf{a})$ satisfies (1.7) when $f(\mathbf{x})$ does. The replacement $f(\mathbf{x}) \to f(\mathbf{x} + \mathbf{a})$ induces the spatial translation $\mathbf{x} \to \mathbf{x} + \mathbf{a}$ in $\phi^f(\mathbf{x})$:

$$\phi^{f(\mathbf{y}+\mathbf{a})}(\mathbf{x}) = \phi^{f(\mathbf{y})}(\mathbf{x} + \mathbf{a}), \tag{3.13}$$

which gives

$$\int d^4y \frac{\delta\phi^f(x)}{\delta f(y)}\partial_i f(y) = \partial_i\phi^f(x), \quad i = 1,2,3. \tag{3.14}$$

Since $\partial_i f(x)$ also satisfies Eq. (1.7) for $f(x)$, (3.4) gives

$$\{\Lambda(\boldsymbol{\partial},\omega) - F'[\phi^f(x)]\}\partial_i\phi^f(\mathbf{x}) = 0, \quad \text{for } \omega = 0. \tag{3.15}$$

We thus see that $\partial_i\phi^f$ $(i = 1,2,3)$ *are the zero energy solutions of the eigenvalue equation* (3.8).

Let us now define

$$V_{ij} = \int d^3x \partial_i\phi^f\partial_j\phi^f. \tag{3.16}$$

We can then write as

$$u_i(\mathbf{x}) = \sum_{j=1}^{3}(V^{-\frac{1}{2}})_{ij}\partial_j\phi^f \quad (i = 1,2,3) \tag{3.17}$$

because these functions satisfy, not only the eigenvalue equation (3.8), but also the orthonormalization condition (3.10). This result shows how the translational invariance of the theory induces the translational modes; as soon as the creation of the extended object breaks the translational symmetry, the dynamics naturally creates the translation modes in order to preserve the translational invariance (this point is further clarified later). *Existence of the translation modes is a particular feature of the self-consistent potential induced by the extended object, which itself is self-consistently created in the quantum many-body system.* When $f(\mathbf{x})$ depends only on two components (say x_1 and x_2) of $\mathbf{x}$, there appear only two translation modes because $\partial_3\phi^f = 0$. When $f(\mathbf{x})$ depends on only one component (say x_1), there appears only one translation mode. The number of translation modes is the same as the number of translation symmetries that are spontaneously broken by the creation of the extended object. In the following, we assume that $f(\mathbf{x})$ depends on x_1, x_2, and x_3. Our notation now is the following: u_i with $i = 1,2,3$ are the wave functions of the translation modes while u_i with $i > 3$ are the wave functions of other bound states.

It is convenient to introduce the field

$$\chi^0(x) = \int d^3k \frac{1}{\sqrt{2\omega_k}} \{u(\mathbf{x},\mathbf{k})\alpha(\mathbf{k})e^{-i\omega_k t} + u^*(\mathbf{x},\mathbf{k})\alpha^\dagger(\mathbf{k})e^{i\omega_k t}\} + \sum_i{}' \frac{1}{\sqrt{2\omega_i}} \{u_i(\mathbf{x})\alpha_i e^{-i\omega_i t} + u_i^*(\mathbf{x})\alpha_i^\dagger e^{i\omega_i t}\} \tag{3.18}$$

where Σ' means that the translation modes are omitted in the summation. It is obvious that

$$\{\Lambda(\partial) - F'[\phi^f(x)]\}\chi^0(x) = 0. \tag{3.19}$$

The fields associated with the translational modes are denoted by

$$\sum_{i=1,2,3} u_i(\mathbf{x})q_i = \sum_{i=1,2,3}\sum_{j=1,2,3} \partial_i\phi^f(\mathbf{x})\eta_{ij}q_j, \tag{3.20}$$

the time derivative of which gives

$$\sum_{i=1,2,3} u_i(\mathbf{x})p_i = \sum_{i=1,2,3}\sum_{j=1,2,3} \partial_i\phi^f(\mathbf{x})\eta_{ij}p_j. \tag{3.21}$$

Here

$$\eta_{ij} = (V^{-\frac{1}{2}})_{ij}. \tag{3.22}$$

The operators q_i and p_i are the canonical coordinates and momenta that satisfy

$$[q_i,p_j] = i\delta_{ij}, \quad i,u = 1,2,3 \tag{3.23}$$

and

$$p_i = \frac{d}{dt}q_i. \tag{3.24}$$

In the following, $q_i (i = 1,2,3)$ are called the "quantum coordinate" of the soliton. The sum rule (3.12) gives the commutation relation

$$\Big[\chi^0(x) + \sum_i u_i(\mathbf{x})q_i, \dot{\chi}^0(y) + \sum_i u_i(\mathbf{y})\dot{q}_i\Big]\delta(t_x - t_y) = i\delta(x-y). \tag{3.25}$$

Now the Hilbert space is enlarged to the direct product of the Fock space associated with $\{\alpha(\mathbf{k}),\alpha_i; i>3\}$ and a representation for three sets of canonical variables $(q_i,p_i; i = 1,2,3)$. The dynamical map (2.1) should be replaced by the following one:

$$\psi^f(x) = \phi^f(x) + [\chi^0(x) + (\mathbf{Q}\cdot\boldsymbol{\partial})\phi^f(x)] + \cdots, \tag{3.26}$$

where

$$Q_i = \sum_j \eta_{ij}q_j. \tag{3.27}$$

The inclusion of new operators α_i $(i>3)$ and $(q_i,p_i; i = 1,2,3)$ is required from the canonical commutation relation

$$\Big[\psi^f(x), \frac{\partial}{\partial t}\psi^f(y)\Big]\delta(t_x - t_y) = i\delta(x-y). \tag{3.28}$$

The simple boson-transformed ψ^f would be satisfactory if we only require the Heisenberg equation to be satisfied. However, the presence of the extended object induces the self-consistent potential that influences the quantum states and creates the bound states of the φ^0-quantum and the extended objects. Then, the canonical relation, that may be interpreted as a completeness relation, requires the inclusion of these bound state modes. Indeed, the dynamical map (3.26) together with (3.25) gives (3.28), since the higher order normal product terms represented by the dots in (3.26) should not contribute to the left-hand side of (3.28) when the tree approximation is used.

Comparing (3.26) with (2.1), we see that *the canonical commutator* (3.28) *requires the following replacement in the dynamical map* (2.1):

$$\int d^4y \frac{\delta\phi^f(x)}{\delta f(y)}\varphi^0(x) \to [\chi^0(x) + (\mathbf{Q}\cdot\boldsymbol{\partial})\phi^f(x)]. \tag{3.29}$$

It is shown later how the Heisenberg equation is satisfied by the new $\psi^f(x)$.

To calculate the $\varphi^0(y)\varphi^0(z)$-term in the dynamical map (2.1), we again replace f by $f + \epsilon g$ in (3.4). We then have

$$\{\Lambda(\partial^x) - F'[\phi^f(x)]\}\int d^4y\, d^4z\, g(y)g(z)\frac{\delta}{\delta f(y)}\frac{\delta}{\delta f(z)} \times \phi^f(x) = \int d^4y \int d^4z\, F''[\phi^f(x)]\frac{\delta\phi^f(x)}{\delta f(z)}\frac{\delta\phi^f(x)}{\delta f(y)}g(y)g(z), \tag{3.30}$$

ıere

$$F''[\phi^f] = \frac{\delta F'[\phi^f]}{\delta\phi^f}. \tag{3.31}$$

Then (3.29) leads to the following replacement in the dynamical map (2.1):

$$\int d^4y \int d^4z\, c(x;y,z){:}\varphi^0(y)\varphi^0(z){:} = 1/2\int d^4y\int d^4z{:}\varphi^0(y)\varphi^0(z){:}\frac{\delta}{\delta f(y)}\frac{\delta}{\delta f(z)}\phi^f(x) \to 1/2\int d^4y\, G(x,y)F''[\phi^f(y)]{:}[\chi^0(y) + (\mathbf{Q}\cdot\boldsymbol{\partial})\phi^f(y)]^2{:}, \tag{3.32}$$

where $G(x,y)$ is the Green's function defined by

$$\{\Lambda(\partial^x) - F'[\phi^f(x)]\}G(x,y) = \delta(x-y). \tag{3.33}$$

On the other hand, since Eq. (3.19) is translationally invariant, it leads to

$$\{\Lambda(\partial) - F'[\phi^f(x)]\}\partial_i\chi^0(x) = F''[\phi^f(x)]\partial_i\phi^f(\mathbf{x})\cdot\chi^0(x). \tag{3.34}$$

Similarly, (3.15) leads to

$$\{\Lambda(\boldsymbol{\partial},\omega) - F'[\phi^f(x)]\}\partial_i\partial_j\phi^f(\mathbf{x}) = F''[\phi^f(x)]\partial_i\phi^f(\mathbf{x})\partial_j\phi^f(\mathbf{x}), \quad \text{for } \omega = 0. \tag{3.35}$$

These equations give

$$\partial_i \chi^0 = \int d^4y\, G(x,y) F''[\phi^f(y)] \partial_i \phi^f(y) \chi^0(y), \qquad (3.36)$$

$$\partial_i \partial_j \phi^f(\mathbf{x}) = \int d^4y\, G(x,y) F''[\phi^f(y)] \partial_i \phi^f(y) \partial_j \phi^f(y). \qquad (3.37)$$

Thus (3.32) becomes

$$\int d^4y \int d^4z\, c(x;y,z) : \varphi^0(y) \varphi^0(z) : \rightarrow 1/2 \int d^4y\, G(x,y) F''[\phi^f(y)] : (\chi^0(y))^2 : + 1/2 (\mathbf{Q}\cdot\boldsymbol{\partial})^2 \phi^f(\mathbf{x}) + (\mathbf{Q}\cdot\boldsymbol{\partial}) \chi^0(x) + \cdots, \qquad (3.38)$$

where the dots stand for terms containing $\dot{\mathbf{Q}}$. Thus, the dynamical map (2.1) is replaced by

$$\psi^f(x) = [1 + (\mathbf{Q}\cdot\boldsymbol{\partial}) + 1/2(\mathbf{Q}\cdot\boldsymbol{\partial})^2 + \cdots]\phi^f(x) + [1 + (\mathbf{Q}\cdot\boldsymbol{\partial}) + \cdots]\chi^0(x) + 1/2 \int d^4y\, G(x,y) F''[\phi^f(y)] : (\chi^0(y))^2 : + \cdots.$$

Repeating similar arguments we can show that

$$\begin{aligned}\psi^f(x) &= \Big[\sum_n \frac{1}{n!} (\mathbf{Q}\cdot\boldsymbol{\partial})^n \Big] \times [\phi^f(\mathbf{x}) + \chi^0(x)] + \tfrac{1}{2} \int d^4y\, G(x;y) F''[\phi^f(y)] : (\chi^0(y))^2 : + \cdots] \\ &= \phi^f(\mathbf{x} + \mathbf{Q}) + \chi^0(\mathbf{x} + \mathbf{Q}, t) + 1/2 \int d^4y\, G(\mathbf{x} + \mathbf{Q}, t; y) F''[\phi^f(y)] : (\chi^0(y))^2 : + \cdots. \end{aligned} \qquad (3.39)$$

Note that *the quantum coordinate* $\mathbf{Q}$ *without time derivative appears only through the form* $\mathbf{x} + \mathbf{Q}$. This means that, as soon as the creation of the soliton breaks the translational symmetry, there appears the quantum coordinate that plays the role of the Goldstone modes. In other words, *the spatial translation* $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{a}$ *is induced by the transformation*:

$$\mathbf{Q} \rightarrow \mathbf{Q} + \mathbf{a}. \qquad (3.40)$$

This is the form of the rearrangement of the spacial translational symmetry. The expression (3.39) shows that the quantum coordinate $\mathbf{Q}$ plays the role of the collective coordinate introduced by Sakita *et al.*[6]

We use the Schrödinger representation for the quantum coordinate. Thus, the vacuum state with the soliton is denoted by $|0;\xi\rangle$ that is the eigenstate of $\mathbf{Q}$:

$$\mathbf{Q}|0;\xi\rangle = \xi|0;\xi\rangle. \qquad (3.41)$$

Similarly, a state of one χ^0-quantum is expressed by $|\mathbf{k};\xi\rangle$ or $|i;\xi\rangle$ that satisfy

$$\mathbf{Q}|\mathbf{k};\xi\rangle = \xi|\mathbf{k};\xi\rangle. \qquad (3.42)$$

The eigenvalue ξ determines the position of the soliton. When the size of the soliton (or the extended object) is so large that the quantum fluctuation of the position becomes negligibly small, the extended object behaves as a classical object, while the extended object behaves as a quantum mechanical object when the fluctuation of quantum coordinate becomes as large as the size of the system.

Let us now ask how the higher order terms in (3.39) are determined. This is formulated as follows. First, define

$$n^f(x) = \psi^f(x) - \phi^f(\mathbf{x}), \qquad (3.43)$$

and rewrite the Heisenberg equation as

$$\{\Lambda(\partial) - F'[\phi^f(x)]\} n^f(x) = F_n[n^f(x)], \qquad (3.44)$$

where

$$F_n[n^f(x)] = F[\psi^f(x)] - F[\phi^f(\mathbf{x})] - F'[\phi^f(x)] n^f(x). \qquad (3.45)$$

This leads to the integral equation

$$n^f(x) = \chi^0(x) + \mathbf{Q}\cdot\boldsymbol{\partial}\phi^f(x) + \int d^4y\, G(x,y) F_n[n^f(y)], \qquad (3.46)$$

where the Green's function $G(x,y)$ is defined in (3.33). Solving (3.46) by successive iteration, we can determine the higher order terms in the dynamical map (3.39).

4. THE QUANTUM COORDINATES. II. GENERAL CONSIDERATION

In this section, we study the interaction between the φ^0-quantum and a static soliton without using any approximation.

To do this we come back to Eq. (2.20) for $u(\mathbf{x},\mathbf{k})$. Recalling (1.8), we see that (2.20) is rewritten as the following integral equation:

$$u(\mathbf{x},\mathbf{k}) = \frac{e^{i\mathbf{k}\mathbf{x}}}{(2\pi)^{3/2}} + \Lambda^{-1}(\boldsymbol{\partial},\omega_k) \int d^3y\, V^f(\mathbf{x},\mathbf{y};\omega_k) u(\mathbf{y},\mathbf{k}), \qquad (4.1)$$

where the asymptotic condition is

$$u(\mathbf{x},\mathbf{k}) \rightarrow \frac{e^{i\mathbf{k}\mathbf{x}}}{(2\pi)^{3/2}}, \quad \text{for } |x| \rightarrow \infty.$$

Equation (2.20) admits also some solutions with discrete energies:

$$\Lambda(\boldsymbol{\partial},\omega_i) u_i(\mathbf{x}) = \int d^3y\, V^f(\mathbf{x},\mathbf{y};\omega_i) u_i(\mathbf{y}). \qquad (4.2)$$

In order to study the orthogonality relations and completeness, we consider the following eigenvalue equation:

$$\Lambda(\boldsymbol{\partial},E) v(\mathbf{x}) = \int d^3y\, V^f(\mathbf{x},\mathbf{y};\omega) v(\mathbf{y}). \qquad (4.3)$$

Here ω is regarded as a parameter. This has a continuous eigenvalue $E_k(\omega)$ and discrete ones $E_i(\omega)$. The eigenfunctions are denoted by $v_k(\mathbf{x},\omega)$ and $v_i(\mathbf{x},\omega)$, respectively. The continuous eigenfunction is obtained from the integral equation

$$v_k(\mathbf{x},\omega) = \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}} + \Lambda^{-1}(\boldsymbol{\partial},E_k) \int d^3y\, V^f(\mathbf{x},\mathbf{y}:\omega) v_k(\mathbf{y},\omega). \qquad (4.4)$$

Equation (4.4) can be regarded as a wave equation with a nonlocal potential $V^f(\mathbf{x},\mathbf{y}:\omega)$. Therefore, the functions

$v_k(x,\omega)$ satisfy the orthogonality relations

$$\int d^3x\, v_k^*(\mathbf{x},\omega)v_q(\mathbf{x},\omega) = \delta(\mathbf{k}-\mathbf{q}). \tag{4.5}$$

We can also normalize $v_i(\mathbf{x},\omega)$ as

$$\int d^3x\, v_i^*(\mathbf{x},\omega)v_j(\mathbf{x},\omega) = \delta_{ij}, \tag{4.6}$$

$$\int d^3x\, v_i^*(\mathbf{x},\omega)v_k(\mathbf{x},\omega) = 0. \tag{4.7}$$

The completeness relation is

$$\int d^3k v_k(\mathbf{x},\omega)v_k^*(\mathbf{y},\omega) + \sum_i v_i(\mathbf{x},\omega)\, v_i^*(\mathbf{y},\omega) = \delta(\mathbf{x}-\mathbf{y}). \tag{4.8}$$

It should be noted that the eigenvalues of Eq. (2.20), i.e., ω_k and ω_i cover all of the solutions of the equation $E(\omega) = \omega$. Comparing (2.20) and (4.1) with (4.3) and (4.4), respectively, we find that

$$u(\mathbf{x},\mathbf{k}) = v_k(\mathbf{x},\omega_k): \tag{4.9}$$

$$u_i(\mathbf{x}) = v_i(\mathbf{x},\omega_i), \quad \text{with } \omega_i = E_i(\omega_i). \tag{4.10}$$

Then $u(\mathbf{x},\mathbf{k})$ and $u_i(\mathbf{x})$ satisfy the orthogonality relation with respect to the following metric tensor Γ:

$$\Gamma(\omega,\omega':\mathbf{x},\mathbf{y}) = \int d^3k\, v_k(\omega,\mathbf{x})v_k^*(\omega',\mathbf{y}) + \sum_i v_i(\omega,\mathbf{x})\, v_i^*(\omega',\mathbf{y}). \tag{4.11}$$

In other words, we have

$$\int d^3x \int d^3y\, u^*(\mathbf{x},\mathbf{k})\Gamma(\omega_k,\omega_q;\mathbf{x},\mathbf{y})u(\mathbf{y},\mathbf{q}) = \delta(\mathbf{k}-\mathbf{q}),$$

$$\int d^3x \int d^3y\, u_i^*(\mathbf{x})\Gamma(\omega_i,\omega_j;\mathbf{x},\mathbf{y})u_j(\mathbf{y}) = \delta_{ij}, \tag{4.12}$$

$$\int d^3x \int d^3y\, u^*(\mathbf{x},\mathbf{k})\Gamma(\omega_k,\omega_i;\mathbf{x},\mathbf{y})u_i(\mathbf{y}) = 0.$$

We also have

$$\delta(\mathbf{x},\mathbf{y}) \equiv \int d^3k\, u(\mathbf{x},\mathbf{k})u^*(\mathbf{y},\mathbf{k}) + \sum_i u_i(\mathbf{x})u_i(\mathbf{y}) \tag{4.13}$$

$$= \int d^3k v_k(\mathbf{x},\omega_k)\, v_k^*(\mathbf{y},\omega_k) + \sum_i v_i(\mathbf{x},\omega_i)\, v_i^*(\mathbf{y},\omega_i). \tag{4.14}$$

Note that, if $V^f(\mathbf{x},\mathbf{y};\omega)$ is independent of ω, $\delta(\mathbf{x},\mathbf{y})$ reduces to the δ-function. The ω-dependence of V^f appears when the loop diagrams are considered. Therefore the nonlocality of $\delta(\mathbf{x},\mathbf{y})$ is due to the contribution from loop diagrams. For convenience, we also introduce a function $\Delta^f(x,y)$ defined by

$$\Delta^f(x,y) = \int d^3k\, u(\mathbf{x},\mathbf{k})u^*(\mathbf{y},\mathbf{k})e^{-i\omega_k(t_x-t_y)} + \sum_i u_i(\mathbf{x})u_i(\mathbf{y})e^{-i\omega_i(t_x-t_y)}. \tag{4.15}$$

Among the discrete levels, there always exist certain zero energy levels. First we recall that, the replacement $f(\mathbf{x}) \to f(\mathbf{x}+\mathbf{a})$ induces the spatial translation $\mathbf{x} \to \mathbf{x}+\mathbf{a}$ in any functional of f, i.e.,

$$O^{f(y+a)}(\mathbf{x}) = O^{f(y)}(\mathbf{x}+\mathbf{a}), \tag{4.16}$$

or in a differential form

$$\int d^4y \frac{\delta O^f(x)}{\delta f(y)} \nabla f(y) = \nabla O^f(x). \tag{4.17}$$

Operating ∂_i $(i = 1,2,3)$ on both sides of (2.5), we obtain

$$\begin{aligned}\Lambda(\partial)\partial_i\phi^f(x) &= \partial_i J^f(x) \\ &= \int d^4y \frac{\delta J^f(x)}{\delta f(y)} \partial_i f(y) \\ &= \int d^4y \int d^4z \frac{\delta J^f(x)}{\delta\phi^f(z)} \times \frac{\delta\phi^f(z)}{\delta f(y)} \partial_i f(y),\end{aligned}$$

which reads as

$$\Lambda(\partial)\partial_i\phi^f(x) = \int d^4y\, V^f(x,y)\partial_i\phi^f(y). \tag{4.18}$$

We thus see that $\partial_i\phi^f(x)$ $(i = 1,2,3)$ are the zero energy solutions of the eigenvalue equation (4.2). These are the translation modes.

Let us now define

$$V_{ij} = \int d^3x\, \partial_i\phi^f(\mathbf{x})\partial_j\phi^f(\mathbf{y}). \tag{4.19}$$

We can then write as

$$u_i(\mathbf{x}) = \sum_{j=1}^{3} (V^{-\frac{1}{2}})_{ij}\partial_i\phi^f \quad (i = 1,2,3), \tag{4.20}$$

since these functions satisfy the eigenvalue equation (4.2) and the orthonormalization condition (4.12) [see also (4.6) and (4.10)]. When $f(\mathbf{x})$ depends only on two components of $\mathbf{x}$, then there appear only two translation modes, while there appears only one translation mode when $f(\mathbf{x})$ depends only on one component of $\mathbf{x}$.

In the following, we assume that $f(\mathbf{x})$ depends on three components of $\mathbf{x}$. The symbols u_i with $i > 3$, denote the wave functions of the bound states except the translation modes.

Having all of the orthonormalized eigenfunctions of Eq. (4.1) and (4.2), we now introduce the boson field χ^0 according to (3.18). Thus, χ^0 contains all the modes except the translation modes. Obviously we have

$$\Lambda(\partial)\chi^0(x) = \int d^4y\, V^f(x,y)\chi^0(y). \tag{4.21}$$

To construct the boson field associated with the translation modes, we introduce the three sets of canonical operators $(q_i,p_i;\, i = 1,2,3)$ which satisfy (3.23) and (3.24). Then, the boson field of the translation field is given by $\Sigma\ u_i(\mathbf{x})q_i$ which is equal to $(\mathbf{Q}\cdot\boldsymbol{\partial})\phi^f(x)$ where the operator $\mathbf{Q}$ is defined by (3.27) and (3.22).

The sum rule (4.14) gives the commutation relation:

$$\Big[\chi^0(x) + \sum_i u_i(\mathbf{x})q_i, \dot{\chi}^0(y) + \sum_i u_i(y)\dot{q}_i\Big]\delta(t_x-t_y) = i\ \delta(\mathbf{x},\mathbf{y})\delta(t_x-t_y), \tag{4.22}$$

which becomes (3.25) when the tree approximation is used.

According to the consideration in the last section, all of the modes including the bound states should participate in the dynamical map ψ^f. Therefore, the dynamical map (2.1) should be replaced by

$$\psi^f(x) = \phi^f(x) + \left[Z_0^{\frac{1}{2}}\left(i\frac{\partial}{\partial t}\right)\chi^0(x) + (\mathbf{Q}\cdot\boldsymbol{\partial})\phi^f(x)\right] + \dots. \tag{4.23}$$

Here $Z_0^{1/2}[i(\partial/\partial t)\chi^0(x)]$ means

$$Z_0^{1/2}\left(i\frac{\partial}{\partial t}\right)\chi^0(x) = \int d^3y\, Z_0^{\frac{1}{2}}\left(\mathbf{x},\mathbf{y};i\frac{\partial}{\partial t}\right)\chi^0(y), \tag{4.24}$$

with

$$Z_0^{1/2}(\mathbf{x},\mathbf{y};\omega) = \int d^3k\, v_k(\mathbf{x},\omega)Z^{\frac{1}{2}}v_k^*(\mathbf{y},\omega) + \sum_i v_i(\mathbf{x},\omega)Z_i^{1/2}v_i^*(\mathbf{y},\omega) \tag{4.25}$$

and Q is defined by

$$Q_i = Z_Q^{1/2}\sum_j \eta_{ij}q_j \tag{4.26}$$

with Z_Q being the renormalization factor.

The dynamical map (4.23) together with (4.22) gives

$$\left[\psi^f(x), \frac{\partial}{\partial t}\psi^f(y)\right]\delta(t_x - t_y) = \delta(t_x - t_y)Z\left(i\frac{\partial}{\partial t_x}\right)\Delta^f(x,y) + \cdots \tag{4.27}$$

where

$$Z(\omega) = Z_0(\omega) + \sum_{i=1}^{3} v_i(\mathbf{x})Z_Q v_i^*(\mathbf{x}) \tag{4.28}$$

and the dots stand for the contributions due to the higher order normal products.

When the loop diagrams are disregarded, we have

$$\delta(t_x - t_y)Z\left(i\frac{\partial}{\partial t_x}\right)\Delta^f(x,y)\underset{\text{tree}}{\rightarrow}\delta(x-y). \tag{4.29}$$

Therefore, the terms denoted by the dots in (4.27) are expected to compensate the difference between the left- and right-hand sides of (4.29).

To calculate the $\varphi^0(y)\varphi^0(z)$-term in the dynamical map (2.1), we take the second functional derivative of (2.5) with respect to f. We then have

$$\Lambda(\partial_x)\frac{\delta^2\phi^f(x)}{\delta f(y)\delta f(z)} = \int d^4\xi\, V^f(x,\xi)\frac{\delta^2\phi^f(\xi)}{\delta f(y)\delta f(z)} + \int d^4\xi_1 d^4\xi_2\, V_2^f(x;\xi_1,\xi_2)\frac{\delta\phi^f(\xi_1)}{\delta f(y)}\frac{\delta\phi^f(\xi_2)}{\delta f(z)}, \tag{4.30}$$

where $V_2^f(x;\xi_1,\xi_2) = \delta^2 J^f(x)/\delta\phi^f(\xi_1)\delta\phi^f(\xi_2)$. On the other hand, when the Hilbert space is enlarged, the rule to include new freedom in the dynamical map is the following replacement:

$$\int d^4y\frac{\delta\phi^f(x)}{\delta f(y)}\varphi^0(y) \rightarrow Z_0^{1/2}\left(i\frac{\partial}{\partial t}\right)\chi^0(x) + \mathbf{Q}\cdot\nabla\phi^f(x). \tag{4.31}$$

Thus (2.30) together with (2.1) leads to

$$\begin{aligned}\int d^4y\int d^4z\, c(x;y,z){:}\varphi^0(y)\varphi^0(z){:} &= \frac{1}{2}\int d^4y\, d^4z{:}\varphi^0(y)\varphi^0(z){:}\, V_2^f(x;y,z)\\ &\rightarrow \frac{1}{2}\int d^4y\, G(x,y)\, V_2^f(y;\xi_1,\xi_2)\\ &\quad {:}\left(Z_0^{1/2}\left(i\frac{\partial}{\partial t_{\xi_1}}\right)\chi^0(\xi_1) + \mathbf{Q}\nabla\phi^f(\xi_1)\right)\\ &\quad\times\left(Z_0^{1/2}\left(i\frac{\partial}{\partial t_{\xi_2}}\right)\chi^0(\xi_2) + \mathbf{Q}\nabla\phi^f(\xi_2)\right){:},\end{aligned} \tag{4.32}$$

where $G(x,y)$ is the Green's function defined by

$$\Lambda(\partial^x)G(x,y) - \int d^4z\, V^f(x,z)G(z,y) = \delta(x-y). \tag{4.33}$$

The inclusion of new operators should be made inside the normal product symbol, since the inclusion of them should not change the normalization factor of scattering states.

By use of the translational invariance of the theory, we obtain from (4.21)

$$\Lambda(\partial_x)\partial_i\chi^0(x) = \int d^4y\, V^f(x,y)\partial_i\chi^0(y) + \int d^4y\, d^4z\, V_2^f(x;y,z)\partial_i\phi^f(z)\chi^0(y). \tag{4.34}$$

Similarly (4.18) leads to

$$\Lambda(\partial)\partial_i\partial_j\phi^f(x) = \int d^4y\, V^f(x,y)\partial_i\partial_j\phi^f(y) + \int d^4y\, d^4z\, V_2^f(x;y,z)\partial_i\phi^f(y)\partial_j\phi^f(z). \tag{4.35}$$

These equations give

$$\partial_i\chi^0(x) = \int d^4y\, d^4\xi_1 d^4\xi_2\, G(x,y)\, V_2^f(y;\xi_1,\xi_2)\times\partial_i\phi^f(\xi_1)\chi^0(\xi_2), \tag{4.36}$$

$$\partial_i\partial_j\phi^f(x) = \int d^4y d^4\xi_1 d^4\xi_2\, G(x,y)\, V_2^f(y;\xi_1,\xi_2)\times\partial_i\phi^f(\xi_1)\partial_j\phi^f(\xi_2). \tag{4.37}$$

Thus (4.32) becomes

$$\int d^4y\int d^4z\, c(x;y,z){:}\,\varphi^0(y)\varphi^0(z){:} \rightarrow \frac{1}{2}G\, V_2^f{:}(Z_0^{1/2}\chi^0)^2{:} + \frac{1}{2}(\mathbf{Q}\cdot\nabla)^2\phi^f(x) + (\mathbf{Q}\cdot\nabla)\, Z_0^{1/2}\chi^0 + \cdots, \tag{4.38}$$

where the integration symbol is omitted and the dots stand for terms containing $\dot{\mathbf{Q}}$. Thus the dynamical map is

$$\psi^f(x) = \left[1 + (\mathbf{Q}\cdot\nabla) + \frac{1}{2}(\mathbf{Q}\cdot\nabla)^2 + \cdots\right]\phi^f(x)$$
$$+ [1 + (\mathbf{Q}\cdot\nabla) + \cdots]Z_0^{1/2}\chi^0(x)$$
$$+ \frac{1}{2}\int d^4\xi_1 d^4\xi_2 d^4y\, G(x,y)\, V_2^f(y;\xi_1,\xi_2)$$
$$:Z_0^{1/2}\chi^0(\xi_1)Z_0^{1/2}\chi^0(\xi_2): + \cdots. \quad (4.39)$$

Repeating similar arguments we can show [cf. (3.39)] that

$$\psi^f(x) = \phi^f(\mathbf{x} + \mathbf{Q},t) + Z_0^{1/2}\chi^0(\mathbf{x} + \mathbf{Q},t)$$
$$+ \frac{1}{2}\int d^4y\, G(\mathbf{x} + \mathbf{Q},t;y)\int d^4\xi_1 d^4\xi_2$$
$$\times V_2^f(y;\xi_1,\xi_2)$$
$$:Z_0^{1/2}\chi^0(\xi_1)Z_0^{1/2}\chi^0(\xi_2): + \cdots. \quad (4.40)$$

Note that $\mathbf{Q}$ *without time derivative appear only through the combination* $\mathbf{x} + \mathbf{Q}$. Therefore,*the spacial translation* $\mathbf{x} \to \mathbf{x} + \mathbf{a}$ *is induced by the* $\mathbf{Q}$*- translation*, i.e., $\mathbf{Q} \to \mathbf{Q} + \mathbf{a}$, dynamical rearrangement of the spatial translational symmetry. Thus, $\mathbf{Q}$ is called the *quantum coordinate.* We use the Schrödinger representation for the quantum coordinate, and therefore, have relations such as (3.41) and (3.42). Thus the soliton,$\phi^f(\mathbf{x} + \mathbf{Q})$, behaves as a classical object or a quantum mechanical object according to its size and the experimental conditions.

A systematic method for determining the dynamical map (4.40) is formulated as follows. Calculate first $\phi^f(x)$ by the simple boson transformation $\varphi^0 \to \varphi^0 + f$. Second, define

$$n^f(x) = \psi^f(x) - \phi^f(\mathbf{x}). \quad (4.41)$$

Then, the Heisenberg equation leads to

$$\int d^4y[\Lambda(\partial^x)\delta(x-y) - V^f(x,y)]n^f(y) = F_n[x:n^f] \quad (4.42)$$

where

$$F_n[x;n^f] \equiv F[\psi^f(x)] - F[\phi^f(\mathbf{x})] - \int d^4y V^f(x,y)n^f(y). \quad (4.43)$$

Use of the Green's function defined in (4.33) rewrites the Heisenberg equation (4.42) in the following integral form:

$$n^f(x) = Z_0^{1/2}\chi^0(x) + (\mathbf{Q}\cdot\partial)\phi^f(x) + \int d^4y(x,y)F_n(y;n^f). \quad (4.44)$$

Solving (4.44) by successive iteration, we can determine the dynamical map of $n^f(x)$, and therefore, also of $\psi^f(x)$.

5. SUMMARY

In this paper, we studied the extended objects (or solitons) created by the condensation of a boson in a quantum many-body system. The boson condensation was mathematically treated by the boson transformation. The soliton thus created coexists and interacts with the quanta. It was shown that the soliton creates a self-consistent potential that influences the dynamics of quanta. It is obvious that the position of the whole soliton system has no relevance because it depends only on the choice of the origin of the coordinate. This fact was manifested in our result that shows that the canonical commutation relation of the Heisenberg field naturally introduces a set of quantum coordinates, q_i $(i = 1,2,3)$ and the conjugate momenta p_i. The translation of the position ($x_i \to x_i + a_i$) is induced by the translation of q_i, i.e., $q_i \to q_i + c_i$ with $a_i = \eta_{ij}c_j$, because x_i and q_i always appear through the combination $x_i + \eta_{ij}q_j$. When the quantum fluctuation of q_i is negligibly small, the soliton behaves as a classical object, while it appears to be a quantum mechanical object when the quantum fluctuation of q_i becomes observable. In this way, both the classical and quantum mechanical objects are naturally created by the condensation of bosons. Depending on the size of the object and experimental conditions, the quantum fluctuation of q_i becomes either observable or unobservable.

We are particularly interested in extended objects of a finite size. An object of this kind carries a *self-maintained* enclosed surface singularity, into which the object is confined. In Ref. 8, we have presented a systematic formalism for treatment of surface singularities that are associated with the boson transformation function $f(x)$. It may be worthwhile to point out that, for example, the crystals are not confined to artificial boxes; their boundary surfaces are self-consistently maintained.

It was shown in Ref. 1 that these kinds of singularities (topological singularities) can be created only when the condensed bosons are massless. It is due to this reason that we observe a variety of topological singularities (including the enclosed surface boundary singularities) in many kinds of ordered states; they appear through the condensation of the Goldstone bosons which are massless.

When an object with a self-maintained boundary surface has a large size (e.g., stars, crystals, etc.), it behaves as a classical object under most of the experimental conditions. The different averaged eigenvalues of $\mathbf{q}$ correspond to different positions of the object. When an object has a very small size and the quantum fluctuation of $\mathbf{q}$ becomes observable, then the system may behave as a quantum object. Many micro domains that appear in a variety of ordered states in solid state physics are objects of this kind (i.e., quantum solitons). The large nucleus may be another example of this kind. When we consider an object of extremely small size, we find the MIT bag-type object. In Ref. 9, it was shown that a small object with an enclosed surface singularity behaves as the MIT bag. There, the quantum coordinate was neglected, and therefore, the system behaved like a classical object. Considerations in this paper suggest that a careful treatment replaces x_i by $x_i + \eta_{ij}q_j$, and therefore that the bag behaves like a quantum object.

Summarizing, we find that there always exists a quantum coordinate $\mathbf{q}$ associated with a soliton. *Depending on the observability or unobservability of the quantum fluctuation associated with* $\mathbf{q}$, *the soliton behaves as a quantum soliton or a classical soliton.*

A merit of the boson theory for the soliton lies in the fact that calculation of quantum effects in soliton dynamics is straightforward. For example, the soliton field $\phi^f(\mathbf{x}$

$+ \mathbf{Q})(Q_i = \eta_{ij} q_i)$ contains the quantum effects due to the loop diagrams that appear through the course of solving the integral equation (1.8) by successive iterations. The quantum effects of this kind are contained in the *classical* behavior of solitons described by $\phi^f(\mathbf{x})$. The real quantum behavior appears through the presence of the quantum coordinate $\mathbf{Q}$.

In this paper we considered a simple model, in which the boson field φ^0 does not carry any degree of freedom associated with an internal symmetry. When φ^0 has, for example, isospin, the $\phi^f(\mathbf{x} + \mathbf{Q})$ is expected to carry not only the quantum coordinate, but also a quantum isospin. This will be a problem of future study.

Let us close this paper with a comment on the energy of the soliton created by the boson condensation. Let H denote the Hamiltonian of the Heisenberg field ψ. It has the form

$$H = \int d^3x\, h\,[x;\psi], \tag{5.1}$$

where $h\,[x;\psi]$, is the Hamiltonian density. The boson transformed Hamiltonian H^f is given by

$$H^f = \int d^3x\, h\,[x;\psi^f]. \tag{5.2}$$

When the tree approximation is used, we can prove

$$\lim_{\hbar\to 0} \langle 0|H^f|0\rangle = \int d^3x\, h\,[x;\phi_0^f]. \tag{5.3}$$

The right-hand side is the classical Hamiltonian for the Euler equation (1.13). A proof for (5.3) is the same as the argument[1] used in derivation of the classical Euler equation (1.13) from the Heisenberg equation (1.2).

When the tree approximation is not used, the soliton energy is given by

$$\langle 0|H^f|0\rangle = \int d^3x\, \langle 0|h\,[x;\psi^f]\,|0\rangle. \tag{5.4}$$

It is important to note here that the space integration should be made only after the vacuum expectation value of *boson transformed* Hamiltonian density is calculated.

As is well known, H is weakly equal to the free Hamiltonian:

$$\langle a|H\,|b\,\rangle = \langle a|H_0\,[\varphi^0]\,|b\,\rangle. \tag{5.5}$$

Here $|a\rangle$ and $|b\,\rangle$ are vectors in the Fock space of physical field φ^0, and $H_0\,[\varphi^0]$ is the free Hamiltonian of φ^0. In (5.5), $\langle a|H\,|b\,\rangle$ precisely means

$$\langle a|H\,|b\,\rangle = \int d^3x\, \langle a|h\,\{x;\psi\}|b\,\rangle. \tag{5.6}$$

Let us recall the relation (5.5) is based on the fact that φ^0 is Fourier transformable. Thus, when $f(x)$ is also Fourier transformable, we can generalize (5.5) as

$$\langle a|H^f|b\,\rangle = \langle a|H_0\,[\varphi^0 + f\,]\,|b\,\rangle,$$

which gives the soliton energy

$$\langle 0|H^f|0\rangle = \langle 0|H_0\,[\varphi^0 + f\,]\,|b\,\rangle. \tag{5.7}$$

However, *these relations are not true when f* has a certain topological singularity which *prohibits the Fourier transform of f*. It has been shown[1] that, unless the energy ω_k of the φ^0-quantum vanishes at a certain nonvanishing value of momentum $\mathbf{k}$, $f(\mathbf{x})$ for a *static* soliton always has a singularity that prohibits the Fourier transform of f. It is an easy task to show, by using the (1 + 1)-dimensional $\lambda\phi^4$— or sine-Gordon model with the tree approximation, that the soliton energy given by (5.3) is different from the one obtained from (5.7).

The moral here is that when $f(x)$ is singular, the soliton energy should be calculated, not by (5.7), but by (5.4). However, (5.7) has been used in all of the past calculations of soliton energies in the boson theory. When the domain of singularity associated with $f(x)$ is confined to a finite domain, the relation (5.7) is supposed to give a reasonably approximate result. Therefore, the calculation of the bag energy presented in Ref. 8 seems to be reliable. In application of the boson theory to superconductivity, there is a good reason[10] to believe that the difference between (5.7) and (5.4), which is called a nonlinear effect, is as small as (energy of collective mode)^2/(plasma energy)^2. This is due to the gauge invariance of the theory. On the other hand, the energy of relativistic string was calculated in Ref. 11, where (5.4) was used. The result contained a ultraviolet divergence. It is an open question if this divergence disappears when (5.4) is used.

ACKNOWLEDGMENTS

One of us (G.O.) wishes to thank the Institute of Theoretical Physics at the University of Alberta for a visit during which time this work was done. This work was supported by the National Sciences and Engineering Research Council, Canada and by the Faculty of Science of the University of Alberta.

[1]H. Matsumoto, P. Sodano, and H. Umezawa, Phys. Rev. D **19**, 511 (1979).
[2]G. Oberlechner, M. Umezawa, and Ch. Zenses, Lett. Nuovo Cimento **23**, 641 (1978).
[3]As for the review of solitons, see A.C. Scott, F. Chin, and D. McLaughlin, Proc. IEEE **61**, 1443 (1976).
[4]R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[5]R.F. Dashen, B. Hasslacher, and A. Neveu, Phys Rev. D **10**, 4114, 4130, 4138 (1974); V. Korepin and L. Faddeev, Teor. Nat. Fiz. **25**, 147 (1975) [Theor. Math. Phys. **25**, 1039 (1976)]. J. Goldstone and R. Jackiw, Phys. Rev. D **11**, 1486 (1975); R. Rajaraman, Phys. Rep. C **21**, 227 (1975); J.L. Gervais and A. Neveu, Phys. Rep. C **23**, 237 (1976).
[6]G.C. Branco, B. Sakita, and P. Senjanovic, Phys. Rev. D **10**, 2573, 2582 (1974). J.L. Gervais and B. Sakita, Phys. Rev. D **11**, 2943 (1975); J.L. Gervais, A. Jevicki, and B. Sakita, Phys. Rev. D **12**, 1038 (1975); Phys. Rep. C **21**, 281 (1976).
[7]G. Wentzel, Helv. Phys. Acta. **13**, 269 (1940); D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).
[8]M. Wadati, H. Matsumoto, and H. Umezawa, Phys. Lett. B **73**, 448 (1978); Phys. Rev. D **18**, 520 (1978).
[9]M. Wadati, H. Matsumoto, and H. Umezawa, Phys. Rev. D **18**, 1192 (1978).
[10]H. Matsumoto and H. Umezawa, Fortschr. Physik **24**, 357 (1976).
[11]H. Matsumoto, N.J. Papastamatiou, and H. Umezawa, Phys. Rev. D **14**, 3536 (1976).

Deformations and spectral properties of merons

Basilis Gidas

Department of Mathematics, Rockefeller University, New York, New York 10021
(Received 30 October 1978)

We consider a meron–antimeron pair located at a, b, $\in \mathbf{R}^4$, and show that the spectrum of its stability operator is not bounded below [in precise mathematical terms: The stability operator defined on $C_0^\infty(\mathbf{R}^4 - \{a,b\})$ has a self-adjoint extension, possibly many, all of which are unbounded below]. We regularize a single meron located at the origin by replacing it inside a sphere of radius R_0 and outside a sphere of radius R by "half instantons," and show that for $R \gg R_0$ the regularized configuration continues to be unstable. For R_0 finite and $R = \infty$, we show that the spectrum of the stability operator continues to extend to $-\infty$. We employ a singular transformation to embed $\mathbf{R}^4$ into $S^3 \times \mathbf{R}$ where the meron pair takes a simple form and its stability operator L becomes $\mathrm{L} = -d^2/d\tau^2 + V$, where $\tau \in \mathbf{R}$, and the potential V can be diagonalized in terms of the angular momenta, spin, and isospin of the vector field. The spectrum of L is continuous and extends from -2 to $+\infty$. We determine the number of (generalized) zero eigenmodes of L, and calculate its spectrum explicitly.

I. INTRODUCTION

By a generally accepted philosophy, the phenomenon of quark confinement in quantum chromodynamics (QCD) is caused by the severe infrared singularities of non-Abelian color gauge theories. Polyakov proposed[1] that the infrared behavior of such theories might be governed by solutions of the classical Euclidean field equations. Two particular solutions of the Euclidean Yang–Mills equations have been discovered: the instanton or pseudoparticle,[2] and the meron,[3] which we study here.

From the physical point of view, the instanton yielded a rich structure of the quantum vacuum,[4] and provided an explanation of the famous U(1) problem.[5] From the mathematical point of view, it implied a deep connection between the Yang-Mills (YM) equations and differential geometry, in particular holomorphic vector bundles.[6,7] Much less is known about merons. Callan, Dashen, and Gross have been considered[8] a mechanism of confinement based on meron-like configurations. Glimm and Jaffe have introduced[9] polymerons on a line (see Ref. 10 for their mathematical existence) and studied confinement in a statistical system of merons.[11] Subsequently, they developed[12] a droplet model for the confinement phase transition. Here we study infinitesimal deformations of a meron pair, and show that the pair is unstable in the sense that its stability operator has a negative spectrum. (Warning: This notion of stability is different from Liapunov's notion of stability; furthermore, both notions of stability are different from the stability in the sense of Poincaré,[13] i.e., the spatial components of the energy–momentum tensor are zero. The meron might be stable in the sense of Liapunov.) In fact, we prove that the spectrum of the stability operator is unbounded below, and that the meron pair "falls" into two pure vacua, one of which is given by a singular gauge transformation, and corresponds to (in Nambu's [13a] terminology) "black instantons." Our analysis shows that in the function space of all configurations, the meron is a saddle point which is unstable at least within configurations with spherical symmetry. We regularize a single meron by replacing it inside a sphere of radius R_0 and outside a sphere of radius R by "half instantons," and show that as long as $R \gg R_0$, the regularized configuration continues to be unstable. For R_0 finite and $R = +\infty$, we show that the stability operator continues to be unbounded below. (See Ref. 14 and Note added in proof at the end of the paper.)

We also prove that the meron-antimeron stability operator has exactly eight zero (generalized, i.e., not square integrable) eigenmodes. The corresponding eight parameters are interpreted as the positions of the two merons. This result shows that a meron pair is a "rigid" dipole, i.e., one meron in a meron pair, cannot rotate in the internal space (isospin space) independently of the other meron. The number of independent, zero (generalized) eigenvectors of the stability operator is obtained by embedding $\mathbb{R}^4$ into $S^3 \times \mathbb{R}$, via a singular transformation. That $S^3 \times \mathbb{R}$ is the natural geometry for merons, is indicated by the fact that one meron is invariant (within gauge equivalence) under the subgroup $O(4) \times O(1,1)$ of the Euclidean conformal group $O(5,1)$. The map $\mathbb{R}^4 \to S^3 \times \mathbb{R}$ is a two-parameter family of transformations. Merons located at points $a = (a_0, \mathbf{0})$, $b = (b_0, \mathbf{0})$, or $a = 0$, $b = \infty$, go into the same solution on $S^3 \times \mathbb{R}$, which in terms of differential geometry, is the standard SU(2) spin bundle over S^3. We call it the *meron bundle*. The stability operator $\mathscr{L}$ of the meron bundle has a simple form, i.e., $-d^2/d\tau^2 + V$, $\tau \in R$, the potential V can be diagonalized in terms of the angular momenta, spin, and isospin of the vector field. Thus the spectrum of $\mathscr{L}$, which is continuous and extends from -2 to $+\infty$, can be calculated explicitly. Our

[a] Supported in part by The National Science Foundation, Grant PHY76-17191 and PHY-77-18762.

 0022-2488/79/102097-13$01.00

$S^3 \times \mathbb{R}$ representation of a meron pair and its stability operator, are suitable for defining appropriate stable approximate meron configurations and determining their contributions to the functional integral.

While the relevance of merons to physics has yet to be determined (their stability makes merons ambivalent!), their mathematical aspects seem to be interesting from the point of view of both differential geometry and analysis. The singular transformation we employ in Sec. IV, to embed $\mathbb{R}^4$ into $S^3 \times \mathbb{R}$, has been motivated by the manifestly conformal covariant formalism[22-24] of the Yang–Mills equations, and it appears to be a special case of the well-known "blow-up" or quadratic transformations in algebraic geometry.[15] This blow up process, performed in a manner that preserves the flat conformal structure of the physical space, may be useful in proving the existence of (or even constructing!) $k \geqslant 2$ meron pair solutions of the YM equations. However, the Riemannian structure of the blown up surface is not unique, and its determination seems to be equivalent to proving the existence of infinite energy (action), nonstable solutions of nonlinear partial differential equations. Also, it should be possible to formulate and prove an Atiyah–Singer index theorem for merons on these (noncompact) blow up surfaces. These problems deserve further investigation.

The organization of this paper is as follows: In Sec. II, we establish the conformal properties of merons. In Sec. III, we show the instability of a single meron as well as of a meron pair. We also study the instability properties of a regularized configuration. In Sec. IV, we embed $\mathbb{R}^4$ into $S^3 \times \mathbb{R}$ via a singular transformation, and construct the meron bundle over S^3. In Sec. V, we study the stability operator $\mathscr{L}$ of the meron bundle, calculate the number of independent, zero (generalized) eigenvectors of $\mathscr{L}$, and determine its spectrum explicitly. Finally, in the Appendix, we summarize some technical calculations that are needed to go from our meron bundle to the standard meron–antimeron pair solution on $\mathbb{R}^4$.

II. CONFORMAL PROPERTIES OF MERONS

The pure Yang–Mills (YM) equations in four space–time dimensions read

$$D_\mu F_{\mu\nu} \equiv \partial_\mu F_{\mu\nu} + [A_\mu, F_{\mu\nu}] = 0, \tag{2.1}$$

where D_μ is the covariant derivative, and the field tensor $F_{\mu\nu}$ (the curvature) is given in terms of the vector field A_μ (connection) by ("structure equation" in the language of differential geometry)

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{2.2}$$

For an SU(2) gauge group, we set $A_\mu = (1/2i) A^a_\mu \sigma^a$, $F_{\mu\nu} = (1/2i) F^a_{\mu\nu} \sigma^a$, $A^a_\mu = i\mathrm{Tr}(A_\mu \sigma^a)$, $F^a_{\mu\nu} = i\mathrm{Tr}(F_{\mu\nu}\sigma^a)$ where σ^a, $a = 1,2,3$ are the Pauli matrices. Equations (2.1) are associated with the action

$$S(A) = \tfrac{1}{2}\int d^4x \,\mathrm{Tr} F_{\mu\nu} F_{\mu\nu} = \tfrac{1}{4}\int d^4x \, F^a_{\mu\nu} F^a_{\mu\nu}. \tag{2.3}$$

(Throughout this paper, Greek indices $\mu,\nu,\kappa,\lambda,\dots,\alpha,\beta,\gamma,\cdots$ run from 0 to 3, and 1 to 4, respectively. Latin indices $a,b,c,\dots,i,j,k,\cdots$ run from 1 to 3. Repeated indices are always summed over.) The finite action configurations fall into homotopic classes labelled by the elements of π_3 [SU(2)$\simeq Z$, which gives rise to the "topological charge," or "Pontryagin index"

$$k = \int d^4x \, Q(x), \tag{2.4a}$$

$$Q(x) = \frac{1}{16\pi^2} \mathrm{Tr} F_{\mu\nu} {}^*F_{\mu\nu} = \frac{1}{32\pi^2} F^a_{\mu\nu} {}^*F^a_{\mu\nu}, \tag{2.4b}$$

where ${}^*F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\kappa\lambda} F_{\kappa\lambda}$ is the dual tensor. The integer k corresponds to the second Chern class.[16]

The massless Euclidean YM equations (for any simple non-Abelian compact Lie group G as a gauge group, and with any number of massless fermions and bosons) are invariant under the action of the Euclidean conformal group O(5,1), which contains rotations, translations, dilatations, and special conformal transformations. They are also invariant under the discrete transformations of reflections, and inversion $J: x_\mu \to x_\mu / x^2$. In this section, we show the simple but important fact, that the one meron solution is invariant (within gauge equivalence) under the O(4)$\times$O(1,1) subgroup of the conformal group O(5,1). This subgroup is generated by rotations and dilatations. The meron is also invariant under reflections and inversions.

The solutin of (2.1) which describes one meron at the origin and one antimeron at infinity reads[3]

$$A^a_\mu(x) = \eta_{a\mu\nu} \frac{x_\nu}{x^2}, \tag{2.5}$$

where $\eta_{a\mu\nu}$ is 't Hooft's symbol.[5] We refer to the appendix of Ref. 5, for properties of $\eta_{a\mu\nu}$ we will use throughout this paper. The field tensor of (2.5) is

$$F^a_{\mu\nu} = -\eta_{a\mu\nu}\frac{1}{x^2} + \eta_{a\nu\kappa}\frac{x_\kappa x_\mu}{x^4} + \eta_{a\mu\kappa}\frac{x_\kappa x_\nu}{x^4} \tag{2.6}$$

From this expression, one sees that the action (2.3) is logarithmically divergent at $x = 0$ and $x = \infty$. Indeed

$$S(A^{\text{meron}}) = \frac{3}{2}\int d^4x \frac{1}{x^4} = 3\pi^2 \int_0^{+\infty} \frac{dr}{r}. \tag{2.7}$$

The topological change (2.4b) is $Q(x) = \frac{1}{2}\delta(x)$.

Under a general transformation

$$x_\mu \to x'_\mu = x'_\mu(x), \tag{2.8}$$

the vector and tensor fields transform like 1- and 2-forms, respectively, i.e.,

$$A'_\mu(x) = \frac{\partial x'_\kappa}{\partial x_\mu} A_\kappa(x'), \tag{2.9}$$

$$F'_{\mu\nu}(x) = \frac{\partial x'_\kappa}{\partial x_\mu}\frac{\partial x'_\lambda}{\partial x_\nu} F_{\kappa\lambda}(x'), \tag{2.10}$$

$${}^*F'_{\mu\nu}(x) = \mathrm{sgn}[J(x';x)]\frac{\partial x'_\kappa}{\partial x_\mu}\frac{\partial x'_\lambda}{\partial x_\nu} {}^*F_{\kappa\lambda}(x'), \tag{2.11}$$

where

$$J(x';x) = \det\left(\frac{\partial x'_\rho}{\partial x_\sigma}\right) \tag{2.12}$$

is the Jacobian of $x'_\mu \to x_\mu$. If (2.8) is a conformal transformation, then

$$\frac{\partial x'_\kappa}{\partial x_\mu}\frac{\partial x'_\lambda}{\partial x_\mu} = \delta_{\kappa\lambda}\,|J(x';x)|^{1/2}. \tag{2.13}$$

For all elements of the conformal group $J(x;x') > 0$ [for the proper Euclidean subgroup $J(x;x') = 1$]. For reflections, and inversion $J < 0$. Equation (2.13) together with (2.10) and (2.11) quickly imply the conformal invariance of the action (2.3) and the topological charge (2.4a). The latter changes sign under reflections and inversion.

Under an infinitesimal rotation

$$x'_\mu = x_\mu - \omega_{\mu\nu}x_\nu,\quad \omega_{\mu\nu} = -\omega_{\nu\mu},\quad (x')^2 = x^2, \tag{2.14}$$

the meron (2.5) transforms into

$$A'^a_\mu(x) = \eta_{a\mu\nu}\frac{x_\nu}{x^2} - \frac{\eta_{a\kappa\nu}\omega_{\kappa\mu}x_\nu + \eta_{a\mu\nu}\omega_{\nu\kappa}x_\kappa}{x^2}. \tag{2.15}$$

We shall see that the last term in (2.15) can be compensated for, by a global gauge transformation. Under a gauge transformation $g \in SU(2)$, the vector fields transform as follows,

$$A'_\mu(x) = g^{-1}(x)A_\mu(x)g(x) + g^{-1}(x)\partial_\mu g(x) \tag{2.16}$$

for an infinitesimal transformation

$$g \simeq I + i\theta^a\sigma^a \tag{2.17}$$

Eq. (2.16) becomes

$$A'_\mu(x) = A_\mu + iD_\mu\theta \equiv A_\mu + i\{\partial_\mu\theta + [A_\mu,\theta]\}. \tag{2.18}$$

From (2.18), one can easily see that the last term in (2.15) is compensated by performing an infinitesimal global gauge transformation given by

$$\theta^a = \frac{i}{2}\eta_{a\mu\nu}\omega_{\mu\nu}. \tag{2.19}$$

This proves single meron invariance under rotations. For dilatations

$$x'_\mu = \lambda x_\mu, \tag{2.20}$$

Eq. (2.9) yields

$$A'^a_\mu(x) = \frac{\partial x'_\nu}{\partial x_\mu}A^a_\nu(x') = \eta_{a\mu\nu}\frac{x_\nu}{x^2}. \tag{2.21}$$

This together with the rotational invariance of the meron, prove the $SO(4)\times O(1,1)$ meron invariance. One can easily check the invariance of (2.5) under reflections and inversion. Clearly, (2.5) is not invariant under translations P_μ, and since the special conformal transformations $\mathscr{K}_\mu$ are equal to $\mathscr{I}P_\mu\mathscr{I}$, the meron cannot be invariant under special conformal transformations. It is also true (and this can be deduced from the results of this paper) that the meron is not invariant under any linear combinations of P_μ and $\mathscr{K}_\mu$

Remark 2.1: As we mentioned in the Introduction, the $O(4)\times O(1,1)$ invariance of the meron indicates that $S^3\times\mathbb{R}$ is the natural geometry for merons. This, we realize in Sec. IV.

Since the YM equations are conformally invariant, if one applies a conformal transformation to (2.5), one obtains again a solution of the YM equations. Thus the special conformal transformation

$$x'_\mu = \frac{x_\mu - a_\mu x^2}{1 - 2ax + a^2x^2} \tag{2.22}$$

acting on (2.1) gives

$$A^a_\mu(x) = \eta_{a\mu\nu}\frac{x_\nu - a_\nu x^2}{x^2(1 - 2a\cdot x + a^2x^2)} - \eta_{a\nu\kappa}\frac{2a_\nu x_\kappa x_\mu}{x^2(1 - 2a\cdot x + a^2x^2)}, \tag{2.23}$$

which describes a meron at the origin and an antimeron at $x_\mu = a_\mu/a^2$. This form of a meron–antimeron pair is gauge equivalent to the familiar form[17]

$$A_\mu(x) = i\frac{x_0 - i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x|}(\tfrac{1}{2}\eta_{a\mu\nu}\sigma^a)\left(\frac{x_\nu}{x^2} - \frac{(x-a)_\nu}{(x-a)^2}\right) \times\frac{(x-a)_0 + i\boldsymbol{\sigma}\cdot(\mathbf{x}-\mathbf{a})}{|x-a|}, \tag{2.24}$$

which can be transformed,[17] via a singular gauge transformation, into a meron-meron pair

$$A^a_\mu = \eta_{a\mu\nu}\left(\frac{x_\nu}{x^2} + \frac{(x-a)_\nu}{(x-a)^2}\right). \tag{2.25}$$

Our meron bundle in Sec. IV will be proven to be equivalent to a meron–antimeron pair (2.25) at $a = (a_0,0), b = (b_0,0)$, or $a = o, b = \infty$.

The field tensor of (2.25) is

$$\begin{aligned}F_{\mu\nu}(x) &= -\eta_{a\mu\nu}\frac{a^2}{x^2(x-a)^2}\eta_{a\kappa\nu}\left(\frac{x_\kappa x_\mu}{x^4} + \frac{(x-a)_\kappa(x-a)_\mu}{(x-a)^4} - \frac{x_\kappa(x-a)_\mu + (x-a)_\kappa x_\mu}{x^2(x-a)^2}\right)\\ &\quad + \eta_{a\mu\kappa}\left(\frac{x_\kappa x_\nu}{x^4} + \frac{(x-a)_\kappa(x-a)_\nu}{(x-a)^4} - \frac{x_\kappa(x-a)_\nu + (x-a)x_\nu}{x^2(x-a)^2}\right)\end{aligned} \tag{2.26}$$

Thus

$$F_{\mu\nu}(x) \underset{|x|\to\infty}{\to} O\left(\frac{1}{|x|^4}\right). \tag{2.27}$$

Therefore, the action is finite at ∞, but logarithmically divergent at $x = 0$ and $x = a$. The same is ture for the configurations (2.23) and (2.24).

Remark 2.2: Since (2.5) is invariant under a seven-parameter subgroup of the 15-parameter conformal group, the most general solution one can get from (2.5) by conformal transformations is an eight-parameter solution. In Sec. V, we prove that none of these parameters can be gauged away, and that the most general two meron solutions has exactly eight-parameters. This eliminates the possibility that one meron, in a meron pair, can have internal degrees of freedom indepenently of the other meron.

III. THE STABILITY OPERATOR ON $\mathbb{R}^4$

Let A^a_μ be a particular solution of the YM equations (2.1). We consider infinitesimal deformations about A^a_μ,

$$A^a_\mu \to A^a_\mu + a^a_\mu. \tag{3.1}$$

The linearized YM equations read

$$D_\mu D_\mu a_\nu - D_\mu D_\nu a_\mu - [F_{\mu\nu}, a_\mu] = 0, \tag{3.2}$$

where D_μ is the covariant derivative w.r.t. the solution A_μ, i.e.,

$$D_\mu a_\nu = \partial_\mu a_\nu + [A_\mu, a_\nu]. \tag{3.3}$$

The stability operator $\mathscr{L} = \mathscr{L}(A)$ for the solution A^a_μ, may be deduced from (3.2),

$$(\mathscr{L}a)_\nu \equiv \mathscr{L}_{\mu\nu} a_\mu = (-D_\kappa D_\kappa \delta_{\mu\nu} + D_\mu D_\nu + \operatorname{ad}F_{\mu\nu}) a_\mu. \tag{3.4}$$

Using the identity

$$[D_\mu, D_\nu] = \operatorname{ad}F_{\mu\nu}, \tag{3.5}$$

$\mathscr{L}$ may be rewritten as follows,

$$(\mathscr{L}a)_\nu = (-D_\kappa D_\kappa \delta_{\mu\nu} + D_\nu D_\mu + 2\operatorname{ad}F_{\mu\nu}) a_\mu$$

$$= (-D_\kappa D_\kappa \delta_{\mu\nu} + 2D_\mu D_\nu - D_\nu D_\mu) a_\mu. \tag{3.6a}$$

Remark 3.1: The stability operator $\mathscr{L}$ could be derived by expanding the action (2.3) in a power series in a_μ, and then calcualting the second variation w.r.t. a_μ at $a_\mu = 0$.

$\mathscr{L}$ acts on the Hilbert space $\mathscr{H}$ defined by the inner product

$$(a,b) = \tfrac{1}{2}\int d^4x \operatorname{Tr} a^*_\mu b_\mu. \tag{3.7}$$

Here a^*_μ is the matrix Hermitian conjugate of a_μ. Whenever the solution $A_\mu(x)$ is locally square integrable, $\mathscr{L} = \mathscr{L}(A)$ may be defined as a quadratic form on C^∞_0 vectors by

$$(a, \mathscr{L}a) = -(D_\kappa a_\nu, D_\kappa a_\nu) + 2(D_\mu a_\nu, D_\mu a_\nu)$$

$$-(D_\nu a_\nu, D_\mu a_\mu) \tag{3.8a}$$

$$= -(D_\kappa a_\nu, D_\kappa a_\nu) + (D_\nu a_\nu, D_\mu a_\mu) + 2(a_\nu, \operatorname{ad}F_{\mu\nu} a_\mu), \tag{3.8b}$$

where (3.8b) is valid whenever $F_{\mu\nu}$ is locally absolutely integrable. The meron (2.5) is locally square integrable, and its $F_{\mu\nu}$, (2.6), is locally absolutely integrable. Therefore, (3.8) defines $\mathscr{L}(A^{\text{meron}})$ as a formally densely symmetric operator on the subset $C^\infty_0(\mathbb{R}^4 - \{0\})$ of $\mathscr{H}$. Similarly, the stability operator for (2.26) [or (2.25)], and (2.24) are symmetric operators defined on the dense domains $C^\infty_0(\mathbb{R}^4 - \{0,a\})$, and $C^\infty_0(\mathbb{R}^4 - \{0, a_\mu/a^2\})$, respectively.

We now study properties of $\mathscr{L}$ under conformal transformations. Let Ω be a conformal transformation

$$\Omega: x_\mu \to x'_\mu = (\Omega x)_\mu. \tag{3.9}$$

Let $a^\Omega_\mu = a'_\mu$ be the transformed field [see (2.10)]. Then

$$(a^\Omega_\mu, a^\Omega_\mu) = \tfrac{1}{2}\int d^4x' \operatorname{Tr} a'_\mu(x') a'_\mu(x')$$

$$= \frac{1}{2}\int d^4x \, |J(x';x)|^{1/2} \operatorname{Tr} a_\mu(x) a_\mu(x). \tag{3.10}$$

Thus the operator

$$U(\Omega): \mathscr{H} \to \mathscr{H} \tag{3.11}$$

$$(U(\Omega)a)_\mu(x') = |J(x';x)|^{-1/4} a'_\mu(x') = |J(x';x)|^{-1/4} \frac{\partial x_\kappa}{\partial x'_\mu} a_\kappa(x) \tag{3.12}$$

is unitary. The stability operator $\mathscr{L}(A)$ transforms like

$$U^{-1}(\Omega)\mathscr{L}(A)U(\Omega) = |J(x';x)|^{-1/2}\mathscr{L}(A^\Omega) \tag{3.13a}$$

$$= |J(x';x)|^{-1/2}\, \Omega^* \mathscr{L}(A)\Omega. \tag{3.13b}$$

In the second line we have used

$$\mathscr{L}(A^\Omega) = \Omega^* \mathscr{L}(A)\Omega \tag{3.14}$$

which can easily be derived from (3.8). For Ω in the Euclidean group we have $J(x';x) = 1$. For dilatations (2.21), and the special conformal transformation (2.23) we have

$$J(\lambda x; x) = \lambda^4 \tag{3.15}$$

$$J(x';x) = (1 - 2a\cdot x + a^2x^2)^{-4}, \tag{3.16}$$

respectively. Thus

$$U^{-1}(\lambda)\mathscr{L}(A)U(\lambda) = \lambda^{-2}\mathscr{L}(A^\lambda), \tag{3.17}$$

$$U^{-1}(\Omega)\mathscr{L}(A)U(\Omega) = (1 - 2a\cdot x + a^2x^2)^2 \mathscr{L}(A^\Omega). \tag{3.18}$$

For the meron (2.5), we have in addition $\mathscr{L}(A^\lambda) = \mathscr{L}(A)$. We shall use (3.17) and (3.18) below when we consider the stability operator for meron pairs.

Theorem 3.1: The formal stability operators of a single meron, defined on $C^\infty_0(\mathbb{R}^4 - \{0\})$, and of a meron pair defined on $C^\infty_0(\mathbb{R}^4 - \{0,a\})$, have numerical range $\mathbb{R}$ and self-adjoint extensions all of which are unbounded below.

Proof: First we consider the single meron (2.5). As we mentioned below (3.8), its stability operator $\mathscr{L}$ is formally symmetric on the dense set $C^\infty_0(\mathbb{R}^4 - \{0\})$. To prove that $\mathscr{L}$ has self-adjoint extensions we use (a slight generalization of) von Neumann's theorem for real operators.[18,19] The map

$$C: a_\mu \to a^*_\mu$$

is a conjugation (i.e., antilinear, norm-preserving, and $C^2 = I$) on $\mathscr{H}$. Obviously C leaves $C^\infty_0(\mathbb{R}^4 - \{0\})$ invariant, and commutes with $\mathscr{L}$. By von Neumann's theorem, the deficiency indices of $\mathscr{L}$ are equal, and therefore $\mathscr{L}$ has self-adjoint extensions. To prove that $\mathscr{L}$ has numerical range $\mathbb{R}$, and all its self-adjoint extensions are unbounded below, we observe that the meron (2.5) may be written as follows

$$A^a = -\eta_{a\mu\nu}\partial_\nu \log \rho_0(x), \tag{3.19a}$$

$$\rho_0(x) = \frac{1}{r}, \quad r = |x|, \tag{3.19b}$$

and for vector fields of the form

$$A^a_\mu = -\eta_{a\mu\nu}\partial_\nu \log \rho(x), \tag{3.20}$$

the YM equations become

$$\partial_\mu\left(\frac{\Delta_\rho}{\rho^3}\right)=0 \qquad (3.21)$$

or

$$\Delta\rho = c\rho^3, \qquad (3.22)$$

where Δ is the four-dimensional Laplacian and C is an arbitrary constant. The meron (3.19) corresponds to

$$C=-1, \quad -\Delta\rho-\rho^3=0. \qquad (3.23)$$

Infinitesimal variations about $\rho_0=1/r$, give rise to the stability operator

$$L(\rho_0)=-\Delta-\frac{3}{r^2}. \qquad (3.24)$$

If we restrict $\mathscr{L}(\rho_0)$ to sphericaly symmetric functions with definite angular momentum $l=0,\frac{1}{2},1,\frac{3}{2},\ldots$, then

$$L(\rho_0)=-\frac{d^2}{dr^2}-\frac{3}{r}\frac{d}{dr}+\frac{4l(l+1)}{r^2}-\frac{3}{r^2}. \qquad (3.25)$$

We will study the operator with $l=0$, i.e.,

$$L(\rho_0)=-\frac{d^2}{dr^2}-\frac{3}{r}\frac{d}{dr}-\frac{3}{r^2}. \qquad (3.26)$$

One easily sees that (3.26) and the original stability operator have the same quadratic form (modulo an immaterial constant of proportionality) on vectors with spherical symmetry. Let $u\in C_0^\infty(0,\infty)\subset L_2(\mathbb{R}^4-\{0\})$ and $u=r^{-1}v$. Then

$$(u,\mathscr{L}(\rho_0)u)=\int d^4x\, u\left(-\frac{d^2u}{dr^2}-\frac{3}{r}\frac{du}{dr}-\frac{3}{r^2}u\right)$$
$$=\int_0^\infty dr\, v\left(-\frac{d^2v}{dr^2}-\frac{1}{r}\frac{dv}{dr}-\frac{2}{r^2}v\right). \qquad (3.27)$$

The equation

$$-\frac{d^2v}{dr^2}-\frac{1}{r}\frac{dv}{dr}-\frac{2}{r^2}v=\lambda v \qquad (3.28)$$

is the Bessel equation with index $n^2=-2$, i.e., $n=i\sqrt{2}$. The two linearly independent solutions of (3.28) with negative λ, are the modified Bessel functions with imaginary index, i.e.,

$$I_{i\sqrt{2}}(\sqrt{|\lambda|}\,r), \qquad (3.29a)$$

$$K_{i\sqrt{2}}(\sqrt{|\lambda|}\,r). \qquad (3.29b)$$

Only $K_{i\sqrt{2}}(\sqrt{|\lambda|}\,r)$ is square integrable at both $r=0$ and $r=\infty$. This implis, by (3.27), that the numerical range of $\mathscr{L}$ is $\mathbb{R}$. The Bessel equation (3.28) can be reduced to a (time-independent) Schrödinger equation by setting $v=r^{-1/2}w$ (or $u=r^{-3/2}w$)

$$-\frac{d^2w}{dr^2}-\frac{9}{4}\frac{1}{r^2}w=\lambda w \qquad (3.30)$$

and

$$(u,\mathscr{L}(\rho_0)u)=\int_0^{+\infty} dr\, w\left(-\frac{d^2w}{dr^2}-\frac{9}{4}\frac{1}{r^2}w\right). \qquad (3.31)$$

The one-dimensional Schrödinger operator $-d^2/dr^2-\eta(1/r^2)$ has been studied exensively in the literature[20,21] (and references given in Ref. 21). For $\eta\leqslant\frac{1}{4}$, the operator has a unique self-adjoint extension (the Friedrichs extension) and is positive. For $\eta>\frac{1}{4}$, the results of Case[20] and Nelson,[21] show that the origin is in the limit circle case, and the operator is not form bounded below; its deficiency indices are (1,1) and there is a one-parameter family of self-adjoint extensions, parametrized by the real numbers and corresponding to the boundary conditions $w'(0)+aw(0)=0$. All these self-adjoint extensions are unbounded below and have discrete negative spectrum. Nelson[21] has constructed an extension, with a physically interesting dynamics, which is unbounded below and has continuous spectrum. This extension is not self-adjoint.

Remark 3.2: Let

$$\tau=\log r, \qquad (3.32)$$

$$f=\frac{d}{d\tau}\log\rho(r)=r\frac{d\rho/dr}{\rho}. \qquad (3.33)$$

Then (3.23) becomes

$$\frac{d^2f}{d\tau^2}-2f(f+1)(f+2)=0. \qquad (3.34)$$

The solutions $f=0$ and $f=-2$ correspond to pure gauges ($f=-2$ is a singular gauge), while $f=-1$ gives (2.5). The meron stability operator coming from (3.34) is

$$\widetilde{\mathscr{L}}=\frac{d^2}{d\tau^2}-2$$

acting on $\widetilde{\mathscr{H}}=L_2(\mathbb{R})$. One easily sees that $f=-1$ is a stable solution of (3.34) in the sense of Liapunov, but $\widetilde{\mathscr{L}}$ has a negative spectrum. Defining

$$T:\mathscr{H}\to\widetilde{\mathscr{H}}$$
$$u\to Tu=\frac{d}{d\tau}(c^\tau u) \qquad (3.35)$$

one finds

$$\mathscr{L}(\rho_0)=e^{-2\tau}T^{-1}\widetilde{\mathscr{L}}T. \qquad (3.36)$$

Although (3.34) and (3.23) (with spherical symmetry) are equivalent, $\mathscr{L}(A^{\text{meron}})$ and $\widetilde{\mathscr{L}}(f^{\text{meron}})$ have different spectra. This phenomenon can be traced to the singular nature of the transformation (3.32). We will encounter a similar phenomenon in Sec. V.

Remark 3.3: One can easily check that the four vectors

$$a_\mu^a(\nu)=\eta_{a\mu\nu}\frac{1}{x^2}-2\eta_{a\mu\kappa}\frac{x_\kappa x_\nu}{x^4}, \quad \nu=0,1,2,3 \qquad (3.37)$$

belong to the null space of $\mathscr{L}(A^{\text{meron}})$. These vectors are obtained by differentiating the solution of a single meron located at $x=z$, i.e.,

$$a_\mu^a(\nu)=-\frac{\partial}{\partial z_\nu}A_\mu^a(x;z)\Big|_{z=0}=-\eta_{a\mu\kappa}\frac{\partial}{\partial z_\kappa}\left(\frac{(x-z)}{(x-z)^2}\right)\Big|_{z=0} \qquad (3.38)$$

They satisfy the "background" gauge

$$D_\mu a_\mu^a\equiv\partial_\mu a_\mu^a+\epsilon^{abc}\eta_{b\mu\nu}\frac{x_\nu}{x^2}a_\mu^c=0 \qquad (3.39)$$

and they are not square integrable (thus the origin is part of the continuum spectrum). In Sec. V, we shall prove that the dimension of the null space (of generalized eigenvectors) of $\mathscr{L}$, for a single meron, is exactly four.

We now consider the meron–antimeron pair described by (2.24), and complete the proof of Theorem 3.1. Let Ω be the conformal transformation (2.23), A_μ^Ω the vector field (2.24), and $\mathscr{L}(A^\Omega)$ its stability operator. Let $a_\mu \in C_0^\infty(\mathbb{R}^4 - \{0\})$, and λ small enough so that $(\Omega^{-1}U^{-1}(\lambda)a)_\mu(x) \in C_0^\infty(\mathbb{R}^4 - \{0, a_\mu/a^2\})$ [recall $U(\lambda)$ and $U(\Omega)$ from (3.12)]. From (3.13), and the fact that A_μ (2.5) is invariant under dilatations, we have

$$(\Omega^{-1}U^{-1}(\lambda)a_\mu, \mathscr{L}(A^\Omega)\Omega^{-1}U^{-1}(\lambda)a_\mu) = \lambda^{-2}(a_\mu, \mathscr{L}(A)a_\mu). \tag{3.40}$$

If we choose a_μ so that $(a_\mu, \mathscr{L}(A)a_\mu) < 0$, then the lhs of (3.40) is also negative, and so A^Ω is unstable. As $\lambda \downarrow 0$, the rhs of (3.40) goes to $-\infty$. To prove that $\mathscr{L}(A^\Omega)$ is form unbounded below it is enough to prove that

$$\lim_{\lambda \downarrow 0} \|\Omega^{-1}U^{-1}(\lambda)a\| = \|a\|. \tag{3.41}$$

Using (3.12) we have

$$(\Omega^{-1}U^{-1}(\lambda)a, \Omega^{-1}U^{-1}(\lambda)a) = \frac{1}{2}\int d^4x |J(\Omega^{-1}\lambda x; \lambda x)|^{1/2} \mathrm{Tr}(U^{-1}(\Omega)U^{-1}(\lambda)a)_\mu(x) (U^{-1}(\Omega)U^{-1}(\lambda)a)_\mu(x). \tag{3.42}$$

From (3.16) one sees that

$$\lim_{\lambda \downarrow 0} J(\Omega^{-1}\lambda x; \lambda x) = 1. \tag{3.43}$$

This together with the unitarity of $U(\Omega)U(\lambda)$ easily imply (3.41). The proof of existence of self-adjoint extensions for $\mathscr{L}(A^\Omega)$ is the same as for $\mathscr{L}(A)$. This completes the proof of Theorem 3.1.

Remark 3.4: The meron–meron pair (2.26) corresponds to the solution

$$\rho^{(mm)}(x) = \frac{|a|}{x|x-a|} \tag{3.44}$$

of (3.23). The corresponding stability operator is

$$\mathscr{L}^{(mm)} = -\Delta - \frac{3a^2}{x^2(x-a)^2}. \tag{3.45}$$

This is obtained from (3.24) by a conformal transformation. One could work directly with (3.45), and prove instability of a meron pair.

Remark 3.5: The stability operator for a meron pair, has eight zero-generalized eigenvectos obtained as in Remark 3.3. In Sec. V, we prove that its null space (generalized eigenvectors!) is exactly eight-dimensional (se Remark 2.2).

We now regularize the single meron (2.5), and study stability properties of the regularized configuration (similar study would be carried through for a meron pair). Let $R_0 < R$. We replace the single meron (2.5) by the configuration

$$A_\mu^a(x) = -\eta_{a\mu\nu}\partial_\nu \log \rho(r), \tag{3.46a}$$

$$\rho(r) = \begin{cases} \dfrac{2R_0^2}{R_0^2 + r^2}, & 0 \leqslant r \leqslant R_0, \\[2ex] \dfrac{1}{r}, & R_0 \leqslant r \leqslant R, \\[2ex] \dfrac{2R^2}{R^2 + r^2}, & R \leqslant r. \end{cases} \tag{3.46b}$$

This amounts to replacing the meron by half-instantons inside a sphere of radius R_0 and outside a sphere of radius R. The stability operators $\mathscr{L}_{R_0,R}$ within deformations coming from $\rho(x)$ in (3.20), reads

$$\mathscr{L}_{R_0,R} = -\Delta - g(r)\frac{1}{r^2}, \tag{3.47}$$

where

$$g(r) = \begin{cases} \dfrac{12R_0^2 r^2}{(R_0^2 + r^2)^2}, & 0 \leqslant r \leqslant R_0, \\[2ex] 3, & R_0 \leqslant r \leqslant R, \\[2ex] \dfrac{12R^2 r^2}{(R^2 + r^2)^2}, & R \leqslant r \end{cases} \tag{3.48}$$

Theorem 3.2: (a) $L_{R_0,R}$ is formally symmetric and densely defined on $C_0^\infty(\mathbb{R}^4)$, and has a self-adjoint extension, possibly many.

(2) For $R \gg R_0$, the numerical range of $\mathscr{L}_{R_0,R}$ contains negative points.

(3) If $R = \infty$ and $R_0 \leqslant C < +\infty$, for a fixed constant C, then $\mathscr{L}_{R_0} \equiv \mathscr{L}_{R_0}$ is form unbounded below.

Proof: The first part of the theorem is proven as in Theorem 3.1. To prove part (2) we follow the procedure of Theorem 3.1. Let $u \in C_0^\infty(0,\infty)$ and $u = r^{-3/2}w$. Then

$$(u, L_{R_0,R}u) = \int d^4x\, u\left(-\Delta - g(r)\frac{1}{r^2}\right)u = \int_0^\infty dr\, w\left(-\frac{d^2}{dr^2} - [g(r) - \tfrac{3}{4}]\right)w. \tag{3.49}$$

One would attempt to solve the equation

$$-\frac{d^2w}{dr^2} - [g(r) - \tfrac{3}{4}]w = \lambda w \tag{3.50}$$

in the three regions of (3.48), and match the solutions at $r = R_0$, and $r = R$. However, the equation does not seem to be explicitly soluable in the inner and outer regions. Thus we settle for a weaker result. Let L be a constant (to be chosen below), and $w(r) a C_0^\infty(0,\infty)$ function which is constant in (R_0, R), zero for $r < R_0 - L$ and $r > R + L$, and monotone in $[R_0 - L, R_0] \cup [R, R + L]$ (e.g., linear and smoothed off at the ends). Provided that $R \ll R_0$, and $L \ll R - R_0$, one can easily check that the rhs of (3.49) is negative. This proves part (2). To prove that (3), we note that it suffices to prove that the operator

$$-\frac{d^2}{dr^2}-[\tilde{g}(r)-\tfrac{3}{4}]\frac{1}{r^2}, \quad (3.51)$$

$$\tilde{g}(r)=\begin{cases}0, & 0\leqslant r<R_0\\ 3, & R_0<r\end{cases}, \quad (3.52)$$

is from unbounded below. Now the equations

$$-\frac{d^2w}{dr^2}+\frac{3}{4}\frac{1}{r^2}w'\lambda w, \quad 0\leqslant r<R_0 \quad (3.53)$$

$$-\frac{d^2w}{dr^2}-\frac{9}{4}\frac{1}{r^2}w=\lambda w, \quad R_0<r \quad (3.54)$$

are soluble, and one could try to match their solutions (and their derivatives). Instead, we proceed as follows. The two linearly independent solutions of (3.54) (for negative λ) are

$$r^{1/2}I_{i\sqrt{2}}(\sqrt{|\lambda|}\,r), \quad (3.55a)$$

$$r^{1/2}K_{i\sqrt{2}}(\sqrt{|\lambda|}\,r). \quad (3.55b)$$

Only (3.55b) is square integrable at infinity. Let $\epsilon>0$ be sufficiently small, and let

$$\zeta(r)\in C^\infty(0,\infty), \quad (3.56a)$$

$$0\leqslant\zeta(r)\leqslant 1, \quad (3.56b)$$

$$\zeta(r)=\begin{cases}1 & \text{for } r>R_0+\epsilon,\\ 0 & \text{for } r<R_0\end{cases}. \quad (3.56c)$$

Let

$$w(r)=r^{1/2}K_{i\sqrt{2}}(\sqrt{|\lambda|}\,r)\zeta(r). \quad (3.57)$$

One easily sees that

$$\left(w\left(-\frac{d^2}{dr^2}-[\tilde{g}(r)-\tfrac{3}{4}]\right)w\right)<0, \quad (3.58)$$

i.e., (3.52) has negative numerical range. Since (3.52) is dilatation invariant, its numerical range extends to $-\infty$; by (3.17). This completes the proof of Theorem 3.2.

IV. EMBEDDING $\mathbb{R}^4$ INTO $S^3\times\mathbb{R}$ AND THE MERON BUNDLE

In this section we embed $\mathbb{R}^4$ into $S^3\times\mathbb{R}$, choose an appropriate basis on the tangent and cotangent spaces, formulate YM equations on $S^3\times\mathbb{R}$, prove that the vector fields transform according to the (1,0) spin representation of SO(4), and determine the meron bundle.

Let

$$a=(a_0,\mathbf{0}), \quad b=(b_0,\mathbf{0}). \quad (4.1)$$

For a meron–antimeron pair at $x=a=(a_0,\mathbf{0})$, and $x=b=(b_0,\mathbf{0})$, we employ the transformation

$$\tau=\log\frac{|x-a|}{|x-b|}, \quad (4.2a)$$

$$z_i=\frac{x_i(b_0-a_0)}{|x-a|\,|x-b|}, \quad i=i,2,3, \quad (4.2b)$$

$$z_4=\frac{-(x-a)(x-b)}{|x-a|\,|x-b|}, \quad (4.2c)$$

and its inverse

$$x_0=\frac{b_0+a_0}{2}+\frac{b_0-a_0}{2}\frac{\sinh\tau}{\cosh\tau+z_4}, \quad (4.3a)$$

$$x_i=\frac{b_0-a_0}{2}\frac{z_i}{\cosh\tau+z_4}, \quad i=1,2,3. \quad (4.3b)$$

For a meron-antimeron pair at $x=a$ and $x=\infty$, we employ the transformation

$$\tau=\log|x-a|, \quad (4.4a)$$

$$z_i=\frac{(x-a)_i}{|x-a|}, \quad i=1,2,3, \quad (4.4b)$$

$$z_4=\frac{x_0-a_0}{|x-a|}, \quad (4.4c)$$

and its inverse

$$x_0=a_0+e^{\tau}z_4, \quad (4.5a)$$

$$x_i=a_i+e^{\tau}z_i, \quad i=1,2,3. \quad (4.5b)$$

From either (4.2) or (4.4) we get

$$z_1^2+z_2^2+z_3^2+z_4^2=1. \quad (4.6)$$

The following correspondence between points is useful in visualizing the map (4.2) [from now on, we will consider case (4.2) only]. We assume $a_0<b_0$:

$$x=o\Rightarrow\tau=\log\frac{|a_0|}{|b_0|}, \quad z_i=0, \quad z_4=1, \quad (4.7a)$$

$$x=\infty\Rightarrow\tau=0, \quad z_i=0, \quad z_4=-1, \quad (4.7b)$$

$$\tau((a_0,\mathbf{0}))=-\infty, \quad \tau((b_0,\mathbf{0}))=+\infty, \quad (4.7c)$$

$$z_i((x_0,\mathbf{0}))=0, \quad (4.7d)$$

$$z_4=((x_0\neq a_0,b_0,\mathbf{0}))=\begin{cases}1 & \text{for } a_0<x_0<b_0,\\ -1 & \text{for } x_0<a_0, \text{ or } x_0<b_0,\end{cases} \quad (4.7e)$$

$$\lim_{x_0\to b_0^+} z_4((x_0,\mathbf{0}))=-1, \quad (4.7f)$$

or

$x_0^-\to a_0^-$

$$\lim_{x_0\to b_0^-} z_4((x_0,\mathbf{0}))=+1. \quad (4.7g)$$

or

$x_0\to a_0^+$

The map (4.2) is graphed in Fig. 1, where $\mathbb{R}^4$ is represented by a two-dimensional space.

Remark: The transformation (4.2) has been motivated by the manifestly conformally covariant formulation of YM equations.[22,24] It is also related to the quadratic (or "blow up") transformations algebraic geometry.[15]

We choose the following basis on the tangent space of $S^3\times\mathbb{R}$:

$$\hat{\partial}_0=\frac{\partial}{\partial\tau}, \quad (4.8a)$$

$$\hat{\partial}_i=\tfrac{1}{2}\eta_{i\alpha\beta}(z_\alpha\partial_\beta-z_\beta\partial_\alpha)=\eta_{i\alpha\beta}\,z_\alpha\partial_\beta, \quad i=1,2,3, \quad (4.8b)$$

and the corresponding basis on the cotangent space:

$$\omega_0=d\tau, \quad (4.9a)$$

$$\omega_i=\tfrac{1}{2}\eta_{i\alpha\beta}(z_\alpha dz_\beta-z_\beta dz_\alpha)$$

$$=\eta_{i\alpha\beta}z_\alpha dz_\beta, \quad i=1,2,3. \quad (4.9b)$$

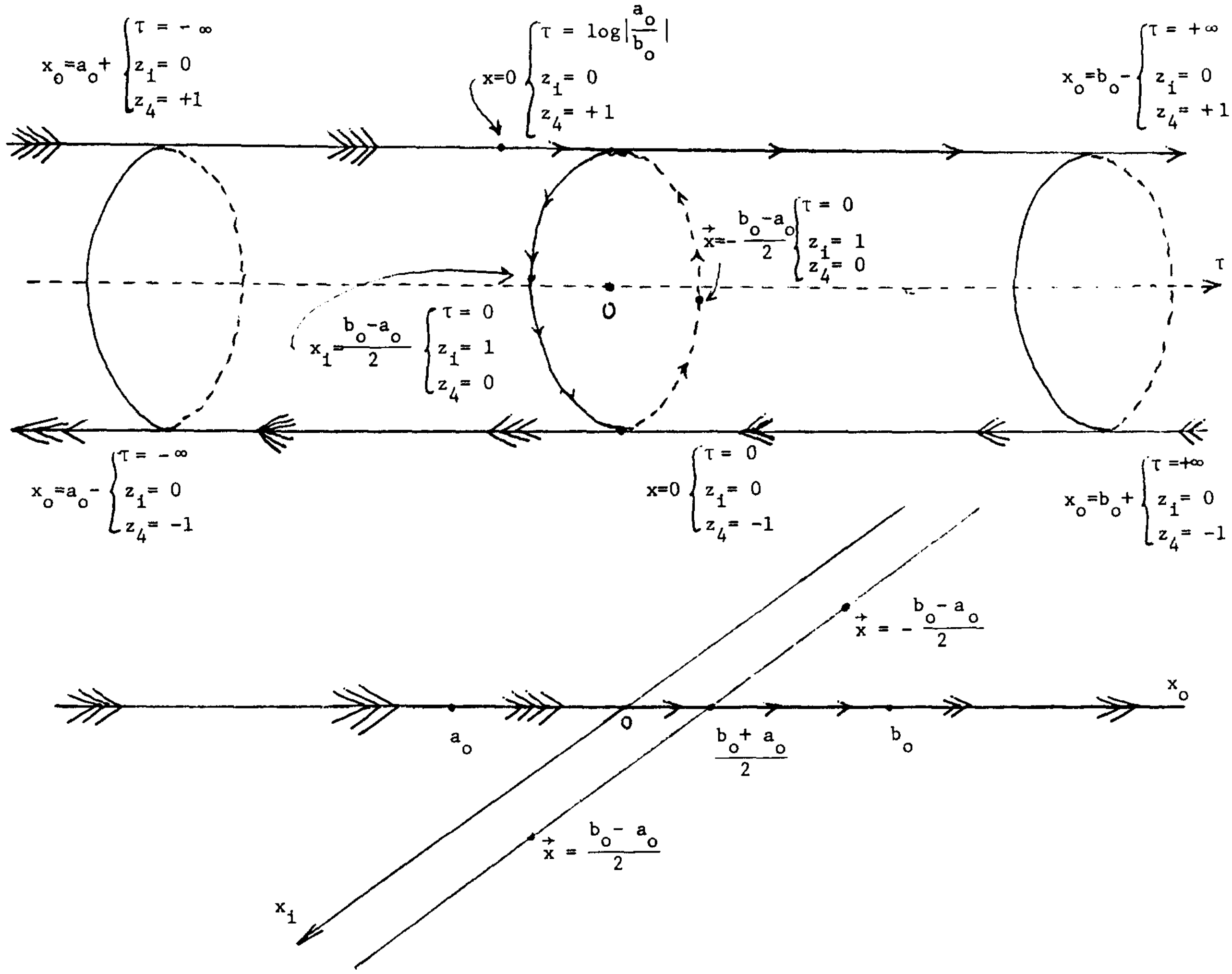


FIG. 1.

Note that

$$[\widehat{\partial}_0,\widehat{\partial}_i] = 0, \qquad (4.10a)$$

$$[\widehat{\partial}_i,\widehat{\partial}_j] = -2\epsilon_{ijk}\widehat{\partial}_k. \qquad (4.10b)$$

The basis (4.8) is chosen so that the tangent space of s^3 (in $S^3\times\mathbb{R}$) is identified, topologically, with the Lie algebra of the first SU(2) component of the rotation group SO(4) $=$ SU(2)×SU(2). The rotations in the two invariant SU(2) subgroups of SO(4) are conveniently defined by Ref. 5;

$$L^i_1 = -\tfrac{1}{4}i\eta_{i\alpha\beta}L_{\alpha\beta} = -\tfrac{1}{2}i\eta_{i\alpha\beta}z_\alpha\partial_\beta, \qquad (4.11a)$$

$$L^i_2 = -\tfrac{1}{4}i\,\overline{\eta}_{i\alpha\beta}L_{\alpha\beta} = -\tfrac{1}{2}i\,\overline{\eta}_{i\alpha\beta}z_\alpha\partial_\beta, \qquad (4.11b)$$

where

$$L_{\alpha\beta} = z_\alpha\partial_\beta - z_\beta\partial_\alpha. \qquad (4.12)$$

The basis (4.8) satisfies

$$\widehat{\partial}_j = 2iL^j_1, \quad j = 1,2,3. \qquad (4.13)$$

The basis has been used previously in general relativity,[25] and on a torus.[26]

We now define vector fields (connections) B_μ, and their field tensors (curvatures) $G_{\mu\nu}$ over $S^3\times\mathbb{R}$. They are related to A_μ and $F_{\mu\nu}$ on $\mathbb{R}^4$ by:

$$A_\mu(x)dx_\mu = B_\mu(z,\tau)\omega_\mu, \qquad (4.14)$$

$$F_{\mu\nu}(x)dx_\mu\wedge dx_\nu = G_{\mu\nu}(z,\tau)\omega_\mu\wedge\omega_\nu, \qquad (4.15)$$

or more explicitly:

$$A_\mu(x) = B_0(z,\tau)\frac{\partial\tau}{\partial x_\mu} + B_j(z,\tau)\tfrac{1}{2}\eta_{j\alpha\beta}M^{(\mu)}_{\alpha\beta}, \qquad (4.16)$$

$$F_{\mu\nu}(x) = G_{0j}(z,\tau)\frac{\partial\tau}{\partial x_\mu}\tfrac{1}{2}\eta_{j\alpha\beta}M^{(\nu)}_{\alpha\beta} + G_{jo}\tfrac{1}{2}\eta_{j\alpha\beta}M^{(\mu)}_{\alpha\beta}\frac{\partial\tau}{\partial x_\nu} + G_{ij}\tfrac{1}{4}\eta_{i\alpha\beta}\eta_{j\gamma\delta}M^{(\mu)}_{\alpha\beta}M^{(\nu)}_{\gamma\delta}, \qquad (4.17)$$

$$M^{(\mu)}_{\alpha\beta} = z_\alpha\frac{\partial z_\beta}{\partial x_\mu} - z_\beta\frac{\partial z_\alpha}{\partial x_\mu} = -M^{(\mu)}_{\beta\alpha}, \qquad (4.18)$$

$$B_0(z,\tau) = A_\mu(x)\widehat{\partial}_0x_\mu = A_\mu(x)\frac{\partial x_\mu}{\partial\tau}, \qquad (4.19)$$

$$B_j(z,\tau) = A_\mu(x)\widehat{\partial}_jx_\mu = A_\mu(x)\tfrac{1}{2}\eta_{j\alpha\beta}N^{(\mu)}_{\alpha\beta}, \qquad (4.20)$$

$$G_{oj}(z,\tau) = F_{\mu\nu}(x)\widehat{\partial}_0x_\mu\widehat{\partial}_jx_\nu = F_{\mu\nu}(x)\frac{\partial x_\mu}{\partial\tau}\frac{1}{2}\eta_{j\alpha\beta}N^{(\mu)}_{\alpha\beta}, \qquad (4.21)$$

$$G_{ij} = F_{\mu\nu}(x)\widehat{\partial}_ix_\mu\widehat{\partial}_jx_\nu = F_{\mu\nu}(x)\tfrac{1}{4}\eta_{i\alpha\beta}\eta_{j\gamma\delta}N^{(\mu)}_{\alpha\beta}N^{(\nu)}_{\gamma\delta}, \qquad (4.22)$$

$$N^{(\mu)}_{\alpha\beta} = Z_\alpha \frac{\partial x_\mu}{\partial z_\beta} - z_\beta \frac{\partial x_\mu}{\partial z_\alpha}. \tag{4.23}$$

The connections B_μ, and its curvature $G_{\mu\nu}$ are related by (the "structure equations")

$$G_{0j} = \hat{\partial}_0 B_j - \hat{\partial}_j B_0 + [B_0, B_j], \tag{4.24a}$$

$$G_{ij} = \hat{\partial}_i B_j - \hat{\partial}_j B_i + [B_i, B_j] + 2\epsilon_{ijk} B_k. \tag{4.24b}$$

In (4.24b), we have made use of the formula (Ref. 27, Chap. III)

$$R(X,Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]} \tag{4.25}$$

for the curvature in terms of covariant derivatives [remember that $\hat{\partial}_i$ and $\hat{\partial}_j$ satisfy (4.10b)].

The action associated with $G_{\mu\nu}$ [and corresponding to (2.3)] is given by

$$S = \frac{1}{2}\int_{-\infty}^{+\infty} d\tau \int_{\mathbb{R}} d^4 z\, \delta(z^2 - 1) \mathrm{Tr} G_{\mu\nu} G_{\mu\nu}$$

$$= \frac{1}{4}\int_{-\infty}^{+\infty} d\tau \int_{\mathbb{R}^4} d^4 z\, \delta(z^2 - 1) G^a_{\mu\nu} G^a_{\mu\nu}. \tag{4.26}$$

The Euler–Lagrange equations of (4.26) give the YM equations on $S^3 \times \mathbb{R}$

$$\hat{D}_j G_{j0} \equiv \hat{\partial}_j G_{j0} + [B_j, G_{j0}] = 0, \tag{4.27a}$$

$$\hat{D}_\mu G_{\mu i} - \epsilon_{ijk} G_{jk} \equiv \hat{\partial}_\mu G_{\mu i} + [B_\mu, G_{\mu i}] - \epsilon_{ijk} G_{jk} = 0. \tag{4.27b}$$

We now study the transformation properties of B_μ under SO(4), construct the spin of the vector fields, prove that the O(4) invariant fields are gauge equivalent to

$$B^a_0(z,\tau) = 0, \quad B_0 = 0 \tag{4.28a}$$

or

$$B^a_i(z,\tau) = f(\tau)\delta^a_i, \quad \mathrm{B_i} = -\tfrac{1}{2}\mathrm{f}(\tau)\mathrm{h}^{-1}\partial_i \mathrm{h}, \tag{4.28b}$$

where $f(\tau)$ is an arbitrary function of τ, and $h = z_4 - i\sigma_j z_j$. Therefore, the O(4)×O(1,1) invariant fields are gauge equivalent to

$$B^a_0(z,\tau) = 0, \tag{4.29a}$$

$$B^a_i(z,\tau) = \rho\delta^a_i, \rho = \text{constant}. \tag{4.29b}$$

Let $\Lambda \in \mathrm{SO}(4)$ (remember $\Lambda_{\alpha\beta}\Lambda_{\alpha\gamma} = \delta_{\beta\gamma}$), and

$$z'^a_\alpha = \Lambda_{\alpha\beta} z_\beta \tag{4.30a}$$

$$\tau' = \tau \tag{4.30b}$$

Then

$$B'^a_\mu(z,\tau)\omega_\mu = B^a_\mu(z',\tau')\omega'_\mu$$

$$= B'_0(z',\tau')\frac{d\tau'}{d\tau} d\tau + B^a_i(z',\tau')(S(\Lambda))_{ij} \tfrac{1}{2}\eta_{j\alpha\beta}$$

$$\times (z_\alpha dz_\beta - z_\beta dz_\alpha), \tag{4.31}$$

where the matrix $S(\Lambda)$ is defined by

$$\eta_{i\alpha\beta}\Lambda_{\alpha\gamma}\Lambda_{\beta\delta} = (S(\Lambda))_{ij}\eta_{j\gamma\delta}. \tag{4.32}$$

From (4.31) we get

$$B'^a_0(z,\tau) = B^a_0(\Lambda z,\tau), \tag{4.33a}$$

$$B'^a_i(z,\tau) = B^a_j(\Lambda z,\tau)(S(\Lambda))_{ji}. \tag{4.33b}$$

We now determine the three-dimensional orthogonal representation $S(\Lambda)$ (spin representation) of SO(4). Let

$$S(\Lambda) = e^{(i/2)\omega_{\alpha\beta}\Sigma_{\alpha\beta}}, \tag{4.34}$$

where $\omega_{\alpha\beta} = -\omega_{\beta\alpha}$, and $\Sigma_{\alpha\beta}$ is to be determined. For infinitesimal Λ,

$$\Lambda_{\alpha\beta} = \delta_{\alpha\beta} + \omega_{\alpha\beta}, \tag{4.35a}$$

$$\omega_{\alpha\beta} = -\omega_{\beta\alpha}, \tag{4.35b}$$

we have

$$S(\Lambda) = I + \frac{i}{2}\omega_{\alpha\beta}\Sigma_{\alpha\beta}. \tag{4.36}$$

Inserting (4.35) and (4.36) into (4.32) we obtain, after some algebra,

$$\eta_{i\gamma\delta} + \frac{i}{2}\omega_{\alpha\beta}(\Sigma_{\alpha\beta})_{ij}\eta_{j\gamma\delta} = \eta_{i\alpha\delta} + \eta_{i\alpha\delta}\omega_{\alpha\gamma} - \eta_{i\gamma\beta}\,\omega_{\delta\beta}. \tag{4.37}$$

This is satisfied if

$$(\Sigma_{\alpha\beta})_{ij} = -i\epsilon_{ijl}\eta_{l\alpha\beta}. \tag{4.38}$$

To find the Casimir operators of the representation, we project $\Sigma_{\alpha\beta}$ into the two SU(2) components of SO(4):

$$S^a_1 = \tfrac{1}{4}\eta_{a\alpha\beta}\Sigma_{\alpha\beta}, \tag{4.39a}$$

$$S^a_2 = \tfrac{1}{4}\bar{\eta}_{a\alpha\beta}\Sigma_{\alpha\beta}, \tag{4.39b}$$

or

$$(S^a_1)_{ij} = -i\epsilon_{aij}, \tag{4.40a}$$

$$(S^a_2)_{ij} = 0. \tag{4.40b}$$

From (4.36) we have

$$(S(\Lambda))_{ij} = \delta_{ij} + \epsilon_{ija}\omega_a, \tag{4.41a}$$

$$\omega_a = \tfrac{1}{2}\eta_{a\alpha\beta}\,\omega_{\alpha\beta}. \tag{4.41b}$$

Summarizing: We have proven

$$(S^a_1)_{ij} = -i\epsilon_{aij} \quad \text{or} \quad S^a_1 B^b_2 = -i\epsilon_{aij}B^b_j, \tag{4.42}$$

$$(S^a_2)_{ij} = 0 \quad \text{or} \quad S^a_2 B^b_2 = 0, \tag{4.43}$$

$$S(\Lambda) = \exp(i\omega_a S^a_1), \quad \omega_a = \tfrac{1}{2}\eta_{a\alpha\beta}\,\omega_{\alpha\beta}. \tag{4.44}$$

Since

$$\vec{S}^2_1 = 2 = 1(1+1), \quad \vec{S}^2_2 = 0, \tag{4.45}$$

the vector fields transform according to the $(s_1,s_2) = (1,0)$ representation of SO(4).

Remark: $s_2 = 0$ could have been expected from the fact that the second SU(2) component of SO(4) has been chosen [see remarks below (4.10)] to be orthogonal to S^3 (in $S^3 \times \mathbb{R}$).

We now prove (4.28). Under gauge transformation $g \in \mathrm{SU}(2)$, we have

$$B_\mu \to B'_\mu = g^{-1}B_\mu g + g^{-1}\hat{\partial}_\mu g. \tag{4.46}$$

For an infinitesimal global gauge transformation

$$g = I + i\theta_a\sigma_a, \tag{4.47}$$

(4.46) becomes

$$B'^a(z,\tau) = B^a_\mu(z,\tau) + i\epsilon^{abc}B^a_\mu(z,\tau)\theta^c. \tag{4.48}$$

On the other hand, for an infinitesimal Λ as in (4.35), we have from (4.33)

$$B'^a_0(z,\tau) = B^a_0(z,\tau) + \omega_{\alpha\beta}z_\beta\partial_\alpha B^a_0(z,\tau), \tag{4.49a}$$

$$B'^a_i(z,\tau) = B^a_i(z,\tau) + \omega_{\alpha\beta}z_\beta\partial_\alpha B^a_i(z,\tau) - \epsilon_{aij}\tfrac{1}{2}\eta_{a\alpha\beta}\,\omega_{\alpha\beta}B^a_j(z,\tau). \tag{4.49b}$$

A careful study of (4.48) and (4.49) shows that they agree if the vector fields satisfy (4.28) and

$$\theta \equiv \frac{1}{2i}\theta_a\sigma_a = -\tfrac{1}{4}\eta_{a\alpha\beta}\delta_{\alpha\beta}S^a_1. \tag{4.50}$$

If we further require O(4)×O(1,1) invariance of the fields, we conclude that $f(\tau)$ must be a constant. The field (4.29) satisfies the YM equations (6.27) if $\rho = 0, -1, -2$. For the vector field (4.28) we have

$$G^a_{0i} = \frac{df}{d\tau}\delta^a_i, \tag{4.51a}$$

$$G^a_{ij} = f(f+2)\epsilon^{aij}. \tag{4.51b}$$

The action (4.26) and the YM equations (4.29) become

$$S = \pi^2\int_{-\infty}^{+\infty} d\tau\left[\frac{1}{2}\left(\frac{df}{d\tau}\right)^2 + \tfrac{1}{2}f^2(f+2)^2\right], \tag{4.52}$$

$$\frac{d^2f}{d\tau^2} - 2f(f+1)(f+2) = 0. \tag{4.53}$$

This has the solutions $f = 0, -1, -2$, and

$$f(\tau) = -2\frac{e^{\tau}}{e^{\tau}+e^{-\tau}}. \tag{4.54}$$

$f = 0$ and -2 are pure gauges, and (4.54) corresponds to the instanton solution. $f = -1$ is a half-gauge with

$$B^a_0 = 0,\quad B^a_i = -\delta^a_i,\quad B_i = \tfrac{1}{2}h^{-1}\partial_i h,\quad h = z_4 - i\boldsymbol{\sigma}\cdot\mathbf{z}, \tag{4.55a}$$

$$G^a_{0i} = 0,\quad G^a_{ij} = -\epsilon^{aij}. \tag{4.55b}$$

This is the standard SU(2) spin bundle[27] over S^3. We call it the *meron bundle*. In the appendix we show that it yields, via (4.2), a meron–antimeron pair at $a = (a_0,\mathbf{0})$, $b = (b_0,\mathbf{0})$, and via (4.4) a meron at $x = a$ and an antimeron at $x = \infty$. We also show that the pure gauge $f = -2$ is given, on $\mathbb{R}^4$, by a singular gauge transformation, and that it corresponds to two instantons [or one instanton for (4.4)] with zero size. The instanton (4.54) translated back to $\mathbb{R}^4$ is in singular gauge; the translation via the maps (4.2) or (4.4) leads to two gauge equivalent forms.

V. THE STABILITY OPERATOR OF THE MERON BUNDLE

In this section, we derive the stability operator of the meron bundle, diagonalize it in terms of the angular momenta, spin, and isospin of the vector fields, prove that the dimension of its null space (a space of generalized eigenvectors) is six, and calculate its spectrum explicitly.

The instability of the meron bundle (like that of the meron on $\mathbb{R}^4$) could easily be sen from Eq. (4.53) [note that the solution $f = -1$ of (4.53) is stable in the sense of ordinary differential equation (Liapunov) but unstable in the sense of the stability operator]. From the same equation (or its double-well potential), one can also see that the meron "decays" into two vacua: $f = 0$ and $f = -2$. The stability operator coming from (4.53) is $(d^2/d\tau^2) - 2$. It has two zero generalized eigenmods.

$$\cos\sqrt{2}\tau,\quad \sin\sqrt{2}\tau. \tag{5.1}$$

If the spin and isospin of the vector field are turned on (keeping the angular momenta zero), it can be shown that the multiplicity of each one of these eigenmodes is three. Thus, together they give six zero-eigenmodes. We shall prove in this section that these are the only zero eigenmodes, i.e., if we turn on the angular momenta (as well as spin and isospin) no zero eigenmodes occur for angular momenta different than zero.

We now consider deformations of a bundle B^a_μ,

$$B_\mu \to B_\mu + b_\mu. \tag{5.2}$$

The linearized YM equations (4.27) read

$$\hat{D}_\mu\hat{D}_\mu b_0 - \hat{D}_\mu\hat{D}_0 b_\mu - [G_{\mu 0},b_\mu] = 0, \tag{5.3}$$

$$\hat{D}_\mu\hat{D}_\mu b_i - \hat{D}_\mu\hat{D}_i b_\mu - [G_{\mu i},b_\mu] + 4\epsilon_{ijk}\hat{D}_k b_j + 4b_i = 0 \tag{5.4}$$

[compare with (3.2)]. Using the identities

$$(\hat{D}_\mu D_0 - \hat{D}_0\hat{D}_\mu)b_\mu = [G_{\mu 0},b_\mu], \tag{5.5}$$

$$(\hat{D}_\mu\hat{D}_i - \hat{D}_i\hat{D}_\mu)b_\mu = [G_{\mu i},b_\mu] + 2\epsilon_{ijk}\hat{D}_k b_j \tag{5.6}$$

[compare with (3.5)] and imposing the "background" gauge

$$\hat{D}_\mu b_\mu \equiv \hat{\partial}_\mu b_\mu + [B_\mu,b_\mu] = 0, \tag{5.7}$$

Eqs. (5.3) and (5.4) become

$$\hat{D}_\mu\hat{D}_\mu b_0 - 2[G_{\mu 0},b_\mu] = 0, \tag{5.8}$$

$$\hat{D}_\mu\hat{D}_\mu b_i - 2[G_{\mu i},b_\mu] - 2\epsilon_{ijk}\hat{D}_k b_j + 4b_i = 0. \tag{5.9}$$

For the meron bundle (4.55), we get

$$-\frac{\partial^2}{\partial\tau^2}b^a_0 - \hat{\partial}_j\hat{\partial}_j b^a_0 - 2\epsilon^{abj}\hat{\partial}_j b^b_0 + 2b^a_0 = 0, \tag{5.10}$$

$$-\frac{\partial^2}{\partial\tau^2}b^a_i - \hat{\partial}_j\hat{\partial}_j b^0_i - 2\epsilon^{abj}\hat{\partial}_j b^b_i - 2\epsilon_{ijk}\hat{\partial}_k b^a_j + 4\epsilon^{abc}\epsilon_{bij}b^c_j + 6b^0_i = 0 \tag{5.11}$$

and the background gauge

$$\frac{\partial}{\partial\tau}b^a_0 + \hat{\partial}_i b^a_i + \epsilon^{abi}b^b_i = 0. \tag{5.12}$$

We now "diagonalize" these equations in terms of angular momenta, spin, and isospin. First, we list some properties of angular momenta, spin, and isospin operators. From (4.11), we get

$$\mathbf{L}^2_1 = \mathbf{L}^2_2 = -\tfrac{1}{4}\hat{\partial}_j\hat{\partial}_j = -\tfrac{1}{4}[\partial_\alpha\partial_\alpha - z_\alpha\cdot(z_\alpha\cdot\partial_\alpha\partial_\alpha + 1)], \tag{5.13a}$$

$$[L^i_p,L^j_q] = i\delta_{pq}\epsilon_{ijk}L^k_p,\quad p = 1,2. \tag{5.13b}$$

From (4.42) and (4.43) we have

$$S^a_1 b^b_i = -i\epsilon_{aij}b^b_j,\quad S^a_2 b^b_i = 0, \tag{5.14a}$$

$$\mathbf{S}_1^2 = 2, \quad \mathbf{S}_2^2 = 0. \tag{5.14b}$$

The isospin T is a three-dimensional orthogonal representation of the isospin group SU(2),a nd is given (like the spin) by

$$T^b b_i^a = i\epsilon^{abc} b_i^c \tag{5.15a}$$

$$\mathbf{T}^2 = 2 = 1(1+1). \tag{5.15b}$$

From (5.13)–(5.15), we get

$$\mathbf{T}\cdot\widehat{L}_1 b_i^a = -\tfrac{1}{2}\epsilon^{abj}\widehat{\partial}_j b_i^b, \tag{5.16a}$$

$$\mathbf{T}\cdot\mathbf{S}_1 b_i^a = \epsilon^{abc}\epsilon_{bij} b_j^c, \tag{5.16b}$$

$$\mathbf{L}_1\cdot\mathbf{S}_1 b_i^a = \tfrac{1}{2}\epsilon_{ijk}\widehat{\partial}_j b_k^a. \tag{5.16c}$$

Then (5.10) and (5.11) become

$$-\widehat{D}_\mu\widehat{D}_\mu b_0^a \equiv \Big(-\frac{d^2}{d^2} + 2\mathbf{L}_1^2 + 2(\mathbf{L}_1+\mathbf{T})^2 - 2\Big) b_0^a = 0, \tag{5.17}$$

$$\Big(-\frac{d^2}{d\tau^2} + 2\mathbf{L}_1^2 + 2(\mathbf{L}_1+\mathbf{S}_1+\mathbf{T})^2 - 2\Big) b_i^a = 0. \tag{5.18}$$

From (5.17) one sees that $-\widehat{D}_\mu\widehat{D}_\mu$ is a strictly positive operator. Therefore, (5.17) implies $b_0^a \equiv 0$. Thus, (5.18) is the only nontrivial deformation equation; it gives rise to the meron bundle stability operator

$$\mathscr{L} = -\frac{d^2}{d\tau^2} + 2\mathbf{L}_1^2 + 2(\mathbf{L}_1+\mathbf{S}_1+\mathbf{T})^2 - 2. \tag{5.19}$$

The gauge condition (5.12) takes the form

$$i2L_1^j b_j^a + \epsilon^{abj} b_j^b = 0. \tag{5.20}$$

The stability operator (5.19) commutes with $\mathbf{L}_1^2$, the total angular momenta $\mathbf{J}_1 = \mathbf{L}_1 + \mathbf{S}_1 + \mathbf{T}$, $\mathbf{L}_2$, and $\mathbf{S}_2(=0)$ (and therefore with $\mathbf{J}_2 = \mathbf{L}_2 + \mathbf{S}_2 + \mathbf{T} = \mathbf{L}_2 + \mathbf{T}$). Eigenvectors of $\mathscr{L}$ can be labelled by the quantum numbers

$$l = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \ldots, j_1, \quad l_2^3 = -l, -l+1, \ldots, l, \tag{5.21}$$

and of course, by $s_1 = 1$, $s_2 = 0$, $t = 1$ (isospin), $s_2^3 = 0$.

The spectrum of $\mathscr{L}$ extends from -2 to $+\infty$ [compare with the stability operator obtained from (4.53)]. The lowest point, -2, corresponds to $\mathbf{L}_1 = 0$, $\mathbf{S}_1 = -\mathbf{T}$. For each set of quantum numbers there is a branch of the spectrum starting from some point in $[-2, +\infty]$. For each point in the branch there are two linearly independent eigenvectors of (5.19). For each eigenvector, the multiplicity of the point is $(2j_1+1)(2j_2+1)$. Thus each point (or better each branch of the spectrum) has total multiplicity

$$2(2j_1+1)(2j_2+1). \tag{5.22}$$

We shall now show that there is only one branch of the spectrum which contains the origin—the one with $l = 0, j_1 = 0$, $j_2 = t = 1$. For $l = 0$ (i.e., $\mathbf{L}_1 = \mathbf{L}_2 = 0$),

$$\mathscr{L} = \frac{d^2}{d\tau^2} + 2(\mathbf{S}_1+\mathbf{T})^2 - 2. \tag{5.23}$$

We decompose the reducible representation $S_1 \otimes T = 1 \otimes 1$ into its irreducible components by the well known process

$$1 \otimes 1 = 0 \oplus 1 \oplus 2. \tag{5.24}$$

Thus the possible values of $(\mathbf{S}_1 + \mathbf{T})^2$ are 0, 2, and 6. Obviously only the branch with $(\mathbf{S}_1 + \mathbf{T})^2 = 0$ contains the origin.

The multiplicity of this branch is $2(0+1)(2\cdot 1+1) = 6$. The six linearly indepenent eigenvectors corresponding to the origin are

$$b_i^a = c_i^a e_{1,m}\cos\sqrt{2}\tau, \quad m = -1,0,1, \tag{5.25a}$$

$$b_i^a = c_i^a e_{1,m}\sin\sqrt{2}\tau, \quad m = -1,0,1, \tag{5.25b}$$

here c_i^a are constant [i.e., (z,τ) independent] vector fields to be chosen so that the gauge condition is satisfied, and $e_{1,m}$ are the basis vectors of the three-dimensional orthogonal representation of SU(2) (isospin representation). The background gauge (5.20) is satisfied by (5.25) if $c_i^a = c_a^i$. With this property, (5.25) are gauge equivalent to

$$b_i^a = \delta_i^a e_{1,m}\cos\sqrt{2}\tau, \quad m = 1,0,1, \tag{5.26a}$$

$$b_i^a = \delta_i^a e_{1,m}\sin\sqrt{2}\tau, \quad m = -1,0,1. \tag{5.26b}$$

For $l = \tfrac{1}{2}$, we have

$$\begin{aligned} L_1 \otimes S_1 \otimes T &= \tfrac{1}{2} \otimes 1 \otimes 1 \\ &= \tfrac{1}{2} \otimes \{0 \oplus 1 \oplus 2\} \\ &= \{\tfrac{1}{2}\} \oplus \{\tfrac{1}{2} \oplus \tfrac{3}{2}\} \oplus \{\tfrac{3}{2} \oplus \tfrac{5}{2}\}. \end{aligned} \tag{5.27}$$

The smallest value of j_1 is $\tfrac{1}{2}$, and the corresponding operator, $-(d^2/d\tau^2) + 1$, is strictly positive. The same is true for $l \geqslant 1$. Thus we have proven that the null space (a space of generalized eigenvectors) of the stability operator of the meron bundle has dimension equal to six, and the corresponding eigenvectors ae given by (5.26). Therefore, the manifold of meron bundle solutions is six-dimensional, i.e., the most general meron bundle solution has six parameters. Since the transformation (4.2) is a two-parameter family of transformations, the most general meron-antimeron pair solution at finite points of $\mathbb{R}^4$, is an eight-parameter solution.

For each value l of the angular momenta, and each value

$$j_1(l) = l-2, l-1, l, l+1, l+2, \quad j_1(l) \geqslant 0, \tag{5.28}$$

the spectrum is given by

$$\lambda = k^2 + 2l(l+1) + 2j_1(l)(j_1(l)+1) - 2, \quad k \in \mathbb{R}.$$

Some of the corresponding eigenvectors will satisfy the gauge condition (5.20), while others will not. Analyzing the transformation properties of $f^a = i2L_1^j b_j^a + \epsilon^{abj} b_j^b$ (via purely group theoretic arguments), one should be able to isolate the eigenvectors that satisfy (5.20). We do not pause to carry out this analysis here.

Note added in proof: After the main results of this paper had been completed, we learned (and subsequently received,[14] that Professor D.N. Williams has proven (by a different method) the unboundedness below of the stability operator of $\mathbb{R}^4$, and has studied regularized meron configurations (different from ours). We thank Professor Williams for communicating his results to us, and for sending Ref. 14 to us.

ACKNOWLEDGMENTS

We wish to thank Professor James Glimm and Dr.

Ricardo Schor for discussions during the completion of this work. It is also a pleasure to thank Professor Arthur Jaffe for his hospitality, extended to us at Harvard University, where part of this work was carried out.

APPENDIX

In this Appendix we summarize the technical formulas which one needs to go from fields on $s^3\times\mathbb{R}$ to fields on $\mathbb{R}^4$. The basic formula for vector fields is (4.16). For fields of the form (4.28), it takes the form

$$A^a(x) = f(\tau)\tfrac{1}{2}\eta_{a\alpha\beta}M^{(\mu)}_{\alpha\beta}$$

$$= f(\tau)[\delta^a_1(M^{(\mu)}_{23} + M^{(\mu)}_{14}) + \delta^a_2(M^{(\mu)}_{31} + M^{(\mu)}_{24}) + \delta^a_3(M^{(\mu)}_{12} + M^{(\mu)}_{34})] \quad \text{(A1)}$$

or

$$A_\mu(x) = f(\tau)\frac{1}{2i}\sigma_a\tfrac{1}{2}\eta_{a\alpha\beta}M^{(\mu)}_{\alpha\beta}$$

$$= -\frac{i}{2}f(\tau)[\sigma_1(M^{(\mu)}_{23} + M^{(\mu)}_{14}) + \sigma_2(M^{(\mu)}_{31} + M^{(\mu)}_{24}) + \sigma_3(M^{(\mu)}_{12} + M^{(\mu)}_{34})] \quad \text{(A2a)}$$

$$= -\tfrac{1}{2}f(\tau)g^{-1}\partial_\mu g, \quad \text{(A2b)}$$

where g is a singular gauge transformation. For the transformation (4.2) we get:

$$M^{(0)}_{kl} = z_k\frac{\partial z_l}{\partial x_0} - z_l\frac{\partial z_k}{\partial x_0} = 0, \quad k,l = 1,2,3, \quad \text{(A3a)}$$

$$M^{(j)}_{kl} = z_k\frac{\partial z_l}{\partial x_j} - z_l\frac{\partial z_k}{\partial x_j} = \frac{(b_0 - a_0)^2(x_k\delta_{lj} - x_l\delta_{kj})}{|x-a|^2|x-b|^2}, \quad \text{(A3b)}$$

$$M^{(0)}_{k4} = z_k\frac{\partial z_4}{\partial x_0} - z_4\frac{\partial z_k}{\partial x_0} = -\frac{x_k(2x_0 - a_0 - b_0)(b_0 - a_0)}{|x-a|^2|x-b|^2}, \quad \text{(A3c)}$$

$$M^{(j)}_{k4} = z_k\frac{\partial z_4}{\partial x_j} - z_4\frac{\partial z_k}{\partial x_j} = \frac{(b_0 - a_0)\{(x-a)\cdot(x-b)\}\delta_{kj}}{|x-a|^2|x-b|^2}. \quad \text{(A3d)}$$

Then (A2) gives

$$A_0(x) = f(\tau)i\,\frac{b_0 - a_0}{2}\,\frac{(2x_0 - a_0 - b_0\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-a|^2|x-b|^2}, \quad \text{(A4a)}$$

$$A_j(x) = f(\tau)i\,\frac{b_0 - a_0}{|x-a|^2|x-b|^2}\Big(-\tfrac{1}{2}(x-a)\cdot(x-b)\sigma_j + \frac{b_0 - a_0}{2}\epsilon_{jkl}x_k\sigma_l + x_j\boldsymbol{\sigma}\cdot\mathbf{x}\Big), \quad \text{(A4b)}$$

$$g(x) = \frac{(x_0 - a_0) - i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-a|}\,\frac{(x_0 - b_0) + i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-b|}, \quad \text{(A4c)}$$

$$g^{-1}(x) = \frac{(x_0 - a_0) + i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-a|}\,\frac{(x_0 - b_0) - i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-b|}. \quad \text{(A4d)}$$

For $f = -1$, we obtain the standard form of a meron–antimeron pair [compare with (2.25) or Ref. 17]. For $f = -2$, we obtain the pure gauge $g^{-1}\partial_\mu g$ which is singular at $x = a$ and $x = b$; formally it has a local charge density $\delta(x-a) - \delta(x-b)$. For

$$f(\tau) = -1\,\frac{e^{-\tau}}{e^{\tau} + e^{-\tau}} = -2\,\frac{(x-b)^2}{(x-a)^2 + (x-b)^2} \quad \text{(A5)}$$

we obtain the instanton at $x = a$ with size 1, in a singular gauge. For the map (4.4), we get

$$M^{(0)}_{kl} = 0, \quad \text{(A6a)}$$

$$M^{(j)}_{kl} = \frac{x_k\delta_{lj} - x_l\delta_{kj}}{|x-a|^2}, \quad \text{(A6b)}$$

$$M^{(0)}_{k4} = \frac{x_k}{|x-a|^2}, \quad \text{(A6c)}$$

$$M^{(j)}_{k4} = -\frac{(x_0 - a_0)\delta_{kj}}{|x-a|^2}. \quad \text{(A6d)}$$

Then formulas (A1) and (A2) give

$$A^a_\mu(x) = -f(\tau)\eta_{a\mu\nu}\frac{(x-a)_\nu}{(x-a)^2}, \quad \text{(A7a)}$$

$$g(x) = \frac{(x_0 - a_0) - i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-a|},$$

$$g^{-1}(x) = \frac{(x_0 - a_0) + i\boldsymbol{\sigma}\cdot\mathbf{x}}{|x-a|}. \quad \text{(A7b)}$$

For $f = -1$, we obtain the single meron solution at $x = a$ [compare with (2.5)]. For

$$f(\tau) = -2\,\frac{e^{-\tau}}{e^{\tau} + e^{-\tau}} = -\frac{2}{1 + (x-a)^2} \quad \text{(A8)}$$

we obtain the instanton solution

$$A^a_\mu = 2\eta_{a\mu\nu}\frac{(x-a)_\nu}{1 + (x-a)^2}\,\frac{1}{(x-a)^2} \quad \text{(A9)}$$

which is gauge equivalent to the standard form

$$A^a_\mu = 2\eta_{a\mu\nu}\frac{(x-a)_\nu}{1 + (x-a)^2}. \quad \text{(A10)}$$

This is also gauge equivalent to the instanton solution one obtains from (A3).

[1]A.M. Polyakov, Phys. Lett. B **59**, 82 (1975).
[2]A.A. Belavin *et al.*, Phys. Lett. B **59**, 85 (1975).
[3]V. de Alfaro, S. Fubini, and G. Furlan, Phys. Lett. B **65**, 163 (1976).
[4]G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976); C. Callan, R. Dashen, and D. Gross, Phys. Lett. B **65**, 354 (1976).
[5]G. 't Hooft, Phys. Rev. D **14**, 3432 (1976).
[6]M.F. Atiyah and R.S. Ward, Commun. Math. Phys. **55**, 17 (1977).
[7]M.F. Atiyah, N.J. Hitchin, and I.M. Singer, "Self-duality in four-dimensional Riemannian Geometry," preprint, 1978 (and references given there).
[8]C. Callan, R. Dashen, and D. Gross, "Toward a Theory of the Strong Interactions," to appear in Phys. Rev. D.
[9]J. Glimm and A. Jaffe, "Multiple Meron Solutions of the Classical Yang–Mills Equations," Phys. Lett. B **73**, 167–70 (1978).
[10]T. Jonsson, O. McBryan, F. Zirilli, and J. Hubbard, "An Existence Theorem for Multimeron Solutions to Classical Yang–Mills Field Equations," Harvard preprint.
[11]J. Glimm and A. Jaffe, Phys. Rev. Lett. **40**, 277–8 (1977).
[12]J. Glimm and A. Jaffe, "A Droplet Model of Quark Confinement," Phys. Rev. D **18**, 463–7 (1978).
[13]I. Bialynicki-Birula and J. Mycielski, in preparation.
[13(a)]Y. Nambu, "Topological Problems in Gauge Theories," Enrico Fermi Institute, preprint (1978).
[14]D.N. Williams, "Mathematical Remarks on Positivity and Self-Adjointness of the Euclidean, Yang–Mills Stability Operator," Michigan preprint.
[15]R. Hartshorne, *Algebraic Geometry* (Springer-Verlag, Berlin, 1977).
[16]S. Chern, *Geometry of Characteristic Classes*, Proceedings of the 13th biannual symposium, Canadian Math. Congress (1972), pp. 1–40.
[17]V. de Alfaro, S. Fubini, and G. Furlen, Phys. Lett. B **72**, 203 (1977).
[18]K. Yoshida, *Functional Analysis* (Springer-Verlag, Berlin 1968).
[19]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-Adjointness* (Academic, New York, 1975).
[20]K.M. Case, Phys. Rev. **80**, 797 (1950).
[21]E. Nelson, J. Math. Phys. **5**, 332 (1964).
[22]P.A.M. Dirac, Ann. Math. **27**, 429 (1936).
[23]G. Mack and A. Salam, Ann. Phys. **53**, 174 (1969).
[24]S. Adler, Phys. Rev. D **6**, 3445 (1972).
[25]M.P. Ryan and L.C. Shepley, *Homogeneous Relativistic Cosmologies* (Princeton U.P., Princeton, New Jersey, 1972).
[26]M. Lüscher, Phys. Lett. B **70**, 321 (1977).
[27]S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. I.

Gauge field singularities and noninteger topological charge[a)]

F. A. Schaposnik
Laboratorio de Física, Teórica Departamento de Física, UNLP C.C.N°67. 1900 La Plata, Argentina
J. E. Solomin
Departamento de Matemática UNLP. C.C.N.°67. 1900 La Plata, Argentina

(Received 17 April 1979; accepted for publication 25 May 1979)

We study the behavior of singular gauge field configurations under gauge transformations and we determine the relation between noninteger topological charge and the possibility of displacing singularities.

I. INTRODUCTION

Since the introduction of the first instanton solution,[1] many important physical effects have been attributed to its existence.[2–4]

At present, big efforts are being made in order to establish a theory of strong interactions based in quantum chromodynamics with a correct account of instanton effects.[5]

However, if only instanton contributions are considered in nonperturbative computations, important problems remain unsolved, namely those related to quark confinement. For this reason it has been stressed that other singular configurations may play a fundamental role.[6]

More recently there have been other suggestions emphasizing the relevance of configurations with noninteger topological charge in the quantization of non-Abelian theories.[7–9]

As it was stressed by Witten[9] these fractional charge configurations maintain the principal characteristics of the instanton picture: existence of an infinite number of degenerate vacua, possibility of resolution of the U(1) problem. At the same time, however, they may overcome its difficulties, in particular, those which arise from the conflict between the boundary condition of pure gauge field at infinity (which leads to instantons) and quark confinement.

Motivated by these remarks, we study in this paper the possible singularities of gauge fields in order to understand their behavior under gauge transformations. Then we will be able to establish our principal results stressing the relation between noninteger topological charge and the possibility of displacing gauge field singularities.

In Sec. II we discuss some interesting Abelian and non-Abelian examples which illuminate our results and which are established in Sec. III.

II. ABELIAN AND NON-ABELIAN EXAMPLES

We shall discuss in this section some examples of gauge field configurations having peculiar properties: both in the Abelian and in the non-Abelian cases they present singularities that one cannot displace by gauge transformations. Another characteristic of the configurations is that their "local" topological charge—to be defined below—is noninteger.

The first example, an Abelian one, resembles to the *c*-instantons introduced by Nielsen and Schroer in the context of the Schwinger model.[10] In the non-Abelian case the configuration we discuss is singular and it behaves at the singularity not as a pure gauge but as an "almost pure gauge".[11] It has nonzero but noninteger local topological charge and, as we will see in Sec. III, this fact is closely related to the gauge invariance of the singularities.

A. The Abelian case

Let us consider an Abelian gauge theory with gauge group U(1) in (Euclidean) two-dimensional space. We shall study a configuration of the form

$$A_\mu^N = \sum_{i=1}^N g_i(x)\epsilon_{\mu\nu}\frac{(x^\nu - a^{i\nu})}{(x^\nu - a^{i\nu})^2} \tag{2.1}$$

with

$$g_i(x) \underset{|x|\to\infty}{=} O(|x|^{-2}), \tag{2.2}$$

$$g_i(x) \underset{x\to a^i}{=} 1 + (x_\nu - a_\nu^i). \tag{2.3}$$

Because of conditions (2.2) and (2.3), A_μ^N is singular at the points a_μ^i $(a_\mu^i \in \mathbb{R}^2)$ and well-behaved at infinity.

We shall now compute the topological charge of A_μ^N from the expression

$$q[A_\mu^N] = \frac{1}{4\pi}\int_{\mathbb{R}^2}\epsilon_{\mu\nu}F^{\mu\nu}\,d^2x = \frac{1}{2\pi}\int_{\mathbb{R}^2}\epsilon_{\mu\nu}\partial^\nu A^\mu\,d^2x. \tag{2.4}$$

Since A_μ^N is $\mathbb{C}^\infty$ (excepting at points $x_\mu = a_\mu^i$), we can apply Stokes' theorem and write expression (2.5) in the form

$$q[A_\mu^N] = \lim_{R\to\infty}\frac{1}{2\pi}\oint_{S_R} A_\mu^N\,dx^\mu - \sum_{i=1}^N \lim_{\epsilon_i\to 0}\frac{1}{2\pi}\oint_{S_{\epsilon_i}} A_\mu^N\,dx^\mu, \tag{2.5}$$

where S_R is a circle of radius R and S_{ϵ_i} are small circles of radius ϵ_i surrounding each singularity. Then,

$$q[A_\mu^N] = \sum_{i=1}^N \lim_{\epsilon_i\to 0}\frac{1}{2\pi}\oint g_i\,d\Theta^i = N. \tag{2.6}$$

Now, near the singularities, A_μ^N behaves like a pure gauge:

$$A_\mu^N \simeq \partial_\mu \tan^{-1}\left(\frac{x_2 - a_2^i}{x_1 - a_1^i}\right), \quad x_\mu \sim a_\mu^i. \tag{2.7}$$

[a)]Financially supported by CIC Buenos Aires, Argentina.

Hence, the singularity can be displaced by a suitable gauge transformation. For example, for the $N = 1$ case, it is evident that the configuration

$$A'_\mu = A_\mu^{N=1} - \partial_\mu \tan^{-1}\left(\frac{x_2 - a_2^i}{x_1 - a_1^i}\right) \tag{2.8}$$

gauge–transformed of $A_\mu^{N=1}$ has the singularity removed to infinity, since we can write A'_μ in the form

$$A'_\mu = [g_1(x) - 1]\,\epsilon_{\mu\nu}\frac{(x_\nu - a_\nu^1)}{(x_\eta - a_\eta^1)^2}, \tag{2.9}$$

regular at $x_\mu = a_\mu^1$.

For general N one applies a chain of gauge transformations of the type (2.8), finally obtaining a regular configuration of the form

$$A'_\mu = \sum_{i=1}^{N} [g_i(x) - 1]\,\epsilon_{\mu\nu}\frac{(x_\nu - a_\nu^i)}{(x_\eta - a_\eta^i)^2}. \tag{2.10}$$

Of course

$$q[A'_\mu] = q[A_\mu^N] = N. \tag{2.11}$$

One can recognize in this procedure the method developed by Giambiagi and Rothe[12] to displace the singularities of 't Hooft's multi-instanton solution.[13] Of course, since our example is Abelian, the prescription in this case is evident.

However, if instead of configuration (2.1) we consider a gauge field of the form

$$A_\mu^{\mathbb{C}} = \mathbb{C}A_\mu^N = \mathbb{C}\sum_{i=1}^{N} g_i(x)\frac{(x_\nu^i - a_\nu^i)}{(x_\eta^i - a_\eta^i)^2}\epsilon_{\mu\nu} \tag{2.12}$$

with g_i defined as in Eqs. (2.2) and (2.3), and we try to displace the singularities as we did before, then we are faced with a problem. The singular gauge transformation that we must perform is of the form

$$A_\mu^{\mathbb{C}'} = A_\mu^{\mathbb{C}} - \sum_{i=1}^{N}\partial_\mu\left[\mathbb{C}\tan^{-1}\left(\frac{x_2^i - a_2^i}{x_1^i - a_1^i}\right)\right]. \tag{2.13}$$

For simplicity, let us consider the case $N = 1$, $a_\mu^1 = 0$. We see that the singular gauge transformation corresponds to the U(1) element

$$g = \exp(i\mathbb{C}\theta), \quad \theta = \tan^{-1}(x_2/x_1) \tag{2.14}$$

which is multivalued whenever $\mathbb{C}\notin Z$. If one excludes this kind of transformations, it is not possible to gauge away the singularities as we did before.

The topological charge of this configuration is

$$q[A_\mu^{\mathbb{C}}] = \mathbb{C}N \tag{2.15}$$

which in general is noninteger. Gauge fields similar to those given by Eq. (2.12) were already studied by Nielsen and Schroer in the context of the Schwinger model.[10] These authors found that configurations with $q = \frac{1}{2}$, the c-instantons, saturate certain functional integrals which show confinement in the Schwinger model.

More recently Witten[9] discussed field theories with noninteger topological charge in the context of the quark confinement problem. In fact, the first model he studied[9] corresponds to the Lagrangian density (in two-dimensions)

$$\mathscr{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\phi^* D^\mu\phi - M^2|\phi|^2 \tag{2.16}$$

with

$$\phi = \phi_1 + i\phi_2 \tag{2.17}$$

and

$$D_\mu = \partial_\mu - iA_\mu. \tag{2.18}$$

Finite action requires $|\phi|_{|x|\to\infty} = 0$ and

$$A_\mu \underset{|x|\to\infty}{=} \partial_\mu\alpha. \tag{2.19}$$

These conditions allow for configurations of the kind (2.12). It is only when $-M^2|\phi|^2$ is replaced by a symmetry breaking potential

$$V(\phi) = -\lambda(|\phi|^2 - a^2)^2 \tag{2.20}$$

that the topological charge must be integer since condition (2.19) is replaced by:

$$|\phi|^2 \underset{|x|\to\infty}{=} a^2, \quad D_\mu\phi \underset{|x|\to\infty}{=} 0, \tag{2.21}$$

thus forbidding $\mathbb{C}\notin Z$, since the contrary is only possible for multivalued ϕ.

B. The non-Abelian case

In analogy with the Abelian case, we will try to gauge away singularities of a particular gauge field configuration $A_\mu^{\mathbb{C}}$, for the non-Abelian case, with gauge group SU(2) in (Euclidean) four-dimensional space.

Let $A_\mu^{\mathbb{C}}$, be

$$A_\mu^{\mathbb{C}} = \sum_{i=1}^{N}\mathbb{C}f_i\,g_i^{-1}\partial_\mu g_i, \tag{2.22}$$

with

$$f_i \underset{x^2\to\infty}{=} \lambda_i^2/x^2, \quad f_i \underset{y_i^2\to 0}{=} 1 - y_i^2/\lambda_i^2, \tag{2.23}$$

$$y_\mu^i = x_\mu - a_\mu^i, \quad \mu = 1,2,3,4,$$

$$g_i = y_4^i - i\frac{\bar{\sigma}\cdot\bar{y}^i}{(y_i)^2}. \tag{2.24}$$

The configuration is singular at points $a_\mu^i \in \mathbb{R}^4$, $i = 1,2,\ldots,N$. For $C = 1$, expression (2.22) corresponds to 't Hooft's multi-instanton solution[13] and since at points a_μ^i

$$A_\mu^{\mathbb{C}=1} \underset{x\to a^i}{\sim} g_i^{-1}\partial_\mu g_i, \tag{2.25}$$

these singularities can be gauged away, as it was shown by Giambiagi and Rothe.[12]

Now, if $C\neq 1$, the problem is completely different: $A_\mu^{\mathbb{C}}$ is no more a pure gauge at the singularities and we will see that it is not possible to gauge away them as it was done in the instanton case.

For simplicity let us consider the case $N = 1$, $a_\mu^1 = 0$. If the singularity was gauge movable, there must exist a gauge transformation h such that

$$A_\mu^{\mathbb{C}} = h^{-1}A_\mu^{\text{reg}}h + h^{-1}\partial_\mu h \tag{2.26}$$

and

$$F_{\mu\nu}^{\mathbb{C}} = h^{-1}F_{\mu\nu}^{\text{reg}}h, \tag{2.27}$$

where A_μ^{reg} is the gauge transformed of $A_\mu^{\mathbb{C}}$, regular at $x_\mu = 0$.

Now,

$$F^{\mathbb{C}}_{\mu\nu} \underset{|x|\to 0}{=} \frac{4i\mathbb{C}(1-\mathbb{C})}{x^4}(x_\mu\sigma_{\nu\rho}x_\rho - x_\nu\sigma_{\mu\rho}x_\rho + x^2\sigma_{\mu\nu}). \tag{2.28}$$

(Here, $\sigma_{ij} = \frac{1}{2}\epsilon_{ijk}\sigma_k$, $\sigma_{i4} = \frac{1}{2}\sigma_i$.) It is easy to see now that $\mathrm{tr}(F^{\mathbb{C}}_{\mu\nu}F^{\mathbb{C}}_{\mu\nu})$ is singular for $\mathbb{C}\neq 0,1$ at $x=0$. But since

$$\mathrm{tr}(F^{\mathbb{C}}_{\mu\nu}F^{\mathbb{C}}_{\mu\nu}) = \mathrm{tr}(F^{\mathrm{reg}}_{\mu\nu}F^{\mathrm{reg}}_{\mu\nu}) \tag{2.29}$$

we have the l.h.s. singular (excepting for $\mathbb{C}\neq 0,1$) and the r.h.s. regular. Then, Eq. (2.27) cannot be satisifed. That is there is no gauge transformation which makes $A^{\mathbb{C}}_\mu$ regular at $x=0$ for $\mathbb{C}\neq 0,1$.

We can compute the topological charge of $A^{\mathbb{C}}_\mu$ from the expression

$$q = \frac{1}{32\pi^2}\,\mathrm{tr}\int d^4x\, F_{\mu\nu}{}^*F_{\mu\nu}. \tag{2.30}$$

we will use the relation[14]

$$\partial_\mu K^\mu = \mathrm{tr}F_{\mu\nu}{}^*F_{\mu\nu} \tag{2.31}$$

with

$$K_\mu = 4\epsilon_{\mu\alpha\beta\gamma}\mathrm{tr}(\tfrac{1}{2}A_\alpha\partial_\beta A_\gamma + \tfrac{1}{3}A_\alpha A_\beta A_\gamma) \tag{2.32}$$

which for regular configurations allows us to write

$$q = \frac{1}{32\pi^2}\lim_{R\to\infty}\int_{\sigma_R} d^3x\, K_\mu n^\mu, \tag{2.33}$$

where σ_R is a surface enclosing a sphere of radius R and n^μ is its outward normal. Of course if A_μ is not regular inside σ_R, expression (2.33) is no longer true. One has to surround each singularity with a small sphere and proceed in analogy with what was done in the Abelian case [see Eq. (2.5)].

It is easy to see that near the singularities

$$K_\mu = \epsilon_{\mu\alpha\beta\gamma}2\mathbb{C}^2(2\mathbb{C}/3-1)\,\mathrm{tr}(g_i^{-1}\partial_\alpha g_i g_i^{-1}\partial_\beta g_i g_i^{-1}\partial_\gamma g_i)$$

thus giving

$$q[A^{\mathbb{C}}_\mu] = \mathbb{C}^2(3-2\mathbb{C})\sum_{i=1}^{N}\lim_{\epsilon_i\to 0}\frac{1}{24\pi^2}\times\int_{\sigma_i}\epsilon_{\mu\alpha\beta\gamma}\, g_i^{-1}\partial_\alpha g_i g_i^{-1}\partial_\beta g_i g_i^{-1}\partial_\gamma g_i n^\mu\, d^3x. \tag{2.34}$$

(There is no contribution of the integral on σ_R since $K^\mu\to 0$ sufficiently fast for $R\to\infty$.) Then

$$q[A^{\mathbb{C}}_\mu] = \mathbb{C}^2(3-2\mathbb{C})\frac{1}{24\pi^2}\sum_{i=1}^{N}\int_{\sigma_i} dg, \tag{2.35}$$

where dg is the invariant volume on the group. Finally, we have

$$q[A^{\mathbb{C}}_\mu] = \mathbb{C}^2(3-2\mathbb{C})N. \tag{2.36}$$

In general q is noninteger for $\mathbb{C}\notin Z$. In particular, for $\mathbb{C}=\frac{1}{2}$ and $N=1$ we obtain $q=\frac{1}{2}$. This case is in fact related to the meron[15] since:

$$A^{\mathrm{meron}}_\mu = \tfrac{1}{2}g_1^{-1}\partial_\mu g_1. \tag{2.37}$$

Of course, expression (2.37) is singular both at the origin and at infinity and then formula (2.34) gives in this case the "local" charge at the origin since it does not take into account the contribution at infinity.

Finally, we want to point out that the action of configurations of the type (2.22) is infinite. This is due to the fact that the (gauge invariant) singularities of $A^{\mathbb{C}}_\mu$ render the action density not locally integrable for $\mathbb{C}=0,1$. In fact, Marino and Swieca[16] have already considered a large class of configurations carrying noninteger topological charge and they proved that they have necessarily infinite action.

From the examples discussed above we see that singularities and noninteger topological charges are closely related. In the next section we will try to understand more deeply the nature of the relation since, as it has been suggested, these types of configurations may play a central role in the confinement mechanism.

III. GAUGE FIELD SINGULARITIES

In this section we shall clarify the peculiar characteristics of the examples discussed before. To this end we will establish a classification of gauge field singularities which will allow us to understand their behavior under gauge transformations.

We then discuss some topological properties of singular gauge configurations and stress the relation between noninteger topological charges and the possibility of displacing an isolate singularity by a gauge transformation.

We will consider a gauge field A_μ on $\mathbb{R}^n$ with gauge group G [For the cases of interest we will take $G = U(1)$ and SU(2).]

We will say that A_μ is regular at $x=x_0$ if it is $\mathbb{C}^\infty$ in a neighborhood of x_0. We can consider that A_μ is defined on S^n [17] with an eventual singularity at the point at infinity. In general A_μ can also present singularities on $\mathbb{R}^n$. We will call S the set of all possible singularities of A_μ.

The behavior of A_μ near its singularities determines whether or not it defines a connection on a principal fibre bundle $P(S^n,G)$ with base manifold S^n. We recall that in the framework of the fibre bundle theory, a gauge transformation corresponds to a change of trivialization of $P(S^n,G)$.[18]

We will distinguish two possible types of singularities of A_μ at $x_0\in S^n$:

(*i*) *Gauge-movable* (*GM*) *singularities*: Those singularities which allow the definition of a connection on a principal fibre bundle $P(S^n,G)$. These type of singularities are produced by the particular trivialization in which one is considering the connection. Then, they can be removed by a gauge transformation.

(*ii*) *Gauge-invariant* (*GI*) *singularities*: Those singularities which prevent the definition of a connection on a principal fibre bundle $P(S^n,G)$ from A_μ. That is, the singularities that are not evitable by a gauge transformation.

If one is faced with a GM singularity of A_μ at x_0, one can obtain an A'_μ, gauge-transformed of A_μ, regular at x_0, by a suitable singular gauge transformation. An example of this kind is the 't Hooft multi-instanton configuration[13] [given by expression (2.22) with $\mathbb{C}=1$] with singularities at $S=\{a^i_\mu, i=1,2,\ldots,N\}$. In fact, in Ref. 12 it is given the explicit gauge transformation, singular at points of S, which sends all the singularities to infinity, thus giving a multi-instanton configuration regular on $\mathbb{R}^n$.

In the fibre bundle language, this corresponds to passing from a trivialization defined on $S^n - S$ to another one defined on $S^n - \{\infty\}$

Of course this procedure is also applicable to the Abelian example defined by Eq. (2.1), since it also presents a GM singularity.

On the other hand, if the singularity is gauge-invariant, it cannot be removed by a change of trivialization. To displace such a singularity it will be necessary to make another kind of transformation associated with other possible invariances of the field theory. An example of this kind is provided by the meron configuration[15] [Eq. (2.38)] which is singular both at the origin and at infinity. The singularities can be displaced only by a conformal transformation to two distinct points $u \neq v$. In general this kind of transformations prevent the displacement of two singularities to the same point.

The Abelian configuration (2.12) and the non-Abelian one, [Eq. (2.22) with $\mathbb{C} \notin Z$ also present GI singularities.

It is well known that a principal fibre bundle $P(S^n, G)$ is topologically characterized by an element of $\Pi_{n-1}(G)$ [20] which, for the case of interest is $\Pi_1(\mathrm{U}(1)) = \Pi_3(\mathrm{SU}(2)) \simeq Z$.

Then, if the gauge field configuration A_μ defines a connection on a principal fibre bundle $P(S^n, G)$, we can compute its topological characterization from the expression

$$q = \int \mathscr{D}, \tag{3.1}$$

where the topological density $\mathscr{D}$ is given by

$$\mathscr{D} = \frac{1}{4\pi} F_{\mu\nu} \epsilon^{\mu\nu} d^2x, \quad n = 2, \quad G = \mathrm{U}(1), \tag{3.2}$$

$$\mathscr{D} = \frac{1}{64\pi^2} \operatorname{tr}(F_{\mu\nu} F_{\alpha\beta} \epsilon^{\mu\nu\alpha\beta}), \quad n = 4, \quad G = \mathrm{SU}(2). \tag{3.3}$$

Even for A_μ with GI singularities [that is, not defining a principal fibre bundle $P(S^n, G)$] the quantity given by expression (3.1) can be considered as a topological invariant associated to A_μ, since its value does not change when A_μ is locally deformed. Of course, this quantity does not correspond in general to the topological characterization of a fibre bundle $P(S^n, G)$.

If A_μ has a singularity at $x = x_0$ $(x_0 \neq \infty)$, applying Stokes' theorem one obtains

$$q = \lim_{\epsilon \to 0} \int_{|x - x_0| = \epsilon} K, \tag{3.4}$$

where

$$K = K_\mu \, dx^\mu, \qquad n = 2$$
$$K_\mu = \frac{1}{2\pi} A_\mu, \qquad G = \mathrm{U}(1); \tag{3.5}$$

$$K = K_\mu n^\mu \, d^3x, \quad n = 4, \quad G = \mathrm{SU}(2),$$

$$K_\mu = \frac{1}{32\pi^2} \epsilon_{\mu\nu\lambda\sigma} \left[A_\nu \left(F_{\lambda\sigma} - \tfrac{2}{3} A_\lambda A_\sigma \right) \right]. \tag{3.6}$$

We can interpretate expression (3.4) saying that the topological charge q is concentrated at $x = x_0$. It is then natural to consider the r.h.s. of Eq. (3.4) as giving the *local topological charge* of A_μ at x_0. (Of course this concept is not gauge invariant. For instance at a GM singularity the local charge can be displaced from x_0 to another point by a singular gauge transformation.)

If A_μ has several singularities, q will be given by the sum of all local topological charges.

We will now establish the following proposition.

Proposition: Let x_0 be an isolated singularity of A_μ. If the local topological charge

$$q = \lim_{\epsilon \to 0} \int_{|x - x_0| = \epsilon} K \tag{3.7}$$

is noninteger ($q \notin Z$), then x_0 is a gauge-invariant singularity. (If $x_0 = \infty$, see Ref. 19.)

Demonstration: Suppose that x_0 is a GM singularity of A_μ and let us call Ψ a gauge transformation, regular on $\mathscr{U} - \{x_0\}$ such that the gauge-transformed field A'_μ is regular at x_0. (Here $\mathscr{U}$ is a neighborhood of x_0 not containing any other singularity of A_μ.)

Let us take a function f, smooth $f: S^n \to \mathbb{R}$ with

$$f \equiv 1 \qquad \text{on } \mathscr{U},$$
$$f \equiv 0 \qquad \text{on a neighborhood of all other singularities of } A_\mu \text{ on } S^n \tag{3.8}$$

Then

$$\widetilde{A}_\mu = f A_\mu \tag{3.9}$$

defines a configuration with a singularity at $x = x_0$. The local charge of $\widetilde{A}_\mu$ is of course equal to that of A_μ. This gauge field A_μ defines a connection on the principal fibre bundle $\widetilde{P}(S^n, G)$ constructed from the transition function Ψ. This means that the local charge of A_μ, Eq. (3.7), coincides with the topological charge of $\widetilde{P}(S^n, G)$. But this charge must be an integer, in contradiction with the proposition's hypothesis. Then, if the local charge of A_μ at x_0 is not an integer, the singularity is gauge invariant, Q.E.D.

In the examples of configurations with noninteger topological charges [Eqs. (2.11) and (2.22)], we have seen that it was not possible to find a gauge transformation in order to displace the singularities. The preceding proposition generalizes this result giving a geometrical interpretation of this fact.

We can apply the proposition to the configuration discussed by Marino and Swieca,[16] already mentioned in Sec. II, which has a noninteger local topological charge (at ∞). Then, the singularity originated by the ill definiteness of A_μ at the point at infinity is not gauge removable.

Finally, we want to comment on the relation between our results and a discussion given by Crewther in Refs. 7 and 8. In Sec. 2 of Ref. 7 there are examples [related to those of Ref. 16] of configurations which cannot be compactified. In the context of our discussion, this corresponds to configurations which do not define a connection on a principal fibre bundle with base S^n.

Then, from our results we see that the configurations that are important from the point of view of Ref. 8 must have GI singularities.

[1]A. Belavin, A. Polyakov, A. Schwartz, and Y. Tyupkin, Phys. Lett. B **59**, 85 (1975).

[2]G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D **14**, 3442 (1976).
[3]R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 1 (1976).
[4]C. Callan, R.F. Dashen, and D.J. Gross, Phys. Lett. B **63**, 334 (1976).
[5]C. Callan, R.F. Dashen, and D.J. Gross, Phys. Lett. B **78**, 307 (1978).
[6]C. Callan, R.F. Dashen, and D.J. Gross, Phys. Lett. B **66**, 375 (1977); Phys. Rev. D **17**, 2717 (1978).
[7]R.J. Crewther, Phys. Lett. B **70**, 349 (1977); Schladming Lectures (1978).
[8]R.J. Crewther, Preprint CERN 2546 (1978).
[9]E. Witten, Harvard Preprint HUTP-78/A042.
[10]N.K. Nielsen and B. Schroer, Phys. Lett. B **66**, 475 (1977).
[11]Y. Nambu, Cal. Tech. Preprint CALT 68-634 (1977).
[12]J.J. Giambiagi and K.D. Rothe, Nucl. Phys. B **129**, 111 (1977).
[13]'t Hooft (unpublished).
[14]See for example, R. Jackiw, C. Nohl, and C. Rebbi, Workshop on Solitons (Salerno, 1977).
[15]V. de Alfaro, S. Fubini, and G. Furlan, Phys. Lett. B **65**, 1631 (1977).
[16]E.C. Marino and J.A. Swieca, P.U.C., Nota Científica 7/78 (1878).
[17]This consideration does not introduce any assumption. If $\lim_{|x|\to\infty} A_\mu$ (and of all its derivatives) is unique, the point at infinity is a regular point; if not, $x = \infty$ will be considered a singular point. For instance, if $A_\mu \to g^{-1}\partial_\mu g$ with g not homotopically trivial, the point at infinity will be considered singular.
[18]See for example, D.A. Popov, Teor. Mat. Fiz. **24**, 347 (1975).
[19]If $x = \infty$, expression (3.4) must be replaced by $q = \lim_{R\to\infty} \int_{|x|=R} K$.
[20]N. Steenrod, *The Topology of Fibre Bundles* (Princeton U.P., Princeton, N.J., 1951).

Casimir invariants and characteristic identities for generators of the general linear, special linear and orthosymplectic graded Lie algebras*

P. D. Jarvis
Department of Physics, University of Southampton, Southampton S09 5NH, U.K.
H. S. Green
Department of Mathematical Physics, University of Adelaide, Adelaide, South Australia 5001

(Received 7 February 1979)

We present the commutation and anticommutation relations, satisfied by the generators of the graded general linear, special linear and orthosymplectic Lie algebras, in canonical two-index matrix form. Tensor operators are constructed in the enveloping algebra, including powers of the matrix of generators. Traces of the latter are shown to yield a sequence of Casimir invariants. The transformation properties of vector operators under these algebras are also exhibited. The eigenvalues of the quadratic Casimir invariants are given for the irreducible representations of ggl($m|n$), gsl($m|n$), and osp($m|n$) in terms of the highest-weight vector. In such representations, characteristic polynomial identities of order ($m + n$), satisfied by the matrix of generators, are obtained in factorized form. These are used in each case to determine the number of independent Casimir invariants of the trace form.

1. INTRODUCTION AND SUMMARY OF RESULTS

The concepts of Fermi–Bose supersymmetry, or "supersymmetry," has provided deep new insights in theoretical particle physics, both in the context of dual theories[1] and space–time symmetries.[2,3] So far, there has been a failure to provide realistic physical models,[4] although the supergravity models are promising candidates.[5] However, apart from applications to models of elementary particles, the same types of structures also arise naturally in other physical settings, for example in the mechanics of classical systems and their quantization,[6] and therefore form an important study in their own right.

Space–time supersymmetry and its analogs are formally described as graded Lie algebras.[7] In such an algebra, there is a grading homomorphism which maps each element, X, into an element (X) belonging to the group Z_2 of integers under addition modulo 2. The conditions for a graded Lie algebra are that the algebra product be graded antisymmetrical,

$$[X,Y] = -(-1)^{XY}[Y,X], \tag{1}$$

and a graded derivation,

$$[X[YZ]] = [[XY]Z] + (-1)^{XY}[Y[XZ]]. \tag{2}$$

Mathematical investigations of graded Lie algebras have been vigorously pursued recently by Kac,[8] Rittenberg and collaborators,[9] Backhouse,[10] and others. It has been found that the sequences of the general linear, special linear, and orthosymplectic graded Lie algebras, denoted here by ggl($m \mid n$), gsl($m \mid n$), and osp($m \mid n$), respectively, share many features in common with the sequences of classical Lie algebras gl(n), sl(n), o(n), and sp(n). Moreover, the space–time conformal and Poincaré supersymmetry algebras are precisely gsl(4 | 1), and a contraction of osp(1 | 4), respectively. A complete classification of all simple graded Lie algebras has now been obtained, and some aspect of the representation theory treated.[8,9] For example, a partial-wave expansion has been developed[11] for products of unitary irreducible representations of Poincaré supersymmetry, in the massive case. Meanwhile generalizations such as Z_p-graded algebras,[12] and trilinear supersymmetry algebras,[13,14] have been considered.

The present paper is concerned with the construction and evaluation of the Casimir invariants of some of the classical graded Lie algebras. A long-established precedent exists from studies of the classical Lie algebras. Racah,[15] generalizing the quadratic operation of Casimir,[16] first constructed invariants C_p of arbitrary order in the generators, and considered the problem of determining a complete set of independent invariants for a semisimple Lie algebra. He gave a solution in the form of certain determinantal invariants[17]; these were also considered, for U(n), by Biedenharn and Louck.[18] Independent of the existence of invariant operators is the question of the generalization, to $n \times n$ matrices over an arbitrary associative algebra, of the Cayley–Hamilton identity. Lehrer-Ilamed[19] showed that a system of $n \times n$ identities does indeed exist. The physical relevance of this was highlighted in connection with symmetry breaking in SU(3).[20–23] In fact, the coefficients occurring in these characteristic identities, for Lie algebras, are more closely related to the Casimir invariants C_p, rather than the determinantal invariants. General hierarchies of characteristic identities were established for gl(n), sl(n), o(n), and sp(n) by Bracken and Green,[24] Green,[25] and others.[26–29] In parallel with these developments, there has been much work on the problem of determining the eigenvalues of the higher-order Casimir invariants.[30]

In the present work, we consruct the Casimir invariants C_p for the graded Lie algebras ggl($m|n$), gsl($m|n$) ($m \neq n$), and osp($m|n$) (n even). The eigenvalues of C_2 are given explicitly, for those irreducible representations possessing a highest-weight vector. We also determine in each case the number of independent invariants C_p. Our work relies on

 0022-2488/79/102115-08$01.00

establishing characteristic polynomial identities satisfied by the generators of the respective graded Lie algebras, acting in such irreducible representations. This is done by an extension of the methods of Bracken and Green.[24] Recently Backhouse[31] has also considered the computation of the invariants of the graded Lie algebras, but along different lines, and without the use of characteristic identities.

Central to our technique is the use of the canonical two-index Gel'fand–Okubo form of the generators. These are thus denoted generically as a matrix $X^A{}_B, A,B = 1,...,m+1$. The graded algebra product is realized by specified commutation and anticommutation relations amongst the generators. In order to handle these defining relations in a uniform way, we develop the index notation consistently with the grading homomorphism. In fact, the latter is applied in the first instance to the index set itself. Thus there are "even" indices $a,b,\cdots = 1,2,...,m$, and "odd" indices $\alpha,\beta\cdots = 1,2,...,n$ on which the grading homomorphism takes the value 0 and 1, respectively. The generators themselves are then graded such that

$$(X^A{}_B) \equiv \binom{A}{B} \equiv ((A) + (B)), \tag{3}$$

with addition modulo 2. Obviously $X^a{}_b, X^\alpha{}_\beta$ are even, while $X^a{}_\beta, X^\alpha{}_b$ are odd. Moreover, the function

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix} \equiv ((A) + (B))\cdot((C) + (D)) \tag{4}$$

is always even, save when $\binom{A}{B}$ and $\binom{C}{D}$ are both odd: It obviously signals the existence of commutation or antiacommutation relations for the generators $X^A{}_B$ and $X^C{}_D$. It is convenient, for each of the functions,

$$(S) = \begin{pmatrix} A & C & E & \cdots \\ B\ \cdot & D & \cdots & \cdots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{pmatrix} \equiv ((A) + (B) + \cdots)\cdot((C) + \cdots)\cdot(\cdots),$$

to associate a parity factor,

$$[S] = (-1)^{(S)}, \tag{5}$$

while some of the formulas simplify using the grading defined by

$$(\bar{S}) = (1 + (S)). \tag{6}$$

In the sequel, the index calculus will prove an invaluable tool in establishing general results without resorting to a tedious separation into even and odd cases.

The main results of the paper are as follows. We formulate the defining relations of the generators $X^A{}_B$, $A,B = 1,...,m+n$, with the help of the index calculus. For gsl($m|n$) ($m \neq n$) the generators are projected from the ggl($m|n$) generators by means of a certain trace condition. For osp($m|n$) this is effected through the introduction of a graded-symmetrical metric tensor G_{AB}. In each case it is shown that the matrix powers

$$(\hat{X}^0)^A{}_B = \delta^A{}_B,$$
$$(\hat{X}^1)^A{}_B = X^A{}_B[B], \tag{7}$$
$$(\hat{X}^{p+1})^A{}_B = X^A{}_C[C](\hat{X}^p)^C{}_B,$$

with a summation convention applied on the index C, with sign $[C]$, are tensor operators of the enveloping algebra, transforming like $X^A{}_B$. The transformation properties of a vector operator V^A (that is, a tensor operator belonging to the defining representation) under the $X^A{}_B$, are also identified. The graded traces of the matrix powers,

$$C_p = (\hat{X}^p)^A{}_A[A], \tag{8}$$

are the required Casimir invariants.

A set of generators, numbering $m+n$ for ggl($m|n$), $m+n-1$ for gsl($m|n$) ($m \neq n$), and $[m/2] + n/2$ for osp($m|n$) (n even, where $[m/2]$ is integral, and $(m-1)/2 \leqslant [m/2] \leqslant m/2$), can be identified as simultaneously diagonalizable, and serve to define weight vectors in irreducible representations. In those cases where a highest weight vector exists, the eigenvalue of C_2 is calculated; higher-order invariants could also be evaluated directly, in principle. Bracken and Green[24] have shown how, for each of gl(n), sl(n), o(n), and sp(n) (n even), a vector operator may be decomposed into a sum of raising and lowering operators, and thus have obtained characteristic polynomial identities. These results are here transferred to the graded Lie algebras under study, and characteristic polynomial identities are obtained in the factorized form

$$\prod_{a=1}^{m} (\hat{X} - m_r I) \prod_{\alpha=1}^{n} (\hat{X} - n_s I) = 0, \tag{9}$$

Involving coefficients of $I^A{}_B \equiv \delta^A{}_B$ depending upon the highest-weight labels, and the matrix $\hat{X}^A{}_B = X^A{}_B[B]$, rather than $X^A{}_B$ itself. The identities (9) show directly that the number of independent invariants C_p is, in general, $(m+n)$ for ggl($m|n$), and $(m+n-1)$ for gsl($m|n$) ($m \neq n$). For osp($m|n$), there are further conditions related to the symmetry of the various matrix powers, leading to results such as

$$C_3 = \tfrac{1}{2}(m - n - 2)C_2 \tag{10}$$

and the number of invariants is, in general, $([m/2] + n/2)$.

In Secs. 2, 3, and 4 below, the program outlined above is carried out for ggl($m|n$), gsl($m|n$) ($m \neq n$), and osp($m|n$) (n even), respectively. Some concluding remarks are made in Sec. 5, and the specific examples of gsl(2|1), osp(1|2), and osp(1|4) are discussed.

2. GRADED GENERAL LINEAR ALGEBRA ggl(m/n)

The graded general linear algebra ggl($m|n$) is generated by $(m+n)^2$ matrices in block form,

$$X = \begin{pmatrix} E_0 & 0_{01} \\ 0_{10} & E_1 \end{pmatrix}, \tag{11}$$

where E_0, E_1 are square matrices of dimensions $m \times m$ and $n \times n$, respectively, and 0_{01}, 0_{10} are $m \times n$, $n \times m$ matrices, respectively. The grading is defined by

$$(X) = \begin{cases} 0 & \text{if } 0_{01} = 0_{10} = 0, \\ 1 & \text{if } E_0 = E_1 = 0, \end{cases} \tag{12}$$

and the set of matrices made into a graded Lie algebra by defining the bracket product in terms of the ordinary matrix product defined by

$$[XY] = XY - (-1)^{(X)(Y)}YX. \tag{13}$$

A suitable set of generators for ggl($m|n$) are the $(m+n)^2$ matrices

$$(e^A{}_B)_C{}^D = \delta^D{}_B\delta^A{}_C, \quad A,B = 1,...,m+n \tag{14}$$

which satisfy the product rule

$$e^A{}_B e^C{}_D = \delta^C{}_B e^A{}_D. \tag{15}$$

The commutation and anticommutation relations of the $e^A{}_B$ may now be written out explicitly, using (13) and (15). Defining "even" and "odd" indices

$$a,b = 1,...,m, \quad \alpha,\beta = 1,...,n, \tag{16}$$

it is easily verified that

$$\begin{aligned}
&[e^a{}_b,e^c{}_d] = \delta^c{}_b e^a{}_d - \delta^a{}_d e^c{}_b,\\
&[e^a{}_b,e^c{}_\delta] = \delta^c{}_b e^a{}_\delta,\\
&[e^a{}_b,e^\gamma{}_d] = -\delta^a{}_d e^\gamma{}_b,\\
&[e^a{}_b,e^\gamma{}_\delta] = 0,\\
&\{e^a{}_\beta,e^c{}_\gamma\} = 0,\\
&\{e^a{}_\beta,e^\gamma{}_d\} = \delta^\gamma{}_\beta\, e^a{}_d + \delta^a{}_d e^\gamma{}_\beta,\\
&[e^a{}_\beta,e^\gamma{}_\delta] = \delta^\gamma{}_\beta\, e^a{}_\delta,\\
&[e^\alpha{}_\beta,e^\gamma{}_c] = \delta^\gamma{}_\beta\, e^\alpha{}_c,\\
&[e^\alpha{}_\beta,e^\gamma{}_\delta] = \delta^\gamma{}_\beta e^\alpha{}_\delta - \delta^\alpha{}_\delta\delta^\gamma{}_\beta,
\end{aligned} \tag{17}$$

where [,] stands for commutator, and { , } for anticommutator. These relations are therefore the defining relations for the generators, $E^A{}_B$, of ggl($m|n$) in any representation. Using the index calculus of (3)–(6), (17) may be rewritten in the unified form

$$[E^A{}_B,E^C{}_D]_{-\left[\begin{smallmatrix}A & \cdot & C\\ B & \cdot & D\end{smallmatrix}\right]} = \delta^C{}_B E^A{}_D - \begin{bmatrix}A & \cdot & C\\ B & \cdot & D\end{bmatrix}\delta^A{}_D E^C{}_B. \tag{18}$$

It follows immediately that if $V^C{}_D$ is a tensor operator transforming under the $E^A{}_B$ in the same way as $E^C{}_D$ (that is, in the adjoint representation), then the trace $V^C{}_C$ is an invariant operator which commutes with the $E^A{}_B$ (since $\left[\begin{smallmatrix}A & \cdot & C\\ B & \cdot & D\end{smallmatrix}\right] = +1$, for any C). Furthermore, if $W^C{}_D$ is a second adjoint operator, then the graded product

$$X^C{}_D = V^C{}_E[E]W^E{}_D \tag{19}$$

also transforms in the adjoint representation.

Thus we can define the matrix powers

$$\begin{aligned}
&(\hat{E}^0)^C{}_D = \delta^C{}_D,\\
&(\hat{E}^{p+1})^C{}_D = E^C{}_E[E](\hat{E}^p)^E{}_D, \quad p = 0,1,\cdots
\end{aligned} \tag{20}$$

all of which transform in the adjoint representation, such that the traces

$$C_p = (\hat{E}^p)^C{}_C[C] \tag{21}$$

are Casimir invariants.

(17) indicates that ggl($m|n$) is associated with an underlying Lie algebra gl(m)×gl(n), generated by the $E^a{}_b$ and $E^\alpha{}_\beta$, respectively. The odd generators $E^a{}_\beta$ and $E^\alpha{}_b$ transform as the $m\times\bar{n}$ and $\bar{m}\times n$ representations, respectively. We can speak of the weights of gl(m)×gl(n) as the weights of ggl($m|n$) itself. We take the $(m+n)$ generators (no summation on a,α):

$$E^a{}_a, \ a = 1,...,m, \quad E^\alpha{}_\alpha, \ \alpha = 1,...,n \tag{22}$$

as the Cartan subalgebra. A vector has weight $(\mu|\nu) = (\mu_1,\mu_2,...,\mu_m|\nu_1,\nu_2,...,\nu_n)$ if it is a simultaneous eigenvector of the $E^a{}_a, E^\alpha{}_\alpha$ with eigenvalues μ_a, ν_α. Connections between the subsets (μ), (ν) are provided by the odd generators $E^a{}_\beta$, $E^\alpha{}_b$, which take weights $(\mu_a|\nu_\alpha)$ into $(\mu_a+1|\nu_\alpha-1)$ and vice versa. An ordering for the weights $(\mu|\nu)$ may be introduced in terms of the ordering of the subsets (μ) and (ν), such that

$$\begin{aligned}
(\mu|\nu) > (\mu'|\nu') \quad &\text{if } (\mu) > (\mu')\\
&\text{or } (\mu) = (\mu') \quad \text{and } (\nu) > (\nu').
\end{aligned} \tag{23}$$

Consider now those representations of ggl($m|n$) processing a highest-weight vector. (For simple graded Lie algebras, it has been established[8,9] that any irreducible representation may be labelled by its highest-weight vector.) We can immediately compute the eigenvalues of

$$\begin{aligned}
&C_1 = E^a{}_a + E^\alpha{}_\alpha,\\
&C_2 = E^a{}_b E^b{}_a - E^\alpha{}_\beta E^\beta{}_\alpha + E^\alpha{}_b E^b{}_\alpha - E^a{}_\beta E^\beta{}_a,
\end{aligned} \tag{24}$$

acting on this vector, by rearranging terms so that it is either annihilated by, or an eigenvector of, the various contributions. The results are

$$\begin{aligned}
C_1 &= \sum_{a=1}^{m}\mu_a + \sum_{\alpha=1}^{n}\nu_\alpha,\\
C_2 &= \sum_{a=1}^{m}\mu_a(\mu_a + m - n + 1 - 2a)\\
&\quad - \sum_{\alpha=1}^{n}\nu_\alpha(\nu_\alpha + m + n + 1 - 2\alpha).
\end{aligned} \tag{25}$$

The defining relations (17), (18) also identify the transformation property of a vector operator V^C under the $E^A{}_B$ to be

$$[E^A{}_B,V^C]_{-\left[\begin{smallmatrix}A & \cdot & C\\ B\end{smallmatrix}\right]} = \delta^C{}_B V^A. \tag{26}$$

V^C provides, in particular, vector operators V^c, V^γ transforming under gl(m)×gl(n) as $m\times 1$, $1\times n$ respectively. Bracken and Green[24] have shown that a vector operator for gl(m) may always be broken into a sum of m terms, each of which changes one of the highest-weight labels by 1, and commutes with the remaining labels. The same is therefore true of V^C, and we may write

$$\begin{aligned}
&V^C = \sum_{a=1}^{m}V^C_{(a)} + \sum_{\alpha=1}^{n}V^C_{(\alpha)},\\
&[\mu_a,V^C_{(b)}]_- = \delta_{ab}V^C_{(b)},\\
&[\nu_\alpha,V^C_{(\beta)}]_- = \delta_{\alpha\beta}V^C_{(\beta)}.
\end{aligned} \tag{27}$$

It therefore follows from (25) that

$$\begin{aligned}
&[C_2,V^C_{(a)}]_- = (2\mu_a + m - n - 2a)\delta^C{}_A V^A_{(a)},\\
&[C_2,V^C_{(\alpha)}]_- = -(2\nu_\alpha + m + n - 2\alpha)\delta^C{}_A V^A_{(\alpha)}.
\end{aligned} \tag{28}$$

On the other hand, by direct computation using (21) and (26),

$$[C_2,V^C]_- = 2E^C{}_A[A]V^A - (m-n)\delta^C{}_A V^A, \tag{29}$$

and this is true also of $V^C_{(a)}$, $V^C_{(\alpha)}$. Thus

$$(E^C{}_A[A] - (\mu_a + m - n - a)\delta^C{}_A)V^A_{(a)} = 0,$$
$$(E^C{}_A[A] + (\nu_\alpha + n + \alpha)\delta^C{}_A)V^A_{(\alpha)} = 0. \quad (30)$$

Since this is true for any vector operator V^C, with components $V^C_{(a)}$, $V^C_{(\alpha)}$, we have, as a matrix equation[24,25]

$$\prod_{a=1}^{m} (\hat{E} - (\mu_a + m - n - a)) \times \prod_{\alpha=1}^{n} (\hat{E} + (\nu_\alpha + n + \alpha)) = 0, \quad (31)$$

where

$$(\hat{E})^A{}_B = E^A{}_B[B].$$

(31) is the desired characteristic polynomial identity, for ggl$(m|n)$, of order $(m + n)$.

It can now be seen how many independent invariants ggl$(m|n)$ of the trace form exist. It is known from the work on U(n)[26] that at most $(m + n)$ independent invariants can be constructed for gl$(m)\times$gl(n); hence there should be at least this number for ggl$(m|n)$. However, from (31), any trace such as C_{m+n+1} may be rewritten in terms of lower order traces. We thus conclude that, for ggl$(m|n)$, $m + n$ independent invariant operators C_p exist.

3. GRADED SPECIAL LINEAR ALGEBRA gsl$(m|n)$ $(m\neq n)$

The $(m + n)^2$-1 generators of gsl$(m|n)$ may be introduced by defining (for $m\neq n$)

$$A^A{}_B = E^A{}_B - \frac{1}{m-n}\delta^A{}_B[B](E^C{}_C). \quad (32)$$

For $m = n$, the algebra of gsl$(m|n)$ is not simple, and gives rise to further so-called classical simple graded Lie algebras as subalgebras,[9] which are beyond the scope of this paper. It can be verified that the $A^A{}_B$ defined by (32) satisfy

$$[A^A{}_B, A^C{}_D] = \delta^C{}_B A^A{}_D - \begin{bmatrix} A & C \\ B & D \end{bmatrix}\delta^A{}_D A^C{}_B, \quad (33)$$

that is, the same relations as the $E^A{}_B$.

It follows that the eigenvalue of C_2, (25), and the polynomial identity, (31), have the same form as for ggl$(m|n)$. Now, however, there is no linear invariant, since

$$A^A{}_A = E^A{}_A - \frac{1}{m-n}\delta^A{}_A[A](E^C{}_C) = 0.$$

Hence the highest-weight labels are restricted by the condition

$$\sum_{a=1}^{m}\mu_a + \sum_{\alpha=1}^{n}\nu_\alpha = 0. \quad (34)$$

Therefore, for gsl$(m|n)$ there are at most $(m + n - 1)$ independent invariants of trace form.

Finally, it is possible to express the traces of the powers of A, in terms of E. Writing the traces as $\langle\hat{A}^p\rangle$, $\langle\hat{E}^p\rangle$, we have for example

$$\langle\hat{A}^2\rangle = \langle\hat{E}^2\rangle - \frac{1}{m-n}\langle\hat{E}\rangle^2$$

and so on.

gsl$(m|n)$ has associated with it an underlying Lie algebra sl$(m)\times$sl$(n)\times$gl(1). This may be seen by writing

$$\bar{A}^a{}_b = A^a{}_b - \frac{1}{m}\delta^a{}_b Z = E^a{}_b - \frac{1}{m}\delta^a{}_b E^c{}_c,$$
$$\bar{A}^\alpha{}_\beta = A^\alpha{}_\beta + \frac{1}{n}\delta^\alpha{}_\beta Z = E^\alpha{}_\beta\,\frac{1}{n}\delta^\alpha{}_\beta E^\gamma{}_\gamma, \quad (35)$$
$$Z = A^a{}_a = -A^\alpha{}_\alpha = -\frac{nE^a{}_a + mE^\alpha{}_\alpha}{m-n}.$$

Obviously the $\bar{A}^a{}_b$, $\bar{A}^\alpha{}_\beta$, Z generate sl(m), sl(n), gl(1) respectively, while the $A^a{}_\alpha$ and $A^\alpha{}_a$ transform under sl$(m)\times$sl$(n)\times$gl(1) as $(m\times\bar{n})_1$ and $(\bar{m}\times n)_{-1}$, respectively.

The invariants C_p for gsl$(m|n)$ may be expressed in terms of $\bar{A}^a{}_b$, $\bar{A}^\alpha{}_\beta$, $A^\alpha{}_a$, $A^a{}_\alpha$, and Z. For example, we have

$$C_2 = \bar{A}^a{}_b\bar{A}^b{}_a - \bar{A}^\alpha{}_\beta\bar{A}^\beta{}_\alpha + A^\alpha{}_b A^b{}_\alpha - A^a{}_\beta A^\beta{}_a - \frac{m-n}{mn}Z^2 \quad (36)$$

with similar expressions for C_3, C_4, etc.

4. GRADED ORTHOSYMPLECTIC ALGEBRA osp$(m|n)$ (n EVEN)

The generators of osp$(m|n)$ are defined in terms of the $E^A{}_B$ through the introduction of a graded-symmetrical, block-diagonal, metric tensor,

$$G_{AB} = [A\cdot B]G_{BA} = -[\bar{A}\cdot\bar{B}]G_{BA}. \quad (37)$$

Thus $G_{ab} = G_{ba}$, $G_{a\beta} = G_{\alpha b} = 0$, and $G_{\alpha\beta} = -G_{\beta\alpha}$. The $\frac{1}{2}((m+2)^2 - (m-1))$ generators

$$\Sigma_{AB} = G_{AC}E^C{}_B + [\bar{A}\cdot\bar{B}]G_{BC}E^C{}_A \quad (38)$$

which are graded-antisymmetrical, that is,

$$\Sigma_{AB} = [\bar{A}\cdot\bar{B}]\Sigma_{BA} \quad (39)$$

or $\Sigma_{ab} = -\Sigma_{ab}$, $\Sigma_{a\alpha} = \Sigma_{\alpha a}$, $\Sigma_{\alpha\beta} = \Sigma_{\beta\alpha}$, satisfy the commutation and anticommutation relations

$$[\Sigma_{AB}, \Sigma_{CD}]_{-\begin{bmatrix} A & C \\ B & D \end{bmatrix}} = G_{CB}\Sigma_{AD} + [\bar{A}\cdot\bar{B}]G_{CA}\Sigma_{BD} + \begin{bmatrix}\bar{C}\cdot\frac{\bar{A}}{\bar{D}}\end{bmatrix}G_{DB}\Sigma_{CA} + [\bar{A}\cdot\bar{B}]\begin{bmatrix}\bar{C}\cdot\frac{\bar{B}}{\bar{D}}\end{bmatrix}G_{DA}\Sigma_{CB}, \quad (40)$$

which are the defining relations of osp$(m|n)$. In particular, the Σ_{ab} generate o(m), the $\Sigma_{\alpha\beta}$ generate sp(n), and the $\Sigma_{a\alpha}$ transform as $(m\times n)$ under the underlying Lie algebra of o$(m)\times$sp(n).

In order to construct invariants it is necessary to introduce the inverse metric tensor G^{AB}, satisfying

$$G^{AC}G_{CB} = \delta^A{}_B \quad (41)$$

and which is therefore also block diagonal and graded symmetrical, as is G_{AB} [cf. (37)]. Also, n must be even, in order for the metric $G_{\alpha\beta}$ of sp(n) to have an inverse.

It is straightforward to verify from (40) and the index calculus (3) – (6), that if V_{CD} and W_{CD} are tensor operators, transforming under Σ_{AB} in the same way as Σ_{CD}, then the operator

$$X_{CD} = V_{CE}[E]G^{EF}W_{FD} \quad (42)$$

also transforms like Σ_{CD}. Moreover, the trace

$$X^C{}_C = G^{CE}X_{EC} \tag{43}$$

of any such tensor operator is an invariant operator, which commutes with the Σ_{AB}.

Thus we may again introduce matrix powers

$$\hat{\Sigma}^0{}_{CD} = G_{CD},$$
$$(\hat{\Sigma}^{p+1})_{CD} = \Sigma_{CE}[E]G^{EF}(\hat{\Sigma}^p)_{FD}, \quad p = 0,1,\cdots \tag{44}$$

all of which transform like Σ_{CD}, and with traces

$$C_p = (\hat{\Sigma}^p)^C{}_C[C] = G^{CE}(\hat{\Sigma}^p)_{EC}[C] \tag{45}$$

which are Casimir invariants.

The weights of the underlying Lie algebra, $o(m)\times sp(n)$, may be regarded as weights of the algebra $osp(m|n)$ itself, following the treatment of weights for $ggl(m|n)$ and $gsl(m|n)$. Specifically, we take the metric tensor to be

$$G_{ab} = \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & m \text{ even}, \\ \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \\ & & 1 \end{pmatrix} & m \text{ odd}, \end{cases} \tag{46}$$

$$G_{a\alpha} = G_{\alpha a} = 0,$$

$$G_{\alpha\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

and the Cartan subalgebra to be the set of generators (in terms of $\Sigma^A{}_B = G^{AC}\Sigma_{CB}$)

$$\Sigma^a{}_a, \; a = 1,\ldots,[m/2], \quad \Sigma^\alpha{}_\alpha, \; \alpha = 1,\ldots,n/2, \tag{47}$$

where $[m/2]$ is the largest whole number less than or equal to $m/2$. A vector thus has weight $(\mu|\nu) = (\mu_1,\ldots,\mu_{[m/2]}| \nu_1,\ldots,\nu_{n/2})$ if it is a simultaneous eigenvector of the $\Sigma^a{}_a, \Sigma^\alpha{}_\alpha$ with eigenvalues μ_a, ν_α. An ordering for the weights $(\mu|\nu)$ may be introduced as in (23). Connections between the subsets (μ), (ν) are provided by the odd generators $\Sigma^a{}_\alpha, \Sigma^\alpha{}_a$. It is convenient to introduce the restricted index sets

$$\underline{a},\underline{b},\ldots = 1,\ldots,[m/2], \quad \bar{a} = \underline{a} + [m/2],\cdots,$$
$$\underline{\alpha},\underline{\beta},\cdots = 1,\ldots,n/2, \quad \bar{\alpha} = \underline{\alpha} + n/2,\cdots. \tag{48}$$

and so on. Then the $\Sigma^{\underline{a}}{}_\alpha(\Sigma^{\bar{a}}{}_\alpha)$ are raising (lowering) operators for $\Sigma^{\underline{a}}{}_{\underline{a}}$, while the $\Sigma^a{}_{\bar{\alpha}}(\Sigma^a{}_{\underline{\alpha}})$ are raising (lowering) operators for $\Sigma^{\underline{\alpha}}{}_{\underline{\alpha}}$. Finally, the $\Sigma^\alpha{}_a$ have the opposite shifting properties to $\Sigma^a{}_\alpha$.

Irreducible representations of $osp(m|n)$ may be labelled by their highest-weight vector.[8,9] It is now a straightforward matter to compute the eigenvalues of the lowest C_p acting on this vector, by rearranging terms so that it is either annihilated by, or an eigenvector of, the various contributions. In fact C_1 is now zero, since

$$C_1 = G^{AB}\Sigma_{BA} = -[\bar{A}\cdot\bar{B}]^2 G^{BA}\Sigma_{AB} = -C_1,$$

while the results for

$$C_2 = \Sigma^a{}_b\Sigma^b{}_a - \Sigma^\alpha{}_\beta\Sigma^\beta{}_\alpha + \Sigma^\beta{}_a\Sigma^a{}_\beta - \Sigma^a{}_\beta\Sigma^\beta{}_a \tag{49}$$

are as follows:

$$C_2 = 2\Big\{\sum_{a=1}^{[m/2]}\mu_a(\mu_a + m - n - 2a) - \sum_{\alpha=1}^{n/2}\nu_\alpha(\nu_\alpha + n + 2 - 2\alpha)\Big\}, \quad m \text{ even},$$

$$C_2 = 2\Big\{\sum_{a=1}^{[m/2]}\mu_a(\mu_a + m - n - 2a) - \sum_{\alpha=1}^{n/2}\nu_\alpha(\nu_\alpha + n + 1 - 2\alpha)\Big\}, \quad m \text{ odd}. \tag{50}$$

The defining relations (40) also identify the transformation property of a vector operator V^C under the Σ_{AB} to be

$$[\Sigma_{AB}, V^C]_{\left[\begin{smallmatrix}A\\B\end{smallmatrix}\cdot C\right]} = \delta^C{}_B V_A + [\bar{A}\cdot\bar{B}]\delta^C{}_A V_B, \tag{51}$$

where $V_C = G_{CE}V^E$.

In particular, V^C has components V, V^γ transforming under $o(m)\times sp(n)$ as $m\times 1$, $1\times n$ respectively. Bracken and Green[24] have shown that a vector operator for $o(m)$ or $sp(n)$ may always be resolved into a sum of m or n terms, each of which changes one of the highest-weight labels by ± 1, and commutes with the remaining labels [and for $o(m)$, if m is odd, there is an additional component which commutes with all the labels]. Thus we may write, for m odd,

$$V^C = \sum_{a=1}^{[m/2]}(V^C_{(a^+)} + V^C_{(a^-)}) + \sum_{\alpha=1}^{n/2}(V^C_{(\alpha^+)} + V^C_{(\alpha^-)}) + V^C_{(0)} \tag{52}$$

with the same expression, omitting $V^C_{(0)}$, for m even, and

$$[\mu_a, V^C_{(a\pm)}]_- = \pm\delta_{ab}V^C_{(b\pm)},$$
$$[\nu_\alpha, V^C_{(\alpha\pm)}]_- = \pm\delta_{\alpha\beta}V^C_{(\beta\pm)}. \tag{53}$$

It therefore follows from (50) that, for example,

$$[C_2, V^C_{(a\pm)}]_- = 2(\pm 2\mu_a - 1 \pm (m - n - 2a))V^C_{(a\pm)}, \tag{54}$$

on the other hand, by direct computation using (45) and (50),

$$[C_2, V^C]_- = 2(2\Sigma^C{}_E[E] - (m - n - 1)\delta^C{}_E)V^E \tag{55}$$

and this is true also of the components $V^C_{(a\pm)}, V^C_{(0)}, V^C_{(\alpha\pm)}$. Thus using $\hat{\Sigma}^C{}_E = \Sigma^C{}_E[E]$,

$$((\hat{\Sigma} - \tfrac{1}{2}(m-n) + 1) \mp (\mu_a + \tfrac{1}{2}(m-n) - a))^C{}_E V^E_{(a\pm)} = 0,$$

with similar equations for $V^C_{(\alpha\pm)}, V^C_{(0)}$. Since these equations are true for arbitrary vector operators V^C, we have as a matrix equation[24,25]

$$\prod_{a=1}^{[m/2]}((\hat{\Sigma} - \tfrac{1}{2}(m-n) + 1)^2 - (\mu_a + \tfrac{1}{2}(m-n) - a)^2) \times \prod_{\alpha=1}^{n/2}((\hat{\Sigma} - \tfrac{1}{2}(m-n))^2 - (\nu_\alpha + \tfrac{1}{2}n + 1 - \alpha)^2) = 0, \quad (m \text{ even}), \tag{56}$$

$$(\hat{\Sigma} - \tfrac{1}{2}(m-n-1))\prod_{a=1}^{[m/2]}((\hat{\Sigma} - \tfrac{1}{2}(m-n) + 1)^2 - (\mu_a + \tfrac{1}{2}(m-n) - a)^2)\prod_{\alpha=1}^{n/2}((\hat{\Sigma} - \tfrac{1}{2}(m-n))^2 - (\nu_\alpha + \tfrac{1}{2}n + \tfrac{1}{2} - \alpha)^2) = 0 \quad (m \text{ odd}).$$

Equations (56) are the desired characteristic polynomial identity, for osp($m|n$), of order ($m+n$).

The question can now be settled of the number of independent invariants of osp($m|n$) which exist, of the trace form. For the underlying Lie algebra o(m)×sp(n), at most $[m/2]+n$ independent invariants can be consructed, and hence there should be at least this number for osp($m|n$). Just as in the o(m) and sp(n) cases, the linear invariant $G^{AB}\Sigma_{BA}$ vanishes for osp($m|n$) because the metric and the matrix of generators have opposite (graded) symmetry properties. The generalization of this result for osp($m|n$), again as is true of the o(m) and sp(n) cases, is the observation that the pth matrix power $(\hat{\Sigma}^p)_{AB}$ is itself symmetric or antisymmetric, depending upon p, up to lower-order terms. It is straightforward to verify that

$$\begin{aligned}
&\Sigma_{AB}=[\bar{A}\cdot\bar{B}]\Sigma_{BA},\\
&(\hat{\Sigma}^2)_{AB}[B]\\
&\quad=-[\bar{A}\cdot\bar{B}](\hat{\Sigma}^2-[\bar{A}\cdot\bar{B}](m-n-2)\hat{\Sigma})_{BA}[A],
\end{aligned} \tag{57}$$

and, in general,

$$(\hat{\Sigma}^{p+1})_{AB}[B]=[\bar{A}\cdot\bar{B}]((-1)^p\hat{\Sigma}^{p+1}+P^p(\hat{\Sigma}))_{BA}[A], \tag{58}$$

where $P^p(\Sigma)$ is a polynomial function of Σ, of order p, defined recursively by (57) and

$$\begin{aligned}
P^p(\hat{\Sigma})_{BA}=&-(\hat{\Sigma}\cdot P^{p-1})_{BA}\\
&-(-1)^p(\hat{\Sigma}^p)_{BA}(m-n-1+(-1)^p)\\
&+\begin{bmatrix}A\\B\end{bmatrix}\cdot A\,\Big]G_{BA}\langle\hat{\Sigma}^p\rangle\\
&+(m-n-1)P^{p-1}(\hat{\Sigma})_{BA}.
\end{aligned} \tag{59}$$

Thus, for example, $(\hat{\Sigma}^3)_{\rm BA}$ is graded-symmetrical, up to lower-order terms, and evaluation of (58) and (59) lead to

$$C_3=\tfrac{1}{2}(m-n-2)C_2. \tag{60}$$

Therefore the odd Casimir invariants $C_1, C_3, C_5,\cdots$ are not independent, for osp($m|n$), but are related directly to the even invariants $C_2, C_4, C_6,\ldots$, just as in the o(m) and sp(n) cases. Moreover, the existence of the polynomial identity (56) of order ($m+n$) shows that any higher-order (even) invariants such as C_{m+n+1}, C_{m+n+2}, may be related to lower ones. Therefore, there are at most $[m/2]+n/2$ independent Casimir invariants for osp($m|n$), namely $C_2, C_4, C_6,\ldots, C_{2[m/2]+n}$.

5. EXAMPLES AND DISCUSSION

As an illustration of the techniques we have developed in the foregoing, it is useful to examine some specific cases of low-dimensional graded Lie algebras of the classical series.

For example the algebra of gsl(2|1) has four even generators [cf. (35)] $\bar{A}^a{}_b$, $a,b=1,2$; Z, and four odd generators $A^a{}_1$, A_{1a}. Adopting the notation

$$\begin{aligned}
&A_i=\tfrac{1}{2}(\tau_i)_a{}^b\bar{A}^a{}_b,\quad i,j=1,2,3,\\
&(\underline{A}\cdot\underline{\tau})_c{}^d=A^d{}_c,\\
&Q_a=A^1{}_a,\\
&\bar{Q}^a=A^a{}_1,
\end{aligned} \tag{61}$$

where the τ_i are the 2×2 Pauli matrices, the defining relations (33) become

$$\begin{aligned}
&[A_i,A_j]=i\epsilon_{ijk}A_k,\\
&[A_i,Q_a]=-\tfrac{1}{2}(\tau_i)_a{}^bQ_b,\\
&[A_i,\bar{Q}^a]=\tfrac{1}{2}\bar{Q}^b(\tau_i)_b{}^a,\\
&[A_i,Z]=0,\\
&[Z,Q_a]=-Q_a,\\
&[Z,\bar{Q}^a]=+\bar{Q}^a,\\
&\{Q_a,\bar{Q}^b\}=(\underline{A}\cdot\tau)_a{}^b-\tfrac{1}{2}\delta^a{}_bZ,
\end{aligned} \tag{62}$$

and the quadratic Casimir operator (36) gives

$$\tfrac{1}{2}C_2=\underline{A}^2-\tfrac{1}{4}Z^2+\tfrac{1}{2}(Q\bar{Q}-\bar{Q}Q). \tag{63}$$

Now irreducible representations of gsl(2|1) may be labelled by their highest weights[8] $\{\mu_1,\mu_2|\nu_1\}$ with $\mu_1+\mu_2+\nu_1=0$. The eigenvalue of C_2, given by (25), is most easily written in terms of

$$l=\tfrac{1}{2}(\mu_1-\mu_2),\quad \nu=-\tfrac{1}{2}(\mu_1+\mu_2)=+\tfrac{1}{2}\nu_1 \tag{64}$$

and leads to

$$\tfrac{1}{2}C_2=(l+\tfrac{1}{2})^2-(\nu+\tfrac{1}{2})^2. \tag{65}$$

For gsl(2|1) there is of course an additional cubic Casimir operator C_3, given by (21).

Formulas (62)–(65) may be checked against results given in studies of gsl(2|1) in the literature.[9] There is an obvious extension to gsl($n|1$) involving the Cartesian basis A_i, $i=1,\ldots,n^2-1$ for sl(n), and $n\times n$ matrices λ_i, $i=1,\ldots,n^2-1$, similar to that considered by Dondi and Sohnius.[32]

The graded orthosymplectic algebra osp(1|2) has three even generators $\Sigma_{\alpha\beta}=\Sigma_{\beta\alpha}$, $\alpha=1,2$ and two odd generators $\Sigma_{1\alpha}=\Sigma_{\alpha 1}$. The metric (46) is

$$G=\begin{pmatrix}1&&\\&0&1\\&-1&0\end{pmatrix}. \tag{66}$$

Again adopting the Cartesian basis

$$\begin{aligned}
&J_i=-\tfrac{1}{2}\Sigma_{\alpha\beta}(\tau_i\epsilon)_{\beta\alpha},\quad i=1,2,3,\\
&\Sigma_{\alpha\beta}=\underline{J}\cdot(\epsilon\underline{\tau})_{\alpha\beta},\\
&\Sigma_{1\alpha}=\Sigma_{\alpha 1}=S_\alpha,
\end{aligned} \tag{67}$$

where the τ_i are again 2×2 Pauli matrices, and $\epsilon_{\alpha\beta}$ is the symplectic part of the metric G, the defining relations (40) become

$$\begin{aligned}
&[J_i,J_j]=i\epsilon_{ijk}J_k,\\
&[J_i,S_\alpha]=-\tfrac{1}{2}(\tau_i)_{\alpha\beta}S_\beta,\\
&\{S_\alpha,S_\beta\}=J_i(\epsilon\tau_i)_{\alpha\beta}
\end{aligned} \tag{68}$$

and the quadratic Casimir (49) gives

$$-\tfrac{1}{2}C_2=\underline{J}^2-S_\alpha\epsilon_{\alpha\beta}S_\beta. \tag{69}$$

Irreducible representations of osp(1|2) may be labelled by their highest weights[8] $[0|\nu]$. The eigenvalue of $-\tfrac{1}{2}C_2$, given by (50), is most easily written in terms of $l=\tfrac{1}{2}\nu$, and is

$$-\tfrac{1}{2}C_2 = 4l(l+\tfrac{1}{2}). \qquad (70)$$

Again, formulas (68)–(70) may be checked against results given in the literature.[9]

The algebra osp(1|4) is the algebra of the graded de Sitter group, and is related by contraction to the algebra of Poincaré supersymmetry. The algebra is [cf. (40)]

$$\begin{aligned}
&[M_{\alpha\beta},M_{\gamma\delta}] = M_{\alpha\delta} + C_{\gamma\alpha} M_{\beta\delta} + C_{\delta\beta}M_{\alpha\gamma} + C_{\delta\alpha}M_{\beta\gamma},\\
&[M_{\alpha\beta},U_\gamma] = C_{\gamma\beta}U_\alpha + C_{\gamma\alpha}U_\beta,\\
&\{U_\alpha,U_\beta\} = M_{\alpha\beta},\\
&M_{\alpha\beta} = M_{\beta\alpha}, \quad C_{\alpha\beta} = -C_{\beta\alpha}, \quad \alpha,\beta = 1,\ldots,4.
\end{aligned} \qquad (71)$$

The isomorphism with the graded so(3,2) algebra is established by identifying $C_{\alpha\beta}$ with the charge conjugation matrix of the Dirac spinor representation, and expanding $M_{\alpha\beta}$ in terms of the symmetric matrices, $(\gamma_\mu C)_{\alpha\beta}$ and $(\sigma_{\mu\nu}C)_{\alpha\beta}$,

$$M_{\alpha\beta} = -(\gamma_\gamma C)_{\alpha\beta}M^\mu - \tfrac{1}{2}(\sigma_{\mu\nu}C)_{\alpha\beta}M^{\mu\nu}, \quad \mu,\nu = 0,1,2,3. \qquad (72)$$

It is found that the $M^{\mu\nu}$ generate an algebra of O(3,1), with the M^μ an additional vector such that

$$[M^\mu,M^\nu] = -iM^{\mu\nu}, \qquad (73)$$

thus enlarging the algebra to o(3,2) with the identification $M^\mu = M^{\mu 4}$, and 5×5 metric with signature $(+,-,-,-,+)$. The remaining commutation relations are

$$\begin{aligned}
&[M^{\mu\nu},U_\alpha] = -\tfrac{1}{2}(\sigma^{\mu\nu})_\alpha{}^\beta U_\beta,\\
&[M^\mu,U_\alpha] = -\tfrac{1}{2}(\gamma^\mu)_\alpha{}^\beta U_\beta.
\end{aligned} \qquad (74)$$

If one defines

$$\begin{aligned}
&\bar{S}_\alpha = \frac{1}{R}U_\alpha,\\
&\bar{P}^\mu = \frac{1}{R^2}M^{\mu 4},\\
&\bar{M}^{\mu\nu} = M^{\mu\nu},
\end{aligned} \qquad (75)$$

and rewrites the algebra (71) in terms of the barred generators, the limit $R\to\infty$ may be taken consistently with the barred generators $\bar{S}_\alpha$, $\bar{P}^\mu$, $\bar{J}^{\mu\nu}$ tending to smooth limits S_α, P^μ, $J^{\mu\nu}$, leaving the graded algebra

$$\begin{aligned}
&[J_{\mu\nu},J_{\rho\sigma}] = i(\eta_{\rho\nu}J_{\mu\sigma} - \eta_{\rho\gamma}J_{\nu\sigma} - \eta_{\sigma\nu}J_{\mu\rho} + \eta_{\sigma\mu}J_{\nu\rho}),\\
&[J_{\mu\nu},P_\rho] = i(\eta_{\nu\rho}P_\mu - \eta_{\mu\rho}P_\nu),\\
&[J_{\mu\nu},S_\alpha] = \tfrac{1}{2}(\sigma_{\mu\nu})_\alpha{}^\beta S_\beta,\\
&[P_\mu,P_\nu] = 0,\\
&[P_\mu,S_\alpha] = 0,\\
&\{S_\alpha,S_\beta\} = -(\gamma_\mu C)_{\alpha\beta}P^\mu,
\end{aligned} \qquad (76)$$

which is, indeed, the space–time Poincaré supersymmetry algebra.

The contraction procedure, (75), may now be applied to the (appropriately rescaled) quadratic and quartic Casimir operators, C_2 and C_4, defined by (46). Explicitly, it is found that

$$C_2 = -4M^\mu M_\mu - 2M^{\mu\nu}M_{\mu\nu} + 2U^\alpha U_\alpha, \qquad (77)$$

so that

$$C_2' = \lim_{R\to\infty} -\frac{1}{4R^4}C_2 = P^\mu P_\mu, \qquad (78)$$

while a similar calculation yields

$$C_2' = (P^\mu P_\mu)(K^\mu K_\mu),$$

where

$$K_r = \epsilon_{\mu\nu\rho\sigma}P^\nu J^{\rho\sigma} + \tfrac{1}{4}\bar{S}_i(\gamma_\mu - P_\mu p^{-1})\gamma_5 S$$

is the usual generalized Pauli–Lubanski vector, whose square provides the additional Casimir invariant for Poincaré supersymmetry.[4]

These results illustrate that the quadratic and quartic Casimir in variants of osp(1|4), as defined by (45), do indeed contract in the required manner to the familiar invariants of Poincaré supersymmetry. Other examples could be studied; for example, according to our analysis, the space-time Conformal supersymmetry algebra,[31] which is a real form of gsl(4|1), should have four Casimir invariants, given by (21). The quadratic and cubic invariants have been given[33]; however with the techniques of the present paper on hand, the labour of explicitly constructing higher invariants is removed, and moreover it is clear that just two further independent Casimir, C_4 and C_5, can be constructed.

The work presented here emphasizes that the sequences ggl($m|n$), gsl($m|n$) ($m\neq n$), osp($m|n$) (n even) of graded Lie algebras are amenable to many of the techniques which have been developed for ordinary Lie algebras. Presumably many of these techniques also carry over to the remaining sequences of simple graded Lie algebras[8,9]; the exceptional algebras are likely to require individual attention. In any case, it is evident that, with the use of the index calculus (3) to (6), and the results obtained here, a complete tensor calculus can be developed for ggl($m|n$), gsl($m|n$), ($m\neq n$), and osp($m|n$) (n even), with eventual applications, for example, to methods of computing the eigenvalues of the higher-order Casimir invariants.[30] Further work along these lines is in progress.

Despite the similarities, however, between the graded Lie algebras and ordinary Lie algebras, many of the theorems applicable to the latter do not carry over. For example, as is evident from (25) and (50), the Killing form is, in general, no longer positive definite, even for the simple graded Lie algebras. Consequently the Casimir invariants are, in general, insufficient to specify the irreducible representations. In fact, Schur's lemma itself is no longer always valid; other complications arise in the representation theory.[9] Nevertheless, the simplicity of the results here presented should be sufficient reason to regard the subject as interesting, both in its own right and for physical applications, and to merit further investigation.

ACKNOWLEDGMENT

P.D.J. wishes to thank P.H. Dondi for discussions and constructive criticism of this work.

Note added in proof: The graded Lie algebras are also known in the literature[9] as "Lie superalgebras."

[1]J. Scherk, Rev. Mod. Phys. **47**, 123 (1975).
[2]J. Wess and B. Zumino, Nuc. Phys. B **70**, 39 (1974).

[3]A. Salam and J. Strathdee, Nucl. Phys. B **76**, 477 (1974).
[4]P. Fayet and S. Ferrara, Phys. Rep. C **32**, 69 (1977).
[5]J. Wess, "Supersymmetry-Supergravity," Proceedings of 8th GIFT Seminar on Theoretical Physics, to appear in Springer Tracts in Modern Physics.
[6]B. Kostant, "Graded Manifolds, graded Lie theory and prequantization," in *Lecture Notes in Mathematics*, **570**, edited by A. Dold and B. Eckmann, **570** (Springer-Verlag Berlin, 1977).
[7]L. Corwin, Y. Ne'eman, and S. Sternberg, Rev. Mod. Phys. **47**, 573 (1975).
[8]V.G. Kac, Commun. Math. Phys. **53**, 31 (1977).
[9]V. Rittenberg, "A Guide to Lie Superalgebras," in *Lecture Notes in Physics* **79**, edited by P. Kramer and A. Rieckers (Springer-Verlag Berlin, 1978).
[10]N. B. Backhouse, J. Math. Phys. **18**, 339 (1977).
[11]P.D. Jarvis, J. Math. Phys. **18**, 551 (1977).
[12]V. Rittenberg and D. Wyler, "Sequences of Generalized Lie Algebras and Superalgebras," Rockefeller University Report C00-2232B-152 (1978), and references therein.
[13]P.D. Jarvis, Australian J. Phys. **31**, 461 (1978).
[14]N. Dragon, "Generalized Symmetry Algebras," talk at Liblice Conference, 1978.
[15]G. Racah, "Group Theory and Spectroscopy," Princeton Institute for Advanced Study Lectures (1951); reprinted in Springer Tracts in Modern Physics **37**, 28 (1965).
[16]H.B.G. Casimir, Proc. Roy. Acad. Amsterdam **34**, 844 (1931).
[17]G. Racah, Rend. Lincei **8**, 108 (1950).
[18]L.C. Biedenharn and J.D. Louck, Commun. Math. Phys. **8**, 89 (1968).
[19]Y. Lehrer-Ilamed, Bull. Res. Counc. Israel **5A**, 197 (1956).
[20]M. Gell-Mann, Phys. Rev. **125**, 1067 (1962).
[21]S. Okubo, Prog. Theor. Phys. (Kyoto) **27**, 949 (1962).
[22]J. Ginibre, J. Math. Phys. **4**, 720 (1963).
[23]Y. Lehrer-Ilamed and H. Goldberg, J. Math. Phys. **4**, 501 (1963).
[24]A.J. Bracken and H.S. Green, J. Math. Phys. **12**, 2099 (1971).
[25]H.S. Green, J. Math. Phys. **12**, 2107 (1971).
[26]J.D. Louck and L.C. Biedenharn, J. Math. Phys. **11**, 2368 (1970).
[27]S. Okubo, J. Math. Phys. **16**, 528 (1975).
[28]M.A. Rashid and C.D. Nwachuku, J. Math. Phys. **17**, 1611 (1976).
[29]S. Okubo, J. Math. Phys. **18**, 2382 (1977).
[30]R.C. King, "Eigenvalues of Generalized Casimir Invariants of Semi-Simple Lie Groups," University of Southampton Report, Mathematical Studies No. 18 (Decamber 1978), and references therein.
[31]N. Backhouse, J. Phys. **12**, 21 (1979).
[32]P.H. Dondi and M. Sohnius, Nucl. Phys. B **81**, 317 (1974).
[33]M. Sohnius, Ph.D. Thesis, University of Karlsruhe (1976).

A note on Sommerfeld's diffraction problem

V. V. S. S. Sastry and A. Chakrabarti

Department of Applied Mathematics, Indian Institute of Science, Bangalore-560012, India
(Received 5 June 1978; revised manuscript received 29 September 1978)

A direct transform technique is found to be most suitable for attacking two-dimensional diffraction problems. As a first example of the application of the technique, the well-known Sommerfeld problem is reconsidered and the solution of the problem of diffraction, by a half-plane, of a cylindrical pulse is made use of in deducing the solution of the problem of diffraction of a plane wave by a soft half-plane.

INTRODUCTION

Various methods of attacking the problem of diffraction by a half-plane (the well known Sommerfeld problem) have been proposed by several authors.[1,2] (The details of the references in this direction are cited in the books of Jones[1] and Noble.[2]) Though the Wiener–Hopf technique is believed to be the most suitable one for two-dimensional diffraction problems, the method has got its own limitations, particularly when one wishes to attack problems under mixed boundary conditions.[3,4] The purpose of the present note is to highlight the technique of Turner,[5] involving the direct application of two kinds of integral transforms, Laplace and Lebedev–Kontorovich, in succession, to the initial and boundary value problem of the diffraction of a cylindrical pulse by a soft half-plane. It is shown here that the solution obtained by Turner[5] can be utilized successfully, to derive the solution of the problem of diffraction of a plane time-harmonic wave by a similar half-plane. It turns out that the total field under consideration can be represented in the form of an infinite series, involving Bessel functions of the first kind. The series could be summed in a closed form by making use of an integral representation of the Bessel function. The final form of the solution then involves the error function complement as is well known.[1]

1. TURNER'S PROBLEM AND ITS SOLUTION

In this section, we briefly present the method of Turner, for the sake of making this note self-contained.

We wish to determine the scalar potential $\phi(r,\theta,t)$ satisfying

$$\nabla^2\phi = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} - \frac{1}{r_0}\delta(r-r_0)\delta(\theta-\theta_0)\delta(t-0^+), \quad (1)$$

subject to the boundary condition

$$\phi = 0 \quad (2)$$

on $\theta = 0$ and 2π; and the initial conditions

$$\phi = 0 \quad \text{and} \quad \frac{\partial\phi}{\partial t} = 0 \quad (3)$$

at $t = 0^+$. We set

$$R = [r^2 + r_0^2 - 2rr_0\cos(\theta-\theta_0)]^{1/2} \quad (4)$$

and

$$R_0 = [x^2 + r_0^2 - 2xr_0\cos\theta_0]^{1/2}. \quad (5)$$

We apply a Laplace transform in t to Eq. (1) and write (denoting the Laplace transform of ϕ by Φ) $\Phi = \Phi^{(1)} + U$, where $\Phi^{(1)} = (1/2\pi)K_0(pR/c)$. The resulting partial differential equation after a suitable change of dependent variable is solved by the application of the Lebedev-Kontorovich transform. After performing the Lebedev-Kontorovich and Laplace inversion transforms in succession, we obtain (see Turner[5]):

$$\phi = \phi^{(i)} + \phi^{(r)} + \phi^{(s)}, \quad (6)$$

$$\phi^{(i)} = -\frac{1}{4\pi}\frac{H(t-R/c)}{(t^2-R^2/c^2)^{1/2}}, \quad (6a)$$

$$\phi^{(r)} = -\frac{1}{4\pi}\frac{H(t-R_1/c)}{(t^2-R_1^2/c^2)^{1/2}}, \quad (6b)$$

$$\phi^{(s)} = \frac{c}{2\pi\sqrt{rr_0}}\sum_{n=0}^{\infty}P_n\left(\frac{r^2+r_0^2-c^2t^2}{2rr_0}\right)\sin(n+\tfrac{1}{2})\theta \times\sin(n+\tfrac{1}{2})\theta_0; \quad |r-r_0|<ct<r+r_0,$$
$$= 0; \quad \text{otherwise}, \quad (6c)$$

ere

$$R_1 = [r^2 + r_0^2 - 2rr_0\cos(\theta+\theta_0)]^{1/2},$$

and $P_n(x)$ is the Legendre polynomial of order n and $H(x)$ is the Heaviside step function.

2. THE TIME HARMONIC PROBLEM

In this section we present a method of deducing the solution for the time harmonic problem from that of the solution of pulse diffraction [Eq. (6)]. Replacing t by $t'-t$, t' being the new time variable, we obtain from Eq. (1)

$$\nabla^2\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t'^2} = -\frac{1}{r_0}\delta(r-r_0)\delta(\theta-\theta_0)\delta(t-t'),$$

where ψ denotes $\phi(r,\theta,t'-t)$. Multiplying this equation by $e^{-i\omega t'}$ and integrating with respect to t' from t to infinity, after using the conditions

$$\psi(r,\theta,t'-t) = 0$$

and

$$\frac{\partial\psi}{\partial t'}(t,\theta,t'-t) = 0 \quad \text{at} \quad t'=t,$$

we readily observe that the function $\tilde{\psi}(r,\theta,t)$ satisfies the partial differential equation

$$(\nabla^2 + k^2)\tilde{\psi} = -\frac{1}{r_0}\delta(r-r_0)\delta(\theta-\theta_0)e^{-i\omega t} \quad (7)$$

 0022-2488/79/102123-03$01.00

and the boundary conditions

$$\tilde{\psi} = 0 \quad \text{on } \theta = 0 \text{ and } 2\pi, \tag{8}$$

where

$$\tilde{\psi}(r,\theta,t) = \int_t^\infty \psi e^{-i\omega t'} dt'$$

and $k = \omega/c$, the wavenumber, ω being the frequency. Equation (7) is precisely the equation governing the time harmonic case.

In the above deductions we have made use of the following property of Dirac's delta function

$$\int_{-\infty}^{\infty} f(t-t')\delta(t')dt' = f(t).$$

Applying the above procedure to Eq. (6), we obtain

$$e^{i\omega t}\tilde{\psi} = \sum_{n=0}^{\infty} K_{n+1/2}(ikr_>)I_{n+1/2}(ikr_<)\sin(n+\tfrac{1}{2})\theta \times \sin(n+\tfrac{1}{2})\theta_0 - \frac{1}{4\pi}\{K_0(ikR) + K_0(ikR_1)\}. \tag{9}$$

This can also be formally checked by noting that the above procedure leads to the Laplace transform field [Eq. (2), p. 68, Turner[5]], the transform parameter being $i\omega$.

We thus observe that the solution of the time harmonic line-source problem is given by Eq. (9).

From now on, we set $r_> = r_0$ and $r_< = r$. Finally, to obtain the scattered field in the case of plane wave incidence, we let r_0 tend to infinity after multiplying Eq. (9) by $(8\pi k r_0)^{1/2} i\exp\{i(kr_0 - \pi/4)\}$ and replacing θ_0 by $\varphi_0 - \pi$, where $0 < \varphi_0 < \pi$. Thus, we obtain, the total field $\tilde{\psi}$, due to the incident plane wave

$$\exp\{-i[kr\cos(\theta-\varphi_0) + \omega t]\}$$

as given by

$$e^{i\omega t}\tilde{\psi} = \sum_{n=0}^{\infty} \exp\left[i(n+\tfrac{1}{2})\frac{\pi}{2}\right] J_{n+1/2}(kr)[\cos(n+\tfrac{1}{2})(\theta-\varphi_0+\pi) - \cos(n+\tfrac{1}{2})(\theta+\varphi_0-\pi)] - \tfrac{1}{2}\{\exp[-ikr \times \cos(\theta-\varphi_0)] + \exp[ikr\cos(\theta+\varphi_0)]\}. \tag{10}$$

3. SUMMATION OF THE SERIES

In this section, we present the details of the summation of the infinite series obtained in the last section. We are required to sum the series

$$S' = 2\sum_{n=0}^{\infty} \exp\left[i(n+\tfrac{1}{2})\frac{\pi}{2}\right] J_{n+1/2}(kr)\cos(n+\tfrac{1}{2})\alpha. \tag{11}$$

It is sufficient to consider

$$S(\varphi) = \sum_{n=0}^{\infty} \exp[i(n+\tfrac{1}{2})\varphi]\, J_{n+1/2}(x). \tag{12}$$

Then, (11) can be expressed as

$$S' = S\left(\frac{\pi}{2}+\alpha\right) + S\left(\frac{\pi}{2}-\alpha\right). \tag{13}$$

We have[6]

$$J_\nu(x) = \frac{2(x/2)^\nu}{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})}\int_0^{\pi/2}\cos(x\cos\alpha)\sin^{2\nu}\alpha\, d\alpha$$

for $\mathrm{Re}\,\nu > -\frac{1}{2}$. Hence on substituting in (12) and interchanging the orders of summation and integration, we obtain

$$S(\varphi) = \frac{2}{\sqrt{\pi}}\int_0^{\pi/2}\cos(x\cos\alpha)\sin\alpha\, d\alpha \times \sum_{n=0}^{\infty}\exp[i(n+\tfrac{1}{2})\varphi]\,\frac{(x/2)^{n+1/2}\sin^{2n}\alpha}{n!} \tag{14}$$

$$= \frac{2}{\sqrt{\pi}} e^{i\varphi/2}\int_0^{\pi/2}\cos(x\cos\alpha)\sin\alpha \times \exp[\tfrac{1}{2}xe^{i\varphi}\sin^2\alpha]d\alpha.$$

On putting $u = \cos\alpha$,

$$S(\varphi) = \sqrt{\frac{2x}{\pi}}\exp[i\varphi/2 + \tfrac{1}{2}xe^{i\varphi}] \times \int_0^1 \cos xu \exp[-\tfrac{1}{2}xe^{i\varphi}u^2]du. \tag{15}$$

The latter integral can easily be expressed in terms of the error function complement. We write

$$\int_0^1 e^{-\beta u^2}\cos xu\, du = \int_0^\infty e^{-\beta u^2}\cos xu\, du - \int_1^\infty e^{-\beta u^2}\cos xu\, du$$

where $\beta = \frac{1}{2}xe^{i\varphi}$.

The first integral can be easily evaluated (see Gradshteyn and Ryzhik,[6] p. 480). Hence

$$\int_0^1 e^{-\beta u^2}\cos xu\, du$$

$$= \tfrac{1}{2}\sqrt{\pi/\beta}\exp(-x^2/4\beta) - \int_1^\infty e^{-\beta u^2}\cos xu\, du$$

$$= \tfrac{1}{2}\sqrt{\pi/\beta}\exp(-x^2/4\beta)\left\{1 - \tfrac{1}{2}\left[\mathrm{erfc}\left(\sqrt{\beta} - \frac{ix}{2\sqrt{\beta}}\right) + \mathrm{erfc}\left(\sqrt{\beta} + \frac{ix}{2\sqrt{\beta}}\right)\right]\right\}, \tag{16}$$

where

$$\mathrm{erfc}(x) = 2\pi^{-1/2}\int_x^\infty \exp(-t^2)dt.$$

Thus, from (15), we obtain

$$S(\varphi) = \exp(ix\sin\varphi)\left(1 - \tfrac{1}{2}\{\mathrm{erfc}[\sqrt{x/2}(e^{i\varphi/2} - ie^{-i\varphi/2})] + \mathrm{erfc}[\sqrt{x/2}(e^{i\varphi/2} + ie^{-i\varphi/2})]\}\right).$$

Making use of Eq. (13), we finally obtain,

$$S' = \exp(-ikr\cos\zeta)[1 - \mathrm{erfc}(e^{i\pi/4}\sqrt{2kr}\,\sin\zeta/2)], \tag{17}$$

where $\zeta = \pi - \alpha$.

The above procedure of summing up the series was suggested by the referee and the authors gratefully acknowledge the suggestion. However, for more general values of the order of the Bessel function appearing in Eq. (11) (which we would come across while considering the problem of mixed boundary conditions, i.e., ϕ prescribed on $\theta = 0$ and $\partial\phi/\partial\theta$ prescribed on $\theta = 2\pi$), an approach similar to that of Khrebet[7] is more useful.

Using the summation formula (17), we immediately obtain from Eq. (10),

$$e^{i\omega t}\tilde{\psi} = \tfrac{1}{2}\mathrm{erfc}\left[e^{i\pi/4}\sqrt{2kr}\,\sin(\theta+\varphi_0)/2\right]\exp[-ikr\cos(\theta + \varphi_0)] - \tfrac{1}{2}\mathrm{erfc}\left[e^{i\pi/4}\sqrt{2kr}\,\sin(\theta-\varphi_0)/2\right] \times\exp[-ikr\cos(\theta-\varphi_0)],$$

which agrees with the well known solution given in Jones.[1]

[1]D.S. Jones, *The Theory of Electromagnetism* (Pergamon, New York, 1964).
[2]B. Noble, *Methods based on the Wiener–Hopf Technique for the Solution of Partial Differential Equations* (Pergamon, New York, 1958).
[3]A.D. Rawlins, "The solutions of mixed boundary value problems in the theory of diffraction by a semi-infinite plane," Proc. R. Soc. London, Ser. A **346**, 469 (1975).
[4]R.A. Hurd, "The Wiener–Hopf–Hilbert method for diffraction problems," Can. J. Phys. **54**, 775 (1976).
[5]R.D. Turner, The diffraction of a cylindrical pulse by a half-plane, Q. A. Math. **14**, 63 (1956).
[6]I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965).
[7]N.G. Khrebet, "Diffraction of plane electromagnetic waves on the edge of a dielectric half plane," Radio Eng. Electron. Phys. **13**, 331 (1968).

Analysis of the dispersion function for anisotropic longitudinal plasma waves[a)]

M. D. Arthur, R. L. Bowden, and P. F. Zweifel

Laboratory for Transport Theory and Mathematical Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

(Received 24 April 1979; accepted for publication 1 June 1979)

An analysis of the zeros of the dispersion function for longitudinal plasma waves is made. In particular, the plasma equilibrium distribution function is assumed to have two relative maxima and is not necessarily an even function. The results of this analysis are used to obtain the Wiener–Hopf factorization of the dispersion function. A brief analysis of the coupled nonlinear integral equations for the Wiener–Hopf factors is also presented.

I. INTRODUCTION

The solution of boundary value problems (fixed frequency waves) as described by the linearized Vlasov equation requires the Wiener–Hopf factorization of the plasma dispersion function Λ.[1,2] For the longitudinal modes which we consider in the present paper, the relevant Vlasov–Maxwell set of equations (in the absence of magnetic fields) can be taken to be

$$\frac{\partial g}{\partial t} + u\frac{\partial g}{\partial z} + \frac{e}{m}(E_A + E)F'(u) = 0, \qquad (1a)$$

$$\frac{\partial E}{\partial t} = -4\pi n_0 e\int_{-\infty}^{\infty} sg(z,s,t)\,ds. \qquad (1b)$$

Here, u represents the longitudinal electron velocity, F is the equilibrium distribution function, g is the deviation of the distribution function from equilibrium, E_A is the applied electric field (z component), E is the z component of the self-consistent electric field, and n_0 is the plasma density. We might note that in the previous analyses, Refs. 1 and 2, it has been customary to use Gauss' Law, instead of Ampere's Law [Eq. (1b)]. After Fourier-transforming in the time variable, this leads to a set of two coupled equations for f_ω and E_ω ($f_\omega(z,u) = \int_{-\infty}^{\infty} e^{i\omega t}g(z,u,t)\,dt$). Since it is much easier to work with a one-component equation, we choose to begin our analysis with Ampere's Law.

Thus, after the aforementioned Fourier transformation, the equation we study is

$$\frac{\partial f}{\partial z} + i\omega Kf = -\frac{e}{i\omega m}E_A\frac{F'(u)}{u}, \qquad (2a)$$

with the unbounded linearized operator K defined by

$$(Kf)(z,u) = \frac{1}{u}f(z,u) + \sigma^2\frac{F(u)}{u}\int_{-\infty}^{\infty} sf(z,s)\,ds \qquad (2b)$$

(we will henceforth not explicitly exhibit the dependence of f on the frequency ω). Here, $\sigma^2 = \omega_p{}^2/\omega^2$, where $\omega_p{}^2 = 4\pi n_0 e^2/m$ is the plasma frequency.

The dispersion function associated with the eigenvalues of K is found to be

$$\Omega(\rho) = 1 + \sigma^2\int_{-\infty}^{\infty}\frac{sF'(s)}{1 - s\rho}\,ds,$$

and is related to the dispersion function Λ of Refs. 1 and 2 by $\Lambda(\rho) = \Omega(1/\rho)$

$$\Lambda(\rho) = 1 - \sigma^2\rho\int_{-\infty}^{\infty}\frac{sF'(s)}{s - \rho}\,ds. \qquad (3)$$

In subsequent papers currently in preparation, we consider the uniqueness of solutions to Eq. (2a) and construct explicit solutions. In the present paper we study Λ.

For the case that F is isotropic, it is well known[1,2] that Λ has no zeros if $\sigma^2 < 1$; K then has no eigenvalues, and the plasma waves are dissipative. However, we are interested in anisotropic plasmas, for example, the "bump on tail" or "two stream" equilibria,[3] for which eigenvalues may indeed exist.

In Sec. II we discuss the zeros of Λ. In Sec. III we present the Wiener–Hopf factorization of Λ by analytic functions X and Y. In Sec. IV we obtain the coupled nonlinear integral equations for these functions and discuss their solutions. Our analysis is then in generalization of the isotropic case considered, for example, in Ref. 3.

II. ZEROS OF Λ

The zeros of the plasma dispersion function for fixed k ($\omega_0 = \omega_0(k)$) have been studied extensively (cf., for example, Ref. 3, Chap. 7). Since our interest is in plasma wave boundary value problems rather than the stability of solutions to the initial value problem, we need $k_0(\omega)$, i.e., the zeros for fixed frequency ω. In this section we sketch the procedure we have used for locating the half-plane in which these zeros can occur and quote the results for "bump on tail" and "two stream" equilibrium distributions. Our procedures can easily be generalized to more complicated equilibrium distributions if desired.

We observe that the zeros occur in complex conjugate pairs. Thus it is sufficient to consider only the zeros in the upper-half plane and on the real axis. We shall use the argument principle to determine the number and location (left or right half-plane) of these zeros. We adopt the terminology "complex zeros" to mean a zero with nonvanishing imaginary part.

Theorem 1: For any "single bumped" distribution F, Λ has no zeros for $\sigma^2 < 1$. (We do not consider the singular case $\sigma^2 = 1$ which represents zeros of Λ at ∞.)

This result is well known.[1,2] and in any event can be seen trivially from the appropriate Nyquist diagram for Λ.

[a)]Supported in part by the National Science Foundation Grant Number ENG75-15882.

Suppose F has two bumps. We distinguish three cases for convenience.

Case 1: "Bump on left tail"; $F(u)$ is an ordinary Maxwellian with a bump for $u<0$.

Case 2: "Bump on right tail"; $F(u)$ is an ordinary Maxwellian with a bump for $u>0$.

Case 3: "Two stream"; $F(u)$ has one peak for $u<0$, and one for $u>0$.

In each case F' will vanish at three finite points. We call them $u_0<u_1<u_2$ (for Case 1, $u_2\leqslant 0$; for Case 2, $u_0\geqslant 0$; for Case 3 $u_0<0<u_2$). Mathematically these three cases could be treated as one, but the above division helps clarify the physics. We need the following results;

Lemma 1: For $u>0$, $\operatorname{Im}\Lambda(iu)<0$ in Case 1 and $\operatorname{Im}\Lambda(iu)>0$ in Case 2.

Proof: From Eq. (3),

$$\operatorname{Im}\Lambda(iu)=u^3\sigma^2\int_{-\infty}^{\infty}\frac{F'(s)}{s^2+u^2}\,ds. \tag{4}$$

For Case 1 the only contribution to the integral comes from the perturbing bump, call it F_1. Decompose F_1' into its even and odd parts, F_{1e}' and F_{1o}'. Then only F_{1e}' contributes to the integral in Eq. (4). F_{1e}' will vanish at two points, $\pm y, y>0$. Furthermore, $F_{1e}'(u)<0$ for $|u|<y$ and $F_{1e}'(u)>0$ for $|u|>y$. It is easily seen that the contribution to the integral in Eq. (4) is negative, since if we add

$$0=\frac{-u^3\sigma^2}{y^2+u^2}\int_{-\infty}^{\infty}F_{1e}'(s)\,ds$$

to Eq. (4), we obtain

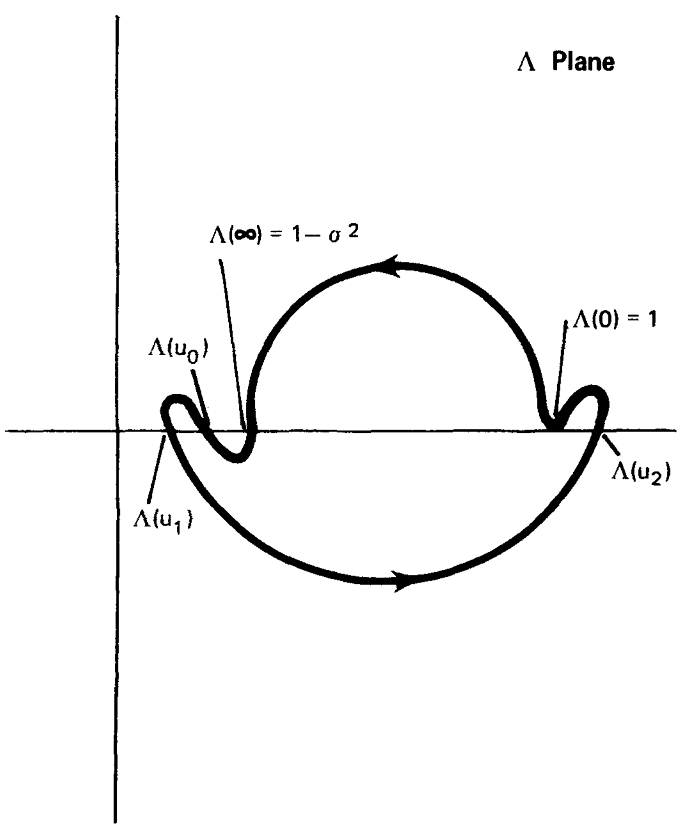


FIG. 1a. Nyquist diagram for Case 1, conditions 1a.

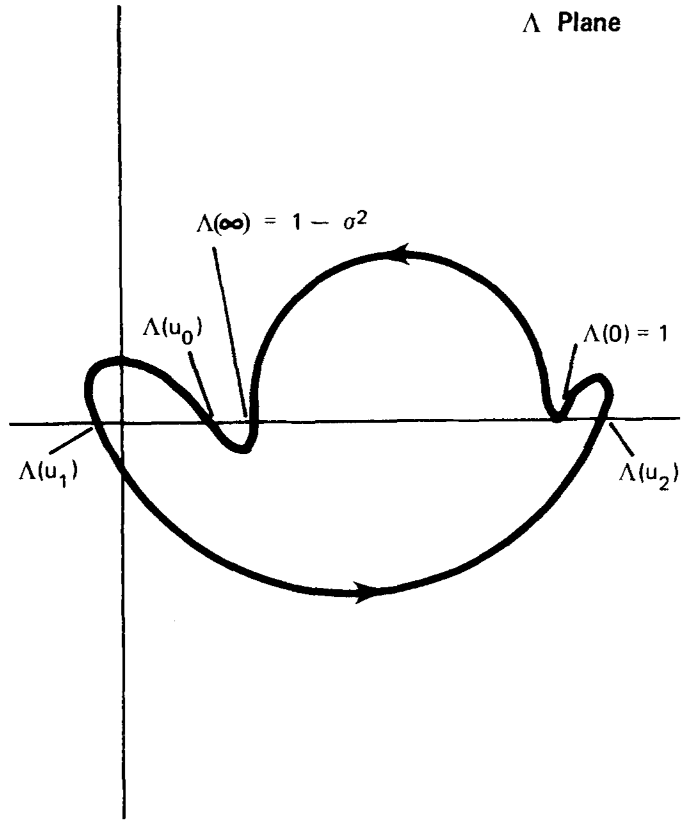


FIG. 1b. Nyquist diagram for Case 1, condition 1b.

$$\operatorname{Im}\Lambda(iu)=u^3\sigma^2\int_{-\infty}^{\infty}\frac{(y^2-s^2)F_{1e}'(s)}{(u^2+s^2)(y^2+u^2)}\,ds, \tag{5}$$

thus completing the proof of the Lemma for Case 1. Case 2 is analogous.

We now let $M=\Lambda(u_0)\Lambda(u_1)\Lambda(u_2)$. We then have *Theorem* 2:

1. $1-\sigma^2>0$:

(a) $M>0$; then for Case 1, 2 and 3, Λ has no zeros;

(b) $M<0$, then, for Case 1, Λ has two zeros in the left half-plane; for Case 2, Λ has two zeros in the right half-plane; for Case 3, Λ has two zeros.

2. $1-\sigma^2<0$:

(a) $M>0$; for Case 1, Λ has two zeros in the right half-plane; for Case 2, Λ has two zeros in the left half-plane; for Case 3, Λ has two zeros;

(b) $M<0$, for Cases 1 and 2, Λ has two zeros in both left and right half-planes; for Case 3, Λ has either no zeros or four zeros.

Proof: We draw the Nyquist diagram for $\Lambda(u)$ as u proceeds from $-\infty$ to $+\infty$ just above the real axis and closes in a semicircular arc in the upper half plane (along the semicircular arch, $\Lambda\sim 1-\sigma^2=$ const, so that portion of the contour makes no change in the argument of Λ). From Eq. (3),

$$\Lambda^{\pm}(u)=\lim_{\epsilon\to 0+}\Lambda(u\pm i\epsilon)$$
$$=1-\sigma^2u\int_{-\infty}^{\infty}\frac{sF'(s)}{s-u}\,ds\mp i\sigma^2u^2F'(u). \tag{6}$$

We show in Figs. 1a and 1b a Nyquist diagram corre-

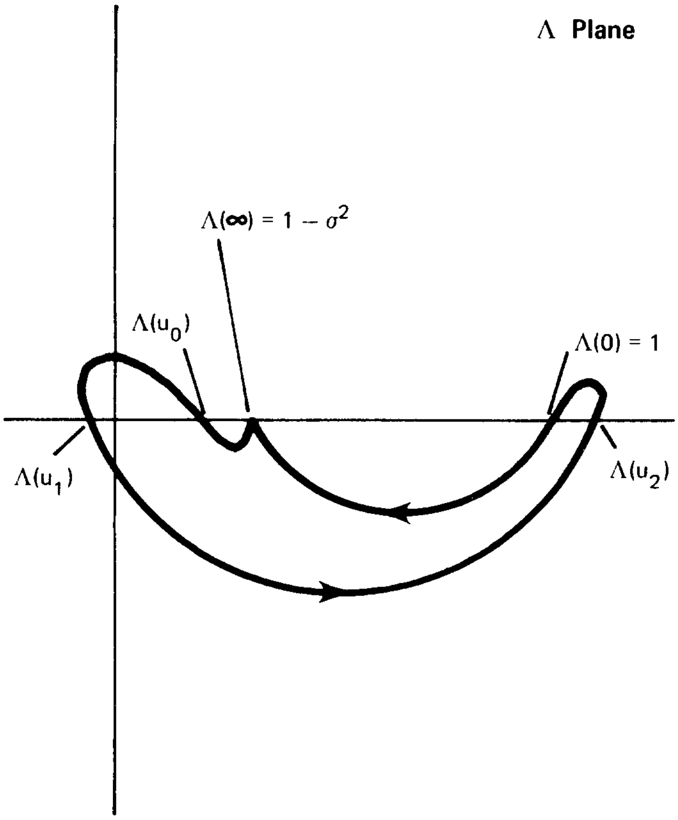


FIG. 2. Nyquist diagram for Case 1, condition 1b, mapping second quadrant only.

sponding to the situations stated in the theorem for Case 1. Note that from Eq. (6), the contour in the Λ plane crosses the real axis at $\pm\infty$, u_1, u_2, u_0, and 0 and since Λ is analytic in the upper half-plane, we must insure that the bounded component of the Λ plane is to the left as one traverses the contour (otherwise the diagram obtained represents a function that has a pole in the upper half-plane). A brief perusal of these figures verifies the assertions of the theorem for Case 1 insofar as the number of zeros is concerned. To prove that the zeros are located in the stated half-plane, consider Lemma 1 and follow the curve in the Λ plane as u advances from $-\infty$ to 0 above the real axis, and from 0 to ∞ upward along the imaginary axis.

Using Lemma 1 we can draw the contour as in Fig. 2, showing that the root does indeed occur in the left half-plane, which completes the proof of the theorem.

The other possibilities for Case 1 are treated in an analogous fashion. In the same way, Cases 2 and 3 can be analyzed. The results are summarized in Tables I and II. We note that real zeros can occur, but only at u_0, u_1, or u_2, where ImΛ vanishes, or at ∞ in the special case $\sigma^2 = 1$.

TABLE I. Zeros of Λ for perturbed Maxwellian. R(L)HP = right (left) half-plane.

	Case 1	Case 2
1. $1-\sigma^2<0$		
a) $M>0$	none	none
b) $M<0$	2 in LHP	2 in RHP
2. $1-\sigma^2>0$		
a) $M>0$	2 in RHP	2 in LHP
b) $M<0$	2 in RHP, 2 in LHP	2 in RHP, 2 in LHP

TABLE II. Zeros of Λ for Case 3.

1. $1-\sigma^2>0$	
a) $M>0$	none
b) $M<0$	2 zeros
2. $1-\sigma^2<0$	
a) $M>0$	2 zeros
b) $M<0$	none or 4 zeros

III. FACTORIZATION OF Λ

For anisotropic plasmas, the factorization of Refs. 1 and 2 is not applicable to the equilibrium distribution functions we consider. We require

$$\Lambda(\rho) = X(\rho)Y(-\rho) \tag{7}$$

with X and Y analytic for Re$\rho<0$. Furthermore, if we let ν_r and ν_l represent the zeros of Λ in the right and left half-planes respectively, then we also require

$$X(\nu_r)=0 \quad \text{for Re}\nu_r>0,$$
$$Y(-\nu_l)=0 \quad \text{for Re}\nu_l<0.$$

If Λ has no zeros, then the following functions are immediately seen to factor Λ:

$$X_0(\rho) = (1-\sigma^2)^{1/2}\exp\left[\frac{1}{2\pi i}\int_0^\infty \ln\left(\frac{\Lambda^+(s)}{\Lambda^-(s)}\right)\frac{ds}{s-\rho}\right], \tag{8a}$$

$$Y_0(\rho) = (1-\sigma^2)^{1/2} \times\exp\left[-\frac{1}{2\pi i}\int_0^\infty \ln\left(\frac{\Lambda^+(-s)}{\Lambda^-(-s)}\right)\frac{ds}{s-\rho}\right]. \tag{8b}$$

To include the zeros of Λ, X_0 and Y_0 must be modified

$$X_1(\rho) = (\rho-\nu_r)(\rho-\bar\nu_r)X_0(\rho), \tag{9a}$$

$$Y_1(-\rho) = (\rho-\nu_r)(\rho-\bar\nu_l)Y_0(-\rho), \tag{9a}$$

X_1 and Y_1 are still not adequate since $\Lambda(\rho)\to 1-\sigma^2$ as $\rho\to\infty$ and the product $X_1(\rho)Y_1(-\rho)$ diverges as $\rho^n(n=2$ or 4). Thus, X_1 and Y_1 must be modified to

$$X(\rho) = X_1(\rho)/\rho^{\epsilon_1}, \tag{10a}$$

$$Y(-\rho) = Y_1(-\rho)/\rho^{\epsilon_2}, \tag{10b}$$

where ϵ_1 (ϵ_2) is either 2 or 0 depending on whether Λ does or does not have a zero in the right (left) half-plane.

To verify that X and Y have no pole at $\rho=0$, we must determine the behavior of $X_0(0)$ and $Y_0(0)$. Since $\rho\sim 0$, the largest contribution in Eq. (8) comes from $s\sim 0$, so that we cut off the range of integration at, say, $a>0$. Then a simple calculation shows for $\rho\to 0$

$$X_0(\rho)\sim\left(\frac{\rho-a}{\rho}\right)^{\theta_1(0)/\pi}\sim\rho^{-\theta_1(0)/\pi}, \tag{11a}$$

$$Y_0(\rho)\sim\left(\frac{\rho+a}{\rho}\right)^{-\theta_2(0)/\pi}\sim\rho^{\theta_2(0)/\pi}, \tag{11b}$$

TABLE III. Values of $\theta_1(0)$ and $\theta_2(0)$.

	Case 1	Case 2
1. $1-\sigma^2>0$		
a) $M>0$	$\theta_1(0)=0$ $\theta_2(0)=0$	$\theta_1(0)=0$ $\theta_2(0)=0$
b) $M<0$	$\theta_1(0)=0$ $\theta_2(0)=2\pi$	$\theta_1(0)=-2\pi$ $\theta_2(0)=0$
2. $1-\sigma^2<0$		
a) $M>0$	$\theta_1(0)=-2\pi$ $\theta_2(0)=0$	$\theta_1(0)=0$ $\theta_2(0)=2\pi$
b) $M<0$	$\theta_1(0)=-2\pi$ $\theta_2(0)=2\pi$	$\theta_1(0)=-2\pi$ $\theta_2(0)=2\pi$

where

$$\theta_1(s)=1/2\ln[\Lambda^+(s)/\Lambda^-(s)],\quad s>0, \tag{11c}$$

$$\theta_2(s)=1/2\ln[\Lambda^+(-s)/\Lambda^-(-s)],\quad s>0, \tag{11d}$$

We observe that

$$\theta_1(0)=\Delta_{(0,\infty)}\arg\Lambda^+ = \Delta_{(0,\infty)}\arg\Lambda^+ + \Delta_{(0,+i\infty)}\arg\Lambda,$$

where the second equality follows from Lemma 1. Similarly,

$$\theta_2(0)=\Delta_{(-\infty,0)}\arg\Lambda^+ = \Delta_{(-\infty,0)}\arg\Lambda^+ + \Delta_{(0,+i\infty)}\arg\Lambda,$$

where again we have used Lemma 1. Clearly, if the root occurs in the left half-plane, $\theta_1(0)=0$ and $\theta_2(0)=2\pi$, whereas if there is a root in the right half-plane, $\theta_1(0)=-2\pi$ and $\theta_2(0)=0$. The results are summarized in Table III. Comparing these results with our definition of ϵ_1 and ϵ_2, we have

$$\epsilon_1=-\theta_1(0)/\pi, \tag{12a}$$

$$\epsilon_2=\theta_2(0)/\pi, \tag{12b}$$

We see, incidentally, that the existence of a zero of Λ in the right or left half-plane induces a double zero in X_0 or Y_0 respectively, so that the notation in Eq. (9) is appropriate.

The result, Eq. (12), directly implies the following theorem.

Theorem 3: The functions X and Y defined by Eqs. (8), (9), and (10) constitute a Wiener–Hopf factorization of Λ given by Eq. (7) and

$X(\rho)\sim$const as $\rho\to 0$.

$Y(\rho)\sim$const as $\rho\to 0$.

Without loss of generality, we may set $X(0)=Y(0)=1$.

IV. COUPLED NONLINEAR INTEGRAL EQUATIONS

For computational purposes, the explicit representation of X and Y obtained in the previous section may not be so convenient as the iterative solution of coupled integral equations. These may easily be determined from Cauchy's theorem. In particular, from Eq. (7)

$$X^+(u)-X^-(u)=\frac{1}{Y(-u)}\times[\Lambda^{+1}(u)-\Lambda^-(u)],\quad u>0, \tag{13a}$$

$$[Y(-u)]^+-[Y(-u)]^-=\frac{1}{X(u)}[\Lambda^+(u)-\Lambda^-(u)],\quad u<0. \tag{13b}$$

Using Eq. (6) and the behavior of X and Y at infinity, Cauchy's theorem yields

$$X(\rho)=(1-\sigma^2)^{1/2}-\int_0^\infty\frac{s^2\sigma^2F'(s)}{Y(-s)(s-\rho)}ds, \tag{14a}$$

$$Y(-\rho)=(1-\sigma^2)^{1/2}-\int_{-\infty}^0\frac{s^2\sigma^2F'(s)}{X(s)(s-\rho)}ds, \tag{14b}$$

These equations can be solved iteratively for the values of $X(\rho)$ and $Y(-\rho)$. A more convenient iteration scheme is defined by taking the limit as $\rho\to 0$ in Eqs. (14). Then

$$1=(1-\sigma^2)^{1/2}-\int_0^\infty s\sigma^2F'(s)/Y(-s)\,ds, \tag{15a}$$

$$1=(1-\sigma^2)^{1/2}-\int_{-\infty}^0 s\sigma^2F'(s)/X(s)\,ds, \tag{15b}$$

and rewriting Eq. (14) as

$$X(-\rho)=1+\rho\sigma^2\int_0^\infty\frac{sF'(s)}{Y(-s)(s+\rho)}ds, \tag{16a}$$

$$Y(-\rho)=1-\rho\sigma^2\int_0^\infty\frac{sF'(-s)}{X(-s)(s+\rho)}ds, \tag{16b}$$

If we make the following change of dependent variable,

$$U_1(\rho)=X^{-1}(-\rho)\rho\sigma^2F'(-\rho), \tag{17a}$$

$$U_2(\rho)=-Y^{-1}(-\rho)\rho\sigma^2F'(\rho), \tag{17b}$$

then Eqs. (14) reduce to the bilinear matrix equation

$$\mathbf{U}=\mathbf{F}+\mathbf{A}(\mathbf{U},\mathbf{U}), \tag{18}$$

where

$$\mathbf{U}=[U_1,U_2], \tag{19a}$$

$$\mathbf{F}(\rho)=\rho\sigma^2[F'(-\rho),-F'(\rho)], \tag{19b}$$

$$\mathbf{A}(\mathbf{U},\mathbf{V})(\rho)=\left[-\rho\int_0^\infty V_1(\rho)U_2(s)\frac{ds}{s+\rho},\rho\int_0^\infty U_1(s)V_2(\rho)\frac{ds}{s+\rho}\right]. \tag{19c}$$

The convergence of the iteration scheme to Eq. (18) has been studied previously.[4] If we define a Banach space with

$$\|\mathbf{U}\|=\max_{i=1,2}\int_0^\infty|U_i(s)|\,ds,$$

then it is shown by fixed point arguments that Eq. (18) has a unique solution in the ball

$$S=\{\mathbf{U}:\|\mathbf{U}-\mathbf{F}\|<\tfrac12\},$$

subject to the condition $\|\mathbf{F}\|<\frac12$, and that an iteration scheme converges if the initial guess is chosen in S (note that if $\mathbf{U}\in S$, then $\|\mathbf{U}\|<1$).

We now show that the solutions to Eq. (18) lying in S is the "physical" solution. We observe that U_1 and U_2 obey

$$U_1(\rho)=F_1(\rho)\left[1+\int_0^\infty U_2(s)\frac{\rho}{s+\rho}ds\right]^{-1}, \tag{20a}$$

$$U_2(\rho) = F_2(\rho)\left[1 - \int_0^\infty U_1(s)\frac{\rho}{s+\rho}\,ds\right]^{-1}, \qquad (20b)$$

Consider Case 1 for situation 1b of Table I. Then Y has zeros in the right half-plane which implies that U_2 has poles in the left half-plane. Thus U_2 must be analytic in the right half-plane and U_1 analytic in $\mathbb{C}\setminus[-\infty,0]$. Since we are dealing with nonlinear integral equations which may have more than one solution, we must prove the following.

Theorem 4: For Case 1, $1-\sigma^2>0$ and $M<0$, the solution to Eq. (18), $[U_1, U_2]$, in the ball S, is analytic in the right half-plane.

Proof: Writing $\rho = \alpha x + i\beta$, we have

$$\left|\frac{\rho}{\rho+s}\right| = \left[\frac{\alpha^2+\beta^2}{(\alpha+s)^2+\beta^2}\right]^{1/2} < 1 \quad \text{for } 0<s<\infty$$

and $\alpha > 0$;

thus

$$\left|\int_0^\infty U_{1,2}(s)\frac{\rho}{\rho+s}\,ds\right| < \int_0^\infty |U_{1,2}(s)|\left|\frac{\rho}{\rho+s}\right| ds < 1.$$

The result follows from Eq. (20).

To get a feeling for the range of parameters for which an iteration scheme corresponding to Eq. (18) converges, we have computed $\|\mathbf{F}\|$ for a bump on tail distribution considered in Ref. 3 and given explicitly by

$$F(u) = (1-\beta)\left(\frac{m}{2\pi kT_1}\right)^{1/2}\exp\left(-\frac{mu^2}{2kT_1}\right) + \beta\left(\frac{m}{2\pi kT_2}\right)^{1/2}\exp\left[-\frac{m(u-V_0)^2}{2kT_2}\right]. \qquad (21)$$

A straightforward integration yields

$$\|\mathbf{F}\| \leqslant \sigma^2\left[\tfrac{1}{2} - \tfrac{1}{2}\beta E_2\left(\frac{mV_0^{\,2}}{2kT_2}\right)^{1/2} + 2\beta\left(\frac{mV_0^{\,2}}{2\pi kT_2}\right)^{1/2}\right], \qquad (22)$$

where E_2 is the error function defined in Ref. 5. A more convenient, if less exact, bound is

$$\|\mathbf{F}\| < \sigma^2\left[\tfrac{1}{2} + 2\beta\left(\frac{mV_0^{\,2}}{2\pi kT_2}\right)^{1/2}\right]. \qquad (23)$$

We conclude that, for certain values of σ^2, β, V_0, and T_2, an iteration scheme converges if the initial guess is chosen in S. For values outside this range, it is necessary to evaluate X and Y from the explicit definitions. We now develop these definitions into a form more useful for computation by a procedure similar to one used in Ref. 2, p. 130.

From Eqs. (8), (9), and (10) we have

$$X(\rho) = (\rho - \nu_r)(\rho - \bar{\nu}_r)\rho^{-\epsilon_1}(1-\sigma^2)^{1/2} \times \exp\left[\frac{1}{\pi}\int \frac{\theta(s)}{s-\rho}\,ds\right], \qquad (24a)$$

$$Y(-\rho) = (\rho - \nu_l)(\rho - \bar{\nu}_l)\rho^{-\epsilon_2}(1-\sigma^2)^{1/2} \times \exp\left[\frac{1}{\pi}\int_{-\infty}^0 \frac{\theta(s)}{s-\rho}\,ds\right], \qquad (24b)$$

where

$$\theta(s) = \arg\Lambda^+(s) = \tan^{-1}\left[\frac{-\pi\sigma^2 s^2 F'(s)}{\lambda(s)}\right], \qquad (25a)$$

$$\lambda(s) = \tfrac{1}{2}[\Lambda^+(s) + \Lambda^-(s)]. \qquad (25b)$$

It is useful to write in Eq. (24)

$$\int_0^\infty \frac{\theta(s)}{s-\rho}\,ds = \int_0^\infty \theta(s)\frac{d}{ds}\ln(s-\rho)\,ds, \qquad (26)$$

and, integrating by parts, we have

$$\int_0^\infty \frac{\theta(s)}{s-\rho}\,ds = \epsilon_1\ln(-\rho) - \int_0^\infty \frac{d\theta}{ds}\ln(s-\rho)\,ds, \qquad (27a)$$

$$\int_{-\infty}^0 \frac{\theta(s)}{s-\rho}\,ds = \epsilon_2\ln(-\rho) - \int_{-\infty}^0 \frac{d\theta}{ds}\ln(s-\rho)\,ds. \qquad (27b)$$

Here we have used Eq. (12) and the fact that $\theta(\infty)=0$. Calculating $d\theta/ds$ from Eq. (25a) and using (24) and (27), we obtain after some algebra

$$X(\rho) = (\rho-\nu_r)(\rho-\bar{\nu}_r)(1-\sigma^2)^{1/2} \times \exp\left\{-\frac{1}{\pi}\int_0^\infty \mathrm{Im}\left[\frac{\Lambda^{+\prime}(s)}{\Lambda^+(s)}\right]\ln(s-\rho)\,ds\right\}, \qquad (28a)$$

$$Y(-\rho) = (\rho-\nu_l)(\rho-\bar{\nu}_l)(1-\sigma^2)^{1/2} \times \exp\left\{-\frac{1}{\pi}\int_{-\infty}^0 \mathrm{Im}\left[\frac{\Lambda^{+\prime}(s)}{\Lambda^+(s)}\right]\ln(s-\rho)\,ds\right\}, \qquad (28b)$$

We have proved that the Eq. (18) has a solution in the ball $S = \{\mathbf{V}: \|\mathbf{V}-\mathbf{F}\| < \tfrac{1}{2} \text{ if } \|\mathbf{F}\| < \tfrac{1}{2}\}$, which we now assume.

V. DISCUSSION

For computational purposes, we investigated the coupled nonlinear integral equation for X and Y in Sec. IV and found that, under certain restrictive conditions on the equilibrium distribution function, we could prove that an iteration scheme for the solution does converge. It might be possible to remove some of these conditions using other analytical techniques although thus far we have not been able to do so.

In Sec. III we found the Wiener–Hopf factorization of Λ, which is necessary in solving the half range problem using the Larsen and Habetler technique,[6] as well as using techniques developed in Ref. 2. That is the problem we are currently pursuing.

[1]F.C. Sure, Ph.D. Dissertation, University of Michigan, 1963.
[2]K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addision-Wesley, Reading, Mass., 1967).
[3]N.A. Krall and A.W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill, New York, 1973).
[4]R.L. Bowden, R. Menikoff, and P.F. Zweifel, J. Math. Phys. **17**, 9 (1976).
[5]E. Jahnke and F. Emde, *Tables of Functions* (Dover, New York, 1945), p. 23.
[6]E.W. Larsen and G.J. Habetler, Commun. Pure. Appl. Math. **26**, 529 (1973).

An approximation method for electrostatic Vlasov turbulence

Alexander J. Klimas
NASA/Goddard Space Flight Center, Laboratory for Extraterrestrial Physics, Greenbelt, Maryland, 20771

(Received 10 April 1979; accepted for publication 4 May 1979)

Electrostatic Vlasov turbulence in a bounded spatial region is considered. An iterative approximation method with a proof of convergence is constructed. The method is nonlinear and applicable to strong turbulence.

I. INTRODUCTION

Consider the one-dimensional Vlasov–Maxwell system of equations,

$$\frac{\partial F}{\partial \tau} + v\frac{\partial F}{\partial x} - E(x,\tau)\frac{\partial F}{\partial v} = 0, \qquad (1)$$

$$\frac{\partial E}{\partial x} = 1 - \int_{-\infty}^{\infty} dv\, F(x,v,\tau) = 1 - n(x,\tau), \qquad (2)$$

$$\frac{\partial E}{\partial \tau} = \int_{-\infty}^{\infty} dv\, vF(x,v,\tau) = u(x,\tau), \qquad (3)$$

for the electron distribution function, $F(x,v,\tau)$, and the electric field $E(x,\tau)$, with a stationary and uniform ion background. Assume the existence of a solution of this system of equations for $-1 \leqslant x \leqslant 1$ (x is dimensionless, and measured in units of an arbitrary length scale L), $-\infty < v < \infty$, and $\tau > 0$. Then a method for constructing approximations to this solution can be developed as follows:

The Fourier–Fourier transform method[1] can be used to transform Eqs. (1)–(3) into an infinite system of first order hyperbolic partial differential equations. This system has been studied before.[1] One further aspect of this system will be considered here; it is that on truncating the infinite system, the resulting finite system is of a standard form which has been used to produce constructive proofs of the existence of solutions to a wide class of such systems.[2] Here, the existence of a solution to the finite system will be assumed. The methods used for proofs of existence will be applied, nevertheless, to produce approximations to the exact solution of the finite system. The existence of the finite system solution and the issue of how well it approximates the solution of the infinite system will be addressed elsewhere.

II. THE FOURIER–FOURIER TRANSFORM

Let

$$f(x,v,\tau) = \sum_{m=-\infty}^{\infty} e^{im\pi x}\frac{1}{2}\int_{-\infty}^{\infty} d\nu\, e^{(-i\pi v\nu)} f_m(\nu,\tau) \quad (4)$$

and

$$\epsilon(x,\tau) = \sum_{m=-\infty}^{\infty} e^{im\pi x}\epsilon_m(\tau),$$

in which

$$f_m(\nu,\tau) = \frac{1}{2}\int_{-1}^{1} dx\, e^{-im\pi x}\int_{-\infty}^{\infty} dv\, e^{i\pi\nu v}F(x,v,\tau) \quad (5)$$

and

$$\epsilon_m(\tau) = \frac{1}{2}\int_{-1}^{1} dx\, e^{-im\pi x}E(x,\tau),$$

then $f = F$ and $\epsilon = E$ for $-1 < x < 1$, and on the boundaries $f(\pm 1,v,\tau) = \frac{1}{2}[F(1,v,\tau) + F(-1,v,\tau)]$ and $\epsilon(\pm 1,\tau) = \frac{1}{2}[E(1,\tau) + E(-1,\tau)]$. Both f and ϵ are periodic in x with period two.

From Eqs. (1)–(3),

$$\frac{\partial f_m}{\partial \tau} + m\frac{\partial f_m}{\partial \nu} + i\pi\nu\sum_{n=-\infty}^{\infty}\epsilon_n f_{m-n} = J_m, \qquad (6)$$

$$im\pi\epsilon_m = \delta_{m,0} - f_m(0,\tau) + (-1)^{m+1}(1 - f_0(0,\tau)), \qquad (7)$$

and

$$\frac{d\epsilon_m}{d\tau} = -\frac{i}{\pi}\frac{\partial f_m}{\partial \nu}\Big|_{\nu=0} \qquad (8)$$

in which $J_m = (-1)^m J_0$, and

$$J_0(\nu,\tau) = \frac{i}{2\pi}\frac{\partial}{\partial \nu}[F(1,\nu,\tau) - F(-1,\nu,\tau)],$$

where

$$F(\pm 1,\nu,\tau) = \int_{-\infty}^{\infty} dv\, e^{i\pi\nu v}F(\pm 1,v,\tau).$$

Equations (7) and (8) are redundant when $m \neq 0$, and Eq. (7) yields no information when $m = 0$. In the following Eq. (7) will be used to determine the $\epsilon_m(\tau)$ when $m \neq 0$, and Eq. (8) will be used to determine $\epsilon_0(\tau)$.

It will be assumed in the following that $f_m = 0$ for $m > M$, where M is arbitrarily large but finite. Since it is expected that those modes which have wavelengths comparable or shorter than the Debye length in the plasma will be strongly damped,[3] there is perhaps some *a priori* justification for the truncation.

In Appendix A it is argued that the solution of Eqs. (1)–(3) is determined by a choice of $F(x,v,0)$ on the initial plane $(-1 \leqslant x \leqslant 1, -\infty < v < \infty, \tau = 0)$, of $\Delta(v,\tau) = F(-1,v,\tau) - F(1,v,\tau)$ in terms of F on the boundaries $(x = \pm 1, -\infty < v < \infty, \tau > 0)$, and of $\epsilon_0(0)$. That argument does not depend on the assumption of a neutral plasma (space averaged). In the following a neutral plasma will be assumed and then it will be shown that the choice of the equivalent quantities, $f_m(\nu,0)$, $J_0(\nu,\tau)$, and $\epsilon_0(0)$ uniquely determines the solution of the truncated Eqs. (6)–(8).

III. THE NEUTRAL PLASMA

The space-averaged electron density is $f_0(0,\tau)$. The restriction to a neutral plasma is affected by setting $f_0(0,\tau) = 1$. This restriction is consistent with Eq. (6) only for a limited class of $J_0(\nu,\tau)$. From Eq. (6),

$$\frac{\partial f_0(0,\tau)}{\partial \tau} = J_0(0,\tau) = \tfrac{1}{2}\delta u(\tau),$$

where $\delta u(\tau) = u(-1,\tau) - u(1,\tau)$. In the following it will be assumed that $J_0(0,\tau) = 0$ and $f_0(0,\tau,) = f_0(0,0) = 1$.

Fyfe and Montgomery[4] have noted that $\epsilon_0(\tau)$ cannot be chosen freely. They have produced an exact solution for $\epsilon_0(\tau)$ and $u_0(\tau)$ [the space average of $u(x,\tau)$] from their model of the one-dimensional Vlasov–Maxwell plasma. Their results apply to the periodic plasma [$\Delta(v,\tau) = 0$]. A generalization to the nonperiodic neutral plasma being considered here is possible.

From Eq. (8) [or Eq. (3)],

$$\begin{aligned}\frac{d\epsilon_0}{d\tau} &= -\frac{i}{\pi}\frac{\partial f_0}{\partial \nu}\bigg|_{\nu} = 0\\ &\equiv -\frac{i}{\pi}f_0'(0,\tau)\\ &= u_0,\end{aligned} \tag{9}$$

and from Eq. (6) [or Eq. (1)],

$$-\frac{i}{\pi}\frac{\partial f_0'(0,\tau)}{\partial \tau} + \epsilon_0(\tau) = J_0'(0,\tau)$$

or

$$\frac{du_0}{d\tau} + \epsilon_0 = \tfrac{1}{2}\delta P, \tag{10}$$

where $\delta P(\tau) = P(-1,\tau) - P(1,\tau)$ and

$$P(x,\tau) = \int_{-\infty}^{\infty} dv\, v^2 F(x,v,\tau).$$

(Notice that since for the neutral plasma $u(-1,\tau) = u(1,\tau)$, δP is actually just the difference in electron plasma temperature at the boundaries.) An exact solution of Eqs. (9) and (10) is available; it is

$$\begin{pmatrix}\epsilon_0(\tau)\\ u_0(\tau)\end{pmatrix} = \begin{pmatrix}\epsilon_0(0)\cos\tau + u_0(0)\sin\tau\\ u_0(0)\cos\tau - \epsilon_0(0)\sin\tau\end{pmatrix} + \frac{1}{2}\int_0^\tau d\lambda\, \delta P(\lambda)\begin{pmatrix}\sin(\tau-\lambda)\\ \cos(\tau-\lambda)\end{pmatrix}. \tag{11}$$

The result of Fyfe and Montgomery is regained when $\delta P = 0$.

Notice that it is possible to obtain large $\epsilon_0(\tau)$ and $u_0(\tau)$ due to an approximately linear growth of the integral in Eq. (11) with increasing τ, if $\delta P(\tau)$ contains harmonic oscillations with period one (the inverse plasma frequency). It is not possible for $\epsilon_0(\tau)$ to be constant in time unless δP is also constant, $\epsilon_0(0) = \tfrac{1}{2}\delta P$ and $u_0(0) = 0$. Under these conditions $u_0(\tau) = u_0(0) = 0$. Since $\epsilon_0(\tau)$ is a measure of the potential difference on the boundaries, it should be noted that the preceding statements concerning $\epsilon_0(\tau)$ also apply to that potential difference. All of the above, and any other consequences of Eq. (11), apply exactly for the neutral plasma no matter what else is occuring in the plasma.

IV. BASIC INTEGRAL EQUATION

Given the solution for $\epsilon_0(\tau)$ and $u_0(\tau)$, a major reduction in the complexity of Eqs. (6)–(8) can be made by introducing a new dependent variable through

$$f_m(\nu,\tau) = K_m(\nu,\tau)\exp\left\{-i\pi\int_0^\tau d\lambda\,[\nu - m(\tau-\lambda)]\times\epsilon_0(\lambda)\right\}.$$

Then

$$\frac{\partial K_m}{\partial \tau} + m\frac{\partial K_m}{\partial \nu} - \nu\sum_{n=-M}^{M}{}'\left(\frac{1}{n}\right)K_n(0,\tau)K_{m-n} = \sigma_m, \tag{12}$$

where

$$\sigma_m(\nu,\tau) = J_m(\nu,\tau)\exp\left\{i\pi\int_0^\tau d\lambda\,[\nu - m(\tau-\lambda)]\epsilon_0(\lambda)\right\},$$

and the prime on the summation symbol indicates that the $n = 0$ term is omitted. Since $\epsilon_0(\tau)$ can be considered a known function of time, a solution of Eq. (12) for K_m is equivalent to a solution of Eqs. (6)–(8) for f_m. Notice that in the special case of a periodic plasma ($J_m = 0$) Eq. (12) becomes independent of $\epsilon_0(\tau)$. Thus, a single solution of Eq. (12), which will be shown to be determined solely by $K_m(\nu,0)$, is equivalent to the entire class of solutions for f_m which contains all possible choices of $\epsilon_0(0)$.

Using the method of characteristics,[5] Eq. (12) can be integrated once to obtain

$$\begin{aligned}K_m(\nu,\tau) = K_m(\nu - m\tau,0) &+ \int_0^\tau d\lambda\,\sigma_m(\nu - m(\tau-\lambda),\lambda)\\ &+ \sum_{n=-M}^{M}{}'\left(\frac{1}{n}\right)\int_0^\tau d\lambda\,[\nu - m(\tau-\lambda)]\\ &\times K_n(0,\lambda)K_{m-n}(\nu - m(\tau-\lambda),\lambda).\end{aligned} \tag{13}$$

Equation (13) will play a central role in the following development of approximations to K_m.

V. THE APPROXIMATION METHOD

The result which will be obtained in this section can be simply stated as follows:

A sequence of functions $K_m(\nu,\tau;\alpha)$ will be introduced with

$$K_m(\nu,\tau;0) = K_m(\nu - m\tau,0) + \int_0^\tau d\lambda\,\sigma_m(\nu - m(\tau-\lambda),\lambda). \tag{14}$$

Since $F(x,v,0)$ will be assumed given, $K_m(\nu,0)$ can be considered a known function of ν which is uniquely related to the initial F; σ_m can be determined from $\Delta(v,\tau)$. The other members of the sequence are to be related to each other through

$$\frac{\partial K_m(\alpha+1)}{\partial \tau} + m\frac{\partial K_m(\alpha+1)}{\partial \nu} - \nu \times\sum_{n=-M}^{M}{}'\left(\frac{1}{n}\right)K_n(0,\tau;\alpha)K_{m-n}(\alpha) = \sigma_m, \tag{15}$$

or

$$K_m(\nu,\tau;\alpha+1) = K_m(\nu,\tau;0)$$

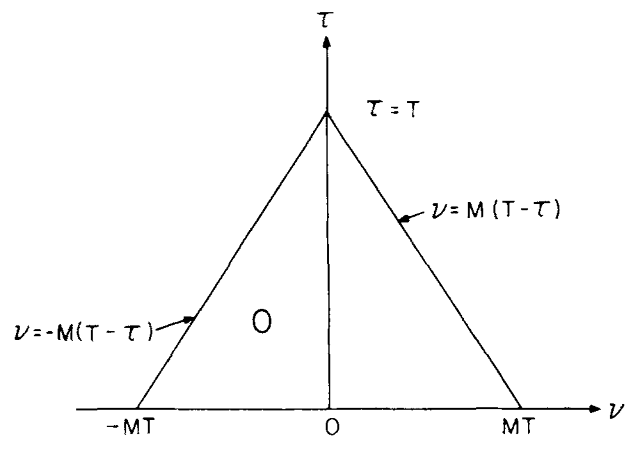


FIG. 1. The domain O on the (ν,τ) plane.

$$+\sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^\tau d\lambda\,[\nu - m(\tau-\lambda)]$$
$$\times K_n(0,\lambda;\alpha)K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha). \quad (16)$$

It will be shown that $\lim_{\alpha\to\infty}K_m(\nu,\tau;\alpha) = K_m(\nu,\tau)$, i.e., it will be shown that the sequence of approximations must converge to the exact solution. This convergence will not depend on the presence of a small parameter for expansion purposes and will apply for any finite value of M. Thus, this approximation method applies to strong turbulence with any finite number of wave modes no matter how large. Notice that the character of the method is to place any member of the sequence in quadrature [through Eq. (16)]; it does not produce equations which must be solved.

A. Preliminaries to proof of convergence

To facilitate the proof of convergence the K_m will be assumed as vector components of a $(2M+1)$-dimensional vector function, $K(\nu,\tau) = (K_{-M}(\nu,\tau), K_{-M+1}(\nu,\tau) \cdots K_M(\nu,\tau))$. Equation (16) can be considered as an integral transformation which relates members of the sequence through $K(\alpha+1) = TK(\alpha)$.

The proof of convergence will be given on the closed domain, O, pictured in Fig. 1. An examination of Eq. (16) will show that knowledge of $K(\alpha)$ on O is necessary and sufficient to determine $K(\alpha+1)$ on O. The domain O is centered on the line $\nu = 0$ since on that line all of the quantities of physical interest (the various moments of F as well as the Fourier components of the electric field) can be found. The time T is any finite time of interest; the solution will be obtained for all $0<\tau\leqslant T$.

A further subdivision of O is necessary. Imagine the strips on the (ν,τ) plane defined by $r\delta<\tau\leqslant(r+1)\delta$ where $r = 0,1,2,\cdots$ and $\delta>0$ is to be determined. Then, let $O_{r\delta}$ be the intersection of the rth strip with O. The transformation T will be shown to be a contraction[6] on each of the $O_{r\delta}$, and then the results for the stationary element of T on each strip will be pieced together to yield Eq. (16).

The following definition of a norm will be used. Let

$$\|K\|_O = \sup_{(\text{On domain } O)} (\max\{|K_{-M}(\nu,\tau)|,\cdots,|K_M(\nu,\tau)|\}).$$

At each point (ν,τ) the absolute values of all of the $K_m(\nu,\tau)$ are to be taken, and then the maximum of these chosen. Then, the supremum, on O, of the resulting function is to be found. The vector K and the domain O have been used here for illustrative purposes. Other vectors and domains will appear, but in each case, the symbol $\|\ \|_O$ has the analogous meaning.

In Appendix B it is shown that in the limit $M=\infty$, $\|K\|_O = 1$. In the following it will be assumed that $\|K\|_O$ exists for finite M. This is not actually a new assumption; on O, it is totally equivalent to the earlier assumption of the existence of a solution to the truncated system. It will also be assumed that the boundary conditions are chosen so that $\|\sigma\|_O$ exists. The number $N = 2(\|K\|_O + \|\sigma\|_O T)$ will be used.

B. Convergence on a narrow strip

Equation (13) can be used to show that

$$K_m(\nu,\tau) = K_m(\nu - m(\tau - r\delta), r\delta)$$
$$+\int_{r\delta}^\tau d\lambda\,\sigma_m(\nu - m(\tau-\lambda),\lambda)$$
$$+\sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_{r\delta}^\tau d\lambda\,[\nu - m(\tau-\lambda)]$$
$$\times K_n(0,\lambda)K_{m-n}(\nu - m(\tau-\lambda),\lambda) \quad (17)$$

for any $\tau\geqslant r\delta$. The proof that on $O_{r\delta}$,

$$K_m(\nu,\tau) = K_m(\nu - m(\tau - r\delta), r\delta)$$
$$+\int_{r\delta}^\tau d\lambda\,\sigma_m(\nu - m(\tau-\lambda),\lambda)$$
$$+\lim_{\alpha\to\infty}\sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_{r\delta}^\tau d\lambda\,[\nu - m(\tau-\lambda)]$$
$$\times K_n(0,\lambda;\alpha)K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha) \quad (18)$$

will now be given with $K_m(\nu,\tau;0)$ determined by Eq. (14). The symbol $T_{r\delta}$ will be used for the integral transformation in Eq. (17) on $O_{r\delta}$. The proof is in three parts.

1. Part one

Let $V(\nu,\tau)$ be any $(2M+1)$-dimensional vector field on $O_{r\delta}$ with its norm bounded such that $\|V\|_{O_{r\delta}}\leqslant N$. Define $V' = T_{r\delta}V$. Then, $\|V'\|_{O_{r\delta}}\leqslant N$ for δ small enough, but finite.

Proof:

$$V'_m(\nu,\tau) = K_m(\nu - m(\tau - r\delta), r\delta)$$
$$+\int_{r\delta}^\tau d\lambda\,\sigma_m(\nu - m(\tau-\lambda),\lambda)$$
$$+\sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_{r\delta}^\tau d\lambda\,[\nu - m(\tau-\lambda)]$$
$$\times V_n(0,\lambda)V_{m-n}(\nu - m(\tau-\lambda),\lambda).$$

Therefore

$$|V'_m(\nu,\tau)|\leqslant\|K\|_O + \delta\|\sigma\|_O + \delta N^2[MT + \tfrac{1}{2}M\delta]$$
$$\times\sum_{n=-M}^{M}{}' \left|\frac{1}{n}\right|.$$

This result holds for any m and any point, (ν,τ). Thus, $\|V'\|_{O_{r\delta}}\leqslant N$ for δ small enough.

Discussion: The significance of this part of the proof lies in the fact that *any* initial vector field which is normed as above can be chosen. Then, the transformation $T_{r\delta}$ can be applied an arbitrary number of times and the result will be a vector field with the same bound on its norm as the initial vector field. Notice that $K(\nu,\tau;0)$, as defined by Eq. (14), is a suitable initial vector field.

2. Part two

Let $V(\nu,\tau)$ and $W(\nu,\tau)$ be any $(2M+1)$-dimensional vector fields on $O_{r\delta}$ such that both $\|V\|_{O_{r\delta}} \leqslant N$ and $\|W\|_{O_{r\delta}} \leqslant N$. Define $V' = T_{r\delta} V$ and $W' = T_{r\delta} W$. Then, $\|V' - W'\|_{O_{r\delta}} \leqslant \frac{1}{2}\|V - W\|_{O_{r\delta}}$ for δ small enough.

Proof:

$$\begin{aligned}
&|V'_m(\nu,\tau) - W'_m(\nu,\tau)| \\
&\leqslant \sum_{n=-M}^{M}{}' \left|\frac{1}{n}\right| \left|\int_{r\delta}^{\tau} d\lambda\, |\nu - m(\tau-\lambda)|\right. \\
&\cdot |V_n(0,\lambda) V_{m-n}(\nu - m(\tau-\lambda),\lambda) \\
&\quad - W_n(0,\lambda) W_{m-n}(\nu - m(\tau-\lambda),\lambda)| \\
&\leqslant \delta\left([MT + \tfrac{1}{2}M\delta] 2N \sum_{n=-M}^{M}{}' \left|\frac{1}{n}\right|\right) \|V - W\|_{O_{r\delta}}.
\end{aligned}$$

Thus, $\|V' - W'\|_{O_{r\delta}} \leqslant \frac{1}{2}\|V - W\|_{O_{r\delta}}$ if δ is small enough.

Discussion: This part of the proof shows that $T_{r\delta}$ is a contraction. When $T_{r\delta}$ is applied to the difference of two normed vector fields, the resulting difference is reduced. In view of part 1. of this proof, $T_{r\delta}$ can be applied an arbitrary number of times to a pair of suitably chosen initial vector fields with the difference between the resulting vector fields reduced each time. In the following, this basic property will be used to find the stationary element of $T_{r\delta}$.

3. Part three

Let $K_m(\nu,\tau;0)$ be defined by Eq. (14). Define $K_m(\nu,\tau;\alpha) = T_{r\delta} K_m(\nu,\tau;\alpha - 1)$ on $O_{r\delta}$. Then, $K_m(\nu,\tau) = \lim_{\alpha\to\infty} K_m(\nu,\tau;\alpha)$ on $O_{r\delta}$.

Proof:

$$|K_m(\nu,\tau;0)| \leqslant \|K\|_O + \|\sigma\|_O T = \tfrac{1}{2}N$$

Therefore, $\|K(0)\|_{O_{r\delta}} \leqslant \frac{1}{2}N < N$. From part 1. then, $\|K(\alpha)\|_{O_{r\delta}} \leqslant N$ for all α. Now, in the result of part 2. of this proof, let $K(\alpha) = V$ and $K(\beta) = W$. Then,

$$\|K(\alpha) - K(\beta)\|_{O_{r\delta}} \leqslant \tfrac{1}{2}\|K(\alpha-1) - K(\beta-1)\|_{O_{r\delta}}$$

Suppose $\alpha = \beta$. Then, $\|K(\alpha) - K(\beta)\|_{O_{r\delta}} = 0$ for all α. Suppose, $\alpha > \beta \geqslant 1$. Then, $\|K(\alpha) - K(\beta)\|_{O_{r\delta}} \leqslant \frac{1}{2}^{\beta}\|K(\alpha-\beta) - K(0)\|_{O_{r\delta}} \leqslant 4N(1/2)^{\beta}$. Thus, $\|K(\alpha) - K(\beta)\|_{O_{r\delta}} \to 0$ as $(\alpha,\beta) \to \infty$. Similarly $\|K(\alpha) - K(\beta)\|_{O_{r\delta}} \to 0$ as $(\alpha,\beta) \to \infty$ when $\beta > \alpha \geqslant 1$. Thus, $K(\alpha)$ is a Cauchy sequence[7] on $O_{r\delta}$. $K(\alpha)$ converges uniformly to K^* on $O_{r\delta}$ where $K^* = \lim_{\alpha\to\infty} K(\alpha) = \lim_{\alpha\to\infty} T_{r\delta} K(\alpha - 1) = T_{r\delta} K^*$. Since $K^* = T_{r\delta} K^*$ it is a stationary element of $T_{r\delta}$. But, it is easy to see that the stationary element of $T_{r\delta}$ is unique.

Suppose there are two stationary elements, K^* and L^*. Then, $K^* - L^* = T_{r\delta} K^* - T_{r\delta} L^*$, and

$$\begin{aligned}
\|K^* - L^*\|_{O_{r\delta}} &= \|T_{r\delta} K^* - T_{r\delta} L^*\|_{O_{r\delta}} \\
&\leqslant \tfrac{1}{2}\|K^* - L^*\|_{O_{r\delta}}
\end{aligned}$$

Thus, $\|K^* - L^*\|_{O_{r\delta}} = 0$.

Since $K(\nu,\tau)$ is a stationary element of $T_{r\delta}$ on $O_{r\delta}$, it is the stationary element given by $K(\nu,\tau) = \lim_{\alpha\to\infty} K(\nu,\tau;\alpha)$. Thus, Eq. (18) follows.

C. Convergence at large

The proof of convergence on a narrow strip given in the preceding section will be used here to construct a proof of convergence at large (on the domain O). It will be shown that on O $\lim_{\alpha\to\infty} K_m(\nu,\tau;\alpha) = K_m(\nu,\tau)$ where $K_m(\nu,\tau;0)$ is given by Eq. (14) and $K_m(\nu,\tau;\alpha)$ is given by Eq. (16). An inductive argument will be given which assumes Eq. (18) on $O_{r\delta}$ as a starting point.

Proof: Notice from Eq. (18) that when $r = 0$,

$$\begin{aligned}
K_m(\nu,\tau) &= K_m(\nu - m\tau,0) + \int_0^{\tau} d\lambda\, \sigma_m(\nu - m(\tau-\lambda),\lambda) \\
&\quad + \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^{\tau} d\lambda\, [\nu - m(\tau-\lambda)] \\
&\quad \times K_n(0,\lambda;\alpha) K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha) \quad (19)
\end{aligned}$$

for $0 < \tau \leqslant \delta$. (It will be assumed that the values of ν under consideration here are always on O.) Assume that for some value of r,

$$\begin{aligned}
K_m(\nu,r\delta) &= K_m(\nu - mr\delta,0) \\
&\quad + \int_0^{r\delta} d\lambda\, \sigma_m(\nu - m(r\delta-\lambda),\lambda) \\
&\quad + \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^{r\delta} d\lambda\, [\nu - m(r\delta-\lambda)] \\
&\quad \times K_n(0,\lambda;\alpha) K_{m-n}(\nu - m(r\delta-\lambda),\lambda;\alpha). \quad (20)
\end{aligned}$$

Notice from Eq. (19) that Eq. (20) is true when $r = 1$. By the induction hypothesis [Eq. (20)],

$$\begin{aligned}
&K_m(\nu - m(\tau - r\delta),r\delta) \\
&= K_m(\nu - m\tau,0) + \int_0^{r\delta} d\lambda\, \sigma_m(\nu - m(\tau-\lambda),\lambda) \\
&\quad + \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^{r\delta} d\lambda\, [\nu - m(\tau-\lambda)] \\
&\quad \times K_n(0,\lambda;\alpha) K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha). \quad (21)
\end{aligned}$$

Equation (21) can be substituted into Eq. (18) to obtain,

$$\begin{aligned}
&K_m(\nu,\tau) = \\
&K_m(\nu - m\tau,0) + \int_0^{\tau} d\lambda\, \sigma_m(\nu - m(\tau-\lambda),\lambda) \\
&\quad + \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^{\tau} d\lambda\, [\nu - m(\tau-\lambda)] \\
&\quad \times K_n(0,\lambda;\alpha) K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha) \quad (22)
\end{aligned}$$

on $O_{r\delta}$, i.e., for $r\delta < \tau \leqslant (r+1)\delta$. In particular, Eq. (22) is true for $\tau = (r+1)\delta$. Thus, if Eq. (20) is true for any value of r, it is true for all larger values of r. Since Eq. (20) is true for $r = 1$, it is true for all values of r and Eq. (22) is true everywhere on O.

Discussion: Equation (22) is the primary result of this paper. From Eq. (22) a sequence of functions can be computed with the understanding that the sequence will converge to the truncated Fourier series expansion of the Vlasov plasma distribution function.

VI. CONVERGING SEQUENCES FOR THE KINETIC AND FIELD ENERGIES

In Appendix B it is shown that in the limit $M = \infty$, K'_m and K''_m ($K'_m = \partial K_m/\partial v$, etc.) can be uniformly bounded on O. In the following it will be assumed that $\|K'\|_O$ and $\|K''\|_O$ exist for finite M. In this case essentially the same procedure as the one given above can be carried out to prove,

$$\begin{aligned}K'_m(\nu,\tau) = {}& K'_m(\nu - m\tau,0) + \int_0^\tau d\lambda\, \sigma'_m(\nu - m(\tau-\lambda),\lambda) \\ &+ \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^\tau d\lambda\, K_n(0,\lambda;\alpha) \\ &\times K_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha) \\ &+ \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^\tau d\lambda\, [\nu - m(\tau-\lambda)] \\ &\times K_n(0,\lambda;\alpha) K'_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha)\end{aligned} \tag{23}$$

and

$$\begin{aligned}K''_m(\nu,\tau) = {}& K''_m(\nu - m\tau,0) + \int_0^\tau d\lambda\, \sigma''_m(\nu - m(\tau-\lambda),\lambda) \\ &+ 2\lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^\tau d\lambda\, K_n(0,\lambda;\alpha) \\ &\times K'_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha) \\ &+ \lim_{\alpha\to\infty} \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)\int_0^\tau d\lambda\, [\nu - m(\tau-\lambda)] \\ &\times K_n(0,\lambda;\alpha) K''_{m-n}(\nu - m(\tau-\lambda),\lambda;\alpha)\end{aligned} \tag{24}$$

as long as the boundary conditions are chosen so that $\|\sigma'\|_O$ and $\|\sigma''\|_O$ exist.

Equations (23) and (24) show that the sequences of functions which are obtained by differentiating the $K_m(\nu,\tau;\alpha)$ converge to the respective derivatives of $K_m(\nu,\tau)$. Thus, $K'_m(\alpha)$ and $K''_m(\alpha)$ can be expected to approximate the exact derivatives. The derivatives, with respect to ν, play an important role in the application of this theory. At $\nu = 0$, K_m and its derivatives are related to the coefficients in Fourier series expansions of the moments of the distribution function. For example,

$$n(x,\tau) = \sum_{m=-\infty}^{\infty} f_m(0,\tau)e^{im\pi x}$$

and

$$u(x,\tau) = -\frac{i}{\pi}\sum_{m=-\infty}^{\infty} f'_m(0,\tau)e^{im\pi x},$$

in which f_m and f'_m can be computed from K_m and K'_m (see Appendix B).

From Eqs. (23) and (24) it can be shown that,

$$\begin{aligned}K''_0(0,\tau;\alpha) = {}& K''_0(0,0) + \int_0^\tau d\lambda\, \sigma''_0(0,\lambda) \\ &- \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)^2 [\,|K_n(0,0)|^2 \\ &- |K_n(0,\tau;\alpha-1)|^2].\end{aligned} \tag{25}$$

This equation bears on energy conservation in the plasma. Using

$$Q(x,\tau) = \int_{-\infty}^{\infty} dv\, v^3 F(x,v,\tau)$$

and $P_0(\tau)$ for the space average of $P(x,\tau)$, Eq. (25) can be rewritten in terms of more familiar notation as

$$\begin{aligned}P_0(\tau;\alpha) = {}& P_0(0) + \epsilon_0^2(0) + \left(\frac{1}{\pi}\right)^2 \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)^2 |K_n(0,0)|^2 \\ &- \left[\epsilon_0^2(\tau) + \left(\frac{1}{\pi}\right)^2 \sum_{n=-M}^{M}{}' \left(\frac{1}{n}\right)^2 |K_n(0,\tau;\alpha-1)|^2\right] \\ &+ \frac{1}{2}\int_0^\tau d\lambda\, \delta Q(\lambda),\end{aligned} \tag{26}$$

where $\delta Q(\lambda) = Q(-1,\lambda) - Q(1,\lambda)$. The αth iterate to the space averaged kinetic energy density is given by $P_0(\tau;\alpha)$. The contributions to $P_0(\tau;\alpha)$ on the right side of Eq. (26) are as follows: The first line in Eq. (26) gives the total energy in the plasma at $\tau = 0$. $P_0(0)$ is the initial kinetic energy, $\epsilon_0^2(0)$ is the initial electric field energy in the space averaged part of the field, and the sum on this line represents the field energy in $E'(x,0)$ averaged over space. The second line of Eq. (26) gives the negative of the total field energy at any time. The sum in the second line is the space average of the field energy in the turbulent part $[E'(x,\tau)]$ of the electric field, as given by the $(\alpha - 1)$th iterate. The last line of Eq. (26) gives the accumulated net transfer of energy into $-1 < x < 1$. Thus, energy is conserved at each iteration. Furthermore, sequences of approximations for the kinetic and electric field energies in the plasma can be computed from Eq. (26). From Eq. (23) and (24) it can be seen that these sequences must converge to, respectively, the kinetic energy in the truncated distribution function and the corresponding turbulent electric field energy.

VII. CONCLUSION

A one-dimensional electrostatic Vlasov–Maxwell plasma model has been considered in a bounded spatial region. Consideration has been limited to a plasma with uniform and stationary ion background and with zero space-averaged charge. An iterative method has been constructed for computing a sequence of approximations to the probability distribution function for the initial-boundary value problem.

The probability function has been Fourier transformed in its velocity variable, and Fourier series expanded in its spatial variable. The Fourier series expansion has been truncated at an arbitrarily large but finite value. It has been assumed that a solution exists to the finite system of partial differential equations which govern the truncated expansion of the distribution function. Under this assumption it has been shown that the sequence of approximations mentioned in the preceding paragraph must converge to the exact solution of the truncated system. Convergence does not depend on the presence of a small expansion parameter for expansion purposes. The degree to which the solution of the truncated system approximates the solution of the infinite system has not been considered, but in view of the arbitrarily large number of Fourier modes that is allowed, a good approximation is anticipated in many applications.

The issue of the rate at which convergence occurs is under investigation at present. In those situations where convergence is rapid enough to make this iterative method useful, it can be viewed as an approximation technique for Vlasov turbulence which is nonlinear and applicable to strong turbulence in a bounded region of space.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge useful discussions with Mr. S. Carchedi and Dr. J. Cooper, Dr. G.V. Ramanathan, and Dr. D. Montgomery.

APPENDIX A

Consider the one-dimensional Vlaxov–Maxwell system of Eqs. (1)–(3). Assume the existence of a solution to this system of equations on the domain D, defined by $-1 < x < 1$, $-\infty < v < \infty$, and $\tau > 0$. Then, what combination of initial and boundary conditions of F and E uniquely, consistently, and conveniently determines that solution?

Given the existence of $E(x,\tau)$, the method of characteristics can be used to solve Eq. (1).[5] The solution is $F(x(s),v(s),\tau(s)) = F(x(0),v(0),\tau(0))$ where $x(s)$, $v(s)$, and $\tau(s)$ are the solutions of the system,

$$\frac{dx(s)}{ds} = v(s),$$

$$\frac{dv(s)}{ds} = -E(x(s),\tau(s)),$$

$$\frac{d\tau(s)}{ds} = 1,$$

subject to $x(0)$, $v(0)$, and $\tau(0)$ for initial conditions. In typical applications of this type of solution the point $(x(0),v(0),\tau(0))$ is known. Then, F is known everywhere along the characteristic line given by $x(s)$, $v(s)$, and $\tau(s)$ for $s \geqslant 0$. The complete solution depends on filling all (x,v,τ) of interest with characteristic lines which are connected to boundaries where $F(x(0),v(0),\tau(0))$ is known. But this typical approach is self-contradictory for the initial-boundary value problem being considered here.

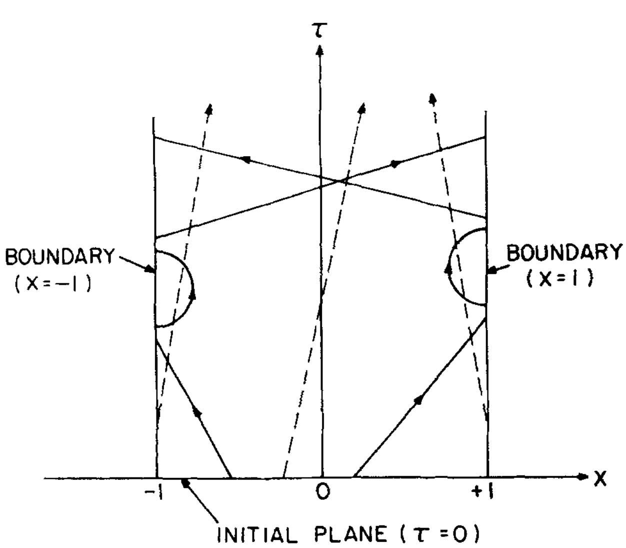


FIG. 2. A schematic representation of all possible characteristic curves on the domain D.

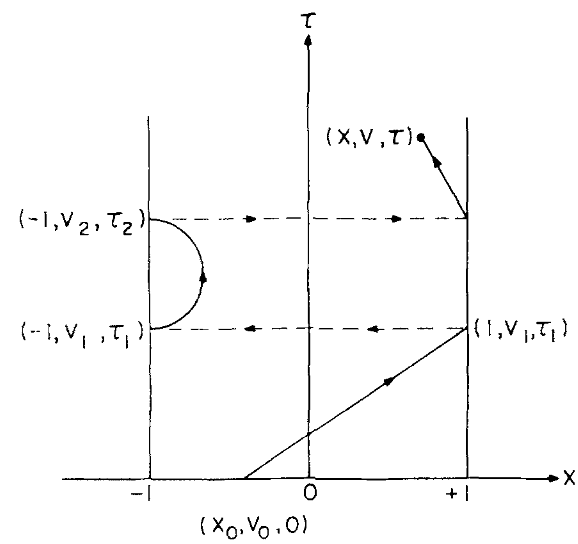


FIG. 3. A possible construction of $F(x,v,\tau)$ using $F(x,v,0)$ and $\Delta(v,\tau)$.

Figure 2 contains schematic representations of projections onto the plane $v = 0$, of various types of possible characteristic lines. Progression along the characteristic lines, with s increasing, is indicated by arrows. The dashed lines indicate characteristics which enter D on the initial plane $(-1 \leqslant x \leqslant 1, -\infty < v < \infty, \tau = 0)$ or on either of the boundaries $(x = \pm 1, -\infty < v < \infty, \tau > 0)$ and then remain trapped in D. These are the only characteristics that can be treated as outlined in the preceding paragraph. All other characteristic lines (solid lines) enter D and then *exit*, with s increasing. The solution of Eq. (1) gives $F =$ constant along each of these characteristic lines. Thus, F must have a single value at every pair of entry and exit points. The initial and boundary values for F are not independent and cannot be chosen arbitrarily.

There is no unique method for choosing the initial-boundary values for F such that the possible contradictions discussed above are avoided. The method used here has been chosen for both mathematical and observational convenience. In Fig. 3 it is demonstrated that the solution at an arbitrary point (x,v,τ), is determined by specifying $F(x,v,0)$ on the initial plane, and $\Delta(v,\tau) = F(-1,v,\tau) - F(1,v,\tau)$ in terms of F on the boundaries. The solid lines indicate possible characteristic lines along which $F =$ const. The dashed lines connect exit and entry points at which the value of F must be related through the use of $\Delta(v,\tau)$. In the example presented in Fig. 3, $F(x,v,\tau) = F(x_0 \cdot v_0, 0) + \Delta(v_1,\tau_1) - \Delta(v_2,\tau_2)$. Given the existence of $E(x,\tau)$, this choice of $F(x,v,0)$ and $\Delta(v,\tau)$ uniquely determines $F(x,v,\tau)$ through Eq. (1). The type of contradictions discussed above are avoided because $\Delta(v,\tau)$ always relates F on an exiting characteristic line to F on an entering characteristic line.

The Δ method for specifying the boundary data allows for the expression of the data in terms of the measurable moments of the distribution function on the boundaries (e.g., the difference in density, current, temperature, etc.,) where others do not. [Mathematically, it is acceptable to specify F

on the boundaries for incoming velocities only. It is also possible to express the boundary conditions in terms of the ratio, $F(x=1)/F(x=-1)$. Neither of these boundary conditions can be expressed in terms of the moments on the distribution function on the boundaries.] Notice that the well studied "periodic plasma" in which $\Delta(v,\tau)=0$ is a special case of the Δ method. More generally, the Δ method plays a natural role in the Fourier series expansion analysis of the nonperiodic plasma. A choice of $F(x,v,0)$ and $\Delta(v,\tau)$ leads to a unique set of $F_m(\nu,0)$ and $J_m(\nu,\tau)$. Thus, the Δ method has been chosen as the basis for this study of the initial-boundary value problem for the Vlasov–Maxwell plasma.

Now, assume $F(x,v,\tau)$ is known. Then, what additional initial-boundary data must be specified to determine $E(x,\tau)$ from Eqs. (2) and (3)?

If $E(x,\tau)$ is separated into its space average $\epsilon_0(\tau)$ plus an x-dependent part through $E(x,\tau)=\epsilon_0(\tau)+E'(x,\tau)$, then Eq. (2) determines $E'(x,\tau)$ only. From Eq. (2),

$$E'(x,\tau)=\mathscr{F}_0(\tau)+x(1-n_0(\tau))-\int_{-1}^{x}dx'\,[n(x',\tau)-n_0(\tau)], \quad \text{(A1)}$$

where $n_0(\tau)$ is the space average of $n(x,\tau)$, and

$$\mathscr{F}_0(\tau)=\tfrac{1}{2}\int_{-1}^{1}dx\,xn(x,\tau). \quad \text{(A2)}$$

From Eq. (3),

$$\frac{d\epsilon_0}{d\tau}=u_0(\tau), \quad \text{(A3)}$$

where $u_0(\tau)$ is the space average of $u(x,\tau)$. Notice if F is known, then $E'(x,\tau)$ is given by Eq. (A1) with no freedom for choosing boundary or initial conditions. Further, given F, $u_0(\tau)$ can be calculated and then $\epsilon_0(\tau)$ can be calculated from Eq. (A3) if $\epsilon_0(0)$ is specified. Thus, only $\epsilon_0(0)$ need be specified to determine $E(x,\tau)$ given F.

By combining the argument given above for determining F given E and for determining E given F, it seems plausible that specifications of $F(x,v,0)$, $\Delta(v,\tau)$ and $\epsilon_0(0)$ uniquely determines a solution of Eqs. (1)–(3). This is *not* a *proof* for the existence, or the uniqueness of that solution; it is at best, a plausibility argument.

APPENDIX B

The Fourier series expansion of F has been truncated in Sec. II, and the resulting truncated version of Eqs. (6)–(8) has been solved in principle in Sec. VI, not for F, but for

$$f(x,v,\tau)=\sum_{m=-M}^{M}e^{im\pi x}\frac{1}{2}\int_{-\infty}^{\infty}dv\,e^{-i\pi\nu v}f_m(\nu,\tau). \quad \text{(B1)}$$

The $f_m(\nu,\tau)$ is the solution of the truncated equations which are still related to $f(x,v,\tau)$ in the usual manner:

$$f_m(\nu,\tau)=\frac{1}{2}\int_{-1}^{1}dx\,e^{-im\pi x}\int_{-\infty}^{\infty}dv\,e^{i\pi\nu v}f(x,v,\tau). \quad \text{(B2)}$$

The degree to which f approximates F has not been investigated in this paper. However, since the theory developed in this paper applies to a truncation for arbitrarily large M it has been assumed that this approximation can be made as good as necessary. In particular, it has been assumed that for some finite M, $\|K\|_O$, $\|K'\|_O$, and $\|K''\|_O$ exist. The conditions under which this assumption is valid are under investigation.

The goal of this appendix is to prove that $\|K\|_O=1$ and both $\|K'\|_O$, and $\|K''\|_O$ exist when $M=\infty$. For this purpose it is convenient to introduce

$$\phi(\tau)=\int_0^{\tau}d\lambda\,\epsilon_0(\lambda). \quad \text{(B3)}$$

In the following it will be assumed that the boundary conditions are chosen such that when the solution of Eq. (11) is substituted into Eq. (B3), the resulting $\phi(\tau)$ can be uniformly bounded on the interval, $0<\tau\leqslant T$. From the definition of the K_m,

$$|K_m|=|f_m|, \quad \text{(B4)}$$

$$|K'_m|\leqslant\pi|\phi|\,|f_m|+|f'_m|, \quad \text{(B5)}$$

and

$$|K''_m|\leqslant\pi^2\phi^2|f_m|+2\pi|\phi|\,|f'_m|+|f''_m|. \quad \text{(B6)}$$

If f_m, f'_m and f''_m can be uniformly bounded on O, then the goal of this appendix will have been achieved.

From Eq. (B2),

$$|f_m(\nu,\tau)|\leqslant\tfrac{1}{2}\int_{-1}^{1}dx\int_{-\infty}^{\infty}dv\,F(x,v,\tau)=f_0(0,\tau),$$

but in Sec. III, it was shown that $f_0(0,\tau)=1$. Thus $|f_m(\nu,\tau)|\leqslant 1$ and $\|K\|_O=1$.

By differentiating Eq. (B2) is can be shown that

$$\left(\frac{1}{\pi}\right)^2|f''_m(\nu,\tau)|\leqslant\frac{1}{2}\int_{-1}^{1}dx\int_{-\infty}^{\infty}dv\,v^2F(x,v,\tau)=-\left(\frac{1}{\pi}\right)^2 f''_0(0,\tau),$$

where $f''_0(0,\tau)\leqslant 0$. But, $-(1/\pi)^2 f''_0(0,\tau)$ is actually the kinetic energy in the plasma. The equation which governs conservation of energy can be obtained from Eqs. (6)–(8) and written,

$$-\left(\frac{1}{\pi}\right)^2 f''_0(0,\tau)=-\left(\frac{1}{\pi}\right)^2 f''_0(0,0)+\epsilon_0^2(0)+\left(\frac{1}{\pi}\right)^2\sum_{n=-\infty}^{\infty}{}'\left(\frac{1}{n}\right)^2|f_n(0,0)|^2-\left[\epsilon_0^2(\tau)+\left(\frac{1}{\pi}\right)^2\sum_{n=-M}^{M}{}'\left(\frac{1}{n}\right)^2\times|f_n(0,\tau)|^2\right]+\frac{1}{2}\int_0^{\tau}d\lambda\,\delta Q(\lambda)$$

where $\delta Q(\lambda)$ is the net rate at which energy is entering the region in x of interest (see Sec. VII). Thus,

$$\left(\frac{1}{\pi}\right)^2|f''_m(\nu,\tau)|\leqslant-\left(\frac{1}{\pi}\right)^2 f''_0(0,0)+\epsilon_0^2(0)+\left(\frac{1}{\pi}\right)^2\sum_{n=-\infty}^{\infty}{}'\left(\frac{1}{n}\right)^2|f_n(0,0)|^2+\frac{1}{2}\int_0^{\tau}d\lambda\,\delta Q(\lambda).$$

The first line of the right side of this equation represents the total initial energy in the plasma. If this initial energy is chosen bounded, and if the rate at which energy is allowed to enter, $\delta Q(\tau)$, is assumed uniformly bounded on $0 \leqslant \tau \leqslant T$, than $|f''_m(\nu,\tau)|$ is uniformly bounded on O.

Given the uniform bounds of $f_0(0,\tau)$ and $f''_0(0,\tau)$ which have been obtained above, it is possible to prove a uniform bound on f'_m. From Eq. (B2)

$$\begin{aligned}
&|f'_m(\nu,\tau)| \\
&\leqslant \frac{\pi}{2}\int_{-1}^{1} dx \int_{-\infty}^{\infty} dv\, |v| F(x,v,\tau) \\
&= \frac{\pi}{2}\int_{-1}^{1} dx \left[\int_{-\infty}^{-1} dv + \int_{-1}^{1} dv + \int_{1}^{\infty} dv\right] |v| F(x,v,\tau) \\
&\leqslant \frac{\pi}{2}\int_{-1}^{1} dx \left[\int_{-\infty}^{-1} dv\, v^2 + \int_{-1}^{1} dv + \int_{1}^{\infty} dv\, v^2\right] F(x,v,\tau) \\
&\leqslant -\frac{1}{\pi} f''_0(0,\tau) + f_0(0,\tau).
\end{aligned}$$

Thus, f'_m is uniformly bounded on O, as well as f_m and f''_m, and from Eqs. (B5) and (B6) $\|K'\|_O$ and $\|K''\|_O$ must exist.

[1]J. Denavit, *Proceedings of the Fourth Conference on Numerical Simulation of Plasma*, edited by J.P. Boris and R.H. Shanny (Naval Research Laboratory, Washington, D.C., 1970), p. 305; T.P. Armstrong, R.C. Harding, G. Knorr, and D. Montgomery, *Methods in Computational Physics, Vol.* 9, *Plasma Physics* (Academic, New York and London, 1970), p. 29; J. Denavit and W.L. Kruer, Phys. Fluids **14**, 1782 (1971); G. Joyce, G. Knorr, and T. Burns, Phys. Fluids **14**, 797 (1971).

[2]R. Courant and D. Hilbert, *Methods of Mathematical Physics, Partial Differential Equations* (Interscience, New York and London, 1962), Vol. II, p. 461.

[3]J.D. Jackson, J. Nucl. Energy C **1**, 171 (1960).

[4]D. Fyfe and D. Montgomery, Phys. Fluids, **21**, 316 (1978).

[5]Ref. 2, p. 28.

[6]R.G. Bartle, *The Elements of Real Analysis* (Wiley, New York, London, and Sydney, 1964), p. 170.

[7]Ref. 6, p. 115; R.C. Buck, *Advanced Calculus* (McGraw-Hill, New York, 1965), 2nd ed., p. 45.

First-stage magnetization and metastable migration field in a type I superconducting slab

André Fortini, Alain Hairie, and Jean-Pierre Girard
Laboratoire de Physique du Solide de l'Université de Caen, associé au C.N.R.S. n°251, 14 032-Caen Cedex, France

(Received 5 March 1979; accepted for publication 24 May 1979)

The penetration of the field in the edge of a type I superconducting slab, placed in a uniform magnetic field, perpendicular to its plane, is analyzed by means of a suited complex potential, derived from general methods used in Dirichlet's type problems. The way of taking into account the singularities of the field distribution at the ends of the edge structure in the intermediate state, and the boundary conditions are discussed. A computational method is described for calculating the potential and flux profiles along the edges and thereby, the thermodynamic potential of the system. The first stage magnetization law together with the equilibrium dimensions of the edge structure and the previously defined migration threshold are deduced from the theory.

I. INTRODUCTION

In a foregoing publication,[1] we have discussed theoretically the existence of a metastable mechanism for explaining the usually observed two-stages flux penetration in the magnetization process of a type I superconducting sample of somewhat arbitrary shape.

The main feature of this mechanism lies in the existence of a thermodynamic threshold value H_t of the applied field H_0, from which the presence of domains in the bulk of the sample is stable. However, because of the fluxoid theorem, flux penetration can only occur when a migration threshold $H_m > H_t$ is reached.

Thus, when $H_0 > H_t$, the magnetic behavior of the sample is completely governed by the metastable mechanism, as if a sort of "magnetic barrier" existed. The above mentioned two-stage penetration corresponds to values of H_0 lower or higher than H_m

(i) $0 < H_0 < H_m$: penetration in the edges, which becomes metastable when $H_0 > H_t$.

(ii) $H_m < H_0 < H_c$: migration in the bulk. H_c is the critical field of the superconducting material.

The interest for research in the domain structure of type I superconductors has been revived in the recent years, in connection with the development of powerful observation techniques.[2–5] The migration of the flux tubes from the sample edges is well confirmed by experiment, together with the irreversible character of the magnetization curve and strong shape effects.[2,5–12] Attempts have been made, on the other hand, to derive the surface energy parameter from the observation of the domain network,[13] or to explain dimensional effects on the critical field.[14]

The purpose of the present paper is to work out the thermodynamic and electromagnetic theory of the metastable first stage, in a slab of rectangular cross section placed in a uniform field H_0, perpendicular to a face (Fig. 1), and assumed of infinite length so that the problem can be reduced to a two-dimensional one. As in Ref. 1 the fine structure of the intermediate state will be averaged out into a phase of continuously varying composition, which is valid if the sample is large enough. In addition the surface energy of the normal-diamagnetic walls will be ignored. The surface energy parameter Δ together with the external dimensions of the sample, determine the equilibrium spacing of the domain structure but is, in turn, negligible in the overall equilibrium of the sample, as far as Δ is much smaller than the spacing.[15]

The second magnetization stage which can be treated in a rather different way will be studied in the next paper.

II. COMPLEX POTENTIAL FOR THE PENETRATION IN THE EDGES

In Fig. 1(a) a half-section of the slab in the $u(x,z)$ plane is represented. In order to avoid a magnitude larger than H_c, the field penetrates into the edges up to a depth 57 (and 7′5′). The edges volume 567, 7′6′5′ are in the intermediate state and the remaining volume 257, 7′5′2′ is in a perfect diamagnetic state.

The calculation of the magnetic moment at equilibirum requires the knowledge of the field distribution inside and outside the sample. This is a very difficult problem in general. It can be appreciably simplified by taking into account the available information about the internal equilibrium. In the intermediate state, the field at equilibrium is of constant magnitude as long as the number of domains is large enough, as this has been shown by different authors.[16,17] This constant value, moreover, turns out to be the critical field H_c, except for very small corrections due to dimensional effects.[17] Strictly speaking a field of constant magnitude is not necessarily uniform but merely straight lines of force. This results in a "fanlike" structure, as shown in Fig. 1(a), matched to the external field distribution by a shallow perturbed sheet, below the free surfaces 56 and 67, in which the orientation of the domain walls changes, branching may occur, etc. Since the internal discrete structure will be smoothed out into a continuous one, this perturbed sheet will be ignored in the following.

Finally, on account of the above mentioned information about the equilibrium in the intermediate state, the problem is now reduced to the determination of an external field distribution obeying the following boundary conditions:

(a) The field at large distance is equal to the applied field $\vec{H}_0(0,H_0)$.

 0022-2488/79/102139-11$01.00

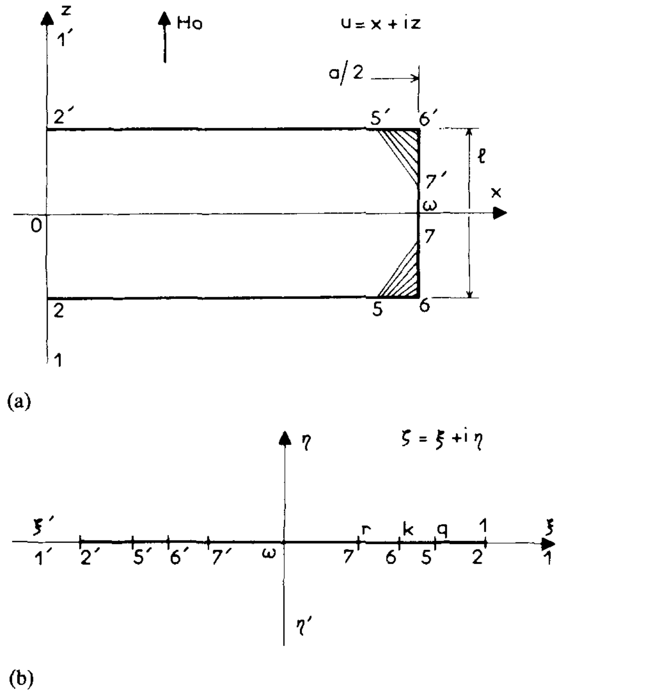


FIG. 1. (a) Half cross-section of the slab in the $u(x,z)$ plane, with fanlike domain structure in the edges, due to field penetration. (b) Conformal representation of the outside of the sample in the $u(x,z)$ plane, into the upper $\zeta(\xi,n)$ plane.

(b) The normal component of the field along 25, 77′, 5′2′ is zero.

(c) the scalar potential along 56 and 67 is matched to the internal potential distribution resulting from the fanlike structure regarded as continuous.

(d) The magnetic flux is conserved throughout the edge so that the flux reaches its maximum value at the vertex 6 (or 6′).

As to the internal equilibrium, the problem is reduced to show that the constant magnitude of the field in the fanlike structure is equal to H_c.

As in Refs. 1,15 the two-dimensional external field distribution will be represented by the complex potential

$$\psi(x,z) = \phi(x,z) + iA(x,z)/\mu_0, \tag{1}$$

$\phi(x,z)$ is the scalar potential, $A(x,z)$ the vector potential. We will choose $\phi = 0$ along the $0x$ and $A = 0$ along the constant flux line $12567\omega7'6'5'2'1'$. $\psi(x,z)$ is an analytic function of the complex variable $u = x + iz$.

The complex field H is derived from the complex potential

$$-H^* = d\psi/du = -H_x + iH_z. \tag{2}$$

To begin with, we will conformally transform the $u(x,z)$-plane into a $\zeta(\xi,\eta)$-plane so as to change the contour $1'2'5'6'7'\omega76521$ into the real $\xi'\xi$ axis (Fig. 1(b)). This is achieved by the Schwarz–Christoffel's formula

$$\frac{du}{d\zeta} = -iB\left(\frac{\zeta^2-k^2}{\zeta^2-1}\right)^{1/2}. \tag{3}$$

The chosen square root is real and positive if ζ is real and larger than 1. The part of the half-plane $x>0$ outside the sample section is transformed into the upper half-plane $\eta>0$. Notations are indicated on Fig. 1(b). ξ_2 can be taken equal to 1, $\xi_6 = k$, etc. k, B are determined by the transverse dimensions a, l of the slab

$$\frac{a}{2} = B\int_k^1\left(\frac{\xi^2-k^2}{1-\xi^2}\right)^{1/2}d\xi = B(E_{k'} - k^2F_{k'}) = BG_{k'},$$

$$\frac{l}{2} = B\int_0^k\left(\frac{k^2-\xi^2}{1-\xi^2}\right)^{1/2}d\xi = B(E_k - k'^2F_k) = BG_k,$$

hence

$$B = a/2G_{k'} = l/2G_k, \quad l/a = G_k/G_{k'}, \tag{4}$$

$k' = (1-k^2)^{1/2}$; $F_k = F(\pi/2,k)$, $E_k = E(\pi/2,k)$ denote the complete elliptic integrals of the first, second kind, and of modulus k.

We are now left with the simpler problem of seeking the complex potential $\psi(\zeta) = \psi(\xi + i\eta)$ obeying the above boundary condition a, at large distance, and b–d on the $\xi'\xi$ axis.

As is well known from Dirichlet's theorem, an analytic function is defined uniquely, in a connected domain, by the limiting values of the real (or imaginary) part on the contour of that domain. The present problem, however, is more involved.

Since $1257\omega7'5'2'1'$, in the u-plane, is an equiflux line, $A(\xi)$ is known, and has been taken equated to 0 on the contour of the diamagnetic matter 1′5′, 7′7, 51 (Fig. 1(b)). Instead, along the edges contour, 5′6′7′ and 765, we only know conditions connecting the scalar potential $\phi(\xi)$ with the vector potential $A(\xi)$.

We will first assume that the scalar potential, or the vector potential, is known along 5′6′7′ and 765, and formulate the subsequent solutions for the complex potential. In case where the vector potential is known, we are concerned with the following problem:

(i) The field at large distance is uniform and equal to $(0,H_0)$ in the u-plane.

(ii) $A(\xi) = 0$ along 1′5′, 7′ω7, 51 (Fig. 1(b)).

(iii) $A(\xi)$ is equal to any given continuous and even function along 5′7′ and 75 (Fig. 1(b)).

(5′7′) $A(\zeta) = A_-(\xi)$,

(75) $A(\zeta) = A_+(\xi) = A_-(-\xi)$.

The solution is given by

$$\psi(\zeta) = \phi(\zeta) + iA(\zeta)/\mu_0 = \frac{1}{\mu_0\pi}\left(\int_{-q}^{-r}\frac{A_-(\xi')d\xi'}{\xi'-\zeta} + \int_r^q\frac{A_+(\xi')d\xi'}{\xi'-\zeta}\right) + BH_0\zeta, \tag{5}$$

or, since $A_+(\xi)$ is an even function

$$\psi(\zeta) = \left(\frac{2\zeta}{\pi}\right)\int_r^q\frac{A_+(\xi')d\xi'}{\xi'^2-\zeta^2} + BH_0\zeta. \tag{6}$$

It can be easily checked that $\psi(\zeta)$ satisfies the required boundary conditions. At large distance $(\zeta\to\infty)$

$$-H^* = -H_x + iH_z = d\psi/du = (d\psi/d\zeta)(d\zeta/du), \tag{7}$$

from (3) and (6)

$$(du/d\zeta) \underset{\zeta\to\infty}{\to} -iB,$$

$$\psi(\zeta) \to -(2/\pi\mu_0\zeta)\int_r^q A_+(\xi')\,d\xi' + BH_0\zeta,$$

whence

$$-H_x + iH_z \to (2/\mu_0\pi\zeta^2)\int_r^q A_+(\xi')\,d\xi' + iH_0.$$

The second term on the right-hand side tends to zero, so that $H_x \to 0$ and $H_z \to H_0$.

If $\zeta = \xi$, outside 5′6′7′ and 765, along $\xi'\xi$ (Fig. 1(b)) (6) is real and so $A(\xi) = 0$.

If ζ tends to a point of 765 from above (i.e., u tends to a point of 765, in Fig. 1(a), from outside), using a well known property of Cauchy's type integrals

$$\psi(\xi) = \frac{1}{\mu_0\pi}\left(\int_{-q}^{-r}\frac{A_-(\xi')\,d\xi'}{\xi'-\xi} + \int_r^q \frac{A_+(\xi')\,d\xi'}{\xi'-\xi}\right) + \frac{iA_+(\xi)}{\mu_0}, \tag{8}$$

as required. Here $r \leqslant \xi \leqslant q$, and the second integral, in the right-hand side, is to be understood in the principal value sense.

The form (5) or (6) for $\psi(\zeta)$ is not quite convenient however, since the c and d boundary conditions are concerned with the potential profile $\phi_+(\xi)$ along 5′6′7′ and 765, rather than the flux profile $A_+(\xi)$. Since

$$\phi(\xi) = \frac{2\xi}{\mu_0\pi}\int_r^q \frac{A_+(\xi')\,d\xi'}{\xi'^2-\xi^2} + BH_0\xi, \tag{9}$$

the problem amounts to solving this equation with respect to $A_+(\xi)$.

Problems of this kind have been extensively studied in the past, for various sets of boundary conditions, in connection with the theory of harmonic functions, and in different fields of mathematical physics. The solution bounded in the neighborhood of the ends of the segments 5′7′ and 75 writes as follows (Ref. 18, Ch. 10):

$$\psi(\zeta) = \frac{[(\zeta^2-q^2)(\zeta^2-r^2)]^{1/2}}{\pi} \times\left\{\int_{-q}^{-r}\frac{\phi_-(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'-\zeta)} - \int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'-\zeta)}\right\}, \tag{10}$$

or, taking account of $\phi_-(\xi') = -\phi_+(-\xi')$

$$\psi(\zeta) = -\frac{2\zeta}{\pi}[(\zeta^2-q^2)(\zeta^2-r^2)]^{1/2} \times\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\zeta^2)}. \tag{11}$$

The square root is chosen real and positive for ζ real and larger than q.

Field at large distance. When $\zeta \to \infty$

$$\psi(\zeta) = \frac{2\zeta}{\pi}\zeta^2\left(1 - \frac{q^2+r^2}{2\zeta^2} + \cdots\right) \times\int_r^q \frac{\phi_+(\xi')}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}} \times\frac{1}{\zeta^2}\left(1+\frac{\xi'^2}{\zeta^2}+\cdots\right)d\xi',$$
$$= \frac{2}{\pi}\int_r^q\left(\zeta + \frac{2\xi'^2-q^2-r^2}{2\zeta}+\cdots\right) \times\frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}}. \tag{12}$$

From (3)

$$\frac{d\zeta}{du} = \frac{i}{B}\left(1 - \frac{k'^2}{2\zeta^2}+\cdots\right),$$

whence

$$-H_x + iH_z = (d\psi/d\zeta)(d\zeta/du) \underset{\zeta\to\infty}{\to} \frac{2i}{\pi B}\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}} - \frac{i}{\pi B\zeta^2}\int_r^q \frac{(2\xi'^2-q^2-r^2+k'^2)\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}} + \cdots. \tag{13}$$

The a-boundary condition is thus fulfilled if

$$\frac{2}{\pi B}\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}} = H_0. \tag{14}$$

The $1/\zeta^2$ term in Eq. (13) will be used later for the calculation of the magnetic momentum.

Flux profile in the edge. For $r \leqslant \xi \leqslant q$, from the theorem on the boundary values of a Cauchy's type integral, Eq. (10) yields

$$\phi_+(\xi) + iA_+(\xi)/\mu_0 = \frac{i}{\pi}[(q^2-\xi^2)(\xi^2-r^2)]^{1/2} \times\left\{\int_{-q}^{-r}\frac{\phi_-(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'-\xi)} - \int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'-\xi)} - \frac{\pi i\phi_+(\xi)}{[(q^2-\xi^2)(\xi^2-r^2)]^{1/2}}\right\}$$
$$= \phi_+(\xi) - \frac{2i\xi}{\pi}[(q^2-\xi^2)(\xi^2-r^2)]^{1/2} \times\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\xi^2)},$$

whence the expression of the flux profile along the edge free surface 567

$$A_+(\xi) = -\frac{2\mu_0\xi}{\pi}[(q^2-\xi^2)(\xi^2-r^2)]^{1/2} \times\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\xi^2)}. \tag{15}$$

The integral must be taken in the principal value sense. $A_+(\xi)$ is now expressed as a function of $\phi_+(\xi)$.

III. SINGULARITIES OF THE FIELD DISTRIBUTION

Equation (10), or (11), represents the complex potential

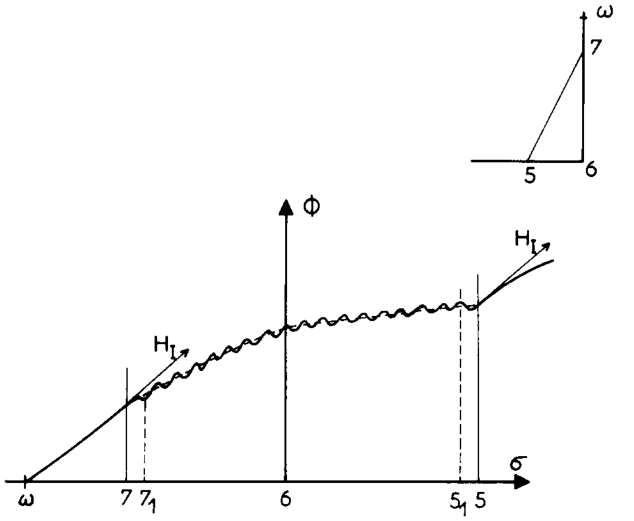


FIG. 2(a) Trend of the real and averaged macroscopic profile of the scalar potential in the edge 567 as a function of the abscissa σ along the contour.

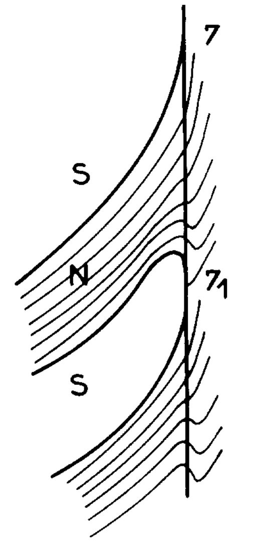


FIG. 2(b) Detailed configuration of the normal (N) and superconducting (S) domains near the end 7 of the edge.

in a very general way. There is no question, however, of taking into account the detailed structure of the intermediate state in the edge. Fortunately the actual structure can be simplified to a large extent.

Whatever may be the exact configuration of the domain walls, as soon as the sample is large enough, the intermediate state structure in the edges exhibits a large number of spatial variations of the field, over some characteristic length which mainly depends on the dimensions and the surface energy parameter.[15] The situation is schematized Fig. 2(a) by a roughly periodic variation of the scalar potential ϕ versus the abscissa σ along the contour of the slab cross section. It can be appreciated in this figure that the averaging of this structure to a macroscopic scale is valid if 56 and 67 are larger than the characteristic spacing. Nevertheless, it is worth noticing an important consequence of this averaging procedure: the value of the field at points 7, 5 in the adjacent diamagnetic matter is no longer directly related to the slope of the averaged macroscopic potential $\phi(\sigma)$. It remains instead, strongly dependent on the detail of the local domain configuration in the nearest edge structure, which is represented by the end periods 77_1 and $5_1 5$ in Fig. 3(a). Thus the averaging procedure of the intermediate state entails a finite discontinuity of the macroscopic field along the surface, in the neighborhood of the last walls.

The question now arises as to how to account for these features in the mathematical treatment, since, in describing the field distribution by means of analytic functions, finite discontinuities never occur around singular points. We will first analyze the field distribution for any continuous profile $\phi_+(\xi)$ represented by a sufficiently regular function.

The expressions of the field components are readily deduced from Eq. (7) by

$$H_x = -\phi'_x = -(d\phi/d\xi)(d\xi/dx),$$
$$H_z = -\phi'_z = -(d\phi/d\xi)(d\xi/dz),$$

or (16)

$$H_x = -A'_z/\mu_0 = -(1/\mu_0)(dA/d\xi)(d\xi/dz),$$
$$H_z = A'_x/\mu_0 = (1/\mu_0)(dA/d\xi)(d\xi/dx).$$

Appropriate expansions of the real and imaginary part of the complex potential in the neighborhood of 5, 6, 7 are calculated in the Appendix. Let us rewrite the result obtianed in the neighborhood of 7. Just above 7 ($\xi < r$) the scalar potential is given by (A9) as follows

$$\phi(\xi) = \phi_+(r) + (2/\pi)(q^2 - r^2)^{1/2}(2r)^{3/2} \times[-\varphi_r(q-r)^{-1/2} + \varphi'_r(q-r)^{1/2} + \theta_r + I_{kq}] \times(r-\xi)^{1/2} + \phi'_+(r)(\xi - r) + \mathcal{O}[(\xi-r)^{3/2}]. \tag{17}$$

$\varphi_r, \varphi'_r, \theta_r, I_{kq}$ are related to $\phi_+(r), \phi'_+(r)$ according to Eqs. (A3), (A6). The first term on the right-hand side of Eq. (17) results from the continuity of the potential at 7. The second one leads to an infinite value of the tangential field at 7 since, from Eqs. (16) and (3)

$$H_z(\xi\to r) = \left(\frac{1-r^2}{k^2-r^2}\right)^{1/2}(\phi'_\xi/B)_{\xi\to r}.$$

The third one expresses the continuity of the field if the second one is zero.

Similarly, the vector potential just below 7 ($\xi > r$) is calculated in the Appendix [cf. Eq. (A12)] as follows

$$A(\xi) = -\frac{2\mu_0}{\pi}(2r)^{3/2}(q^2-r^2)^{1/2} \times[-\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq}] \times(\xi-r)^{1/2} + \mathcal{O}[(\xi-r)^{3/2}]. \tag{18}$$

The first term on the right-hand side leads to an infinite value of the H_x component of the field at 7 since

$$\mu_0 H_x(\xi\to r) = \left(\frac{1-r^2}{k^2-r^2}\right)^{1/2}\left(\frac{A'_\xi}{B}\right).$$

The following term, instead, of order $(\xi - r)^{3/2}$, leads to $\mu_0 H_x(\xi\to r) = 0$.

Physically the *local* value of $\mu_0 H_x$, which measures the flux density, starts from zero when one gets from 7 towards 6, and next, oscillates around a slowly varying mean value. We are thus led to conclude that the profile function $\phi_+(\xi)$

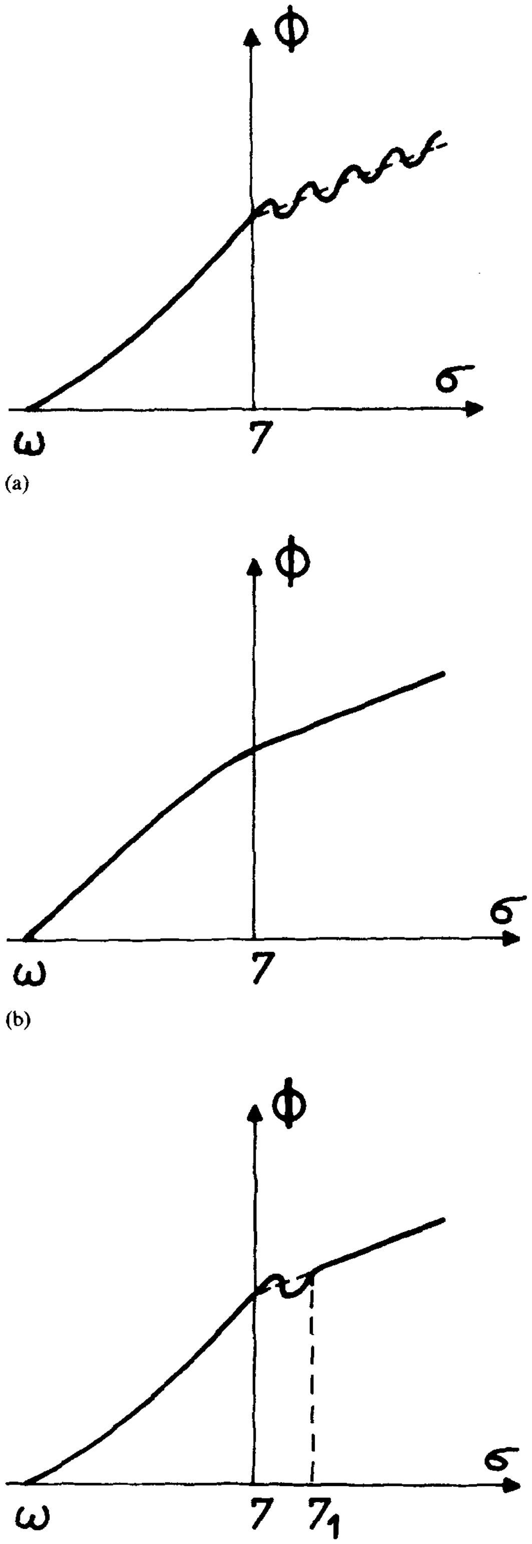


FIG. 3. Comparison of realistic and averaged profiles in the neighborhood of 7. (a) real profile. (b) averaged macroscopic profile with continuity of the field at 7. (c) restoration of the real field between ω and 7 with a small increase $\Delta\phi_+$ of ϕ_+.

necessarily satisfies the relation

$$\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq} = 0 \qquad (19)$$

which, at once, entails $H_x(7) = 0$ and finite and continuous values of the tangential field H_z around the edge end 7.

A similar relation must be obeyed at point 5

$$\varphi_q(q-k)^{-1/2} + \varphi'_q(q-k)^{1/2} + \theta_q + I_{rk} = 0. \qquad (20)$$

It expresses that the normal induction $\mu_0 H_z$ at point 5 starts from zero, and that the tangential field H_x is both finite and continuous in the neighborhood of 5.

These conclusions are valid for any realistic field distribution. The question now arises as whether they could be applied to the macroscopic distribution defined above. This does not appear necessary since, because of the dependence of Eq. (19) on the derivative $\phi'_+(r)$ (in the second and third terms), a very small local change of the profile, near $\xi = r$, suffices for satisfying the latter equation, without any restriction on the macroscopic parameters. It will be shown however, in Sec. V, from a thermodynamic argument, that (19) and (20) must be fulfilled. Therefore, Eqs. (19) and (20) will be regarded as determining the dimensions 56 and 67, through the parameters q, r.

An apparent difficulty, remains, due to the resultant continuity of the field (Fig. 3(b)). As stressed above, the continuity of the macroscopic field cannot hold at 7. The local magnitude of H_z where the last wall meets the free surface (Fig. 2(b)), is the same as inside the normal domains, i.e., H_c at equilibrium. But this departure of the macroscopic distribution from the real one is of no consequence. The reason is that the real field can again be restored, along the diamagnetic free surface $\omega 7$, by means of an infinitely small increase $\Delta\phi_+$ of ϕ_+, in the small interval 77_1 (Fig. 3(c)). It can be verified with the help of the general expressions (11) and (13) that anywhere else, the field distribution, as well as the conditions (19) and (20), undergo only infinitely small changes. This is the case in particular, for the field at large distance and thereby, for the magnetic momentum (see Sec. V).

In summary, in seeking the equilibrium of the structure as a whole, in the macroscopic description, we will have to satisfy conditions (19) and (20) which lead to the profile of Fig. 3(b). Then, the more realistic profile of Fig. 3(c) can always be regarded as achieved by a small rearrangement of the detail of the adjacent domain structure. Since such a rearrangement results in a finite variation of H_z at 7 (and H_x at 5), but in small changes in the thermodynamic potential, i.e., in the overall equilibrium, the exact value of H_z along 7ω can be ignored.

Flux conservation in the edge. The flux agreement between 56 and 67 is intimately related to the behavior of the field in the neighborhood of the vertex 6. It can be obtained by writing that the flux A reaches its maximum value at the vertex.

Denoting by H_6^h and H_6^v the tangential components of the field on each side, close to the vertex 6 (Fig. 4), the scalar potential along the edge is given by

$$d\phi_+/d\xi = -H_6^v(dz/d\xi) = BH_6^v[(k+\xi)/(1-\xi^2)]^{1/2}(k-\xi)^{1/2} \qquad (\xi<k),$$

$$d\phi_+/d\xi = -H_6^h(dx/d\xi) = BH_6^h[(\xi+k)/(1-\xi^2)]^{1/2}(\xi-k)^{1/2} \qquad (\xi>k),$$

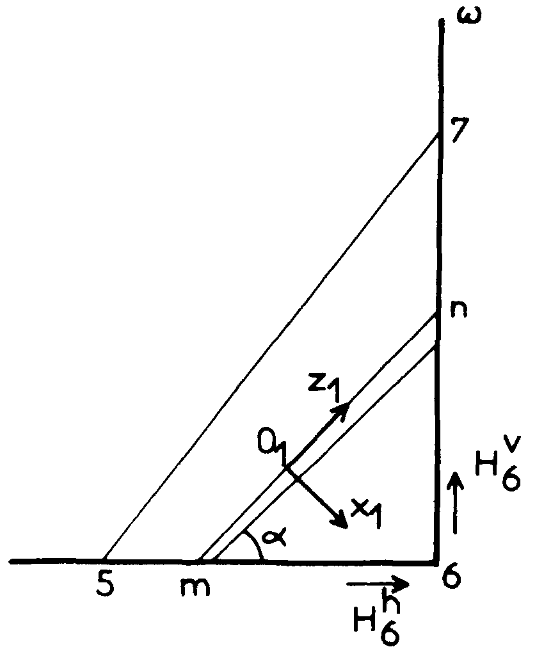


FIG. 4. Sketch of a straight line domain mn in the edge 567.

whence

$$\phi_+(\xi) = \phi_6 - (2/3)BH_6^v[2k/(1-k^2)]^{1/2}(k-\xi)^{3/2} + \cdots \quad (\xi < k)$$

$$\phi_+(\xi) = \phi_6 + (2/3)BH_6^k \times [2k/(1-k^2)]^{1/2}(\xi - k)^{3/2} + \cdots \quad (\xi > k). \tag{21}$$

The analysis (Appendix) of the function $A_+(\xi)$ resulting from the latter expansions of $\phi_+(\xi)$ yields the following expansion of $A_+(\xi)$ near the vertex 6 [Eq. (A17)],

$$A_+(\xi) = A_6 + A_6'(\xi - k) - (2/3)\mu_0 B \times [2k/(1-k^2)]^{1/2} \begin{cases} H_6^h(k-\xi)^{3/2}, & (\xi < k) \\ H_6^v(\xi - k)^{3/2}, & (\xi > k), \end{cases} \tag{22}$$

A_6, A_6' are coefficients which depend on the parameters of the problem. Thus, $A_+(\xi)$ presents a maximum located at $\xi = k$ if

$$A_6' = 0. \tag{23}$$

Assuming this condition is satisfied, $(dA_+/d\xi)_k$ behaves like $(k-\xi)^{1/2}$, $(\xi - k)^{1/2}$ so that

$$\mu_0 H_x = -\left(\frac{dA}{d\xi}\frac{d\xi}{dz}\right)_{\xi = k} = \frac{1}{B}\frac{dA}{d\xi}\left(\frac{1-k^2}{k^2-\xi^2}\right)^{1/2} = \mu_0 H_6^h \quad (\xi < k),$$

$$\mu_0 H_z = -\left(\frac{dA}{d\xi}\frac{d\xi}{dx}\right)_{\xi = k} = -\frac{1}{B}\frac{dA}{d\xi}\left(\frac{1-k^2}{\xi^2-k^2}\right)^{1/2} = \mu_0 H_6^v \quad (\xi > k). \tag{24}$$

Thus, the flux conservation condition (23) entails finite and continuous values of the induction around the vertex 6. In addition, the slopes of the flux profile on both sides of 6 are in accordance with the slopes of the potential profile (Fig. 5). This expresses the continuity of the external field components around the vertex 6.

IV. COMPUTATIONAL DETERMINATION OF THE POTENTIAL AND FLUX PROFILES ALONG THE EDGES

We are now faced with the problem of seeking the potential profile $\phi_+(\xi)$ and matching it to the internal requirement of a field distribution with constant magnitude H_I and obeying the relations (19), (20), (23). This would be, of course, an extremely difficult problem to overcome by purely analytic means. For this reason a computational method has been worked out which automatically takes into account both the boundary conditions and the Eq. (23). The dimensions of the edge will be chosen next so as to satisfy Eqs. (19) and (20).

According to the method discussed in Ref. 1 for the construction of the thermodynamic potential, the internal field is, at any time, proportional to the external field during the magnetization process. We will put

$$H_I = \gamma_I(q,r)H_0$$

where γ_I does not explicitly depend on H_0.

Consider a regular sequence of points m along 56 (Fig. 4). Each of them is put in correspondance with a point n of

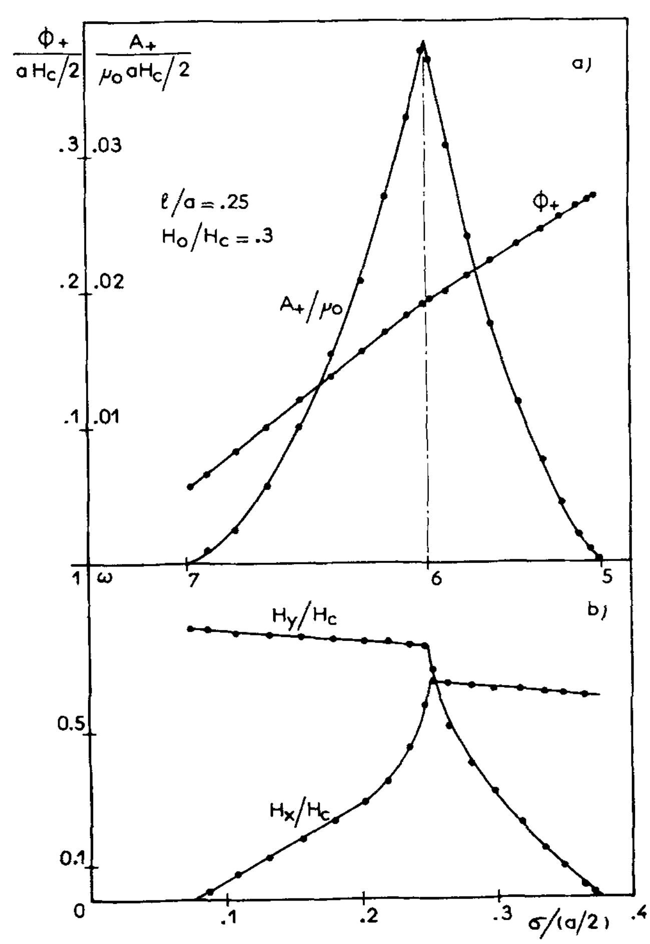


FIG. 5. (a) Potential ϕ_+ and flux profile A_+/μ_0 in the edge 567, at equilibrium, resulting from the computational method, for $l/a = 0{,}25$, in units of $aH_c/2$. (b) Related values of the reduced fields components H_x/H_c, H_y/H_c along the edge. ● Computed points.

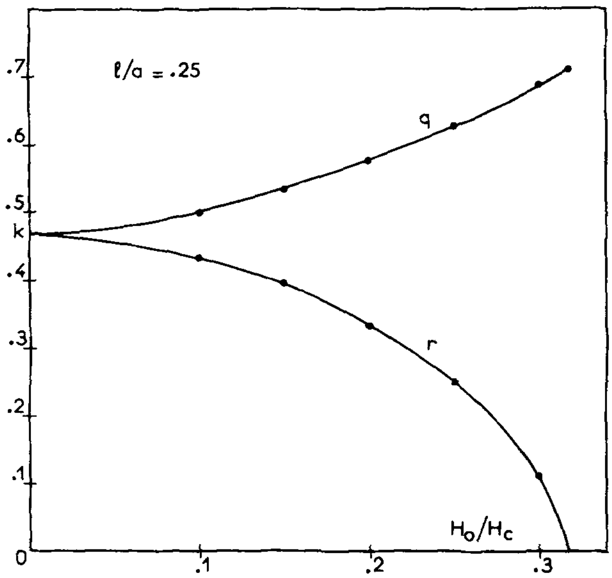


FIG. 6(a) Computed edges parameters q, r, for typical values of the reduced field H_0/H_c, up to the migration field.

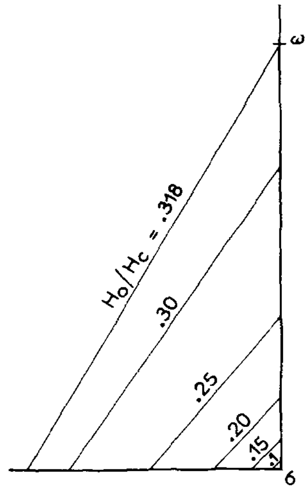


FIG. 6(b) Corresponding positions of the limiting walls 57.

67. The sequence of the n is initially arbitrary. Due to the internal fanlike structure, with a uniform value of the field, the potential profile which results from any choice for the n is known and given in terms of the angle α by

$$\phi_+(\xi) = \phi_6 + \gamma_I H_0 \int_{m6} \cos\alpha(x)\,dx, \quad \text{along } 56,$$
$$\phi_+(\xi) = \phi_6 - \gamma_I H_0 \int_{6n} \sin\alpha(z)\,dz, \quad \text{along } 67. \tag{25}$$

The resulting values of the flux at each point m or n is then calculated from Eq. (15). On defining the profile function $\beta(\xi)$ by

$$\beta(\xi) = -\int_{6n} \sin\alpha(x)\,dx, \quad \text{if } r \leqslant \xi \leqslant k,$$
$$\beta(\xi) = \int_{m6} \cos\alpha(x)\,dx, \quad \text{if } k \leqslant \xi \leqslant q, \tag{26}$$

and putting

$$R(\xi) = [(q^2 - \xi^2)(\xi^2 - r^2)]^{1/2},$$

Eq. (15) will be written in the following more detailed form:

$$A_+(\xi) = -(2\mu_0/\pi)\xi R(\xi) \int_r^q \frac{\phi_6 + \gamma_I H_0 \beta(\xi')}{R(\xi')(\xi'^2 - \xi^2)}\,d\xi'. \tag{27}$$

The potential ϕ_6 at the vertex 6 is related to the external applied field H_0 through Eq. (14) which will be rewritten, similarly

$$\int_r^q [\phi_6 + \gamma_I H_0 \beta(\xi)] R^{-1}(\xi)\,d\xi = \pi B H_0/2,$$

whence

$$\phi_6 = H_0\left[\pi B/2 - \gamma_I \int_r^q \beta(\xi) R^{-1}(\xi) d\xi\right] \times \left[\int_r^q R^{-1}(\xi)\,d\xi\right]^{-1}. \tag{28}$$

On eliminating ϕ_6 from (28), $A_+(\xi)$ can now be written as

$$A_+(\xi) = -\mu_0 B H_0 q \xi \Pi_{rq}(\xi) R(\xi)/F_\chi + (2\mu_0/\pi)\gamma_I H_0 \xi R(\xi) \times \left[(q\,\Pi_{rq}(\xi)/F_\chi) \int_r^q \beta(\xi') R^{-1}(\xi')\,d\xi' - \int_r^q \frac{\beta(\xi')\,d\xi'}{R(\xi')(\xi'^2 - \xi^2)}\right], \tag{29}$$

with

$$\int_r^q R^{-1}(\xi)\,d\xi = F_\chi/q,$$
$$\chi = (1 - r^2/q^2)^{1/2}, \tag{30}$$
$$\int_r^q \frac{d\xi'}{R(\xi')(\xi'^2 - \xi^2)} = \Pi_{rq}(\xi).$$

By means of an iterative procedure, the computer permits the determination the n sequence such that the detailed flux agreement between the two sides (Fig. 4) is satisfied, i.e.,

$$A_m = A_n.$$

The method yields correlated functions $\phi_+(\xi) - \phi_6$ and $A_+(\xi)$, matched to the internal fanlike structure and which, obviously, satisfy flux conservation, together with the condition (23) (since field components are finite at the vertex 6).

Results of such computations are shown in Fig. 5 for a typical value of the reduced applied field H_0/H_c, and $H_I = H_c$. The trend of the fields on both sides of 6 is found as expected from Eqs. (21) and (24).

The dimensions 56 and 67 of the edge can now be calculated by using Eqs. (19) and (20), which will be taken in the following form, more convenient for numerical computations

$$\{[\phi_7 - \phi_+(\xi)](r - \xi)^{-1/2}\}_{\xi \to r} = (1/\mu_0)[A_+(\xi)(\xi - r)^{-1/2}]_{\xi \to r} = 0,$$
$$\{[\phi_+(\xi) - \phi_5](\xi - q)^{-1/2}\}_{\xi \to q} = (1/\mu_0)[A_+(\xi)(q - \xi)^{-1/2}]_{\xi \to q} = 0. \tag{31}$$

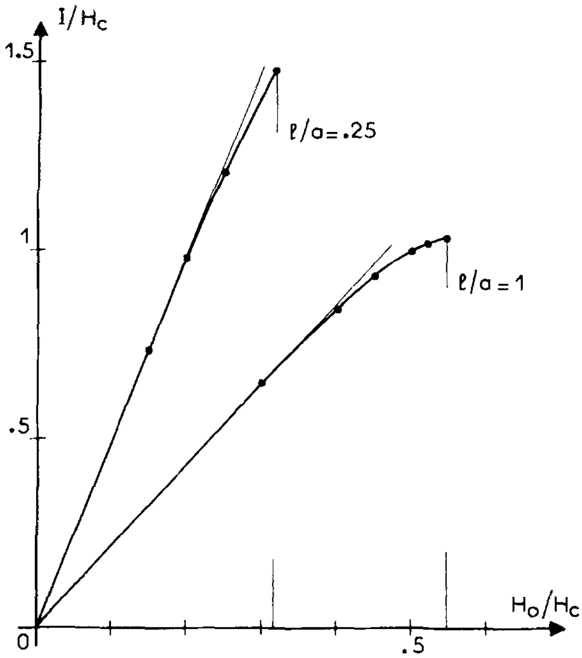


FIG. 7. Initial stage of the magnetization of the slab (in reduced units I/H_c) versus the reduced field H_0/H_c, up to the migration field, for two selected dimensions ratio ● Computed points.

In Fig. 6 are shown the computed field dependance of the edge parameters q, r, along with the related positions of the inside limiting wall of the edge, up to the migration field H_m.

V. THERMODYNAMIC POTENTIAL AND MAGNETIZATION LAW

The thermodynamic equilibrium of the sample, in the applied field $\vec{H}_0$, is determined by the minimum of the suited thermodynamic potential. The essence of the contruction of this potential was analyzed in Ref. 1. Let us recall that the intermediate state structures are initially assumed "frozen" in a definite configuration, as the field is increased from zero up to the value H_0 under consideration. The result consists of the sum of the condensation energy W_c and the magnetic contribution, which includes the magnetization energy and the coupling energy with the external field

$$G = W_c - \mu_0 \int_o^{H_0} M \cdot dH_0, \tag{32}$$

M is the magnetic momentum of the slab, obviously parallel to $0z$.

Magnetic momentum. The magnetic momentum can be conveniently derived from the behavior of the polarization field of the sample at large distance. For the presently considered slab of infinite length in the y dimension, the polarization field at large distance, on the z axis behaves like $M_z/2\pi z^2$, so that the z component of the momentum can be deduced from $H_z(z\to\infty)$ as

$$M_z = \lim_{z\to\infty} 2\pi z^2 [H_z(z) - H_0]. \tag{33}$$

On the other hand, Eq. (13) gives the behavior of the field at large ζ. Since, from Eq. (3), $\zeta \sim -z/B$ when $z\to\infty$, we obtain, per unit length of the slab

$$M = -2B \int_r^q (2\xi'^2 - q^2 - r^2 + k'^2) R^{-1}(\xi') \phi_+(\xi')\, d\xi'.$$

On substituting $\phi_+(\xi')$ from (25), (26), and (28), we obtain for the magnetic polarization intensity per unit volume

$$\begin{aligned} I &= (M/la) H_0 \\ &= -(\pi B^2/la)(2q^2 E_\chi/F_\chi - q^2 - r^2 + k'^2) \\ &\quad + (4B\gamma_I/la) \int_r^q (q^2 E_\chi/F_\chi - \xi^2) \beta(\xi) R^{-1}(\xi)\, d\xi \end{aligned} \tag{34}$$

with

$$qE_\chi = \int_r^q \xi^2 R^{-1}(\xi)\, d\xi.$$

The magnetization law of the first stage can now be derived from Eq. (34), on substituting q and r from Eq. (31) and taking $H_I = H_c$ (see below). Typical results are shown in Fig. 7. Notice the increasing curvature of the line $I(H_0/H_c)$, as it approaches the transition field, which obviously results from flux penetration in the edges.

Condensation energy. The condensation energy W_c can be deduced from the flux profile A_+ in the edge. Denoting by s the local fraction of matter in the superconductive state

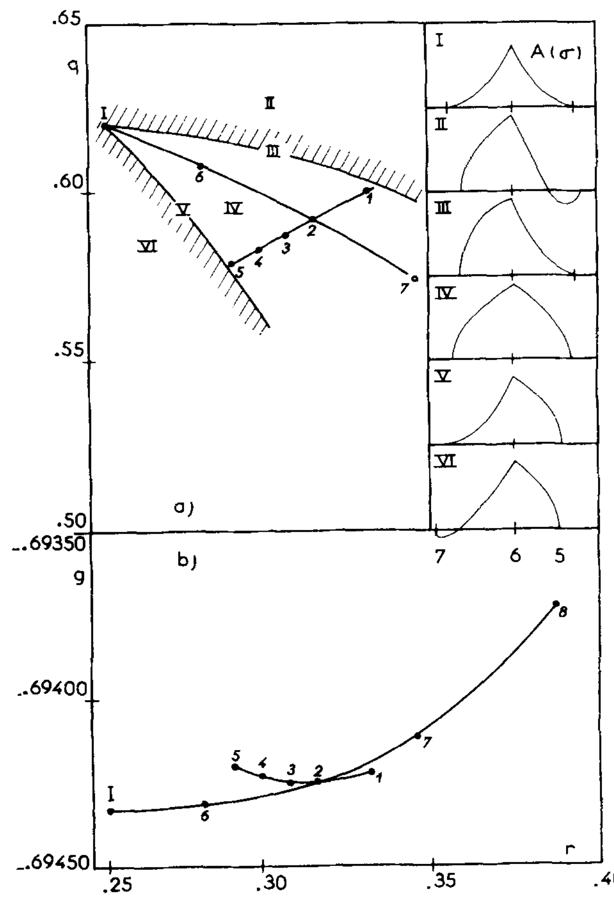


FIG. 8. (a) Evolution of the trend of the flux profile along the edge 567 (in the inset) versus the values of the parameters q, r, represented in the (q,r) plane. (b) Variations of the reduced thermodynamic potential along the lines 12345 and I6278. Note the minimum at I. ● Computed points.

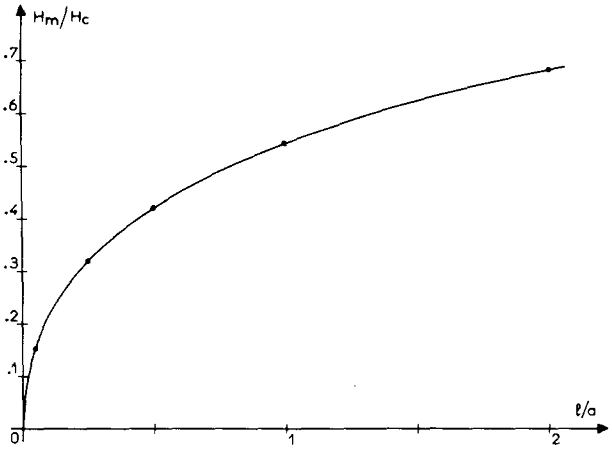


FIG. 9. Theoretical plot of the reduced migration threshold H_m/H_c, versus the dimensions ratio l/a of the slab.

in the edges, the reduced value of W_c per unit volume (in units of $\mu_0 H_c^2/2$) is given by

$$w_c = -(2/la)\int\int_{\text{half cross section}} s\,dx\,dz$$
$$= -1 + (2/la)\int\int_{\text{edges}} (1-s)\,dx\,dz.$$

The integral in the right-hand member which represents the volume in the normal state, is conveniently calculated, inside the fanlike structure, with the help of the $x_1 0 z_1$ axis, along a straight line of force (Fig. 4)

$$\int\int_{\text{edges}} (1-s)\,dx\,dz = \int\int_{\text{edges}} (1-s)\,dx_1\,dz_1$$
$$= 2\int\int_{765} B\,dx_1\,dz_1/\mu_0 H_I$$
$$= 2\int_{76} mn\cdot dA_+/\mu_0 H_I$$
$$= 2\int_{76} (\phi_m - \phi_n)(dA_+/\mu_0 H_I^2). \quad (35)$$

We have used the relations $1 - s = B/\mu_0 H_I$, $\phi_m - \phi_n = mn\cdot H_I$. We thus obtain for w_c

$$w_c = -1 - (4/la\,\mu_0 H_I)\int_r^q A_+\,d(mn). \quad (36)$$

We can now derive the complete expression of the reduced thermodynamic potential for fixed dimensions and a fixed internal structure in the edges (see Ref. 1). From Eqs. (32), (35) and the definition $I = 2M/laH_0$

$$g = -1 - (4/la\,\mu_0 H_I)\int_r^q A_+\,d(mn) - IH_0^2/H_c^2,$$

whence, by using Eqs. (29) and (34)

$$g = -1 + (4/la)\int_r^q \Big\{ (B_q/F_\chi \gamma_I)\xi R(\xi)\Pi_{rq}(\xi) + (2/\pi)\xi R(\xi)\Big[\frac{q\Pi_{rq}(\xi)}{F_\chi}\int_r^q \frac{B(\xi')\,d\xi'}{R(\xi')} + \int_r^q \frac{\beta(\xi')\,d\xi'}{R(\xi')(\xi'^2 - \xi^2)}\Big]\Big\} d(mn)$$
$$+ (B/la)(H_0^2/H_c^2)\Big[\pi B(2q^2E_\chi/F_\chi - q^2 - r^2 + k'^2) - 4\gamma_I\int_r^q (q^2E_\chi/F_\chi - \xi'^2)\beta(\xi')R^{-1}(\xi')\,d\xi'\Big]. \quad (37)$$

Internal equilibrium. The calculation of the equilibrium value of γ_I is straightforward. For fixed values of q, r a minimum of g with respect to γ_I occurs as

$$\gamma_I = \left\{\frac{\int_r^q q\xi R(\xi)\Pi_{rq}(\xi)\,d(mn)}{\int_r^q (\xi^2 F_\chi - q^2E_\chi)\beta(\xi)R^{-1}(\xi)\,d\xi}\right\}^{1/2}\frac{H_c}{H_0}. \quad (38)$$

Numerical computations for various sets of q, r, inside the allowed range defined below, show that the expression $\{\ \}^{1/2}$ is very close to unity. Thus, the internal field H_I is found, as expected, to be equal to H_c at equilibrium.

Dimensional equilibrium. Taking $H_I = H_c$, Eq. (37) defines a function of the dimensional parameters q and r. The equilibrium values of the dimensions 56 and 67 are thus defined by the equations

$$\frac{\partial g}{\partial q} = 0, \quad \frac{\partial g}{\partial r} = 0. \quad (39)$$

The trend of flux profile functions $A_+(\xi)$ obtained for various sets of q, r are represented in Fig. 8. It is shown that the physically acceptable profiles are only obtained inside a definite region of the (q,r) plane.

On studying variations of the $g(q,r)$ function along a line such as $I\,627$, inside the allowed area, it is found that a minimum only occurs at the end I for which the flux profile exhibits the characteristic trend shown in the inset (Fig. 8). This is precisely the flux behavior resulting from Eq. (31). We conclude that, as stated above, the equilibrium values of the edge dimensions are determined by the field continuity conditions (31).

VI. MIGRATION FIELD

It has been shown in Ref. 1 that the magnetic barrier which prevents domains of the edge structure from migrating into the bulk vanishes when the applied field is such that 7 and 7′ meet at ω, i.e., $r = 0$. This is the definition of the migration threshold H_m. The latter is quite easily observable in experiments, by the sharp decay which takes place in the magnetization curve, at the beginning of the migration process.

From the above definition, the migration field and the q_m dimensional parameter of the edge, just before migration, are obtained from Eq. (31), by taking $H_I = H_c$ and $r = 0$.

On Fig. 9 are plotted computed values of H_m versus the ratio l/a of the sample dimensions. The so obtained dependance of H_m upon l/a is strongly supported by experiments performed in our laboratory, which will be presented in a forthcoming paper.

VII. CONCLUSION

The initial flux penetration in a type I superconductor has been investigated in the convenient geometry of an infinite slab of rectangular cross section, in a perpendicular magnetic field. In constructing the analytic solution of the relevant Dirichlet's problem, in a macroscopic model, we have been able to compute the field profiles along the sides of the edges, bounding the volume in the intermediate state, and the related external field distribution. Finite and even infinite differences between the macroscopic model distribution and the real one are shown to have no bearing on the thermodynamic potential which permits the derivation of the equilibrium dimensions of the flux penetrated regions, and the increasing part of the magnetization law. The most important result is the calculation of the migration threshold of the domains from the edges into the bulk, which can be easily compared with experiment.

APPENDIX: Behavior of the complex potential in the neighborhood of points 5, 6, 7

Scalar potential at 5, 7. Equation (11) gives for the scalar potential just above 7 ($\xi < r$)

$$\phi(\xi) = (2\xi/\pi)[(q^2-\xi^2)(\xi^2-r^2)]^{1/2} \times \int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\xi^2)}. \quad \text{(A1)}$$

If the potential profile $\phi_+(\xi')$ is represented by a sufficiently regular function in the vicinity of $\xi' = r$, the function

$$\varphi(\xi',\xi) = \frac{\phi_+(\xi')}{[(q^2-\xi'^2)(\xi'+r)]^{1/2}(\xi'+\xi)}, \quad \text{(A2)}$$

is a regular function of ξ and ξ' too, and can be expanded as follows

$$\varphi(\xi',\xi) = (q^2-r^2)^{-1/2}(2r)^{-3/2} \times\Big\{\phi_+(r) + \Big[\phi_+(r)\frac{7r^2-3q^2}{4r(q^2-r^2)} + \phi'_+(r)\Big](\xi'-r) - \phi_+(r)(\xi-r)/2r + \cdots\Big\}.$$

We will put, for brevity in the following,

$$\varphi_r = \phi_+(r)(q^2-r^2)^{-1/2}(2r)^{-3/2}[1-(\xi-r)/2r],$$
$$\varphi'_r = (q^2-r^2)^{-1/2}(2r)^{-3/2}[\phi_+(r)(7r^2-3q^2)/4r(q^2-r^2) + \phi'_+(r)], \quad \text{(A3)}$$

so that

$$\varphi(\xi',\xi) = \varphi_r + \varphi'_r(\xi'-r) + \mathscr{O}[(\xi'-r)^2, (\xi'-r)(\xi-r), (\xi-r)^2]. \quad \text{(A4)}$$

We now proceed to separate the singular part of the integrand in the integral of the left-hand member of (A1).

$$\int_r^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\xi^2)} = \int_r^k \frac{\varphi_r + \varphi'_r(\xi'-r)}{(\xi'-r)^{1/2}(\xi'-\xi)}d\xi' + 2\theta_r + 2I_{kq} + \mathscr{O}[(\xi-r)],$$

$$\theta(\xi) = \frac{1}{2}\int_r^k \frac{\varphi(\xi',\xi)-\varphi_r-\varphi'_r(\xi'-r)}{(\xi'-r)^{1/2}(\xi'-\xi)}d\xi', \quad \text{(A5)}$$

$$I_{kq}(\xi) = \frac{1}{2}\int_k^q \frac{\phi_+(\xi')\,d\xi'}{[(q^2-\xi'^2)(\xi'^2-r^2)]^{1/2}(\xi'^2-\xi^2)}.$$

It can be easily checked that $\theta(\xi)$ and $I_{kq}(\xi)$ are regular functions in the vicinity of $\xi = r$

$$\theta(\xi)_{\xi\to r} = \theta_r + \mathscr{O}[(\xi-r)],$$
$$I_{kq}(\xi)_{\xi\to r} = I_{kq} + \mathscr{O}[(\xi-r)]. \quad \text{(A6)}$$

We will use the elementary integrals

$$\int_r^k \frac{(\xi'-r)^{n-1/2}}{\xi'-\xi}d\xi' = \int_{r-\xi}^{k-\xi}(x+\xi-r)^{n-1/2}\frac{dx}{x} = \sum_{m=1}^{n}\frac{2(\xi-r)^{n-m}(k-r)^{m-1/2}}{2m-1} + \begin{cases} 2(-)^n(r-\xi)^{n-1/2}\tan^{-1}\Big(\dfrac{k-r}{r-\xi}\Big)^{1/2} & (\xi<r),\\ (\xi-r)^{n-1/2}\ln\dfrac{(k-r)^{1/2}-(\xi-r)^{1/2}}{(k-r)^{1/2}+(\xi-r)^{1/2}} & (\xi>r),\end{cases} \quad \text{(A7)}$$

which lead to

$$\int_r^k \frac{\varphi_r+\varphi'_r(\xi'-r)}{(\xi'-r)^{1/2}(\xi'-\xi)}d\xi' = 2\varphi_r(r-\xi)^{-1/2}\tan^{-1}\Big(\frac{k-r}{r-\xi}\Big)^{1/2} + 2\varphi'_r\Big[(k-r)^{1/2}-(r-\xi)^{1/2}\tan^{-1}\Big(\frac{k-r}{r-\xi}\Big)^{1/2}\Big],$$

by expanding, next,

$$\tan^{-1}\Big(\frac{k-r}{r-\xi}\Big)^{1/2} = \pi\varphi_r(r-\xi)^{-1/2} - 2\varphi_r(k-r)^{-1/2} + 2\varphi'_r(k-r)^{1/2} - \pi\varphi'_r(r-\xi)^{1/2} + \mathscr{O}[(\xi-r)]. \quad \text{(A8)}$$

On substituting in (A5), then replacing the whole expression (A5) into (A1) and expanding the function of ξ outside the integral, we obtain

$$\begin{aligned}\phi(\xi) &= \frac{1}{\pi}(2r)^{3/2}(q^2-r^2)^{1/2}(r-\xi)^{1/2} \times\Big\{1-\frac{5q^2-9r^2}{4r(q^2-r^2)}(r-\xi)+\mathscr{O}[(\xi-r)^2]\Big\}\\ &\quad\times\{\pi\varphi_r(r-\xi)^{-1/2} - 2\varphi_r(k-r)^{-1/2} + 2\varphi'_r(k-r)^{1/2} + 2\theta_r + 2I_{kq} - \pi\phi'_r(r-\xi)^{1/2} + \mathscr{O}[(\xi-r)]\},\\ &= (2r)^{3/2}(q^2-r^2)^{1/2}\Big\{\varphi_r + \frac{2}{\pi}[-\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq}](r-\xi)^{1/2}\\ &\quad + \Big[\frac{5q^2-9r^2}{4r(q^2-r^2)}\varphi_r + \varphi'_r\Big](\xi-r) + \mathscr{O}[(r-\xi)^{3/2}]\Big\}.\end{aligned}$$

Substituting φ_r, φ'_r from (A.3) it is found that the coefficient of $(\xi - r)$ is $\phi'_+(r)$. We finally obtain

$$\phi(\xi) = \phi_+(r) + (2/\pi)(2r)^{3/2}(q^2-r^2)^{1/2}[-\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq}] \times(r-\xi)^{1/2} + \phi'_+(\xi-r) + \mathscr{O}[(r-\xi)^{3/2}]. \quad \text{(A9)}$$

Similarly, we can derive the expansions of $\phi(\xi)$ just beyond point 5, on the left-hand side ($\xi > q$)

$$\phi(\xi) = \phi_+(q) - (2/\pi)(2q)^{3/2}(q^2 - r^2)^{1/2} \times [\varphi_q(q-k)^{-1/2} + \varphi'_q(q-k)^{1/2} + \theta_q + I_{rk}] \times (\xi - q)^{1/2} + \phi'_+(\xi - q) + \mathscr{O}[(\xi - q)^{3/2}], \quad (A10)$$

with definitions of $\varphi_q, \varphi'_q, \theta_q, I_{rk}$ similar to (A3) and (A6).

Vector potential at 5, 7. We start from the expression (15) of $A(\xi)$ and introduce, as before, in the vicinity of 7, the regular function $\varphi(\xi',\xi)$ defined by (A2). The integral in (A5) has now to be calculated in the case where $\xi > r$. By using (A7)

$$\int_r^k \frac{\varphi_r + \varphi'_r(\xi' - r)}{(\xi' - r)^{1/2}(\xi' - \xi)} d\xi' = -\varphi_r(\xi - r)^{-1/2} \ln \frac{(k-r)^{1/2} + (\xi - r)^{1/2}}{(k-r)^{1/2} - (\xi - r)^{1/2}} + \varphi'_r\left[2(k-r)^{1/2} - (\xi - r)^{1/2} \times \ln \frac{(k-r)^{1/2} + (\xi - r)^{1/2}}{(k-r)^{1/2} - (\xi - r)^{1/2}}\right],$$

and expanding, next, the ln, for $\xi \to r$

$$= -2\varphi_r(k-r)^{-1/2} + 2\varphi'_r(k-r)^{1/2} + \mathscr{O}[(\xi - r)]. \quad (A11)$$

On substituting into (A5), then the whole expression (A5) into Eq. (15), the following expansion of the vector potential is obtained

$$\begin{aligned} A(\xi) = & -(2\mu_0/\pi)(2r)^{3/2}(q^2 - r^2)^{1/2}(\xi - r)^{1/2} \\ & \times \left\{1 + \frac{5q^2 - 9r^2}{4r(q^2 - r^2)}(\xi - r) + \mathscr{O}[(\xi - r)]\right\} \\ & \times \{-\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq} + \theta[(\xi - r)]\} \\ = & -(2\mu_0/\pi)(2r)^{3/2}(q^2 - r^2)^{1/2} \\ & \times [-\varphi_r(k-r)^{-1/2} + \varphi'_r(k-r)^{1/2} + \theta_r + I_{kq}](\xi - r)^{1/2} + \mathscr{O}[(\xi - r)^{3/2}]. \end{aligned} \quad (A12)$$

Notice that the coefficient of $(\xi - r)^{1/2}$ is the same as the coefficient of $(r - \xi)^{1/2}$ in (A9).

Similarly, we would get, just on the right-hand side of 5 ($\xi < q$)

$$A(\xi) = -(2\mu_0/\pi)(2q)^{3/2}(q^2 - r^2)^{1/2} \times [\varphi_q(q-k)^{-1/2} + \varphi'_q(q-k)^{1/2} + \theta_q + I_{rk}](q - \xi)^{1/2} + \mathscr{O}[(q - \xi)^{3/2}]. \quad (A13)$$

Vector potential at 6. On account of Eqs. (15) and (21)

$$A(\xi) = -(2\mu_0\xi/\pi)[(q^2 - \xi^2)(\xi^2 - r^2)]^{1/2} \times \left\{\int_r^k \frac{\phi_6 - (2/3)BH_6^v[2k/(1-k^2)]^{1/2}(k - \xi')^{1/2} + \cdots}{[(q^2 - \xi'^2)(\xi'^2 - r^2)]^{1/2}(\xi'^2 - \xi^2)} d\xi' + \int_k^q \frac{\phi_6 + (2/3)BH_6^k[2k/(1-k^2)]^{1/2}(\xi' - k)^{3/2} + \cdots}{[(q^2 - \xi'^2)(\xi'^2 - r^2)]^{1/2}(\xi'^2 - \xi^2)} d\xi'\right\}. \quad (A14)$$

We introduce the regular function at $\xi' = k, \xi = k$

$$\chi(\xi',\xi) = [(q^2 - \xi'^2)(\xi'^2 - r^2)]^{-1/2}(\xi' + \xi)^{-1} = \chi_k + \chi'_k(\xi' - k) + \cdots. \quad (A15)$$

Although the calculations are much more involved than in previous cases, they do not present any serious difficulty. They are conducted by using the following elementary integrals:

$$\begin{aligned} \int_r^k \frac{(k - \xi')^{n-1/2}}{\xi' - \xi} d\xi' &= \int_{r-\xi}^{k-\xi} (k - \xi - x)^{n-1/2} \frac{dx}{x} \\ &= -\sum_{m=1}^{n} \frac{2(k-\xi)^{n-m}(k-r)^{m-1/2}}{2m-1} \\ &\quad - \begin{cases} (k-\xi)^{n-1/2} \ln \dfrac{(k-r)^{1/2} + (k-\xi)^{1/2}}{(k-r)^{1/2} - k - \xi)^{1/2}} & (\xi < k), \\ 2(-)^n(\xi - k)^{n-1/2} \tan^{-1}\left(\dfrac{k-r}{\xi - k}\right)^{1/2} & (\xi > k), \end{cases} \end{aligned} \quad (A16)$$

and the integrals deduced from (A7) in the substitution $r \to k$, $k \to q$.

The result presents integer and half-integer powers of $\xi - k$ and $k - \xi$. As discussed in the text, the most important ones are the lowest half-integer powers which appear in a rather simple way, as follows

$$\begin{aligned} A(\xi) = & A_6 + A'_6(\xi - k) \\ & - \frac{2\mu_0 k}{\pi}[(q^2 - k^2)(k^2 - r^2)]^{1/2} \\ & \times \frac{2\pi\beta}{3}\left(\frac{2k}{1-k^2}\right)^{1/2} \\ & \times \chi_k \begin{Bmatrix} H_6^k(k - \xi)^{3/2}, & (\xi < k) \\ H_6^r(\xi - k)^{3/2}, & (\xi > k) \end{Bmatrix} + \mathscr{O}[(\xi - k)^2]. \end{aligned}$$

By replacing $\chi_k = [(q^2 - k^2)(k^2 - r^2)]^{-1/2}/2k$, it remains

$$A(\xi) = A_6 + A'_6(\xi - k) - \frac{\mu_0 B}{3}\left(\frac{2k}{1-k^2}\right)^{1/2} \times \begin{cases} H_6^h(k - \xi)^{3/2} & (\xi < k) \\ H_6^v(\xi - k)^{3/2} & (\xi > k). \end{cases} \quad (A17)$$

[1]A. Fortini and E. Paumier, Phys. Rev. B **14**, 55 (1976).
[2]F. Haensler and L. Rinderer, Helv. Phys. Acta **40**, 659 (1967).
[3]H. Trauble, U. Essmann, Phys. Stat. Solidi **25**, 395 (1969).
[4]H. Kirchner, Phys. Stat. Sol. A **4**, 531 (1971).
[5]H. Kirchner and A. Kiendl, Phys. Lett. A **39**, 293 (1972).
[6]D.C. Baird, Can. J. Phys. **42**, 1682 (1964); **43**, 971 (1965).
[7]J.P. Girard, E. Paumier, and A. Fortini, J. Phys. **32**, 895 (1971).
[8]R. Olafsson and J.F. Allen, J. Phys. F **2**, 123 (1972).
[9]A. Fortini, E. Paumier, J. Provost, and J.P. Girard, Rev. Phys. Appl. **8**, 427 (1973).
[10]J. Provost, E. Paumier, and A. Fortini, J. Phys. F **4**, 439 (1974).
[11]H.U. Habermeier, Phys. Stat. Sol. **39**, 203 (1977).
[12]U. Essmann, W. Wiethaup, and H.U. Habermeier, Phys. Stat. Sol. **43**, 151 (1977).
[13]J.J. Krempasky and D.E. Farrell, Phys. Rev. B **9**, 2894 (1974).
[14]T. Miyazak, J. Phys. Soc. Jap. **44**, 63 (1978).
[15]A. Fortini and E. Paumier, Phys. Rev. B **5**, 1850 (1972).
[16]F. London, *Superfluids* (Dover, New York), Vol. 1.
[17]P.G. De Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, Amsterdam, 1966).
[18]N.I. Muskhelishvili, *Singular Integral Equations*, translated by J.R.M. Radok (Noordhoff, Leyden, 1958).